学位論文

Search for gluinos in final states with jets and large missing transverse momentum using 36fb^{-1} data observed in the ATLAS detector

(ATLAS検出器で測定された 36fb^{-1} データを用いたジェットと横方向消失運動量を持つ終状態でのグルイーノ探索)

平成29年12月博士（理学）申請

東京大学理学系研究科
物理学専攻
安達 俊介

Ph.D Thesis

Search for gluinos in final states with jets and large missing transverse momentum using 36fb^{-1} data observed in the ATLAS detector

Shunsuke Adachi

[1]*The University of Tokyo*

Abstract

Supersymmetry is one of well-motivated theoretical frameworks beyond the Standard Model ("SM") in the elementary particle physics. The supersymmetric models predict new particles accompanying all of the SM particles. It is promising that gluino is discovered at the LHC because of its relatively light mass and its large production cross-section in a proton-proton collision. In this thesis, gluino is searched for with the ATLAS detector in the proton-proton collision at the LHC in final states with multiple jets and large missing energy ($E_{\mathrm{T}}^{\mathrm{miss}}$) originating from undetected supersymmetric particles $\tilde{\chi}_1^0$ (the lightest neutralino). In order to improve the sensitivity especially for the high $\tilde{\chi}_1^0$ mass around 1 TeV, new techniques of quark/gluon separation and multivariate analysis are introduced. For a direct decay signal with a high gluino mass and a high $\tilde{\chi}_1^0$ mass, they make a gain in the background rejection power by factor 2 from the previous study.

As a result of this search using $36.1\mathrm{fb}^{-1}$data recorded in 2015 and 2016 at the center-of-mass energy $\sqrt{s} = 13$ TeV, there is no significant excess indicating the gluino. Strongest exclusion limits on the gluino and $\tilde{\chi}_1^0$ masses are obtained for gluino direct and one-step decay models. Especially, the high $\tilde{\chi}_1^0$ mass region in the high gluino mass range is more effectively searched than the previous study. The $\tilde{\chi}_1^0$ mass is excluded up to 1 TeV in the gluino mass range of 1.50–1.80 TeV for the gluino direct decay, and excluded up to 0.85 TeV in the gluino mass range of 1.25–1.85 TeV for the gluino one-step decay at 95% confidence level.

Preface

The ATLAS experiment is taken part in by the ATLAS collaboration of a few thousands of people. In this dissertation, the analysis uses data recorded by the ATLAS detector and based on the ATLAS software framework. The LHC and ATLAS detector are described in Section 2, which are developed and operated by the CERN laboratory and ATLAS collaboration. The content in that section is based on papers published by them, but a sub-section of "data quality monitoring in the LAr calorimeter" in Section 2.2.4 is based on my own study. The analysis is based on the reconstruction algorithm and Monte Carlo simulations described in Sections 3 and 4 developed by the ATLAS collaboration. I performed the calibration of quark/gluon separation variable explained in Section 6 and the main analysis, in which quark/gluon separation and multivariate analysis are introduced, described in Sections 7–10.

Contents

1 Supersymmetry

1.1 Introduce supersymmetry

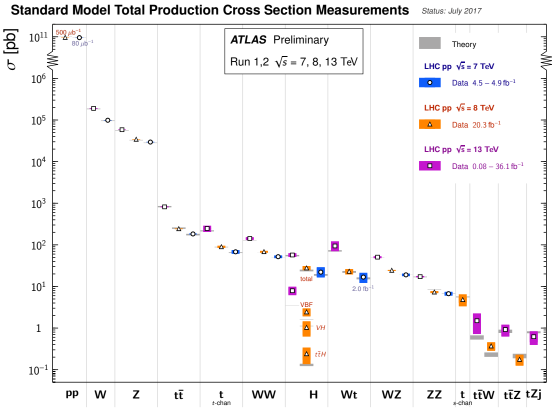


Figure 1.1: Summary of comparison between measurements with the ATLAS detector and theoretical predictions in several SM total production cross section measurements [1]. Theoretical predictions are obtained with NLO or higher order calculation.

The Standard Model (SM) of particle physics is extremely successful in describing the phenomena of elementary particles and their interactions. For instance, the predicted total production cross-sections of the SM process in p-p collision perfectly agree with the measured ones with the ATLAS detector (Figure 1.1). Nevertheless, it is believed to be a low energy realization of a more general theory because there are serious problems in the SM as follows:

- Higgs mass problem ("Fine tuning" problem)
- No unification of electroweak and strong interactions
- No Dark Matter ("DM") candidate

These problems could be explained by new physics appearing at the TeV scale. "Supersymmetry (SUSY)" is an attractive candidate for the new physics. SUSY is introduced to solve the Higgs mass problem, which is related to the stabilization of the Higgs mass against radiative corrections from Planck scale physics.

The Higgs mass is measured at the LHC experiments as $m_H = 125.09 \pm 0.21(\text{stat}) \pm 0.11(\text{syst})$ GeV [2]. The Higgs mass is calculated from both of bare mass of the Higgs and the higher order corrections. These corrections are calculated from the sum of one-loops with all of the fermions (quarks and leptons) in the SM. Fermion (f) one-loop correction as shown in Figure 1.2 contributes to these corrections. The calculation of the Higgs mass with this correction is written in Eqs. 1.1 and 1.2.

$$m_H^2 = m_{H,0}^2 + \Delta m_{H,f}^2 \tag{1.1}$$

$$\Delta m_{H,f}^2 = -\tfrac{|y|^2}{8\pi^2}\Lambda_{\mathrm{cut-off}}^2 + \cdots , \tag{1.2}$$

where $m_{H,0}$ is the bare Higgs mass, y is the Higgs coupling with a fermion and $\Lambda_{\mathrm{cut-off}}$ is cut-off scale emerging in the loop calculation, which is necessary to be enough large to cover considered physics process and here it's about 10^{15} GeV(GUT scale[Section 1.2]). m_H^2 is 125^2 GeV^2 and the second term on the right-hand side is $\mathcal{O}(10^{15\times 2})$ GeV^2. Thus, $m_{H,0}^2$ needs to be $\mathcal{O}(10^{30})$ GeV^2 and the upper 26 digits of the $m_{H,0}^2$ need to be the same as the second term. This unnatural tuning on the $m_{H,0}^2$ is called as "Fine tuning problem".

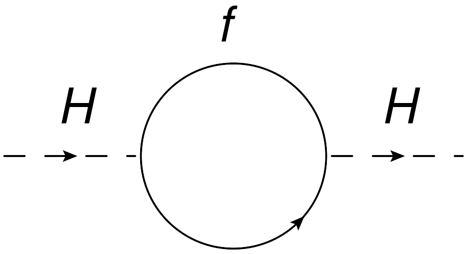


Figure 1.2: Loop correction of fermions on the Higgs mass.

If a new scalar particle S exists, a loop correction represented in Figure 1.3 also contributes to the Higgs mass corrections. This correction[1] is given as

$$\Delta m_{H,S}^2 = \frac{|\lambda_S|}{16\pi^2}\Lambda_{\mathrm{cut-off}}^2 + \cdots . \tag{1.3}$$

Fermi particles such as quarks and leptons satisfy the anticommutation relations of $f(\boldsymbol{x})f(\boldsymbol{x}') = -f(\boldsymbol{x}')f(\boldsymbol{x})$. In contrast, scalar particles follow boson statistics, which satisfy the commutation relations of $S(\boldsymbol{x})S(\boldsymbol{x}') = +S(\boldsymbol{x}')S(\boldsymbol{x})$. This difference makes opposite sign between Eqs. 1.2 and 1.3. Due to this opposite sign, if a new scalar particle with $|\lambda_S| = 2|y|^2$ such as in Fig 1.3 is introduced, the $\Lambda_{\mathrm{cut-off}}$ contributions in the Higgs mass calculation can be canceled.

In this case, the symmetry (supersymmetry, SUSY) between fermions and bosons can cancel the $\Lambda_{\mathrm{cut-off}}$ contributions in the Higgs mass calculation from all of the fermions in the SM. A supersymmetry transformation Q changes a bosonic state to a fermionic state and vice versa.

$$Q|\mathrm{fermion}\rangle = |\mathrm{boson}\rangle, \qquad Q|\mathrm{boson}\rangle = |\mathrm{fermion}\rangle \tag{1.4}$$

This symmetry introduces new particles ("supersymmetric(SUSY) particle") accompanying all of the SM particles. The tables of the SM particles and the SUSY particles are shown in Figure 1.4.

---

[1] The relative difference by 2 between new scalar particle's correction and fermion's correction is due to spin sum. A fermion has spin $\frac{1}{2}$ but a scalar particle has spin 0. The spin makes $2s + 1$(s is spin) factor in summing all spin states in the calculation.

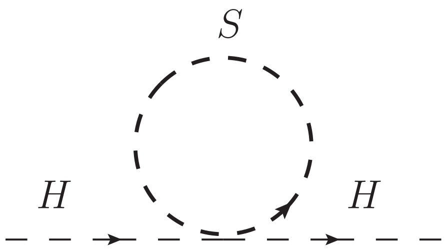


Figure 1.3: Loop correction of a scalar particle on the Higgs mass.

The new particle is not observed at the same mass as that of the SM particle. Therefore, the mass of each new particle is not exactly the same as that of the corresponding SM particle and the SUSY is considered to be broken softly in order to lift up the mass of the SUSY particles to the heavier mass that are not ruled out by experiments until now.

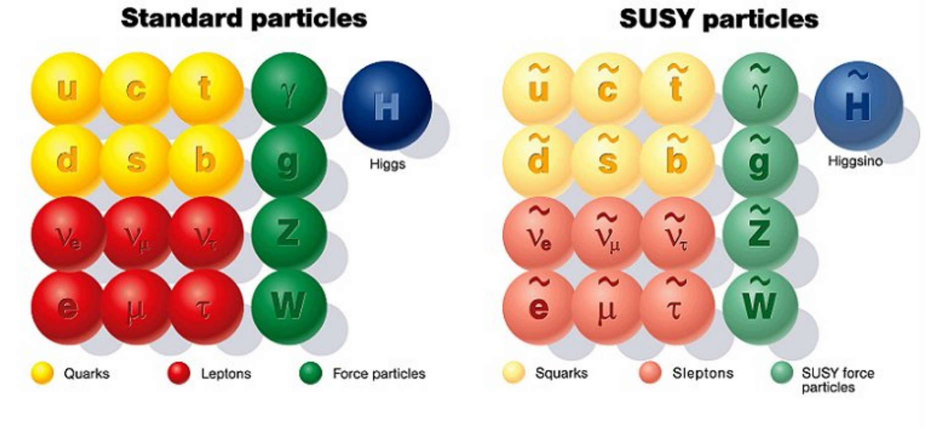


Figure 1.4: The Standard Model (SM) particles (Left) and the supersymmetric particles accompanying each particle in the SM (Right).

$\widetilde{W}^0$, $\widetilde{B}^0$, and $\widetilde{H}$ are mixed to form four mass eigenstates. They are electrically neutral and called "neutralinos". $\tilde{\chi}_1^0$, $\tilde{\chi}_2^0$, $\tilde{\chi}_3^0$, and $\tilde{\chi}_4^0$ denote these four neutralinos in the ascending mass order. $\widetilde{W}^{\pm}$, $\widetilde{H}^{+}$, and $\widetilde{H}^{-}$ make four mass eigenstates with electric charge ± 1, which are called "charginos". They are denoted as $\tilde{\chi}_1^{\pm}$ and $\tilde{\chi}_2^{\pm}$. $\tilde{\chi}_1^{\pm}$ is lighter than $\tilde{\chi}_2^{\pm}$. The neutralinos, charginos, and gluino $\tilde{g}$, which is a SUSY particle corresponding to the gluon, are referred to as "gauginos". The other SUSY particles accompanying the fermions in the SM are called "sfermions".

1.2 Benefits of the supersymmetry

1.2.1 Unification of the electroweak and strong interactions

The SUSY is introduced in order to make the light Higgs mass naturally. However, the SUSY has also other benefits. One of the important benefits is a unification of the electroweak- and strong-coupling constants at a high energy scale (GUT scale). The theory to unify these interactions is referred to as Grand Unified Theory (GUT). The SM is a gauge theory of $SU(3) \times SU(2) \times U(1)$. Their coupling constants $\alpha_i(Q)$ $(i = 1, 2, 3)$ are running by following a renormalization group equation ("RGE") depending on energy scale Q (RG energy scale). The RGE is written as [3]

$$\frac{\mathrm{d}\alpha_i(Q)}{\mathrm{d}\ln Q} = -\frac{b_i}{2\pi}\alpha_i(Q)^2, \tag{1.5}$$

where b_i are the coefficients obtained from loop calculations contributing to each gauge boson. By solving Eq. 1.5, the evolution of the coupling constants are described as

$$\frac{1}{\alpha_i(Q)} = \frac{1}{\alpha_i(m_W)} + \frac{b_i}{2\pi}\ln\left(\frac{Q}{m_W}\right) + \cdots, \tag{1.6}$$

In the naive $SU(5)$ GUT case without the SUSY, the three coupling constants cannot be unified. In contrast, if the SUSY particles are taken into account in Eq. 1.6 and located at $O(1)$ TeV, these particles contribute to b_i to change the running of the coupling constants, which makes unification of three coupling constants at an energy scale of $O(10^{15-16})$ GeV as shown in Figure 1.5. To realize this unification, the gaugino masses are loosely constrained to around $O(1)$ TeV [4].

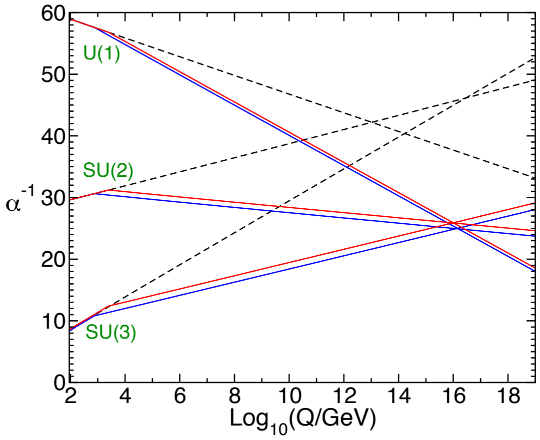


Figure 1.5: Running of the coupling constants $\alpha_i(Q)$ following renormalization group equations without (dashed lines) and with (solid lines) supersymmetry. Here, the Minimal Supersymmetric Standard Model is used as a model of the supersymmetry.

1.2.2 Dark Matter

In addition to the unification at the GUT scale, the SUSY can provide a good DM candidate.

The existence of Dark Matter (DM) is established by cosmic measurements, for example, a measurement of galactic rotation curves as shown in Figure 1.6. The DM can make an interaction with matters only at a low rate since it does not measure in luminous cosmic observation. In the SM particles, there is no candidate to explain such a non-interactive matter except neutrino. However, neutrino also cannot be a DM candidate since neutrino is inconsistent with the relic DM density explained later due to its too light mass. The relic density of the DM Ω_{DM} is obtained by the recent measurement of the anisotropy of the cosmic microwave background ("CMB" [6]). The measured value is

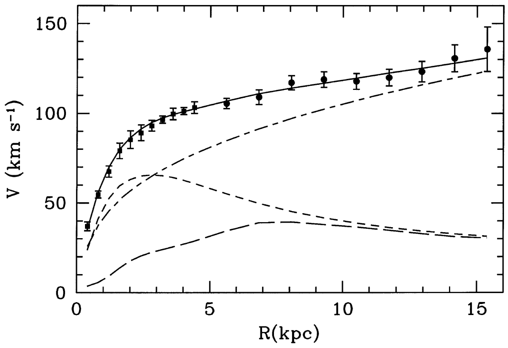


Figure 1.6: Measured rotation curve (points) of galaxy M33 with the fitted model (continuous line) [5]. The other lines show the dark halo contribution (dot-dashed line), the disc in galaxy contribution (short-dashed line) and the gas contribution (long-dashed line), respectively. The y-axis v is a rotational velocity of an object on a stable Kepler orbit with a radius R around a galaxy. The dark halo contribution is necessary to make a good agreement between the fitted model and measurement. This is one important piece of evidence for the Dark Matter.

$$\Omega_{\text{DM}} H_0^2 = 0.1186 \pm 0.0020, \tag{1.7}$$

$$\Omega_X \equiv \rho_X / \rho_{\text{crit}}, \tag{1.8}$$

where H_0 is the Hubble constant in units of 100km/(s $\cdot$ Mpc) and ρ_{crit} is the critical mass density[2]. If the DM is a Weakly Interacting Massive Particle (WIMP), the DM was in thermal equilibrium with the ordinary SM particles in the early universe, and the time evolution of the Ω_{DM} can be calculated. From the present DM density Ω_{DM} and its evolution equation, the DM mass is estimated to be 10 – a few 1000 GeV [7].

In the SUSY particles, the lightest neutralino $\tilde{\chi}_1^0$ can be stable and only weakly interact with other particles under an assumption of "R-parity" conservation. Thus, the $\tilde{\chi}_1^0$ can be a good candidate for the DM and it implies that the mass of the $\tilde{\chi}_1^0$ is 10 – a few 1000 GeV, considering the DM relic density.

---

[2] If total mass density of all matters ρ_{total} is ρ_{crit}, the universe is a flat universe, which is the border between close universe and open universe

R-parity conservation

The R-parity conservation is often assumed in the SUSY, which is mainly motivated by the experimental result of the proton-decay. The observed lower limit on the proton lifetime is

$$\tau(p \to e^+ + \pi^0) > 1.6 \times 10^{34} \text{ years} \tag{1.9}$$

at 90% confidence level. In the SUSY, if baryon and boson numbers are not conserved, proton decay can easily occur via a SUSY particle, which is inconsistent with the above experimental result. For example, Figure 1.7 shows one of proton decay processes via a SUSY particle. The decay process in Figure 1.7 will be forbidden if baryon- and lepton-numbers are conserved in the SUSY as in the SM. In the SM, these numbers are accidentally conserved as a consequence that there is no possible renormalizable term which violates lepton- or baryon-number. However, in the SUSY, there are renormalizable terms which can violate these numbers. Thus, the conservation of these numbers does not naturally occur. In order to forbid the proton decay process, alternative weaker conservation is introduced. That is the R-parity defined as

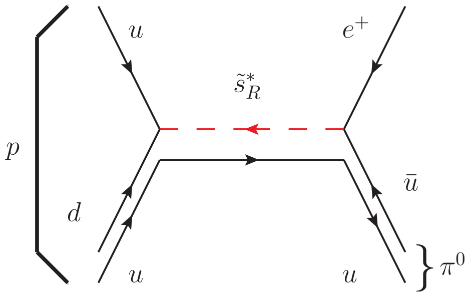


Figure 1.7: $p \to e^+ + \pi^0$ decay via a supersymmetric particle $\tilde{s}^*_R$ without any conservation related to lepton- and baryon-numbers.

$$R = (-1)^{3(B-L)+2s}, \tag{1.10}$$

where B and L are baryon- and lepton-numbers, respectively, and s indicates spin. In Figure 1.7, u and d in the initial state have $R = +1$ (even), $\tilde{s}^*_R$ has $R = -1$ (odd), and e^+ and $\bar{u}$ in the final state have $R = +1$ (even). Thus, both of $ud \to \tilde{s}^*_R$ and $\tilde{s}^*_R \to e^+\bar{u}$ are forbidden under the R-parity conservation. Furthermore,

$$R = \begin{cases} +1 & \text{(SM particles)} \\ -1 & \text{(SUSY particles)} \end{cases} \tag{1.11}$$

is established.

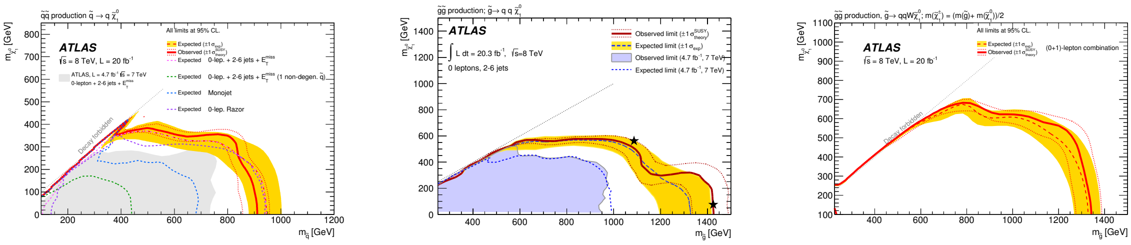
In the SUSY models with the R-parity conservation, the following two constraints on the process of the SUSY particles can be said:

1. The SUSY particles are produced from the SM particles by pair-production.
2. One SUSY particle cannot decay to only the SM particles.

By the constraint 2, the Lightest Supersymmetric Particle ("LSP") cannot decay to any SM particle, which means it is stable. Thus, if the $\tilde{\chi}^0_1$ is the LSP, $\tilde{\chi}^0_1$ could be a good DM candidate.

1.3 Run1 results

The SUSY provides great benefits described in the previous section. To obtain these benefits the gaugino mass needs to be around 1 TeV typically. In contrast, the sfermion mass is know to be further heavier than the gaugino mass. The results of the squark (and gluino) searches in Run1 at the ATLAS experiment are shown in Figures 1.8. The squark mass is excluded up to 900 GeV already. In addition, the squark mass is strongly predicted to be heavier than ~ 5 TeV as explained in the next section, which is consistent with this Run1 result.



(a) Exclusion limits on squark pair-production decaying to $qq\tilde{\chi}_1^0$ in Run1 by 0-lepton analysis [8]. Red solid and break lines show observed and expected exclusion limits, respectively. The expected exclusion limit is obtained from predicted signal yield by the simulation.

(b) Exclusion limits on gluino pair-production decaying to $qq\tilde{\chi}_1^0$ in Run1 by 0-lepton analysis. A red solid line shows an observed exclusion limit and a blue break line shows an expected exclusion limit obtained from the signal yield expected by the simulation. Black star symbols indicate two benchmark mass points which are studied in detail in the paper [8]. The shaded area (light blue) shows an older result at $\sqrt{s}$ = 7 TeV.

(c) Exclusion limits on gluino pair-production decaying to $qqW\tilde{\chi}_1^0$ via a $\tilde{\chi}_1^\pm$ in Run1 by 0-lepton and 1-lepton analysis combination [8]. The 1-lepton analysis is analysis dedicated to the final states with one isolated lepton (muon or electron). Red solid and break lines show observed and expected exclusion limits, respectively. The expected exclusion limit is obtained from predicted signal yield by the simulation. $\tilde{\chi}_1^\pm$ mass is fixed as $m(\tilde{\chi}_1^\pm) = \left(m(\tilde{\chi}_1^0) + m(\tilde{g})\right)/2$ in each mass point.

Figure 1.8: Observed and expected exclusion limits by Run1 analysis with 20.3fb^{-1}data at $\sqrt{s}$ = 8 TeV for (a) squark pair-production with direct decay to $qq\tilde{\chi}_1^0$, (b) gluino pair-production with direct decay to $qq\tilde{\chi}_1^0$, and (c) gluino pair-production with one-step decay to $qqW\tilde{\chi}_1^0$ via $\tilde{\chi}_1^\pm$ [8]. The x-axis is squark/gluino mass, and the y-axis is $\tilde{\chi}_1^0$ mass. The inside of a red contour indicates the mass region excluded at 95% confidence level (C.L.). The 0-lepton analysis is a search with the same final states as in this search.

1.4 Mass constraint on the sfermions

1.4.1 The Higgs mass calculation

The Higgs mass in the SM is a very important role in searching for SUSY particles since it implies the SUSY mass scale, especially sfermion mass. In the Minimal Supersymmetric Standard

Model ("MSSM"), which is a minimal supersymmetric extension of the SM, the Higgs mass is limited at tree-level by an upper bound of

$$m_H \leq m_Z |\cos 2\beta|, \tag{1.12}$$

where m_Z is the Z boson mass. The MSSM requires two higgs doublet (H_u and H_d) in order to cancel the infinity term of the higgsino-loop ("triangle anomaly"), and β is defined as $\tan\beta \equiv v_d/v_u$, where v_d and v_u are VEVs ("vacuum expectation value", v) of the two higgs. Thus, in order to lift the Higgs mass from the Z mass to 125 GeV, large radiative corrections are needed. A large correction can be provided by a large SUSY mass scale M_S or a large mixing X_t between a stop and the Higgs as shown in Figures 1.9 and 1.10. These imply that the M_S is roughly larger than $\sim$ 5 TeV if the mixing X_t does not make a large correction. Even if the X_t correction large, M_S would be larger than a few TeV. This constraint impacts on the common scalar mass m_0, which gives the same mass to the sfermions at the GUT scale, and it is suspected that the sfermion mass is also above 5 TeV or more.

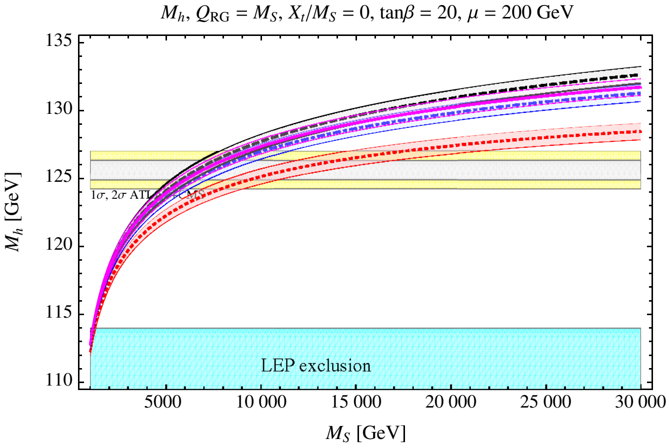


Figure 1.9: Constraint on SUSY mass scale M_S from Higgs mass (M_h = 125 GeV) [9]. The straight band around M_h = 125 GeV is an error band of the measured Higgs mass. The three curved bands with lines are constraints on the M_S and M_h. The different color lines show the resummed calculation (the solid magenta line) and the four-loop (the black dotted line), three-loop (the blue dot-dashed line), and two-loop (the red dotted line) fixed-order calculations. The overlapped regions of the Higgs mass band and the constraint bands are allowed region. Here, the parameters in the MSSM are fixed: the stop-stop-Higgs trilinear coupling $X_t = 0$, which makes stop mixing correction to the Higgs mass, the ratio of VEVs ("vacuum expectation value") of H_u and H_d in the MSSM $\tan\beta = v_d/v_u = 20$, and higgsino mass μ = 200 GeV.

1.4.2 Flavor mixing and CP violation

In addition to the Higgs mass, other experimental results related to the flavor mixing and CP violation support such a heavy sfermion mass. If the SUSY particles exist, the flavor mixing and CP violating processes that are very suppressed in the SM can be possible.

The flavor mixing process $\mu \longrightarrow e\gamma$ is very suppressed in the SM because it is forbidden at tree-level due to lepton flavor conservation. If the SUSY model has a mass mixing term ($m_{\tilde{\mu}_R^* \tilde{e}_R}$) between $\tilde{e}_R$ and $\tilde{\mu}_R$, whose superpartners are the right-handed parts of the SM mass eigenstates e and μ, the

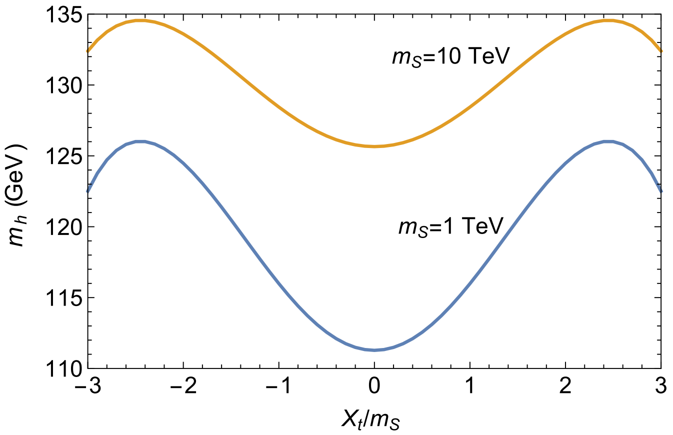


Figure 1.10: Relation between X_t/m_S and the Higgs mass m_h at $\tan\beta = 50$ in case that the m_S is 1 TeV (blue) and 10 TeV (orange) [9]. m_S is the SUSY mass scale and X_t is a stop-stop-Higgs trilinear coupling, which makes stop mixing correction to the Higgs mass.

$\mu \to e\gamma$ process is enhanced by the diagram shown in Figure 1.11. This branching ratio is given as [10]

$$\mathrm{Br}(\mu \to e\gamma) = \left(\frac{|m_{\tilde{\mu}_R^* \tilde{e}_R}|^2}{m_{\tilde{l}_R}^2}\right)^2 \left(\frac{100\ \mathrm{GeV}}{m_{\tilde{l}_R}}\right)^4 10^{-6} \times \begin{cases} 15 & \text{for} \quad m_{\tilde{B}} \ll m_{\tilde{l}_R}, \\ 5.6 & \text{for} \quad m_{\tilde{B}} = 0.5 m_{\tilde{l}_R}, \\ 1.4 & \text{for} \quad m_{\tilde{B}} = m_{\tilde{l}_R}, \\ 0.13 & \text{for} \quad m_{\tilde{B}} = 2 m_{\tilde{l}_R}, \end{cases} \tag{1.13}$$

where it is assumed that both $\tilde{e}_R$ and $\tilde{\mu}_R$ are nearly mass eigenstates and degenerate to be $m_{\tilde{l}_R}$ and that the bino $\tilde{B}$ is nearly a mass eigenstate. However, the upper limit on this branching ratio is set by MEG experiment [11]. The upper limit is $\mathrm{Br}(\mu \to e\gamma) < 4.2 \times 10^{-13}$. Thus, if the $|m_{\tilde{\mu}_R^* \tilde{e}_R}|$ is close to $m_{\tilde{l}_R}$ and $m_{\tilde{B}} \ll m_{\tilde{l}_R}$, the slepton mass is heavier than ~ 7.7 TeV.

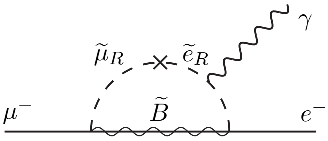


Figure 1.11: One of the diagrams contributing to the $\mu^- \to e^-\gamma$ in the MSSM with lepton flavor-violating term indicated by " × " [10].

Furthermore, CP violation in $K^0 \leftrightarrow \bar{K}^0$ mixing can be enhanced via the SUSY particles as well. The K^0 and $\bar{K}^0$ are mixed to be mass eigenstates of K_S^0 and K_L^0 via weak interactions and there is a very small mass difference between them. This mass difference can be enhanced by the process via the SUSY particles such as Figure 1.12 if the SUSY has the mixing term between different flavor squarks. However, from the experimental result that $\Delta m = m_{K_L^0} - m_{K_S^0}$ is only 3.5×10^{-15} GeV, a constraint on the $m_{\tilde{s}_R^* \tilde{d}_R}$ is obtained as [10]

$$\frac{|\mathrm{Re}[(m^2_{\tilde{s}_R^* \tilde{d}_R})^2]|^{1/2}}{m^2_{\tilde{q}}} < \left(\frac{m_{\tilde{q}}}{1000\ \mathrm{GeV}}\right) \times \begin{cases} 0.04 & \text{for} \quad m_{\tilde{g}} = 0.5 m_{\tilde{q}}, \\ 0.10 & \text{for} \quad m_{\tilde{g}} = m_{\tilde{q}}, \\ 0.22 & \text{for} \quad m_{\tilde{g}} = 2 m_{\tilde{q}}. \end{cases} \tag{1.14}$$

Hence, if the mass mixing term is the same size of the mass $m_{\tilde{q}}$, the squark mass is much larger than 10 TeV.

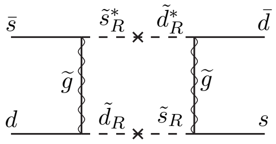


Figure 1.12: One of the diagrams contributing to $K^0 \leftrightarrow \bar{K}^0$ mixing in the MSSM with strangeness violating term indicated by "×" [10].

Since there is no mechanism to make the mixing terms zero in the SUSY, it is natural that the sfermion masses are very heavy to suppress these enhancement caused by the SUSY particles.

1.5 Production of supersymmetry particles

As remarked in the previous section, several things support the heavy sfermions. Thus, this search focuses on the gluino expected to have a few TeV mass, which has a large production cross-section as described below.

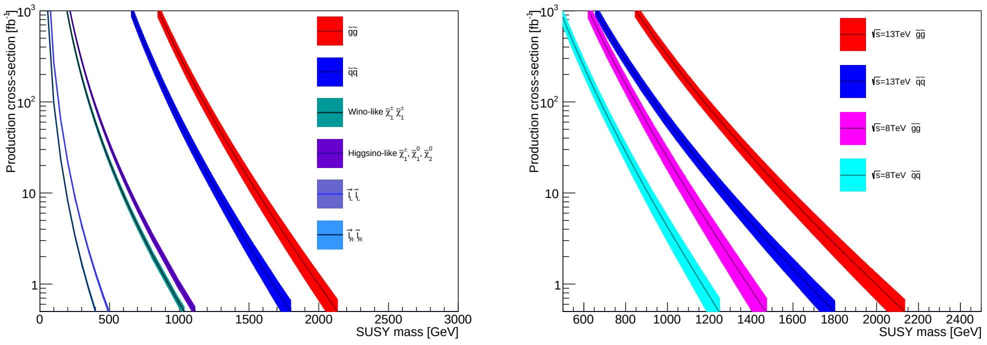
Figures 1.13 show SUSY pair-production cross-sections for various SUSY particles in the proton-proton collision. The gluino pair-production ($\tilde{g}\tilde{g}$) and squark pair-production ($\tilde{q}\tilde{q}$) have much more cross-section than $\tilde{\chi}^0\tilde{\chi}^0$, $\tilde{\chi}^\pm\tilde{\chi}^\pm$, and $\tilde{l}\tilde{l}$ productions because gluinos and squarks can be produced via the strong coupling. Therefore, the LHC, the largest proton-proton collider in the world, is the only and the best experiment at the present day to search for heavy gluino and squark productions. In Run2 of the LHC from 2015, the production cross-section of gluinos increased much from Run1 (2010–2012) by increasing the center-of-mass energy from 8 TeV to 13 TeV[3]. For example, if the gluino mass is 1.5 TeV, the gluino cross-section at $\sqrt{s}$ = 13 TeV is ten or more times higher than that at $\sqrt{s}$ = 8 TeV. Thus, this search is very important in the early stage of Run2. The Feynman diagrams of the gluino production process is shown in Figures 1.14.

1.6 Gluino decay

In this analysis, two kinds of gluino decay shown in Figures 1.15 are assumed as benchmark signal models. In these models, the following assumptions are made;

---

[3] The data acquisition periods of Run1 and Run2 at the LHC is described in Section 2.1.



(a) (b)

Figure 1.13: Production cross-sections of SUSY particles as a function of the SUSY particle mass: (a) production cross-sections of various SUSY particles at $\sqrt{s}$ = 13 TeV and (b) strong pair-production cross-sections at $\sqrt{s}$ = 8 TeV and 13 TeV. The lines are center values and the color bands indicate uncertainties. The cross-sections of electroweak productions and strong (gluino-gluino $\tilde{g}\tilde{g}$ and squark-squark $\tilde{q}\tilde{q}$) pair-productions are drawn. Wino-like chargino-chargino pair production ($\tilde{\chi}_1^\pm \tilde{\chi}_1^\pm$) cross-section is quoted from [12] and [13]. The sum cross-section of higgsino-like productions from charginos and neutralinos ($\tilde{\chi}_1^\pm, \tilde{\chi}_1^0, \tilde{\chi}_2^0$) assuming the degenerate mass of them is quoted from [14] and [15]. Left- or right-handed slepton-slepton pair production ($\tilde{l}_L^+ \tilde{l}_L^-$ or $\tilde{l}_R^+ \tilde{l}_R^-$) are quoted from [16].

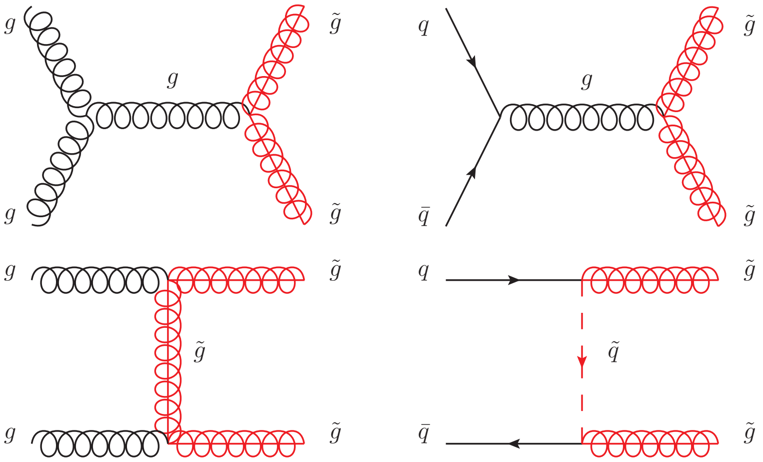


Figure 1.14: Feynman diagrams of gluino pair-production. Top two figures show s-channel production via a gluon. Bottom two figures show t-channel production.

- The LSP is the lightest neutralino and it's unseen in the detector. (It is a good candidate for the Dark Matter.)
- The gluino mass is much lighter than the squark mass. (Typically, the squark mass is assumed to be more than 5 TeV.)
- The light-flavor squark (the first and the second generations, $\tilde{u}$, $\tilde{d}$, $\tilde{c}$ and $\tilde{s}$) masses are lighter

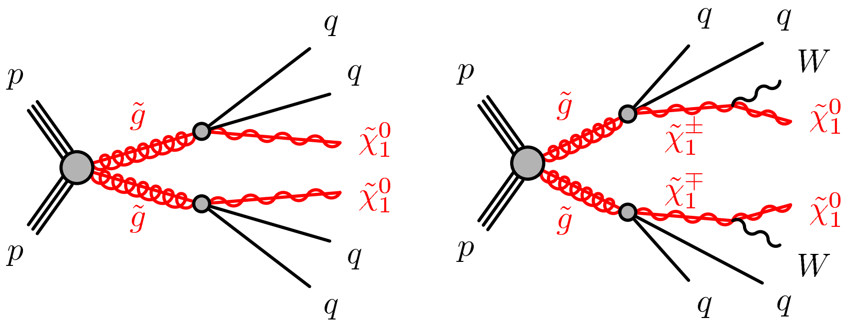
> than third generation squarks (stop and sbottom). Thus, a gluino decays to a neutralino or a chargino via a virtual light-flavor squark accompanying two light-flavor quarks (u, d, c, and s) as shown in Figure 1.15(c). There are no heavy flavor quarks (top or bottom quarks) from this decay process. (However, it's possible that a heavy flavor quark caused by a boson decay in the final states in Figure 1.15(b).)

For Figure 1.15(a), in addition to the above, it is assumed that only the LSP is lighter than the gluino. This decay is referred to as **"direct decay"** hereafter. For Figure 1.15(b), it is assumed that the LSP and the lightest chargino $\tilde{\chi}_1^{\pm}$, are lighter than the gluino and the mass difference between the LSP and $\tilde{\chi}_1^{\pm}$ is not small and can make a W boson in the decay from the LSP to $\tilde{\chi}_1^{\pm}$. Here, the mass of $\tilde{\chi}_1^{\pm}$ is fixed to the exact center value between the LSP mass and the gluino mass, i.e. $m(\tilde{\chi}_1^{\pm}) = (m(\tilde{g})+m(\tilde{\chi}_1^0))/2$. This decay is referred to as **"one-step decay"** hereafter. In the benchmark models, 100% branching fraction is assumed in each decay process. These signals are described using the simplified models [17–19], which are defined by an effective Lagrangian describing the interactions of a small number of new particles as above.

1.7 Target signal mass range

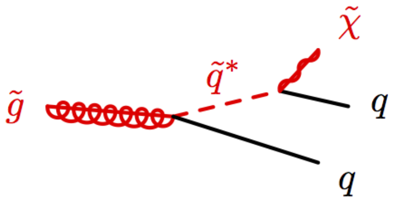
The LHC is the highest energy accelerator in the world, which is the best machine to search the heavy SUSY particles. Especially, gluinos are the most suitable target to be discovered because of its large cross-section and its relatively light expected mass. Due to the large gain in the cross-section of gluinos from Run1 to Run2, the large improvement in the high gluino mass and the low $\tilde{\chi}_1^0$ mass region is expected. In such large $\Delta M \equiv m(\tilde{g}) - m(\tilde{\chi}_1^0)$ region ($\Delta M \sim 1.5$–2 TeV), the quarks emitted from gluino decay have extremely large momenta, which make it very easy to distinguish the gluino signal from the SM background process. Therefore, the gluino search in the large ΔM region is made effectively and it is only limited by the gluino production cross-section.

However, in the high $\tilde{\chi}_1^0$ mass region such as $m(\tilde{\chi}_1^0) \sim 1$ TeV, the quarks have smaller momenta, which are not easy to be separated from the background and the search is not improved much by more luminosity or more center-of-mass energy. In the theoretical point of view, the 1 TeV $\tilde{\chi}_1^0$ is not too heavy and should be explored as well because the $\tilde{\chi}_1^0$ mass is expected to be $O(10)$ GeV– a few TeV from the DM relic density, Thus, this analysis is designed to improve the sensitivity in the high $\tilde{\chi}_1^0$ mass region around 1 TeV and the new techniques of quark/gluon separation and multivariate analysis are introduced to realize it.



(a) Direct decay.

(b) One-step decay.



(c) Detail of the gluino decay.

Figure 1.15: (a) Gluino pair-production decaying to $qq\tilde{\chi}_1^0$. (b) Gluino pair-production decaying to $qqW\tilde{\chi}_1^0$ via a $\tilde{\chi}_1^\pm$. (c) Detail of the gluino decaying to $qq\tilde{\chi}$ in (a) and (b). Gluino can decay to $qq\tilde{\chi}$ only via a virtual squark $\tilde{q}^*$ whose mass is assumed to be much heavier than the gluino mass. Here, $\tilde{\chi}$ denotes $\tilde{\chi}_1^0$ in (a) and $\tilde{\chi}_1^\pm$ in (b).

2 LHC and ATLAS detector

2.1 Large Hadron Collider

The Large Hadron Collider ("LHC", [20]) is a two-ring-superconducting-hadron accelerator and collider installed in the existing 26.7 km tunnel that was constructed for the CERN "LEP" (Large Electron-Positron collider [21]) machine. The tunnel has eight straight sections and eight arcs and lies between 45 m and 170 m below the surface.

The accelerator complex at CERN (Figure 2.1) is a succession of machines with increasing energies. Each machine injects the beam into the next one, which takes over to bring the beam to an even higher energy. The starting point for the protons is the Linear Accelerator 2 ("LINAC2"), taking hydrogen atoms from a bottle of hydrogen gas. Its energy is reached to 50 MeV in the LINAC2. The protons are transferred to the Proton Synchrotron Booster ("PSB" or "BOOSTER"), the Proton Synchrotron ("PS") and the Super Proton Synchrotron ("SPS") in this order. In these transfer steps, proton energy is increased to 1.4 GeV, 25 GeV, and 450 GeV, respectively. Finally, the 450 GeV protons are transferred to the LHC, where they are accelerated to 6.5 TeV.

The LHC started in 2010 and continued to run until 2012. In 2010 and 2011, its center-of-mass energy (represented by "$\sqrt{s}$ ") was 7 TeV and increased up to 8 TeV from 2012. This run period is called as "Run1". After Run1, the LHC was shut down until 2015 in order to upgrade both of the accelerators and the detectors. It restarted to run at the center-of-mass energy of 13 TeV from 2015. At present, it's also running and this period is called as "Run2".

In Run2, the center-of-mass energy is 13 TeV and peak luminosity is approximately $1{\times}10^{34}\mathrm{cm}^{-2}\mathrm{s}^{-1}$. The protons are stored in a bunch structure. In the LHC, more than 2000 proton bunches are stored and the time distance between bunches is 25 ns. (This time distance was 50 ns in Run1 and changed from Run2.) The one bunch contains 1×10^{11} protons. These LHC parameters and others concerning beam shape are summarized in Table 2.1.

In the LHC, there are four detectors: ATLAS, CMS, ALICE, and LHCb. The ATLAS and CMS are the general-purpose particle detectors, the ALICE is a heavy-ion detector, and the LHCb is a detector specialized in investigating *b*-quark physics. This analysis uses the data recorded with the ATLAS detector.

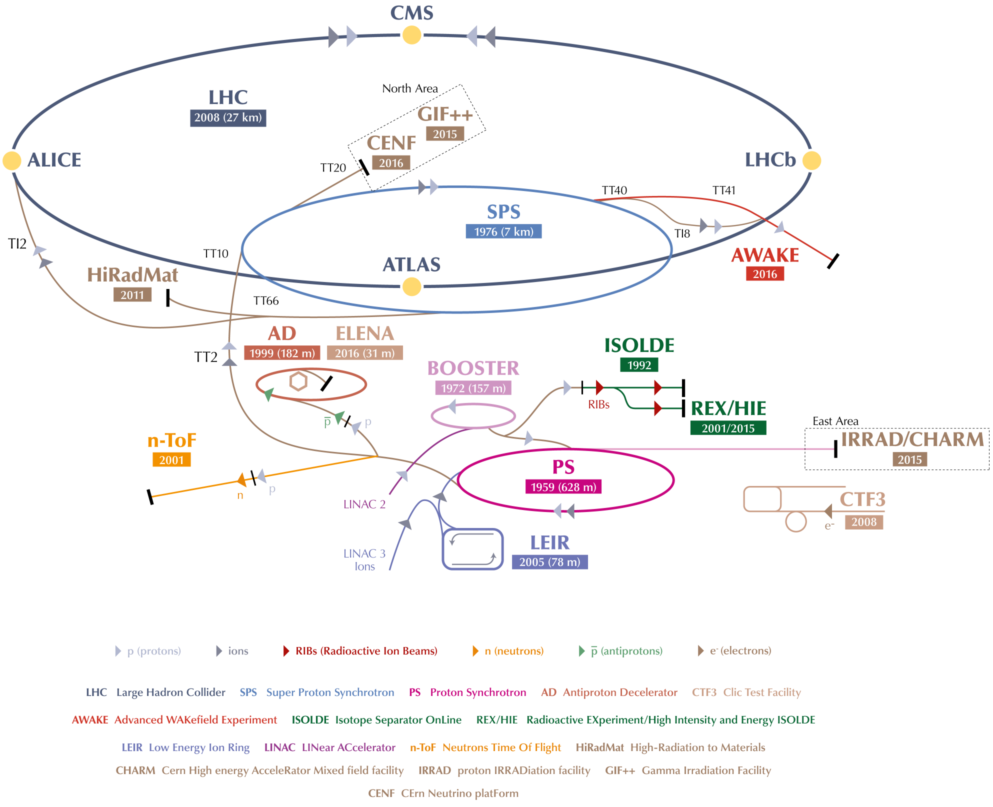


Figure 2.1: The schematic view of the accelerator complex at CERN [22]. Each name in the figure shows an abbreviation of the machine name and the full name is shown in the bottom of the figure. The direction of a triangle along a line shows the flow of accelerated particle and its color shows the type of particle. Protons obtained from LINAC2 are accelerated succeedingly by BOOSTER, PS, SPS, and LHC.

Table 2.1: The LHC parameters. The design values are cited from [20]. The values in 2015 and 2016 Runs are typical ones [23].

	[unit]	Design	2015 Run	2016 Run
Energy	[TeV]	14	13	
Peak luminosity	$[10^{34}\text{cm}^{-2}\text{s}-1]$	1	1.2	1.1
Number of bunches		2808	2448	2076
Bunch spacing	[ns]	25		
Proton intensity per bunch	[p/bunch]	1.15×10^{11}	1.1×10^{11}	1.18×10^{11}
Transverse emittance (rms)	[μm]	3.75	2.5	2.6
Bunch length (4σ)	[ns]	1.0	1.25	1.05

2.2 The ATLAS detector

A Toroidal LHC ApparatuS ("ATLAS") detector is a general purpose detector at the LHC. The ATLAS detector covers nearly 4π around the collision point and it is 44 m long and 25 m in diameter, and it weighs about 7000 tonnes.

The ATLAS detector consists of an ensemble of magnet systems and sub-detectors with a cylindrical shape placed around the interaction point ("IP") of a proton-proton collision. The layout of the ATLAS detector is shown in Figure 2.2. The ensemble of sub-detectors consists of the inner detector ("ID"), which is used for tracking of charged particles, the calorimeters, which are designed for measuring the energy of particles, and the muon spectrometer, which is utilized for identifying muons and measuring their momenta. The magnet system is designed to bend the charged particles to measure their momenta. There are a thin superconducting solenoid magnet surrounding the ID and three large superconducting toroids (one barrel and two endcaps) arranged around the calorimeters.

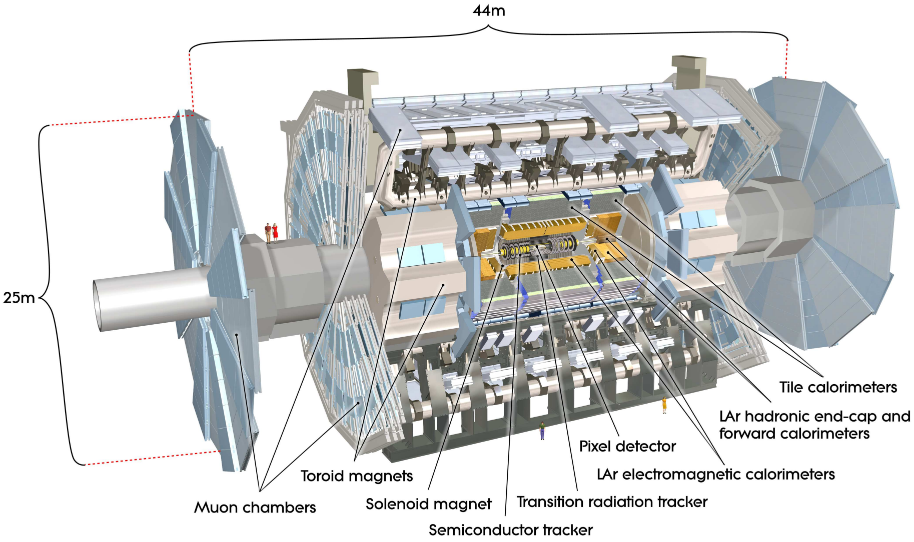


Figure 2.2: The layout of the ATLAS detector [24]. The length is 44 m and the height is 25 m. The overall weight is approximately 7000 tonnes.

2.2.1 Coordinate system

The coordinate system used to describe the ATLAS detector and the particles emitted from the collisions is a right-handed Cartesian coordinate system (x, y, z) with its origin at the center of the detector (the nominal IP). The beam axis is defined as z-axis and the x-y plane is transverse to the beam direction. The positive x-axis is defined as pointing from the interaction point to the center of the LHC ring and the positive y-axis is defined as pointing upwards. A spherical coordinates system (r, θ, ϕ) is also used. The radial distance $r = \sqrt{x^2 + y^2}$ is a distance from the IP, the azimuthal angle $\phi = \arctan \frac{y}{x}$ is measured around the beam axis, and the polar angle $\theta = \arctan \frac{r}{z}$ is the angle from the beam axis.

Rapidity (y_{rap}) is defined as

$$y_{\mathrm{rap}} = \frac{1}{2} \ln \left(\frac{E + p_z}{E - p_z} \right), \tag{2.1}$$

where E and p_z are the energy and the momentum along the beam direction, respectively. The pseudo-rapidity ("η"), which is equal to the rapidity in the relativistic limit, is often used in physics analysis. It's defined as,

$$\eta = -\ln \tan \left(\frac{\theta}{2} \right). \tag{2.2}$$

A distance ΔR in the two dimensional plane of the pseudo-rapidity (η) and azimuthal angle (ϕ) is also used as the distance between particles. It is defined as

$$\Delta R = \sqrt{(\Delta\eta)^2 + (\Delta\phi)^2}, \tag{2.3}$$

where $\Delta\eta$ and $\Delta\phi$ are the distance in the η and in the ϕ, respectively.

2.2.2 Magnet

The magnet system is composed of four large superconducting magnets. Figure 2.3 shows the layout of the four magnets. The solenoid magnet is aligned on the beam axis surrounding the ID, and the three large toroids located outside the calorimeters.

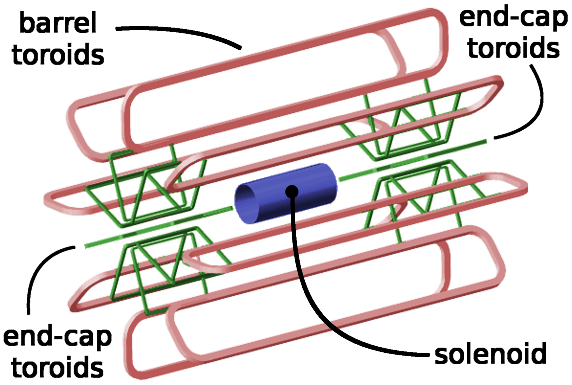


Figure 2.3: The layout of the four magnets [25]. The solenoid magnet is a layer coil located at the center. The each of the barrel and two endcap toroid magnets are composed of eight coils.

Solenoid magnet

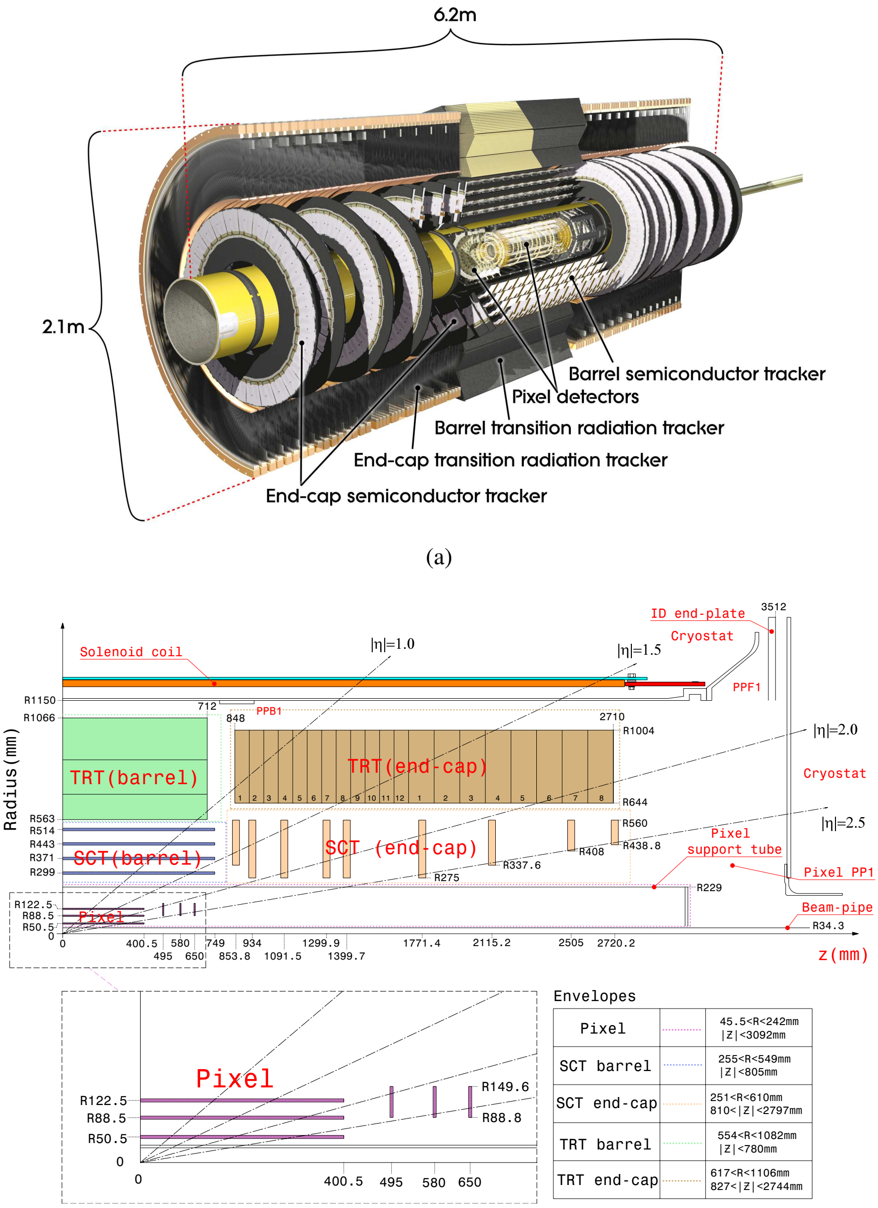
The solenoid magnet provides a 2 T axial magnetic field for the ID. It is a thin layer coil made of an Al-stabilized NbTi conductor. It is located inside the calorimeters and the flux of magnetic field is returned by the steel of the hadronic calorimeter. In order to suppress its impact on the energy measurement in the calorimeters, it is designed to keep the material thickness as low as possible, resulting in the solenoid assembly contributing to approximately 0.66 radiation lengths (X_0).

Toroidal magnet

The toroidal magnets are located outside the calorimeters and within the muon spectrometer. The barrel and two endcap toroids produce approximately 0.5 T and 1 T magnetic field, respectively. The conductor is based on a pure Al-stabilized Nb/Ti/Cu conductor.

2.2.3 Inner detector

The inner detector is composed of pixel trackers, silicon microstrip ("SCT"), and Transition Radiation Tracker ("TRT"). In the barrel region, they are arranged on concentric cylinders around the beam axis, while in endcap regions they are arranged on disks perpendicular to the beam axis. These are immersed in a 2 T magnetic field generated by the central solenoid. From a collision, approximately 1000 particles will emerge every 25 ns within $|\eta| < 2.5$, making a large number of tracks in the detector. To achieve enough momentum and vertex resolution, high-precision measurements must be made with fine detector granularity. The layout of the ID is illustrated in Figures 2.4.



(a)

(b)

Figure 2.4: The layout of the ATLAS inner detector [24]. (a) Three-dimensional view of the inner detector and (b) x-y plane view of its quarter-section. The beam pipe has one pixel-layer of an Insertable B-Layer (IBL).

Pixel tracker

The pixel tracker is the most inner part of the ID. There are three barrel layers in the barrel and 3×2 endcap disks in the forward and backward of the ID (Figure 2.5). The most inner barrel layer is referred to as "B-layer" or "Layer 0". B-layer is important to measure vertex position and tag long-lived heavy flavor particle like b-quarks. In addition to them, from Run2, a pixel barrel layer called as Insertable B-Layer ("IBL") is installed inside B-layer. It's attached to a new beam-pipe and it was simultaneously installed when the beam pipe was replaced to the new one. The cross-section of the new beam-pipe is shown in Figure 2.6. The main motivation of the IBL is to make the b-quark tagging efficiency higher.

The pixel tracker is designed to provide at least three precise measurement points for tracks with $|\eta| < 2.5$. The IBL and the three barrel layers lie at radii of 25.7, 50.5, 88.5 and 122.5 mm around the beam axis covering the central region up to $|\eta| = 1.9$. The three endcap disks lie at $|z|$ of 495, 580 and 650 mm in each of the forward and backward endcaps. The three barrel layers are segmented in R-ϕ and z. All pixel sensors in the three barrel layers and three disks are identical and have 47232 pixels of $50 \times 400 \mu\text{m}^2$ nominally or of $50 \times 600 \mu\text{m}^2$ in the regions at the front-end chips on a module. The size of the pixel in the IBL is $50 \times 250 \mu\text{m}^2$. The pixel sensor is composed of an oxygen-rich n-type bulk segmented into 47232 n^+-in-n pixels. The position resolution is 10μm (R-ϕ) in the barrel and endcaps, and 115μm in z-direction in the barrel or in R-direction in the endcaps.

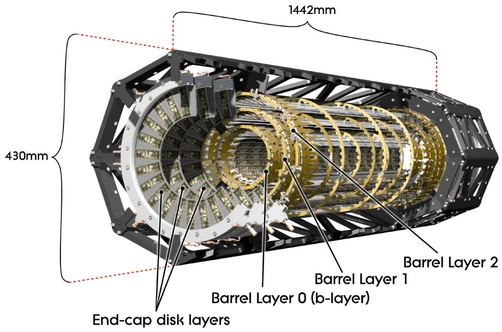


Figure 2.5: The overview of pixel tracker [26]. The Insertable B-Layer is not shown here. It's attached on the beam pipe.

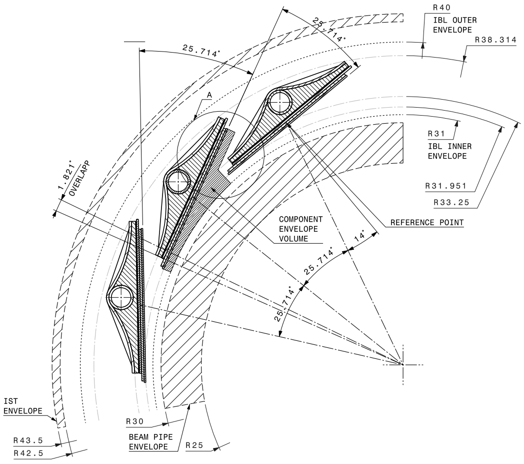


Figure 2.6: The cross-section of the new beam-pipe containing the Insertable B-Layer [26].

Semiconductor tracker

The SCT (Semiconductor tracker) is composed of four barrel coaxial cylindrical layers and 9×2 endcap disks. The barrel layers are located at radii 299, 371, 443 and 514 mm around the beam axis, covering the central region up to $|\eta| = 1.1$. The endcap disks lie in the $|z|$ range between 853.8 and 2720.2 mm, providing a total coverage up to $|\eta| = 2.5$. Barrel layers consist of tiling SCT modules which have 80 μm pitch micro-strip sensors. The module has the four sensors, each of which is composed of 768 microstrips. The two of them are rotated by +20 mrad and the other two are rotated by −20 mrad around the geometrical center of the sensors, which makes a small angle (40 mrad) between the strips. By this angle, one set of strips in each module can measure both coordinates of strip direction and its perpendicular one. The illustration of the module is shown in Figure 2.7. In the barrel layers, the strips are 6.4 cm long and parallel to the beam axis. In the endcap disks, the strips are running radially. The nominal resolution in both cases is 17μm in R-ϕ plane and 580μm in z in the barrel or in R in the endcaps.

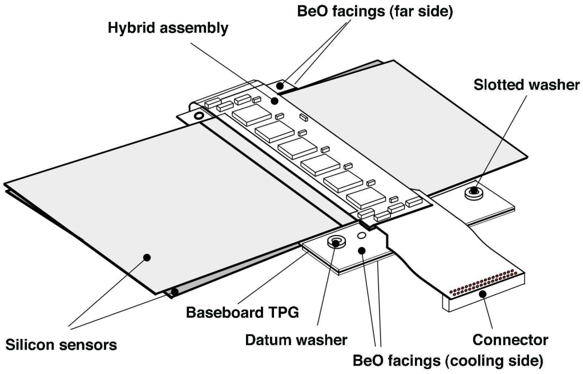


Figure 2.7: The illustration of SCT module in the barrel layers [24]. It has four sensors, each two on the top and bottom sides. The microstrips are running from the hybrid assembly to the outer (right and left) of the silicon sensors.

Transition radiation tracker

The TRT is a straw-tube tracker. Transition radiation is a phenomenon that a relativistic particle with charge ze emits photons when it crosses a boundary between two media with different dielectric constants. When a relativistic particle passes through the boundary between vacuum and the medium, it emits photons with the energy $\hbar\omega$,

$$\hbar\omega = \frac{\alpha z^2 \gamma \hbar\omega_p}{3}, \tag{2.4}$$

where α, γ, and ω_p are the fine structure constant, the Lorentz boost factor of the particle and a plasma frequency, respectively. The plasma frequency ω_p is defined as

$$\hbar\omega_p = \frac{m_e c^2}{\alpha}\sqrt{4\pi N_e r_e^3} = 28.81\ \mathrm{eV} \times \sqrt{\rho(\mathrm{in\ g/cm}^3)\langle Z/A\rangle} \tag{2.5}$$

where N_e, r_e and m_e are the electron density of the medium, the classical electron radius ($= \frac{e^2}{4\pi\epsilon_0 m_e c^2}$), and the electron mass, respectively. The ρ, Z and A are density, proton number and mass number of the medium. As shown in Eq. 2.4, an emitted energy is proportional to a Lorentz boost factor of the charged particle. Thus, to obtain enough hard photon to be detected, the particle needs to be boosted. For a particle with $\gamma = 10^3$, the radiated photons are in the soft x-ray range from 2 to 40 keV. Due to this feature, the transition radiation is used to distinguish electrons from other heavier charged particles.

The number of photons with energy $\hbar\omega > \hbar\omega_0$ is given by [27]

$$N_\gamma(\hbar\omega > \hbar\omega_0) = \frac{\alpha z^2}{\pi}\left[\left(\ln\frac{\gamma\hbar\omega_p}{\hbar\omega_0} - 1\right)^2 + \frac{\pi}{12}\right], \tag{2.6}$$

within corrections of order $\left(\hbar\omega_0/\gamma\hbar\omega_p\right)^2$. Thus, for example, the number of photons with energy $\hbar\omega > \hbar\omega_0 = \gamma\hbar\omega_p/10$ is $2.519\alpha z^2/\pi = 0.59\% \times z^2$. Since the photon yield from a single interface is low, in the TRT, it is enhanced by using a stack of many foil radiators separated by gas-filled gaps. The gas-filled gaps are formed in straw-tubes of 4 mm diameter. The straw-tube is made from multi-layer films of polyimide film, aluminum, and graphite-polyimide layers. The straw-tube works as cathodes and has an anode wire at the center, which is formed by a gold-plated tungsten with 31 μm diameter. The spaces between the straws are filled with polymer fibers in the barrel region and foils in the endcap regions as radiators. The straws are filled with Xe/CO_2/O_2 gases with the proportions of 70/27/3%. The xenon gas is used for a good X-ray absorption, and CO_2 and O_2 are used to provide a constant electron drift velocity and photon-quenching.

There are two kinds of signals from the TRT. One of them is the signals caused by the transition radiation from electrons. When an electron passes through the TRT, transition radiation photons are emitted with typical energy of 5–30 keV. These soft X-rays can be absorbed by the xenon gas, depositing energy in the gas and leading to significantly higher electric signals. The other kind of signals is caused by gas ionization. When a charged particle traverses the TRT, it ionizes the gas inside the tubes. The electrons generated by the ionization drift towards the wire and make signals. The TRT is designed so that charged particles with $p_T > 0.5$ GeV and $|\eta| < 2.0$ cross at least 32 straws, except in the transition region ($0.8 < |\eta| < 1.2$) where this number decreases to at least 22 straws.

The TRT detector is composed of a barrel and two endcap regions. The TRT barrel consists of three kinds of modules (Type-1, Type-2 and Type-3). There are 32 modules in each type and each module has hundreds of straws. The straws are parallel to the beam axis, and they occupy the region between $563 < r < 1066$ μm and $|z| < 712$ μm, corresponding to a coverage of $|\eta| < 0.7$. There are total 73 straw layers in the barrel. In the endcaps, there are two types of modules, called Type-A and Type-B. There are total 160 layers of straws in each TRT endcaps along the z-axis, and each layer contains 768 radially oriented straws of 37 cm length with uniform azimuthal spacing in ϕ. The TRT provides only R-ϕ information with the resolution of 130 μm per straw.

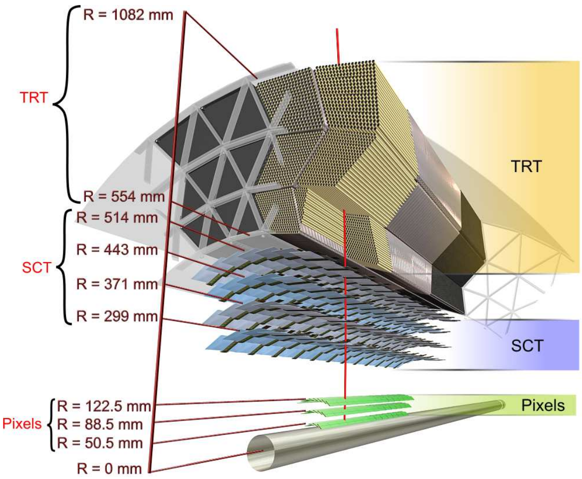


Figure 2.8: The illustration of the transition radiation tracker (TRT) and the other inner detector sub-systems in the barrel region [24].

2.2.4 Calorimeter

A schematic view of calorimeters is shown in Figure 2.9. The calorimeters cover the wide range ($|\eta| < 4.9$), using different techniques. There are two kinds of calorimeters in terms of techniques, Liquid Argon ("LAr") calorimeter and tile calorimeter. The calorimeters are also separated into electromagnetic ("EM") calorimeters and hadronic calorimeters in terms of their purpose. The EM calorimeters are aiming to measure electrons and photons, and the hadronic calorimeters are to measure hadronic jets. The EM calorimeters are the LAr calorimeter. For the hadronic calorimeters, the tile calorimeter and LAr calorimeter are used in the barrel region and endcap regions, respectively. Over the η region covered by the inner detector, fine granularity of the EM calorimeter is ideally suited for precision measurements of electrons and photons. For the rest of the calorimeters, coarser granularity is used, which is sufficient for jet reconstruction and measurement of missing transverse momentum $E_{\mathrm{T}}^{\mathrm{miss}}$.

Colorimeters must provide good containment for electromagnetic and hadronic showers, and must also suppress punch-through into the muon system. Hence, calorimeter depth is an important value. In the barrel, the total thickness is above 11 interaction lengths (λ) at $\eta = 0$. In the endcaps, the total thickness is above 10 λ.

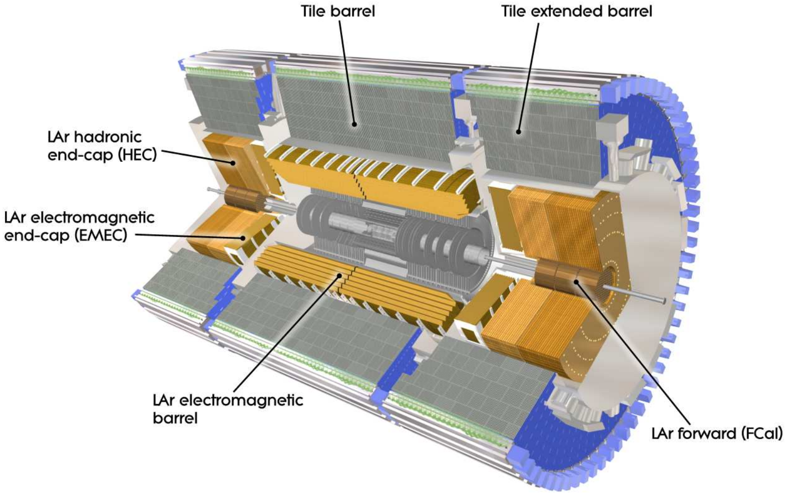


Figure 2.9: A schematic view of the calorimeters [24].

Electromagnetic calorimeter

There are two types of the EM calorimeter: LAr EM Barrel ("EMB") and LAr EM endcap ("EMEC")[1]. The EMB and the two EMECs are located in the barrel region and the endcap regions, respectively. The EM calorimeter is a lead-LAr detector with liquid Argon as active material, accordion-shaped electrodes, and lead absorber plates. Liquid Argon has been chosen due to its stability of response and its intrinsic radiation-hardness. The accordion geometry provides complete ϕ symmetry without azimuthal cracks at $z = 0$ and between the EMB and the EMECs. The readout electrodes are located in the gaps between the absorbers. The EM calorimeter has three layers and pre-sampler (only in $|\eta| < 1.8$) in front of the layers. The first layer is finely segmented in η for an accurate position measurement. The second layer is the thickest layer and collects the largest fraction of the energy of the electromagnetic shower, and the third layer collects only the tail of the electromagnetic shower. The segmentation of the calorimeter in η and in depth is obtained by etched

---

[1] A part of LAr Forward calorimeter ("FCal") is also used as the EM calorimeter, but it will be described in Section 2.2.4.

patterns on the different layers, and ϕ is segmented by different electrodes. The segmentation of the EMB is illustrated in Figure 2.10.

The EM calorimeter has the total thickness of > 22 radiation lengths (X_0) in the barrel and > $24X_0$ in the endcaps. The radiation length is the flight distance in which a high-energy electron loses its energy to $1/e$ of its original energy by bremsstrahlung, which is corresponding to 7/9 of the mean free path for pair production by a high-energy photon.

The energy resolution at energy E[GeV] is

$$\frac{\sigma}{E} = \frac{a}{\sqrt{E}} + \frac{b}{E} + c, \tag{2.7}$$

where the first term a is the stochastic term, the second term b is the electronics and pile-up noise term, and the third term c is the constant term. The first term a is a direct reflection of the statistical fluctuation. "a" is 0.08–0.11. "b" is, for instance, 0.4 GeV at the beam intensity of $1 \times 10^{34}\text{cm}^{-2}\text{s}^{-1}$ near $\eta = 0$. "c" comes from geometry uncertainties of the accordion shape and mechanics of the modules, which is kept to be below 0.7%. It is the most important resolution term in the energy measurement for high energy electrons and photons.

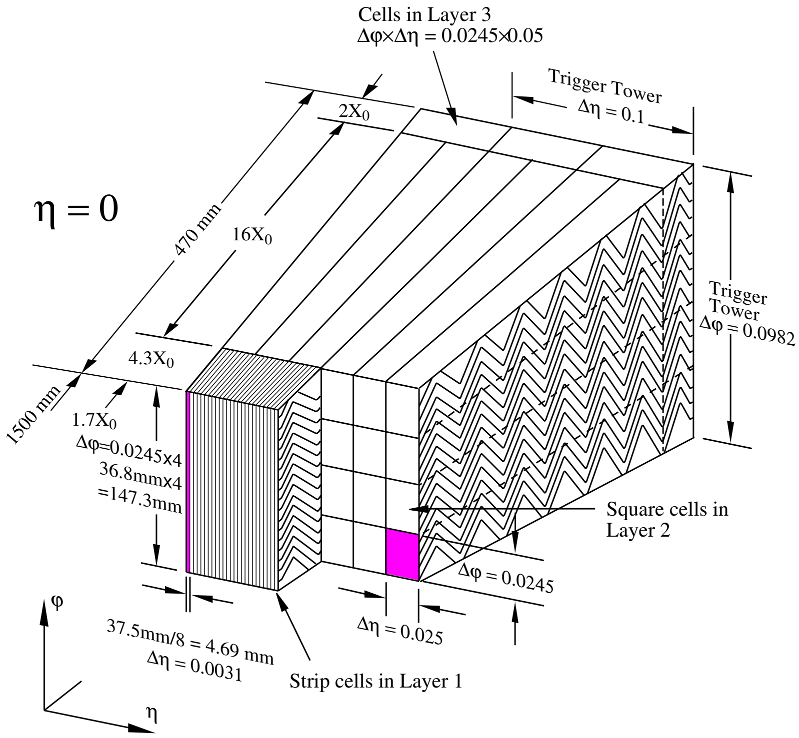


Figure 2.10: Segmentation of the LAr EM barrel calorimeter (EMB) [24]. The pre-sampler of the EMB is not shown here.

Data quality monitoring in the LAr calorimeter

The LAr calorimeter has approximately 1.8×10^5 cells making enough fine granularity for precise measurements of electrons and photons. Each cell measures an energy deposit in it from an electric signal pulse height. The height is obtained from four sampling points in the pulse as shown in Figure 2.11. The cells are not always stable throughout the data acquisition period and each noisy cell is masked collision-by-collision or throughout a run[2]. Thus, monitoring to find noisy cells from all the cells is necessary. The noisy cell finding is based on a factor called as "quality factor" given as

$$\begin{aligned} Q\text{−factor} &= \sum_{i=1}^{4} [s_i - P(t_i)]^2 \qquad (2.8) \\ P(t_i) &= A\,[g(t_i) + \xi_i - \tau g'(t_i)], \qquad (2.9) \end{aligned}$$

where s_i is i-th sampling height, t_i is i-th time, and $P(t)$ is predicted ideal height at a sampling time t. $P(t)$ is calculated from normalized prediction shape and its differentiation, $g(t_i)$ and $g'(t_i)$, which are obtained from a calibration pulse [28]. "A" and τ are a pulse height and a time shift, respectively. They are calculated from the four sampling points. ξ_i is a "residual correction" on the $g(t_i)$. The residual correction is introduced in order to reproduce real data sampling points more correctly, which is calculated from the difference between the prediction shape and averaged observed sampling shape in good quality data for each cell. A cell having sampling data with bad quality factor above 4000 is labeled as a bad cell and considered whether should be masked or not.

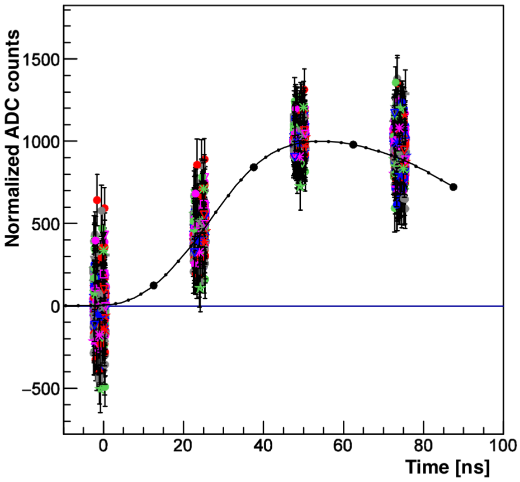


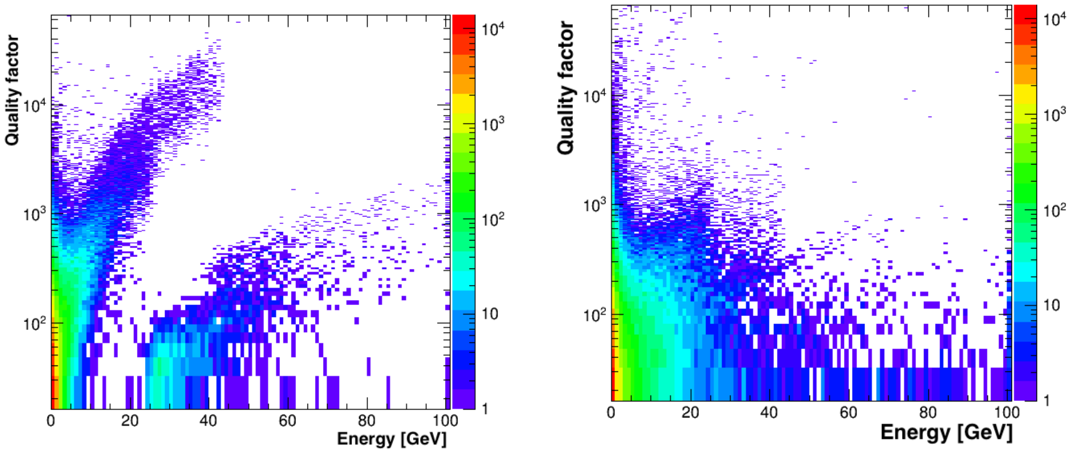
Figure 2.11: Sampling four points in several collisions in a LAr calorimeter cell and its predicted pulse shape (black line). The y-axis corresponds to electric voltage measured by ADC. The pulse height is normalized to 1000.

An old residual correction is obtained from the Run1 data. However, the correction is possibly invalid in Run2 due to the difference in the pileup condition between Run1 and Run2 caused by

---

[2] "Run" is defined as a period from injecting protons until bumping proton beam in the LHC (a few–10 hours).

increasing luminosity and shortening proton bunch interval. The quality factor with the old correction is checked in the good quality sampling data. The good data is obtained from $Z \rightarrow ee$ events. The Z events are selected by opposite-charge two electrons and Z mass window cut on the invariant mass of the electrons. The good sampling data is obtained from cells located in path of the two electrons in the Z events. Two-dimensional distribution of quality-factor with the old residual correction obtained from the Run1 data and deposit energy in the good sampling data in Run2 is shown in Figure 2.12(a). This figure indicates the old correction is not suitable for Run2 data because there are a lot of data with quality-factor above 4000 in the deposit energy from 10 to 40 GeV. Therefore, the residual correction is recomputed from the Run2 data in 2015. The two-dimensional distribution of quality-factor with the new residual correction and energy deposit in the same sampling data is shown in Figure 2.12(b). The data with large quality-factor in the energy range of 10 and 40 GeV is vanished in this figure by this new residual correction. This new residual correction can decrease mis-labeling as bad cell on good quality cell.



(a) Old residual correction

(b) New residual correction

Figure 2.12: Two-dimensional distribution of quality-factor and deposit energy [GeV] in good cells, in which the quality-factor is corrected by (a) old residual correction or (b) new residual correction. The good cell events are defined as sampling data of a cell located in the path of electrons coming from $Z \rightarrow ee$ decay. The old correction was obtained in Run1 and the new correction is obtained from the Run2 data.

Hadronic calorimeter

Jets and missing transverse momentum $E_{\mathrm{T}}^{\mathrm{miss}}$ are obtained mainly in the hadronic calorimeter. The ATLAS hadronic calorimeter consists of three types of detectors: the tile calorimeter, the LAr hadronic endcap calorimeter ("HEC"), and the LAr forward calorimeter ("FCal").

Tile calorimeter The tile calorimeter is a sampling calorimeter with 14 mm thick steel plates as absorber material and 3 mm thick plastic scintillators as active material. It is located outside of the EM calorimeter in the barrel region. The photons from the scintillators are detected by photomultipliers ("PMTs") as shown in Figure 2.13. The readout cells are built by grouping fibers into the photomultipliers and segmented as shown in Figure 2.14. It consists of central barrel ($|\eta| < 1.0$) and two extended barrels ($0.8 < |\eta| < 1.7$). It is segmented in depth in three layers, approximately 1.5, 4.1 and 1.8 λ thick for the barrel, and 1.5, 2.6 and 3.3 λ for the extended barrels.

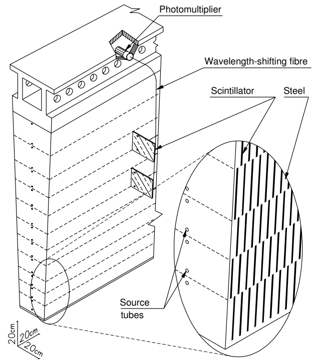


Figure 2.13: Schematic view of a Tile calorimeter module [24]. Each module covers 5.625 degrees in azimuth, and 64 modules are used in each z position to cover all ϕ range.

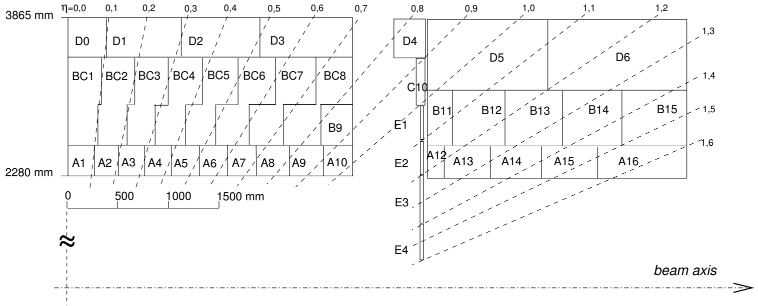


Figure 2.14: Segmentation of the Tile calorimeter in depth and η [24].

HEC The HEC is a copper/LAr sampling calorimeter with a flat-plate design, which covers the range $1.5 < |\eta| < 3.2$. It consists of two independent wheels per endcap: a front wheel ("HEC1") and a rear wheel ("HEC2"). It is located directly behind the EMECs. Each wheel is divided into two segments in depth. Hence, totally there are four segments per endcap. Each of the four HEC wheels is constructed of 32 identical wedge-shaped modules, as shown in Figure 2.15. The modules of the HEC1 have 24 copper plates, each 25 mm thick, plus a 12.5 mm thick front plate. For the HEC2, the modules are made of 16 copper plates, each 50 mm thick, plus a 25 mm thick front plate. The geometry of the HEC and FCal is shown in Figure 2.16.

FCal The FCal is the most forward calorimeter in the endcaps using LAr as active material. Its benefits are uniformity of the calorimetric coverage up to $|\eta| < 4.9$ and reduced radiation background levels in the muon spectrometer. It consists of three modules in each endcap; the first ("FCal1"), made of copper, is aiming at electromagnetic measurements (EM calorimeter), and the other two ("FCal2,3") are made of tungsten in order to measure the energy of the hadronic interactions (Hadronic calorimeters). Each module consists of a metal matrix, with regularly spaced longitudinal channels filled with the electrode structure consisting of concentric rods and tubes parallel to the beam axis as shown in Figure 2.17. Liquid Argon is filled in the gaps between the rod and the tube.

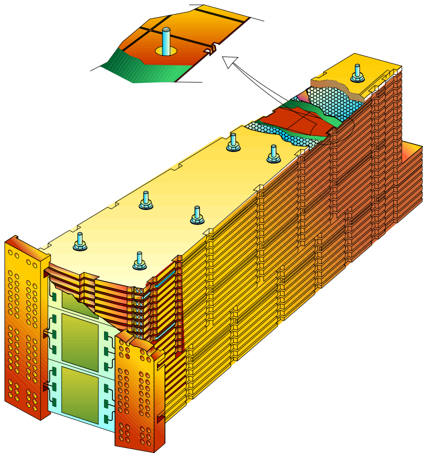


Figure 2.15: Illustration of a module of hadronic endcap calorimeters [24].

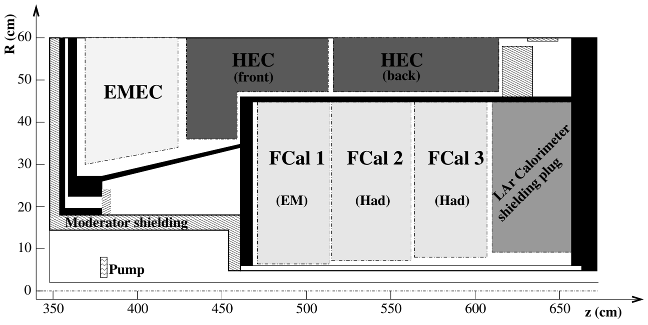


Figure 2.16: Geometry of hadronic endcap calorimeters and forward calorimeters [24].

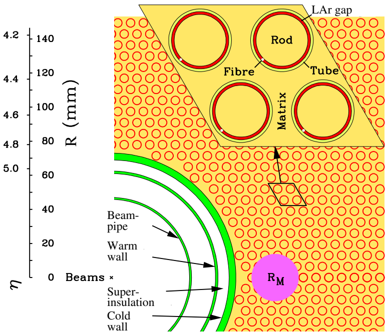


Figure 2.17: Electrode structure of forward calorimeter 1 (the most inner FCal) [24]. Molière radius, R_M, is also represented by a purple solid disk.

2.2.5 Muon spectrometer

The muon spectrometer ("MS") is the outermost part of the ATLAS detector, designed to detect muons which can pass through calorimeters and to measure their momenta in the pseudo-rapidity range of $|\eta| < 2.7$. It is also designed to trigger on these particles in the range of $|\eta| < 2.4$. In the barrel region, the muon chambers are located between and on the eight coils of the superconducting barrel toroid magnet, while, in the endcap regions, they are located in front and behind the two endcap toroid magnets. The chambers in the barrel are arranged in three concentric cylindrical shells around the z-axis at $r \approx$ 5m, 7.5m, and 10m. In each of the endcap regions, the chambers form four large wheels, perpendicular to the z-axis and located at $|z| \approx$ 7.4m, 10.8m, 14m, and 21.5m. Figures 2.18 and 2.19 show an overview of the MS and the positions of the muon chambers, respectively.

There are four types of chambers in the MS: Monitored Drift Tube ("MDT"), Cathode Strip Chamber ("CSC"), Resistive Plate Chamber ("RPC"), and Thin Gap Chamber ("TGC"). The MDT and CSC are aiming at precision measurement of the muon track coordinates (precision-tracking chambers), and the RPC and TGC are trigger chambers for three purposes: providing bunch-crossing identification, providing well-defined p_T thresholds, and measuring the muon coordinate in the direction orthogonal to that determined by the precision-tracking chambers.

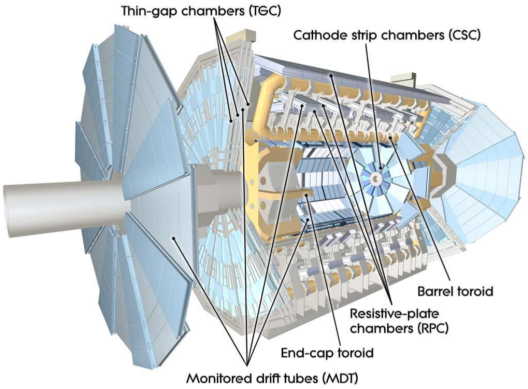


Figure 2.18: The overview of the muon spectrometer [24].

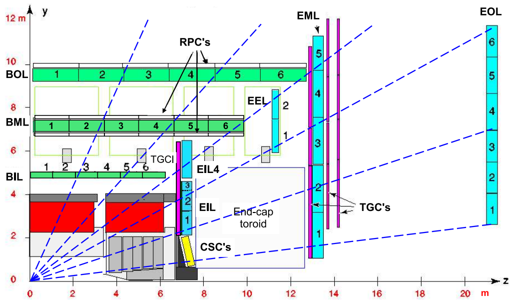


Figure 2.19: A cross-section of the muon chambers in a plane containing the beam axis [24]. The chambers with a name beginning with "B" or "E" represents Monitored Drift Tube ("MDT").

Precision-tracking chambers

The MDT chambers are located in both of the barrel (three cylindrical shells) and endcap (four wheels per endcap) regions, which cover the pseudo-rapidity range of $|\eta| < 2.7$. At large $|\eta|$, in order to withstand the high rate and background conditions, the CSC, which has high granularity, is used over $2 < |\eta| < 2.7$.

MDT The MDT chambers consist of two groups of three or four drift tube layers as shown in Figure 2.20, which achieve an average resolution of 80μm per tube, or approximately 35μm per chamber. The diameter of the drift tube is 29.970 mm. Ar/CO$_2$ (93/7%) gas at 3 bar is filled in it. The direction of tubes in the barrel and endcaps is along ϕ, i.e. the center points of the tubes are tangential to circles around the beam axis. Therefore, the MDT measure only z coordinate in the barrel and only r coordinate in the endcaps. For accurate measurement of the track coordinate, precision alignment of muon chambers is necessary. In the MDT chambers, an internal chamber alignment system with optical alignment rays is implemented, which continuously monitors potential deformations of the frame.

CSC The 16 CSC chambers form one wheel in each endcap as shown in Figure 2.21. The CSC is multi-wire proportional chamber with the wires oriented in the radial direction and two cathodes filled with Ar/CO$_2$(80/20%) gas. Both cathodes are segmented, one with the strips perpendicular to the wires and the other parallel to the wires. From these two cathodes, η and ϕ coordinates of tracks can be measured[3]. With strip width of 1.5–1.6 mm and the readout pitch of 5-6 mm in the cathodes perpendicular to the wires, the CSC reaches a resolution of 60μm in the radial direction.

---

[3] The CSC wire signals are not read out and cannot contribute to the coordinate measurement.

The cathodes parallel to the wires have coarser segmentation, resulting in a resolution of 5 mm in the ϕ-direction.

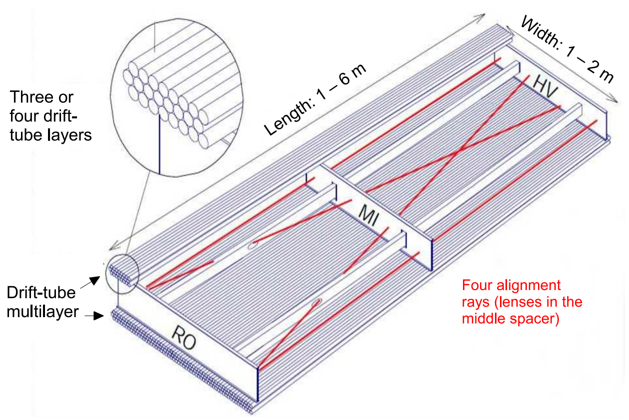


Figure 2.20: Structure of a Monitored Drift Tube chamber [24]. Four optical alignment rays, two parallel and two diagonal, allow for monitoring of the internal geometry of the chamber.

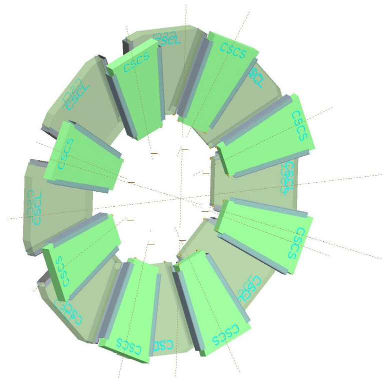


Figure 2.21: The layout of a wheel of Cathode-Strip Chambers [24].

Trigger chambers

The barrel region ($|\eta| < 1.05$) is covered by the three RPC layers. The two RPC layers (RPC1&2) sandwich the middle MDT layer, while the third one (RPC3) is located close to the outer MDT layer. The endcap regions ($1.05 < |\eta| < 2.4$) are covered by the four TGC layers. The three layers are in font of (TGCM1) and behind (TGCM2&M3) the second MDT wheel, while the fourth one (TGCI) is located in front of the innermost MDT wheel.

RPC The RPC is a gaseous parallel electrode-plate (i.e. no wire) detector. Two resistive plates (phenolic-melaminic plastic laminate) are kept parallel to each other at a distance of 2 mm by insulating spacers. The electric field between the plates is approximately 4.9 kV/mm. The signal is read out via capacitive coupling to metallic strips mounted on the resistive plates. The filled gas is a $C_2H_2F_4$/Iso-C_4H_{10}/SF_6 (94.7/5/0.3%). An RPC chamber has two independent detector layers, each of which has two orthogonal sets of pick-up strips with a width of 25–35 mm. Thus, RPC can measure both of the z and ϕ coordinates. The resolutions of these coordinates are 10 mm (RMS). The timing resolution is 1.5 ns.

TGC The TGC is multi-wire proportional chamber. The TGC chamber has two or three sets of two cathode planes and anode wire plane between the two cathode planes as shown in Figure 2.22. The wire-to-cathode distance is 1.4 mm, which is smaller than the wire-to-wire distance of 1.8 mm. The chamber is filled with a high quenching gas mixture of CO_2 and n-pentane (n-C_5H_{12}). The wires with diameter of 50 μm are arranged parallel to the MDT tubes (ϕ-direction) and the cathodes are segmented in ϕ-direction. By using these cathode strips, the TGC measures ϕ coordinate in order to complement the measurement of the MDT in the radial direction. The radial coordinate is also measured by the anode wires. The resolution of ϕ-direction is 3–7 mm (RMS), and that of radial direction is 2–6 mm (RMS). The timing resolution is 4 ns.

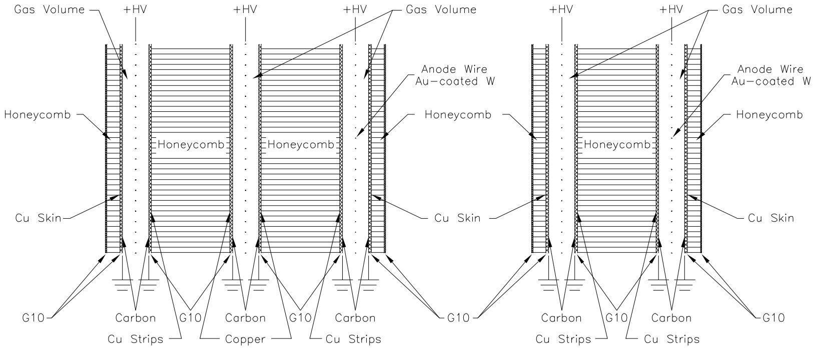


Figure 2.22: A cross-section of double- and triple-layer modules of Thin Gap Chambers [24].

2.2.6 Luminosity detector

"Luminosity measurement using Cerenkov integrating detector" ("LUCID") is a detector for online luminosity monitoring. It consists of twenty aluminum tubes which surround the beam pipe and point toward the interaction point. The tubes filled with C_4F_{10} at 1.2–1.4 bar, providing a Cerenkov threshold of 2.8 GeV for pions and 10 MeV for electrons. The detector is located at a distance of approximately ±17m from the IP in each of the endcap regions. The radial distance from the beam line is 10 cm. The Cerenkov light emitted by a passing particle is reflected on the inside polished walls of the tubes to be collected by PMTs located at the end of tubes.

Luminosity is calculated from number of signals from PMTs as

$$\mathcal{L} = \frac{\mu_{\mathrm{vis}} n_b f_r}{\sigma_{\mathrm{vis}}}, \tag{2.10}$$

where μ_{vis} is a mean number of interaction measured by the LUCID, f_r is a revolution frequency of the LHC ring, n_b is a number of bunch pairs colliding per revolution, and σ_{vis} is the total inelastic cross-section multiplied by the efficiency of the LUCID. σ_{vis} is measured in van der Meer (vdM) scans, in which the absolute luminosity can be inferred from direct measurements of the beam parameters. In the vdM scans, the beams are separated by steps of a specific distance, which allows a direct measurement of beam widths. σ_{vis} can be obtained as a calibration constant from the luminosity measured by the vdM scans and μ_{vis} measured by the LUCID at that time.

2.2.7 Trigger system

The trigger is system deciding whether to keep an event from a given bunch-crossing interaction or not. The Trigger and Data Acquisition ("TDAQ") system is shown in Figure 2.23. There are two main components of a hardware-based first-level trigger ("L1 trigger") and a software-based high-level trigger ("HLT").

The L1 trigger decision is performed by inputs from the L1 calorimeter (L1Calo) and L1 muon (L1Muon) triggers. The L1Calo trigger is based on the calorimeters and aiming to find electron/photon, tau, jet and missing transverse momentum by using signals exceeds a predefined threshold in each segmentations for trigger ("trigger towers") defined in the calorimeters. The trigger tower size is 0.1×0.1 in $\eta \times \phi$ as shown in Figure 2.10. The L1Muon trigger is based on the trigger chambers in the MS (the RPCs and TGCs). It requires a coincidence of hits in the different layers within a road that tracks the path of a muon from the IP through the detector.

The minimum time between two consecutive L1 accepts ("simple dead-time") is limited, and also the number of L1 accepts allowed in a given number of bunch-crossings ("complex dead-time") is restricted in order to avoid front-end buffers from overflowing. For instance, in 2015 run, the simple dead-time was set to 4 bunch-crossings (100 ns) and the complex dead-time was set to around 10 triggers in 350–400 bunch-crossings [29].

After the L1 trigger acceptance, the events are buffered in the Read-Out System (ROS) and processed by the HLT. The HLT receives Region-of-Interest (RoI) information from L1 triggers, which defines where should be considered to reconstruct particles in the HLT algorithms. After the events are accepted by the HLT, they are transferred to storage for offline reconstruction.

The maximum L1 trigger rate is limited to 100 kHz. The HLT trigger rate is $\sim$ 1 kHz in average.

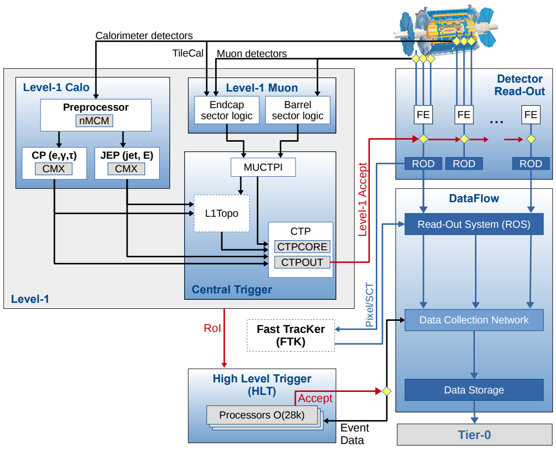
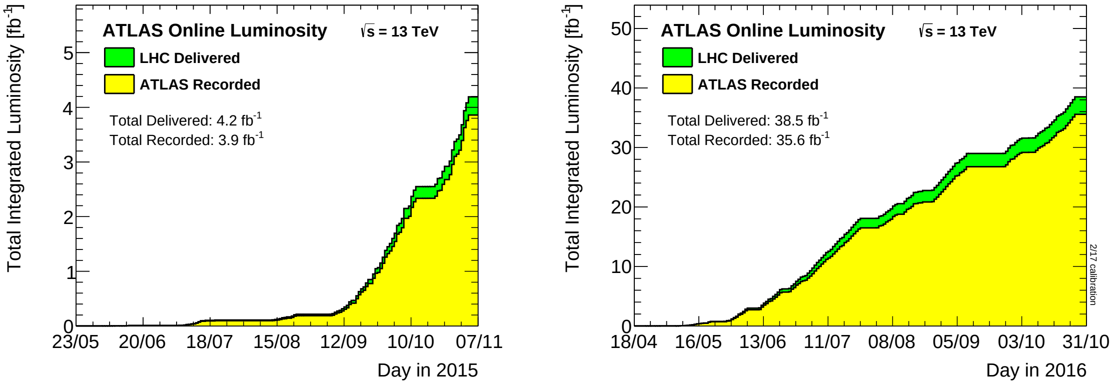


Figure 2.23: The flowchart of data in the Trigger and Data Acquisition system [29]. L1 Topo and FTK are being commissioned and not used.

2.3 Recorded data

2.3.1 Luminosity

The total integrated luminosity recorded in ATLAS is 39.5fb^{-1} in 2015+2016 run as shown in Figures 2.24. The peak luminosity in 2015 is $5.0\times10^{33}\text{cm}^{-2}\text{s}^{-1}$ and that in 2016 is $13.8\times10^{33}\text{cm}^{-2}\text{s}^{-1}$. Since 2015 was the first year of Run2, the LHC didn't run as it was designed, but it ran stably in 2016 and reached the design luminosity of $1\times10^{34}\text{cm}^{-2}\text{s}^{-1}$. The peak luminosity in each LHC p-p fill is shown in Figures 2.25. For physics analysis, detector conditions have to be good. Thus, time periods with high inefficiency due to not-working detector modules were dropped from the data for physics analysis. In this analysis, $36.07\pm0.03\text{fb}^{-1}$ data recorded in 2015 and 2016 with enough good detector-condition is used.



(a) 2015

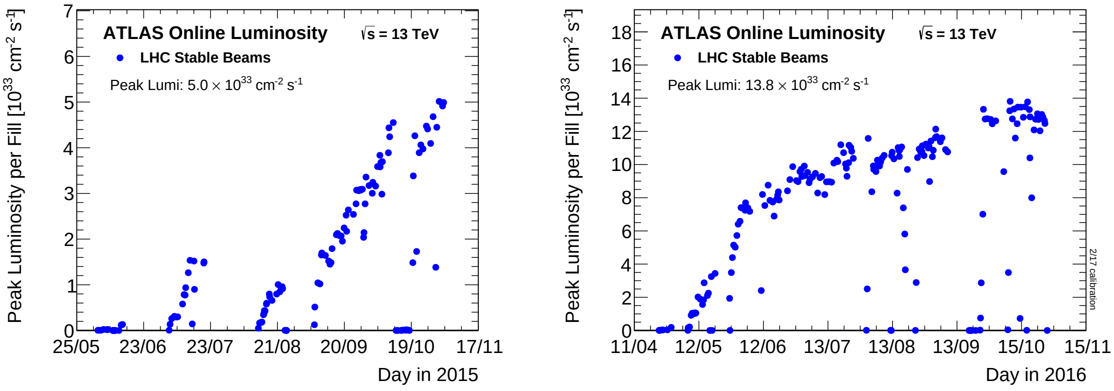
(b) 2016

Figure 2.24: The total integrated luminosity [fb^{-1}] delivered to and recorded with the ATLAS during stable beams for p-p collisions at the center-of-mass energy $\sqrt{s}$ = 13 TeV in (a) 2015 and (b) 2016 [30].

2.3.2 Pileup

In the p-p collisions, there are not only interesting hard interactions but also many non-interesting inelastic interactions. The cross-section (σ_{inel}) of the inelastic interaction is about 80 mb for 13 TeV collisions [30]. At the design luminosity of $1\times10^{34}\text{cm}^{-2}\text{s}^{-1}$, the average number of inelastic events is a few of tens per bunch crossing. Figure 2.26(a) shows the luminosity-weighted distribution of the mean number (μ) of interactions per bunch crossing for the 2015 and 2016 p-p collision data at 13 TeV. The "mean" means that the μ is not the number of real occurred interactions in each bunch crossing but it is the mean of the Poisson distribution of the number of interactions per bunch crossing. It is calculated from the instantaneous per bunch luminosity $\mathcal{L}_{\text{bunch}}$ measured by the luminosity detector as

$$\mu = \mathcal{L}_{\text{bunch}}\frac{\sigma_{\text{inel}}}{f_r}, \tag{2.11}$$



(a) 2015

(b) 2016

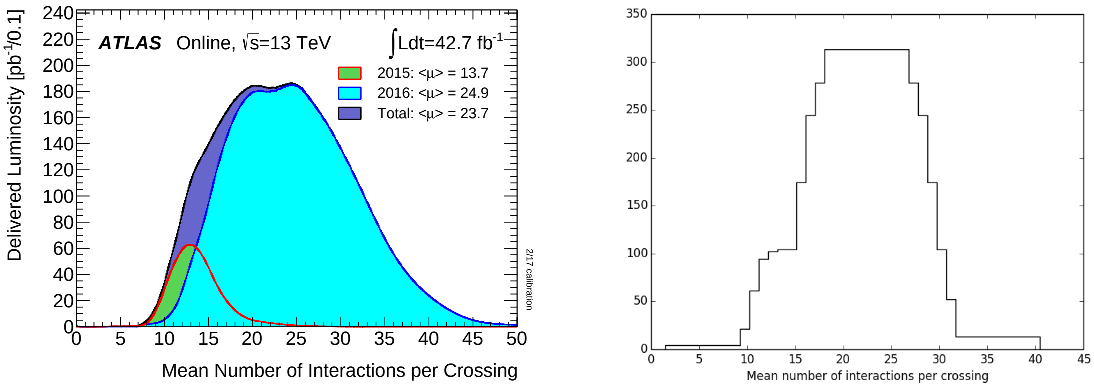
Figure 2.25: The peak instantaneous luminosity [$10^{33}\text{cm}^{-2}\text{s}^{-1}$] delivered to ATLAS during stable beams for p-p collisions at the center-of-mass energy $\sqrt{s}$ = 13 TeV in (a) 2015 and (b) 2016 [30].

where f_r is the LHC revolution frequency. These non-interacting inelastic interactions are called as "pileup". These interactions are superimposed to an interesting physics process in Monte Carlo ("MC") simulation. The distribution of the μ in the MC is shown in Figure 2.26(b). The pileup is composed of "in-time pileup" and "out-of-time pileup". The in-time pileup is an interaction caused by the other protons in the same bunch in which the interesting physics process occurs. It is simulated by superimposing a simulation of additional p-p collisions at an appropriate rate during the bunch-crossing. The out-of-time pileup is caused by bunch-crossings before or after the bunch-crossing of interest because detector response time is longer than bunch-crossing period (25 ns). Thus, additional p-p collision simulation for pileup is overlaid also before and after the bunch-crossing and the time window in which the simulation is overlaid is varied in each detector depending on the response time of each detector. Such a pileup simulation is overlaid as additional energy deposits in each detector before the conversion from energy to detector signal (digitization step) is made.

In order to match μ distributions between the data and MC, a weight in each event of the MC simulation is reweighted ("pileup reweighting") [4].

---

[4] To be precise, the distribution of the μ in data used for pileup reweighting is different from Figure 2.26(a). The practically used one is the distribution of the luminosity-weighted average of the μ over all of the bunches in each Lumi Blocks ("LBs"). The LB is a unit used in ATLAS as minimum data taking time period of about 1 minute. Since this averaging removes smearing on the μ over bunches, the distribution is shaper than Figure 2.26(a).



(a)

(b)

Figure 2.26: (a) The luminosity-weighted mean number (μ) of interactions per bunch crossing in data, and (b) the distribution of the μ in Monte Carlo before pileup reweighting [30].

3 Physics object reconstruction and identification

The measured energy deposit in each detector is reconstructed to be objects: tracks, clusters, jets, b-jets (jets originating from b-quarks), electrons, photons, muons, and missing transverse momentum $E_{\mathrm{T}}^{\mathrm{miss}}$. In this section, the algorithms to reconstruct tracks, clusters, jets, and $E_{\mathrm{T}}^{\mathrm{miss}}$ are explained. Electrons, photons, muons, and b-jets are used only for the background estimation (Section 8.2). Their reconstruction and identification are described in Appendix A.

3.1 Track

3.1.1 Track reconstruction

In the track reconstruction for a charged particle in the Inner Detector (ID), there are several strategies as described in [31, 32]. The main strategy is called as inside-out track finding, which has the following steps:

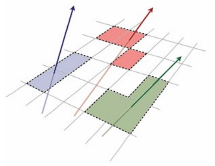


Figure 3.1: Illustration of single-particle pixel clusters in the pixel detector. Each color shows each cluster that has energy deposit from single charged particle represented by an arrow.

1. **Space point formation**
   The initial step is clustering hits (Figure 3.1) and transforming the clusters in the pixel and SCT detectors into three-dimensional measurements called as "space points".

2. **Seed track finding and extension in the SCT**
   The second step is finding seed tracks. The seed track is built from three space points in the pixel and SCT layers. These seeds are then extended throughout the silicon detectors by choosing additional space-points in the seed track direction from the remaining layers, in which a combinatorial Kalman filter [33] algorithm is used. This extended trajectory is referred to as a track candidate.

3. **Ambiguity solving**
   The collection of the track candidates has random hits combinations referred to as "fake tracks" or duplicated tracks which have a shared segment with another track candidate. In this step, these ambiguity is resolved based on a "track score" that is calculated from fit quality in the Kalman filter and the number of shared hits and holes. The hole means a missing cluster, which is expected there, in the road of the track candidate.

4. **Extension into the TRT**
   Track candidates passing the ambiguity solving step are extended into the TRT.

Then, these reconstructed tracks are identified by track quality selection. There are two sets of selection of "Loose" and "TightPrimary":

Loose

- $p_{\mathrm{T}} > 400$ MeV
- $|\eta| < 2.5$
- Number of pixel and SCT clusters on track (referred to as "silicon hits") ≥ 7
- Number of shared clusters in the pixel and SCT ≤ 1
- Number of pixel and SCT holes ≤ 2
- Number of pixel holes ≤ 1

TightPrimary
This requires the following selection in addition to the Loose selection

- Number of silicon hits ≥ 9 if $|\eta| \leq 1.65$
- Number of silicon hits ≥ 11 if $|\eta| \geq 1.65$
- At least one hit on one of the two innermost pixel layers
- No pixel holes
- $|d_0^{\mathrm{BL}}| < 2.0$mm
- $|z_0^{\mathrm{BL}} \sin\theta| < 3.0$mm

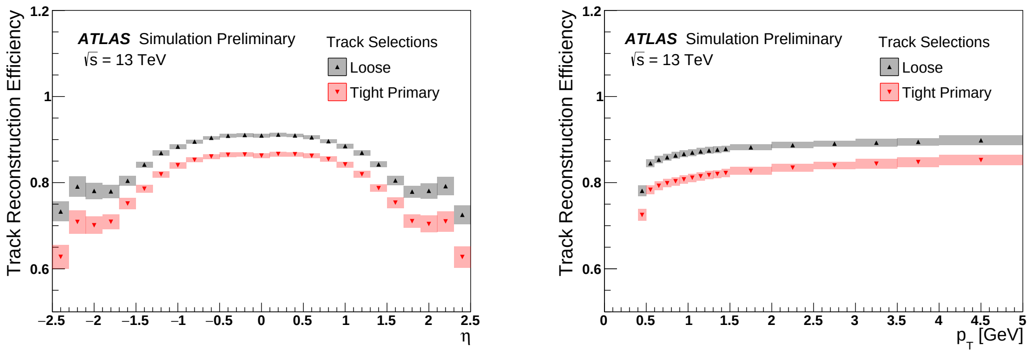
d_0^{BL} is the transverse distance between the track and the measured beam-line position, z_0^{BL} is the longitudinal distance along the beam between the track and the primary vertex (described in Section 3.1.3), and θ is a polar angle of the track.

3.1.2 Track reconstruction efficiency and fake rate

The Loose track selection is a default track selection used to reconstruct jets, electrons, and muons , which has high efficiency but also has a non-negligible fraction of fake tracks. The fake track is a track reconstructed by a mis-combination of hits of several different particles. The TightPrimary selection aims to reconstruct primary tracks, which are defined as tracks of a charged particle with a mean lifetime $\tau \geq 300$ ps. The TightPrimary is optimized to reject fake tracks, but it has more track reconstruction inefficiency than the Loose selection as shown in Figures 3.2. The efficiency at $\eta = 0$ and $p_{\mathrm{T}} > 5$ GeV is $\sim 90\%$ and $\sim 85\%$ for the Loose and the TightPrimary selections, respectively. The rate of the fake tracks is estimated from the behavior of the averaged number of tracks $\langle N_{\mathrm{tracks}}(\mu)\rangle$ in data as a function of a number of pileup μ (Figure 3.3(a)). Here, assuming that the number of real (i.e. non-fake) tracks is proportional to the μ to first order, the number of real tracks is estimated by the fit with a linear function $f(\mu) = m\mu$ in the range of $10 \geq \mu \geq 15$. The fake track rate $R_{\text{fake}}(\mu)$ is given as

$$R_{\text{fake}}(\mu) = \frac{\langle N_{\text{tracks}}(\mu)\rangle - f(\mu)}{\langle N_{\text{tracks}}(\mu)\rangle}. \tag{3.1}$$

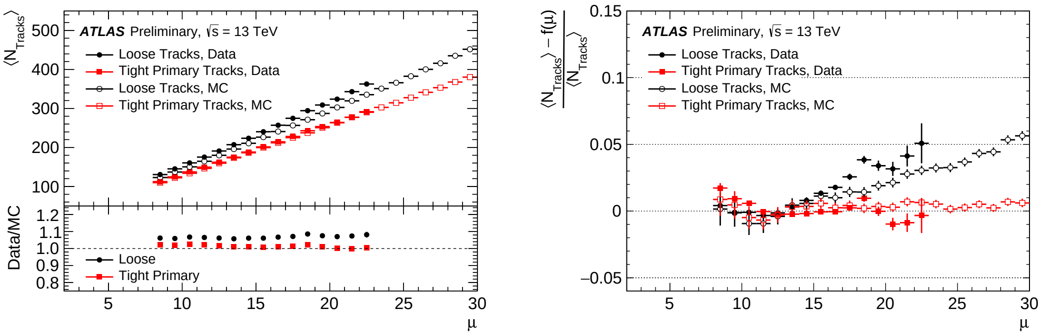
This is shown in Figure 3.2(b). The fake rate of the TightPrimary selection is almost zero at any number of pileup, but that of the Loose selection is increasing to 5% at $\mu = 22$ in the data.



(a)

(b)

Figure 3.2: Track reconstruction efficiency calculated in the simulation of the Loose and TightPrimary selections as a function of (a) η and (b) p_{T} [32]. The bands show the total systematic uncertainty.



(a)

(b)

Figure 3.3: Average numbers of tracks $\langle N_{\text{tracks}}(\mu)\rangle$ in the data and simulation as a function of a number of pileup μ in the Loose and TithtPrimary selections (a), and the fake track rates estimated from the linear fits (b) [32].

3.1.3 Vertices

Vertices are reconstructed from the tracks by using their z coordinates along the beamline. A vertex candidate is reconstructed by Iterative Vertex Finding [34] procedure. All tracks are refitted to the vertex candidates and it is determined which track is associated with the vertex. The vertex with the largest sum of squared transverse momenta of the associated tracks is defined as a "primary vertex".

3.2 Topological cluster

A topological cluster ("topo-cluster") is a group of cells with energy deposit in the calorimeters, which is used to reconstruct a jet. There are two steps to reconstruct the topological cluster as described in the following:

1. **Proto-cluster formation**
   At first, the grouping of cells into "proto-clusters" is performed based on the cell signal significance $\varsigma_{\text{cell}}^{\text{EM}}$ in each cell defined as,

   $$\varsigma_{\text{cell}}^{\text{EM}} = \frac{E_{\text{cell}}^{\text{EM}}}{\sigma_{\text{noise,cell}}^{\text{EM}}}, \tag{3.2}$$

   where $E_{\text{cell}}^{\text{EM}}$ is the energy measured in the cell on the electromagnetic (EM) energy scale, and $\sigma_{\text{noise,cell}}^{\text{EM}}$ is an expected noise level in the cell. The noise level is varied in $O(10)$–$O(10^3)$ MeV depending on the calorimeter position and pileup [35]. The EM scale energy is defined by an energy calibration reproducing the energy deposited by electrons and photons correctly, in which there is no correction on the loss of signal for hadrons. Proto-clusters are formed by a growing-volume algorithm starting from a calorimeter cell with a highly significant seed signal. The seeding and growing are controlled by three parameters $\{S, N, P\}$, which define three kinds of threshold on the $\varsigma_{\text{cell}}^{\text{EM}}$ as $|\varsigma_{\text{cell}}^{\text{EM}}| > S, N, P$. The values of $\{S, N, P\}$ are set to $\{4,2,0\}$ in this proto-clustering. The seeds are defined as the cells with $|\varsigma_{\text{cell}}^{\text{EM}}|$ above S. A found seed is growing by adding three-dimensional neighbor cells with $|\varsigma_{\text{cell}}^{\text{EM}}| > N$ into the proto-cluster of the seed. This growth is stopped when there is no neighbor cell passing this requirement. At the last, the neighbor cells with $|\varsigma_{\text{cell}}^{\text{EM}}| > P$ are added to the proto-cluster. This clustering for each seed is made in $|\varsigma_{\text{cell}}^{\text{EM}}|$-ordering of the seeds.

2. **Cluster splitting**
   The proto-cluster is too large to provide a good measurement of the energy from one particle because spatial signal structures inside the cluster are not considered. Therefore, as a next step, the cluster is split based on local signal maximum cells. The local signal maximum is defined by $E_{\text{cell}}^{\text{EM}} > 500$ MeV and topological requirements that there is no neighbor with larger energy and at least four neighbor cells inside the cluster. The split clusters are the topological clusters.

(ϕ, η) of the topological cluster is defined as a simple energy-weighted sum of the (ϕ, η) of the inside cells. The energy of the cluster is given as the sum of deposit energy of the cells.

3.3 Jet

A jet is a bundle of particles produced by a hadronization of a parton, which is observed as a shower in the calorimeters. A jet reconstructed from the topo-clusters is referred to as "calo jet". The tracks associated with the jet has information of the parton origin of it. An identification of the jet originating from b-quark using track information ("b-tagging") is described in Appendix A.1. The track information inside a jet to separate light-quark (u, d, c, and s) and gluon jets is described in Section 6. It is a key technique to improve this gluino search.

3.3.1 Jet reconstruction

In this analysis, anti-k_t algorithm [36] with $R = 0.4$ is used to reconstruct jets. The R is a parameter used in the algorithm and indicates a typical jet cone size in the $\phi - \eta$ plane. In the anti-k_t algorithm, the following variable is defined between two topo-clusters i and j:

$$d_{i,j} = \min\left(p_{T,i}^{-2}, p_{T,j}^{-2}\right)\frac{\Delta R_{i,j}^2}{R^2} \tag{3.3}$$

$$d_{i,B} = p_{T,i}^{-2}, \tag{3.4}$$

where $\Delta R_{i,j}$ is a distance between i and j in the $\phi - \eta$ plane. The $d_{i,j}$ of all combinations of input clusters are calculated. The two clusters with the minimum $d_{i,j}$ among them and the $d_{i,B}$ are merged into one. If the minimum one is $d_{i,B}$, the cluster i is considered as a jet and removed from the input list. This merging step is repeated until no cluster is left in the list. Truth jets are also reconstructed by the anti-k_t algorithm using truth particles in the simulation, which is used in the jet energy calibration as a reference described in Section 3.3.2.

3.3.2 Jet energy calibration

To calibrate a jet momentum from the EM-scale momenta of the topo-clusters, there are three steps:

1. Pileup correction
2. Absolute jet energy scale ("JES") calibration
3. *In situ* calibration (only to data)

Pileup correction

The pileup correction removes the additional energy caused by the in-time and out-of-time pileups. The correction has two components of an area-based p_T density subtraction, in which a ghost-association technique [37] is used, and a residual correction obtained from the simulation.

- **Area-based p_T density subtraction**
  This method subtracts pileup contribution to the p_T of each jet according to its area in each event. The pileup contribution is calculated from the median p_T-density $\rho = \mathrm{median}\{p_\mathrm{T}^{\mathrm{jet},j}/A^{\mathrm{jet},j}\}$, where $A^{\mathrm{jet},j}$ is an area of the jet j and the index j enumerates the jets reconstructed by the k_T algorithm [38] with $|\eta| < 2$, which is chosen due to its good sensitivity to soft radiation and used only in this correction. In this method, simulated "ghost" particles with infinitesimal

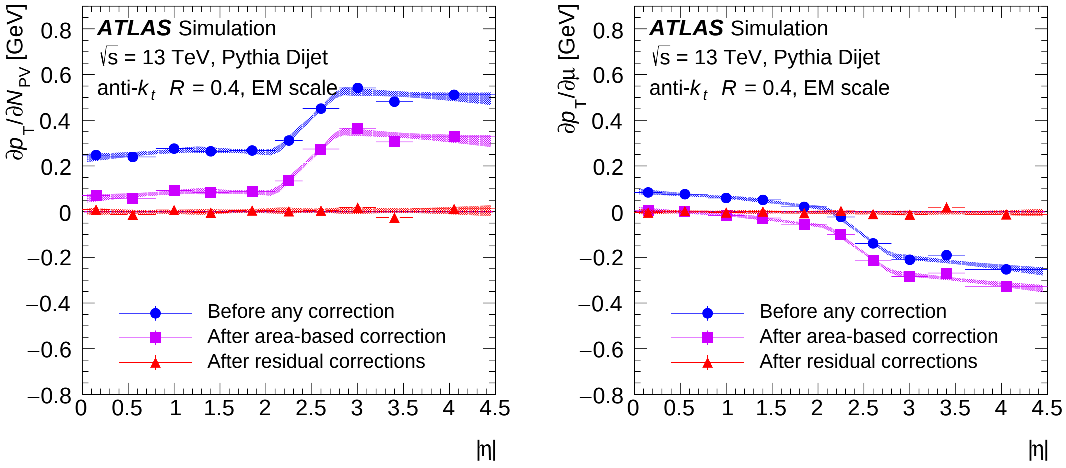
momentum are added uniformly in solid angle to the events before jet reconstruction. The number of ghost particles N_g^j included in the jet j in the reconstruction is used to calculate the $A^{\mathrm{jet},j}$ given as $A^{\mathrm{jet},j} = N_g^j / A_g$, where A_g is the number of ghost particles in a unit area. From the obtained median ρ, the p_{T} subtraction for the jet j is defined as $-\rho \times A^{\mathrm{jet,j}}$.

- **Residual correction**
  Residual correction is necessary because only the central jets are considered in the ρ calculation and some dependence of the anti-k_t jet p_{T} on the pileup amount can be seen after the area-based p_{T} subtraction. There are dependences of the p_{T} on the number of primary vertices N_{PV}[1] and on the number of interactions per bunch-crossing μ, which are sensitive to the in-time and out-of-time pileups, respectively. These dependences are linear and independent of each other. Therefore, they are separately evaluated by the linear fits in the simulation in each (p_{T},η) range.

Both corrections are described in one equation given as

$$p_{\mathrm{T}}^{\mathrm{corr}} = p_{\mathrm{T}}^{\mathrm{reco,EM}} - \rho \times A^{\mathrm{jet},j} - \alpha(p_{\mathrm{T}},\eta) \times (N_{\mathrm{PV}} - 1) - \beta(p_{\mathrm{T}},\eta) \times \mu, \tag{3.5}$$

where $p_{\mathrm{T}}^{\mathrm{reco,EM}}$ is p_{T} at the EM scale before any correction, and $\alpha(p_{\mathrm{T}},\eta)$ and $\beta(p_{\mathrm{T}},\eta)$ are the coefficients in the residual correction depending on N_{PV} and μ, respectively, in each p_{T} and η range. The dependencies before and after corrections in the simulation are shown in Figures 3.4.



(a) (b)

Figure 3.4: Dependencies of the EM scale p_{T} on (a) the number of primary vertices N_{PV} (averaged over μ) and (b) pileup (averaged over N_{PV}) as a function of $|\eta|$ [39]. Blue, purple, and red points show the dependence before any correction, after area-based correction, and after all the pileup corrections, respectively.

Absolute JES calibration

The absolute jet energy scale calibration transfers the reconstructed jet momentum to the particle level momentum. This calibration is made by using the Pythia Monte Carlo ("MC") simulation after

---

[1] The primary vertex here means a vertex close to the beam axis in the transverse plane.

the pileup corrections. In the MC simulation, the Gaussian fit is performed on the

$$R = \frac{E_{\mathrm{T}}^{\mathrm{jet}}}{E_{\mathrm{T}}^{\mathrm{truth}}}, \tag{3.6}$$

where $E_{\mathrm{T}}^{\mathrm{reco}}$ and $E_{\mathrm{T}}^{\mathrm{truth}}$ are the reconstructed transverse energy and truth-level p_{T} of a jet. This fit is performed in each $p_{\mathrm{T}}^{\mathrm{truth}}$ and η_{det}[2] range, and the energy response $\langle R \rangle$ defined as the mean of the Gaussian fit is used to calibrate the absolute jet energy scale. The $\langle R \rangle$ in each η_{det} is shown in Figure 3.5.

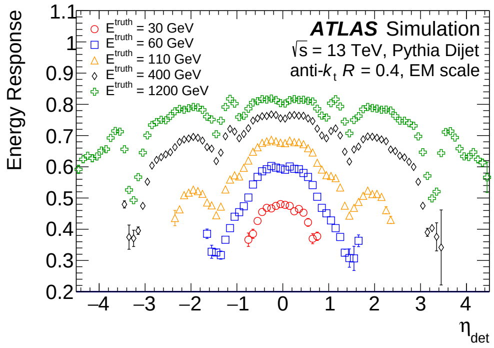


Figure 3.5: Energy response as a function of η_{det} for a different truth-level transverse energy of jets ($E_{\mathrm{T}}^{\mathrm{truth}}$) [39].

In situ calibration (only to data)

The last correction is *in situ* calibration of data. This correction is aiming to compensate the difference between the data and MC simulation. The response $\mathcal{R}_{in-situ}$ is defined in the data and MC simulation as the average ratio of jet p_{T} to reference object p_{T}, binned in range of the reference object p_{T}. The response is obtained in $Z\,(\rightarrow ll)$+jet, γ+jet, and multi-jet events. Z, γ and the vector sum of recoiling jets are treated as reference objects in each case. The recoiling objects are the other jets than the leading jet in the multi-jet events. The correction factor is obtained as

$$c = \frac{\mathcal{R}_{in-situ}^{\mathrm{data}}}{\mathcal{R}_{in-situ}^{\mathrm{MC}}}. \tag{3.7}$$

The combined factor calculated from Z+jet, γ+jet, and multi-jet events is shown in Figure 3.6.

3.3.3 Uncertainties on jet

Uncertainties on the jet energy scale coming from all of the above calibrations are shown in Figures 3.7. The total uncertainty strongly depends on the jet p_{T} and varies from 1% to 4%.

---

[2] The η_{det} is defined to η pointing to the geometric center of the detector in order to avoid any ambiguity as to which region of the detector is measuring the jet.

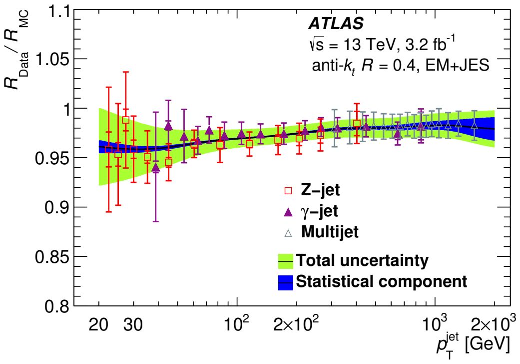
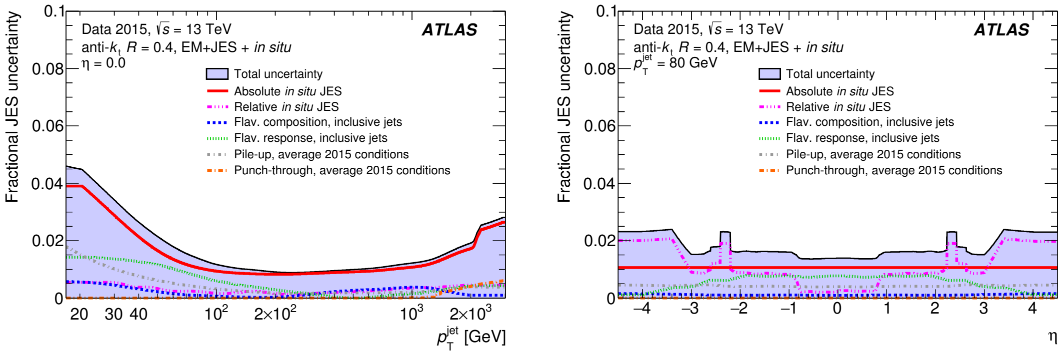


Figure 3.6: Ratio of the jet response in data to that in the nominal MC simulation after the absolute jet energy scale calibration as a function of jet p_T for Z+jet, γ+jet, and multi-jet *in situ* calibrations [39].



(a) (b)

Figure 3.7: Combined uncertainty on the jet energy scale after all of the calibrations as a function of (a) jet p_T and (b) jet η at p_T = 80 GeV [39].

3.3.4 Requirements for jet in this analysis

Jets used in this analysis in apart from jets utilized for the calibration of a quark/gluon separation variable in Section 6 are required the following selection:

- $p_T > 50$ GeV
- $|\eta| < 2.8$

The requirements on jets in Section 6 are described in that section.

3.4 Missing transverse momentum E_T^{miss}

In searches for the SUSY particles, the most important feature of the signals is a large missing momentum derived from the undetectable $\tilde{\chi}_1^0$. In p-p collisions, the initial momentum in the beam direction is unknown since the momentum contributing to the collision is parton momentum, which is a part of proton momentum and changed in each collision. However, the initial momentum in the plane transverse to the beam direction is zero. Therefore, a missing momentum of the collision products in the transverse plane can be calculated as a negative sum of visible transverse momenta. In practice, the missing transverse momentum $E_{\mathrm{T}}^{\mathrm{miss}}$ is calculated from the hard physics objects (jets, electrons, muons, and photons) and "track-based soft term" (TST). The TST is a remaining energy deposit calculated from the reconstructed tracks associated with the primary vertex but not associated with the hard objects. The hard object is fully calibrated with an algorithm dedicated to each kind of the object. Thus, the TST is a key term for a good $E_{\mathrm{T}}^{\mathrm{miss}}$ performance.

3.4.1 $E_{\mathrm{T}}^{\mathrm{miss}}$ reconstruction

The $x(y)$ component of the $E_{\mathrm{T}}^{\mathrm{miss}}$ is given as

$$E_{x(y)}^{\mathrm{miss}} = E_{x(y)}^{\mathrm{miss},e} + E_{x(y)}^{\mathrm{miss},\mu} + E_{x(y)}^{\mathrm{miss},\gamma} + E_{x(y)}^{\mathrm{miss,jets}} + E_{x(y)}^{\mathrm{miss,soft}}, \tag{3.8}$$

where $E_{x(y)}^{\mathrm{miss},e,\mu,\gamma,\mathrm{jets}}$ is the negative vector sum of the momenta of the respective calibrated objects and the $E_{x(y)}^{\mathrm{miss,soft}}$ is the TST. The magnitude $E_{\mathrm{T}}^{\mathrm{miss}}$ and azimuthal angle ϕ^{miss} of the $\vec{E}_{\mathrm{T}}^{\mathrm{miss}}$ are calculated as

$$E_{\mathrm{T}}^{\mathrm{miss}} = \sqrt{\left(E_x^{\mathrm{miss}}\right)^2 + \left(E_y^{\mathrm{miss}}\right)^2}, \tag{3.9}$$

$$\phi^{\mathrm{miss}} = \arctan\left(\frac{E_y^{\mathrm{miss}}}{E_x^{\mathrm{miss}}}\right). \tag{3.10}$$

The $E_{x(y)}^{\mathrm{miss,soft}}$ is calculated from tracks satisfying the "TightPrimary" selection described in Section 3.1 but not associated with any hard object.

3.4.2 $E_{\mathrm{T}}^{\mathrm{miss}}$ performance

The $E_{\mathrm{T}}^{\mathrm{miss}}$ performance is checked in $Z \to ll(Zll)$ events and $W \to l\nu(Wl\nu)$ events. The Zll events are used as no $E_{\mathrm{T}}^{\mathrm{miss}}$ events, and $E_{\mathrm{T}}^{\mathrm{miss}}$ in the $Wl\nu$ events are considered as "genuine" $E_{\mathrm{T}}^{\mathrm{miss}}$. There are three important variables to check the performance:

- $E_{\mathrm{T}}^{\mathrm{miss}}$ resolution is defined as the RMS of the $E_{x,y}^{\mathrm{miss}}$ in Zll events.
- $E_{\mathrm{T}}^{\mathrm{miss}}$ response R is defined in Zll events as

  $$R = \left\langle \vec{E}_{\mathrm{T}}^{\mathrm{miss}} \cdot \vec{A}_Z \right\rangle = \left\langle \vec{E}_{\mathrm{T}}^{\mathrm{miss}} \cdot \frac{\vec{p}_{\mathrm{T}}^{Z}}{|p_{\mathrm{T}}^{Z}|} \right\rangle, \tag{3.11}$$

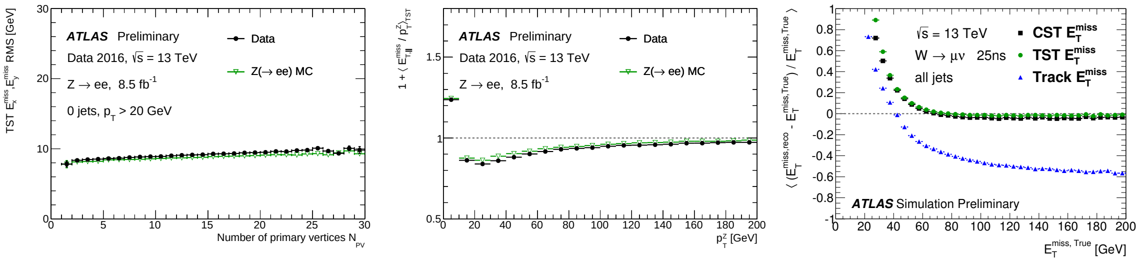
  where $\vec{p}_{\mathrm{T}}^{Z}$ is a transverse momentum of the Z-boson calculated from the vector sum of the transverse momenta of the two leptons, and $\vec{A}_Z$ is the unit vector of it. The bracket represents the mean of the inside variable. The R is sensitive to the balance between the Z-boson and the soft hadronic recoil. If they are balanced exactly, this would be zero.

- $E_{\mathrm{T}}^{\mathrm{miss}}$ linearity L is defined in the $Wl\nu$ simulation as

$$L = \left\langle \frac{E_{\mathrm{T}}^{\mathrm{miss}} - E_{\mathrm{T}}^{\mathrm{miss,True}}}{E_{\mathrm{T}}^{\mathrm{miss,True}}} \right\rangle, \tag{3.12}$$

where $E_{\mathrm{T}}^{\mathrm{miss,True}}$ is the truth-level $E_{\mathrm{T}}^{\mathrm{miss}}$ in the simulation (and the $E_{\mathrm{T}}^{\mathrm{miss}}$ in the equation is the reconstructed one). The L measures the consistency between the truth-level and reconstructed $E_{\mathrm{T}}^{\mathrm{miss}}$. If $E_{\mathrm{T}}^{\mathrm{miss}}$ were reconstructed correctly, the L would be zero.

These variables are shown in Figures 3.8. The resolution is less than 10 GeV in the range of the number of primary vertices N_{PV} between 0 and 30. The response is varied from −15 % to almost zero of the p_{T}^{Z} in the range of $20 < p_{\mathrm{T}}^{Z} < 200$ GeV. This negative value indicates an underestimation of the soft recoil. The linearity has a good performance and is zero for $E_{\mathrm{T}}^{\mathrm{miss,True}} > 70$ GeV.



(a) Resolution (b) Response (R) (c) Linearity (L)

Figure 3.8: (a) $E_{\mathrm{T}}^{\mathrm{miss}}$ resolution (RMS) of $E_{x}^{\mathrm{miss,soft}}$ and $E_{\mathrm{y}}^{\mathrm{miss,soft}}$ as a function of number of primary vertices N_{PV} and (b) $E_{\mathrm{T}}^{\mathrm{miss}}$ response divided by Z p_{T} (p_{T}^{Z}) as a function of p_{T}^{Z} in 8.5fb^{-1}data and MC simulation of $Z \longrightarrow ee$ events [40]. (c) $E_{\mathrm{T}}^{\mathrm{miss}}$ linearity as a function of $E_{x}^{\mathrm{miss,True}}$ in the $W \longrightarrow \mu\nu$ MC simulation [41]. In the linearity, green points show the TST $E_{\mathrm{T}}^{\mathrm{miss}}$, which is used in this analysis, and the other color points show $E_{\mathrm{T}}^{\mathrm{miss}}$ calculated by other algorithms.

3.4.3 Uncertainties on $E_{\mathrm{T}}^{\mathrm{miss}}$

In Eq. 3.8, the uncertainties on the hard objects terms are provided in each object calibrations and their uncertainties are propagated into the each $E_{\mathrm{T}}^{\mathrm{miss,obj}}$ term. Thus, the uncertainties on the soft term are focused on here.

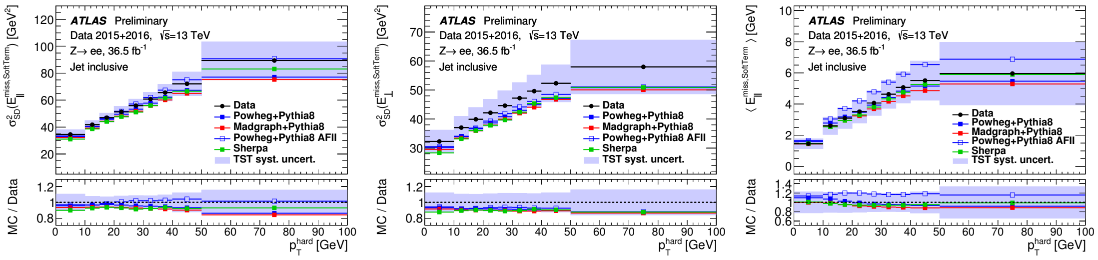
The main uncertainty is obtained from the differences between several MC generators because the difference between the data and MC simulation was smaller than the MC differences in Run1. In addition to this modeling uncertainty, the difference between the MC simulations in the different detector and running conditions provides small but non-negligible uncertainty.

The systematic uncertainties on the soft term are taken into account as three uncertainties defined based on the axis of $\vec{p}_{\mathrm{T}}^{\mathrm{hard}}$, which is the vector sum of transverse momenta of the hard objects in an event. The three uncertainties are:

- uncertainty on the resolution longitudinal to $\vec{p}_{\mathrm{T}}^{\mathrm{hard}}$,

- uncertainty on the resolution transverse to $\vec{p}_{\mathrm{T}}^{\,\mathrm{hard}}$,
- uncertainty on the soft term scale longitudinal to $\vec{p}_{\mathrm{T}}^{\,\mathrm{hard}}$.

These uncertainties are quantified by using the balance between the hard term and soft term in $Z \to ll$ events. The obtained uncertainties are shown in Figures 3.9. The uncertainties are ~ 10 % on the resolutions and ~ 20 % on the scale.



(a) Longitudinal resolution

(b) Transverse resolution

(c) Longitudinal scale

Figure 3.9: The uncertainties for the $E_{\mathrm{T}}^{\mathrm{miss}}$ soft term as a function of $\vec{p}_{\mathrm{T}}^{\,\mathrm{hard}}$ on (a) the longitudinal resolution, (b) the transverse resolution, and (c) the longitudinal scale [42]. They are obtained by using 36.1fb^{-1}data in 2015+2016 with inclusive-jet selection (no selection on the jet multiplicity).

4 Monte Carlo simulation

4.1 Monte Carlo process

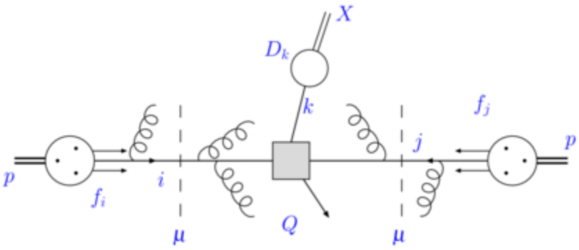


Figure 4.1: The schematic view of each Monte Carlo generator step in a proton-proton collision evolving to hadronic final state X [43]. Here, $\mu = \mu_F$ (factorization scale) $= \mu_R$ (renormalization scale). The middle region sandwiched by two break lines above μ is a process at a momentum transfer $Q^2 > \mu^2$. The shaded box shows a hard process which is calculated by matrix-element method. Winding lines show gluons coming from initial state radiations. D_k represents hadronization process from parton k to X.

Following four steps are simulated in the Monte Carlo ("MC") simulation:

1. Parton distribution function ("PDF")
2. Hard process
3. Parton shower (ISR and FSR)
4. Hadronization

The PDF (an example is shown in Figure 4.2) is a distribution of each parton in a proton as a function of momentum fraction x. It is obtained from various measurements. The hard process is a target process in each MC simulation at large momentum transfer Q^2. At such a large transfer momentum, strong interaction coupling α_S becomes enough small for perturbation theory to be valid. Thus, the hard process is predicted by perturbative calculations of a matrix-element method based on Feynman diagrams at a specific order of α_S. Parton shower simulates the process of branching external partons from two incoming partons or outgoing partons (gluon emission or quark pair-production), in which Q^2 is evolved. This process is separately calculated from the hard process because it is difficult to calculate perturbatively due to a large α_S. The hadronization is a process in that a parton evolves to numerous color-neutral hadrons. A formalization for a cross-section $\sigma_{pp\to X}$ of the process from a

p-p collision to a final state X reflects this separation of the steps as

$$\begin{aligned}\sigma_{pp\to X} \quad &= \sum_{i,j,k} \int \mathrm{d}x_1 \mathrm{d}x_2 \mathrm{d}z f_i(x_1, \mu_F^2) f_j(x_2, \mu_F^2) \times \\ &\quad \hat{\sigma}_{ij\to k}\left(x_1, x_2, Q^2, \alpha_s(\mu_R^2), \mu_F^2\right) D_{k\to X}(z, \mu_F^2), \end{aligned} \tag{4.1}$$

where i, j and k stand for the parton species, $x_{1,2}$ are momentum fractions of two partons in the protons, $f_{i,j}$ are the PDFs, $\hat{\sigma}_{ij\to k}$ is a cross-section of the process of $ij \longrightarrow k$, and $D_{k\to X}$ is a hadronization function. The μ_F is an arbitrary energy scale which is a boundary between the PDF and the hard process or between the hard process and hadronization process. The μ_R is also an arbitrary scale of the renormalization scale in the hard process. Usually, μ_F and μ_R is set to the same in the MC simulation.

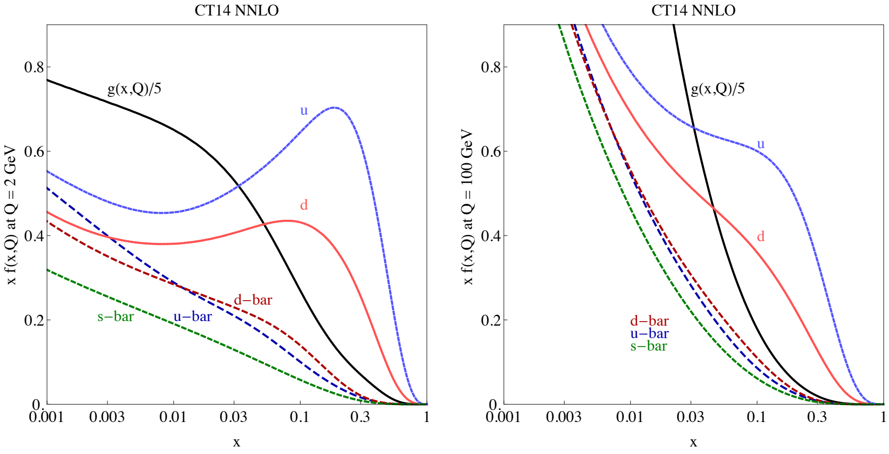


Figure 4.2: Parton Distribution Function (PDF) $f(x, Q)$ at energy scale (left) $Q = 2$ GeV and (right) 100 GeV as a function of momentum fraction x at next-to-next-to-leading order (NNLO). The PDF is obtained from the CTEQ-TEA global analysis (CT14, [44]). The various lines show different flavor partons.

4.2 Monte Carlo samples

The MC event generators used for the SUSY signals and SM background in this analysis are summarized in Table 4.1. There are two kinds of software for MC event generators. One is general-purpose Monte Carlo generators such as Pythia, Sherpa, and Herwig++, which can simulate all of the four steps of the PDF, hard process, parton shower, and hadronization. In contrast, the other is simulator only for the matrix-element calculation of the hard process, which is interfaced to parton shower and hadronization part of a general-purpose MC generator. In the main background processes, boson+jets process is simulated by Sherpa, and $t\bar{t}$ and single top processes are simulated by Powheg-box+Pythia6. The details of the SM background process will be described in Section 8.1.

Table 4.1: The MC simulation used for the SUSY signals and SM background in this analysis. The PDF sets, generators for a hard process, simulator of parton showers, and the order of α_s in cross-section calculations to obtain yield normalization are shown.

Physics process	PDF set	Generator (Hard process)	Parton shower & Hadronization	Cross-section normalization
SUSY signals	NNPDF2.3LO	MG5_aMC@NLO	Pythia8	NLO+NLL
$Z/\gamma^{*}(\rightarrow \ell\bar{\ell}, \nu\nu)$ + jets	NNPDF3.0NNLO	Sherpa2.2.1	Sherpa2.2.1	NNLO
$W(\rightarrow \ell\nu)$ + jets	NNPDF3.0NNLO	Sherpa2.2.1	Sherpa2.2.1	NNLO
γ + jets	CT10	Sherpa2.1.1	Sherpa2.1.1	LO
$t\bar{t}$	CT10	Powheg-box	Pythia6	NNLO+NNLL
Single top (Wt-channel)	CT10	Powheg-box	Pythia6	NNLO+NNLL
Single top (s-channel)	CT10	Powheg-box	Pythia6	NLO
Single top (t-channel)	CT10	Powheg-box	Pythia6	NLO
Diboson(WW, WZ, ZZ)	CT10	Sherpa2.2.1	Sherpa2.2.1	NLO
Multi-jet	NNPDF2.3LO	Pythia8	Pythia8	LO

4.2.1 Signal Monte Carlo

In a signal MC event, pair-gluons are generated by using MADGRAPH 5 and they are replaced by gluinos, which are interfaced to Pythia8, reproducing their decay process. Only light-flavor quarks are considered in the decay of the simplified models as described in Section 1.6. For the direct decay (one-step decay), a simplified model is defined such that all other superpartners except the lightest neutralino (and the lightest chargino) are decoupled, and the free parameters are $m_{\tilde{g}}$ and $m_{\tilde{\chi}_1^0}$ (and $m_{\tilde{\chi}_1^\pm}$). The MC sample is produced at each grid point in the two-dimensional mass plane of $m_{\tilde{g}}$ and $m_{\tilde{\chi}_1^0}$ with an interval of $O(100)$ GeV. The $m_{\tilde{\chi}_1^\pm}$ in the one-step decay is fixed to $m_{\tilde{\chi}_1^\pm} = (m_{\tilde{g}} + m_{\tilde{\chi}_1^0})/2$. The PDF in the signal MC is NNPDF2.3LO.

5 Analysis strategy

In order to search for the SUSY particles, event selection that maximizes the expected signal yield relative to the SM background yield is prepared. The selected kinematic phase space is referred to as "signal region (SR)" hereafter.

For an improvement in the signal sensitivity of the SR, a new information that can separate the signal and background more effectively is necessary for the event selection. In the signal decay, there are at least four quarks. However, the quark-jets are treated as just jets and no information related to the quark/gluon identification is used in the previous study [45]. In fact, the quark-jet fraction in the leading four jets in p_{T} ordering is larger than 80% in large $\Delta M(\tilde{g}, \tilde{\chi}_1^0)$ signals at the preselection, which is a loose SR-like selection described in Section 7.3, as shown in Figure 5.1.

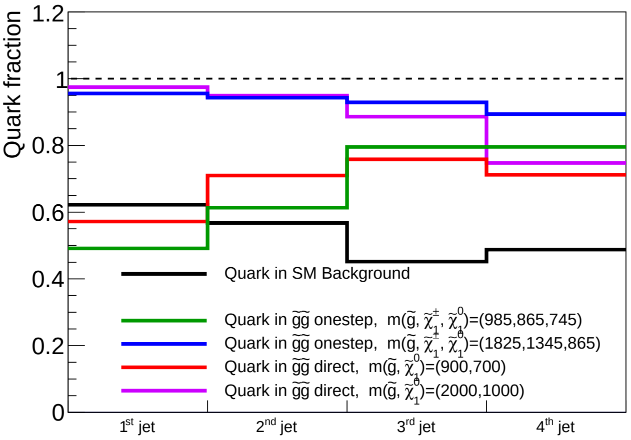
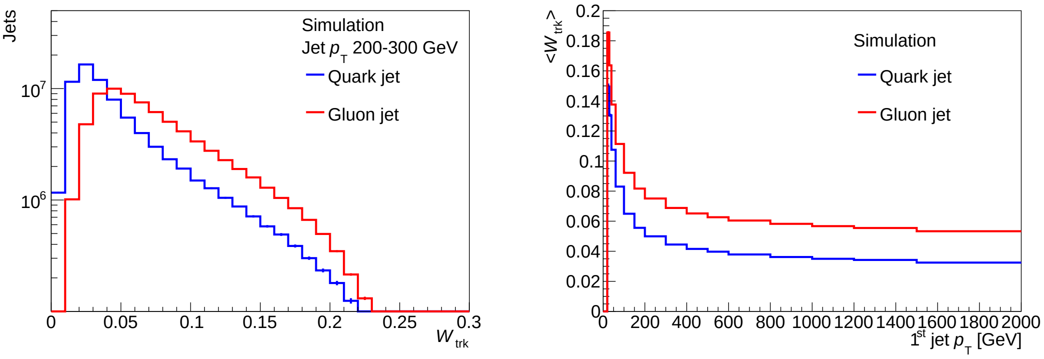


Figure 5.1: Quark-jet fractions in each of the leading four jets at the **PreDHigh** preselection defined in Table 7.2. Each bin presents a fraction that a jet with a specific number of the p_{T}-order originates from a quark. A black line shows the fraction of the SM total background. Other colored lines show the quark-jet fractions in different signals. In both of the gluino direct decay and one-step decay signals, the quark fractions are high for the large $\Delta M(\tilde{g}, \tilde{\chi}_1^0)$ signals (purple and blue lines). However, it is not the case for the first and second p_{T} jets in the low ΔM signals (red and green lines) because the quark-jet originating from the gluino decay has low p_{T} in such a low ΔM signal and it is often lower than ISR and FSR gluons.

In this analysis, the quark/gluon separation is utilized for the improvement. For quark/gluon separation, track information inside a jet is useful and several variables, e.g. number of tracks and jet width calculated from the associated tracks W_{trk}, are known as discriminating variables [46–49]. However, there are two difficulties to use quark/gluon separation variable:

- **Correlation between the quark/gluon separation variable and jet p_T**
  The variable is usually strongly correlated to the jet p_{T}. Figures 5.2 show the distribution of the W_{trk}, which is employed in this analysis, in quark- and gluon-jets and its mean value in each jet p_{T} range. The definition of the W_{trk} is described later in Section 6.1. The quark-jet has a lower value than the gluon-jet in W_{trk}, and the mean value will be decreased according to the increasing jet p_{T}. Hence, it is necessary to optimize the W_{trk} selection according to the jet p_{T} in order to obtain the maximum sensitivity.

- **Quark/gluon separation for the leading four jets**
  The quark/gluon separation needs to be considered for all of the four jets to obtain the improvement because the discriminating power of the quark/gluon separation variable is not much large. For instance, the gluon rejection is 60% at the quark acceptance of 80% for the $p_{\mathrm{T}} \sim 200$ GeV jet in Run1 [46]. The selection on the four jets makes the selection more complicated.



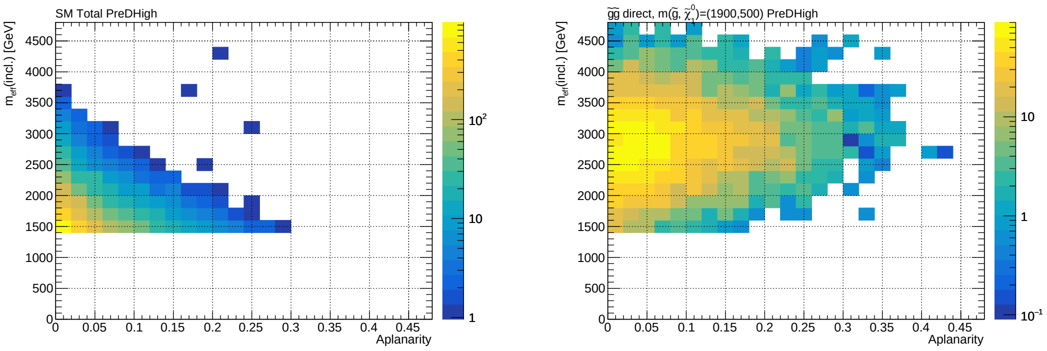
(a) W_{trk} distributions of quark- and gluon-jets

(b) Mean of W_{trk} in each jet p_{T} range

Figure 5.2: (a) Distributions of W_{trk} in the jet p_{T} range between 200 GeV and 300 GeV, and (b) mean of W_{trk} in each jet p_{T} range for quark-(blue line) and gluon-jets(red line) in the multi-jet MC simulation. The selection applied here is the multi-jet sample selection for two-process extraction defined in Table 6.2 of Section 6.

To accomplish the complicated selection taking into account the correlation between the p_{T}s and W_{trk}s of the four jets, a multivariate analysis technique, Boosted Decision Tree ("BDT",[50]), is employed in this analysis. Among the multivariate analysis techniques, the BDT is generally used in the ATLAS experiment, for example, b-tagging, $H \to bb$ search [51], etc. The BDT provides one score (BDT score) indicating if an event is signal-like or background-like according to measurement variables ("input variables") of the event and correlations between them.

In apart from the quark/gluon separation, the BDT itself is expected to provide an improvement if the conventional variables used in the previous study [45] are also given as input variables of the BDT analysis because the BDT can take into account the correlation between the input variables. The correlation is not considered in the previous analysis, in which only the cut on each variable is used. For example, in the previous analysis, aplanarity, which is a variable related to the multi-jet topology, and m_{eff}, which is a variable indicating the hardness of the event, are used. The details of the two variables will be described in Section 7.2. In the SR of the previous analysis, the fixed cut on each of the aplanarity and m_{eff} is required. However, the two-dimensional distribution of them is very different between the signal and background as shown in Figures 5.3. Obviously, the varying lower cut on the m_{eff} according to the aplanarity, in which the cut value is decreased with the increasing aplanarity, is better than the fixed-cuts, i.e. rectangular cut, on them.



(a) SM background

(b) Gluino direct decay signal

Figure 5.3: Two-dimensional distributions between m_{eff} and aplanarity in (a) the SM background and (b) the gluino direct decay signal with $(\tilde{g}, \tilde{\chi}_1^0)$ = (1900, 500) GeV at the preselection **PreDHigh** defined in Table 7.2. The total number of the signal is normalized to that of the SM background.

Therefore, in order to obtain the best improvement, this analysis uses a *BDT analysis* in which the *quark/gluon separation variables (*W_{trk}*)* of the leading four jets and the conventional variables used in the previous analysis are taken into account as input variables. To accomplish this analysis, calibration and uncertainties of the W_{trk} are necessary because the W_{trk} is sensitive to the hadronization, which is difficult to be predicted precisely in the simulation. The calibration of the W_{trk} will be explained in Section 6, which will provide a "scale factor" ("SF") as a jet-by-jet correction factor on the simulation and its up and down variations as systematic uncertainties.

6 Calibration of quark/gluon separation variable

6.1 Quark/gluon separation variable

In order to distinguish quark jets from gluon jets, the information of the track activity inside jets is important because the color factor of gluons is larger than that of quarks by factor 9/4 ("Casimir ratio"), which makes gluons emit more particles in the hadronization than quarks. Thus, a gluon jet has more charged tracks in it and the jet width is larger than that of a quark-jet.

Here, jet width computed from the associated tracks W_{trk} is used as a quark/gluon separation ("q/g separation") variable, which is a track-p_{T}-weighted width of the jet divided by the scalar sum of the track transverse momenta. It is defined as

$$W_{\mathrm{trk}} = \frac{\sum_{\mathrm{trk}\in\mathrm{jet}} p_{\mathrm{T,trk}} \Delta R_{\mathrm{trk,jet}}}{\sum_{\mathrm{trk}\in\mathrm{jet}} p_{\mathrm{T,trk}}}, \tag{6.1}$$

where $p_{\mathrm{T,trk}}$ is a p_{T} of a charged track reconstructed by the inner detector (ID) and $\Delta R_{\mathrm{trk,jet}}$ is a distance in the η-ϕ plane between the track and the jet axis. This variable is insensitive to the track inefficiency because it is defined as a ratio. The charged tracks used here are required to have $p_{\mathrm{T,trk}} > 1$ GeV and identified by the "TightPrimary" selection described in Section 3.1 in order to remove fake tracks which make η-dependence in W_{trk}.

The calibration of this variable and estimation of its uncertainties are necessary since such jet substructure information is not used in the conventional SUSY searches and mis-modeling of the simulation, especially in gluon jets, is known in the previous study for q/g separation in Run1 [47]. The calibration of the q/g separation variable is performed by applying binned jet-by-jet scale factor in the simulation for quark- and gluon-jets, respectively. The scale factor is obtained from the W_{trk} distributions in quark- and gluon-jets from data in order to match the shape of the simulation to that of the data, in which the shape is obtained in each jet p_{T} range because it depends on jet p_{T} strongly. The jet used in this calibration is restricted to jets with $p_{\mathrm{T}} > 40$ GeV and $|\eta| < 2.1$. The slightly tighter $|\eta|$ requirement than a usual jet identification ($|\eta| < 2.8$) is in order to avoid $|\eta|$ dependence of W_{trk} caused by the ID coverage ($|\eta| < 2.5$).

6.2 Method to extract quark/gluon from data

To extract the shape of W_{trk} distributions for quark- and gluon-jets from data separately, a "matrix method" of two samples with different quark/gluon fractions is used. The matrix method can extract pure quark or gluon jets from quark-enriched and gluon-enriched samples under the assumption that the two samples have the same shape of W_{trk} distributions in each of quark and gluon jets. The matrix method is performed for each p_{T} bin defined in Table 6.1 and the quark-enriched and gluon-enriched samples are different between lower p_{T} ranges and higher p_{T} ranges.

Table 6.1: The p_{T} range division for the calibration of W_{trk} and samples used in extraction of pure quark- and gluon-jets.

Lower p_{T} bin boundary [GeV]													
40	60	100	150	200	300	400	500	600	800	1000	1200	1500	2000
Z+jets & multi-jet samples (2-process extraction)							Higher & lower $\lvert\eta\rvert$ jet samples in multi-jet						

For the lower p_{T} ranges (40–500 GeV), the quark-enriched and gluon-enriched samples are obtained from the leading p_{T} jet in Z+jets events and multi-jet events, respectively. The quark in Z+jets is an associated quark in the Feynman diagram given in Figures 8.1 of Section 8.1. The gluon in multi-jet comes from a gluon emission in QCD process. This extraction method is referred to as "2-process extraction" hereafter.

For the higher p_{T} ranges (500–2000 GeV), the quark-enriched and gluon-enriched samples are obtained from higher and lower $\lvert\eta\rvert$ jets of the leading two jets in multi-jet events, respectively. The higher $\lvert\eta\rvert$ jet comes from high x (momentum fraction) of a PDF. Since the PDF in the high-x range has a high probability of valence-quarks as shown in Figure 4.2 of Section 4.1, the higher $\lvert\eta\rvert$ jet sample has more quark jets. In contrast, the lower $\lvert\eta\rvert$ jet has lower x and more gluon jets.

The matrix in 2-process extraction is given as

$$\begin{pmatrix} p_{\mathrm{Z+jets}}(W_{\mathrm{trk}}) \\ p_{\mathrm{Multi-jet}}(W_{\mathrm{trk}}) \end{pmatrix} = \underbrace{\begin{pmatrix} f_{\mathrm{Z+jets,Q}} & f_{\mathrm{Z+jets,G}} \\ f_{\mathrm{Multi-jet,Q}} & f_{\mathrm{Multi-jet,G}} \end{pmatrix}}_{\equiv F} \begin{pmatrix} p_{\mathrm{Q}}(W_{\mathrm{trk}}) \\ p_{\mathrm{G}}(W_{\mathrm{trk}}) \end{pmatrix} \tag{6.2}$$

$$\Leftrightarrow \begin{pmatrix} p_{\mathrm{Q}}(W_{\mathrm{trk}}) \\ p_{\mathrm{G}}(W_{\mathrm{trk}}) \end{pmatrix} = F^{-1} \begin{pmatrix} p_{\mathrm{Z+jets}}(W_{\mathrm{trk}}) \\ p_{\mathrm{Multi-jet}}(W_{\mathrm{trk}}) \end{pmatrix}, \tag{6.3}$$

and the matrix using the higher and lower $\lvert\eta\rvert$ jets is

$$\begin{pmatrix} p_{\mathrm{H}}(W_{\mathrm{trk}}) \\ p_{\mathrm{L}}(W_{\mathrm{trk}}) \end{pmatrix} = \underbrace{\begin{pmatrix} f_{\mathrm{H,Q}} & f_{\mathrm{H,G}} \\ f_{\mathrm{L,Q}} & f_{\mathrm{L,G}} \end{pmatrix}}_{\equiv F'} \begin{pmatrix} p_{\mathrm{Q}}(W_{\mathrm{trk}}) \\ p_{\mathrm{G}}(W_{\mathrm{trk}}) \end{pmatrix} \tag{6.4}$$

$$\Leftrightarrow \begin{pmatrix} p_{\mathrm{Q}}(W_{\mathrm{trk}}) \\ p_{\mathrm{G}}(W_{\mathrm{trk}}) \end{pmatrix} = F'^{\,-1} \begin{pmatrix} p_{\mathrm{H}}(W_{\mathrm{trk}}) \\ p_{\mathrm{L}}(W_{\mathrm{trk}}) \end{pmatrix}, \tag{6.5}$$

where $p_{\mathrm{Q,G}}(W_{\mathrm{trk}})$ show W_{trk} distributions in pure quark- and gluon-jet samples, $p_{\mathrm{Z+jets/Multi-jet/H/L}}(W_{\mathrm{trk}})$ show W_{trk} distributions in Z+jets, multi-jet, higher $\lvert\eta\rvert$ jet, and lower $\lvert\eta\rvert$ jet samples, respectively, and $f_{\mathrm{X,Q/G}}$ are fractions of quark and gluon jets in sample X. Between Eqs. 6.2 and 6.3 or Eqs. 6.4 and 6.5, the inverse matrix of F or F' is calculated and used to extract pure quark/gluon ($p_{\mathrm{Q,G}}$). The distribution of quark/gluon-enriched sample is obtained from data and the fraction of quarks and gluons in them are obtained from the MC simulations. This matrix is computed in each W_{trk} bins and each jet p_{T} ranges.

To avoid a systematic uncertainty coming from the parton shower modeling, this calibration is performed for Pythia8 and Sherpa separately. For Sherpa, the Z+jets MC is Sherpa2.2.1 and the multi-jet MC is Sherpa2.1.1. This difference will be taken into account in the systematic uncertainties. In Sections 6.1–6.4, the figures using Pythia8 (Sherpa) show results with NNPDF3.0NNLO (NNPDF3.0NNLO) and NNPDF2.3LO (CT10) PDF sets for 2-process extraction and for higher/lower $|\eta|$ jet extraction, respectively.

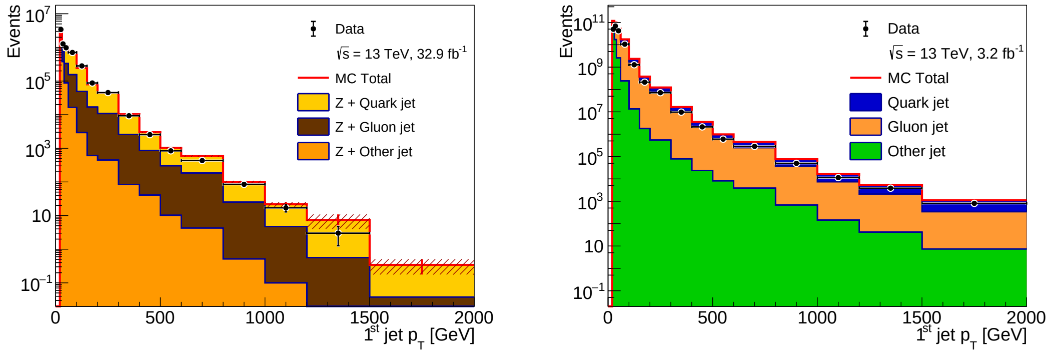
The selections for all the samples used here are summarized in Table 6.2. The pure Z+jets events are obtained by requirements of two opposite charge leptons and Z mass cut between 75 GeV and 105 GeV in the invariant mass of the two leptons. The multi-jet sample consists of no-lepton events. There is $|\eta| < 2.1$ requirement because the ID covers $|\eta|$ up to 2.5.

Table 6.2: The selections to retrieve quark/gluon-enriched samples. The requirements in the upper half are in order to obtain Z+jets/multi-jet events. The ones in the lower half are to enrich quark jets or gluon jets. "j_i" represents the i-th jet in p_{T}-ordering.

	Selection	Z+jets sample	Multi-jet sample		
			For 2-process extraction	Higher $\lvert\eta\rvert$ jet sample	Lower $\lvert\eta\rvert$ jet sample
Preselection	Trigger	Single lepton trigger	Or of single jet triggers		
	Object	Two opposite charge leptons	No lepton		
	$m_{\ell\ell}$	$75 < m_{\ell\ell} < 105$ GeV	-		
	Number of jets	≥ 1	≥ 2		
	b-veto	$j_1 \neq b$-jet	$j_1 \neq b$-jet	$j_1 \neq b$-jet and $j_2 \neq b$-jet	
	$\lvert\eta(j_1)\rvert$	< 2.1	< 2.1	< 2.1	
	$\lvert\eta(j_2)\rvert$	-	-	< 2.1	
Quark/gluon enhance	Target parton	Quark	Gluon	Quark	Gluon
	$p_{\mathrm{T}}(Z)/p_{\mathrm{T}}(j_2)$	<1.5	-	-	
	$p_{\mathrm{T}}(j_1)/p_{\mathrm{T}}(j_2)$	-	-	< 1.5	
	$p_{\mathrm{T}}(j_2)$	$<\max(30\ \mathrm{GeV}, 0.5p_{\mathrm{T}}(Z))$	> 20 GeV	> 20 GeV	
	$\Delta\phi(Z, j_1)$	> 2.5	-	-	
	$\Delta\phi(j_1, j_2)$	-	> 2.5	-	
	$\lvert\eta(j_1)\rvert$	-	$< \lvert\eta(j_2)\rvert$	-	
	Used jet in j_1 or j_2	Only j_1	Only j_1	Higher $\lvert\eta\rvert$ jet	Lower $\lvert\eta\rvert$ jet

The jet p_{T} distribution for each sample with Pythia8 is shown in Figures 6.1 and 6.2. The partonic flavor label (quark[u, d, s, or c], gluon, b-quark or "other") of a jet in the simulation is defined by a flavor of the highest-energy parton in the parton shower within $\Delta R = 0.4$ with the jet. 0.4 is equal to the radius parameter of the jet algorithm. The fractions of quark and gluon jets in each sample are shown in Figures 6.3. In the lower p_{T} ranges ($<$ 500 GeV), the quark fraction of Z+jets is high ($\sim$ 75%) and the difference between the quark fractions in Z+jets and multi-jet is large (30–50%), but the quark fraction of higher $|\eta|$ jet is low ($\lesssim$ 50%). Thus, the Z+jets and multi-jet are used as quark/gluon-enriched samples in the lower p_{T} ranges. In the higher p_{T} ranges ($>$ 500GeV), higher

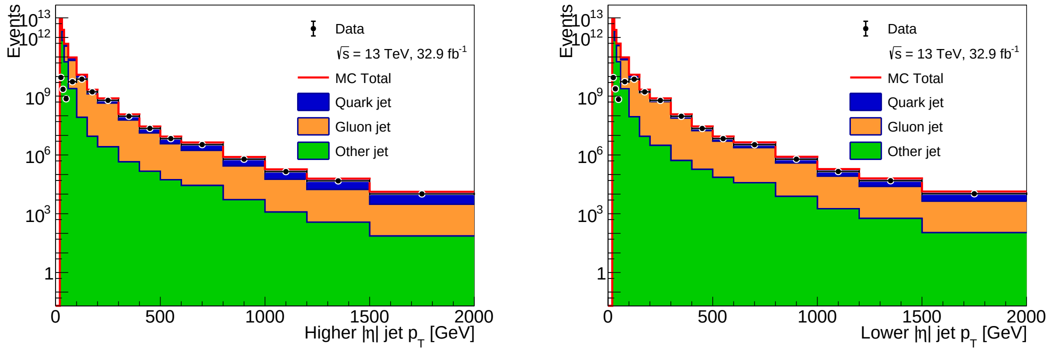
$|\eta|$ jet has a large fraction of quarks (> 60%). Thus, in the higher p_{T} ranges, the higher/lower $|\eta|$ jet samples are used[1]. The difference between the MC event generators in the fractions is known to be small in the previous study [48]. In the whole p_{T} range, b-quark jets and jets labeled "other" exist, but it is suppressed to be lower than a few %, which can be ignored. The jets labeled "other" are jets mainly originating from pileup.



(a) Z+jets (Pythia8)

(b) Multi+jet (Pythia8)

Figure 6.1: The p_{T} distributions of the leading jets with Pythia8 MC in (a) Z+jets and (b) multi-jet for 2-sample process extraction. Data for (a) is 32.9fb^{-1}in 2016 and that for (b) is 3.2fb^{-1}in 2015. The normalization of the simulation is decided by cross-section.



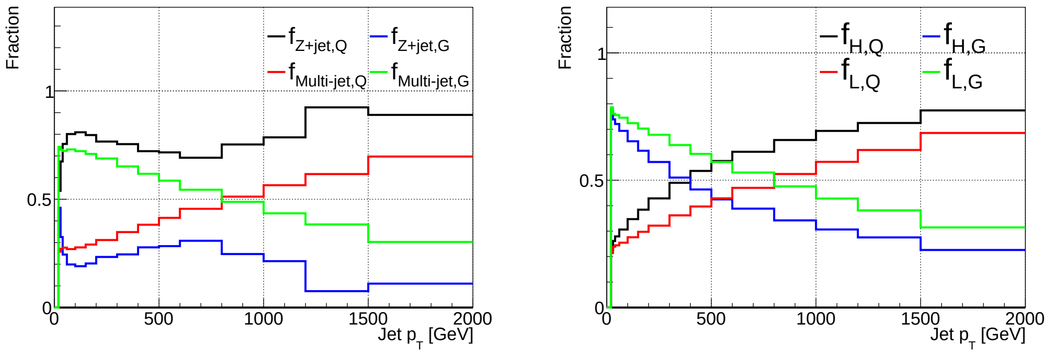
(a) Higher $|\eta|$ jet (Pythia8)

(b) Lower $|\eta|$ jet (Pythia8)

Figure 6.2: The p_{T} distributions of (a) the higher $|\eta|$ jet and (b) the lower $|\eta|$ jet of the leading two jets in multi-jet events with Pythia8 MC. Data for both figures are 32.9fb^{-1}in 2016. In the low p_{T} range (< 100 GeV), the number of data is much smaller than the simulation due to the event skimming only on data in the data processing, but the low p_{T} range is not used for the calibration. The normalization of the simulation is decided by cross-section.

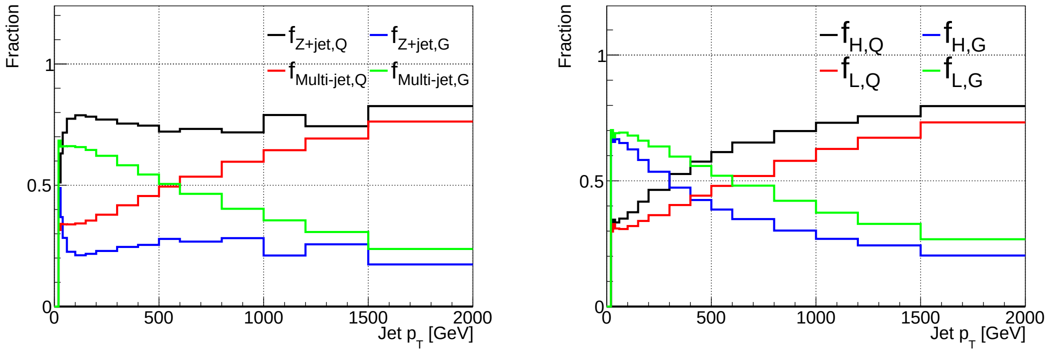
---

[1] The low statistic of Z+jets sample in the higher p_{T} ranges is another reason to use the higher/lower $|\eta|$ jet samples.



(a) For 2-process extraction (Pythia8)

(b) For higher/lower $|\eta|$ jet extraction (Pythia8)



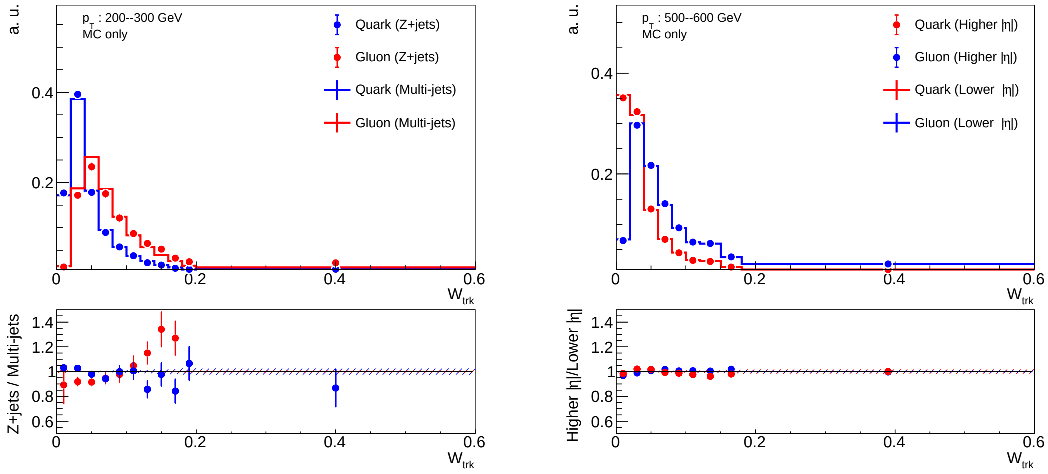
(c) For 2-process extraction (Sherpa)

(d) For higher/lower $|\eta|$ jet extraction (Sherpa)

Figure 6.3: Fractions of quark- and gluon-jets in each of (a)(c) Z+jets and multi-jet for 2-process extraction and (b)(d) higher and lower $|\eta|$ jets (H and L represent higher and lower $|\eta|$ jets.). Pythia8 and Sherpa MCs are used in the top and bottom figures, respectively. These values are used as elements in the F (F') matrix in Eq. 6.2 (6.4).

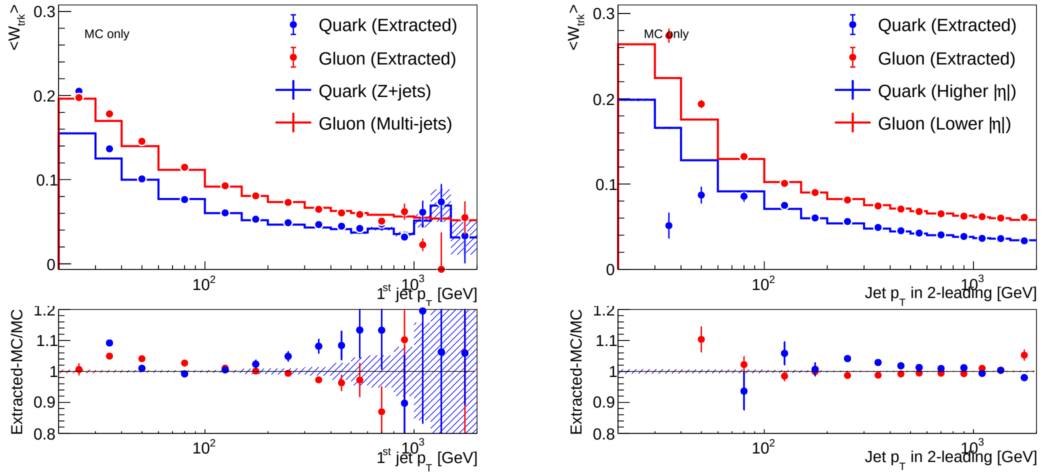
6.3 MC closure

The matrix method (Eq. 6.3) is valid only if the shape of the W_{trk} distribution is the same between the quark- and gluon-enriched samples, respectively for quark and gluon jets. The validation of this assumption is performed in the MC simulation by injecting the MC samples as $p_{\mathrm{Q/G-rich}}(W_{\mathrm{trk}})$ in Eq. 6.3. The difference between pure quark/gluon samples defined by the partonic flavor label in the MC and the extracted pure quark/gluon samples by Eq. 6.3 is defined as an MC non-closure. Since there is a bit difference between Z+jets and multi-jet or between higher $|\eta|$ and lower $|\eta|$ jets in the MC (Figures 6.4), there is 10% MC non-closure at maximum in the mean of W_{trk} in both of Pythia8 and Sherpa as shown in Figures 6.5. In Pythia8, the maximum MC non-closure exists in quark W_{trk} in the jet p_{T} range between 300 and 400 GeV in the 2-process extraction. In this range, the gluon W_{trk} in multi-jet is lower than that in Z+jets as shown in Figure 6.4(a). This difference causes the MC non-closure in the extraction. Overall, the MC non-closure in Sherpa is larger than that in Pythia8. This MC non-closure is taken into account as a systematic uncertainty in Section 6.5.

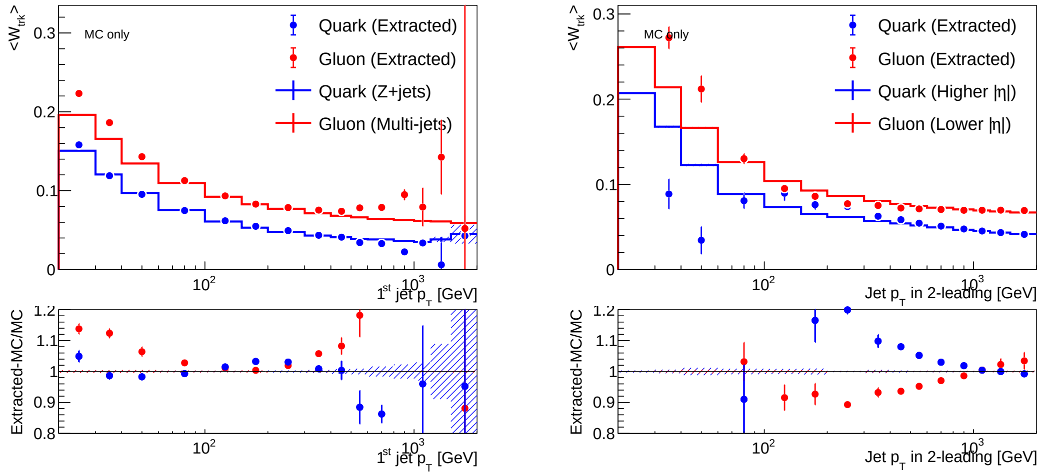


(a) Z+jets v.s. Multi-jet (Pythia8) (b) Higher $|\eta|$ v.s. Lower $|\eta|$ (Pythia8)

Figure 6.4: Comparisons in the shape of W_{trk} distributions in quark or gluon jets between (a) the leading jets of Z+jets events and multi-jet events in the jet p_{T} range of 200–300 GeV, and between (b) higher $|\eta|$ and lower $|\eta|$ jets in the jet p_{T} range of 500–600 GeV. The two samples in the comparison are equalized in the total number of events. The MC is Pythia8.



(a) For 2-process extraction (Pythia8) (b) For higher/lower $|\eta|$ jet extraction (Pythia8)



(c) For 2-process extraction (Sherpa) (d) For higher/lower $|\eta|$ jet extraction (Sherpa)

Figure 6.5: The difference in the mean of W_{trk} between the pure quark or gluon jets defined by the partonic flavor label and the extracted pure quark or gluon jets by the matrix method in each p_{T} bins. (a) and (c) are comparisons in 2-process extraction, which is used in the lower jet p_{T} ranges below 500 GeV. (b) and (d) are comparisons in the matrix method using the higher and lower $|\eta|$ jets, which is used in the higher jet p_{T} ranges above 500 GeV. Pythia8 and Sherpa MC simulations are used in the top and bottom figures, respectively.

6.4 Scale factor

The extracted pure quark and gluon distributions are shown in Figures 6.6. The "scale factor" ("SF") in each W_{trk} bin and in each p_{T} bin for quark and gluon jets is respectively calculated from these distributions in order to correct the shape of W_{trk} distributions. The SF for a quark/gluon jet with a p_{T} and a W_{trk} is given as

$$SF_{\mathrm{Q/G}}\left(W_{\mathrm{trk}}; p_{\mathrm{T,j}}\right) = \frac{p_{\mathrm{Q/G,\ Ext.Data}}\left(W_{\mathrm{trk}}; p_{\mathrm{T,j}}\right)}{p_{\mathrm{Q/G,\ Ext.MC}}\left(W_{\mathrm{trk}}; p_{\mathrm{T,j}}\right)}, \tag{6.6}$$

where Q/G indicates quark/gluon, $p_{\mathrm{T,j}}$ represents the j-th jet p_{T} bin including p_{T}, and $p_{\mathrm{Q/G,\ Ext.Data/MC}}$ is a distribution of the extracted pure quark/gluon jets from data/MC shown as in Figures 6.6. The denominator on the right-hand side is the distribution of the extracted quark/gluon jets from the MC by the same matrix method in order to suppress the influence of the MC non-closure. This SF is used as a jet-by-jet weight in the analysis when the W_{trk} of the jet is used in the selection. Systematic uncertainties are considered as "SF up/down" with an up/down variation from the nominal SF.

6.5 Systematic uncertainties

The systematic uncertainties are listed below:

- (Parton shower modeling)
- The MC non-closure
- PDF uncertainties
- Tracking uncertainties
- Data statistical uncertainty

To avoid an additional systematic uncertainty coming from the parton shower modeling, this calibration is performed for Pythia8 and Sherpa separately as described before. The MC non-closure is the difference between the pure quark/gluon distribution in the MC defined by the jet parton label and the extracted pure quark/gluon distribution in the MC described in Section 6.3. The half of the non-closure is added as a systematic uncertainty to the SF up/down symmetrically. PDF uncertainties are obtained by using LHAPDF-6.1.5 package [52], which provides other PDF sets and the PDF internal variations for each PDF set as weight variations depending on the momentum fraction x and the partonic flavor of reacted partons in the collided protons. By changing a nominal PDF weight to the systematic variation, the MC with the PDF variation is obtained. The variation on the SF from the PDF uncertainties is calculated from the difference between the nominal SF and the SF computed from the MC with the PDF variation weight. The MC inputs are $F^{(\prime)}$ matrix in Eqs. 6.2 and 6.4, and $p_{\mathrm{Q/G,Ext.MC}}$ in Eq. 6.6. Here, NNPDF3.0 [53], CT10 [54], and MMHT2014 [55] PDF sets and their internal systematic variations are considered. The total up and down variations coming from all the PDF uncertainties are determined from the envelope of the variations of the three PDF sets. For the tracking uncertainties, there are five sources of systematic uncertainties:

Track reconstruction efficiency

The uncertainty on the track reconstruction efficiency caused by the uncertainty of the ID material distributions, which is less than 1% in the efficiency.

Fake track rate

The uncertainty on the rate of reconstructed fake tracks passing the track ID selection.

Impact parameter resolution

The uncertainty on the transverse (d_0) and longitudinal (z_0) impact parameter resolution. This reflects the difference in the resolution between the data and MC.

Detector distortion

The uncertainty on the reconstructed Q/p (charge over momentum), d_0, and z_0 caused by the detector distortion that is not able to be considered in the alignment of the ID.

Lost track in a dense environment

The uncertainty on the probability of losing a track in a core of jets due to the track-dense environment inside the jets.

These uncertainties are considered by randomly dropping a track or changing its parameter before W_{trk} computation.

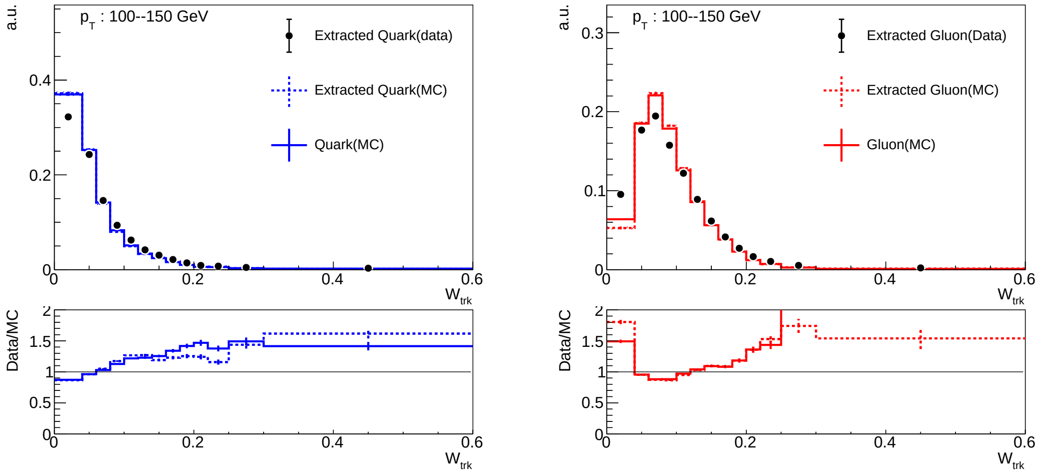
The total uncertainty of the above (quadrature sum of them) and the total systematic uncertainties on the SF using Pythia8 are shown in Figures 6.7, and also the breakdown of the systematic uncertainties for Pythia8 is shown in Figures 6.8. In Pythia8, up or down uncertainty at the peak of W_{trk} is $\sim$ 10% for both of quark and gluon jets in the p_{T} range between 100–150 GeV and for gluon jets in the p_{T} rage of 800–1000 GeV. It is mainly caused by the MC non-closure. For quark jets with p_{T} between 800 and 1000 GeV, the uncertainty at the peak is $\sim$ 5% mainly contributed by the PDF uncertainties. The uncertainties obtained by Sherpa are shown in Figures 6.9 and Figures 6.10. If Sherpa samples are used, there is also 5–10% uncertainties at the peak as in Pythia8.

The obtained calibrations are valid only for quark- and gluon-jets. Therefore, the calibrations cannot be used for b-quark jets and jets labeled "other". For Sherpa, in the lower p_{T} ranges, the calibrations are valid only for quark jets in Sherpa2.2.1 and gluon jets in Sherpa2.1.1 because the quark- and gluon-enriched samples are Sherpa2.2.1 Z+jets MC and Sherpa2.1.1 multi-jet MC, respectively. In the higher p_{T} ranges, only Sherpa2.1.1 multi-jet MC is used and both of the quarks and gluons are valid only for Sherpa2.1.1. Hence, for the other parton flavors (b-quark or label of "other") or the other parton shower modelings, only the uncertainties are defined and the nominal SF is set to 1. The uncertainty for the b-quarks is defined by the envelope of the uncertainties for the quarks and gluons. The uncertainty for a different-version shower modeling (Pythia6 for Pythia8, Sherpa2.2.1 for Sherpa2.1.1, and Sherpa2.1.1 for Sherpa2.2.1) is defined by the envelope of 1 and SF up/down of the obtained calibration for the corresponding shower modeling (Pythia8 or Sherpa). The computation algorithms for the calibrations and the uncertainties on W_{trk} SF for each parton flavor and each parton shower modelings are summarized in Table 6.3.

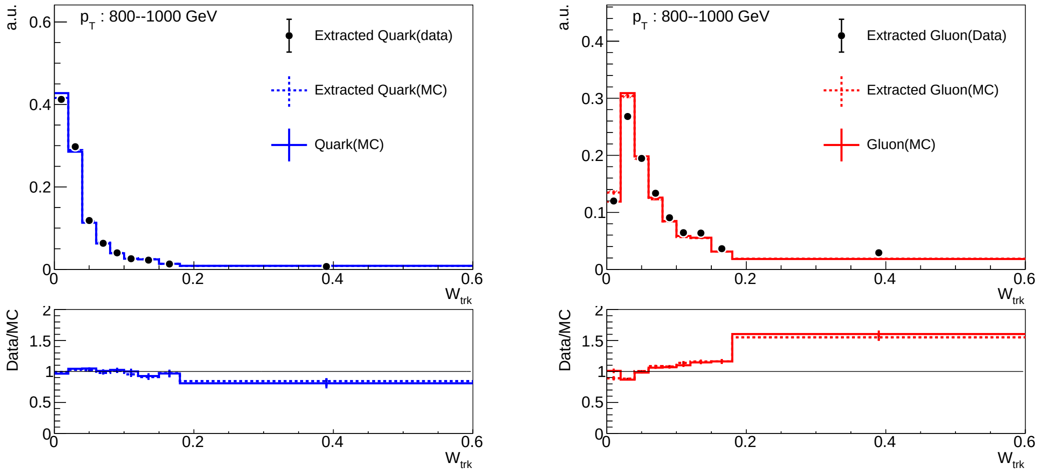
The systematic uncertainties listed here are valid for the Z+jets/multi-jet MC samples used in the calibration. Generally, the W_{trk} distribution is supposed to depend mainly on the jet p_{T} and parton label. However, in order to use this calibration for the other MC samples used in the analysis, it is necessary to check that there is no process dependence in the W_{trk} distribution (especially gluon jets for background process and quark jets for the signal process). The validation and additional systematic uncertainty associated with the validation will be described in Section 9.3.4.

Table 6.3: The summary of computation algorithms of W_{trk} scale factor (SF) and its up/down variations. P(S) means the value obtained in the calibration for Pythia8(Sherpa) is used. In case that nominal SF is 1 and SF up/down is P or S, the SF up and down are obtained from the envelope of 1 and SF up/down of the obtained calibration for Pythia8 or Sherpa. The "Q or G" means the SF up and down is defined by the envelope of the SF ups and SF downs calculated for quark jets and gluon jets. The "P or S" means the SF up and down is defined by the envelope of the SF ups and SF downs calculated with Pythia8 and Sherpa. The jets labeled "other" have 100% uncertainties on the SF.

Parton flavor		Quark	Gluon	*b*-quark	Other
Pythia8	Nominal SF	P	P	1.	1.
	SF up/down	P	P	Q or G	2/0
Pythia6	Nominal SF	1.	1.	1.	1.
	SF up/down	P	P	Q or G	2/0
Sherpa 2.2.1	Nominal SF	S	1.	1.	1.
(In the lower p_{T} ranges)	SF up/down	S	S	Q or G	2/0
Sherpa 2.1.1	Nominal SF	1.	S	1.	1.
(In the lower p_{T} ranges)	SF up/down	S	S	Q or G	2/0
Sherpa 2.2.1	Nominal SF	1.	1.	1.	1.
(In the higher p_{T} ranges)	SF up/down	S	S	Q or G	2/0
Sherpa 2.1.1	Nominal SF	S	S	1.	1.
(In the higher p_{T} ranges)	SF up/down	S	S	Q or G	2/0
Others	Nominal SF	1.	1.	1.	1.
	SF up/down	P or S	P or S	Q or G × P or S	2/0

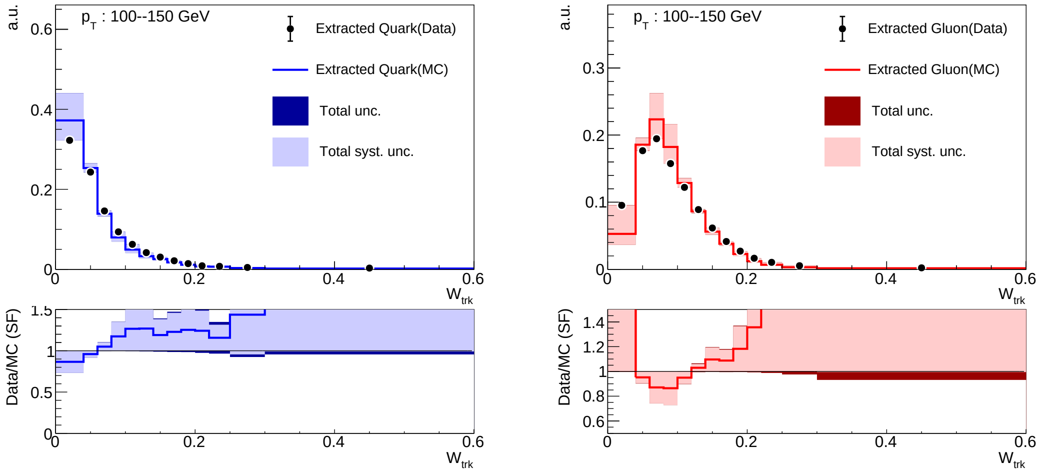


(a) Quark: $100 < p_{\mathrm{T}} < 150$ GeV (Pythia8) (b) Gluon: $100 < p_{\mathrm{T}} < 150$ GeV (Pythia8)

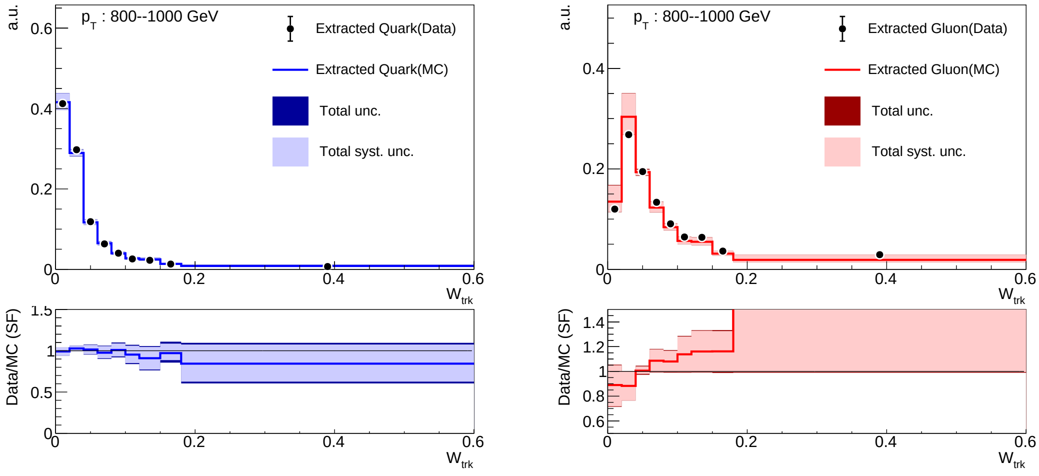


(c) Quark: $800 < p_{\mathrm{T}} < 1000$ GeV (Pythia8) (d) Gluon: $800 < p_{\mathrm{T}} < 1000$ GeV (Pythia8)

Figure 6.6: The W_{trk} distributions of pure (a)(c) quark jets and (b)(d) gluon jets extracted by the matrix method from the data and MC. The top two figures show W_{trk} in the jet p_{T} range between 100 and 150 GeV, and the bottom two figures show W_{trk} in the jet p_{T} range between 800 and 1000 GeV. Solid-line histograms show the W_{trk} distributions of quark or gluon jets defined by the jet parton flavor label in the MC. A bottom panel in each figure shows the ratio of the extracted data to the extracted MC by the matrix method (break line) or the MC defined by the parton label (solid line). The MC is Pythia8.

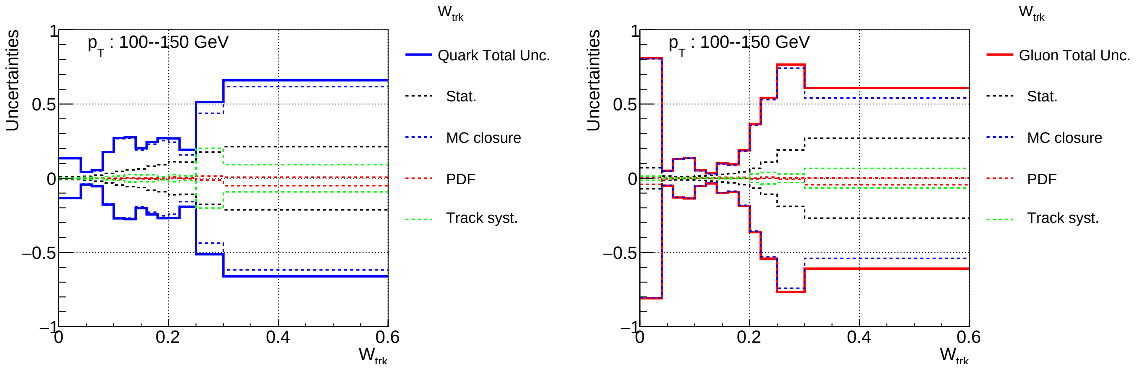


(a) Quark: $100 < p_T < 150$ GeV (Pythia8) (b) Gluon: $100 < p_T < 150$ GeV (Pythia8)



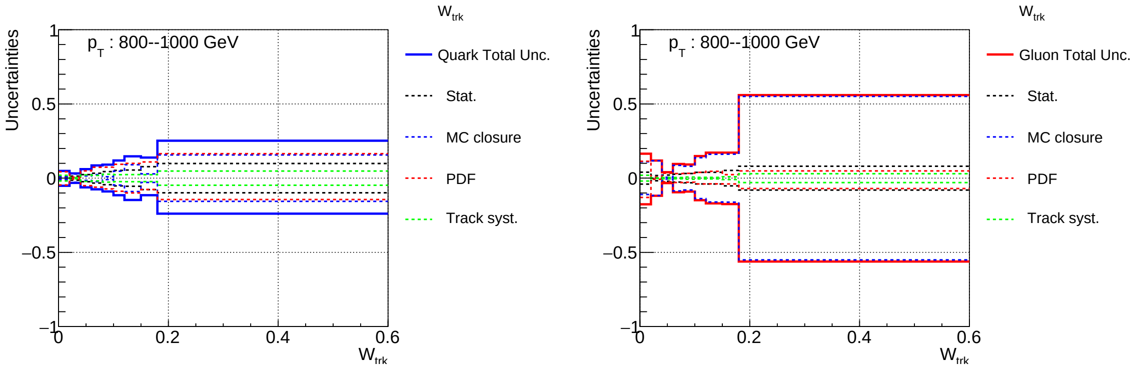
(c) Quark: $800 < p_T < 1000$ GeV (Pythia8) (d) Gluon: $800 < p_T < 1000$ GeV (Pythia8)

Figure 6.7: The W_{trk} distributions of the extracted (a)(c) quark jets and (b)(d) gluon jets from the data and MC with the total uncertainties and total systematic uncertainties obtained by Pythia8 MCs. The top two figures show the distributions in the jet p_T range between 100 and 150 GeV. The bottom two show that in the jet p_T range between 800 and 1000 GeV. The lower panel in each figure shows the extracted data divided by the extracted MC in the upper panels, which corresponds to the W_{trk} SF.



(a) Quark : $100 < p_{\mathrm{T}} < 150$ GeV (Pythia8)

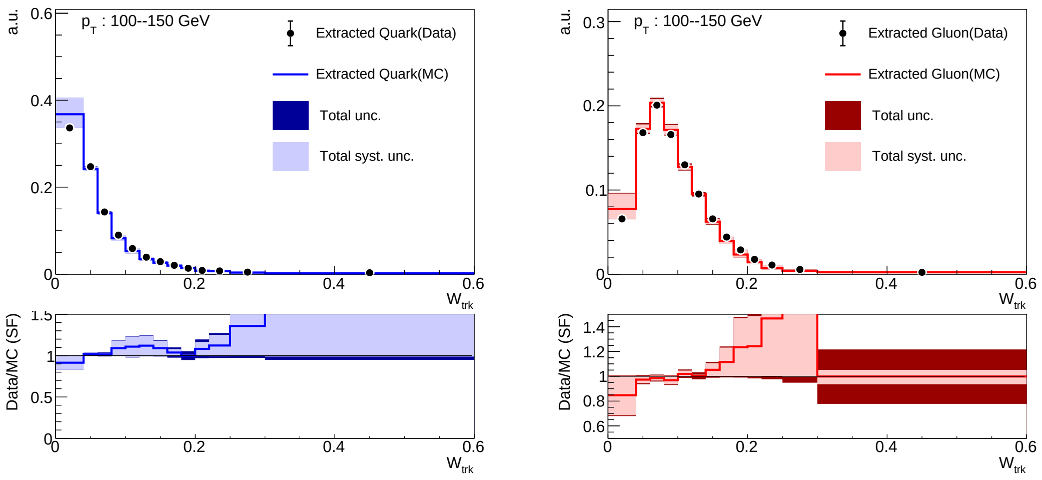
(b) Gluon : $100 < p_{\mathrm{T}} < 150$ GeV (Pythia8)



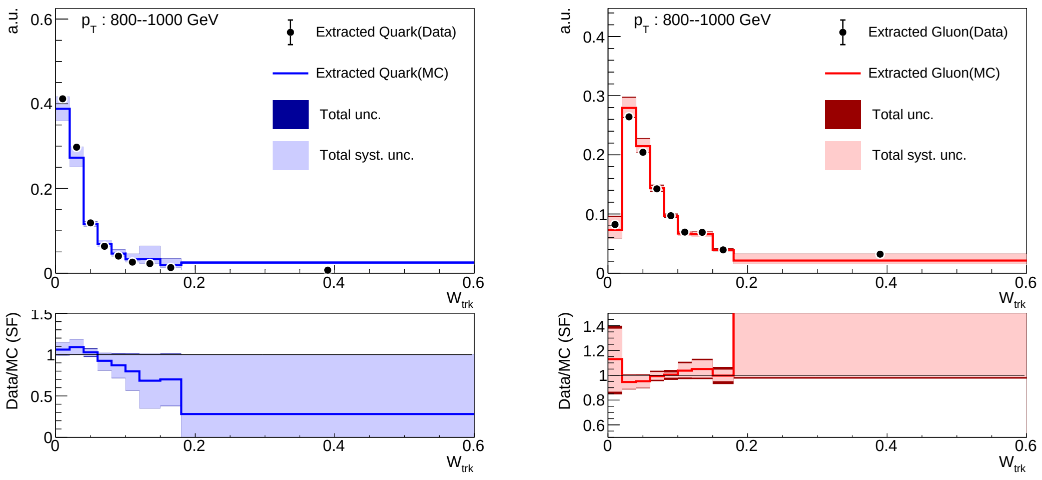
(c) Quark : $800 < p_{\mathrm{T}} < 1000$ GeV (Pythia8)

(d) Gluon : $800 < p_{\mathrm{T}} < 1000$ GeV (Pythia8)

Figure 6.8: The fraction of each uncertainty on the SF obtained by Pythia8 MCs as a function of W_{trk} for (a)(c) quark jets and (b)(d) gluon jets. The top two figures show the uncertainties in the jet p_{T} range between 100 and 150 GeV. The bottom two show that in the jet p_{T} range between 800 and 1000 GeV. The PDF up/down shows the envelope of all the PDF uncertainties. The track systematic up/down shows the quadrature sum of five sources of track systematic uncertainties.

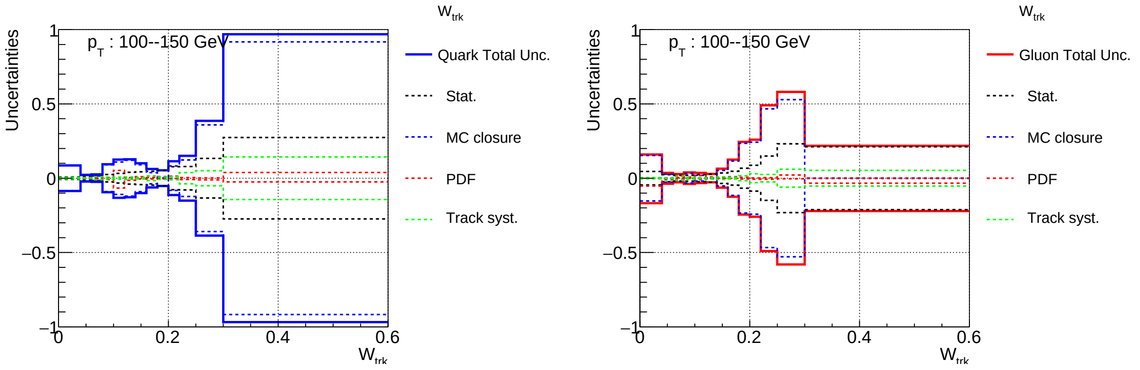


(a) Quark : $100 < p_{\mathrm{T}} < 150$ GeV (Sherpa) (b) Gluon : $100 < p_{\mathrm{T}} < 150$ GeV (Sherpa)

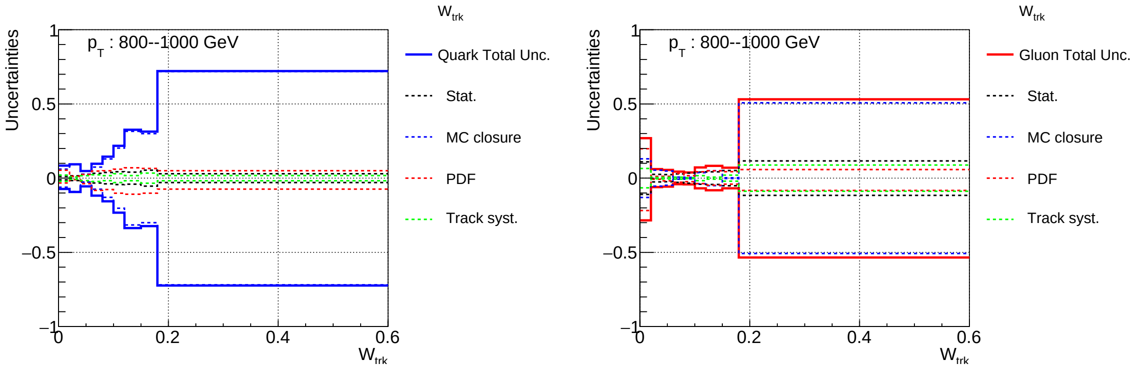


(c) Quark : $800 < p_{\mathrm{T}} < 1000$ GeV (Sherpa) (d) Gluon : $800 < p_{\mathrm{T}} < 1000$ GeV (Sherpa)

Figure 6.9: The W_{trk} distribution of the extracted (a)(c) quark jets and (b)(d) gluon jets from data and MC with the total uncertainties and the total systematic uncertainties obtained by Sherpa MCs. The top two figures show the distribution in the jet p_{T} range between 100 and 150 GeV. The bottom two show that in the jet p_{T} range between 800 and 1000 GeV. The lower panel in each figure shows the extracted data divided by the extracted MC in the upper panels, which corresponds to the W_{trk} SF.



(a) Quark : $100 < p_T < 150$ GeV (Sherpa) (b) Gluon : $100 < p_T < 150$ GeV (Sherpa)



(c) Quark : $800 < p_T < 1000$ GeV (Sherpa) (d) Gluon : $800 < p_T < 1000$ GeV (Sherpa)

Figure 6.10: The fraction of each uncertainty on the SF obtained by Sherpa MCs as a function of W_{trk} for (a)(c) quark jets and (b)(d) gluon jets. The top two figures show the uncertainties in the jet p_{T} range between 100 and 150 GeV. The bottom two show that in the jet p_{T} range between 800 and 1000 GeV. The PDF up/down shows the envelope of all of the PDF uncertainties. The track systematic up/down shows the quadrature sum of five sources of track systematic uncertainties.

7 Event selection

In SUSY searches, there is one difficulty that the optimization is needed to cover wide signal mass range because the mass of the SUSY particles is unknown. In this search, the various mass range of the gluino mass and the lightest neutralino ($\tilde{\chi}_1^0$) mass should be covered. The kinematics of the detected particles in the signal decay strongly depends on the mass difference between the gluino and $\tilde{\chi}_1^0$ since the momenta of the emitted four quarks via decay from gluino to $\tilde{\chi}_1^0$ or $\tilde{\chi}_2^\pm$ come from the extra energy due to the mass difference. To cover such a wide mass range of the signals, ten signal regions (SRs) are defined for each of the mass difference in the direct decay and one-step decay.

7.1 Trigger

An event is decided to be recorded by the High-Level Trigger (HLT), each of which is dedicated to capturing a specific feature in the event. In this search, the large missing transverse energy is one of the important features in the signal event. Therefore, the large $E_{\mathrm{T}}^{\mathrm{miss}}$ trigger is used to pick up signal candidate events. The triggers used for this analysis in the 2015 and 2016 data-sets are listed in Table 7.1. In this table, the triggers used for control regions ("CRs") and validation regions ("VRs") are also listed. These regions are defined to estimate contributions of the SM background process in the SR and check the estimation, which are described in Section 8. The "xe", "g", "e", and "mu" indicate $E_{\mathrm{T}}^{\mathrm{miss}}$, γ, electron, and muon, respectively. The numerical value following these characters represents an energy threshold of the object of interest in a unit of GeV. If there are two ore more triggers listed in one row, the logical OR of them are used to pick up the events. The $E_{\mathrm{T}}^{\mathrm{miss}}$ triggers used for the SR in 2016 have several $E_{\mathrm{T}}^{\mathrm{miss}}$ thresholds because the increasing peak luminosity over the 2016 run had led to unacceptably high rates of the lower threshold triggers.

The trigger efficiency of `HLT_xe110_mht_L1XE50` (the highest $E_{\mathrm{T}}^{\mathrm{miss}}$ threshold trigger) as a function of reconstructed (offline) $E_{\mathrm{T}}^{\mathrm{miss}}$ is shown in Figure 7.1. The efficiency is calculated from the number of events passing the $E_{\mathrm{T}}^{\mathrm{miss}}$ trigger and single lepton triggers divided by the number of events passing single lepton triggers[1]. Generally, the $E_{\mathrm{T}}^{\mathrm{miss}}$ trigger has slow turn-on (increasing part of the efficiency) with respect to the offline $E_{\mathrm{T}}^{\mathrm{miss}}$, in which 200 GeV is needed to assure the plateau efficiency despite the much lower trigger threshold (110 GeV). This is due to the deteriorated resolution of the online $E_{\mathrm{T}}^{\mathrm{miss}}$ calculation only using calorimeter energy deposits, while the offline calculation takes into account of contributions of muons and soft tracks as well. In the figure, the muon contribution is removed as in the offline $E_{\mathrm{T}}^{\mathrm{miss}}$ calculation. This efficiency curve without the muon contribution is similar to the offline $E_{\mathrm{T}}^{\mathrm{miss}}$ in the SR because the SR has no muon due to its no lepton requirement. In this analysis, the SR has 100% trigger efficiency because at least 300 GeV $E_{\mathrm{T}}^{\mathrm{miss}}$ is required.

---

[1] The single lepton triggers are ones used for CRW/T, VRW/T, VRZ.

Table 7.1: The summary of the High-Level Triggers used for each region in this analysis for 2015 or 2016 data.

Region	2015 data	2016 data

SR (0-lepton)	`HLT_xe70_mht`	`HLT_xe90_mht_L1XE50`
		`HLT_xe100_mht_L1XE50`
		`HLT_xe110_mht_L1XE50`

CRY (1-photon)	`HLT_g120_loose`	`HLT_g140_loose`

CRW, CRT (1-lepton)
VRW, VRT (1-lepton)
VRZ (2-lepton)
	`HLT_e60_lhmedium`
	`HLT_e120_lhloose`
	`HLT_e24_lhmedium`
	`HLT_mu20_iloose_L1MU15`
	`HLT_mu50`
		`HLT_e60_lhmedium_nod0`
		`HLT_e140_lhloose_nod0`
		`HLT_e26_lhtight_nod0_ivarloose`
		`HLT_mu26_ivarmedium`
		`HLT_mu50`

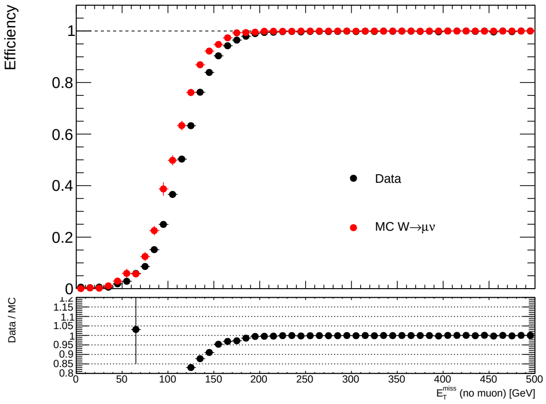


Figure 7.1: The trigger efficiency of `HLT_xe110_mht_L1XE50` for 2016 data and $W \to \mu\nu$ MC as a function of reconstructed (offline) $E_{\mathrm{T}}^{\mathrm{miss}}$ without muon contribution. The data shown here is recorded in a period when `HLT_xe110_mht_L1XE50` is used as a nominal $E_{\mathrm{T}}^{\mathrm{miss}}$ trigger.

7.2 Discriminating variables

The discriminating variables between the signal and SM background used in the event selection are picked up from the conventional SUSY search [45]. They are listed in the following:

N_{jet}

Number of jets reconstructed by the anti-k_t algorithm (Section 3.3) with $p_{\mathrm{T}} > 50$ GeV and $|\eta| < 2.8$.

1st–4th jet p_{T}, $|\eta|$

p_{T} and $|\eta|$ of the leading four jets in p_{T} ordering with $p_{\mathrm{T}} > 50$ GeV and $|\eta| < 2.8$.

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

The missing transverse momentum described in Section 3.4 obtained from the negative vector sum of all hard object momenta and soft term in the transverse plane of the ATLAS detector.

m_{eff}

The scalar sum of $E_{\mathrm{T}}^{\mathrm{miss}}$ and p_{T}s of all jets with $p_{\mathrm{T}} > 50$ GeV and $|\eta| < 2.8$ given as

$$m_{\mathrm{eff}} \equiv E_{\mathrm{T}}^{\mathrm{miss}} + H_{\mathrm{T}} \qquad \left(H_{\mathrm{T}} \equiv \sum_{j}^{N_{\mathrm{jet}}} p_{\mathrm{T}}(j) \right) \tag{7.1}$$

This has been used as the most discriminating variable in the conventional SUSY searches. The signal with a large $\Delta M(\tilde{g}, \tilde{\chi}_1^0)$ has a large m_{eff} and emerges as a broad peak in its distribution, which corresponds to the sum of the momenta carried by the four quarks and two $\tilde{\chi}_1^0$s (and two W bosons) in the final states of the direct (one-step) decay signal.

$E_{\mathrm{T}}^{\mathrm{miss}}/\sqrt{H_{\mathrm{T}}}$, $E_{\mathrm{T}}^{\mathrm{miss}}/m_{\mathrm{eff}}(4\mathrm{j})$

$m_{\mathrm{eff}}(4\mathrm{j})$ uses the leading four jets instead of all jets. Thus, $E_{\mathrm{T}}^{\mathrm{miss}}/m_{\mathrm{eff}}(4\mathrm{j})$ is defined as $E_{\mathrm{T}}^{\mathrm{miss}}/\left(E_{\mathrm{T}}^{\mathrm{miss}} + \sum_j^4 p_{\mathrm{T}}(j)\right)$. $E_{\mathrm{T}}^{\mathrm{miss}}/\sqrt{H_{\mathrm{T}}}$ and $E_{\mathrm{T}}^{\mathrm{miss}}/m_{\mathrm{eff}}(4\mathrm{j})$ mean the $E_{\mathrm{T}}^{\mathrm{miss}}$ largeness relative to the total event hardness (H_{T} or m_{eff}). $E_{\mathrm{T}}^{\mathrm{miss}}/\sqrt{H_{\mathrm{T}}}$ is used for SRs with $N_{\mathrm{jet}} \geq 2$ and $E_{\mathrm{T}}^{\mathrm{miss}}/m_{\mathrm{eff}}(4\mathrm{j})$ is used for SRs with $N_{\mathrm{jet}} \geq 4, 5, 6$. These variables are very powerful to remove multi-jet background since the multi-jet background has only $E_{\mathrm{T}}^{\mathrm{miss}}$ caused by the mis-measurement of the jet momenta and its $E_{\mathrm{T}}^{\mathrm{miss}}$ is small relative to the jet hardness. The details of the multi-jet process will be described in Section 8.1.3.

$\Delta\phi(\mathrm{jet}_{1,2,(3)}, \vec{E}_{\mathrm{T}}^{\mathrm{miss}})_{\mathrm{min}}$, $\Delta\phi(\mathrm{jet}_{i>3}, \vec{E}_{\mathrm{T}}^{\mathrm{miss}})_{\mathrm{min}}$

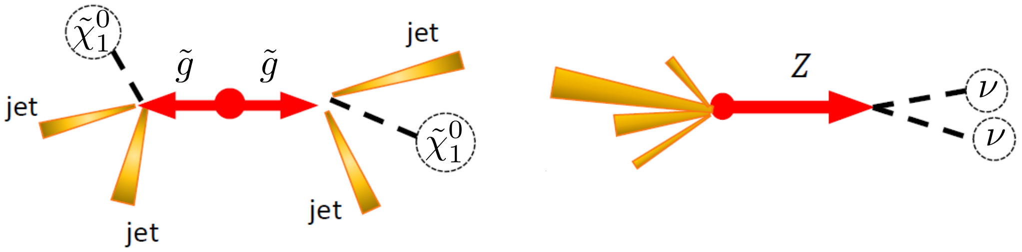
$\Delta\phi(\mathrm{jet}, \vec{E}_{\mathrm{T}}^{\mathrm{miss}})$ is the distance in ϕ direction between a jet and $\vec{E}_{\mathrm{T}}^{\mathrm{miss}}$, which is also powerful to remove the multi-jet background because the $E_{\mathrm{T}}^{\mathrm{miss}}$ in the multi-jet is near to the mis-measured jet and $\Delta\phi(\mathrm{jet}, \vec{E}_{\mathrm{T}}^{\mathrm{miss}})$ is small. $\Delta\phi(\mathrm{jet}_{1,2,(3)}, \vec{E}_{\mathrm{T}}^{\mathrm{miss}})_{\mathrm{min}}$ and $\Delta\phi(\mathrm{jet}_{i>3}, \vec{E}_{\mathrm{T}}^{\mathrm{miss}})_{\mathrm{min}}$, which are the minimum $\Delta\phi(\mathrm{jet}, \vec{E}_{\mathrm{T}}^{\mathrm{miss}})$s in the leading three jets and the remaining jets respectively, are prepared to use $\Delta\phi(\mathrm{jet}, \vec{E}_{\mathrm{T}}^{\mathrm{miss}})$ information in the selection.

Aplanarity

In the signal decay, there are two heavy gluinos each of which produces several jets and the topology of all the jets is isotropic as shown in Figure 7.2(a). In contrast, the topology of the jets in the SM background event is planar. The main SM background process in the SR is $Z(\rightarrow \nu\nu)$+jets or $W(\rightarrow l\nu)$+jets, in which the Z or W boson tends to be boosted to a direction due to the large $E_{\mathrm{T}}^{\mathrm{miss}}$ requirement of the SR selection and the jets are boosted to the opposite direction, which forms a planar shape as shown in Figure 7.2(b). Therefore, the topology of multiple jets is different between the signal and background. This difference is considered in aplanarity, which is the third eigenvalue of the tensor of jets momenta (the definition is written in [56]). This variable has a large value if the directions of the jets spread spherically like the signal.

1st–4th jet W_{trk}

W_{trk} has discriminating power to separate quark- and gluon-jets. In large $\Delta M(\tilde{g}, \tilde{\chi}_1^0)$ signals, this variable is useful because the four quarks emitted from the decay of gluinos have large momenta to become the leading four jets. This variable is calibrated in Section 6. The W_{trk} Scale Factor (SF) obtained there is applied in SRs using this variable.



(a) Signal (Gluino pair-production) (b) The main SM background ($Z \to \nu\nu$)

Figure 7.2: Schematic views of the distribution of jets and $E_{\mathrm{T}}^{\mathrm{miss}}$ sources in (a) signal (gluino pair-production) and (b) the main SM background ($Z \to \nu\nu$).

7.3 Preselection

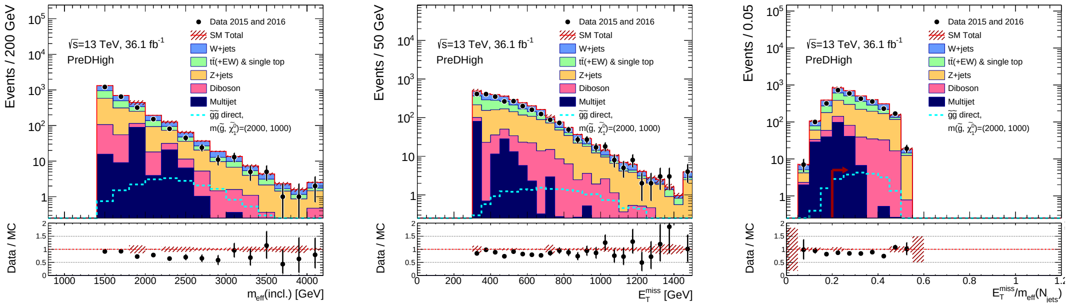
After trigger selection, four sets of selection are defined as preselections, which are referred to as "**PreDHigh/PreDLow**(**PreOHigh/PreOLow**)" aiming for the high and low $\Delta M(\tilde{g}, \tilde{\chi}_1^0)$ in the direct (one-step) decay, respectively. The common cuts among them are lepton veto, $E_{\mathrm{T}}^{\mathrm{miss}} >$ 300 GeV, the leading and the second leading jet p_{T} requirements. N_{jet}, m_{eff}, $\Delta\phi(\mathrm{jet}_{1,2,(3)}, \vec{E}_{\mathrm{T}}^{\mathrm{miss}})_{\mathrm{min}}$, and $E_{\mathrm{T}}^{\mathrm{miss}}/m_{\mathrm{eff}}(4\mathrm{j})$ requirements are varied depending on the target signal $\Delta M(\tilde{g}, \tilde{\chi}_1^0)$. Generally, one-step decay signals have more jets than direct decay signals due to $W \to qq$ decay in the signal decay chain, and the higher $\Delta M(\tilde{g}, \tilde{\chi}_1^0)$ signal has more jets and larger m_{eff} because the jets have larger momenta. These signal features are taken into account in the cuts of the preselection. The $E_{\mathrm{T}}^{\mathrm{miss}}$ and $\Delta\phi(\mathrm{jet}_{1,2,(3)}, \vec{E}_{\mathrm{T}}^{\mathrm{miss}})_{\mathrm{min}}$ cuts can suppress multi-jet background to around 10% or less as shown in Table 7.3. The details of each background process will be described in Section 8.1. The distribution of each variable in PreDHigh is shown in Figures 7.3. For the other preselections, the distributions are shown in Appendix B.1.

Table 7.2: Four sets of selection aiming for different signal $\Delta M(\tilde{g}, \tilde{\chi}_1^0)$ ranges.

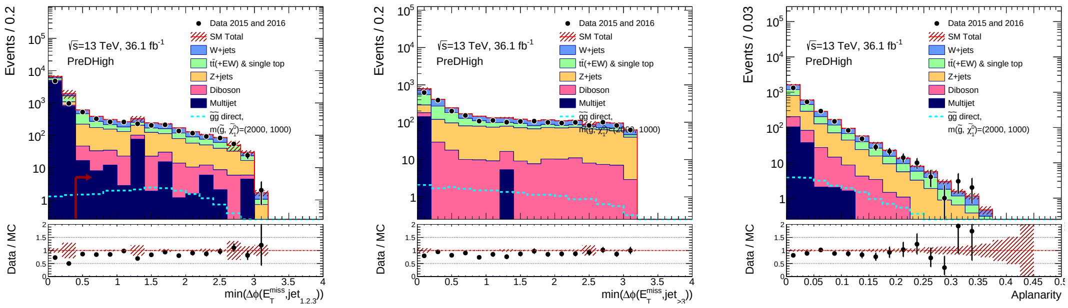
Preselection		PreDHigh	PreDLow	PreOHigh	PreOLow
lepton veto		0 lepton			
$E_{\mathrm{T}}^{\mathrm{miss}}$		> 300 GeV			
1st jet p_{T}		> 200 GeV			
2nd jet p_{T}		> 50 GeV			
Number of jets		≥ 4	≥ 2	≥ 6	≥ 5
m_{eff}			> 1400 GeV		> 800 GeV
$\Delta\phi(\mathrm{jet}_{1,2,(3)}, \vec{E}_{\mathrm{T}}^{\mathrm{miss}})_{\mathrm{min}}$		> 0.4	> 0.2	> 0.4	> 0.2
$E_{\mathrm{T}}^{\mathrm{miss}}/m_{\mathrm{eff}}(4\mathrm{j})$		> 0.2	-	> 0.2	-
Target signal	decay	Gluino Direct decay		Gluino One-step decay	
	$\Delta M(\tilde{g}, \tilde{\chi}_1^0)$	High $\Delta M(\tilde{g}, \tilde{\chi}_1^0)$	Low $\Delta M(\tilde{g}, \tilde{\chi}_1^0)$	High $\Delta M(\tilde{g}, \tilde{\chi}_1^0)$	Low $\Delta M(\tilde{g}, \tilde{\chi}_1^0)$

Preselection	**P**reDHigh	**P**reDLow	**P**reOHigh	**P**reOLow
Diboson	200 ± 8 (6.8%)	625 ± 14 (5.3%)	51 ± 4 (5.7%)	328 ± 10 (4.5%)
Z/γ^{*}+jets	1046 ± 6 (35.7%)	4643 ± 15 (39.1%)	189 ± 3 (21.3%)	1584 ± 11 (21.9%)
W+jets	809 ± 19 (27.6%)	3183 ± 26 (26.8%)	205 ± 4 (23.2%)	1641 ± 17 (22.7%)
$t\bar{t}$ + single top	728 ± 8 (24.8%)	2381 ± 14 (20.0%)	336 ± 5 (38.0%)	2755 ± 16 (38.1%)
Multi-jet	146 ± 68 (5.0%)	1053 ± 317 (8.9%)	104 ± 67 (11.7%)	929 ± 577 (12.8%)
Total MC	2929 ± 72	11884 ± 319	884 ± 67	7238 ± 578

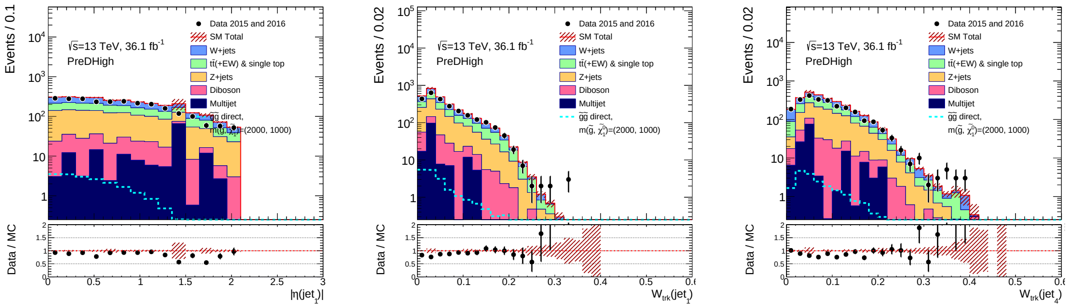
Table 7.3: Numbers of the SM background processes corresponding to 36.1 fb^{-1} at the preselection **PreDHigh**, **PreDLow**, **PreOHigh**, and **PreOLow** expected by the MC simulations.



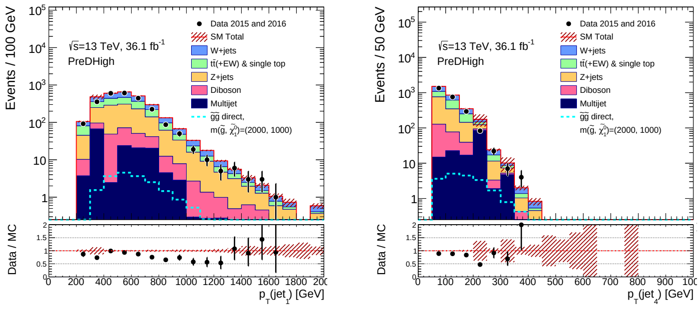
(a) m_{eff}(inc)　(b) $E_{\mathrm{T}}^{\mathrm{miss}}$　(c) $E_{\mathrm{T}}^{\mathrm{miss}}/m_{\mathrm{eff}}$(4j)



(d) $\Delta\phi(\mathrm{j}_{1,2,(3)}, \vec{E}_{\mathrm{T}}^{\mathrm{miss}})_{\mathrm{min}}$　(e) $\Delta\phi(\mathrm{j}_{i>3}, \vec{E}_{\mathrm{T}}^{\mathrm{miss}})_{\mathrm{min}}$　(f) Aplanarity



(g) 1st jet $|\eta|$　(h) 1st jet W_{trk}　(i) 4th jet W_{trk}



(j) 1st jet p_{T}　(k) 4th jet p_{T}

Figure 7.3: Distributions of m_{eff}(inc), $E_{\mathrm{T}}^{\mathrm{miss}}$, $E_{\mathrm{T}}^{\mathrm{miss}}/m_{\mathrm{eff}}$(4j), $\Delta\phi(\mathrm{j}_{1,2,(3)}, \vec{E}_{\mathrm{T}}^{\mathrm{miss}})_{\mathrm{min}}$, $\Delta\phi(\mathrm{j}_{i>3}, \vec{E}_{\mathrm{T}}^{\mathrm{miss}})_{\mathrm{min}}$, Aplanarity, 1st jet $|\eta|$, 1st jet W_{trk}, 4th jet W_{trk}, 1st jet p_{T}, and 4th jet p_{T} in SR of **PreDHigh**. Black points show the $36.1\mathrm{fb}^{-1}$data and filled histograms show the SM background MCs normalized by cross-section. A large bin of the multi-jet is caused by an event with an extreme high event weight reflecting the skimming weight in the MC event generation. This is unphysical value to be ignored.

7.4 Boosted Decision Tree

In this analysis, Boosted Decision Tree (BDT) is used in order to improve the search by taking into account the correlation between the variables including $W_{\rm trk}$s of the leading four jets as discussed in Section 5. BDT has not been used in the SUSY searches because it is necessary to fix the signal kinematics for the BDT training while the possible kinematic phase space of the SUSY signal is wide due to its unknown mass. In this analysis, to cover the various mass range, the BDT training is separated into ten kinds (**D1–D5** and **O1–O5**) of training each of which is aiming for a specific $\Delta M(\tilde{g}, \tilde{\chi}_1^0)$ range. The D1–D5 and O1–O5 BDT trainings are aiming for the gluino direct decay and one-step decay, respectively. The BDT training with a lower number in the name is performed with larger $\Delta M(\tilde{g}, \tilde{\chi}_1^0)$ signal MCs.

7.4.1 BDT training

The kinematic phase space of the signal is similar between the same $\Delta M(\tilde{g}, \tilde{\chi}_1^0)$ signals. For example, Figures 7.4 show the distributions of $\Delta M(\tilde{g}, \tilde{\chi}_1^0) = 1000$ GeV direct decay signals at the preselection PreDHigh. Thus, in each training, signal MCs having similar $\Delta M(\tilde{g}, \tilde{\chi}_1^0)$ and gluino mass are merged. The five target signal mass ranges are prepared for each of the direct decay and one-step decay as shown in Figures 7.5. For the background in the training, all of the SM background MCs are taken into account except for the multi-jet background: Z+jets, W+jets, $t\bar{t}$, single top, and diboson MCs. For both of the signal and background, the MC samples after a specific preselection determined for each training are used. Hence, the multi-jet background is already suppressed as described in Section 7.3. The preselection of each BDT training is shown in Table 7.4.

7.4.2 Input variables

To determine the input variables for each target signal mass range, several kinds of BDT trainings are performed with different sets of input variables and the expected sensitivity of a new SR with a cut on the BDT score is roughly estimated. At first, a method of the estimation is described below.

Expected-sensitivity calculation

The expected sensitivity is simply calculated from three numbers: numbers of the signal and the total SM background estimated by the MC simulation in a given signal region and the systematic uncertainty on the number of the background. For this calculation, approximation formula to calculate the significance (σ) of signal null-hypothesis against the signal plus background is used [58, 59][2].

Here, the uncertainty on the number of the background is estimated from results of 23 SRs ($m_{\rm eff}$-based SRs) in the previous study published in Moriond 2017 conference [57]. The systematic uncertainties in the results strongly depend on the statistic of the SR (=tightness of the selection). In Figure 7.6, 23 points are drawn, each of which corresponds to one of the 23 SRs. Each

---

[2] In practice, a function of ROOT [60], `RooStats::NumberCountingUtils::BinomialObsZ`, is used.

point indicates a square-root of the number of predicted background and total systematic uncertainty for each SR. This graph can be fit well by the following equation,

$$\frac{\Delta N_{\mathrm{bkg}}}{N_{\mathrm{bkg}}} = \mathrm{Max}\left(\frac{a}{\left(\sqrt{N_{\mathrm{bkg}}}\right)^{b}} + c\, ,\, 0.10 \right). \tag{7.2}$$

$$\begin{cases} a = 0.46 \\ b = 1.45 \\ c = 0.085 \end{cases} \tag{7.3}$$

The total systematic uncertainty in a new SR built using each BDT is estimated from the above equation.

Determination of the input variables

The input variables for each BDT training are determined based on 3σ sensitivity contours on the gluino direct or one-step decay signal. A BDT score is obtained from BDT training with each set of input variables and a new BDT SR is defined by a cut on the BDT score, the preselection, and additional cuts on variables not used as input variables of the BDT. The cut values of the BDT score and the additional cuts are optimized to maximize the sensitivity of one of the target signal mass points.

The following rules are defined in determining the input variables:

- $\Delta\phi(\mathrm{jet}_{1,2,(3)}, \vec{E}_{\mathrm{T}}^{\mathrm{miss}})_{\mathrm{min}}$ and $\Delta\phi(\mathrm{jet}_{i>3}, \vec{E}_{\mathrm{T}}^{\mathrm{miss}})_{\mathrm{min}}$ are not included in the input variables because they are used to define a control region for multi-jet background ("CRQ") as described in Section 8.2 , where the multi-jet events are enriched.

- N_{jet} is not used as an input variable due to its large mis-modeling. (For example, it can be seen in Figures 8.11(b) and 8.12(b) in Section 8.2.)

In any BDT training, most of the variables listed in Section 7.2 make some improvement in the sensitivity. For example, Figures 7.7 show difference of the sensitivity contours with and without one variable in D3. However, it is found that some variables are unnecessary in specific BDT trainings:

- $E_{\mathrm{T}}^{\mathrm{miss}}/m_{\mathrm{eff}}(4\mathrm{j})$ does not improve the sensitivity in D1–D4 and O1–O3. This can be seen in Figures 7.8.

- Aplanarity is not necessary for D5 and O4–O5. This can be seen in Figures 7.9.

- W_{trk}s of the leading four jets:
  - The W_{trk}s do not make any improvement in D5 due to the low quark fraction in the low gluino mass signal as shown in Figure 5.1 in Section 5.

- In D1 and O1, the W_{trk}s are also dropped from the input variables because they do not improve the sensitivity significantly in the gluino mass direction as can be seen in Figures 7.8(a) and 7.8(e). The reason for the small improvement is that the m_{eff} is enough to distinguish the signal from the SM background for the highest $\Delta M(\tilde{g}, \tilde{\chi}_1^0)$ signals and only the signal cross-section determines the upper limit on the sensitivity.
- In contrast, the W_{trk}s are added as input variables in O5 for the one-step decay because $\sim$ 50 GeV improvement in the $\tilde{\chi}_1^0$ mass direction can be seen as shown in Figure 7.9(c). The improvement is caused by higher quark multiplicity than the direct decay due to $W \to qq$ decay.

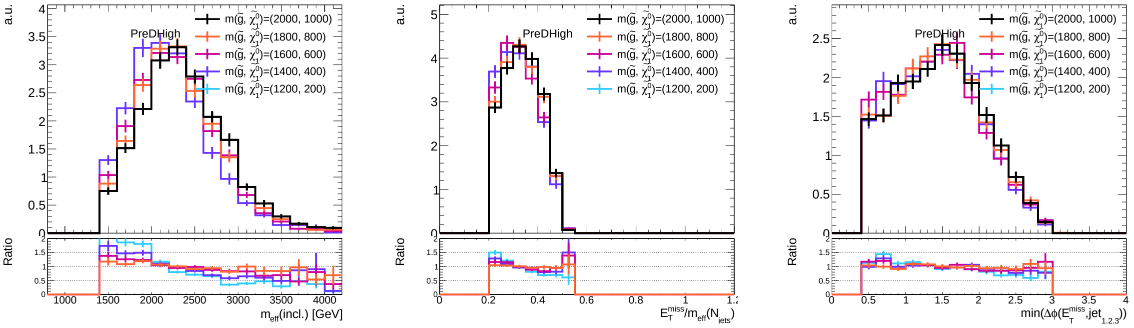
From the above, input variables for each BDT training are determined as in Table 7.4.

Table 7.4: The summary of the preselection and input variables used in each BDT training. The $\mathrm{j_i}$ represents the i th leading p_{T} jet.

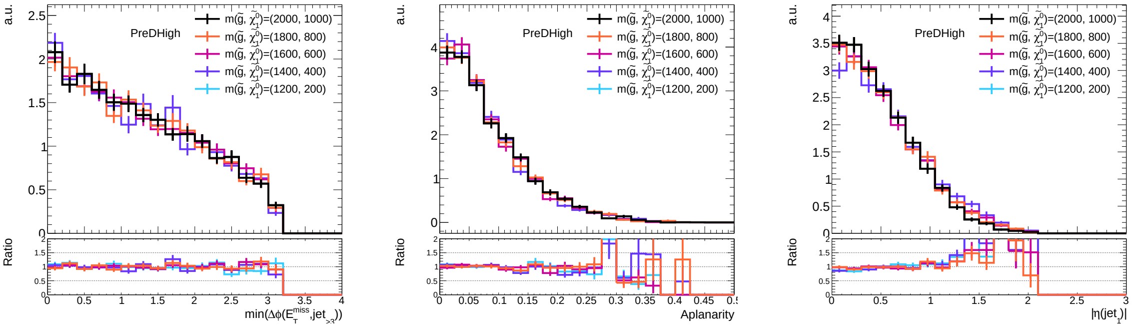
Signal region	D1	D2, D3, D4	D5	O1	O2, O3	O4, O5
Preselection	PreDHigh ($N_{\mathrm{jet}} \geq 4$)		PreDLow ($N_{\mathrm{jet}} \geq 2$)	PreOHigh ($N_{\mathrm{jet}} \geq 6$)		PreOLow ($N_{\mathrm{jet}} \geq 5$)
$p_{\mathrm{T}}(\mathrm{j}_1)$, $p_{\mathrm{T}}(\mathrm{j}_2)$, $p_{\mathrm{T}}(\mathrm{j}_3)$, $p_{\mathrm{T}}(\mathrm{j}_4)$	○	○	○	○	○	○
$\lvert\eta(\mathrm{j}_1)\rvert$, $\lvert\eta(\mathrm{j}_2)\rvert$, $\lvert\eta(\mathrm{j}_3)\rvert$, $\lvert\eta(\mathrm{j}_4)\rvert$	○	○	○	○	○	○
m_{eff}	○	○	○	○	○	○
$E_{\mathrm{T}}^{\mathrm{miss}}/\sqrt{H_{\mathrm{T}}}$ or $E_{\mathrm{T}}^{\mathrm{miss}}/m_{\mathrm{eff}}(4\mathrm{j})$	-	-	$E_{\mathrm{T}}^{\mathrm{miss}}/\sqrt{H_{\mathrm{T}}}$	-	-	$E_{\mathrm{T}}^{\mathrm{miss}}/m_{\mathrm{eff}}(4\mathrm{j})$
Aplanarity	○	○	-	○	○	-
W_{trk} (j_1), W_{trk} (j_2), W_{trk} (j_3), W_{trk} (j_4)	-	○	-	-	○	○
Number of variables	10	14	10	10	14	14

7.4.3 BDT score distribution in the signal region

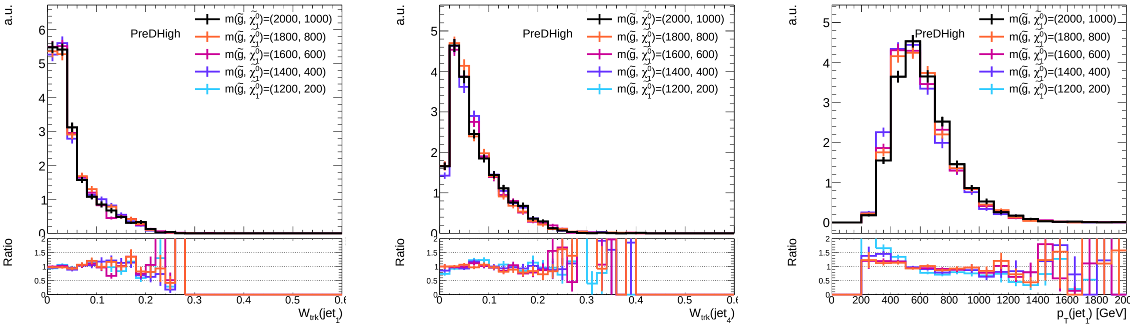
The distributions of the BDT scores are shown in Figures 7.10 and 7.11. The SM background is the MC simulation normalized by the cross-section. A BDT score of a signal in each the target mass range is superimposed on each figure. The signal and background can be separated by the BDT score successfully. However, as a tendency of the discriminating power of the BDT scores, a BDT aiming for smaller $\Delta M(\tilde{g}, \tilde{\chi}_1^0)$ has smaller separation power because the smaller ΔM signal has softer jets and it is more difficult to separate the signal and background. For example, in Figures 7.10(e) and 7.11(e) for D5 and O5, the signals spread to the low BDT score.



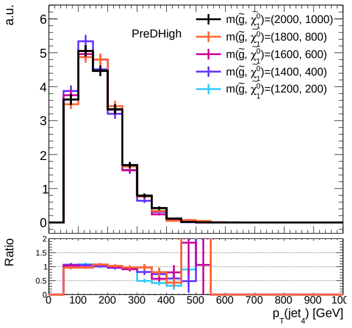
(a) m_{eff}(inc) (b) $E_{\mathrm{T}}^{\mathrm{miss}}/m_{\mathrm{eff}}(4\mathrm{j})$ (c) $\Delta\phi(\mathrm{j}_{1,2,(3)}, \vec{E}_{\mathrm{T}}^{\mathrm{miss}})_{\mathrm{min}}$



(d) $\Delta\phi(\mathrm{j}_{i>3}, \vec{E}_{\mathrm{T}}^{\mathrm{miss}})_{\mathrm{min}}$ (e) Aplanarity (f) 1st jet $|\eta|$

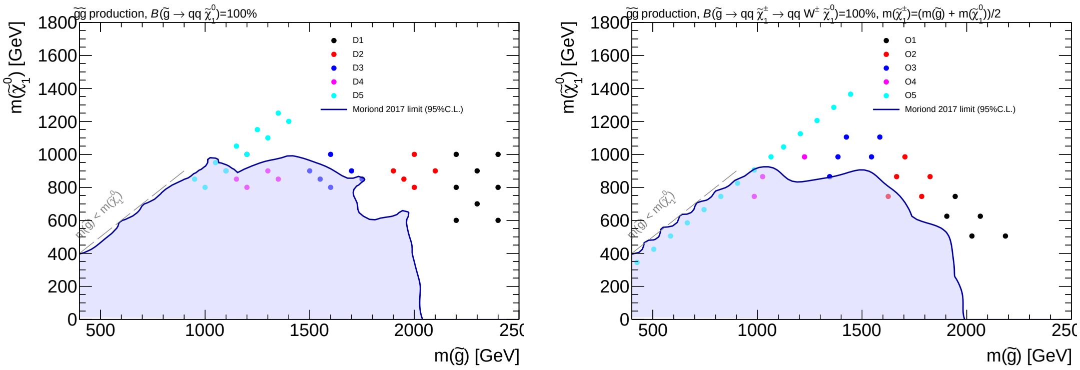


(g) 1st jet W_{trk} (h) 4th jet W_{trk} (i) 1st jet p_{T}



(j) 4th jet p_{T}

Figure 7.4: Distributions of m_{eff}(inc), $E_{\mathrm{T}}^{\mathrm{miss}}/m_{\mathrm{eff}}(4\mathrm{j})$, $\Delta\phi(\mathrm{j}_{1,2,(3)}, \vec{E}_{\mathrm{T}}^{\mathrm{miss}})_{\mathrm{min}}$, $\Delta\phi(\mathrm{j}_{i>3}, \vec{E}_{\mathrm{T}}^{\mathrm{miss}})_{\mathrm{min}}$, Aplanarity, 1st jet $|\eta|$, 1st jet W_{trk}, 4th jet W_{trk}, 1st jet p_{T}, and 4th jet p_{T} in $\Delta M(\tilde{g}, \tilde{\chi}_1^0) = 1000$ GeV signals at the preselection **PreDHigh**. The total numbers of each signals are normalized to be the same.



(a) Gluino direct decay

(b) Gluino one-step dcay

Figure 7.5: Signal mass points used for each of the ten BDT trainings shown in the two-dimensional mass plan of the gluino mass and the $\tilde{\chi}_1^0$ mass. Five BDT trainings (D1–D5) are aiming for the direct decay and the other five trainings (O1–O5) are for the one-step decay. The same colored points are used for the same BDT training in each signal decay. The blue line shows the observed exclusion limit in the previous study [57] at 95% confidence level.

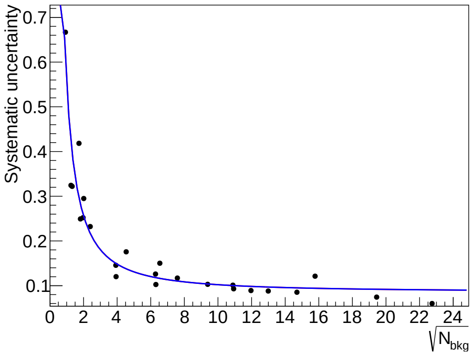


Figure 7.6: Ratio of total systematic uncertainty on a number of predicted background $\Delta N_{\mathrm{bkg}}/N_{\mathrm{bkg}}$ v.s. square-root of the number $\sqrt{N_{\mathrm{bkg}}}$ of 23 SRs (m_{eff}-based SRs) in the previous study [57], and a fitted function given by Eq. 7.2. A point corresponds one SR of the 23 SRs.

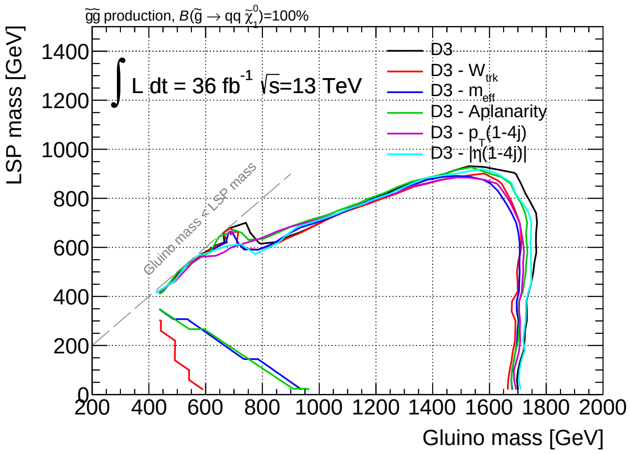
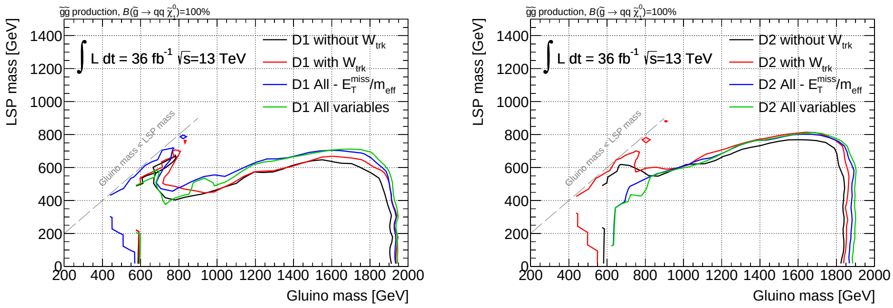
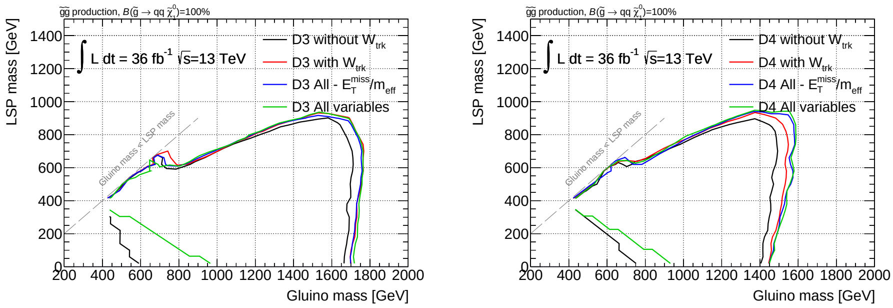


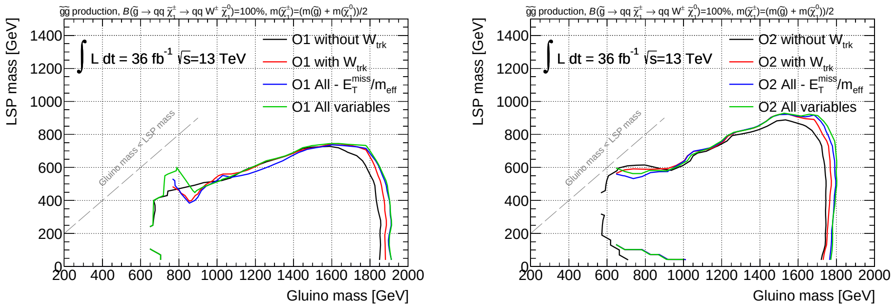
Figure 7.7: 3σ sensitivity contours of BDT SRs with different sets of input variables for D3 in the gluino direct decay signal. A black line shows the contour using input variables listed in Table 7.4 (The final set of input variables). Red, blue, green, purple, and cyan lines show the contours removing W_{trk} of the four jets, m_{eff}, aplanarity, p_{T}s, and $|\eta|$s of the four jets from the black line, respectively. There is some deterioration around the target mass region ($m(\tilde{g}) \sim 1.7$ TeV and $m(\tilde{\chi}_1^0) \sim 0.9$ TeV) in any BDT SRs with reduced input variables from the black line. The x-axis is gluino mass and the y-axis is LSP mass, i.e. $\tilde{\chi}_1^0$ mass.



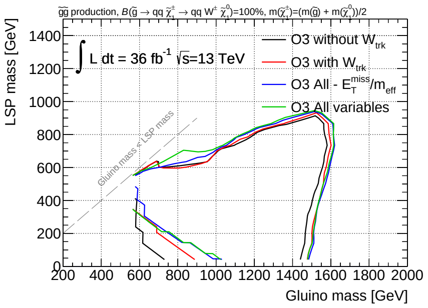
(a) D1 (b) D2



(c) D3 (d) D4

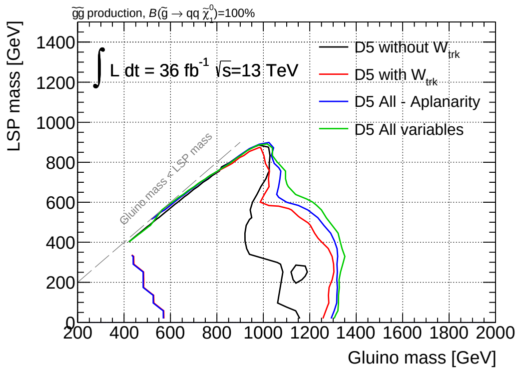


(e) O1 (f) O2

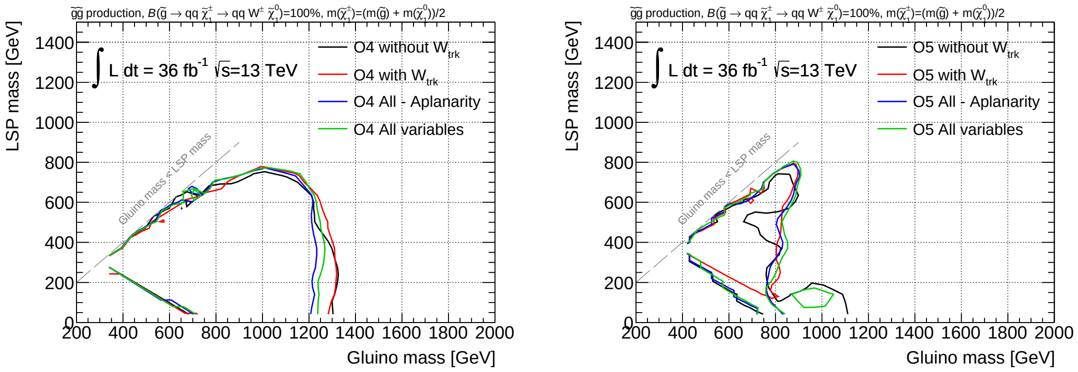


(g) O3

Figure 7.8: 3σ sensitivity contours of BDT SRs with different sets of input variables for (a) D1, (b) D2, (c) D3, (d) D4, (e) O1, (f) O2, and (g) O3 target signals. Black lines show the contours using input variables listed in Table 7.4 except W_{trk}s of the leading four jets. Red lines show the contours using W_{trk}s in addition. Green lines show the contours using all the variables listed in Section 7.2. Blue lines show the contours using all the variables except $E_{\mathrm{T}}^{\mathrm{miss}}/m_{\mathrm{eff}}(4\mathrm{j})$. The difference between the green and blue lines are very small. The x-axis is gluino mass and the y-axis is LSP mass, i.e. $\tilde{\chi}_1^0$ mass.



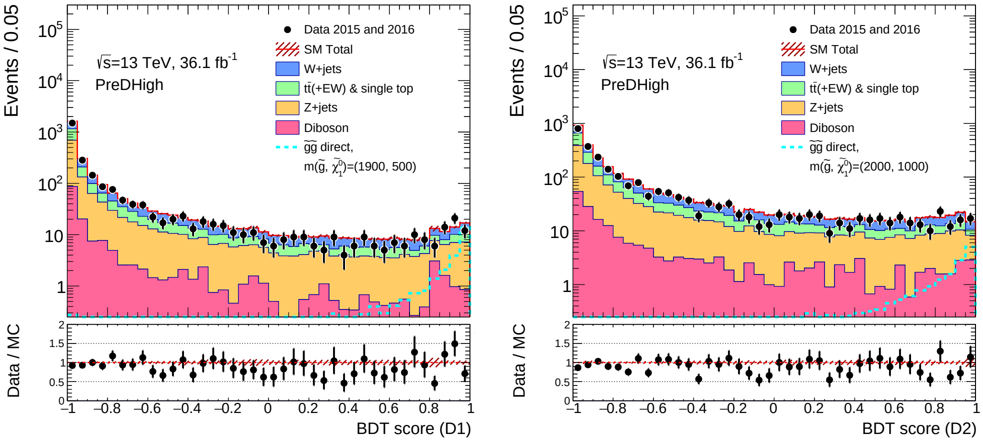
(a) D5



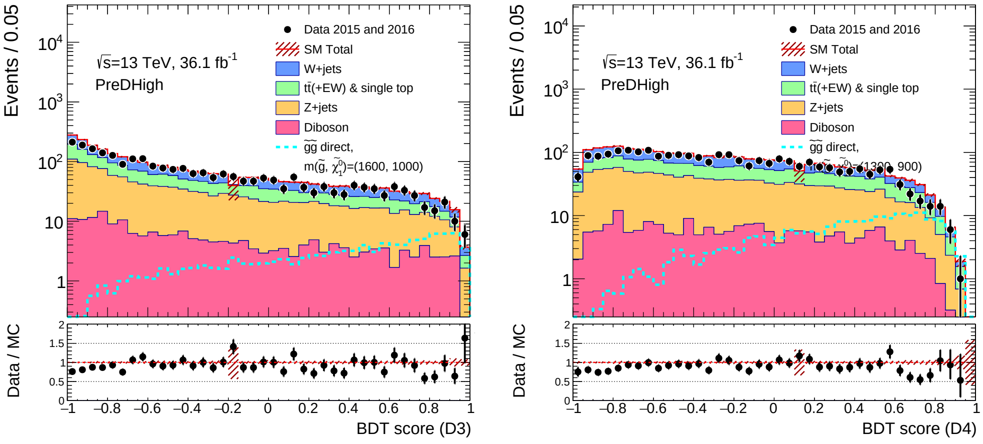
(b) O4

(c) O5

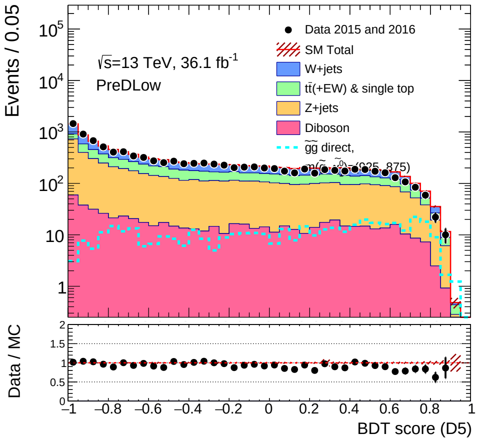
Figure 7.9: 3σ sensitivity contours of BDT SRs with different sets of input variables for (a) D5, (b) O4, and (c) O5 target signals. Black lines show the contours using input variables listed in Table 7.4 except W_{trk}s of the leading four jets. Red lines show the contours using W_{trk}s in addition. Green lines show the contours using all the variables listed in Section 7.2. Blue lines show the contours using all the variables except aplanarity. The difference between the green and blue lines are very small. The x-axis is gluino mass and the y-axis is LSP mass, i.e. $\tilde{\chi}_1^0$ mass.



(a) D1 (b) D2

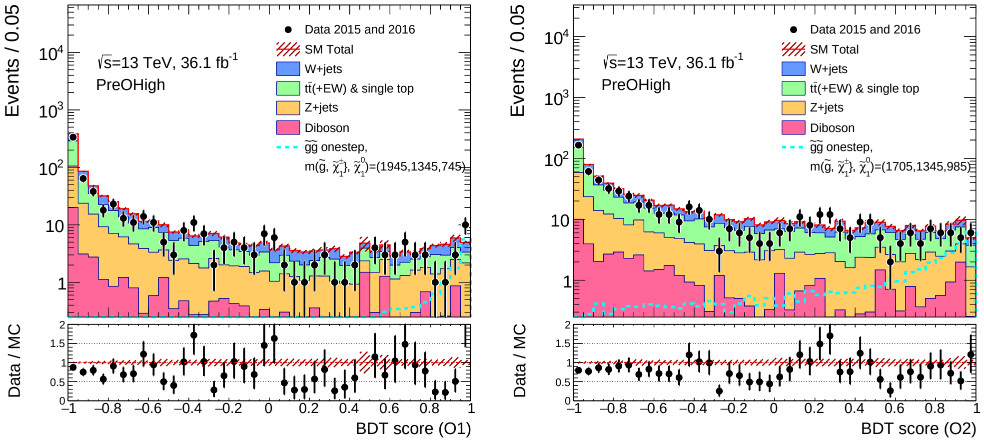


(c) D3 (d) D4

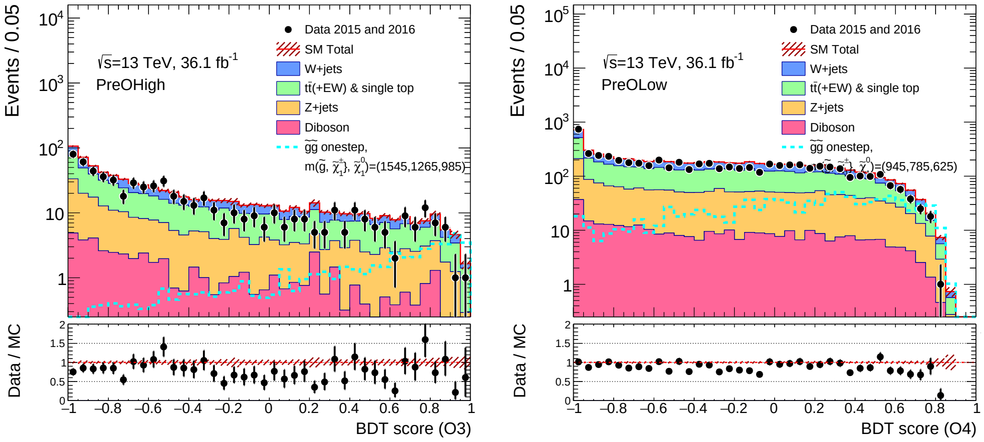


(e) D5

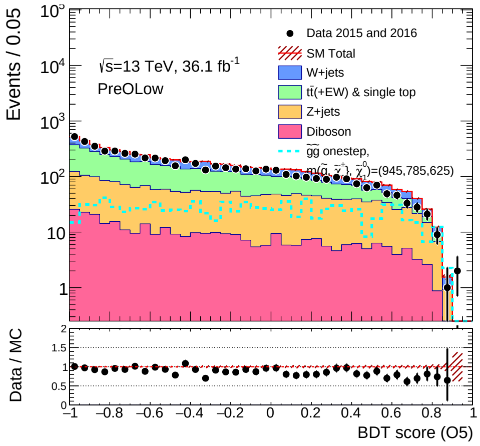
Figure 7.10: Distributions of BDT scores for (a)D1, (b)D2, (c)D3, (d)D4, and (e)D5 in the 36.1fb^{-1}data, the SM background MC, and signal MC that is one of each target mass points. The selections for D1–D4 and D5 are the preselection PreDHigh and PreDLow, respectively. The background and signal MCs are normalized by the cross-section. The multi-jet process is removed here because its number of events is small but some statistical high fluctuation exists, which should be ignored.



(a) O1 (b) O2



(c) O3 (d) O4



(e) O5

Figure 7.11: Distributions of BDT scores for (a)O1, (b)O2, (c)O3, (d)O4, and (e)O5 in the 36.1fb^{-1}data, the SM background MC, and signal MC that is one of each target mass points. The selections for O1–O3 and O4–O5 are the preselection PreOHigh and PreOLow, respectively. The background and signal MCs are normalized by the cross-section. The multi-jet process is removed here because its number of events is small but some statistical high fluctuation exists, which should be ignored.

7.4.4 Correlation between two input variables

The BDT analysis uses the correlation between input variables. Thus, if the input MC has large mis-modeling in the correlation, the behavior of the BDT score should have mis-modeling as well. The correlations between two of the input variables are checked here by the comparison between the data and MC in the γ control region (CRY) defined in Section 8.2, which is used to estimate the main background process $Z(\rightarrow \nu\nu)$+jets. Table 7.5 shows the comparison in five input variables after PreDHigh preselection. There is no significant difference between the data and MC, and also it was checked that PreDLow, PreOHigh, and PreOLow have no significant difference in the correlations as shown in Appendix B.2.

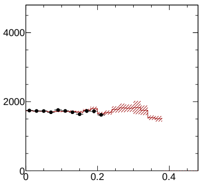
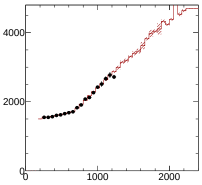
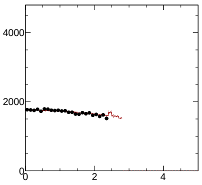
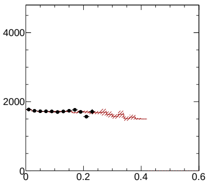
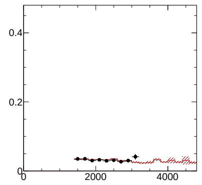
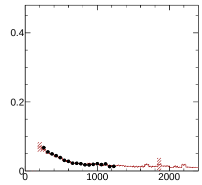
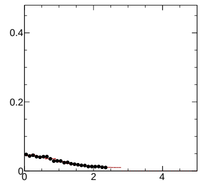
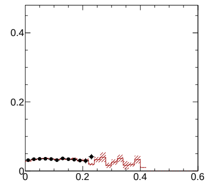
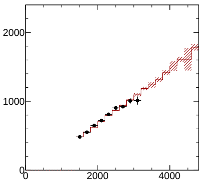
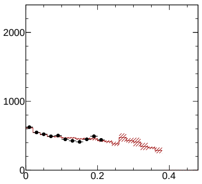
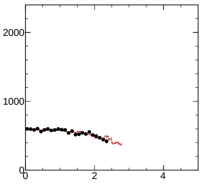
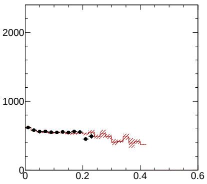
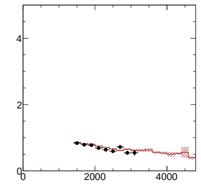
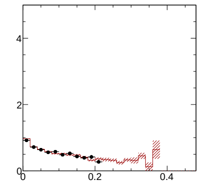
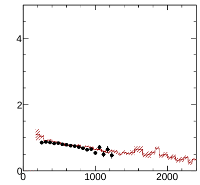
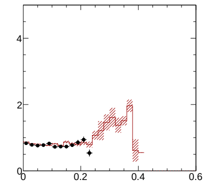
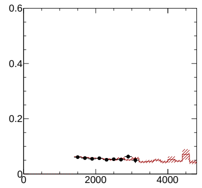
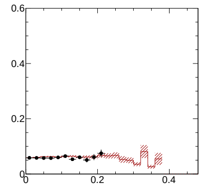
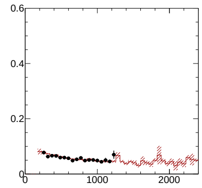
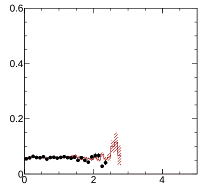
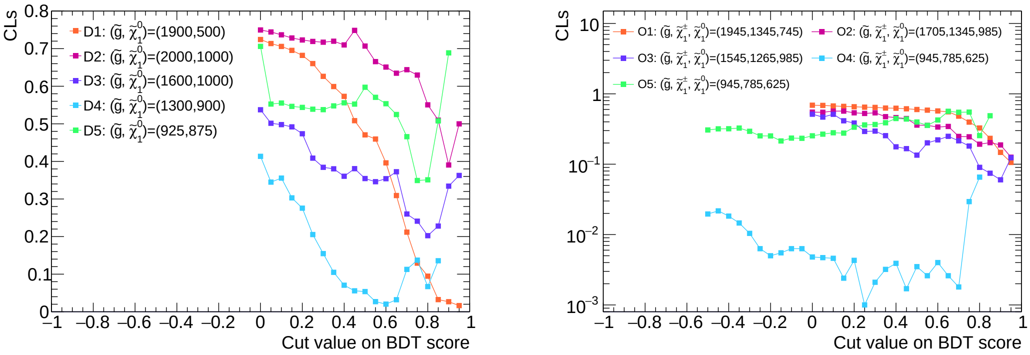
	$m_{\text{eff}}(\text{inc})$	Aplanarity	1^{st} jet p_{T}	1^{st} jet $\lvert\eta\rvert$	1^{st} jet W_{trk}
$m_{\text{eff}}(\text{inc})$					
Aplanarity					
1^{st} jet p_{T}					
1^{st} jet $\lvert\eta\rvert$					
1^{st} jet W_{trk}					

Table 7.5: This table shows profiles between two variables of five input variables in CRY at the preselection **PreDHigh**. The y-axis in each figure is mean of one variable in each x bin of the other variable. Each top title indicates a variable of the x-axis of figures in each column. Each left title indicates a variable of the y-axis of figures in each row. Black point shows 36.1fb^{-1}data and the red line shows the total SM backgound, which is normalized by cross-section.

7.5 Signal region definitions



(a) Gluino direct decay

(b) Gluino one-step dcay

Figure 7.12: BDT-cut scans for optimization of the BDT score cuts in the five SRs for (a) the gluino direct and (b) one-step decay signals, respectively. The x-axis is lower-cut value on the BDT score and the y-axis is expected CLs of each BDT SR with the given cut. The signal mass used the CLs calculation is one of target signal masses for each BDT.

Ten SRs are defined by a cut on each BDT score, the preselection, and additional cuts on $\Delta\phi(\mathrm{jet}_{1,2,(3)}, \vec{E}_\mathrm{T}^\mathrm{miss})_\mathrm{min}$, $\Delta\phi(\mathrm{jet}_{i>3}, \vec{E}_\mathrm{T}^\mathrm{miss})_\mathrm{min}$, and N_jet. Each of SR has a BDT score dedicated to the specific signal mass range. All sets of the selection for the direct decay and one-step decay are summarized in Tables 7.7 and 7.7, respectively.

The $|\eta|$ cut reflects the valid η range of the W_trk calibration (Section 6). The $\Delta\phi(\mathrm{jet}_{1,2,(3)}, \vec{E}_\mathrm{T}^\mathrm{miss})_\mathrm{min}$, $\Delta\phi(\mathrm{jet}_{i>3}, \vec{E}_\mathrm{T}^\mathrm{miss})_\mathrm{min}$, N_jet and BDT score cuts are optimized to maximize the expected exclusion sensitivity (CL_s, Section 9.2.2) calculated by the exact same background estimation, statistical treatment, and systematic uncertainties described in Sections 8–9. BDT-cut scans in ten SRs are shown in Figures 7.12. Some BDT cuts are loosened from the optimal cut value to increase statistics in control regions, which are defined by the exact same BDT cut as in the SR and used to estimate the background as described in Section 8.2. As a tendency for the BDT cut, a BDT aiming for a smaller $\Delta M(\tilde{g}, \tilde{\chi}_1^0)$ signal has a lower BDT cut because it is more difficult to separate the signal and background and the cut is more loosened to keep more signal events.

Table 7.6: The summary of the SR definitions for the gluino direct decay. A BDT score for each SR is trained for the target signal mass shown in Figure 7.5(a).

Signal region	D1	D2	D3	D4	D5
Preselection	PreDHigh				PreDLow
$\lvert\eta(\mathrm{j}_{1-4})\rvert$	-	< 2.1			-
$\Delta\phi(\mathrm{jet}_{1,2,(3)}, \vec{E}_{\mathrm{T}}^{\mathrm{miss}})_{\mathrm{min}}$	>0.4	>0.6	>0.6	>0.4	>0.2
$\Delta\phi(\mathrm{jet}_{i>3}, \vec{E}_{\mathrm{T}}^{\mathrm{miss}})_{\mathrm{min}}$	>0.2	>0.4	>0.4	>0.2	>0.1
BDT score	>0.90	>0.80	>0.80	>0.60	>0.75
Training signal mass $\Delta M(\tilde{g}, \tilde{\chi}_1^0)$	$\sim$ 1.5 TeV	$\sim$ 1 TeV	$\sim$ 500 GeV	$\sim$ 300 GeV	$\sim$ 150 GeV

Table 7.7: The summary of the SR definitions for the gluino one-step decay. A BDT score for each SR is trained for the target signal mass shown as in Figure 7.5(b).

Signal region	O1	O2	O3	O4	O5
Preselection	PreOHigh			PreOLow	
$\lvert\eta(\mathrm{j}_{1-4})\rvert$	-	< 2.1		< 2.1	
$\Delta\phi(\mathrm{jet}_{1,2,(3)}, \vec{E}_{\mathrm{T}}^{\mathrm{miss}})_{\mathrm{min}}$	>0.4			>0.4	
$\Delta\phi(\mathrm{jet}_{i>3}, \vec{E}_{\mathrm{T}}^{\mathrm{miss}})_{\mathrm{min}}$	>0.2			>0.2	
Number of jets (N_{jet})	-	-	-	≥ 6	-
BDT score	>0.80	>0.70	>0.50	>0.00	>-0.15
Training signal mass $\Delta M(\tilde{g}, \tilde{\chi}_1^0)$	$\sim$ 1.3 TeV	$\sim$ 900 GeV	$\sim$ 500 GeV	$\sim$ 200 GeV	$\sim$ 80 GeV

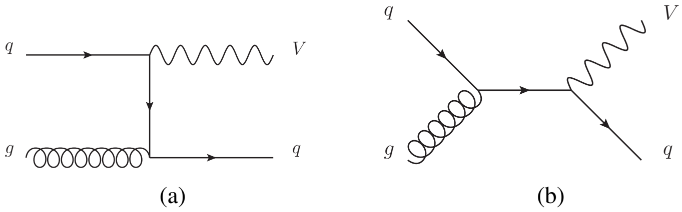
8 Background estimation

8.1 SM background process

The main background for this analysis is Z+jets, W+jets, and top-quark production composed from top and anti-top quark production ($t\bar{t}$) and single top quark production after the preselection as can be seen in Table 7.3. Multi-jet process is sub-dominant after the preselection but it will be killed by the BDT score cut in the SRs. Diboson production process is not dominant but also considered in the analysis.

8.1.1 Vector boson + jets background (Z, W, or γ+jets)

Figures 8.1 show Feynman diagrams of the main process of vector boson + one jet productions. Background process in the SR has at least one more jet than these diagrams due to the requirement of the number of jets. The additional jet is often produced as a gluon-jet since a gluon can radiate from quark lines of the Feynman diagrams as an initial state radiation ("ISR") or final state radiation ("FSR").



(a) (b)

Figure 8.1: Main process of one vector boson ($V = Z, W$, or γ) plus one parton (quark q or gluon g) production at tree-level.

Z decay The decay process of a Z boson is shown in Figure 8.2. The decay branching ratio is approximately 20 % to neutrinos, 10 % to leptons, and 70 % to hadrons. If the Z is boosted (has high p_{T}), $Z \to \nu\nu$ process can make large $E_{\mathrm{T}}^{\mathrm{miss}}$. Hence, $Z \to \nu\nu$ process is one of the dominant processes in the SM background after the selection of this analysis.

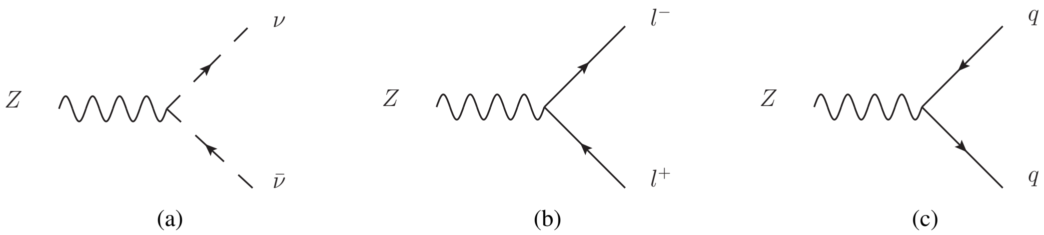


Figure 8.2: Z boson decay process. l stands for e, μ, or τ.

***W* decay** The decay process of W boson is shown in Figure 8.3. The decay branching ratio is approximately 11 % to a lepton plus a neutrino for each lepton flavor and 67 % to hadrons. The neutrino in the $W \to l\nu$ process makes $E_{\mathrm{T}}^{\mathrm{miss}}$. Thus, the $W \to l\nu$ process is a possible remaining background process after the selection. However, because this analysis requires also no reconstructed lepton (only e or μ), the lepton in this decay is restricted to hadronic decaying τ, or an undetected or not-reconstructed lepton due to out-of-acceptance or inefficiency of the particle reconstruction. τ has two decay modes of leptonic decay and hadronic decay. In the leptonic decay, τ decays to a lepton and two neutrinos via a W boson, whose branching ratio is $\sim$ 35%. In the hadronic decay, it decays to one or three charged pions and neutral particles or neutrinos, which is detected as a jet with one or three tracks in the inner detector (ID). The branching ratio of hadronic decay is $\sim$ 65%.

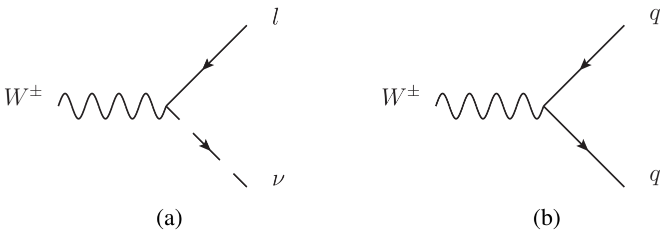


Figure 8.3: W boson decay process. l stands for e, μ, or τ.

γ+jets γ+jets process is not the remaining process after the selection. However, because its production process is similar to that of Z+jets in apart from the difference between γ and Z, it is used to estimate the yield of Z+jets events in this analysis. The γ can be directly identified by using the ID and electromagnetic calorimeter information (Appendix A.2.2).

8.1.2 Top background

In the top quark production, the dominant process is top and anti-top pair production ($t\bar{t}$). The sub-dominant production is single top production. The $t\bar{t}$ production is produced by the process from two partons in collided protons to two partons. The single top production is produced by the Feynman diagrams shown in Figure 8.4. The top quark decays to $W + b$-quark, which possibly makes $E_{\mathrm{T}}^{\mathrm{miss}}$ in case that the W decays to $l + \nu$. In the $t\bar{t}$ decay, only the decay process with leptonic decaying Ws such as shown in Figure 8.5 is possible to remain after the $E_{\mathrm{T}}^{\mathrm{miss}}$ requirement. As a feature of the $t\bar{t}$ decay, the number of jets in the $t\bar{t}$ process is large since the final state has two b-quarks and two quarks coming from a hadronic decaying W.

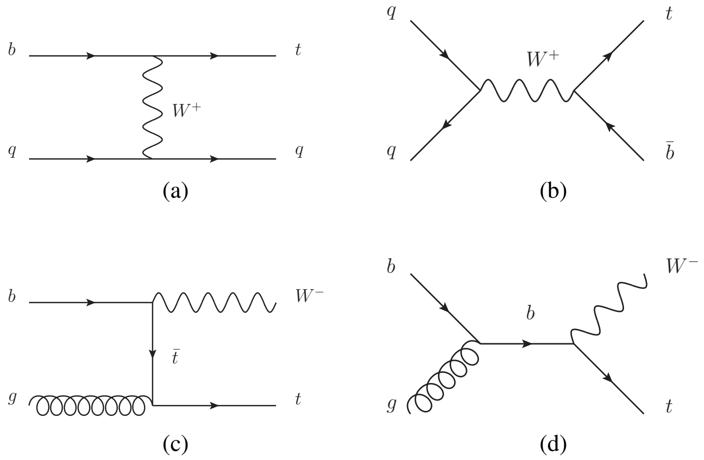


Figure 8.4: Main process of single top-quark production at tree-level. (a)(b) One top-quark + another flavor quark production. (c)(d) One top-quark associated with a W boson production.

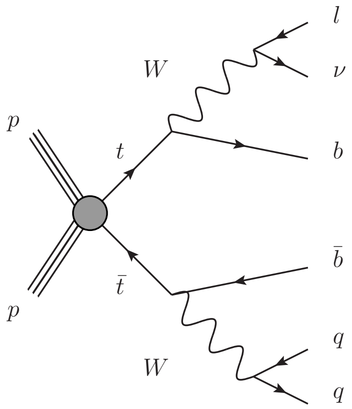


Figure 8.5: Decay process of top (t) and anti-top ($\bar{t}$) pair production with one lepton in the final state. In one hand of $t\bar{t}$, W boson decays to a lepton plus a neutrino. In the other hand of $t\bar{t}$, W decays hadronically.

8.1.3 Multi-jet background

The main process of dijet production, which has two jets in the final states and is the simplest case of the multi-jet process, is shown in Figure 8.6. In the process from two partons to two partons, the dominant process is $gg \to gg$ and the sub-dominant is $qg \to qg$, because the probability of initial gluons is larger than that of initial quarks due to the PDF and a gluon-gluon-gluon coupling is larger than a gluon-quark-quark coupling by the difference of the color factors. The color factor of gluons and quarks is $C_A = 3$, $C_F = 4/3$, respectively. Since multi-jet process have no neutrino, there is no $E_{\mathrm{T}}^{\mathrm{miss}}$-origin in the physics process. However, $E_{\mathrm{T}}^{\mathrm{miss}}$ can be caused by the detector resolution or miss-measurement of the jet momenta. The illustration of $E_{\mathrm{T}}^{\mathrm{miss}}$ coming from miss-measurement of a jet p_{T} is shown in Figure 8.7. The direction of this "fake" $E_{\mathrm{T}}^{\mathrm{miss}}$ is close to that of the jet in ϕ. This feature is used in the $\Delta\phi(\mathrm{jet}, \vec{E}_{\mathrm{T}}^{\mathrm{miss}})$cuts of the preselection to reduce multi-jet background.

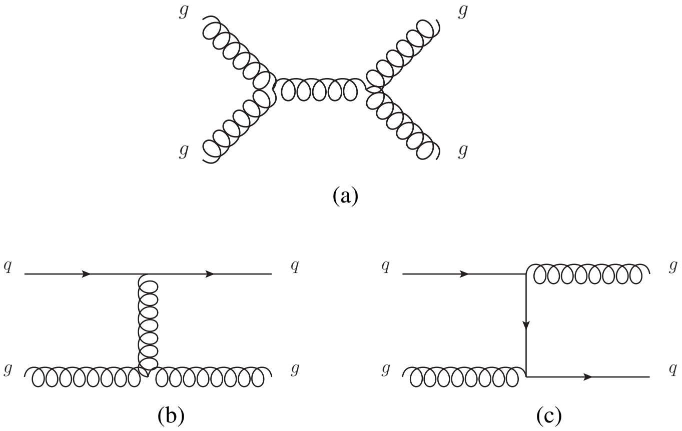


Figure 8.6: Main process of dijet production at tree-level of (a) $gg \rightarrow gg$, (b)(c) $qg \rightarrow qg$ (q: quark, g: gluon).

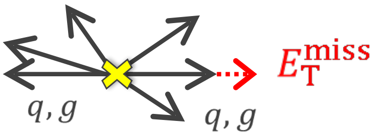


Figure 8.7: A schematic view of multi-jet process with missing transverse momentum ($E_{\mathrm{T}}^{\mathrm{miss}}$) in the transverse plane. The $E_{\mathrm{T}}^{\mathrm{miss}}$ (red break arrow) is caused by the detector resolution or miss-measurement of a jet momentum. The direction of this "fake" $E_{\mathrm{T}}^{\mathrm{miss}}$ is close to that of the jet in the transverse plane.

8.1.4 Diboson background

The diboson production means the process having two bosons of W, Z, and γ, i.e. production of WW, ZZ, $\gamma\gamma$, WZ, $W\gamma$, or $Z\gamma$. Due to the large $E_{\mathrm{T}}^{\mathrm{miss}}$ requirement of this analysis, only the production process with a leptonic decaying W or Z decaying to neutrinos remains after the selection.

8.2 Background estimation method

In order to estimate the numbers of events of the four processes in the SR, four dedicated Control Regions ("CR") are prepared: CRY (for $Z(\rightarrow \nu\nu)$), CRW (for W+jets), CRT(for $t\bar{t}$ and single top),

Table 8.1: A summary of main selection and target background process in each Control Region.

Region	SR	CRY	CRW	CRT	CRQ
Target process	Signal	$Z(\rightarrow \nu\nu)$ +jets	$W(\rightarrow l\nu)$ +jets	$t\bar{t}$ & single top	Multi-jet
Main selection	0 lepton + large $E_{\mathrm{T}}^{\mathrm{miss}}$	1 photon	1 lepton & b-veto	1 lepton & At least 1 b-jet	Low $\Delta\phi(\mathrm{jet}_{1,2,(3)}, \vec{E}_{\mathrm{T}}^{\mathrm{miss}})_{\mathrm{min}}$

and CRQ (for multi-jet). Each CR is designed to enhance the contribution of the background process of interest, which is defined to be orthogonal to the SR. The main selection in each CR is listed in Table 8.1. The CRY requires one photon. In the CRY, the γ is taken into account as $\nu\nu$, i.e. $E_{\mathrm{T}}^{\mathrm{miss}}$. Due to the similar topology between Z+jets and γ+jets, the kinematic phase space of γ+jets is similar to that of $Z(\rightarrow \nu\nu)$ +jets if the γ+jets if the γ is treated as $E_{\mathrm{T}}^{\mathrm{miss}}$. The CRW and CRT are required to have exact one lepton (one electron or one muon). The background process estimated by the CRT includes the single top process since its contribution is much smaller than $t\bar{t}$ and the influence of this treatment on the background estimation is also small. Hereafter, $t\bar{t}$ and single top processes are collectively referred to as "top process". To enrich the top process, b-jet, which can be produced in the decay of the top-quark, is required in the CRT. In contrast, it is required in the CRW that b-jet does not exist in order to suppress the top process. The $\Delta\phi(\mathrm{jet}_{1,2,(3)}, \vec{E}_{\mathrm{T}}^{\mathrm{miss}})_{\mathrm{min}}$ and $E_{\mathrm{T}}^{\mathrm{miss}}/m_{\mathrm{eff}}(4\mathrm{j})$ cuts, which are effective to suppress the multi-jet process, are inverted at CRQ to enrich the multi-jet process. They are not utilized for the inputs of the BDT analysis in order to be used in the definition of the CRQ.

The event yield in the SR is predicted from the number of observed data in the CR and the acceptance difference between the SR and CR calculated from the MC; that is, the predicted number of events $N_{SR}^{x,\mathrm{pred}}$ in the SR of a background process x is given as

$$N_{SR}^{x,\mathrm{pred}} = N_{\mathrm{CR}x}^{x,\mathrm{data}} \times \frac{N_{\mathrm{SR}}^{x,\mathrm{MC}}}{N_{\mathrm{CR}x}^{x,\mathrm{MC}}}, \tag{8.1}$$

where CRx is a CR dedicated to the background process x, $N_{\mathrm{CR}x}^{x,\mathrm{data}}$ is the number of observed data events in the CRx, and $N_{\mathrm{Y}}^{x,\mathrm{MC}}$ is the number of the MC events for the background process x in a region Y. In another point of view, $N_{\mathrm{CR}x}^{x,\mathrm{data}}/N_{\mathrm{CR}x}^{x,\mathrm{MC}}$ is behaved as a normalization factor ("$\mu(x)$") on each MC in the SR for the background x.

In general, the kinematic phase space of the CR is set to be close to that of the SR under the assumption that the discrepancy between the data and MC is similar between the SR and CR. Especially, the range of m_{eff} in the SR and CR is kept to be similar by requiring the same BDT cut in the CR as in the SR because the m_{eff} distribution has mis-modeling, which is caused by the mis-modeling in the jet activity predicted by the QCD. The comparison of the kinematic phase space between the SR and CR is shown in Section 8.2.5. The validation of this method is performed by preparing Validation Regions ("VR"). The VR is also dominated by a specific background process, but the selection is different from that of the CR. By comparison between the observed data and the background prediction corrected by the CRs, the validation for the background process is established. In the VR, the prediction is obtained by multiplying $N_{\mathrm{CR}x}^{x,\mathrm{data}}/N_{\mathrm{CR}x}^{x,\mathrm{MC}}$ to the pure[1] MC

---

[1] The pure MC prediction means the MC is normalized by cross-section.

prediction $N^{x,\mathrm{MC}}_{\mathrm{VR}x}$. In addition to this validation, various kinds of uncertainties on the acceptance difference ($N^{x,\mathrm{MC}}_{\mathrm{SR}}/N^{x,\mathrm{MC}}_{\mathrm{CR}x}$) obtained from the MCs are taken into account, which will be described in Section 9.3.

The details of the CRs and VRs definitions are listed in Tables 8.2 and 8.3. The estimation and validation of each background process are described in the following sections.

8.2.1 $Z(\to \nu\nu)$ +jets

Z can be tagged with two opposite charged leptons. However, the statistic of the $Z \to ll$ process is too small to estimate the number of Z+jets events in the SR. Thus, the Z+jets is estimated from the γ+jets process in the CRY, which has more statistic. Here, the mass difference between Z and γ can be ignored because the event kinematics in the SR and CRY is much harder than the Z mass.

The CRY events are selected by one photon requirement. The p_{T} of the γ is required to be larger than 150 GeV because the CRY uses a single photon trigger whose efficiency reaches 100% around $p^{\gamma}_{\mathrm{T}} \sim 140$ GeV [61]. In the CRY, to obtain the similar kinematic phase space of $Z(\to \nu\nu)$ in the SR, the γ is treated as $E^{\mathrm{miss}}_{\mathrm{T}}$; that is, the momentum vector in the transverse plane of the γ is added to the $E^{\mathrm{miss}}_{\mathrm{T}}$ as $\vec{E}^{\mathrm{miss}\,\prime}_{\mathrm{T}} = \vec{E}^{\mathrm{miss}}_{\mathrm{T}} + \vec{p}^{\gamma}_{\mathrm{T}}$. The $E^{\mathrm{miss}}_{\mathrm{T}} > 300$ GeV cut at the preselection is applied on the $E^{\mathrm{miss}\,\prime}_{\mathrm{T}}$ instead of the original $E^{\mathrm{miss}}_{\mathrm{T}}$. Thus, this region has only small real $E^{\mathrm{miss}}_{\mathrm{T}}$ if p^{γ}_{T} is large, which

Table 8.2: Control Region (CR) definitions for each background process.

Cut	Control Region			
	CRY	CRW	CRT	CRQ
Trigger	Single photon trigger	OR of single lepton triggers		As for SR cut ($E^{\mathrm{miss}}_{\mathrm{T}}$ trigger)
Particles	≥ 1 signal photon with $p_{\mathrm{T}} > 150$ GeV No e/μ	Exactly 1 electron or muon with $p_{\mathrm{T}}(e) > 27$ GeV or $p_{\mathrm{T}}(\mu) > 27$ GeV		As for SR cut (No e/μ)
b-jet (with $p_{\mathrm{T}} > 50$ GeV and $\lvert\eta\rvert < 2.5$)	–	No b-jet	$\geq 1b$-jet	–
m_{T} cut for W mass window	–	30 GeV $< m_{\mathrm{T}}(\ell, E^{\mathrm{miss}}_{\mathrm{T}}) <$ 100 GeV		–
QCD cut	–	$E^{\mathrm{miss}}_{\mathrm{T}} > 200$ GeV		–
Particle treatment	Treat a photon as invisible	Treat a lepton as invisible		–
Use below:	$E^{\mathrm{miss}\,\prime}_{\mathrm{T}} = \lvert\vec{E}^{\mathrm{miss}}_{\mathrm{T}} + \vec{p}^{\gamma}_{\mathrm{T}}\rvert$	$E^{\mathrm{miss}\,\prime}_{\mathrm{T}} = \lvert\vec{E}^{\mathrm{miss}}_{\mathrm{T}} + \vec{p}^{\mathrm{lepton}}_{\mathrm{T}}\rvert$		–
BDT score	As for SR cut			
$\Delta\phi(\mathrm{jet}_{1,2,(3)}, \vec{E}^{\mathrm{miss}}_{\mathrm{T}})_{\mathrm{min}}$, $\Delta\phi(\mathrm{jet}_{i>3}, \vec{E}^{\mathrm{miss}}_{\mathrm{T}})_{\mathrm{min}}$	No cut	As for SR cut		$\Delta\phi(\mathrm{jet}_{1,2,(3)}, \vec{E}^{\mathrm{miss}}_{\mathrm{T}})_{\mathrm{min}} < X$ or X : $\Delta\phi(\mathrm{jet}_{1,2,(3)}, \vec{E}^{\mathrm{miss}}_{\mathrm{T}})_{\mathrm{min}}$ cut in SR
$E^{\mathrm{miss}}_{\mathrm{T}}/m_{\mathrm{eff}}(4\mathrm{j})$	As for SR cut			$0.14 < E^{\mathrm{miss}}_{\mathrm{T}}/m_{\mathrm{eff}}(4\mathrm{j}) < 0.20$ if the SR has $E^{\mathrm{miss}}_{\mathrm{T}}/m_{\mathrm{eff}}(4\mathrm{j}) > 0.20$
$E^{\mathrm{miss}}_{\mathrm{T}}$	As for SR cut			
m_{eff}(incl.)				
$p_{\mathrm{T}}(j_{1,2,3,4})$				
Number of jets				
m_{eff}				
$E^{\mathrm{miss}}_{\mathrm{T}}/\sqrt{H_{\mathrm{T}}}$				
$\lvert\eta(j_{1,2,3,4})\rvert$				

Table 8.3: Validation Region (VR) definitions for VRY/Z/W/T. The definitions of the VRQs is shown in Figure 8.13.

Cut	Validation Region				
	for Z+jets			for W+jets & Top	
	VRYdPhi	VRZ	VRZdPhi	VRW/TJ	VRW/TJdPhi
Trigger	Single photon trigger	OR of single lepton triggers			
Particles	As for CRY cut	Exact 2 opposite charged electrons or muons: $p_{\mathrm{T}}(e) > 27, 7$ GeV or $p_{\mathrm{T}}(\mu) > 27, 7$ GeV		As for CRW/T cut	
b-jet with $p_{\mathrm{T}} > 50$ GeV and $\lvert\eta\rvert < 2.5$	–			No b-jet for VRWXs / $\geq 1b$-jet for VRTXs	
$m_{\ell\ell}$, m_{T} cut for Z, W mass window	–	66 GeV $< m_{\ell\ell} <$ 116 GeV		30 GeV $< m_{\mathrm{T}}(\ell, E_{\mathrm{T}}^{\mathrm{miss}}) <$ 100 GeV for VRWXs	
Particle treatment	Treat a photon as invisible	Treat leptons as invisible		Treat a lepton as a jet	
Use below:	$E_{\mathrm{T}}^{\mathrm{miss}\,\prime} = \lvert\vec{E}_{\mathrm{T}}^{\mathrm{miss}} + \vec{p}_{\mathrm{T}}^{\,\gamma}\rvert$	$E_{\mathrm{T}}^{\mathrm{miss}\,\prime} = \lvert\vec{E}_{\mathrm{T}}^{\mathrm{miss}} + \vec{p}_{\mathrm{T}}(ll)\rvert$		A lepton with $p_{\mathrm{T}} > 50$ GeV as a jet	
BDT	As for SR cut				
$\Delta\phi(\mathrm{jet}_{1,2,(3)}, \vec{E}_{\mathrm{T}}^{\mathrm{miss}})_{\mathrm{min}}$, $\Delta\phi(\mathrm{jet}_{i>3}, \vec{E}_{\mathrm{T}}^{\mathrm{miss}})_{\mathrm{min}}$	As for SR cut	No cut (As CR cut)	As SR cut	No cut	As SR cut
$E_{\mathrm{T}}^{\mathrm{miss}}$	As for SR cut				
m_{eff}(incl.)					
$p_{\mathrm{T}}(j_{1,2,3,4})$					
Number of jets					
$E_{\mathrm{T}}^{\mathrm{miss}}/\sqrt{H_{\mathrm{T}}}$					
$E_{\mathrm{T}}^{\mathrm{miss}}/m_{\mathrm{eff}}(4\mathrm{j})$					
$\lvert\eta(j_{1,2,3,4})\rvert$					

makes this region orthogonal to the SR. The variables[2] using $\vec{E}_{\mathrm{T}}^{\mathrm{miss}}$ are recalculated by the $\vec{E}_{\mathrm{T}}^{\mathrm{miss}\,\prime}$ instead of the original $\vec{E}_{\mathrm{T}}^{\mathrm{miss}}$. In order to increase the statistic of the CRY, the $\Delta\phi(\mathrm{jet}_{1,2,(3)}, \vec{E}_{\mathrm{T}}^{\mathrm{miss}})_{\mathrm{min}}$ and $\Delta\phi(\mathrm{jet}_{i>3}, \vec{E}_{\mathrm{T}}^{\mathrm{miss}})_{\mathrm{min}}$ cuts, which are applied in the SR, are not required here. These variables are checked at the preselection. The distribution of the $\Delta\phi(\mathrm{jet}_{1,2,(3)}, \vec{E}_{\mathrm{T}}^{\mathrm{miss}})_{\mathrm{min}}$ is shown in Figures 8.8. The shapes of them are well modeled by the MC.

The γ+jets MC is generated by the LO generator but the Z+jets MC is generated at the NNLO level. This difference affects the cross-section of the MCs significantly. Thus, in this analysis, the cross-section of the γ+jets is corrected by a factor, called as a "κ-factor ", obtained from a loose CRY ("CRYL") and a loose $Z \to ll$ region ("CRZL"). The κ-factor is defined as

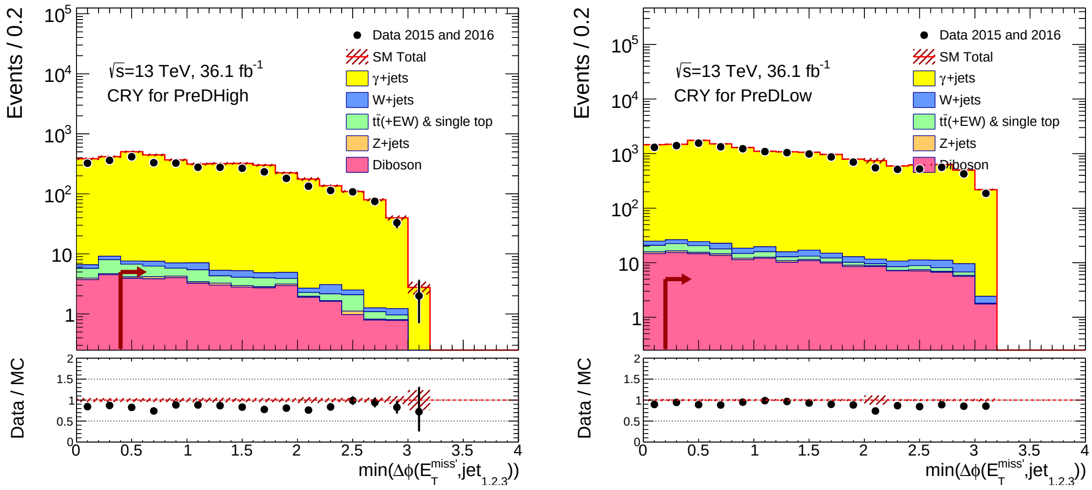
$$\kappa = \frac{N_{\mathrm{CRYL}}^{\gamma+\mathrm{jets,data}} / N_{\mathrm{CRYL}}^{\gamma+\mathrm{jets,MC}}}{N_{\mathrm{CRZL}}^{Zll,\mathrm{data}} / N_{\mathrm{CRZL}}^{Zll,\mathrm{MC}}} \tag{8.2}$$

$$= \frac{(N_{\mathrm{CRYL}}^{\mathrm{data}} - N_{\mathrm{CRYL}}^{\mathrm{Other\ bkg}}) / N_{\mathrm{CRYL}}^{\gamma+\mathrm{jets,MC}}}{(N_{\mathrm{CRZL}}^{\mathrm{data}} - N_{\mathrm{CRZL}}^{\mathrm{Other\ bkg}}) / N_{\mathrm{CRZL}}^{Zll,\mathrm{MC}}}, \tag{8.3}$$

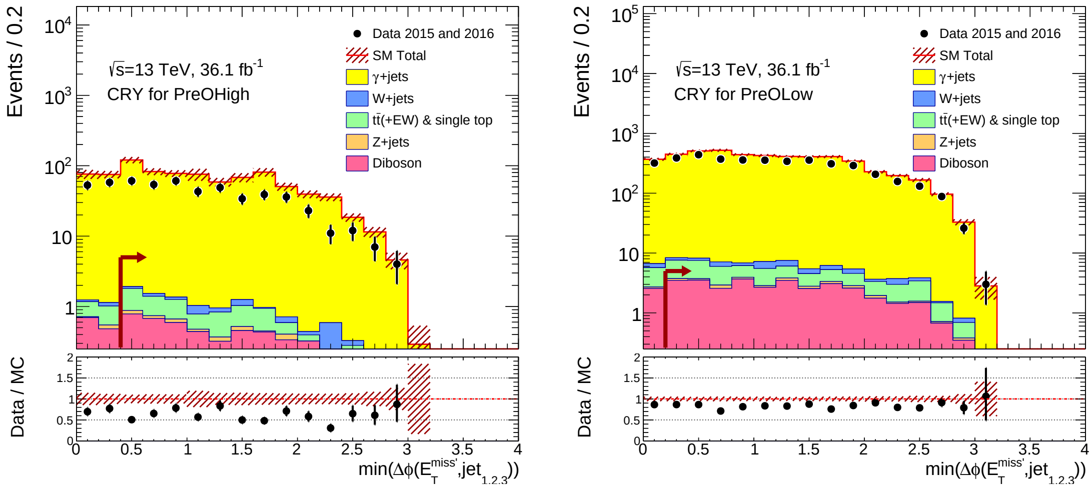
where $N_{\mathrm{CRYL}}^{\gamma+\mathrm{jets,data/MC}}$ is the number of γ+jets events in the CRYL of the data/MC and $N_{\mathrm{CRZL}}^{Zll,\mathrm{data/MC}}$ is

---

[2] m_{eff}, $E_{\mathrm{T}}^{\mathrm{miss}}/\sqrt{H_{\mathrm{T}}}$, $E_{\mathrm{T}}^{\mathrm{miss}}/m_{\mathrm{eff}}(4\mathrm{j})$, $\Delta\phi(\mathrm{jet}_{1,2,(3)}, \vec{E}_{\mathrm{T}}^{\mathrm{miss}})_{\mathrm{min}}$, and $\Delta\phi(\mathrm{jet}_{i>3}, \vec{E}_{\mathrm{T}}^{\mathrm{miss}})_{\mathrm{min}}$.



(a) CRY for PreDHigh (b) CRY for PreDLow



(c) CRY for PreOHigh (d) CRY for PreOLow

Figure 8.8: Distributions of $\Delta\phi(\mathrm{j}_{1,2,(3)}, \vec{E}_{\mathrm{T}}^{\mathrm{miss}})_{\mathrm{min}}$ in CRYs of (a) **PreDHigh**, (b) **PreDLow**, (c) **PreOHigh**, and (d) **PreOLow**. Black point shows 36.1fb^{-1}data and filled histograms show the SM background events, which are normalized by cross-section except the γ +jets process. The normalization of the γ +jets is multiplied by κ-factor in addition. The red arrow shows the preselection but it is not used in CRY.

the number of $Z \rightarrow ll$ events in the CRZL of the data/MC. The CRZL is defined to enrich Z decaying to two opposite charged leptons by requiring the Z mass range of the invariant mass calculated from the two leptons. The $E_{\mathrm{T}}^{\mathrm{miss}}$ and m_{eff} cuts for the CRZL are loosened to recover the statistic of the $Z \rightarrow ll$ events and the BDT cut is not required. On the other variables, SR-like cuts are applied. The cuts for the CRYL are also loosened as in CRZL. The κ-factor is calculated for each jet multiplicity of $N_{\mathrm{jet}} \geq 2, 4, 5,$ and 6 corresponding to the requirements in the ten SRs. The definition of the CRYL and CRZL are summarized in Table 8.4.

Table 8.4: CRYL and CRZL definitions. They are used to compute the κ-factor.

Cut	Control Region	
	CRYL	CRZL
Trigger	Single photon trigger	OR of single lepton triggers
Particles	≥ 1 signal photon with $p_{\mathrm{T}} > 150$ GeV No e/μ	Exact 2 opposite charged electrons or muons: $p_{\mathrm{T}}(e) > 27, 7$ GeV or $p_{\mathrm{T}}(\mu) > 27, 7$ GeV
$m_{\ell\ell}$ cut	–	66 GeV $< m_{\ell\ell} <$ 116 GeV
Particle treatment	Treat photon as invisible	Treat leptons as invisible
Use below:	$E_{\mathrm{T}}^{\mathrm{miss}\,\prime} = \lvert\vec{E}_{\mathrm{T}}^{\mathrm{miss}} + \vec{p}_{\mathrm{T}}^{\gamma}\rvert$	$E_{\mathrm{T}}^{\mathrm{miss}\,\prime} = \lvert\vec{E}_{\mathrm{T}}^{\mathrm{miss}} + \vec{p}_{\mathrm{T}}(ll)\rvert$
BDT score	–	
$E_{\mathrm{T}}^{\mathrm{miss}}$	>250	
$\Delta\phi(\mathrm{jet}_{1,2,(3)}, \vec{E}_{\mathrm{T}}^{\mathrm{miss}})_{\mathrm{min}}$,	>0.4	
m_{eff}(incl.)	>1200	
$p_{\mathrm{T}}(j_1)$	>200	
Number of jets	$\geq 2, 4, 5, 6$	
$E_{\mathrm{T}}^{\mathrm{miss}}/\sqrt{H_{\mathrm{T}}}$	>14	

The κ-factor is applied in Eq. 8.1 to correct the $Z(\rightarrow \nu\nu)$ estimation as

$$N_{SR}^{Z\nu\nu,\mathrm{pred}} = N_{\mathrm{CRY}}^{\gamma\text{+jets},\mathrm{data}} \times \frac{N_{\mathrm{SR}}^{Z\nu\nu,\mathrm{MC}}}{N_{\mathrm{CRY}}^{\gamma\text{+jets},\mathrm{MC}} \times \kappa}. \tag{8.4}$$

In Eq. 8.3, to retrieve the number of γ+jets or $Z \rightarrow ll$ events in the data, the simultaneous fit including CRT, CRW, and CRQ is performed to estimate the other SM background $\left(N_{\mathrm{CRYL/CRZL}}^{\mathrm{Other\ bkg}}\right)$, in which the kinematic cuts of $E_{\mathrm{T}}^{\mathrm{miss}}$, $\Delta\phi(\mathrm{jet}_{1,2,(3)}, \vec{E}_{\mathrm{T}}^{\mathrm{miss}})_{\mathrm{min}}$, m_{eff}, jet p_{T}, N_{jet}, and $E_{\mathrm{T}}^{\mathrm{miss}}/\sqrt{H_{\mathrm{T}}}$ in the CRs are the same as in the CRYL and CRZL. The fit procedure is the same as in the "background-only fit" described in Section 9. The obtained κ-factor for each jet multiplicity is listed in Table 8.5.

The CRYL and CRZL are much lower m_{eff} range than the CRY. However, the κ-factor obtained in the CRYL and CRZL is also valid in the high m_{eff} range after the BDT cut as CRY since there

Table 8.5: The κ-factors for each jet multiplicity obtained from the CRYL and CRZL in the background-only fit including CRT, CRW, and CRQ.

Jet multiplicity	κ-factor
$N_{\text{jet}} \geq 2$	1.560 ± 0.035
$N_{\text{jet}} \geq 4$	1.668 ± 0.068
$N_{\text{jet}} \geq 5$	1.680 ± 0.116
$N_{\text{jet}} \geq 6$	2.257 ± 0.261

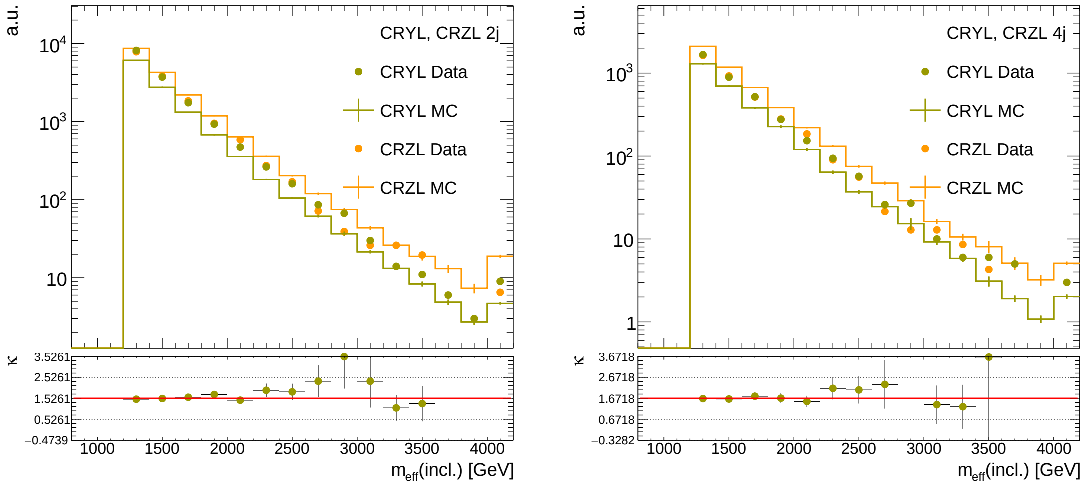
is no significant dependence on the m_{eff} above statistical fluctuation as shown in Figure 8.9 [3]. The validation of the κ-factor after the BDT cut is also performed in the VRs.

The VR for the estimation of the $Z(\rightarrow \nu\nu)$ process is prepared in the two lepton region (VRZ). It has a similar selection to that of CRZL, but the kinematic requirements and BDT cut are the same as in the SR. Thus, it can be confirmed that this estimation using κ-factor is valid also in the harder kinematic phase space. There are three VRs for the CRY and CRZ. The VRZ and VRZdPhi (VRZdPhi has $\Delta\phi$ cuts[4] as in the SR) are $Z \rightarrow ll$ region to see the estimated Z events. The VRYdPhi is a one-photon region with the $\Delta\phi$ cuts as in the SR. It is prepared to confirm that the loosened $\Delta\phi$ cuts in the CRY does not affect the estimation.

---

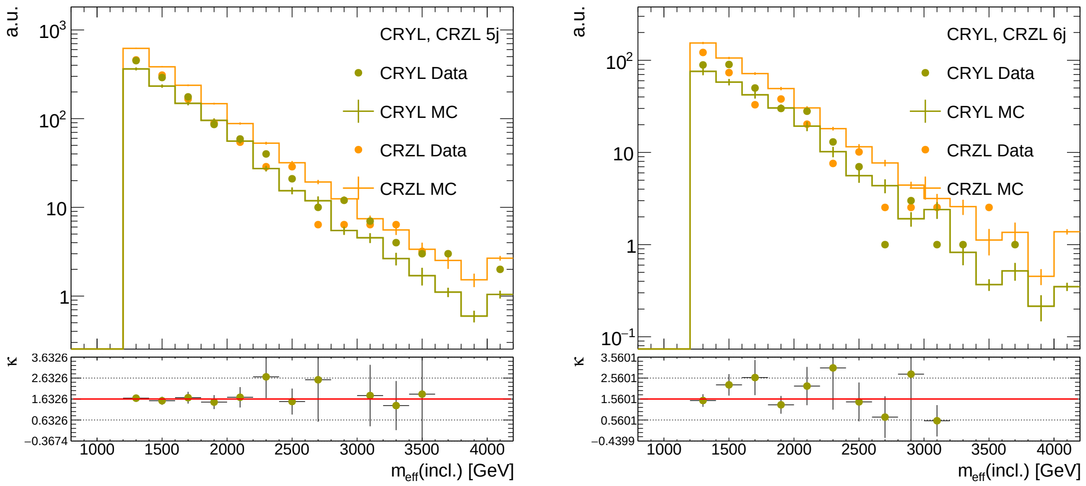
[3] A bit discrepancy around 3000 GeV in $N_{\text{jet}} \geq 2$ does not affect the estimation in the SR with $N_{\text{jet}} \geq 2$ (SRD5) because the m_{eff} range in the SR is much lower than it.

[4] The $\Delta\phi$ cuts mean the $\Delta\phi(\text{jet}_{1,2,(3)}, \vec{E}_{\text{T}}^{\text{miss}})_{\text{min}}$ and $\Delta\phi(\text{jet}_{i>3}, \vec{E}_{\text{T}}^{\text{miss}})_{\text{min}}$ cuts.



(a) $N_{\text{jet}} \geq 2$

(b) $N_{\text{jet}} \geq 4$



(c) $N_{\text{jet}} \geq 5$

(d) $N_{\text{jet}} \geq 6$

Figure 8.9: Distributions of m_{eff} (inc) in CRYL and CRZL with $N_{\text{jet}} \geq$ (a)2, (b)4, (c)5, and (d)6 in the data and total SM background MC. The MC in the CRYL is normalized by cross-section. The data and MC in the CRZL have scaled arbitrarily with the same scale. In the bottom panel, the kappa factor and a fitted constant line are shown.

8.2.2 $W(\rightarrow l\nu)$ +jets / $t\bar{t}$ (and single top)

There are two categories of no lepton (including tau here) events in W+jets and top ($t\bar{t}$ and single top) events in the SR. The one is the missing electron or muon due to the out-of-acceptance or the inefficiency in its identification. The out-of-acceptance is mainly occurred by the low lepton p_{T} less than 7 GeV (lepton p_{T} threshold). The other is the hadronic decaying tau. This tau is reconstructed as a jet if its p_{T} is larger than 50 GeV. The fraction of the missing electron or muon is approximately 40% and the hadronic tau is 60% at the preselection (Tables 8.6–8.9).

The CRW and CRT have a requirement of one lepton. There are two ways to treat the lepton in order to obtain similar kinematics to that in the SR. The one is to treat the lepton as an invisible particle like a missing lepton, and the other is to treat it as a jet like a hadronic tau. In the CRW/T, the lepton is treated as an invisible particle, in which the momentum vector in the transverse plane of the lepton is added to the original $E_{\mathrm{T}}^{\mathrm{miss}}$ as $E_{\mathrm{T}}^{\mathrm{miss}}{}' = |\vec{E}_{\mathrm{T}}^{\mathrm{miss}} + \vec{p}_{\mathrm{T}}^{\mathrm{lepton}}|$. In contrast, in the VRs, the lepton is treated as a jet and its four-momentum is taken into account in the computations of all of the kinematic variables as a jet.

Table 8.6: Fraction of leptons in W+jets and $t\bar{t}$ events at the preselection **PreDHigh**.

Process	Fraction (%)
W with 1 lepton	20.4
$t\bar{t}$ with 1 lepton	14.3
No lepton	0.0
2 leptons	1.7
Hadronic tau	63.6

Table 8.7: Fraction of leptons in W+jets and $t\bar{t}$ events at the preselection **PreDLow**.

Process	Fraction (%)
W with 1 lepton	23.4
$t\bar{t}$ with 1 lepton	15.3
No lepton	0.0
2 leptons	1.9
Hadronic tau	59.5

Table 8.8: Fraction of leptons in W+jets and $t\bar{t}$ events at the preselection **PreOHigh**.

Process	Fraction (%)
W with 1 lepton	13.6
$t\bar{t}$ with 1 lepton	19.1
No lepton	0.0
2 leptons	1.7
Hadronic tau	65.6

Table 8.9: Fraction of leptons in W+jets and $t\bar{t}$ events at the preselection **PreOLow**.

Process	Fraction (%)
W with 1 lepton	12.5
$t\bar{t}$ with 1 lepton	21.6
No lepton	0.0
2 leptons	1.9
Hadronic tau	64.2

The CRs and VRs for the W+jets and top process are defined to enrich the W boson. In the top process, a W produced in the decay of the top quark is tagged. In addition to one lepton (electron or muon) requirement, there is also m_{T} cut to enrich the W. The m_{T} is defined as

$$m_{\mathrm{T}} = \sqrt{2 p_{\mathrm{T}}^{\ell} E_{\mathrm{T}}^{\mathrm{miss}} (1 - \cos[\Delta\phi(\vec{p}_{\mathrm{T}}^{\ell}, \vec{E}_{\mathrm{T}}^{\mathrm{miss}})])}, \tag{8.5}$$

where p_{T}^{ℓ} is a lepton p_{T} and $\Delta\phi(\vec{p}_{\mathrm{T}}^{\ell}, \vec{E}_{\mathrm{T}}^{\mathrm{miss}})$ is the ϕ-distance between the lepton and the $\vec{E}_{\mathrm{T}}^{\mathrm{miss}}$. The distribution of this variable has an endpoint in the mass of the parent particle, which is assumed to be the W in this case. Thus, the m_{T} cut below the W mass can enrich the W boson. Its distribution in the CRW at the preselection PreOLow is shown in Figure 8.10. A discrepancy in the height between the data and MC can be seen, but it will be corrected by the normalization factor in Eq. 8.1 in the background estimation.

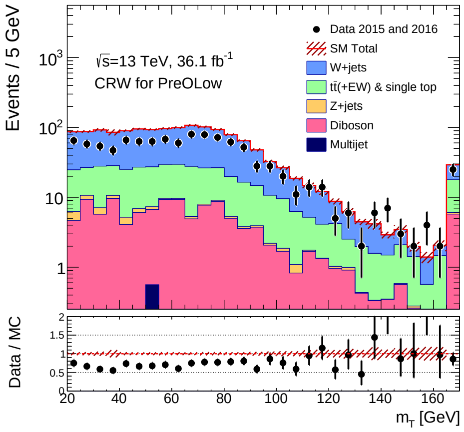
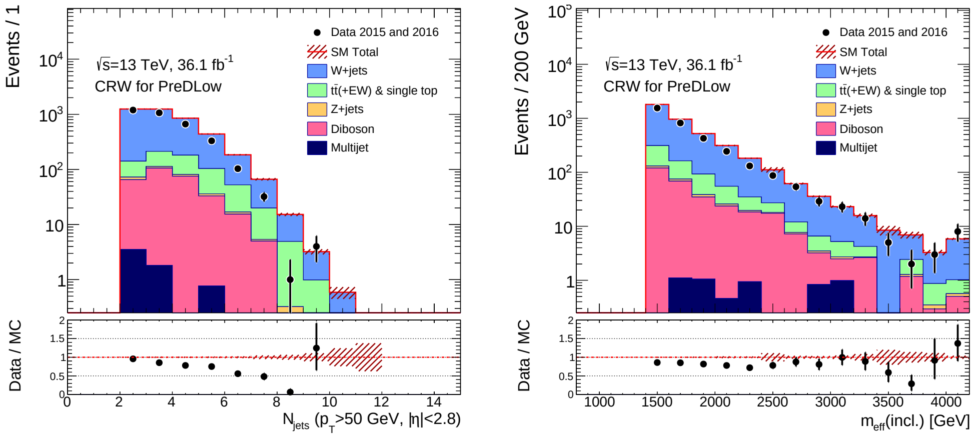


Figure 8.10: Distribution of m_{T} in CRW at **PreOLow** in the data and total SM background MC. The MC is normalized by cross-section.

For the VRs, two selections with and without $\Delta\phi(\mathrm{jet}_{1,2,(3)}, \vec{E}_{\mathrm{T}}^{\mathrm{miss}})_{\mathrm{min}}$ and $\Delta\phi(\mathrm{jet}_{i>3}, \vec{E}_{\mathrm{T}}^{\mathrm{miss}})_{\mathrm{min}}$ cut as in the SR are prepared. The VR for W/Top with the $\Delta\phi$ cuts as for SRs is referred to as VRW/TJdPhi. The VR for W/Top without the $\Delta\phi$ cuts is referred to as VRW/TJ, which is to validate the estimation with more statistic.

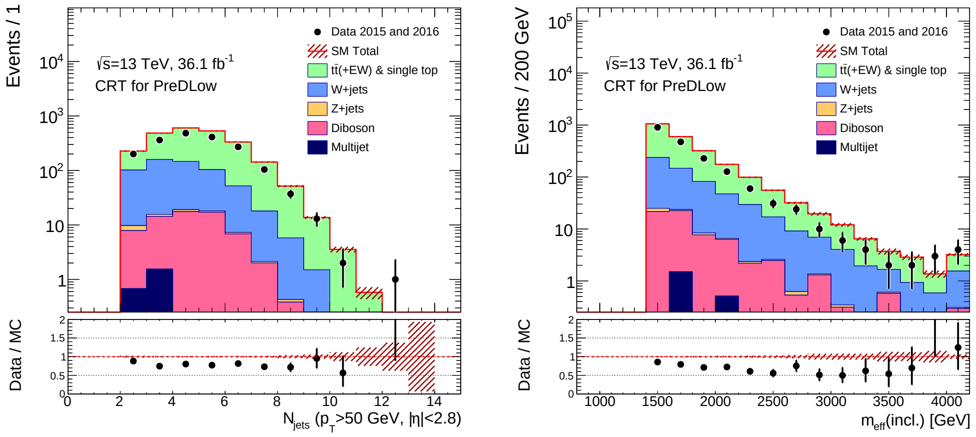
In the CRW and the CRT, large mis-modeling in N_{jet} and m_{eff} distributions are seen as shown in Figures 8.11 and 8.12. Due to the mis-modeling in the shapes, the normalization obtained by cross-section is less than the observed data in the CRW/Ts and VRW/Ts. To avoid the mis-modeling, the CRs (and also VRs) have the same cuts on N_{jet}, m_{eff}, and BDT score as for SRs.



(a) N_{jets}

(b) m_{eff}(inc)

Figure 8.11: Distributions of N_{jets} and m_{eff}(inc) in CRW of **PreDLow**. Black points show the $36.1\mathrm{fb}^{-1}$data and filled histograms show the SM background events, which is normalized by cross-section.



(a) N_{jets}

(b) $m_{\text{eff}}(\text{inc})$

Figure 8.12: Distributions of N_{jets} and $m_{\text{eff}}(\text{inc})$ in CRT of **PreDLow**. Black points show the 36.1fb^{-1}data and filled histograms show the SM background events, which is normalized by cross-section.

8.2.3 Multi-jet

The multi-jet process is significantly removed by the $E_{\text{T}}^{\text{miss}}$, $E_{\text{T}}^{\text{miss}}/\sqrt{H_{\text{T}}}$ or $E_{\text{T}}^{\text{miss}}/m_{\text{eff}}(4\text{j})$, $\Delta\phi(\text{jet}_{1,2,(3)}, \vec{E}_{\text{T}}^{\text{miss}})_{\text{min}}$, and BDT cuts. In this analysis, the **SRD1–4** and **SRO1–2** have $\Delta\phi(\text{jet}_{1,2,(3)}, \vec{E}_{\text{T}}^{\text{miss}})_{\text{min}}$ cut of 0.4 or 0.6 and $E_{\text{T}}^{\text{miss}}/m_{\text{eff}}(4\text{j})$ cut of 0.2 (required at the preselection). The inverted cuts of them are used to enrich the multi-jet events. The CR for the multi-jet is defined by the inverted $\Delta\phi(\text{jet}_{1,2,(3)}, \vec{E}_{\text{T}}^{\text{miss}})_{\text{min}}$ cut and the intermediate $E_{\text{T}}^{\text{miss}}/m_{\text{eff}}(4\text{j})$ cut. The VRs for multi-jet ("VRQdPhi" and "VRQm") have only one of the two cuts as described in Figure 8.13.

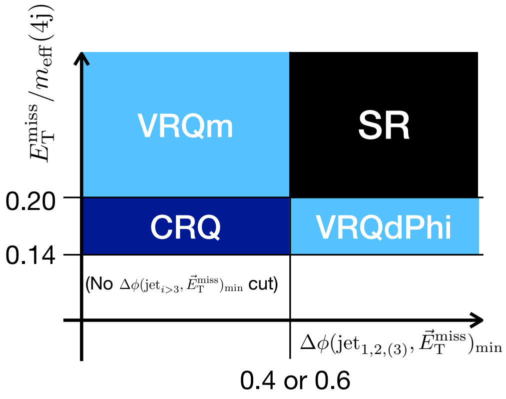


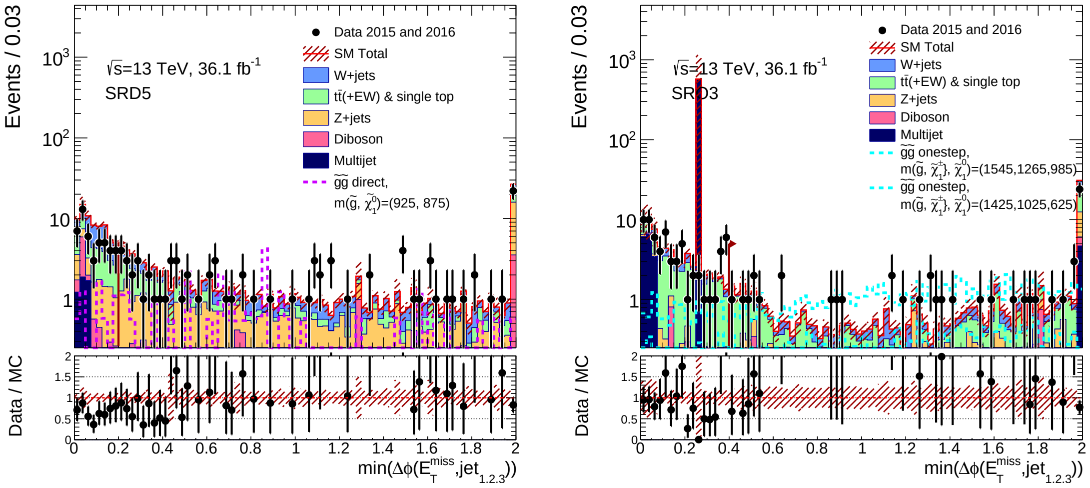
Figure 8.13: A schematic diagram of the definitions of the CRQ and VRQs. The CRQ have the inverted $\Delta\phi(\text{jet}_{1,2,(3)}, \vec{E}_{\text{T}}^{\text{miss}})_{\text{min}}$ cut (0.4 or 0.6) and $0.14 < E_{\text{T}}^{\text{miss}}/m_{\text{eff}}(4\text{j}) < 0.20$ cut. The VRQdPhi has the same $\Delta\phi(\text{jet}_{1,2,(3)}, \vec{E}_{\text{T}}^{\text{miss}})_{\text{min}}$ cut as for the SR. The VRQm has the same $E_{\text{T}}^{\text{miss}}/m_{\text{eff}}(4\text{j})$ cut as for the SR. If the inverted $\Delta\phi(\text{jet}_{1,2,(3)}, \vec{E}_{\text{T}}^{\text{miss}})_{\text{min}}$ cut is applied, the $\Delta\phi(\text{jet}_{i>3}, \vec{E}_{\text{T}}^{\text{miss}})_{\text{min}}$ cut is not applied.

For the other SRs of **SRD5** and **SRO3–5**, the CRQ is not prepared because the multi-jet process is negligible in the SR predicted by the MC as shown in Figure 8.14, and the CRQ has also only small multi-jet events as shown in Figure 8.15. Although the CRQ is not prepared for these SRs, the multi-jet MC normalized by cross-section is included in the estimation.

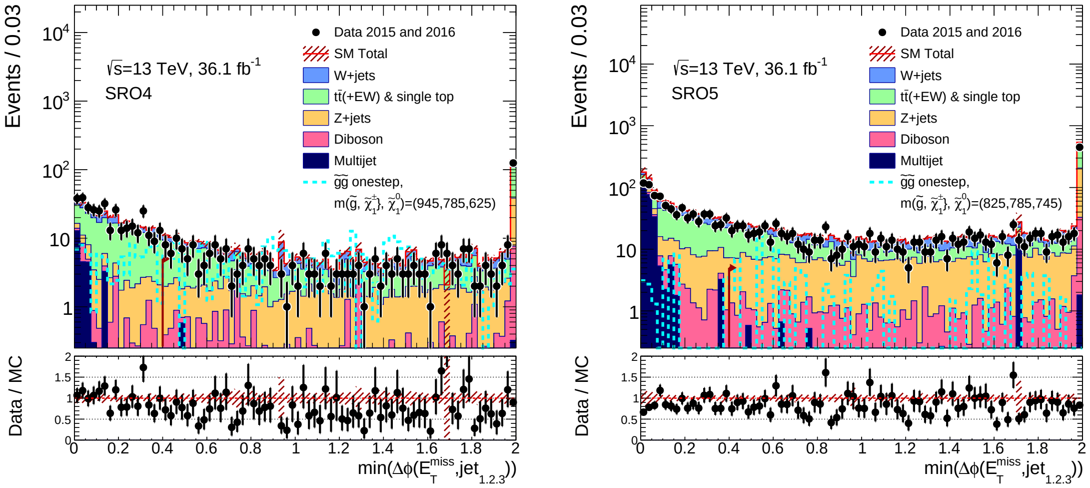
A difference in the kinematic phase space between the CRQ and SR can be seen in Section 8.2.5. This is caused by the different $\Delta\phi(\mathrm{jet}_{1,2,(3)}, \vec{E}_{\mathrm{T}}^{\mathrm{miss}})_{\mathrm{min}}$ and $E_{\mathrm{T}}^{\mathrm{miss}}/m_{\mathrm{eff}}(4\mathrm{j})$ cuts. The difference does not influence the background estimation of the multi-jet process because its contribution to the SR is very small. However, to be conservative, a flat 100% systematic uncertainty is added on the multi-jet event yield as described in Section 9.3.3.

8.2.4 Diboson

The diboson process has only small contribution to the SR and the enriched region is unable to be prepared. In this analysis, the diboson events in the SR are estimated by the MCs normalized by the cross-section. In order to cover possible variations, a flat systematic uncertainty of 30% is added to the diboson MCs as described in Section 9.3.3.

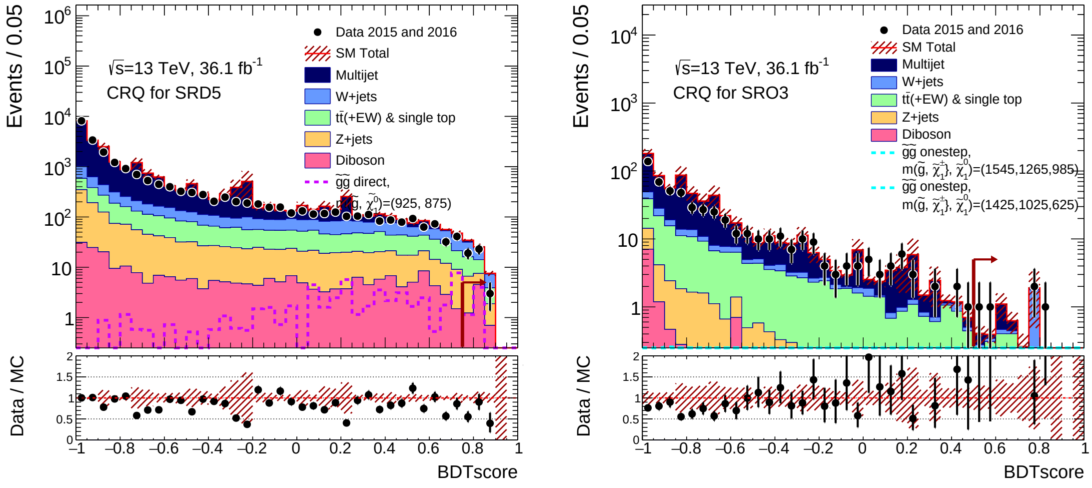


(a) SRD5 (b) SRO3

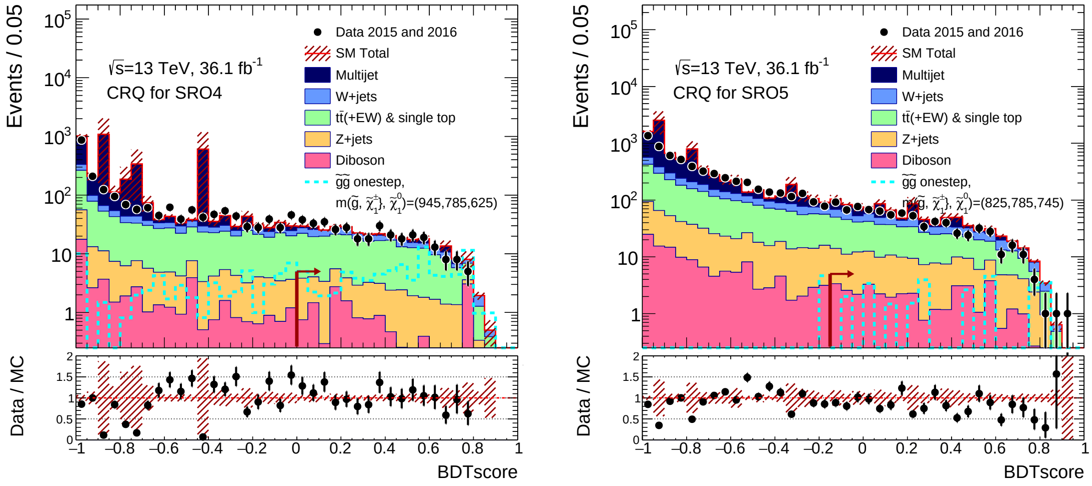


(c) SRO4 (d) SRO5

Figure 8.14: Distributions of $\Delta\phi(\mathrm{j}_{1,2,(3)}, \vec{E}_{\mathrm{T}}^{\mathrm{miss}})_{\mathrm{min}}$ in SRs of (a) **SRD5**, (b) **SRO3**, (c) **SRO4**, and (d) **SRO5**. Black point shows $36.1\mathrm{fb}^{-1}$data and filled histograms show the SM background MCs normalized by cross-section. A large bin of the multi-jet consists of only one event with an extreme high event weight reflecting the skimming weight in the MC event generation. This is unphysical value to be ignored here. The red arrow shows the cut of this variable.



(a) CRQ for SRD5 (b) CRQ for SRO3

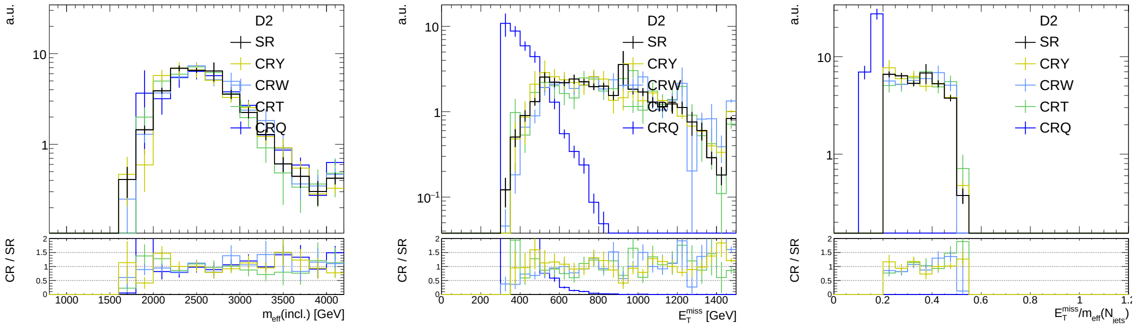


(c) CRQ for SRO4 (d) CRQ for SRO5

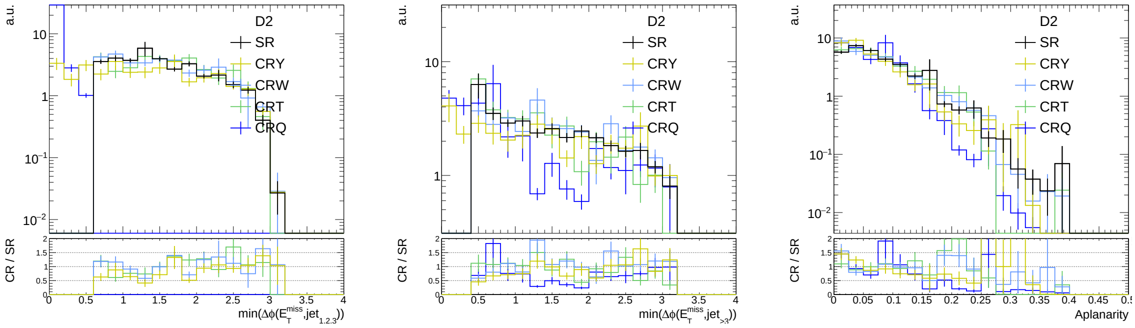
Figure 8.15: Distributions of BDT score in CRQs of (a) **SRD5**, (b) **SRO3**, (c) **SRO4**, and (d) **SRO5**. Black point shows 36.1fb^{-1}data and filled histograms show the SM background MCs normalized by cross-section. A large bin of the multi-jet consists of only one event with an extreme high event weight reflecting the skimming weight in the MC event generation. This is unphysical value to be ignored here. The red arrow shows the cut of this variable.

8.2.5 Kinematic comparison between SR and CR

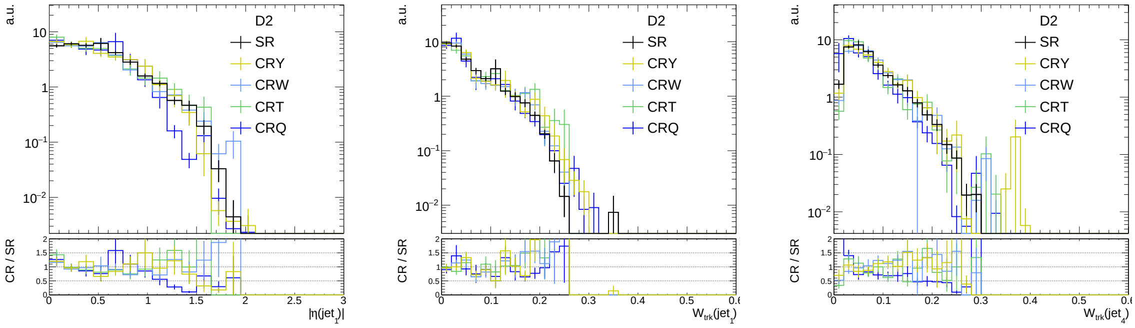
This section shows comparisons of the kinematic phase space between the SR and the CRs (Figures 8.16 shows the SRD2 and the corresponding CRs). The CRQ has different phase space due to the different $\Delta\phi(\text{jet}_{1,2,(3)}, \vec{E}_{\mathrm{T}}^{\mathrm{miss}})_{\mathrm{min}}$ or $E_{\mathrm{T}}^{\mathrm{miss}}/m_{\mathrm{eff}}(4\mathrm{j})$ cut, but the kinematics in the other CRs are seen to be very similar to that in the SR. This fact confirms that the MC simulations in the CRs have similar mis-modeling as in the SR and the normalization factor μs obtained in the CRs can correct the mis-modeling in the SR.



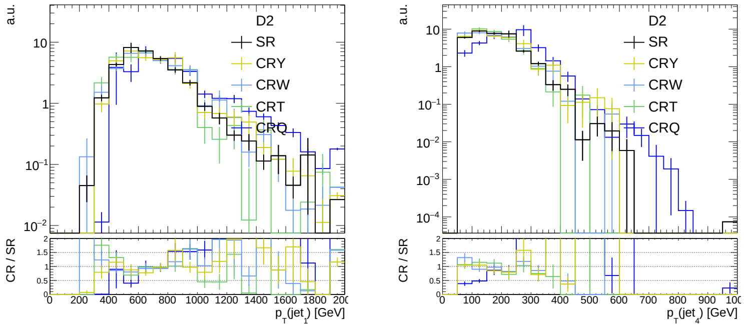
(a) $m_{\text{eff}}(\text{inc})$ (b) $E_{\mathrm{T}}^{\text{miss}}$ (c) $E_{\mathrm{T}}^{\text{miss}}/m_{\text{eff}}(4\text{j})$



(d) $\Delta\phi(\text{j}_{1,2,(3)}, \vec{E}_{\mathrm{T}}^{\text{miss}})_{\text{min}}$ (e) $\Delta\phi(\text{j}_{i>3}, \vec{E}_{\mathrm{T}}^{\text{miss}})_{\text{min}}$ (f) Aplanarity



(g) 1st jet $|\eta|$ (h) 1st jet W_{trk} (i) 4th jet W_{trk}

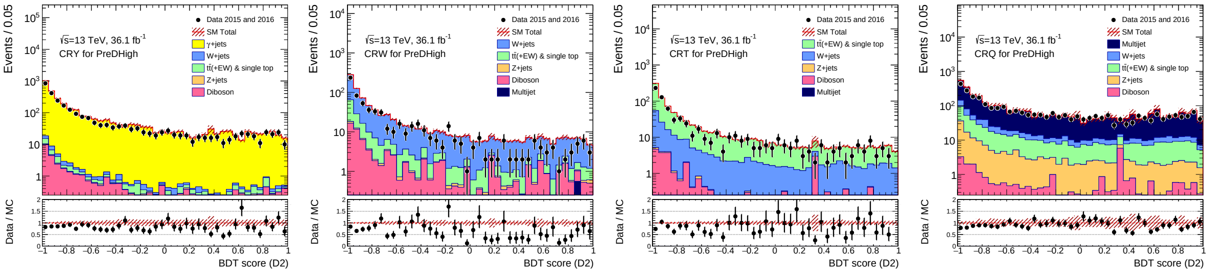


(j) 1st jet p_{T} (k) 4th jet p_{T}

Figure 8.16: Distributions of $m_{\text{eff}}(\text{inc})$, $E_{\mathrm{T}}^{\text{miss}}$, $E_{\mathrm{T}}^{\text{miss}}/m_{\text{eff}}(4\text{j})$, $\Delta\phi(\text{j}_{1,2,(3)}, \vec{E}_{\mathrm{T}}^{\text{miss}})_{\text{min}}$, $\Delta\phi(\text{j}_{i>3}, \vec{E}_{\mathrm{T}}^{\text{miss}})_{\text{min}}$, Aplanarity, 1st jet $|\eta|$, 1st jet W_{trk}, 4th jet W_{trk}, 1st jet p_{T}, and 4th jet p_{T} in the CRY, CRW, CRT, and CRQ of **SRD2** in the MC simulation. The total number of events in each CR is normalized to that in the SR in each plot.

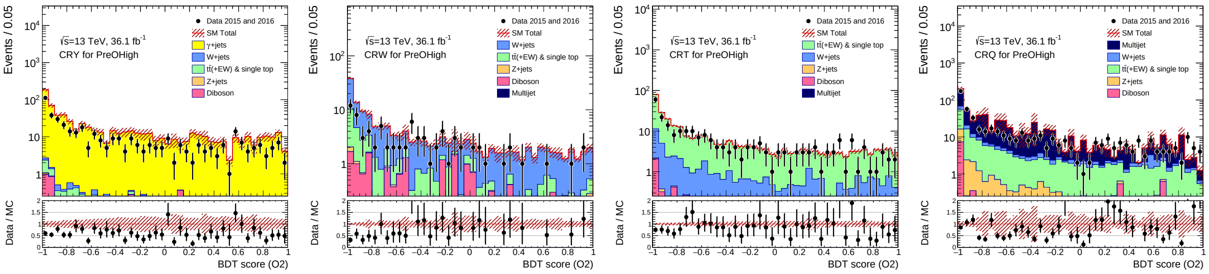
8.3 BDT score distribution in CR

In this section, distributions of the BDT scores before the BDT cut in the CRs are shown. Figures 8.17 and 8.18 show the distributions in SRD2 and SRO2, respectively. The discrepancy between the data and MC in each CR after the BDT cut is taken into account as a normalization factor in Eq. 8.1. The MC in the CRW and CRT has a tendency of underestimation, which is caused by mis-modeling in N_{jet} and m_{eff} as described in Section 8.2.2. However, there is no significant shape mis-modeling in all the CRs, which ensures that the BDT score cut makes no additional discrepancy between the data and MC. The other BDT scores in the CRs are shown in Appendix B.3. There is no significant shape mis-modeling as well.



(a) CRY (b) CRW (c) CRT (d) CRQ

Figure 8.17: Distributions of BDT score (D2) in the CRY, CRW, CRT, and CRQ of the preselection **PreDHigh**. Black point shows 36.1fb^{-1} data and filled histograms show the SM background MC normalized by cross-section except for the γ +jets process in the CRY. The normalization of the γ +jets is multiplied by κ-factor in addition.



(a) CRY (b) CRW (c) CRT (d) CRQ

Figure 8.18: Distributions of BDT score (O2) in the CRY, CRW, CRT, and CRQ of the preselection **PreOHigh**. Black point shows 36.1fb^{-1} data and filled histograms show the SM background MC normalized by cross-section except for the γ +jets process in the CRY. The normalization of the γ +jets is multiplied by κ-factor in addition.

9 Statistical treatment

In order to estimate the signal and background yields in the SR, a profile Log Likelihood Ratio (LLR) approach is used. There are two kinds of fit configurations: "background-only fit" and "exclusion fit" (Table 9.1). The background-only fit is performed to estimate the SM background in the SR by extrapolating the fit result in the CR. For the validation of the background-only fit, the fit result is extrapolated to the VR instead of the SR. In the background-only fit, signal contamination in the CR is assumed to be negligible. The exclusion fit introduces hypothetical signal yields in the SR and CRs, which are numbers of the signal events in the SR and CRs obtained from the MC times a signal strength parameter μ_{sig}. It treats μ_{sig} as a free parameter to be determined in the fit. If there is no significant excess in the background-only fit, the exclusion fit calculates the upper limit on μ_{sig} at each signal mass point.

9.1 The Likelihood function

The likelihood function is defined as

$$L(\boldsymbol{n}|\boldsymbol{\mu}, \boldsymbol{s}, \boldsymbol{b}, \boldsymbol{\theta}) = \mathcal{P}_{\mathrm{SR}} \times \mathcal{P}_{\mathrm{CRY}} \times \mathcal{P}_{\mathrm{CRW}} \times \mathcal{P}_{\mathrm{CRT}} \times \mathcal{P}_{\mathrm{CRQ}} \times C_{\mathrm{Syst}}(\boldsymbol{\theta}^0, \boldsymbol{\theta}), \tag{9.1}$$

where

- $\boldsymbol{n}$ is a set of the observed numbers of events in each region;
- $\boldsymbol{\mu}$ is a set of the normalization factors for each process, i.e. $\mu_j = N^{j,\mathrm{data}}/N^{j,\mathrm{MC}}$ for a given sample j, which is mainly constrained in the dedicated CR;
- $\boldsymbol{s}$ and $\boldsymbol{b}$ are sets of the predicted numbers of the signal and background events in each region, respectively, normalized to the cross-sections and integrated luminosity;
- $\boldsymbol{\theta}$ is a set of nuisance parameters describing systematic uncertainties, and $\theta_i = \pm 1$ corresponds to a source of systematic uncertainties varied by $\pm 1\sigma$ in the fit, where 1σ is the standard deviation;

Table 9.1: Samples and regions used in each fit procedure of the statistical tests to discover or exclude the target signal.

Fit procedure	Background-only fit	Exclusion fit
Samples used	The SM background	The SM background + Signal
Fit regions	CRs	SR + Corresponding CRs

The change of θ_j can change $s(\boldsymbol{\theta})$ and $b(\boldsymbol{\theta})$. In practice, they are continuously interpolated between the nominal value and its variations because the inputs of the fit are only $\pm 1\sigma$ variations of $s(\boldsymbol{\theta})$ and $b(\boldsymbol{\theta})$. The $\mathcal{P}_\mathrm{X}$ is a Poisson term describing the probability to observe n_X events when a number of events λ_X is expected in a given region X. It is given as

$$\mathcal{P}_\mathrm{X} = P\left(n_\mathrm{X} | \lambda_\mathrm{X}(\boldsymbol{\mu}, s_\mathrm{X}, \boldsymbol{b}_\mathrm{X}, \boldsymbol{\theta})\right) \tag{9.2}$$

$$\lambda_\mathrm{X}(\boldsymbol{\mu}, s_\mathrm{X}, \boldsymbol{b}_\mathrm{X}, \boldsymbol{\theta}) = s_\mathrm{X}(\theta) \cdot \mu_\mathrm{sig} + \sum_{j}^{\mathrm{W,Z,Top,Multijet}} b_{\mathrm{X},j}(\boldsymbol{\theta}) \cdot \mu_j + b_\mathrm{X,Diboson}(\boldsymbol{\theta}). \tag{9.3}$$

"$b_\mathrm{X,Diboson}$" is used without a normalization factor since there is no CR dedicated to the di-boson process. In addition to it, the multi-jet process in the **SRD5** and **SRO3–5** has no normalization factor as well (only $b_\mathrm{X,Multijet}$). A probability of each systematic uncertainty becoming a specific value θ_j is included using the probability density function $C_\mathrm{Syst}(\boldsymbol{\theta})$. It is a product of the Gaussian distributions G for each systematic uncertainties, which is given as

$$C_\mathrm{Syst}(\boldsymbol{\theta}) = \prod_{j \in \mathrm{Syst}} G(\theta_j). \tag{9.4}$$

where "Syst" is a set of the systematic uncertainties and the standard deviation of the Gaussian is set to 1 because $\theta = \pm 1$ corresponds to $\pm 1\sigma$.

9.2 Fit procedure

9.2.1 Background-only fit

For the background-only fit, the $\mathcal{P}_\mathrm{SR}$ term in Eq. 9.1 is removed and the signal strength μ_sig is fixed to zero. The maximum likelihood fit is performed to the number of the observed data in CRs and the best-fit background normalization factors $\boldsymbol{\mu}$ are extrapolated to the SR and VRs.

9.2.2 Exclusion fit

In the exclusion fit, first, the maximum likelihood fit is performed to the observed data in the SR and CRs to determine the best-fit μ_sig, $\boldsymbol{\mu}$, and $\boldsymbol{\theta}$, hereafter denoted as $\hat{\mu}_\mathrm{sig}$, $\hat{\boldsymbol{\mu}}$, and $\hat{\boldsymbol{\theta}}$, respectively. Then, a p-value for a hypothetical signal strength μ_sig is obtained. A test statistic q_{μ_sig} is defined as:

$$q_{\mu_\mathrm{sig}} = -2 \log \left(\frac{L(\mu_\mathrm{sig}, \hat{\hat{\boldsymbol{\theta}}})}{L(\hat{\mu}_\mathrm{sig}, \hat{\boldsymbol{\theta}})} \right), \tag{9.5}$$

where $\hat{\hat{\boldsymbol{\theta}}}$ is a set of nuisance parameters to maximize the likelihood function at the tested μ_sig value. The probability density function (PDF) $f(q_{\mu_\mathrm{sig}} | \mu_\mathrm{sig})$ is obtained by pseudo-experiments. It is known to be approximated by an asymptotic formula [62], which is used in the exclusion fit. The p-values in signal plus background and background only hypotheses are calculated by integrating the PDF from a test statistic $q_{\mu_{\text{obs}}}$, which is obtained from the actual data, to infinite:

$$p_{s+b} = \int_{q_{\mu_{\text{obs}}}}^{\infty} f(q_{\mu_{\text{sig}}}|\mu_{\text{sig}} = 1)\, \mathrm{d}q_{\mu_{\text{sig}}} \tag{9.6}$$

$$p_b = \int_{q_{\mu_{\text{obs}}}}^{\infty} f(q_{\mu_{\text{sig}}}|\mu_{\text{sig}} = 0)\, \mathrm{d}q_{\mu_{\text{sig}}}. \tag{9.7}$$

To calculate the significance of the observed data in the SR on the given signal, CL_s [63] is used, which is defined as

$$CL_s \equiv \frac{p_{s+b}}{1 - p_b}. \tag{9.8}$$

The penalty factor of $1/(1 - p_b)$ is introduced to avoid downward fluctuations of the SM background that would give unexpectedly strong exclusion power. The upper limit on the signal mass is set at the mass with 95% confidence level on $\mu_{\text{sig}} = 1$, where $CL_s = 5\%$. The expected upper limit, in which the data in the SR is unknown, is obtained from the p-values calculated by integrating the PDF from the median of the background-only PDF $f(q_{\mu_{\text{sig}}}|\mu_{\text{sig}} = 0)$ instead of $q_{\mu_{\text{obs}}}$ in Eqs. 9.6 and 9.7.

9.3 Systematic uncertainties

For the treatment of the nuisance parameter, there are three ways:

- Fully correlated across the different regions and the different physics processes. (ex. luminosity uncertainty, uncertainty on the jet energy scale)
- Fully correlated across the different regions but independent per process. (ex. theory uncertainties on each MC)
- Fully uncorrelated variables with one parameter per a region. (ex. MC statistical uncertainties)

In the background estimation of Eq. 8.1, the acceptance difference between the SR and CR is corrected by $N^{\text{MC}}_{\text{SR}}/N^{\text{MC}}_{\text{CR}}$. This correction factor is referred to as a "transfer factor" and the systematic uncertainties are estimated on this transfer factor. Thus, in the case of a fully correlated systematic uncertainty between the SR and CR, the systematic variation could be canceled.

All the systematic uncertainties are summarized in Table 9.2. Each of them is described in the following sections.

9.3.1 General uncertainties

Luminosity $36.1 \pm 1.1\,\text{fb}^{-1}$ (3.2%). This is measured by the *vdM* scan and the luminosity detector (see Section 2.2.6).

Table 9.2: Summary of systematic uncertainties included in the likelihood fit.

Name	Name in systematic table	Process	Nuisance parameter Treatment	Comment
General				
Luminosity		All	Fully correlated	±3.2%
Pileup reweighting	pileUp			
MC statistic	stat		Uncorrelated	
CR statistic	mu_X	W/Z/γ+jets/Top	Free parameter	
Object modeling				
Jet energy scale (JES)	JET_GroupedNP_(1-3)	All	Fully correlated	
Jet energy resolution (JER)	JER			
$E_{\mathrm{T}}^{\mathrm{miss}}$ soft term scale	MET_SoftTrk_Scale			
$E_{\mathrm{T}}^{\mathrm{miss}}$ soft term resolution	MET_SoftTrk_ResoPerp/Para			
b-tagging efficiency (b/c/light)	EFF_Y			Only in CRW and CRT
Physics process modeling				
Scale variation	Scale_renom/fact/qsf/ckkw	$W/Z/\gamma$+jets	Correlated per process	SHERPA scale variation samples
Generator comparison	GeneratorTop	$t\bar{t}$		POWHEG-BOX vs aMC@NLO
Radiation	radiationTop	$t\bar{t}$	Correlated per process	High vs Low
Parton shower	Pythia8/HerwigppTop	$t\bar{t}$		PYTHIA6 P2011 vs HERWIG++
Cross-section and acceptance				
- for Z+jets (κ-factor)	Kappa	Z/γ	Uncorrelated	
- for di-boson	FlatDiboson	Diboson	Uncorrelated	±30%
- for multi-jet	QCDError	Multijets	Uncorrelated	±100%
- for W+jets	FlatW	W+jets	Uncorrelated	+100% only for SRO4
- for Top	FlatTop	$t\bar{t}$ +single top	Uncorrelated	+100% only for SRD4
W_{trk} calibration (only for SRD2–4 and SRO2–5)				
PDF uncertainty for quark/gluon/b-quark/other	Wtrk_PDF_Quark/Gluon/Bquark/Other	All	Fully correlated	
Calibration systematic uncertainties for quark/gluon/b-quark/other in i-th jet ($i = 1, 2, 3, 4$)	Wtrk_Stat_NonClosure_Trk_Quark/Gluon/Bquark/Other_i-j	All	Fully correlated	
Process dependency (Boosted W)	Wtrk_Signal_EventTopology_Quark_12 Wtrk_Signal_EventTopology_Quark_i-j ($i = 3, 4$)	One-step decay signal	Uncorrelated	
Signal modeling				
Acceptance		signal	Uncorrelated	Varying from 30% to 0% depending on $\Delta M(\tilde{g}, \tilde{\chi}_1^0)$

Pileup reweighting As described in Section 2.3.2, pileup reweighting is applied based on the distributions of the mean number of interactions per bunch crossing $<\mu>$. After the reweighting, the discrepancy in the number of vertices N_{vtx} is found[1]. Since the number of vertices is more effective to describe the physics in p-p collisions, $<\mu>$ distribution of data is scaled by 1.09 to take into account the correlation with N_{vtx} distribution. The scale factor of $<\mu>$ is varied to 1.0 and 1.18 to obtain the systematic uncertainty on the pileup reweighting.

CR statistic The uncertainty on the μ_j in Eq. 9.2[2]. This uncertainty is caused by the statistical uncertainty of the number of the observed data in the CRs. This kind of uncertainty is dominant in the all the uncertainties in this search because the tight BDT cut applied in the CRs as in the SR decreases their statistics.

9.3.2 Object modeling uncertainties

Jet energy scale (JES) The uncertainty on the jet energy scale calibration (see Section 3.3.2).

Jet energy resolution (JER) The uncertainty on the jet energy resolution calibration (see Section 3.3.2).

---

[1] The $<\mu>$ is calculated from the instance measured luminosity, but the N_{vtx} is calculated from the observed tracks.

[2] This is not a nuisance parameter, but a free parameter.

E_T^{miss} soft term scale The uncertainty on the energy scale of the soft term in the E_T^{miss} (see Section 3.4).

E_T^{miss} soft term resolution The uncertainty on the energy resolution of the soft term in the E_T^{miss} (see Section 3.4).

b-tagging efficiency (b/c/light) The uncertainty on the efficiency correction of the b-tagging [64]. This uncertainty is only taken into account in the CRW and CRT because the b-tagging is used only in them.

There are also uncertainties on the other objects such as electrons, muons, and photons. But they are ignored because their impacts are very small. All the object modeling uncertainties are treated as fully correlated uncertainties because they are object-by-object uncertainties and not affected by the event topology or other objects.

9.3.3 Physics process modeling uncertainties

Scale variation in SHERPA ($Z/W/\gamma$+jets) The uncertainties of the renormalization scale, factorization scale, resummation scale, and CKKW matching scale in the SHERPA V+jets MCs are considered. These uncertainties are estimated by the study in the simulated samples produced with different scale variations in truth level. The variations are:

- Renormalization ("renom") scale (see Section 4.1): ×2 and ×1/2 from the nominal.
- Factorization ("fact") scale (see Section 4.1): ×2 and ×1/2 from the nominal.
- Resummation ("qsf") scale [65]: ×2 and ×1/2 from the nominal.
- CKKW ("ckkw") scale [66]: 15 GeV and 30 GeV while the nominal is 20 GeV.

Generator comparisons ($t\bar{t}$) A variation of the $t\bar{t}$ MC generated by aMC@NLO from the nominal MC generated by POWHEG-BOX is considered to take into account the possible difference on the matrix element calculation, conservatively.

Radiation ($t\bar{t}$) Variations of the simulated events generated by the different factorization and renormalization scale (×2 and ×1/2) from the nominal sample are considered.

Parton shower ($t\bar{t}$) Possible uncertainty of the parton shower modeling is evaluated by comparing simulated samples using PYTHIA6 and HERWIG++.

Cross-section and acceptance There are uncertainties on the cross-section and acceptance in each background process that are not correlated between the SR and CR.

- For Z/γ+jets (κ-factor): The uncertainty on the κ-factor calculation described in Section 8.2.1. It is dominated by the statistical error in the CRZL.
- For diboson: An additional uncertainty conservatively covering the scale and other-possible variations is introduced since there is no dedicated CR for the diboson process. (±30%)

- For multi-jet: An additional uncertainty as a conservative way to cover possible mis-modeling of the QCD simulation is considered. (±100%)
- For W+jets (Only SRO4): Conservatively, an additional uncertainty is considered in SRO4 to cover the observed difference between the data and background prediction in the VRWJdPhi of SRO4; see Section 10.1. (+100%)
- For $t\bar{t}$ +single top (Only SRD4): An additional uncertainty is introduced conservatively to cover the observed difference in the VRTJdPhi of SRD4; see Section 10.1. (+100%)

9.3.4 W_{trk} calibration uncertainties (Only in SRD2–4 and SRO2–5)

The systematic uncertainties associated with the W_{trk} calibration (Section 6) are calculated as variations of the scale factor (SF). The SF is prepared to be applied jet-by-jet depending on the jet p_{T} and W_{trk}. In this analysis, because the W_{trk}s of the leading four jets are used, the four SFs are multiplied by the event weight of each event in the MC. In the W_{trk} calibration, four sources of the systematic uncertainty are considered: PDF uncertainty, MC non-closure, data statistical error, and tracking uncertainties. The MC non-closure, data statistical error, and tracking uncertainties are merged to one set of up/down SF variations ("W_{trk} calibration uncertainty") by quadrature sum of their variations in order to simplify the systematic uncertainties. The PDF uncertainty is taken into account separately. The up/down variations of the PDF uncertainty are correlated between the leading four jets conservatively because the PDF uncertainty is a common source for all the four jets. In contrast, the merged W_{trk} calibration uncertainty coming from the calibration procedure are considered on each jet.

Both the PDF and calibration uncertainties are considered separately for each parton flavor since the calibration is done for each flavor (quark, gluon, b-quark or "other"). Hence, there are 4 PDF uncertainties and 4×4[3] W_{trk} calibration uncertainties in total.

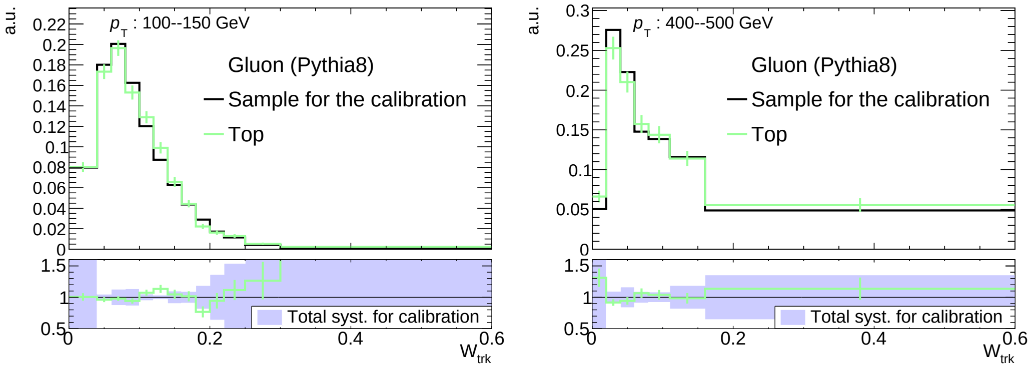
W_{trk} uncertainties on the process dependence

In the W_{trk} calibration, the W_{trk} distributions obtained from data are compared with Z+jets and multi-jet MC simulations ("calibration samples") in the looser kinematic regions than the SRs as explained in Section 6. Thus, if there is a difference in the W_{trk} distribution between the calibration samples and the samples used in the analysis, an additional uncertainty on the difference is necessary. The difference in the W_{trk} distribution between them is checked in the MC simulations as shown in Figures 9.1 for gluon jets in the background process and in Figures 9.2 for quark jets in the signal process. Here, top ($t\bar{t}$ + single top) and signal simulations use PYTHIA8 parton showering model, and Z+jets and W+jets simulations use SHERPA parton showering model. The selection of the samples used in the analysis is the preselection (PreDHigh or PreOHigh) in order to ensure enough MC statistic. The difference between the calibration samples and the samples used in the analysis is

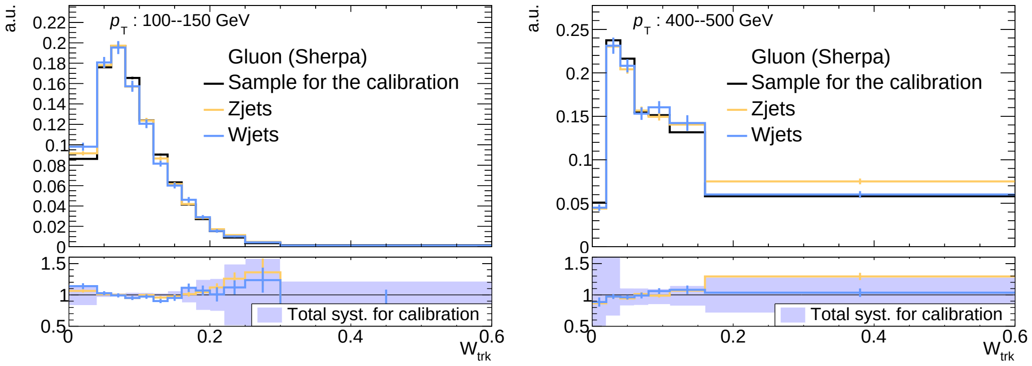
---

[3] The first factor 4 represents a separation of the uncertainty to each flavor (quark, gluon, b-quark,or "other"), and the second factor 4 represents an additional separation to each jet of the leading four jets. The number of total systematic uncertainties for the W_{trk} calibration is sixteen.

smaller than the SF uncertainties obtained in the calibration and it is negligible except the one-step decay signal. This statement can be valid in the real SR because the calibration samples in the loose kinematic region and the samples in the preselection have only negligible difference. Thus, the uncertainty on the process dependence is not taken into account in this analysis for the background processes and the direct decay signal.



(a) Top : $100 < p_T < 150$ GeV (Pythia8) (b) Top : $400 < p_T < 500$ GeV (Pythia8)

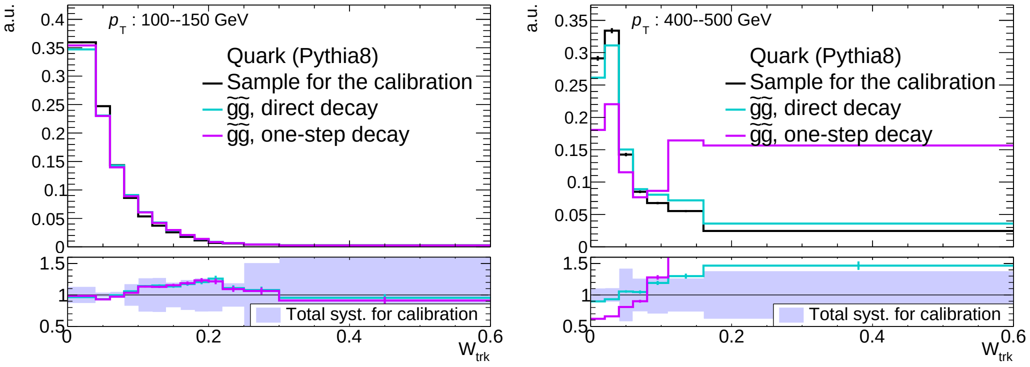


(c) Z/W+jets : $100 < p_T < 150$ GeV (Sherpa) (d) Z/W+jets : $400 < p_T < 500$ GeV (Sherpa)

Figure 9.1: Distributions of W_{trk} for gluon jets in the MC samples used in the quark/gluon calibration (Section 6) and (a)(b) top and (c)(d) Z+jets/W+jets process. The samples for the calibration are at the loose selection as in the calibration, and the samples of top and Z+jets/W+jets are at the preselection **PreDHigh**. The left two figures are in jet p_{T} range between 100 and 150 GeV. The right two figures are in jet p_{T} range between 400 and 500 GeV. In the bottom panel in each figure, each colored line shows the ratio to the gluon sample for the calibration and light blue area shows the magnitude of the total uncertainties of the SF obtained in Section 6. The total number of events in each sample is normalized to 1.

The difference between the calibration samples and one-step decay signal is caused by the W-boson in the one-step decay. As explained Section 7.3, lepton veto is required so that W decays hadronically to two quarks. In the high p_{T} regions, the W is boosted and the two quarks are emitted collinearly and reconstructed as one jet. An angular ΔR_{qq} between the two quarks is approximately [67]

$$\Delta R_{qq} \sim \frac{m_W}{2p_{\mathrm{T},W}}, \tag{9.9}$$



(a) Siganls : $100 < p_{\mathrm{T}} < 150$ GeV (Pythia8) (b) Signals : $400 < p_{\mathrm{T}} < 500$ GeV (Pythia8)

Figure 9.2: Distributions of W_{trk} for quark jets in the MC samples used in the quark/gluon calibration (Section 6) and gluino direct and one-step decay signals in (a) jet p_{T} range between 100 and 150 GeV and (b) the range between 400 and 500 GeV. The samples for the calibration are at the loose selection as in the calibration, and the samples of the direct decay and one-step decay signals are at the preselection **PreDHigh** and **PreOHigh**, respectively. The signal samples are merged ones with different mass points in each decay. In the bottom panel in each figure, each colored line shows the ratio to the quark sample for the calibration and light blue area shows the magnitude of the total uncertainties of the SF obtained in Section 6. The total number of events in each sample is normalized to 1.

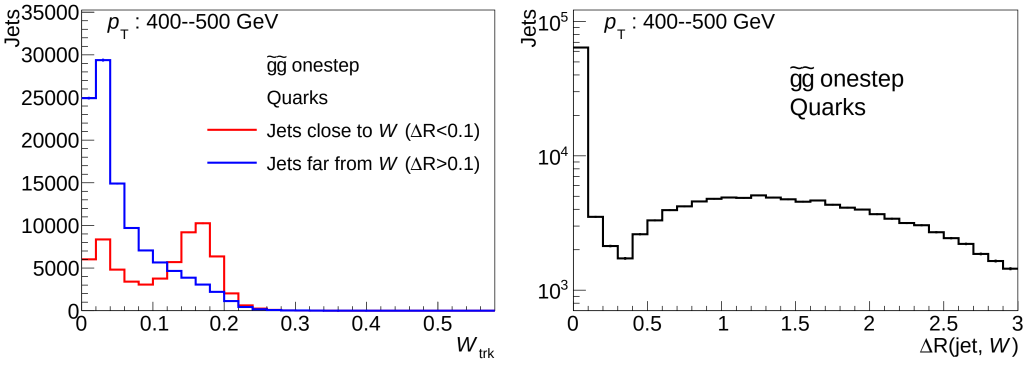
where $p_{\mathrm{T},W}$ and m_W are transverse momentum and mass (80.379 ± 0.012 GeV, [27]) of the W, respectively. Thus, in the jet reconstruction of anti-k_t algorithm with $R = 0.4$ (Section 3.3) adopted in this analysis, the two quarks are possibly merged in the $p_{\mathrm{T},W} > 400$ GeV. Figure 9.3(a) shows the W_{trk} distributions of jets close to and far from the W boson in the one-step decay signal. A jet with $\Delta R < 0.1$ to the W is defined as a jet close to W in the figure. The ΔR distribution is shown in Figure 9.3(b). The W_{trk} distribution of the jets far from the W has similar to that of isolated quark jets as seen in calibration samples. However, that of the jets close to the W has a different distribution from it.

An additional W_{trk} uncertainty on the MC modeling related to the boosted W is necessary for the one-step decay signal sample. The boosted $W \to qq$ decay is better-described process in the simulation than the fragmentation of quark jets, which describes the W_{trk} distribution. The mis-modeling in the W_{trk} distribution for the boosted W in the MC is at least smaller than the difference between the isolated quark-jets and one-step decay jets including the boosted W as shown in Figure 9.2. Then, the W_{trk} uncertainty on the boosted W jet ("boosted W uncertainty") in the one-step decay signal is evaluated from the difference in Figure 9.2. The SF up and down variations for this uncertainty on quark jets are defined as

$$\begin{cases} SF_{Q,\mathrm{up}}(W_{\mathrm{trk}}; p_{\mathrm{T,j}}) & = & SF_{Q,\mathrm{nom}}(W_{\mathrm{trk}}; p_{\mathrm{T,j}}) + \Delta SF(W_{\mathrm{trk}}; p_{\mathrm{T,j}}) \\ SF_{Q,\mathrm{down}}(W_{\mathrm{trk}}; p_{\mathrm{T,j}}) & = & SF_{Q,\mathrm{nom}}(W_{\mathrm{trk}}; p_{\mathrm{T,j}}) - \Delta SF(W_{\mathrm{trk}}; p_{\mathrm{T,j}}) \text{ or 0 if it is negative} \end{cases} \tag{9.10}$$

$$\Delta SF(W_{\mathrm{trk}}; p_{\mathrm{T,j}}) = \left| 1 - \frac{p_{Q,\mathrm{onestep}}(W_{\mathrm{trk}}; p_{\mathrm{T,j}})}{p_{Q,\mathrm{calib}}(W_{\mathrm{trk}}; p_{\mathrm{T,j}})} \right|, \tag{9.11}$$

where $SF_{Q,\mathrm{nom/up/down}}$ are nominal-, up-, and down-scale-factors for quark jets, $p_{\mathrm{T,j}}$ is j-th jet p_{T} range, and $p_{Q,\mathrm{onestep}}$ and $p_{Q,\mathrm{calib}}$ are the W_{trk} distributions of quark jets in the one-step decay signal



(a)

(b)

Figure 9.3: (a) Distributions of W_{trk} for jets close to and far from a W boson and (b) a distribution of $\Delta R(\mathrm{jet}, W)$ in the one-step decay signal at the preselection **PreOHigh**. If there are two hadronic decay W bosons, the ΔR is defined as a smaller one in the two Ws. The jets close to the W boson are defined as the jets with $\Delta R(\mathrm{jet}, W) < 0.1$ and the other jets are defined as jets far from the W. The jets are quark jets and their p_{T} range is from 400 to 500 GeV. The signal sample is merged one with different mass points.

and calibration MC samples as shown in Figure 9.2, respectively. The up and down variations of the signal yield in each SR are obtained by using the $SF_{\mathrm{up/down}}$ instead of SF_{nom} over all the events in the signal MC sample. The systematic variation in the W_{trk} distribution of the leading jet in the SRO2 is shown in Figure 9.4. The light blue area is the range between the up and down variations for the leading jet. The number of signal events with this systematic uncertainty on the leading jet is $24.8^{+31.2}_{-11.3}(^{+128\%}_{-46\%})$ events for the one-step decay signal at the $(m_{\tilde{g}}, m_{\tilde{\chi}_1^\pm}, m_{\tilde{\chi}_1^0}) = (1665, 1265, 865)$ GeV. This boosted W uncertainty is large in the one-step decay signals with a large $\Delta M(\tilde{g}, \tilde{\chi}_1^0)$ and covers the second broad peak caused by the boosted $W \to qq$ jets around $W_{\mathrm{trk}} = 0.16$ as shown in Figure 9.4. Thus, the systematic uncertainty is enough conservative.

The calibration in Section 6 is only for isolated jets and not valid for the boosted W jets. Thus, such a large systematic uncertainty on it, which is close to 100% of the number of the boosted W jet events, is taken into account. However, the BDT training is optimal if there is no mis-modeling in the boosted W jet of the signal MC sample because the signal sample used in the BDT training simulates the boosted W jet as well.

This systematic uncertainty is considered for the four leading jets of the one-step decay signal only in SRO2–5, where W_{trk}s of the leading four jets are used in the BDT training. A jet close to W ($\Delta R < 0.1$) is often the first or second p_{T} jet as shown in Figure 9.5 and events having two jets close to W ($N_{W-\mathrm{jet}} = 2$) is dominated by the events in which the first and second jets are close to W. Thus, the systematic uncertainties on the first and second jets are correlated considering the events having two jets close to W. The remaining part of the $N_{W-\mathrm{jet}} = 1$ events consists of events in which the third or fourth jet is close to W. The remaining part (4.4%) of the $N_{W-\mathrm{jet}} = 2$ events consists of events in which the first or second jet is close to W and the third or fourth jet is close to the other W. The correlation between the two boosted W jets in that events should be considered but the fraction is enough small respect to the first and second jet uncertainty ($^{+249}_{-57}\%$) to ignore it. Thus, the systematic

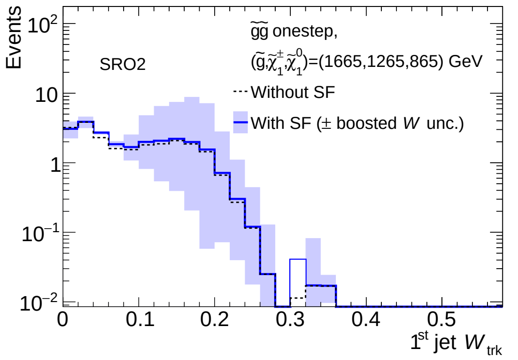


Figure 9.4: W_{trk} distributions of the leading jet in the one-step decay signal MC at the $(m_{\tilde{g}}, m_{\tilde{\chi}_1^\pm}, m_{\tilde{\chi}_1^0}) = (1665, 1265, 865)$ GeV in the SRO2. A black break line is a distribution before applying the SF, and a blue solid line is after applying the SF. The light blue area shows the range between the up and down variations of the uncertainty on the boosted W uncertainty.

uncertainties on the third and fourth jets are taken into account separately and not correlating to the uncertainty on the other jets.

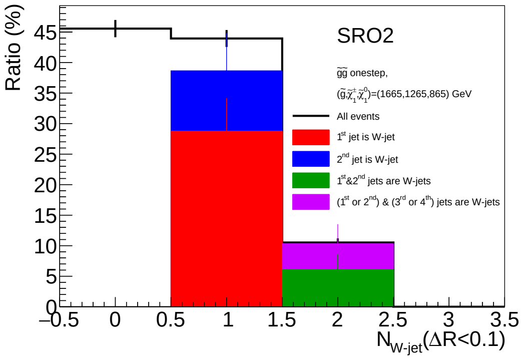


Figure 9.5: Event fraction in each number of boosted $W \to qq$ jets in the leading four jets ($N_{W\text{-jet}}$) for the one-step decay signal at the $(m_{\tilde{g}}, m_{\tilde{\chi}_1^\pm}, m_{\tilde{\chi}_1^0}) = (1665, 1265, 865)$ GeV in the SRO2. The boosted W jet is defined by $\Delta R(\text{jet}, W) < 0.1$. $N_{W\text{-jet}} = 1$ events are composed of three types of events in which the first jet (red-filled fraction), second jet (blue-filled fraction), or the other jet in the leading four jets is a boosted W jet. In the $N_{W\text{-jet}} = 2$, green-filled fraction shows events in which the first and second jets are boosted W jets and violet-filled fraction shows events in which the first or second jet is a boosted W jet and the third or fourth jet is the other one.

9.3.5 Signal modeling uncertainties

For the signal MC, the general uncertainties and object modeling uncertainties are taken into account. In addition to them, the scale variations such as the renormalization scale and factorization scale of the MC generator are also considered. However, their scale variations do not impact on the signal acceptance to the high $\Delta M(\tilde{g}, \tilde{\chi}_1^0)$ signals. They affect only the low $\Delta M(\tilde{g}, \tilde{\chi}_1^0)$ signals below

100 GeV. The uncertainty is studied in the MCs with varied scales at the truth level and parametrized as a function of $\Delta M(\tilde{g}, \tilde{\chi}_1^0)$ to be $6.51 \times \exp(-0.04\Delta M + 1.44)$ %.

10 Result and Interpretation

10.1 Validation

The normalization factors $\mu(Z\text{+jets})$, $\mu(W\text{+jets})$,$\mu(\text{Top})$, and $\mu(\text{Multijet})$ obtained by the background-only fit are summarized in Figure 10.1. In the SRD1–3 and SRO1, $\mu(\text{Top})$ is much lower than 1 due to the mis-modeling in the high m_{eff} range of the top MC. In the SRO1–O5, $\mu(W\text{+jets})$ has also lower value due to the mis-modeling in the high N_{jet} of the W+jets MC. They are the same tendency as shown in Figures 8.11 and 8.12 of Section 8.2.2. $\mu(Z\text{+jets})$ in the SRO1–O5 is also smaller than 1. This is validated by the VRZs later. The details of the event yields in the CRs are listed in Appendix C.2.

The validation of the normalization factors is performed in the VRs. The summary of the difference between the observed data and total SM background prediction corrected by the normalization factors is shown in Figure 10.2. The difference between the data and background prediction in each VR is shown in each cell and the numbers of the difference are normalized by the standard deviation σ_{tot} defined as $\sigma_{\text{tot}} \equiv \sqrt{\sigma_{\text{obs}}^2 + \sigma_{\text{pred}}^2}$, where σ_{obs} and σ_{pred} are the uncertainties on the data and background prediction, respectively. There is no significant difference in all the VRs. Numbers of the observed data and background prediction in the several VRs are shown in Figures 10.3. The statistics in the VRZs, where two leptons of $Z \to ll$ are tagged, are too limited, and some VRZs have no observed event but it is consistent within the error. VRZs with loosened BDT score (BDT$>$ 0.0) are also checked in Figure 10.3(b). In these regions, the normalization factors obtained in the CRs are the same as in VRZs. The predicted number of the background have a good agreement with the number of observed events there. The VRWJs have a good agreement between the data and background prediction as shown in Figure 10.3(c). In the VRTJs for SRD2–3, there are pulls of the data from the background prediction (Figure 10.3(d)). However, the pull is covered by the large uncertainties of the background prediction and the top process does not impact on the predictions in the SRs because the SRs have only small top-process contribution. In the VRWJdPhi for SRO4 and the VRTJdPhi for SRD4, which have fewer statistics than the VRW/TJ, there were non-negligible pulls of 2.2σ without additional uncertainties on the W+jets and top process; the numbers of the observed data and background prediction were 49 (12) and 27.84 ± 8.30 (5.13 ± 1.88) in the VRWJdPhi (VRTJdPhi). It was possible that the pulls were just caused by fluctuations. However, conservatively, flat 100% systematic uncertainties are added to the W+jets and top MC events in the regions of SR, CRs, and VRs for SRO4 and SRD4, respectively. Now, the data and background prediction in all the VRWJdPhis and VRTJdPhis are consistent within the uncertainties. The other VRs are shown in Appendix C.1. There is also no significant difference between the data and background prediction.

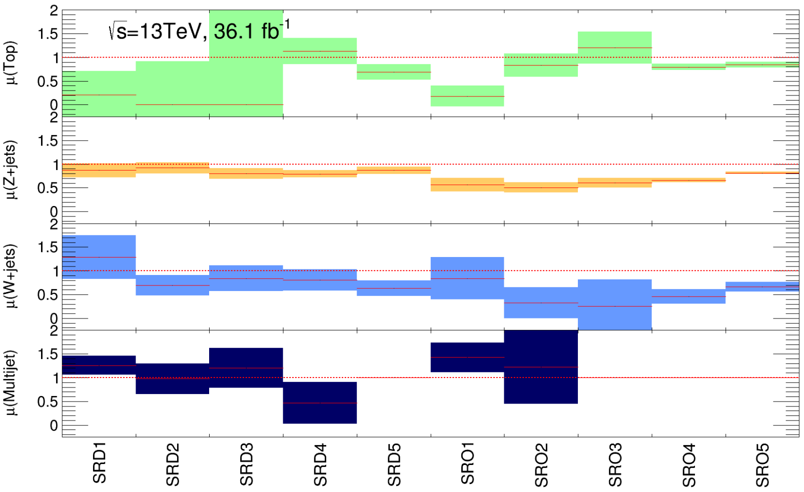
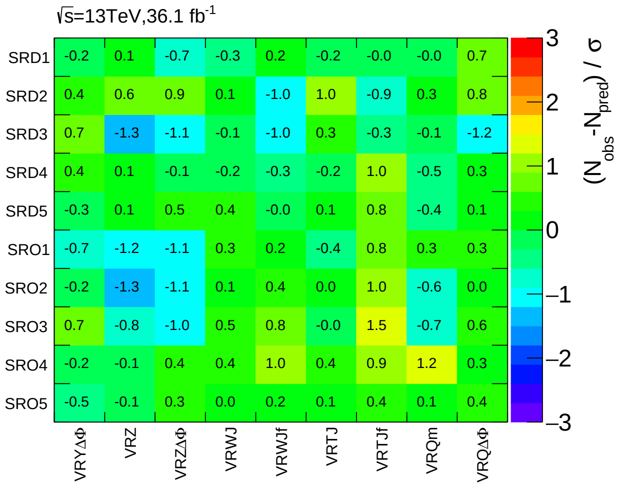
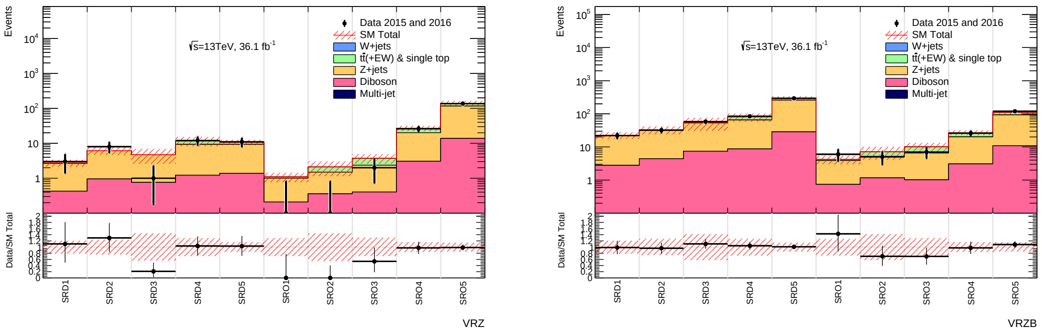


Figure 10.1: A summary of the normalization factors μs obtained by the background-only fit: μ(Top), μ(Z+jets), μ(W+jets), and μ(Multi − jet). Each bin indicates each SR and each row shows each μ. A solid red line and colored band in each bin and each row show the nominal value and its uncertainty of each μ, respectively. μ(Multi − jet) is not fitted as a free parameter in SRD5, SRO3–SRO5 because there is no CRQ and it is fixed to 1.



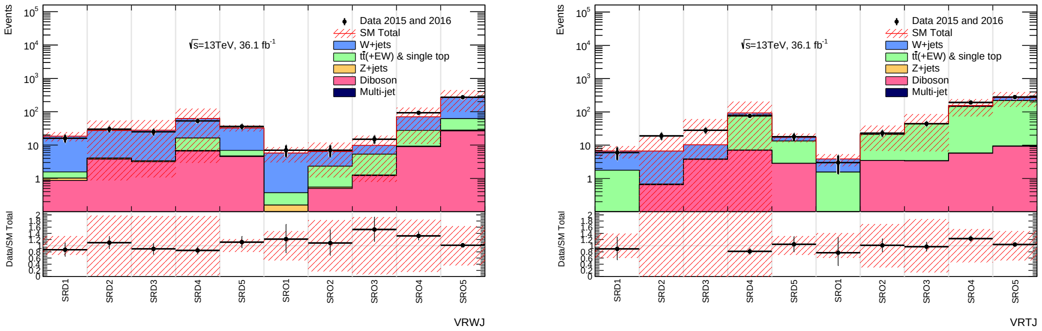
	VRYΔΦ	VRZ	VRZΔΦ	VRWJ	VRWJf	VRTJ	VRTJf	VRQm	VRQΔΦ
SRD1	-0.2	0.1	-0.7	-0.3	0.2	-0.2	-0.0	-0.0	0.7
SRD2	0.4	0.6	0.9	0.1	-1.0	1.0	-0.9	0.3	0.8
SRD3	0.7	-1.3	-1.1	-0.1	-1.0	0.3	-0.3	-0.1	-1.2
SRD4	0.4	0.1	-0.1	-0.2	-0.3	-0.2	1.0	-0.5	0.3
SRD5	-0.3	0.1	0.5	0.4	-0.0	0.1	0.8	-0.4	0.1
SRO1	-0.7	-1.2	-1.1	0.3	0.2	-0.4	0.8	0.3	0.3
SRO2	-0.2	-1.3	-1.1	0.1	0.4	0.0	1.0	-0.6	0.0
SRO3	0.7	-0.8	-1.0	0.5	0.8	-0.0	1.5	-0.7	0.6
SRO4	-0.2	-0.1	0.4	0.4	1.0	0.4	0.9	1.2	0.3
SRO5	-0.5	-0.1	0.3	0.0	0.2	0.1	0.4	0.1	0.4

Figure 10.2: A summary of pulls in the VRs. Each cell corresponds one VR. The x-axis shows the type of the VR in each column. The y-axis shows the corresponding SR to the VRs in each row. The z-axis is a difference between the observed data yield and predicted background yield divided by the quadrature sum of the statistical error on the data and total uncertainties on the background prediction; $z = \left(N_{\text{obs}} - N_{\text{pred}}\right) / \sqrt{\sigma_{\text{obs}}^2 + \sigma_{\text{pred}}^2}$. The text in each cell shows the z-axis value in each VR.



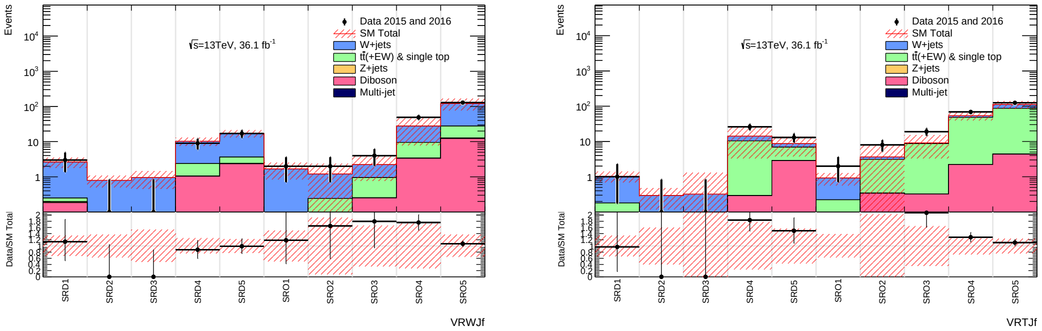
(a) VRZ

(b) VRZ with a loose BDT cut (BDT>0.0)



(c) VRWJ

(d) VRTJ



(e) VRWJdPhi

(f) VRTJdPhi

Figure 10.3: Numbers of the observed data and SM background prediction in (a) VRZ, (b) VRZ with a loose BDT cut, (c) VRWJ, (d) VRTJ, (e) VRWJdPhi, and (f) VRTJdPhi after the background-only fit. Each bin indicates an observed number of data (black point) and a predicted number of the SM background (color filled bars). Each color shows one kind of background process. The red shaded band to the SM total background represents its total uncertainties. The bottom panel shows the number of the data divided by the number of the total background prediction.

10.2 Result

The results of the ten SRs are shown in Figure 10.4. There is no significant excess in all the SRs. The observed event yield and the predicted background event yields are summarized in Table 10.4 and breakdowns of the dominant systematic uncertainties on the background prediction are summarized in Tables 10.2 and 10.3. The main background process is $Z(\to \nu\nu)$ events, except SRO2–4, where the top process is dominant due to the high jet multiplicity requirement ($N_{\mathrm{jet}} \geq 6$). The dominant systematic uncertainties are the uncertainty on μs due to the not large statistics in the CRs. The systematic uncertainties of the Z+jets scale variation ("zScale_XXX") and the PDF variations on the gluon W_{trk} ("Wtrk_PDF_Gluon") are also major in some SRs.

The distribution of the most discriminating variable m_{eff} in the previous analysis is shown in Figures 10.5 and 10.6. There is no signal-like excess in the distributions as well.

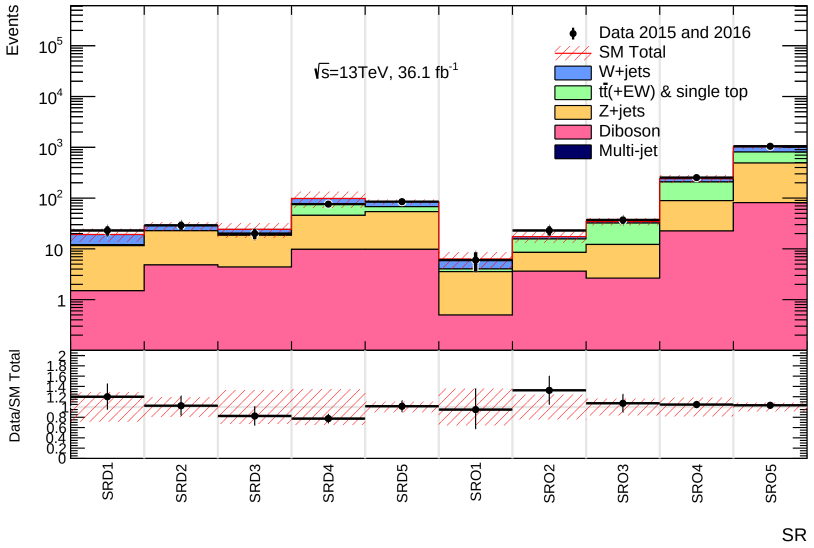


Figure 10.4: Numbers of the observed data and SM background prediction in the ten SRs after the background-only fit. Each bin indicates an observed number of data (black point) and a predicted number of the SM background (color filled bars). Each color shows one kind of background process. The red shaded band to the SM total background represents its total uncertainties. The bottom panel shows the number of the data divided by the number of the total background prediction.

Table 10.1: Numbers of events observed in the SRs compared with the background predictions obtained from the background-only fits. Empty cells (indicated by a '-') mean the prediction is lower than 0.01.

Signal Region	SRD1	SRD2	SRD3	SRD4	SRD5
MC expected events					
Diboson	1.46	4.81	4.39	9.78	9.78
Z/γ^*+jets	11.59	19.42	17.53	45.18	50.87
W+jets	5.51	8.04	6.80	25.15	25.33
$t\bar{t}$ + single top	2.49	2.26	2.93	28.32	19.66
Multi-jet	0.03	–	–	–	–
Fitted background events					
Diboson	1.5 ± 0.5	4.8 ± 2.6	4.4 ± 2.3	10 ± 4	9.8 ± 3.0
Z/γ^*+jets	10 ± 5	18 ± 4	14 ± 4	36 ± 7	44 ± 6
W+jets	7.1 ± 2.7	5.6 ± 1.8	5.7 ± 2.8	20 ± 6	16 ± 4
$t\bar{t}$ + single top	$0.52^{+1.21}_{-0.52}$	$0.00^{+1.04}_{-0.00}$	$0.01^{+5.85}_{-0.01}$	$32.05^{+33.22}_{-32.05}$	14 ± 5
Multi-jet	$0.04^{+0.04}_{-0.04}$	–	–	–	–
Total MC	21.08	34.52	31.64	108.42	105.64
Total bkg	19 ± 6	28 ± 5	24 ± 8	98 ± 34	84 ± 9
Observed	23	29	20	76	85

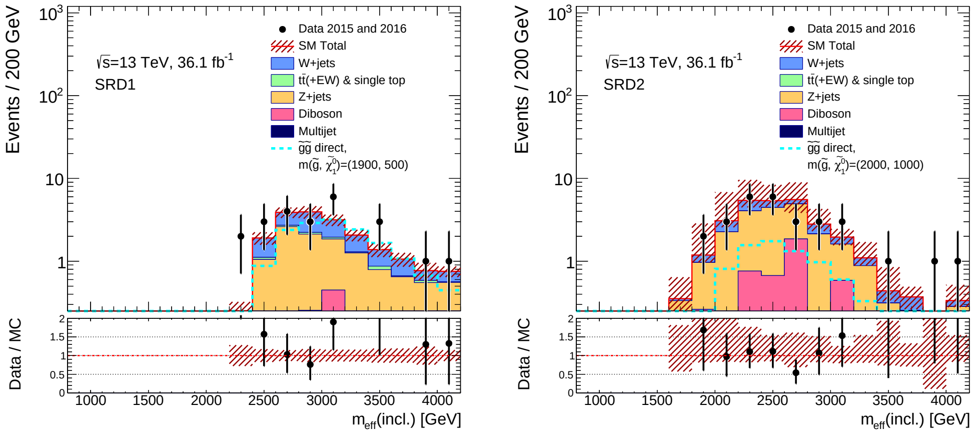
Signal Region	SRO1	SRO2	SRO3	SRO4	SRO5
MC expected events					
Diboson	0.50	3.61	2.65	22.57	81.22
Z/γ^*+jets	5.43	9.62	15.68	101.02	502.19
W+jets	2.72	5.21	9.06	73.45	311.52
$t\bar{t}$ + single top	2.55	8.57	16.70	149.36	376.36
Multi-jet	–	–	–	–	–
Fitted background events					
Diboson	0.5 ± 0.17	3.6 ± 2.5	2.6 ± 0.9	23 ± 9	81 ± 32
Z/γ^*+jets	3.1 ± 2.0	4.9 ± 1.5	9.5 ± 2.9	66 ± 14	410 ± 70
W+jets	2.3 ± 1.2	$1.71^{+1.72}_{-1.71}$	$2.28^{+3.72}_{-2.28}$	$33.69^{+35.73}_{-33.69}$	210 ± 40
$t\bar{t}$ + single top	$0.46^{+0.53}_{-0.46}$	7.2 ± 3.2	20 ± 7	119 ± 21	320 ± 40
Multi-jet	–	–	–	–	–
Total MC	11.20	27.01	44.09	346.41	1271.28
Total bkg	6.3 ± 2.3	17 ± 4	35 ± 6	240 ± 40	1010 ± 80
Observed	6	23	37	253	1047

Table 10.2: Breakdowns of the dominant systematic uncertainties on the background predictions in SRD1, SRD2, SRD3, SRD4, and SRD5. Note that the individual uncertainties can be correlated, and do not necessarily add up quadratically to the total background uncertainty. The percentages show the size of the uncertainty relative to the total expected background.

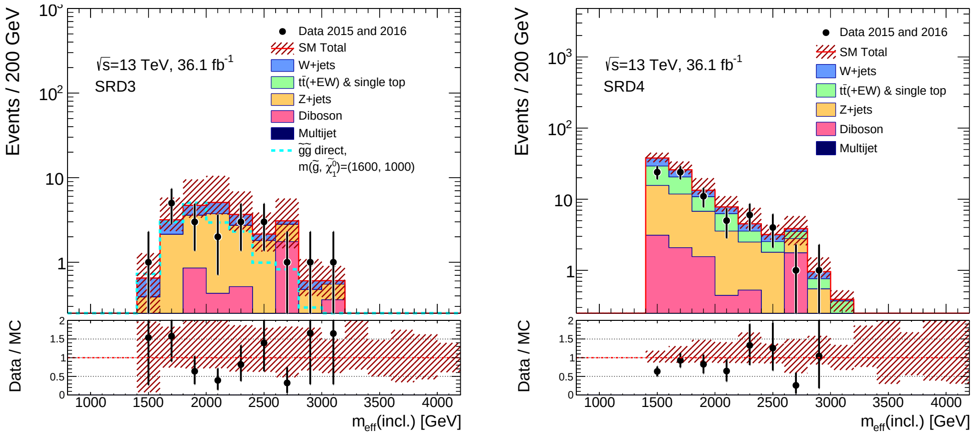
SRD1		SRD2		SRD3		SRD4		SRD5	
$N_{\mathrm{pred}}^{\mathrm{bkg}}$	19.17	$N_{\mathrm{pred}}^{\mathrm{bkg}}$	28.30	$N_{\mathrm{pred}}^{\mathrm{bkg}}$	24.24	$N_{\mathrm{pred}}^{\mathrm{bkg}}$	98.12	$N_{\mathrm{pred}}^{\mathrm{bkg}}$	83.85
$\sqrt{N_{\mathrm{pred}}^{\mathrm{bkg}}}$	±4.38	$\sqrt{N_{\mathrm{pred}}^{\mathrm{bkg}}}$	±5.32	$\sqrt{N_{\mathrm{pred}}^{\mathrm{bkg}}}$	±4.92	$\sqrt{N_{\mathrm{pred}}^{\mathrm{bkg}}}$	±9.91	$\sqrt{N_{\mathrm{pred}}^{\mathrm{bkg}}}$	±9.16
$\Delta N_{\mathrm{syst.}}^{\mathrm{bkg}}$	±5.52 [28.79%]	$\Delta N_{\mathrm{syst.}}^{\mathrm{bkg}}$	±5.42 [19.17%]	$\Delta N_{\mathrm{syst.}}^{\mathrm{bkg}}$	±7.82 [32.28%]	$\Delta N_{\mathrm{syst.}}^{\mathrm{bkg}}$	±33.87 [34.52%]	$\Delta N_{\mathrm{syst.}}^{\mathrm{bkg}}$	±8.91 [10.63%]
zScale_renorm	±4.00 [20.9%]	Wtrk_PDF_Gluon	±3.35 [11.9%]	mu_Top	±5.85 [24.1%]	FlatTop	±31.89 [32.5%]	mu_W+jets	±4.15 [5.0%]
mu_W+jets	±2.55 [13.3%]	mu_Z+jets	±2.27 [8.0%]	Wtrk_PDF_Gluon	±3.41 [14.1%]	mu_Top	±7.74 [7.9%]	mu_Z+jets	±3.89 [4.6%]
mu_Z+jets	±1.74 [9.1%]	mu_W+jets	±1.73 [6.1%]	mu_Z+jets	±2.00 [8.3%]	mu_W+jets	±5.69 [5.8%]	zScale_fact	±3.49 [4.2%]
zScale_fact	±1.72 [9.0%]	stat_SR	±1.49 [5.3%]	mu_W+jets	±1.82 [7.5%]	Wtrk_PDF_Gluon	±4.82 [4.9%]	mu_Top	±3.23 [3.9%]
zScale_qsf	±1.05 [5.5%]	FlatDiboson	±1.45 [5.1%]	wScale_CT14	±1.69 [7.0%]	mu_Z+jets	±3.48 [3.5%]	PartonTop	±3.21 [3.8%]
zScale_ckkw	±0.95 [5.0%]	zScale_fact	±1.37 [4.8%]	stat_SR	±1.39 [5.7%]	zScale_fact	±3.20 [3.3%]	FlatDiboson	±2.93 [3.5%]
mu_Top	±0.88 [4.6%]	pileUp	±1.19 [4.2%]	FlatDiboson	±1.32 [5.4%]	FlatDiboson	±2.93 [3.0%]	stat_SR	±2.50 [3.0%]
stat_SR	±0.71 [3.7%]	Wtrk_PDF_Quark	±1.13 [4.0%]	Wtrk_Calib_Gluon_2j	±1.30 [5.4%]	PartonTop	±2.79 [2.8%]	GeneratorTop	±2.14 [2.6%]
GeneratorTop	±0.66 [3.5%]	mu_Top	±1.04 [3.7%]	zScale_fact	±1.12 [4.6%]	stat_SR	±2.58 [2.6%]	zScale_qsf	±1.93 [2.3%]
pileUp	±0.47 [2.4%]	zScale_qsf	±0.91 [3.2%]	wScale_MMHT2014	±1.07 [4.4%]	RadiationTop	±2.31 [2.4%]	κ-factor	±0.99 [1.2%]

Table 10.3: Breakdowns of the dominant systematic uncertainties on the background predictions in SRO1, SRO2, SRO3, SRO4, and SRO5. Note that the individual uncertainties can be correlated, and do not necessarily add up quadratically to the total background uncertainty. The percentages show the size of the uncertainty relative to the total expected background.

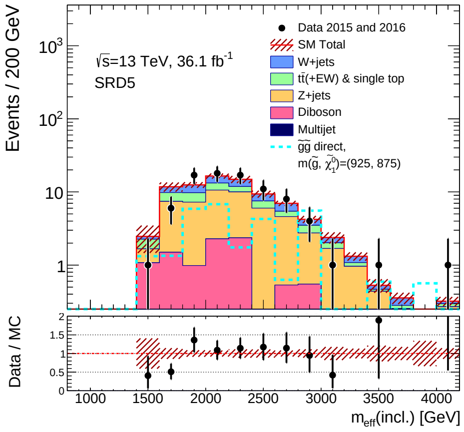
SRO1		SRO2		SRO3		SRO4		SRO5	
$N_{\mathrm{pred}}^{\mathrm{bkg}}$	6.32	$N_{\mathrm{pred}}^{\mathrm{bkg}}$	17.38	$N_{\mathrm{pred}}^{\mathrm{bkg}}$	34.53	$N_{\mathrm{pred}}^{\mathrm{bkg}}$	241.60	$N_{\mathrm{pred}}^{\mathrm{bkg}}$	1014.17
$\sqrt{N_{\mathrm{pred}}^{\mathrm{bkg}}}$	±2.51	$\sqrt{N_{\mathrm{pred}}^{\mathrm{bkg}}}$	±4.17	$\sqrt{N_{\mathrm{pred}}^{\mathrm{bkg}}}$	±5.88	$\sqrt{N_{\mathrm{pred}}^{\mathrm{bkg}}}$	±15.54	$\sqrt{N_{\mathrm{pred}}^{\mathrm{bkg}}}$	±31.85
$\Delta N_{\mathrm{syst.}}^{\mathrm{bkg}}$	±2.27 [35.90%]	$\Delta N_{\mathrm{syst.}}^{\mathrm{bkg}}$	±4.13 [23.79%]	$\Delta N_{\mathrm{syst.}}^{\mathrm{bkg}}$	±5.58 [16.15%]	$\Delta N_{\mathrm{syst.}}^{\mathrm{bkg}}$	±42.99 [17.79%]	$\Delta N_{\mathrm{syst.}}^{\mathrm{bkg}}$	±80.67 [7.95%]
zScale_renorm	±1.34 [21.2%]	mu_Top	±2.14 [12.3%]	mu_Top	±5.67 [16.4%]	FlatW	±33.52 [13.9%]	Wtrk_PDF_Gluon	±37.41 [3.7%]
mu_W+jets	±1.21 [19.2%]	mu_W+jets	±1.71 [9.9%]	mu_W+jets	±3.72 [10.8%]	GeneratorTop	±17.64 [7.3%]	zScale_fact	±36.37 [3.6%]
mu_Z+jets	±0.80 [12.7%]	Wtrk_PDF_Gluon	±1.59 [9.2%]	RadiationTop	±2.22 [6.4%]	mu_W+jets	±11.11 [4.6%]	mu_W+jets	±31.91 [3.1%]
zScale_fact	±0.72 [11.4%]	PartonTop	±1.39 [8.0%]	mu_Z+jets	±1.62 [4.7%]	mu_Top	±10.24 [4.2%]	GeneratorTop	±30.56 [3.0%]
zScale_ckkw	±0.67 [10.6%]	GeneratorTop	±1.28 [7.4%]	stat_SR	±1.27 [3.7%]	κ-factor	±7.36 [3.0%]	κ-factor	±28.27 [2.8%]
zScale_qsf	±0.52 [8.2%]	stat_SR	±1.12 [6.5%]	zScale_fact	±1.22 [3.5%]	zScale_fact	±7.14 [3.0%]	FlatDiboson	±24.37 [2.4%]
mu_Top	±0.51 [8.1%]	FlatDiboson	±1.09 [6.3%]	κ-factor	±1.06 [3.1%]	Wtrk_PDF_Gluon	±6.81 [2.8%]	mu_Top	±23.32 [2.3%]
κ-factor	±0.34 [5.4%]	Wtrk_Calib_Gluon_2j	±1.06 [6.1%]	JET_GroupedNP_1	±0.92 [2.7%]	FlatDiboson	±6.75 [2.8%]	zScale_qsf	±20.77 [2.0%]
stat_SR	±0.31 [4.9%]	mu_Z+jets	±1.05 [6.1%]	PartonTop	±0.87 [2.5%]	PartonTop	±6.31 [2.6%]	pileUp	±15.95 [1.6%]
JET_GroupedNP_3	±0.24 [3.8%]	Wtrk_Calib_Quark_2j	±0.84 [4.8%]	zScale_qsf	±0.82 [2.4%]	stat_SR	±5.53 [2.3%]	Wtrk_PDF_Quark	±15.33 [1.5%]



(a) SRD1 (b) SRD2

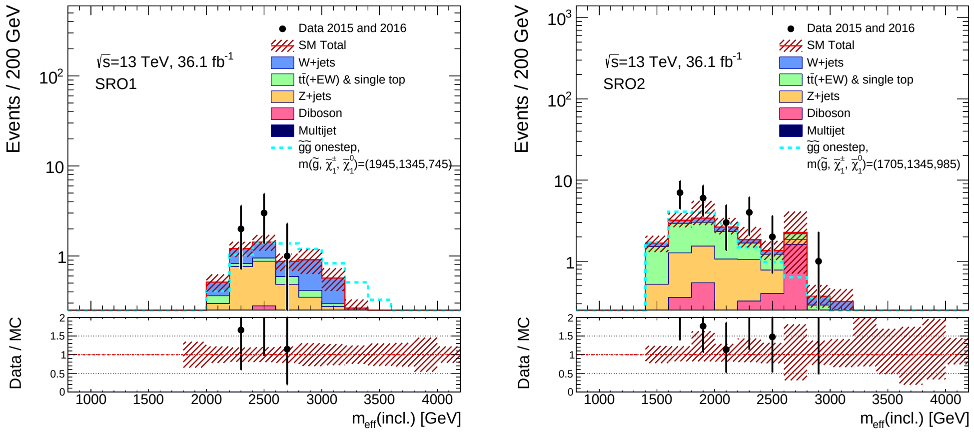


(c) SRD3 (d) SRD4

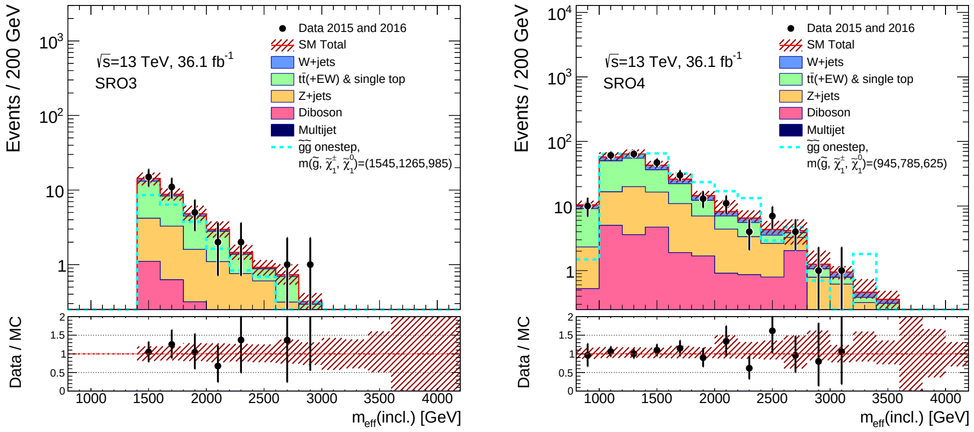


(e) SRD5

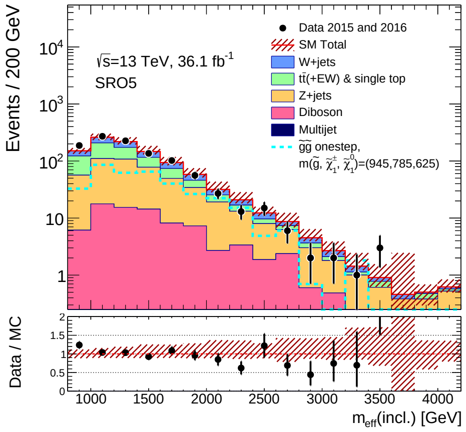
Figure 10.5: Distributions of m_{eff}(inc) in the SRs of (a) **SRD1**, (b) **SRD2**, (c) **SRD3**, (d) **SRD4**, and (e) **SRD5**. Black point shows 36.1fb^{-1}data and filled histograms show the SM background events, which is corrected by the normalization factors (μs). A red shaded area shows the uncertainties on the background prediction. The statistical uncertainty, the systematic uncertainties on the normalization factors, the Z+jets scale variations, and the modeling uncertainties on the top process are considered here.



(a) SRO1 (b) SRO2



(c) SRO3 (d) SRO4



(e) SRO5

Figure 10.6: Distributions of m_{eff}(inc) in the SRs of (a) **SRO1**, (b) **SRO2**, (c) **SRO3**, (d) **SRO4**, and (e) **SRO5**. Black point shows 36.1fb^{-1}data and filled histograms show the SM background events, which is corrected by the normalization factors (μs). A red shaded area shows the uncertainties on the background prediction. The statistical uncertainty, the systematic uncertainties on the normalization factors, the Z+jets scale variations, and the modeling uncertainties on the top process are considered here.

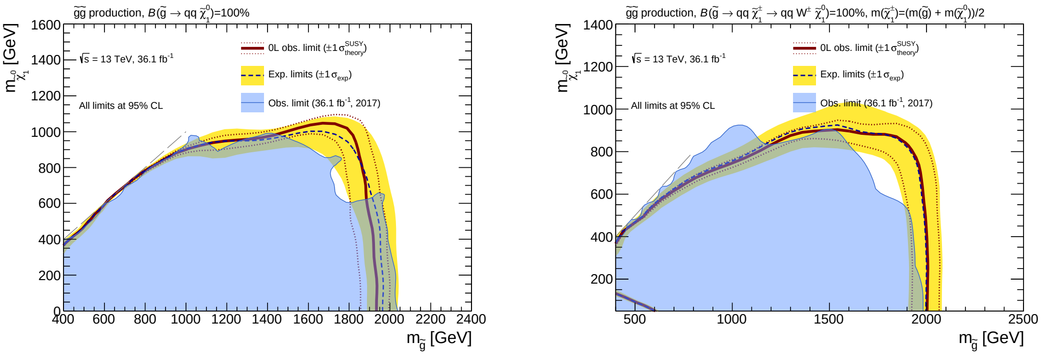
10.3 Interpretation

Since there is no significant excess in the SRs, the exclusion fit is performed to investigate their exclusion power on the benchmark signal models, i.e. the gluino direct decay and one-step decay signals. For each mass point in the two-dimensional plane of the gluino and $\tilde{\chi}_1^0$ masses, the exclusion fit described in Section 9.2.2 is performed, in which the signal is taken into account in the SR and also in the CRs. In each mass point, the best sensitive SR is selected based on the expected CLs, which is calculated using the SM background prediction instead of the data in the SR. From the observed CLs (calculated using the observed data) in the selected SR, a 95% confidence level exclusion limit is drawn as in Figures 10.7. The superimposed blue area is the observed exclusion limit by the previous study published for the Moriond 2017 conference [57] using the same data. The analysis in this thesis has more exclusion power than the previous study especially by $\sim$ 200 GeV of the $\tilde{\chi}_1^0$ mass in the high gluino mass range around 1.8 TeV in both of the direct and one-step decay signals due to the new techniques of the quark/gluon separation and the multivariate analysis (BDT). Table 10.4 lists background and signal yields in the best-sensitive SR at a high gluino and high $\tilde{\chi}_1^0$ mass point in each decay signal. The SRD3 has more background reduction power by factor 2 than the SR in the previous study with a similar signal yield of the direct decay signal. In the SRO2, a signal yield in a one-step decay signal is increased by more than factor 3. In both case, exclusion sensitivity (σ_{excl}) calculated from CLs is also increased by factor 2.5 or more. By this result, the signal mass region in which the previous analysis had small moderate excess is completely excluded. For the gluino direct decay model, the gluino mass is excluded up to **1.93, 1.92, 1.89, 1.86, and 1.85 TeV** at the $\boldsymbol{m(\tilde{\chi}_1^0) = 0.0, 0.4, 0.6, 0.8,}$ **and 0.9 TeV**, respectively. For the gluino one-step decay model, the gluino mass is excluded up to **2.00, 2.00, 1.97, and 1.93 TeV** at the $\boldsymbol{m(\tilde{\chi}_1^0) = 0.0, 0.5, 0.7,}$ **and 0.8 TeV**, respectively. Hence, the $\tilde{\chi}_1^0$ mass is excluded up to **1 TeV** in the gluino mass range of **1.50–1.80 TeV** for the gluino direct decay, and excluded up to **0.85 TeV** in the gluino mass range of **1.25–1.85 TeV** for the gluino one-step decay.

Table 10.4: Comarison between the best BDT SR in this analysis and the best SR in the previous study [45] for two signal mass point of (left) $(\tilde{g}, \tilde{\chi}_1^0) = (1700, 900)$ GeV and (right) $(\tilde{g}, \tilde{\chi}_1^\pm, \tilde{\chi}_1^0) = (1705, 1345, 985)$ GeV.

Signal Region	**SRD3**	**Previous study**
Total predicted bkg	24.2 ± 7.8	48.6 ± 7.2
Signal $(\tilde{g}, \tilde{\chi}_1^0) = (1700, 900)$ GeV	26.3 ± 6.3	25.8 ± 1.3
CLs	0.0034	0.13
Significance (σ_{excl})	2.70	1.10

Signal Region	**SRO2**	**Previous study**
Total predicted bkg	17.3 ± 4.1	5.1 ± 1.7
Signal $(\tilde{g}, \tilde{\chi}_1^\pm, \tilde{\chi}_1^0) = (1705, 1345, 985)$ GeV	20.7 ± 5.8	5.9 ± 0.7
CLs	0.21	0.49
Significance (σ_{excl})	0.77	0.014



(a) Gluino direct decay

(b) Gluino one-step decay

Figure 10.7: Observed exclusion limits at 95% confidence level based on CLs calculation for (a) the gluino direct decay model and (b) the gluino one-step decay model. A red solid line is the observed exclusion limit calculated by the nominal signal cross section, and the red break lines show the limits with the up and down signal cross-section variations. A blue break line is the expected exclusion limit. Its $\pm 1\sigma$ deviations are given by a yellow band. Observed exclusion limits in the previous study using the same data of 36.1fb^{-1}is represented by a blue shaded area [57].

11 Conclusion

A search for the gluino in final states with jets and large missing transverse momentum using 36fb^{-1} data recorded in 2015 and 2016 at the center-of-mass energy $\sqrt{s} = 13$ TeV is performed. In this search, new techniques of the quark/gluon separation variable and multivariate analysis are introduced, which picks up the four quark-jets emerging in the gluino decay chain in order to separate the signal and background further.

There is no significant excess implying the gluino. However, powerful exclusion limits on the gluino and $\tilde{\chi}_1^0$ masses are obtained for gluino direct and one-step decay models due to the new techniques. This search has a gain in the background rejection power in the signal region by factor 2 from the previous study for a signal mass point in the high gluino and $\tilde{\chi}_1^0$ mass region in the direct decay signal, and also exclusion sensitivity (σ_{excl}) is increased by factor 2.5 for the mass point. Especially, the high $\tilde{\chi}_1^0$ mass region in the high gluino mass range is more effectively searched than the previous analysis. The $\tilde{\chi}_1^0$ mass is excluded up to **1 TeV** in the gluino mass range of **1.50–1.80 TeV** for the gluino direct decay, and excluded up to **0.85 TeV** in the gluino mass range of **1.25–1.85 TeV** for the gluino one-step decay at 95% confidence level. This exclusion mass range indicates that the typical $\tilde{\chi}_1^0$ mass $\sim$ 1 TeV is almost ruled out in the gluino mass below 1.8 TeV, which is supposed to be reasonable mass considering the GUT and DM relic density.

Acknowledgements

I would like to express the deepest appreciation to professor Shoji Asai for supervising me for six years from my fourth year in the undergraduate school until my graducation of the Ph.D course concerning not only the ATLAS experiment but also ortho-positronium experiment that I performed during the master course. And I would like to also appreciate much to Professor Junichi Tanaka for supervising me at CERN as a analysis team coordinator of *the International Center for Elementary Particle Physics* ("ICEPP"). For the technical aspect, I am deeply grateful to Takashi Yamanaka, who was a research assistant professor until 2017 and a coordinator of the 0-lepton analysis for the SUSY strong production in the ATLAS ("SUSY 0-lepton analysis").

The following is a list of people who I appreciate concerning this research and my living at CERN for three years in the Ph.D course.

B. Peterson	Coordinating CERN analysis team for SUSY. He gave me a good discussion of the BDT analysis and helped me to scan pMSSM signal models for the BDT analysis.
T. Eifert, **L. Vivarelli**	The SUSY analysis coordinators in the ATLAS.
Y. Nakahama, **P. Jackson,** **C. Rogan,** **L. Duflot**	The analysis coordinators of the SUSY 0-lepton analysis.
N. Makovec	An analyser in the SUSY 0-lepton analysis team. He helped me to study non-collision background.
Y. Sano, **M. ROnzani**	Analysers in the SUSY 0-lepton analysis team. They participated in the analysis for Moriond 2017 together.
Y. Enari	Supervisor of my detector work for the LAr calorimeter.
W. Lampl	Technical supervisor of my detector work for the LAr calorimeter.
N. Berger	Technically advising my detector work for the LAr calorimeter.
M. Aleksa	The coordinator of the LAr calorimeter team.
B. Trocme, **E. Kuwertz,** **P. Strizenec**	Experts of data quality monitoring shift for the LAr calorimeter. They advised me in finding noisy cells of the LAr calorimeter.
H. Okawa	Supervisor of the quark/gluon tagging calibration.

F. Rubbo, **B. Nachman**	Group coordinators for the quark/gluon tagging in the ATLAS.
T. Saito	Advising on this dissertation for main analysis procedure and a captain of Tokyo-hiphopper relay team, leading the team win in 2015 and 2017.
T. Nobe	Advising on this dissertation for statistical treatment and job application, and telling me how to have a fun at CERN. He is a member of "enclosure".
Y. Okumura	Advising on the job application and showing me how hard we should work at CERN as one of hardest workers in the ICEPP staffs.
T. Kawamoto	One of ideal models to live at CERN over the whole life.
K. Terashi	Advising on the quark/gluon tagging calibration as a "God of Jet".
T. Sakamoto	Leading computing team in ICEPP for the ATLAS.
T. Mashimo, **T. Kishimoto**	Helping me to use the login.icepp server.
M. Ishino	ATLAS run coordinator.
R. Sawada	Helping me in the computing as an administrator of the ICEPP server at CERN.
K. Nakamura	Giving me good smell coffee and advising me how to maintenance my car.
T. Masubuchi	Advising me analysis procedure and providing comedic stories about his experience.
C. Kato	The leader of AJ(Atlas-Japan)-sangaku-bu (Mountain club).
Y. Yamaguchi	Taking me to fishing.
A. Miyazaki	Giving me some alcohol and homemade delicious food.
K. Yoshihara	Inviting me seafood party at his house and showing me his cute baby.
M. Morinaga	Buying my car NISSAN z350.
S. Chen	Funny broccoli-guy and an Iceland-mate.
Y. Minami	Discussing the 0-lepton analysis with me, who was a previous 0-lepton analyser in my institute.
A. Kobayashi	Maybe, she will get angry if she notice that here is no her name.
C. Kozakai	Colleague to deal with a dissertation together.
Y. Minegishi	A room-mate at CERN, who had borrow money of 700,000 yens from me for the house rent. He is an Iceland-mate.
S. Suzuki	Fat boy.
K. Onogi	Feeding me as a famous cooker in Japanese students at CERN.
T. Kunigo	A strange colleague. He argued hotly with H. Otono at his house when I was also there. It was a curious experience for me.

H. Otono A funny research associate at Kyushu University.

S. Shirabe Shirabe Shirabe. A attendant of H. Otono. He founded "Go Go Radio" in the "7. Radio+CM" of the album of "yes" created by enclosure. He is one of members in the Tokyo-hiphopper relay team in 2017.

K. Todome Advising me the method to estimate fake photon background in γ+jets samples.

D. Yamaguchi,
D. Sekihata, Participating in CERN summer-student program with me in 2013.
S. Ieki

M. Hirose Willing to agree our abrupt visiting Freiburg.

M. Saito Telling me some information in tracking and his strong emotion on *ramen*.

Y. Kano One of officemates at CERN and missionary of "t's up".

T. Yamazaki Disturbing me in working.

M. Yamatani Showing me how hard to live with a wife at CERN. He is one of the Tokyo-hiphopper members for the CERN relay race, and I expect his good run next year to win.

M. Tanaka One of the Tokyo-hiphopper members, and I expect his good run next year to win.

Y. Yang One of officemates, who knows Japanese well especially regarding to Japanese animation.

C. I-Huan One of officemates, who trains hard his muscle at CERN.

T. Nitta One of AJ-sangaku-bu members, who is a specialist of rock climbing.

K. Uno The 0-lepton analyser following me, who will sophisticated the analysis at the end of Run2. He showed me how to eat a lot of food and get weight.

R. Iguchi A rare kind of guy who doesn't get weight at CERN in contrast to K. Uno.

K. Uchida A strong guy in ICEPP.

K. Kimura A one of the craziest guys who I have ever met. He had traveled around the world as a backpacker.

A. Ishida Super drunker.

T. Yamaji A colleague in the same laboratory for a long time from fourth year in the undergraduate school.

K. Shu, X. Fan, Y. Seino, S. Kamioka
The talented juniors in our laboratory. K. Shu resembles a guy who had confined a girl in his house in Chiba. X. Fan is a strong guy to overwhelm his colleagues and juniors. Y. Seino looks like normal guy relative to the other juniors at first. S. Kamioka helped me to attend the final examination of ALPS course.

A Reconstruction and identification of the other particles

A.1 b-tagging

b-tagging is an algorithm to determine whether a jet is a jet originating from b-quark (b-jet) or not. The b-tagging is used in this analysis to make a top-quark enriched sample for the background estimation. It uses charged particle tracks to calculate discriminating variable between different jet flavors (light-flavor [u, d, or c-quarks or gluons], c-quarks, and b-quarks).There are three algorithms to provide complementary information:

- Impact parameter based variable algorithms: IP2D, IP3D
- Inclusive secondary vertex reconstruction algorithm: SV
- Decay chain multi-vertex reconstruction algorithm: JetFitter

The provided variables in them are combined by a multivariate analysis (MV2) which provides the best separation between different flavors.

Impact parameter based variables: IP2D, IP3D

The b-quark decays to b-hadrons having a long lifetime around $c\tau = 450\mu\text{m}$. Due to this feature, their distance from the primary vertex tends to have a large value. Here, the transverse and longitudinal impact parameters ("IPs") d_0 and z_0 are defined as a distance of closest approach in the $r - \phi$ plain and a distance along the beam axis between the track of the b-hadron and the primary vertex, respectively. The IP2D uses only the transverse IP, while the IP3D uses the both of them as inputs. Their output variable is a log-likelihood ratio ("LLR") between three jet-flavor hypothesis (b, c, light) which is calculated from probability density function of their input variables obtained from the MC simulation.

Secondary vertex: SV

A secondary vertex is a vertex of two tracks with a distance from the primary vertex. A secondary vertex finding algorithm is described in [69]. It finds a secondary vertex from all pairs of the track candidates in the jet. This algorithm ignores tracks originating from a decay of a long-lived particle (e.g. K_s and Λ), a photon conversion described in Section **??**, or a hadronic interaction with a

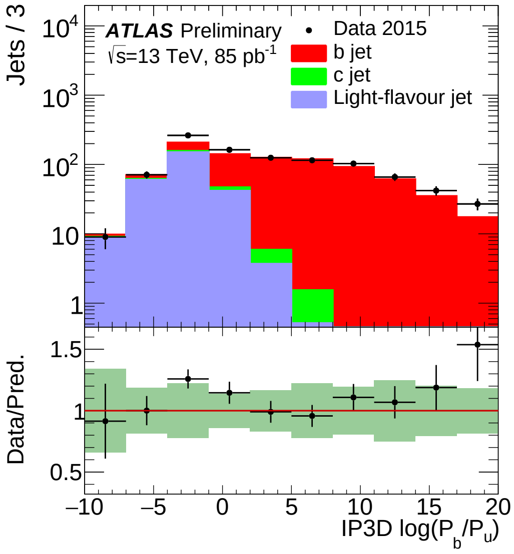


Figure A.1: The log likelihood ratio of the IP3D b-tagging algorithm applied to jets in a $t\bar{t}$ dominated sample [68]. The sample is required to have e and μ to enrich $t\bar{t} \rightarrow e\mu b\bar{b}$. Black points represent the data of 85pb^{-1}recorded in 2015. The bottom panel shows the ratio of the data to the simulation. The dark green shaded area is the total systematic uncertainty on the simulation, and the error bar of the points show the statistical errors of the data.

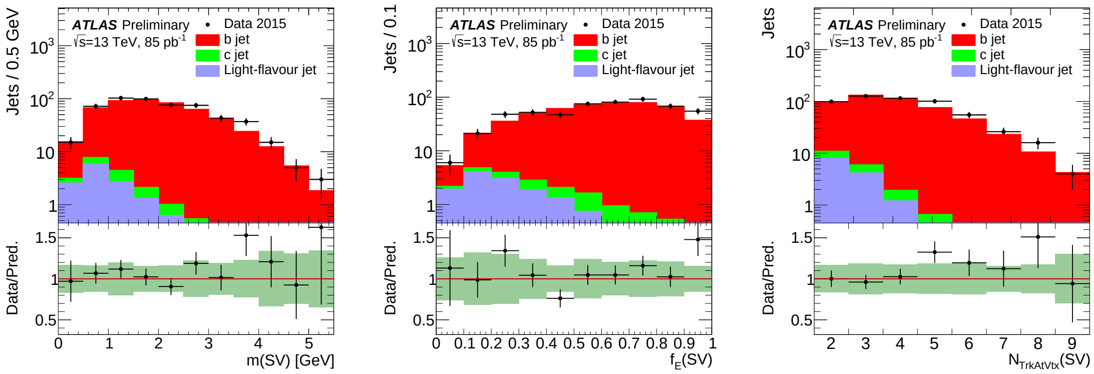
material. From the reconstructed secondary vertex with the two tracks, eight properties of it are retrieved and used for b-tagging in the MV2 [70]. Three of them are shown in Figures A.2.

Decay chain multi-vertex: JetFitter

The decay chain multi-vertex reconstruction is performed by JetFitter algorithm [71]. It reconstructs the full decay chain of the primary vertex→ b−hadron → c−hadron. From the full reconstruction of the decay chain with multi-vertex, eight properties are obtained and used for b-tagging in the MV2 [70]. Three of them are shown in Figures A.3.

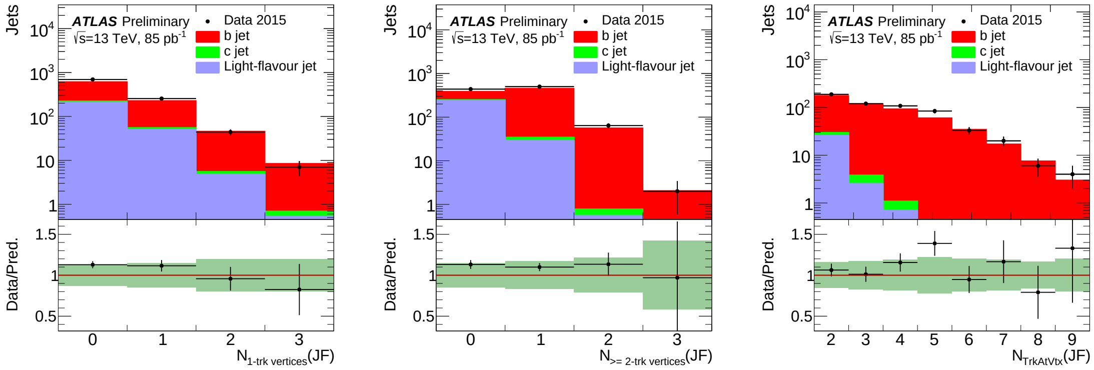
Final discriminating variable: MV2

The 24 input variables obtained from the three algorithms explained above are combined by using a boosted decision tree ("BDT") algorithm, which is one of the multivariate algorithms, in order to discriminate b-jets from light-flavor jets and c-jets. The BDT output used in this analysis is referred to as "MV2c10". It is defined as an output of the BDT trained with b-jets as signal and a mixture of 90% light-flavor jets and 10% c-jets as background. The output of the MV2c10 in $t\bar{t}$-enriched sample is shown in Figure A.4. The b-jet used in this analysis is defined by a jet with the MV2c10 output above the cut where the b-tagging efficiency is 77%.



(a) (b) (c)

Figure A.2: Properties obtained from the secondary vertex with two tracks in a jet of a $t\bar{t}$ dominated sample [68]: (a) The invariant mass of the two tracks, (b) an energy fraction of the two tracks to that of all tracks within the jet, and (c) a number of tracks associated with the secondary vertex. The sample is required to have e and μ to enrich $t\bar{t} \to e\mu b\bar{b}$. Black points represent the data of 85pb^{-1}recorded in 2015. The bottom panel shows the ratio of the data to the simulation. The dark green shaded area is the total systematic uncertainty on the simulation, and the error bar of the points show the statistical errors of the data.



(a) (b) (c)

Figure A.3: Properties obtained from the decay chain multi-vertex reconstruction in a jet of a $t\bar{t}$ dominated sample [68]: (a) the number of vertices with one track, (b) the number of vertices with at least two tracks, and (c) the number of tracks from vertices with at least two tracks. The sample is required to have e and μ to enrich $t\bar{t} \to e\mu b\bar{b}$. Black points represent the data of 85pb^{-1}recorded in 2015. The bottom panel shows the ratio of the data to the simulation. The dark green shaded area is the total systematic uncertainty on the simulation, and the error bar of the points show the statistical errors of the data.

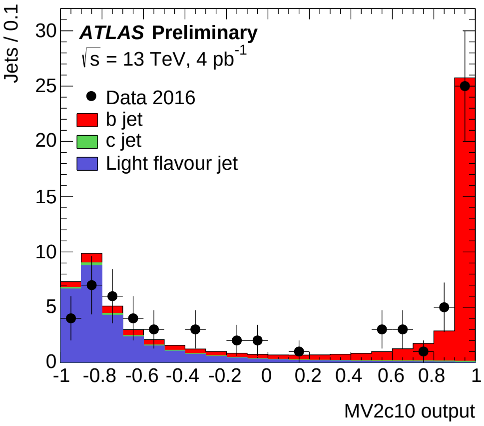


Figure A.4: The MV2c10 output in jets of a $t\bar{t}$ dominated sample [72]: The sample is required to have e and μ to enrich $t\bar{t} \to e\mu b\bar{b}$. Black points represent the data of 4pb^{-1}recorded in 2016. The bottom panel shows the ratio of the data to the simulation. The dark green shaded area is the total systematic uncertainty on the simulation, and the error bar of the points show the statistical errors of the data.

A.2 Electron and photon

Electron and photon are reconstructed from the tracks in the ID and cells in the EM calorimeters. Both of them are used in the background estimation to prepare specific-process enriched samples.

A.2.1 Electron

Electron reconstruction

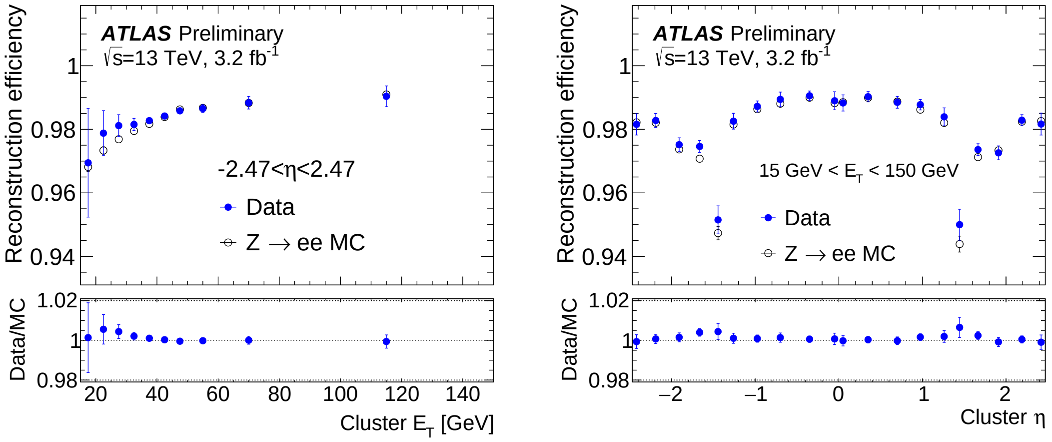
The electron reconstruction in the $|\eta| < 2.47$ has the following steps:

- **Seed-cluster reconstruction**
  A sliding window clustering algorithm described in [73] is used to find seed clusters in the electron reconstruction. In this algorithm, a cluster size of 3×5 in units of 0.025×0.025, corresponding to the granularity of the calorimeter middle layer, in ϕ-η is searched and that a cluster with total cluster transverse energy above 2.5 GeV is defined as a seed cluster.
- **Track reconstruction**
  This step is performed with the standard track reconstruction described in Section 3.1.1. In the standard track reconstruction, to extend the seed tracks, a pattern recognition and a track fit are made with a pion hypothesis for energy loss due to interactions with the detector material. However, it is not suitable for the electron track. Thus, if a seed track cannot be extended to a full track under the pion hypothesis but falls into one of the EM cluster regions, a second attempt, in which the pattern recognition with an electron hypothesis allowing for larger energy loss, is performed. In the track fit of the track candidates, a special fit with the electron hypothesis is also performed if the standard fit with the pion hypothesis is failed. These additional attempts can improve the performance of electrons.
- **Track fit for electrons**
  The tracks are checked if they are loosely matched to the EM clusters or not, by using the distance in the ϕ-η plane between the extrapolated track and the cluster. The track loosely associated with the cluster is refit using an optimized Gaussian Sum Filter ("GSF") [74].
- **Electron candidate reconstruction**
  At last, the similar matching is repeated with stricter conditions after the GSF refit.

The efficiency of this reconstruction procedure is calculated using $Z \to ee$ events in both of the data and the MC (Figures A.5). The ratio between the efficiencies of them is used as a correction factor called as "Scale Factor" (SF) in order to correct the MC.

Electron identification

Electron identification discriminates signal-like electrons from other objects (background) of hadronic jets or converted photons. The identification uses many kinds of properties: properties of the electron cluster and the tracks, quantities related to track-cluster matching, the calorimeter shower



(a)

(b)

Figure A.5: Electron reconstruction efficiencies in $Z \longrightarrow ee$ events of the 2015 data and the MC as a function of (a) the cluster E_T in the full η range and (b) the η in the E_T range between 15 GeV and 150 GeV [75].

shape, and information from the TRT, which is designed to distinguish electrons from hadrons by a transition radiation signature described in Section 2.2.3. The algorithm for the identification is based on the likelihood (LH) method, in which the probability density function (PDF) of these properties for the signal (electron) and the background are used. The discriminant of the LH algorithm is

$$d_\mathcal{L} = \frac{\mathcal{L}_S}{\mathcal{L}_S + \mathcal{L}_B}, \qquad \mathcal{L}_{S(B)}(\vec{x}) = \prod_{i=1}^{n} P_{s(b),i}(x_i), \tag{A.1}$$

where $\vec{x}$ is the vector of the discriminating variables mentioned above and $P_{s,i}(x_i)$ is the value of the signal (or background) PDF of the i^th variable evaluated at x_i. The three kinds of selections using this LH discriminant are prepared with different efficiencies of the electron identification and the background rejection powers referred to as "Loose", "Medium", and "Tight". The efficiency of each selection is 97, 95, and 91 % at $E_\mathrm{T} = 70$ GeV, respectively (Figures A.6).

Electron isolation

Electron isolation is in order to further discriminate the signal-like electron from the background by using the energy deposit around the electron candidate. There are two discriminating variables for the isolation:

- **A calorimetric isolation energy $E_\mathrm{T}^\mathrm{cone0.2}$**: The sum of transverse energies of topological clusters at the EM scale within a cone of $\Delta R = 0.2$ around the electron candidate.
- **A track isolation $p_\mathrm{T}^\mathrm{varcone0.2}$**: The sum of transverse momenta of the tracks within a cone of $\Delta R = \min(0.2, 10\ \mathrm{GeV}/E_\mathrm{T})$ around the candidate electron track. The tracks need to originate

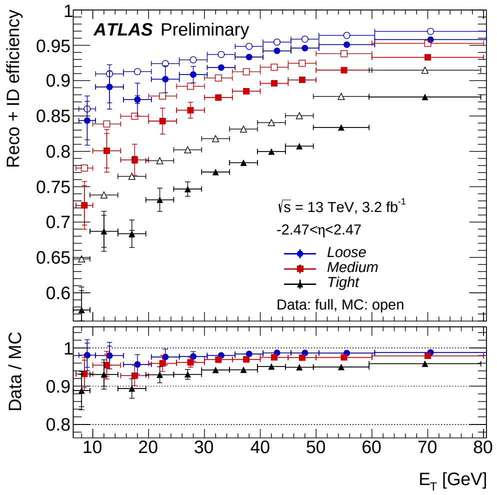


Figure A.6: Efficiency of electron reconstruction and identification (ID) with different ID selection as a function of E_T in $Z \to ee$ simulation [75].

from the primary vertex of the hard collision and satisfy the quality requirements of $E_\mathrm{T} > 1$ GeV, a number of hits requirements, and $|z_0 \sin\theta| < 3\mu$m.

The isolation cut is defined by selection on $E_\mathrm{T}^\mathrm{cone0.2}/E_\mathrm{T}$ and $E_\mathrm{T}^\mathrm{varcone0.2}/E_\mathrm{T}$. In this analysis, "GradientLoose" selection is used, in which the isolation efficiency is fixed to $0.057\% \times E_\mathrm{T} + 95.57\%$ by varying the cut values.

Requirements for electron in this analysis

In this analysis, the electron identification is used to veto electrons to search signal and require electrons in the background estimation. In each case, a slightly different requirement is used. For the electron veto,

- $p_\mathrm{T} > 7$ GeV
- $|\eta| < 2.47$
- Loose ID
- No isolation

are required. For the electron tagging, the ID requirement is tightened to Tight selection and the GradientLoose isolation is required in addition to the above.

A.2.2 Photon

Photon reconstruction

The photon reconstruction procedure is the same as in electron reconstruction (Appendix A.2.1) until the track fit by the GSF because photon and electron has very similar signatures in the EM calorimeters.

After the GSF fit, for the photon reconstruction, "conversion vertex" candidates are reconstructed. The conversion vertex is a point where paired electrons with opposite charges are produced from a photon. Thus, these vertex candidates are reconstructed from two opposite charged tracks. The two tracks need to satisfy geometrical requirements, number-of-track hits requirements, and a good fit quality in a fit of the two tracks with a constraint that the tracks are parallel at the conversion vertex. The paired tracks with a conversion vertex are loosely matched to the EM clusters.

After matchings of the EM clusters to the electron track candidates and to the conversion vertex, the EM clusters are classified into the electron, the unconverted photon, or the converted photon. A cluster without a matched electron track nor a matched conversion vertex candidate has been classified into an unconverted photon. A cluster matched to an electron track or a conversion vertex only is considered as an electron or a converted photon, respectively.

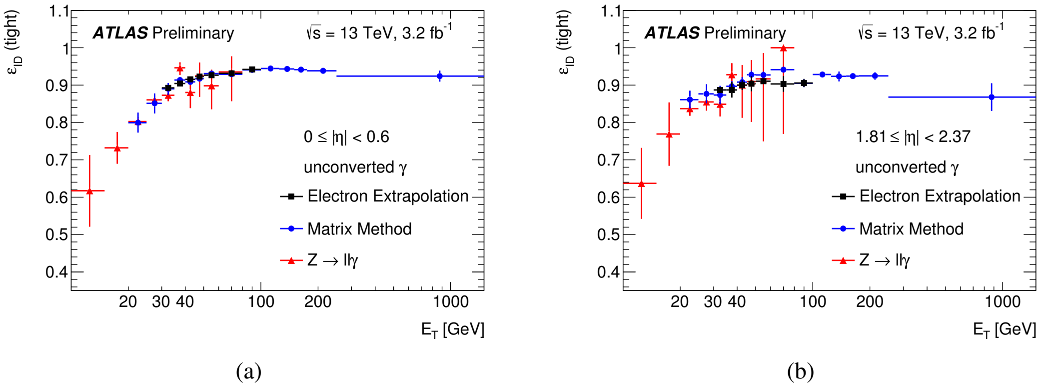
Photon identification

The photon identification is aiming to distinguish a prompt photon, which is produced at the interaction point, from background photons, which is mainly caused by $\pi^0 \rightarrow \gamma\gamma$ in a hadron decay. This identification uses cuts on several discriminating variables related to the lateral and longitudinal shower development in the EM calorimeters and the shower leakage to the hadronic calorimeters. The prompt photon typically has a narrower shower and smaller leakage to the hadronic calorimeters than the hadronic background. In the background photons of the $\pi^0 \rightarrow \gamma\gamma$ decay, the two photons are characterized by the two separate local energy maxima in the $\phi - \eta$ plane of the shower. This feature is captured by the first layer of the EM calorimeters having finely segmented strips. In this analysis, "Tight" selection is used, in which the photon identification efficiency is 88-92% for $E_{\mathrm{T}} > 100$ GeV unconverted photons as shown in Figures A.7.

Photon isolation

Electron isolation provides an additional background suppression by using a feature that background photons are surrounded by a sizable hadronic activity, that is a poor isolation. The isolation selection is defined by two variables in the following, which are the same as in the electron isolation (A.2.1) except the cone size (ΔR) definition.

- **A calorimetric isolation energy with $\Delta R = 0.4$ $E_{\mathrm{T}}^{\mathrm{cone0.4}}$**: The sum of transverse energies of topological clusters at the EM scale within a cone of $\Delta R = 0.4$ around the photon candidate.



(a) (b)

Figure A.7: Photon identification efficiencies by tight selection obtained from three different data-driven measurements using 3.2fb^{-1}data in 2015 for unconverted photons in the η range between (a) $0 < |\eta| < 0.6$ and (b) $1.81 < |\eta| < 2.37$ [76]. Each method of the data-driven measurements is described in [76].

- **A track isolation with $\Delta R = 0.2$ $p_{\mathrm{T}}^{\mathrm{cone0.2}}$**: The sum of transverse momenta of the tracks within a cone of $\Delta R = 0.2$ around the candidate photon track.

In this analysis, a set of isolation cuts called as "FixedCutTight" is used, which is given as

$$E_{\mathrm{T}}^{\mathrm{cone0.4}} < 0.022 p_{\mathrm{T},\gamma} + 2.45 \quad \& \quad p_{\mathrm{T}}^{\mathrm{cone0.2}}/p_{\mathrm{T},\gamma} < 0.05. \tag{A.2}$$

Requirements for photon in this analysis

In this analysis, photons are used only in the background estimation to prepare γ+jets enriched sample. In this sample, one photon satisfying the following selection is required;

- $p_{\mathrm{T}} > 25$ GeV
- $|\eta| < 2.37$
- Tight ID
- FixedCutTight isolation

A.2.3 Energy calibration and resolution correction for electron and photon

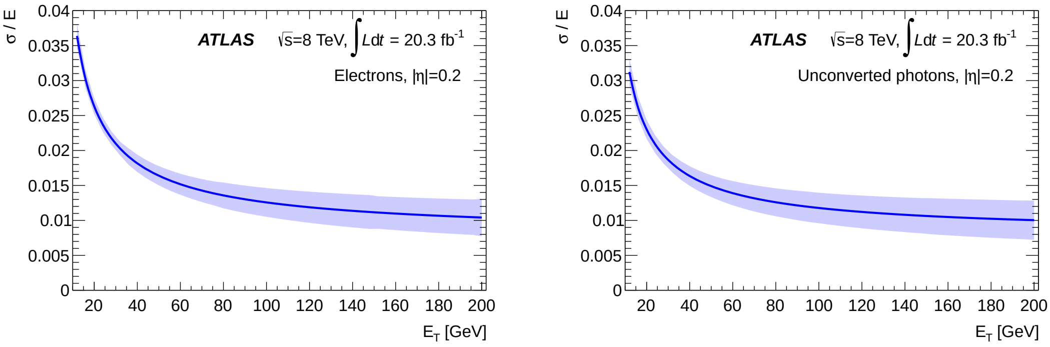
The energy calibration is performed in three main steps:

1. Data-driven corrections in order to equalize the response of the different longitudinal layers in the EM calorimeter in data with respect to the simulation. (These corrections are applied only to data.)

2. MC-based calibration applied to both of data and MC. This calibration is aiming to correct the energy deposit in front of the calorimeter (typically a few–20% of the electron energy for 100 GeV electrons.) and outside of the cluster (~ 5%), and also correct the variation of the energy response depending on the particle's incident position on the calorimeter.

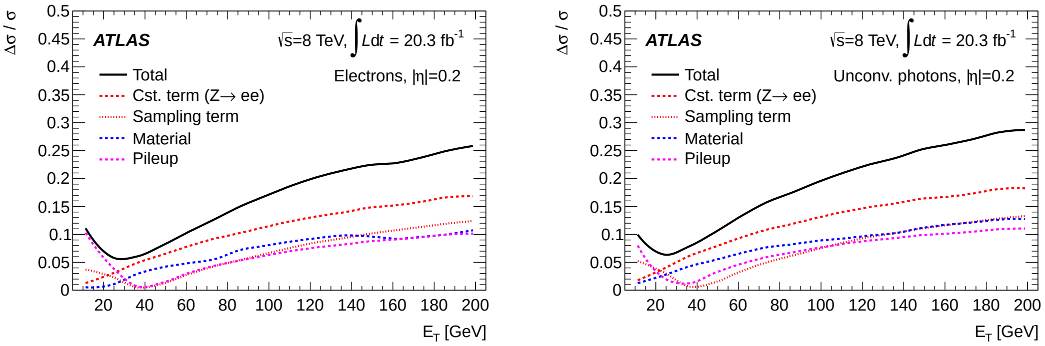
3. Data-driven (*In-situ*) correction on the energy scale applied only to data and correction on the resolution applied only to MC. These corrections are implemented as two constants α_i and c_i' given as
$$E_i^{\text{data}} = E_i i^{\text{MC}} (1 + \alpha_i) \tag{A.3}$$
for the energy scale correction and
$$\left(\frac{\sigma(E)}{E_i^{\text{data}}} \right) = \left(\frac{\sigma(E)}{E_i^{\text{MC}}} \right) + c_i' \tag{A.4}$$
for the energy resolution correction. The $E^{\text{data/MC}}$ is the energy scale in data and MC, respectively, and the label i indicates a given η range. These constants are obtained from $Z \to ee$ events in the data and MC. The constant c_i' is the dominant resolution uncertainty in the E_{T} above 30 GeV as shown in Figures A.8(c) and A.8(d).

The energy resolution and its uncertainties including one coming from the calibration are shown in Figures A.8.



(a) Energy resolution of electrons

(b) Energy resolution of photons



(c) Resolution uncertainties for electrons

(d) Resolution uncertainties for photons

Figure A.8: Energy resolution of (a) electrons and (b) photons, and its resolution uncertainties for (c) electrons and (d) photons in Run1 [77]. In the figures of the resolution uncertainties, the cst. term is the uncertainty on the c_i' obtained in the *in situ* calibration (Eq. A.4), and the sampling term is the uncertainty on the a in Eq. 2.7.

A.3 Muon

The muon has no energy deposit in the calorimeters because it is a minimum ionizing particle and it cannot make bremsstrahlung radiations due to its much heavier mass than the electron. In the ATLAS detector, the muon spectrometers (MS) are located outside of the calorimeters and immersed in a ~ 0.5T magnetic field. The muon can be reconstructed as a track at low misidentification rate by combining the information in the ID and the MS, and its momentum can be measured precisely due to the large size of the muon spectrometers.

Muon reconstruction

In the muon reconstruction, the tracks reconstructed in the ID and the MS are combined. The track in the ID is reconstructed by the procedure described in Section 3.1.1. The track in the MS is reconstructed in the following procedures:

1. In each MDT chamber and nearby trigger chambers, a segment is reconstructed, which is a set of hits aligned on a trajectory in the bending plane. This formation is performed by a Hough transform method [78].
2. Muon track candidates are built by fitting hits of the segments in different layers. In this fit, the segment generated in the middle layer is used as a seed at first, and it extrapolated to outer and inner layers by combining the other segments. At least two matching segments in the track candidate are required for the track candidate except in the transition region between the barrel and the endcap.
3. The hits in the track candidate are fitted by a global χ^2 fit. The track candidate is accepted as an MS track if the fit quality satisfies selection criteria.

Then, the reconstructed MS track is combined with the ID track. A combined track is formed by a global refit with the hits from both tracks. At first, muon tracks are reconstructed by outside-in pattern recognition, in which an MS track is extrapolated inward and matched to an ID track. Then, inside-out pattern recognition is performed, in which an ID track is extrapolated to an MS track. The combined tracks are treated as muon candidates. In addition to the combined tracks, the MS tracks whose direction is matched to the interaction point are treated as muon candidates in order to recover the acceptance in the range of $2.5 < |\eta| < 2.7$ where the ID does not exist.

Muon identification

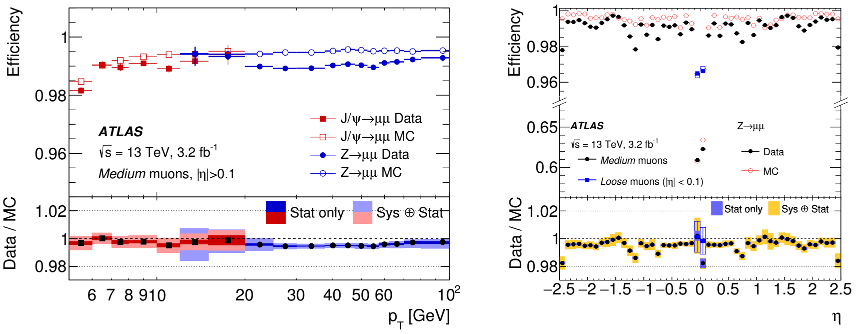
Muon identification is performed in order to suppress background mainly caused by pion and kaon decays. The background track of the in-flight decays of charged hadrons in the ID has a distinctive "kink" topology, which makes the poor fit quality. In addition, it often has a disagreement in measured momenta between the ID track and the MS track. Such features are taken into account in the identification by using the following properties in addition to numbers of hist and holes in the MDT:

- ***q*/*p* significance**: An absolute value of the difference between the ratios of the muon charge and momentum in the ID and the MS divided by the quadrature-sum of the corresponding uncertainties.

- **ρ'**: An absolute value of the difference between the transverse momenta measured in the ID and the MS divided by the p_{T} of the combined track.

- **Normalized χ^2** of the combined track fit.

In this analysis, "Medium" identification selection is used, which minimizes the systematic uncertainties of the muon reconstruction and calibration. The reconstruction and identification efficiency of the medium muon is shown in Figures A.9. The efficiency is larger than 99% for $p_{\mathrm{T}} > 6$ GeV muons.



(a) (b)

Figure A.9: Reconstruction and identification efficiencies for medium muon in $Z \rightarrow \mu\mu$ and $J/\psi \rightarrow \mu\mu$ events of the 2015 data and the MC as a function of (a) η in the range of $p_{\mathrm{T}} > 10$ GeV (including the efficiency of Loose identified muons in $|\eta| < 0.1$) and (b) p_{T} in the range of $0.1 < |\eta| < 2.5$ [79].

Muon isolation

Isolation selection for muons is also prepared like electrons and photons. It is a powerful tool to reject the background muon caused by the hadron decay. The isolation selection is defined by two variables in the following, which are the same as in the electron isolation (A.2.1) except the cone size ΔR for a calorimetric isolation.

- **A calorimetric isolation energy with $\Delta R = 0.3$ $E_{\mathrm{T}}^{\mathrm{cone0.3}}$**: The sum of transverse energies of topological clusters at the EM scale within a cone of $\Delta R = 0.3$ around the muon candidate.

- **A track isolation $p_{\mathrm{T}}^{\mathrm{varcone0.2}}$**: The sum of transverse momenta of the tracks within a cone of $\Delta R = \min(0.2, 10\ \mathrm{GeV}/E_{\mathrm{T}})$ around the candidate muon track.

In this analysis, an isolation selection criteria referred to as "GradientLoose" is used, in which the cuts on the two isolation variables are defined to realize the ≥ 95(90) % efficiency at muon $p_{\mathrm{T}} = 25(60)$ GeV. The efficiency is shown in Figure A.10.

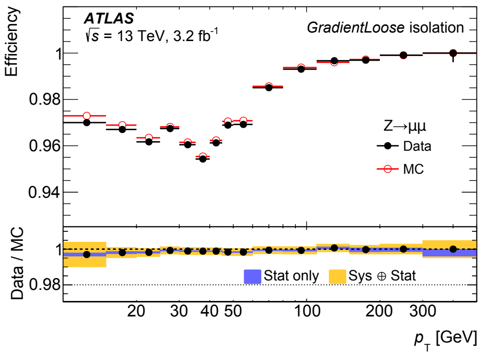


Figure A.10: Efficiency of "GradientLoose" isolation for muons as a function of p_{T} in $Z \to \mu\mu$ events [79].

Muon momentum calibration and resolution

In order to have a good agreement between data and MC, a set of corrections is applied to the muon momentum in the MC. It is performed as corrections on each momentum of the ID track and the MS track. The corrected transverse momenta $p_{\mathrm{T}}^{\mathrm{Cor,Det}}$ (Det=ID, MS) is given as,

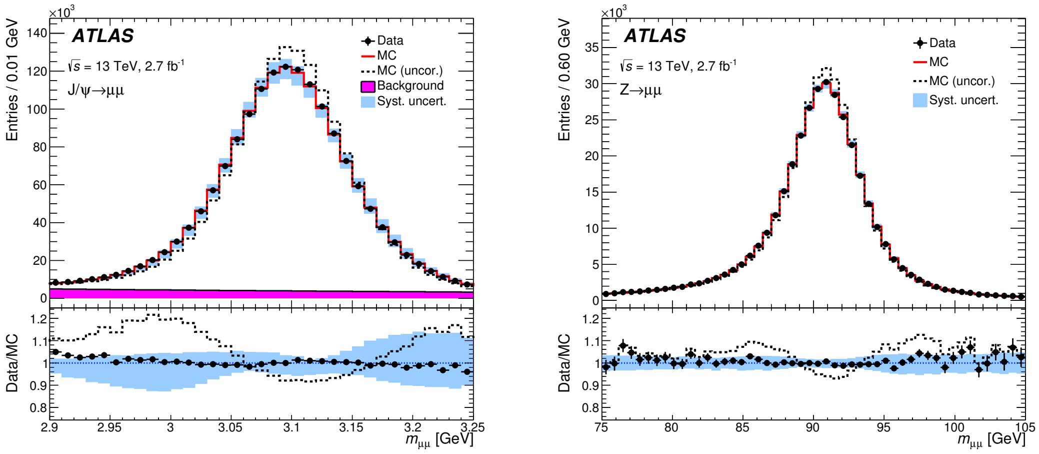
$$p_{\mathrm{T}}^{\mathrm{Cor,Det}} = \frac{p_{\mathrm{T}}^{\mathrm{MC,Det}} + \sum_{i=0}^{1} s_n^{\mathrm{Det}}(\eta,\phi)\left(p_{\mathrm{T}}^{\mathrm{MC,Det}}\right)^n}{1 + \sum_{m=0}^{2} \Delta r_m^{\mathrm{Det}}(\eta,\phi)\left(p_{\mathrm{T}}^{\mathrm{MC,Det}}\right)^{m-1} g_m}, \tag{A.5}$$

where $p_{\mathrm{T}}^{\mathrm{MC,Det}}$ is uncorrected transverse momentum in the MC, and g_m are Gaussian distributions with mean=0 and width=1. The terms of $s_n^{\mathrm{Det}}(\eta,\phi)$ and $\Delta r_m^{\mathrm{Det}}(\eta,\phi)$ are the momentum scale corrections and the momentum resolution smearing corrections in each (η,ϕ) region, respectively. The correction factors of s_n^{Det} and $\Delta r_m^{\mathrm{Det}}$ are determined from data using a binned maximum-likelihood fit with templates of the invariant mass distribution for $J/\psi \to \mu\mu$ and $Z \to \mu\mu$ in the data and MC. The invariant mass distributions are shown in Figures A.11, in which the mass distribution of the MC after the corrections has good agreement with that of the data.

Requirements for muon in this analysis

In this analysis, the muon identification is used to veto muons to search signal and require muons in the background estimation. In each case, a slightly different requirement is used. For the muon veto,

- $p_{\mathrm{T}} > 7$ GeV
- $|\eta| < 2.7$
- Medium ID
- No isolation



(a) $J/\psi \rightarrow \mu\mu$ (b) $Z \rightarrow \mu\mu$

Figure A.11: Invariant mass distributions for (a) $J/\psi \rightarrow \mu\mu$ and (b) $Z \rightarrow \mu\mu$ using 2.7fb^{-1}data in 2015 [79]. Black break lines and red solid lines show the distributions of the MC before and after the muon momentum scale and resolution corrections, respectively.

are required. For the muon tagging,

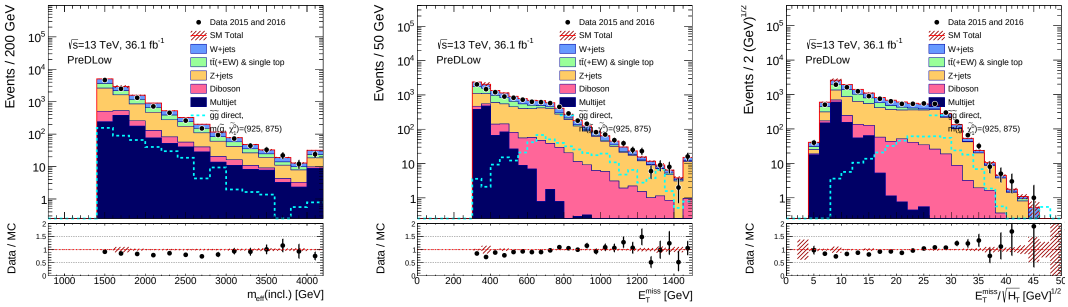
- GradientLoose isolation
- $|d_0| < 3\text{mm}$
- $|z_0 \sin\theta| < 0.5\text{mm}$

are required additionally.

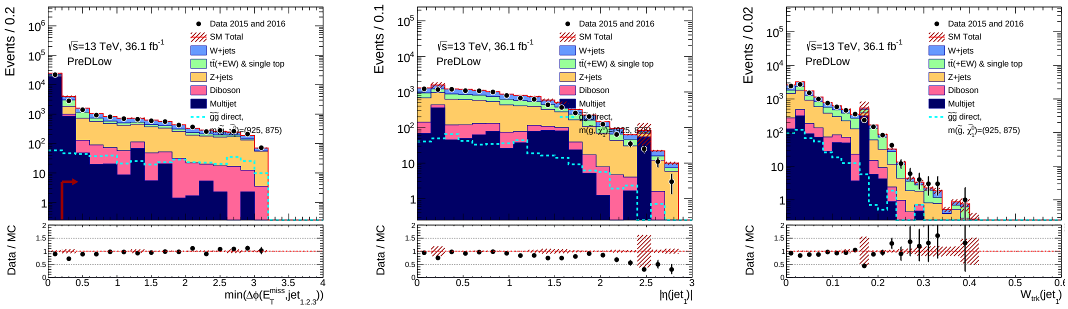
B Distributions of the discriminating variables

B.1 Distribution in the preselection

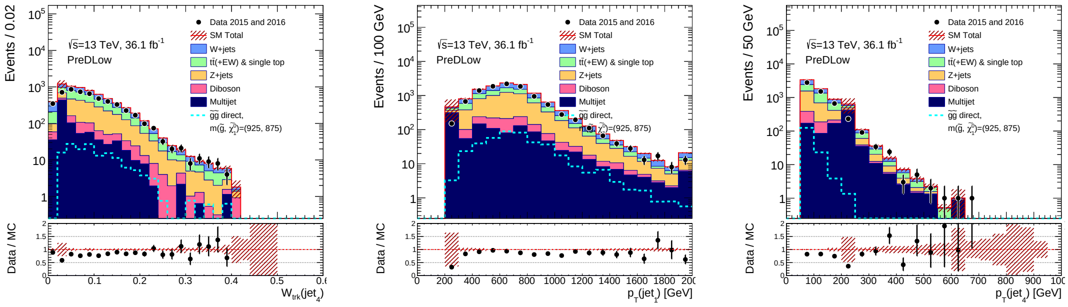
Here, the distributions of the discriminating variables in the preselection (loose SR-like selection) are shown in Figures B.1–B.3. The distributions in PreDHigh is shown in Section 7.3.



(a) $m_{\mathrm{eff}}(\mathrm{inc})$ (b) $E_{\mathrm{T}}^{\mathrm{miss}}$ (c) $E_{\mathrm{T}}^{\mathrm{miss}}/\sqrt{H_{\mathrm{T}}}$

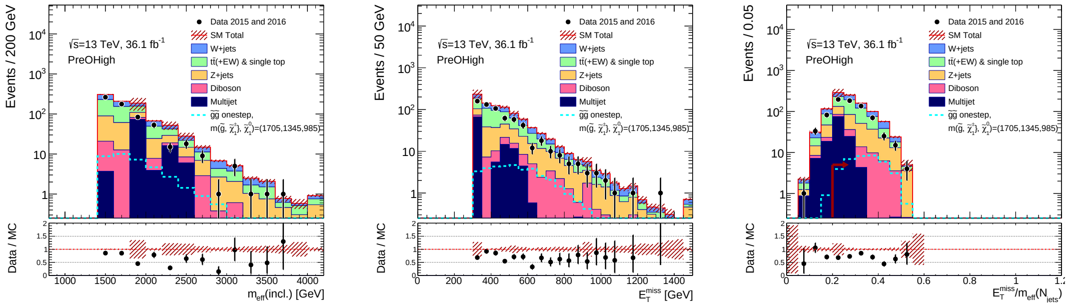


(d) $\Delta\phi(\mathrm{j}_{1,2,(3)}, \vec{E}_{\mathrm{T}}^{\mathrm{miss}})_{\mathrm{min}}$ (e) 1st jet $|\eta|$ (f) 1st jet W_{trk}

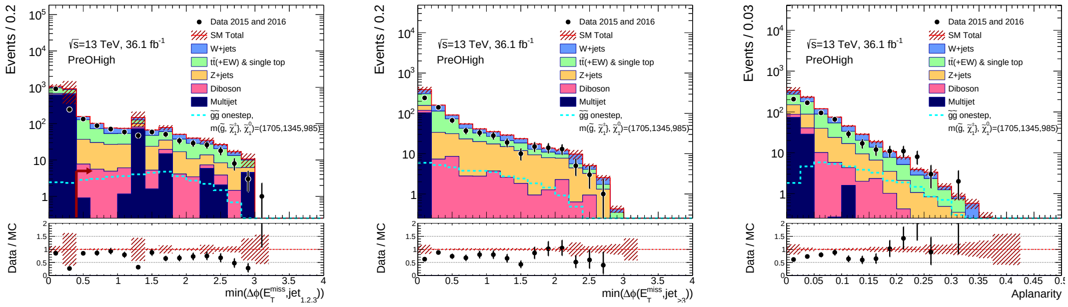


(g) 4th jet W_{trk} (h) 1st jet p_{T} (i) 4th jet p_{T}

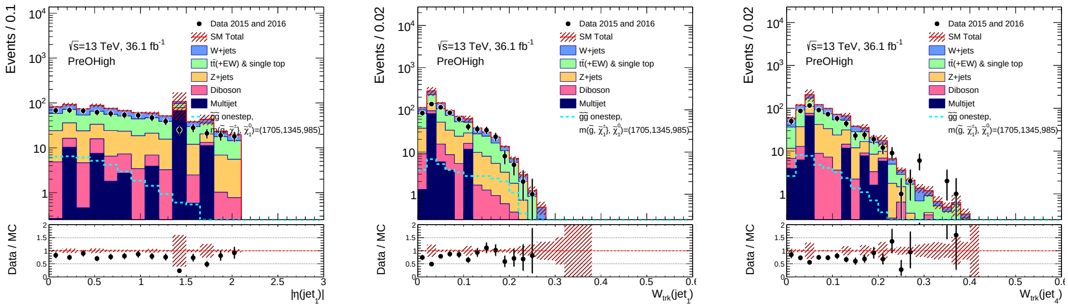
Figure B.1: Distributions of $m_{\mathrm{eff}}(\mathrm{inc})$, $E_{\mathrm{T}}^{\mathrm{miss}}$, $E_{\mathrm{T}}^{\mathrm{miss}}/\sqrt{H_{\mathrm{T}}}$, $\Delta\phi(\mathrm{j}_{1,2,(3)}, \vec{E}_{\mathrm{T}}^{\mathrm{miss}})_{\mathrm{min}}$, 1st jet $|\eta|$, 1st jet W_{trk}, 4th jet W_{trk}, 1st jet p_{T}, and 4th jet p_{T} in SR of **PreDLow**. Black points show the $36.1\mathrm{fb}^{-1}$data and filled histograms show the SM background MCs normalized by cross-section. A large bin of the multi-jet is caused by an event with an extreme high event weight reflecting the skimming weight in the MC event generation. This is unphysical value to be ignored.



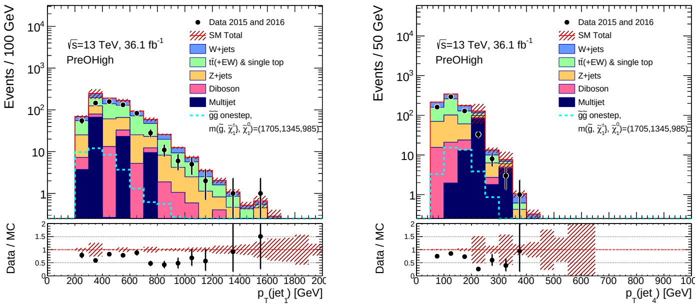
(a) $m_{\rm eff}$(inc) (b) $E_{\rm T}^{\rm miss}$ (c) $E_{\rm T}^{\rm miss}/m_{\rm eff}$(4j)



(d) $\Delta\phi(\mathrm{j}_{1,2,(3)}, \vec{E}_{\rm T}^{\rm miss})_{\rm min}$ (e) $\Delta\phi(\mathrm{j}_{i>3}, \vec{E}_{\rm T}^{\rm miss})_{\rm min}$ (f) Aplanarity

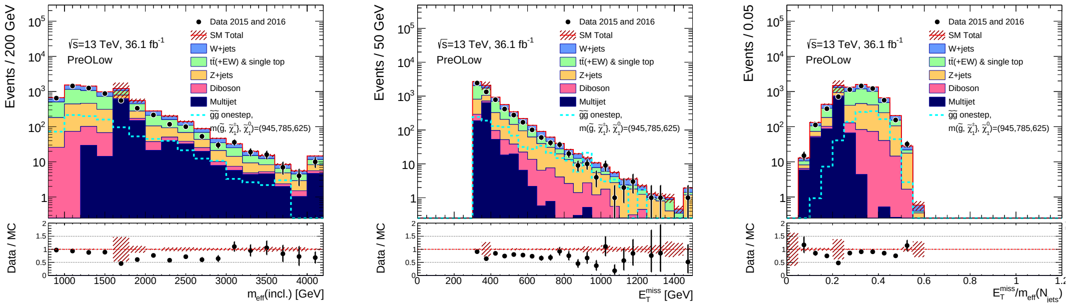


(g) 1st jet $|\eta|$ (h) 1st jet $W_{\rm trk}$ (i) 4th jet $W_{\rm trk}$

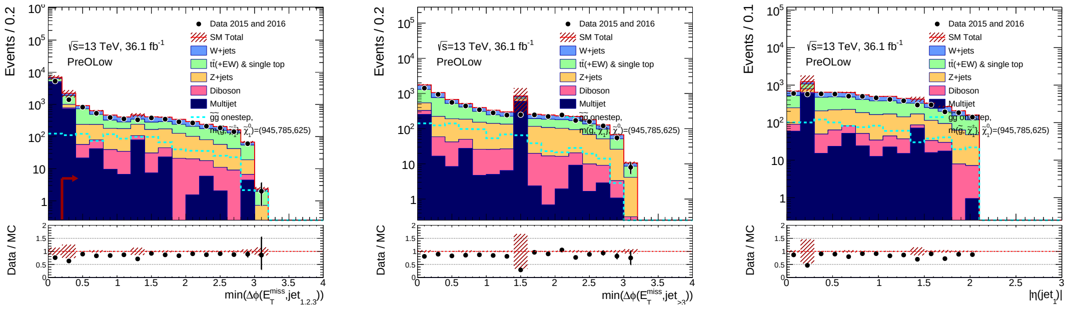


(j) 1st jet $p_{\rm T}$ (k) 4th jet $p_{\rm T}$

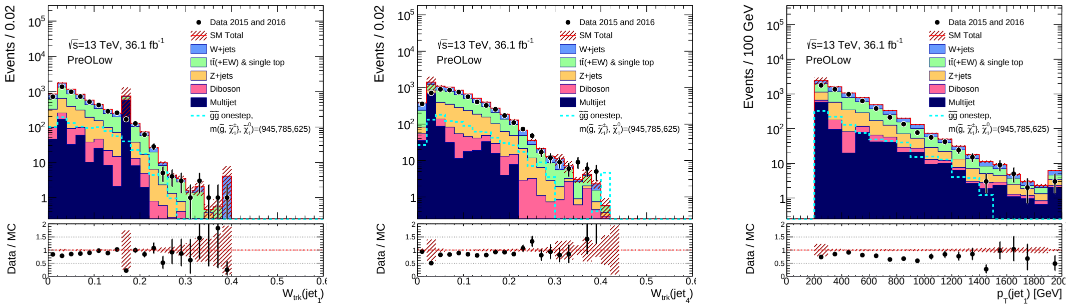
Figure B.2: Distributions of $m_{\rm eff}$(inc), $E_{\rm T}^{\rm miss}$, $E_{\rm T}^{\rm miss}/m_{\rm eff}$(4j), $\Delta\phi(\mathrm{j}_{1,2,(3)}, \vec{E}_{\rm T}^{\rm miss})_{\rm min}$, $\Delta\phi(\mathrm{j}_{i>3}, \vec{E}_{\rm T}^{\rm miss})_{\rm min}$, Aplanarity, 1st jet $|\eta|$, 1st jet $W_{\rm trk}$, 4th jet $W_{\rm trk}$, 1st jet $p_{\rm T}$, and 4th jet $p_{\rm T}$ in SR of **PreOHigh**. Black points show the $36.1\,{\rm fb}^{-1}$ data and filled histograms show the SM background MCs normalized by cross-section. A large bin of the multi-jet is caused by an event with an extreme high event weight reflecting the skimming weight in the MC event generation. This is unphysical value to be ignored.



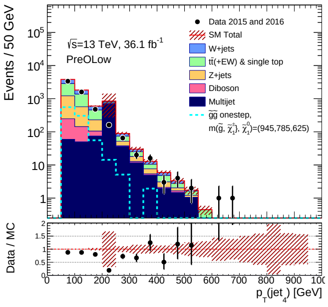
(a) $m_{\mathrm{eff}}(\mathrm{inc})$ (b) $E_{\mathrm{T}}^{\mathrm{miss}}$ (c) $E_{\mathrm{T}}^{\mathrm{miss}}/m_{\mathrm{eff}}(4\mathrm{j})$



(d) $\Delta\phi(\mathrm{j}_{1,2,(3)}, \vec{E}_{\mathrm{T}}^{\mathrm{miss}})_{\mathrm{min}}$ (e) $\Delta\phi(\mathrm{j}_{i>3}, \vec{E}_{\mathrm{T}}^{\mathrm{miss}})_{\mathrm{min}}$ (f) 1st jet $|\eta|$



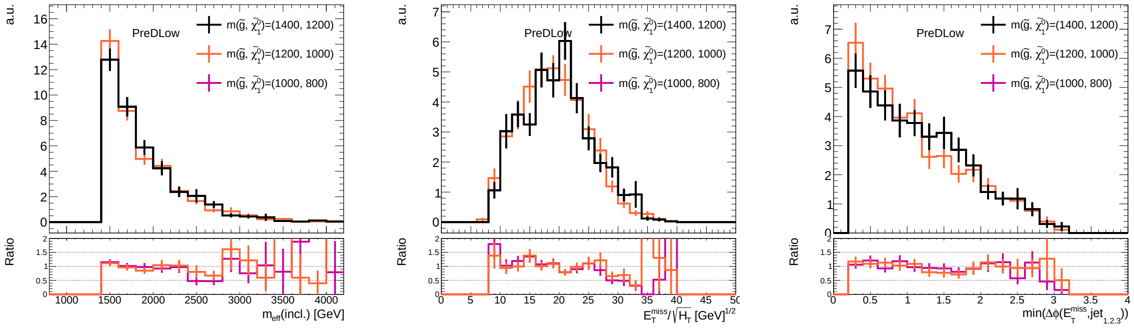
(g) 1st jet W_{trk} (h) 4th jet W_{trk} (i) 1st jet p_{T}



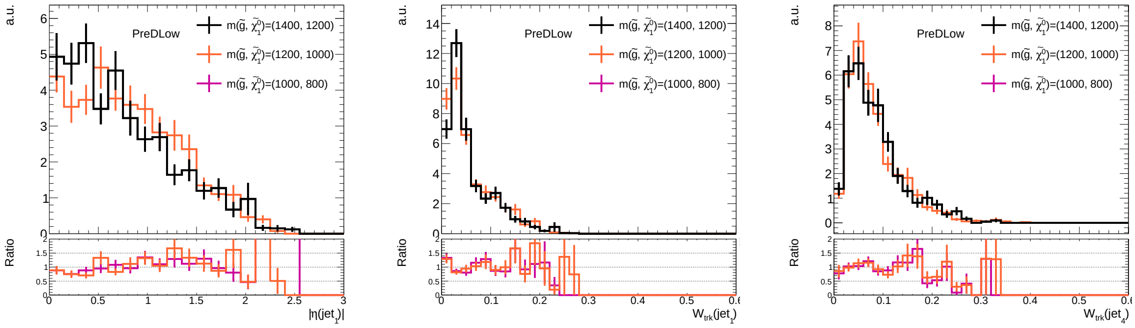
(j) 4th jet p_{T}

Figure B.3: Distributions of $m_{\mathrm{eff}}(\mathrm{inc})$, $E_{\mathrm{T}}^{\mathrm{miss}}$, $E_{\mathrm{T}}^{\mathrm{miss}}/m_{\mathrm{eff}}(4\mathrm{j})$, $\Delta\phi(\mathrm{j}_{1,2,(3)}, \vec{E}_{\mathrm{T}}^{\mathrm{miss}})_{\mathrm{min}}$, $\Delta\phi(\mathrm{j}_{i>3}, \vec{E}_{\mathrm{T}}^{\mathrm{miss}})_{\mathrm{min}}$, 1st jet $|\eta|$, 1st jet W_{trk}, 4th jet W_{trk}, 1st jet p_{T}, and 4th jet p_{T} in SR of **PreOLow**. Black points show the 36.1fb^{-1}data and filled histograms show the SM background MCs normalized by cross-section. A large bin of the multi-jet is caused by an event with an extreme high event weight reflecting the skimming weight in the MC event generation. This is unphysical value to be ignored.

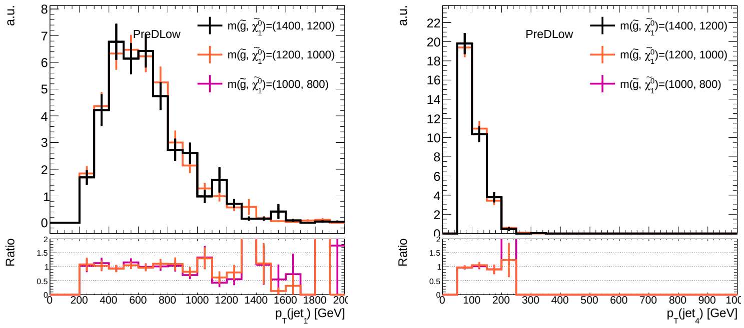
Here, the distributions of the discriminating variables in the preselection (loose SR-like selection) for the same $\Delta M(\tilde{g}, \tilde{\chi}_1^0)$ signals are shown in Figures B.4–B.6.



(a) $m_{\rm eff}$(inc) (b) $E_{\rm T}^{\rm miss}/\sqrt{H_{\rm T}}$ (c) $\Delta\phi({\rm j}_{1,2,(3)}, \vec{E}_{\rm T}^{\rm miss})_{\rm min}$

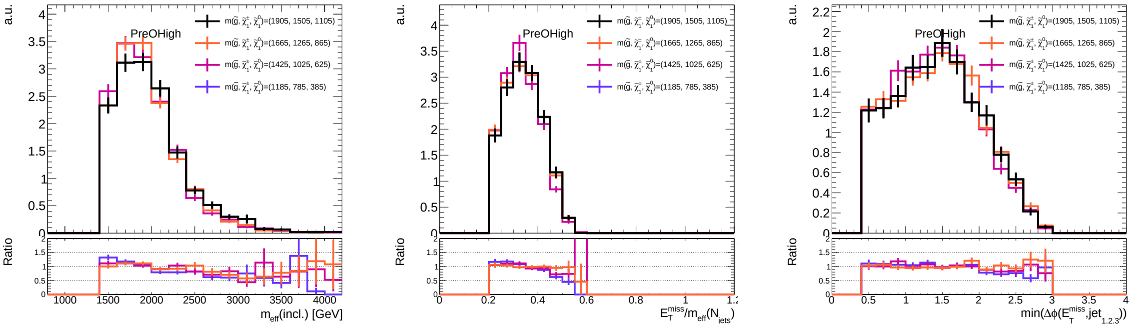


(d) 1st jet $|\eta|$ (e) 1st jet $W_{\rm trk}$ (f) 4th jet $W_{\rm trk}$

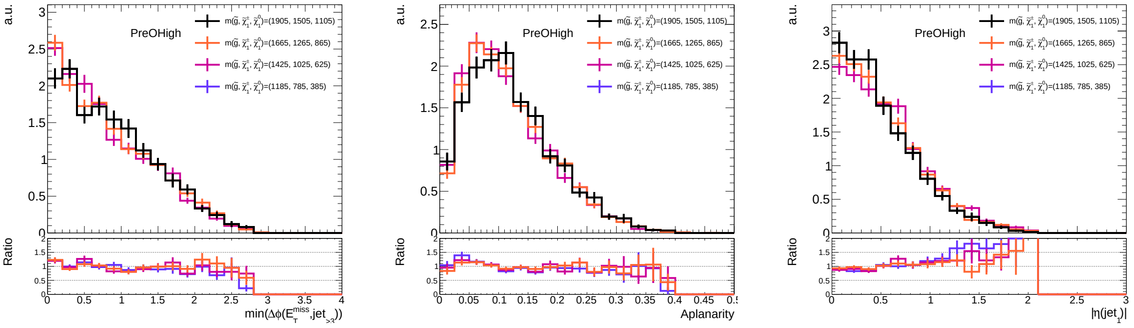


(g) 1st jet $p_{\rm T}$ (h) 4th jet $p_{\rm T}$

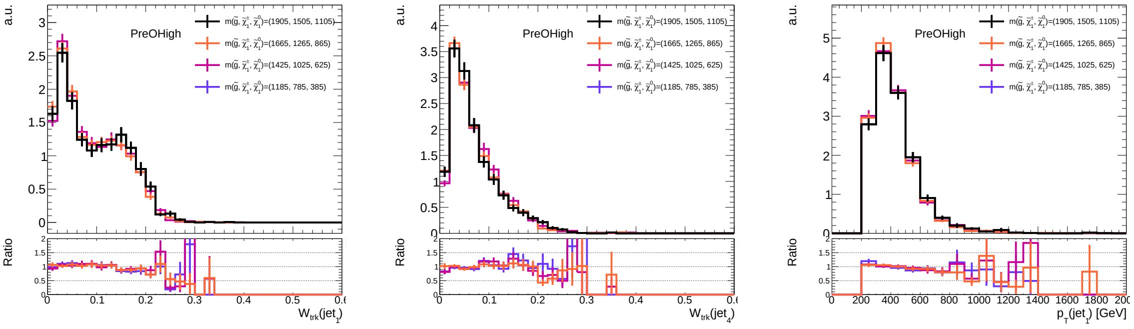
Figure B.4: Distributions of $m_{\rm eff}$(inc), $E_{\rm T}^{\rm miss}/\sqrt{H_{\rm T}}$, $\Delta\phi({\rm j}_{1,2,(3)}, \vec{E}_{\rm T}^{\rm miss})_{\rm min}$, 1st jet $|\eta|$, 1st jet $W_{\rm trk}$, 4th jet $W_{\rm trk}$, 1st jet $p_{\rm T}$, and 4th jet $p_{\rm T}$ in $\Delta M(\tilde{g}, \tilde{\chi}_1^0) = 200$ GeV signals at the preselection **PreDLow**. The total numbers of each signals are normalized to be the same.



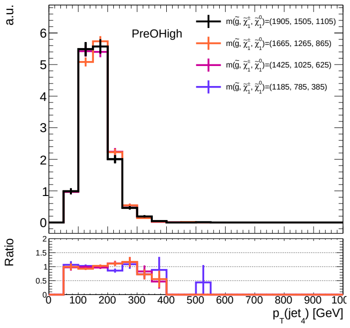
(a) m_{eff}(inc) (b) $E_{\mathrm{T}}^{\mathrm{miss}}/m_{\mathrm{eff}}$(4j) (c) $\Delta\phi(\mathrm{j}_{1,2,(3)}, \vec{E}_{\mathrm{T}}^{\mathrm{miss}})_{\mathrm{min}}$



(d) $\Delta\phi(\mathrm{j}_{i>3}, \vec{E}_{\mathrm{T}}^{\mathrm{miss}})_{\mathrm{min}}$ (e) Aplanarity (f) 1st jet $|\eta|$

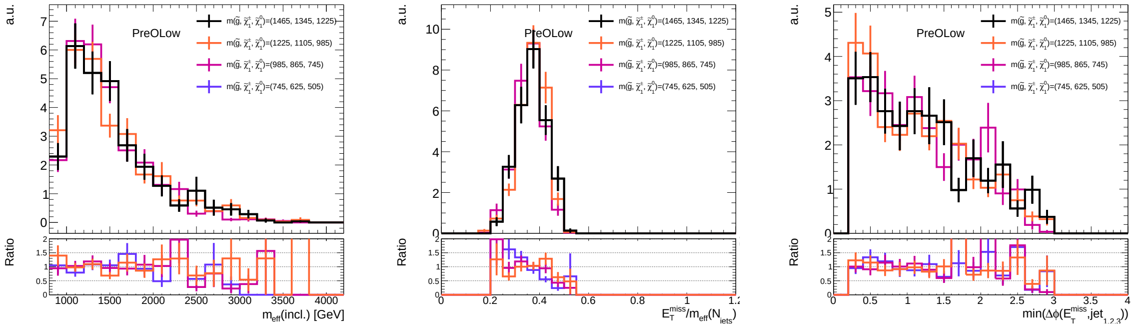


(g) 1st jet W_{trk} (h) 4th jet W_{trk} (i) 1st jet p_{T}

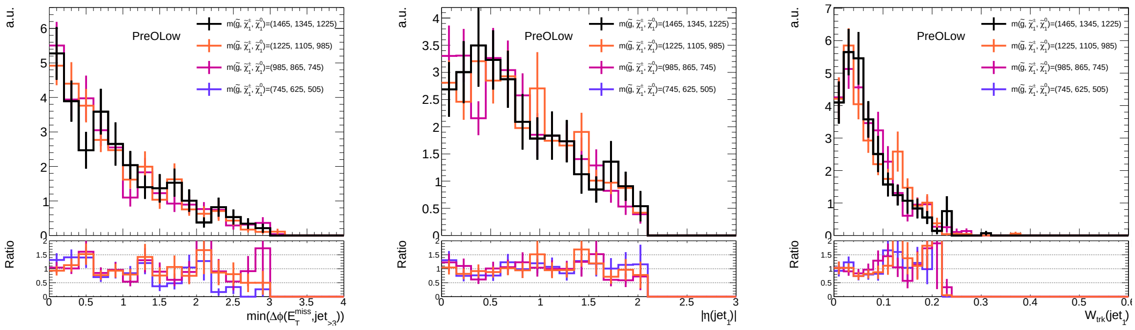


(j) 4th jet p_{T}

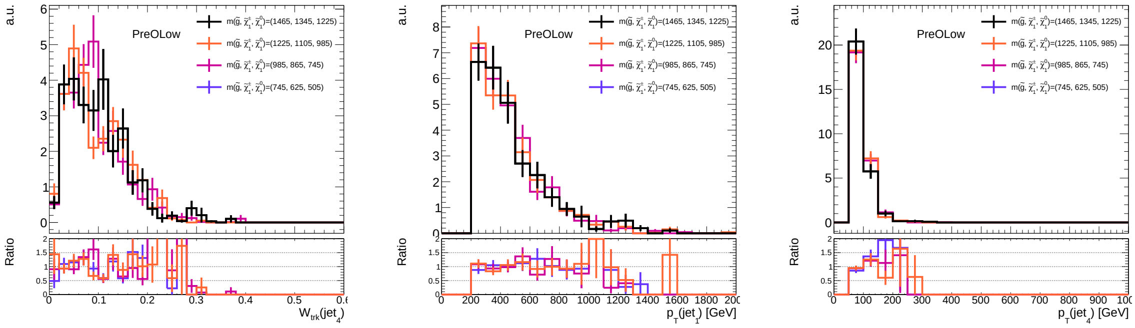
Figure B.5: Distributions of m_{eff}(inc), $E_{\mathrm{T}}^{\mathrm{miss}}/m_{\mathrm{eff}}$(4j), $\Delta\phi(\mathrm{j}_{1,2,(3)}, \vec{E}_{\mathrm{T}}^{\mathrm{miss}})_{\mathrm{min}}$, $\Delta\phi(\mathrm{j}_{i>3}, \vec{E}_{\mathrm{T}}^{\mathrm{miss}})_{\mathrm{min}}$, Aplanarity, 1st jet $|\eta|$, 1st jet W_{trk}, 4th jet W_{trk}, 1st jet p_{T}, and 4th jet p_{T} in $\Delta M(\tilde{g}, \tilde{\chi}_1^0) = 800$ GeV signals at the preselection **PreOHigh**. The total numbers of each signals are normalized to be the same.



(a) $m_{\text{eff}}(\text{inc})$ (b) $E_{\text{T}}^{\text{miss}}/m_{\text{eff}}(4\text{j})$ (c) $\Delta\phi(\text{j}_{1,2,(3)}, \vec{E}_{\text{T}}^{\text{miss}})_{\text{min}}$



(d) $\Delta\phi(\text{j}_{i>3}, \vec{E}_{\text{T}}^{\text{miss}})_{\text{min}}$ (e) 1st jet $|\eta|$ (f) 1st jet W_{trk}



(g) 4th jet W_{trk} (h) 1st jet p_{T} (i) 4th jet p_{T}

Figure B.6: Distributions of $m_{\text{eff}}(\text{inc})$, $E_{\text{T}}^{\text{miss}}/m_{\text{eff}}(4\text{j})$, $\Delta\phi(\text{j}_{1,2,(3)}, \vec{E}_{\text{T}}^{\text{miss}})_{\text{min}}$, $\Delta\phi(\text{j}_{i>3}, \vec{E}_{\text{T}}^{\text{miss}})_{\text{min}}$, 1st jet $|\eta|$, 1st jet W_{trk}, 4th jet W_{trk}, 1st jet p_{T}, and 4th jet p_{T} in $\Delta M(\tilde{g}, \tilde{\chi}_1^0) = 240$ GeV signals at the preselection **PreOLow**. The total numbers of each signals are normalized to be the same.

B.2 Correlation between two input variables

Tables B.1–B.3 show the comparison of five input variables in CRY after PreDLow, PreOHigh, and PreOLow preselection, respectively.

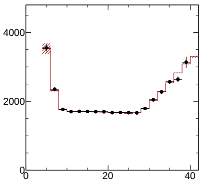
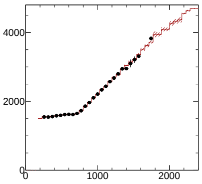
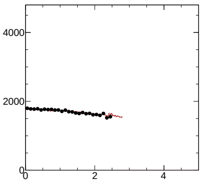
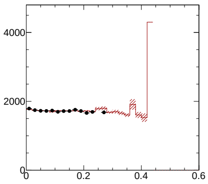
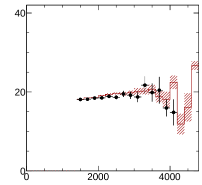
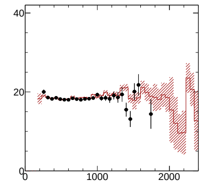
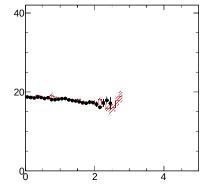
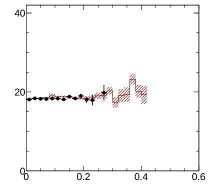
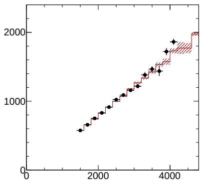
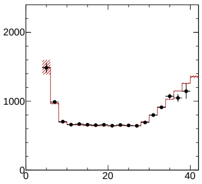
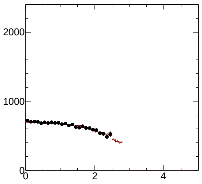
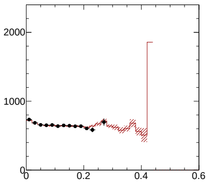
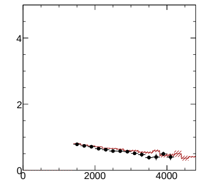
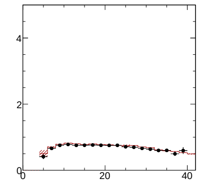
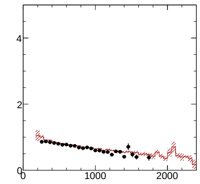
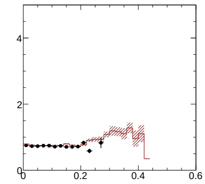
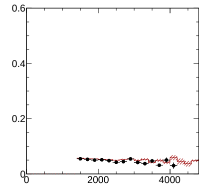
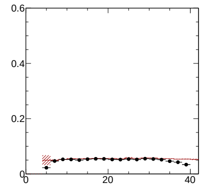
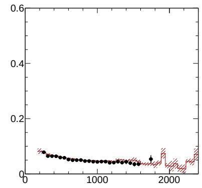
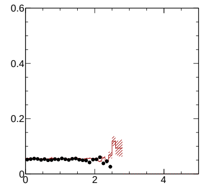
	m_{eff}(inc)	$E_{\mathrm{T}}^{\mathrm{miss}}/\sqrt{H_{\mathrm{T}}}$	1$^{\mathrm{st}}$ jet p_{T}	1$^{\mathrm{st}}$ jet $\lvert\eta\rvert$	1$^{\mathrm{st}}$ jet W_{trk}
m_{eff}(inc)					
$E_{\mathrm{T}}^{\mathrm{miss}}/\sqrt{H_{\mathrm{T}}}$					
1$^{\mathrm{st}}$ jet p_{T}					
1$^{\mathrm{st}}$ jet $\lvert\eta\rvert$					
1$^{\mathrm{st}}$ jet W_{trk}					

Table B.1: This table shows profiles between two variables of five input variables in CRY at the preselection **PreDLow**. The y-axis in each figure is mean of one variable in each x bin of the other variable. Each top title indicates a variable of the x-axis of figures in each column. Each left title indicates a variable of the y-axis of figures in each row. Black point shows 36.1fb^{-1}data and the red line shows the total SM background, which is normalized by cross-section.

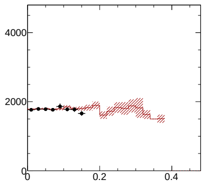
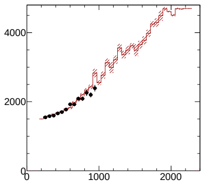
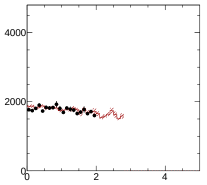
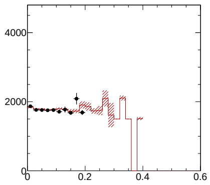
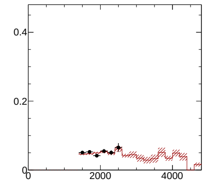
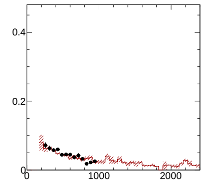
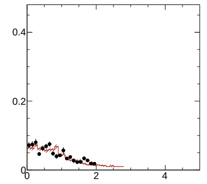
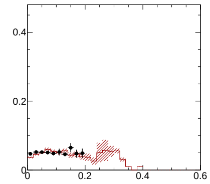
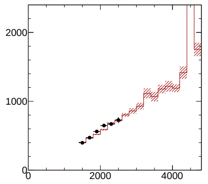
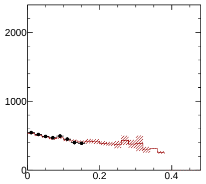
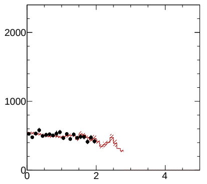
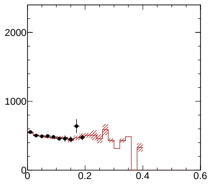
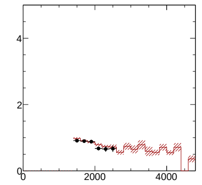
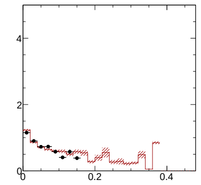
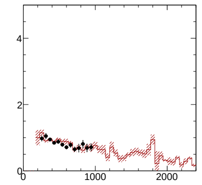
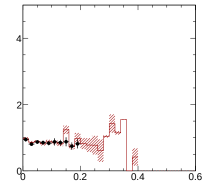
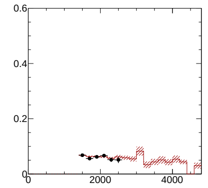
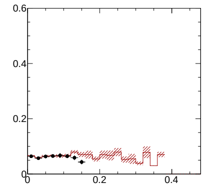
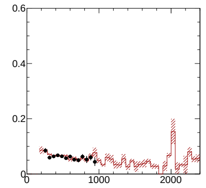
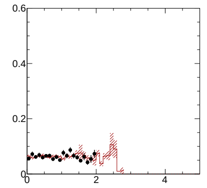
	m_{eff}(inc)	Aplanarity	1$^{\text{st}}$ jet p_{T}	1$^{\text{st}}$ jet $\lvert\eta\rvert$	1$^{\text{st}}$ jet W_{trk}
m_{eff}(inc)					
Aplanarity					
1$^{\text{st}}$ jet p_{T}					
1$^{\text{st}}$ jet $\lvert\eta\rvert$					
1$^{\text{st}}$ jet W_{trk}					

Table B.2: This table shows profiles between two variables of five input variables in CRY at the preselection **PreOHigh**. The y-axis in each figure is mean of one variable in each x bin of the other variable. Each top title indicates a variable of the x-axis of figures in each column. Each left title indicates a variable of the y-axis of figures in each row. Black point shows 36.1fb^{-1}data and the red line shows the total SM background, which is normalized by cross-section.

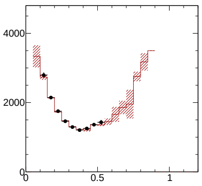
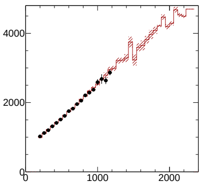
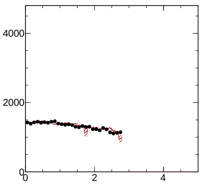
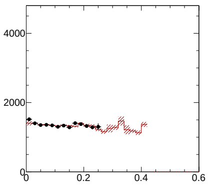
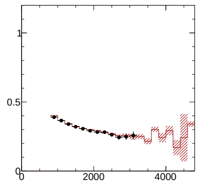
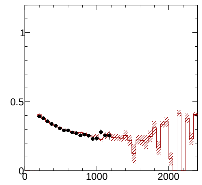
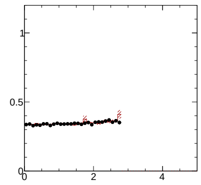
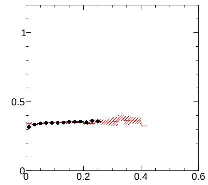
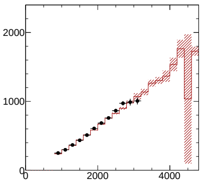
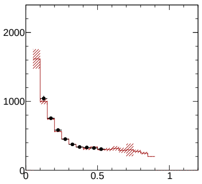
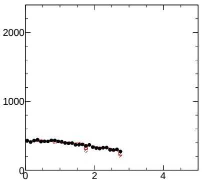
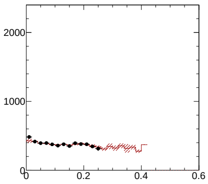
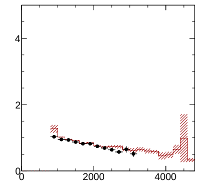
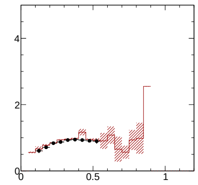
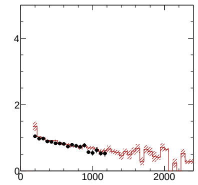
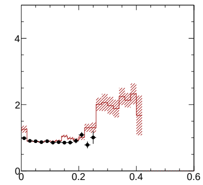
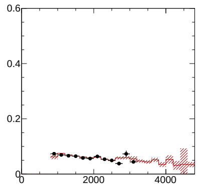
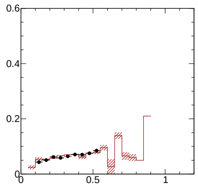
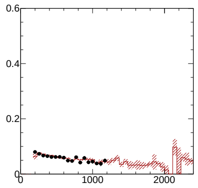
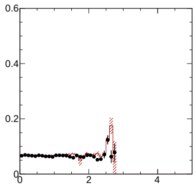
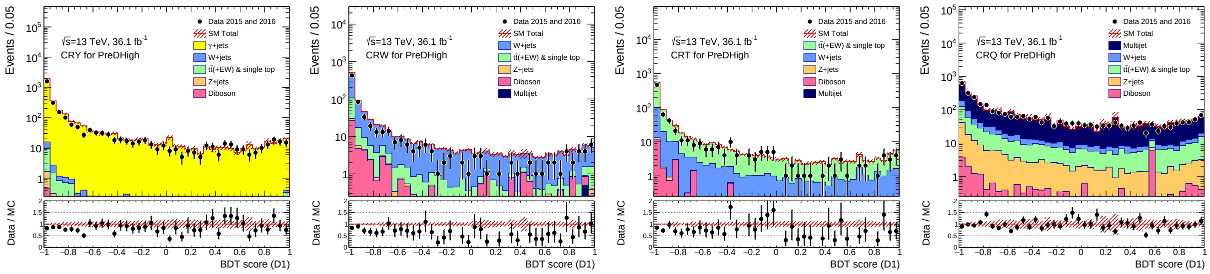
	$m_{\mathrm{eff}}(\mathrm{inc})$	$E_{\mathrm{T}}^{\mathrm{miss}}/m_{\mathrm{eff}}(4\mathrm{j})$	1st jet p_{T}	1st jet $\lvert\eta\rvert$	1st jet W_{trk}
$m_{\mathrm{eff}}(\mathrm{inc})$					
$E_{\mathrm{T}}^{\mathrm{miss}}/m_{\mathrm{eff}}(4\mathrm{j})$					
1st jet p_{T}					
1st jet $\lvert\eta\rvert$					
1st jet W_{trk}					

Table B.3: This table shows profiles between two variables of five input variables in CRY at the preselection **PreOLow**. The y-axis in each figure is mean of one variable in each x bin of the other variable. Each top title indicates a variable of the x-axis of figures in each column. Each left title indicates a variable of the y-axis of figures in each row. Black point shows 36.1fb^{-1}data and the red line shows the total SM background, which is normalized by cross-section.

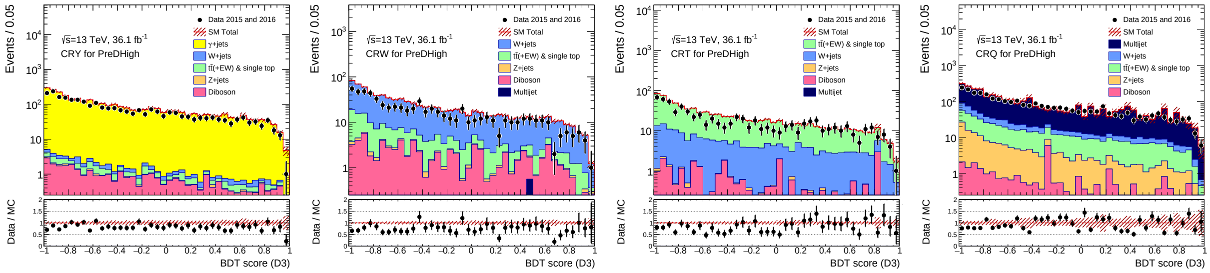
B.3 BDT scores in the CRs

Figures B.7–B.14 show the distributions of the BDT scores in the CRs for SRD1, SRD3, SRD4, SRO1, SRO3, and SRO4.



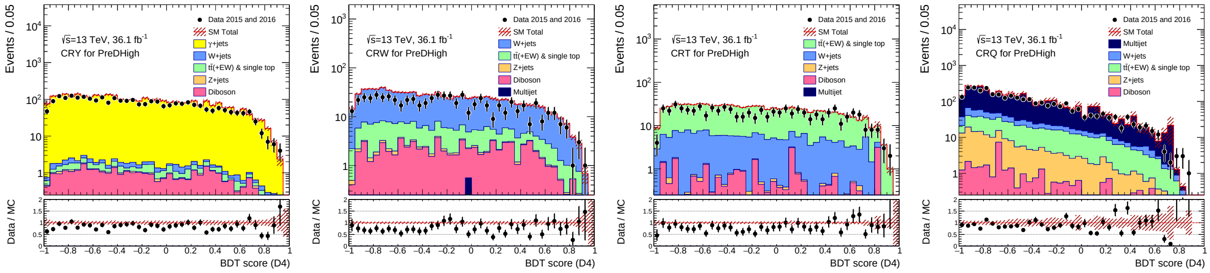
(a) CRY (b) CRW (c) CRT (d) CRQ

Figure B.7: Distributions of BDT score (D1) in the CRY, CRW, CRT, and CRQ of the preselection **PreDHigh**. Black point shows 36.1fb^{-1}data and filled histograms show the SM background MC normalized by cross-section except for the γ +jets process in the CRY. The normalization of the γ +jets is multiplied by κ-factor in addition.



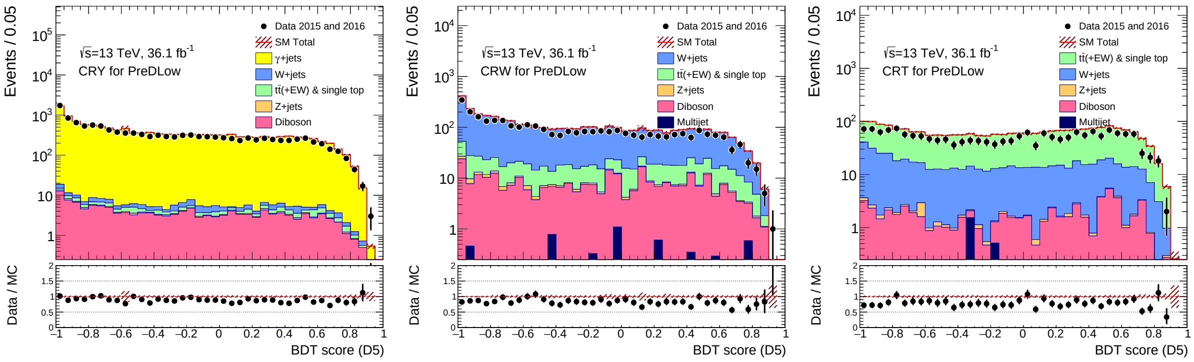
(a) CRY (b) CRW (c) CRT (d) CRQ

Figure B.8: Distributions of BDT score (D3) in the CRY, CRW, CRT, and CRQ of the preselection **PreDHigh**. Black point shows 36.1fb^{-1}data and filled histograms show the SM background MC normalized by cross-section except for the γ +jets process in the CRY. The normalization of the γ +jets is multiplied by κ-factor in addition.



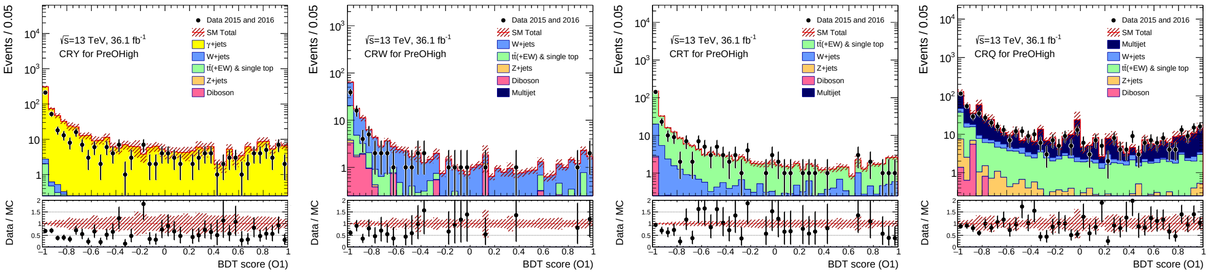
(a) CRY (b) CRW (c) CRT (d) CRQ

Figure B.9: Distributions of BDT score (D4) in the CRY, CRW, CRT, and CRQ of the preselection **PreDHigh**. Black point shows 36.1fb^{-1}data and filled histograms show the SM background MC normalized by cross-section except for the γ +jets process in the CRY. The normalization of the γ +jets is multiplied by κ-factor in addition.



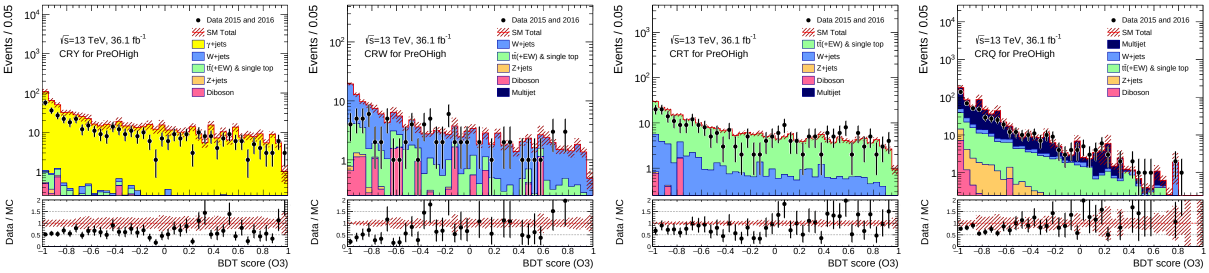
(a) CRY (b) CRW (c) CRT

Figure B.10: Distributions of BDT score (D5) in the CRY, CRW, and CRT of the preselection **PreDLow**. Black point shows 36.1fb^{-1}data and filled histograms show the SM background MC normalized by cross-section except for the γ +jets process in the CRY. The normalization of the γ +jets is multiplied by κ-factor in addition.



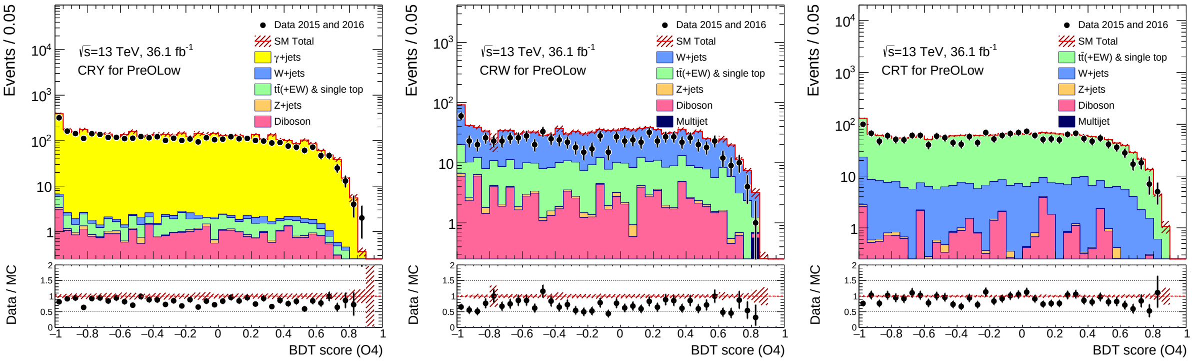
(a) CRY (b) CRW (c) CRT (d) CRQ

Figure B.11: Distributions of BDT score (O1) in the CRY, CRW, CRT, and CRQ of the preselection **PreOHigh**. Black point shows 36.1fb^{-1}data and filled histograms show the SM background MC normalized by cross-section except for the γ +jets process in the CRY. The normalization of the γ +jets is multiplied by κ-factor in addition.



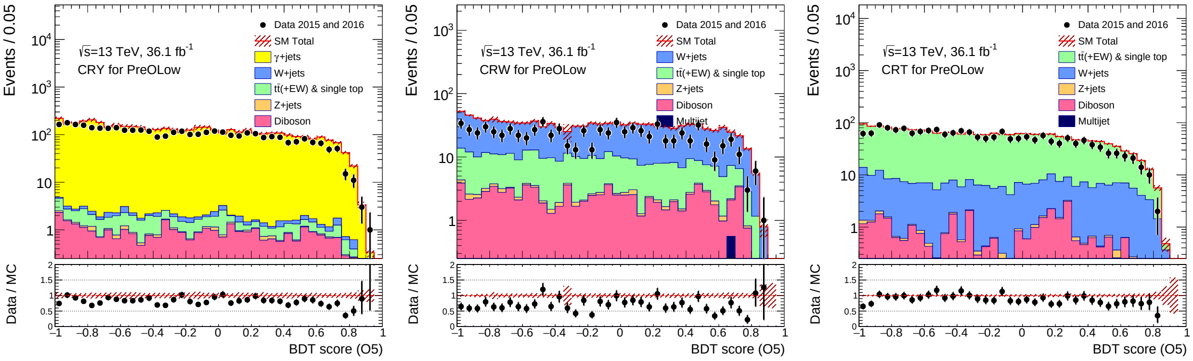
(a) CRY (b) CRW (c) CRT (d) CRQ

Figure B.12: Distributions of BDT score (O3) in the CRY, CRW, CRT, and CRQ of the preselection **PreOHigh**. Black point shows 36.1fb^{-1}data and filled histograms show the SM background MC normalized by cross-section except for the γ +jets process in the CRY. The normalization of the γ +jets is multiplied by κ-factor in addition.



(a) CRY (b) CRW (c) CRT

Figure B.13: Distributions of BDT score (O4) in the CRY, CRW, and CRT of the preselection **PreOLow**. Black point shows 36.1fb^{-1}data and filled histograms show the SM background MC normalized by cross-section except for the γ +jets process in the CRY. The normalization of the γ +jets is multiplied by κ-factor in addition.



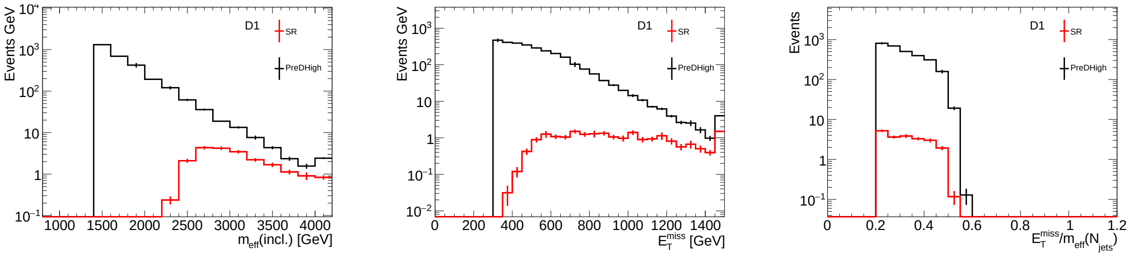
(a) CRY (b) CRW (c) CRT

Figure B.14: Distributions of BDT score (O5) in the CRY, CRW, and CRT of the preselection **PreOLow**. Black point shows 36.1fb^{-1}data and filled histograms show the SM background MC normalized by cross-section except for the γ +jets process in the CRY. The normalization of the γ +jets is multiplied by κ-factor in addition.

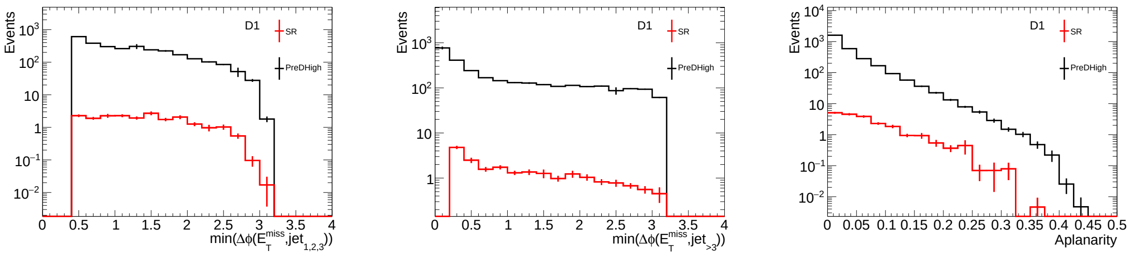
B.4 Distributions before and after BDT cut

B.4.1 One-dimensional distributions

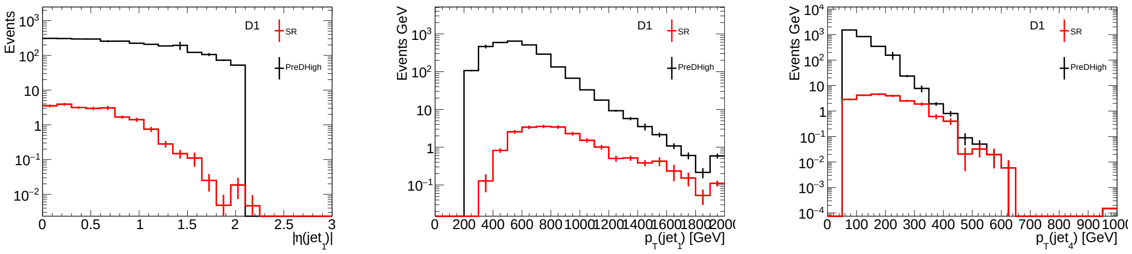
Figures B.15–B.24 show the distributions of the kinematic variables in the preselection and SR for SRD1–5 and SRO1–5.



(a) m_{eff}(inc) (b) $E_{\mathrm{T}}^{\mathrm{miss}}$ (c) $E_{\mathrm{T}}^{\mathrm{miss}}/m_{\mathrm{eff}}$(4j)

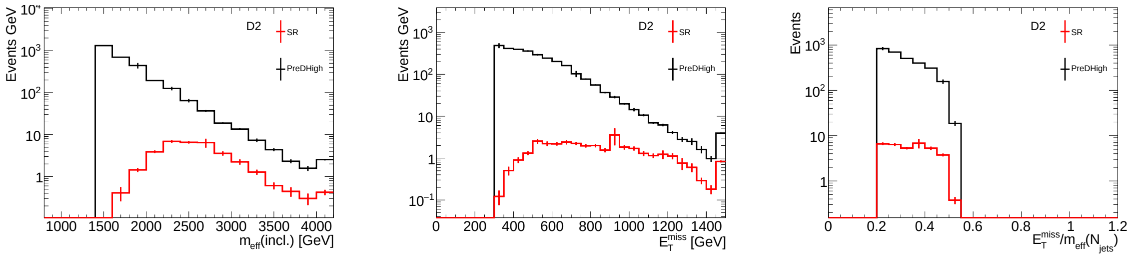


(d) $\Delta\phi(\mathrm{j}_{1,2,(3)}, \vec{E}_{\mathrm{T}}^{miss})_{\mathrm{min}}$ (e) $\Delta\phi(\mathrm{j}_{i>3}, \vec{E}_{\mathrm{T}}^{miss})_{\mathrm{min}}$ (f) Aplanarity

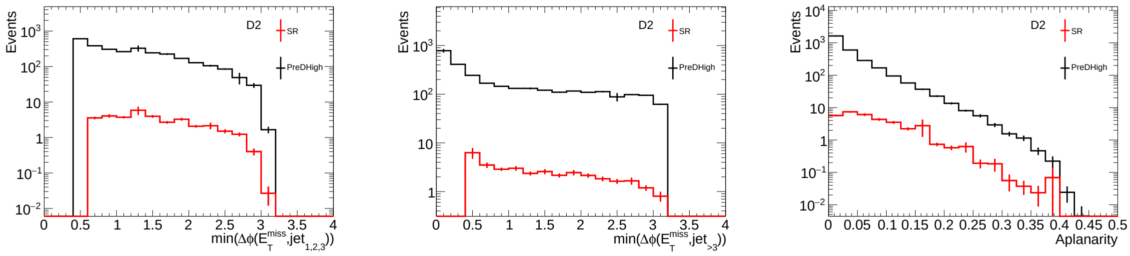


(g) 1st jet $|\eta|$ (h) 1st jet p_{T} (i) 4th jet p_{T}

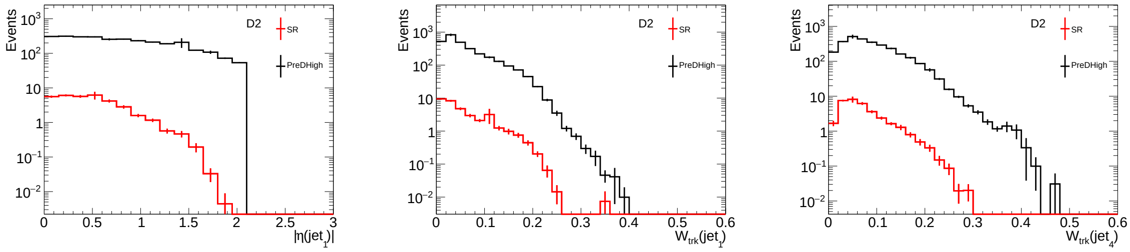
Figure B.15: Distributions of m_{eff}(inc), $E_{\mathrm{T}}^{\mathrm{miss}}$, $E_{\mathrm{T}}^{\mathrm{miss}}/m_{\mathrm{eff}}$(4j), $\Delta\phi(\mathrm{j}_{1,2,(3)}, \vec{E}_{\mathrm{T}}^{miss})_{\mathrm{min}}$, $\Delta\phi(\mathrm{j}_{i>3}, \vec{E}_{\mathrm{T}}^{miss})_{\mathrm{min}}$, Aplanarity, 1st jet $|\eta|$, 1st jet p_{T}, and 4th jet p_{T} of **SRD1** in Monte Carlo simulation. The total number of events is normalized by cross-section.



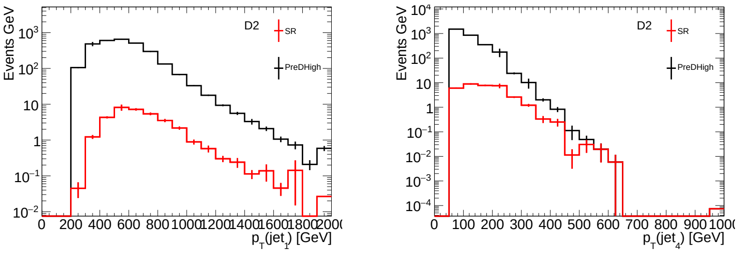
(a) $m_{\text{eff}}(\text{inc})$ (b) $E_{\text{T}}^{\text{miss}}$ (c) $E_{\text{T}}^{\text{miss}}/m_{\text{eff}}(4\text{j})$



(d) $\Delta\phi(\text{j}_{1,2,(3)}, \vec{E}_{\text{T}}^{miss})_{\text{min}}$ (e) $\Delta\phi(\text{j}_{i>3}, \vec{E}_{\text{T}}^{miss})_{\text{min}}$ (f) Aplanarity

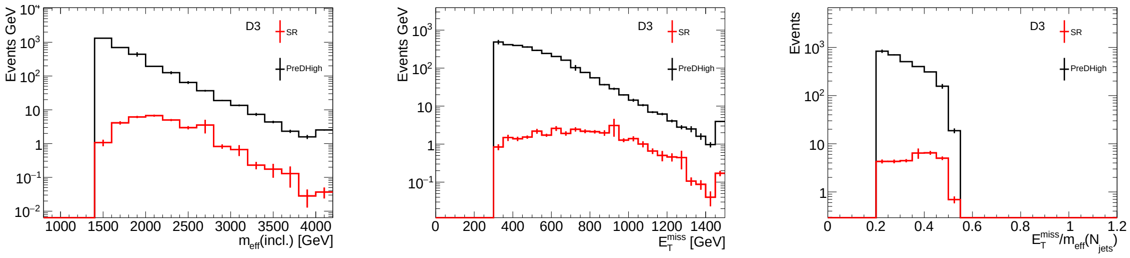


(g) 1st jet $|\eta|$ (h) 1st jet W_{trk} (i) 4th jet W_{trk}

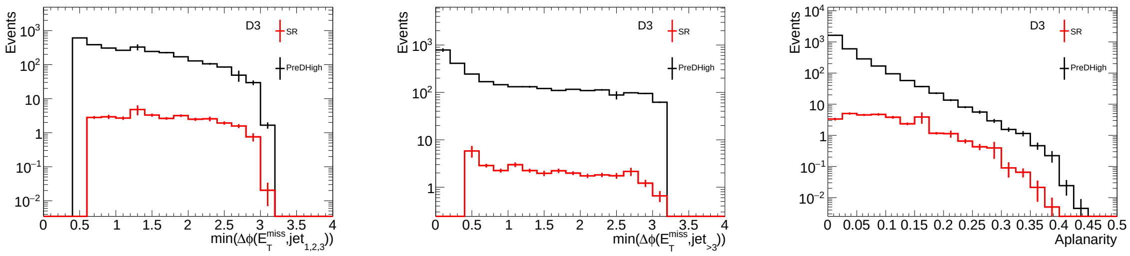


(j) 1st jet p_{T} (k) 4th jet p_{T}

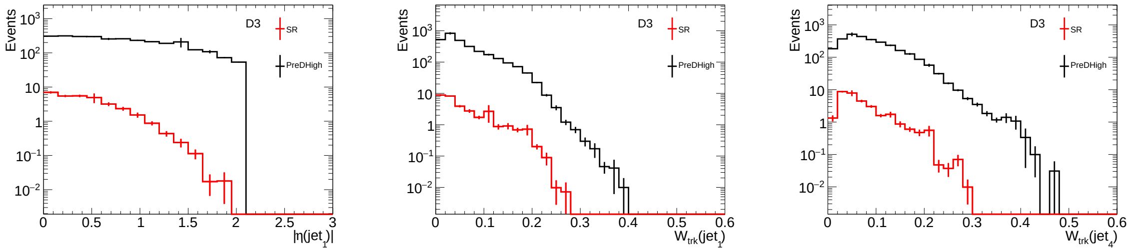
Figure B.16: Distributions of $m_{\text{eff}}(\text{inc})$, $E_{\text{T}}^{\text{miss}}$, $E_{\text{T}}^{\text{miss}}/m_{\text{eff}}(4\text{j})$, $\Delta\phi(\text{j}_{1,2,(3)}, \vec{E}_{\text{T}}^{miss})_{\text{min}}$, $\Delta\phi(\text{j}_{i>3}, \vec{E}_{\text{T}}^{miss})_{\text{min}}$, Aplanarity, 1st jet $|\eta|$, 1st jet W_{trk}, 4th jet W_{trk}, 1st jet p_{T}, and 4th jet p_{T} of **SRD2** in Monte Carlo simulation. The total number of events is normalized by cross-section.



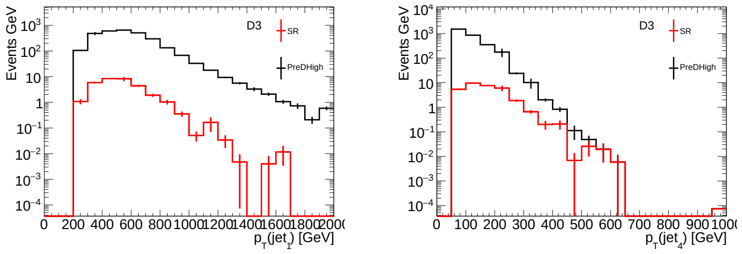
(a) $m_{\mathrm{eff}}(\mathrm{inc})$ (b) $E_{\mathrm{T}}^{\mathrm{miss}}$ (c) $E_{\mathrm{T}}^{\mathrm{miss}}/m_{\mathrm{eff}}(4\mathrm{j})$



(d) $\Delta\phi(\mathrm{j}_{1,2,(3)}, \vec{E}_{\mathrm{T}}^{miss})_{\mathrm{min}}$ (e) $\Delta\phi(\mathrm{j}_{i>3}, \vec{E}_{\mathrm{T}}^{miss})_{\mathrm{min}}$ (f) Aplanarity

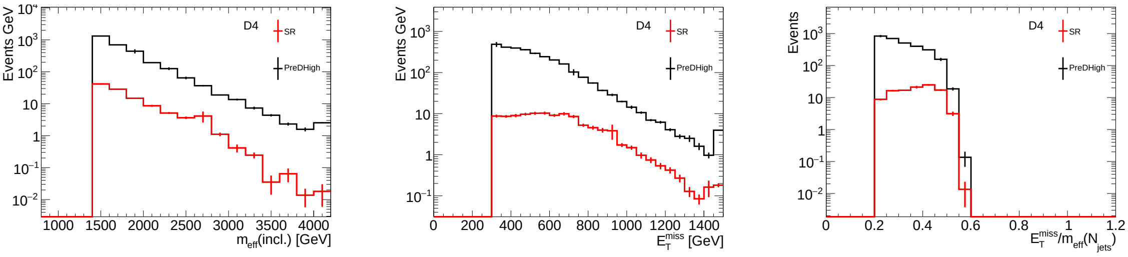


(g) 1st jet $|\eta|$ (h) 1st jet W_{trk} (i) 4th jet W_{trk}

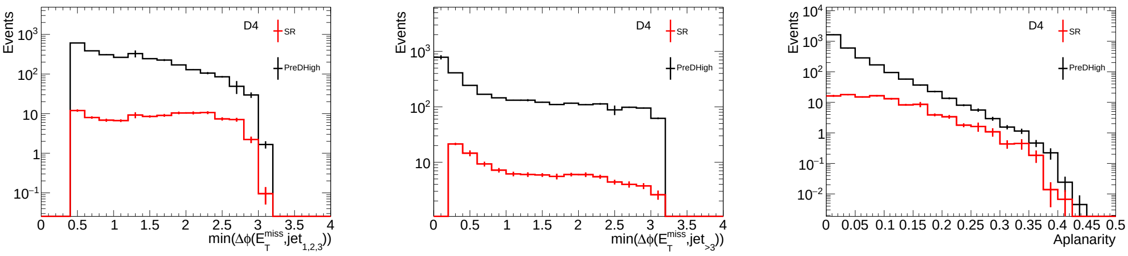


(j) 1st jet p_{T} (k) 4th jet p_{T}

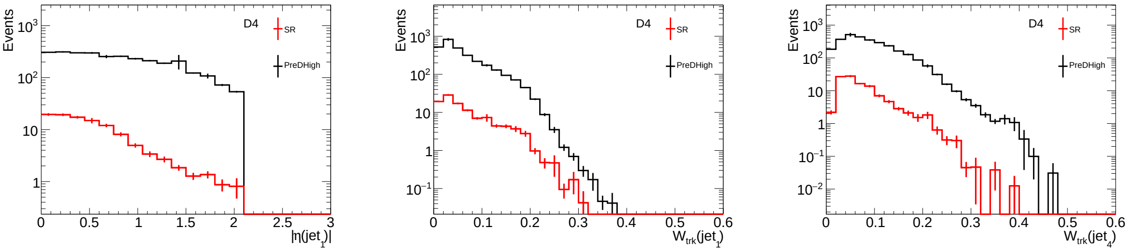
Figure B.17: Distributions of $m_{\mathrm{eff}}(\mathrm{inc})$, $E_{\mathrm{T}}^{\mathrm{miss}}$, $E_{\mathrm{T}}^{\mathrm{miss}}/m_{\mathrm{eff}}(4\mathrm{j})$, $\Delta\phi(\mathrm{j}_{1,2,(3)}, \vec{E}_{\mathrm{T}}^{miss})_{\mathrm{min}}$, $\Delta\phi(\mathrm{j}_{i>3}, \vec{E}_{\mathrm{T}}^{miss})_{\mathrm{min}}$, Aplanarity, 1st jet $|\eta|$, 1st jet W_{trk}, 4th jet W_{trk}, 1st jet p_{T}, and 4th jet p_{T} of **SRD3** in Monte Carlo simulation. The total number of events is normalized by cross-section.



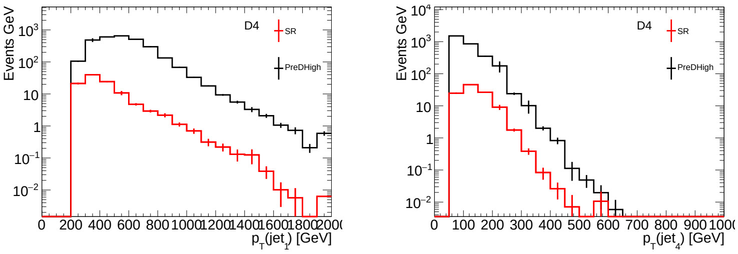
(a) $m_{\mathrm{eff}}(\mathrm{inc})$ (b) $E_{\mathrm{T}}^{\mathrm{miss}}$ (c) $E_{\mathrm{T}}^{\mathrm{miss}}/m_{\mathrm{eff}}(4\mathrm{j})$



(d) $\Delta\phi(\mathrm{j}_{1,2,(3)}, \vec{E}_{\mathrm{T}}^{miss})_{\mathrm{min}}$ (e) $\Delta\phi(\mathrm{j}_{i>3}, \vec{E}_{\mathrm{T}}^{miss})_{\mathrm{min}}$ (f) Aplanarity

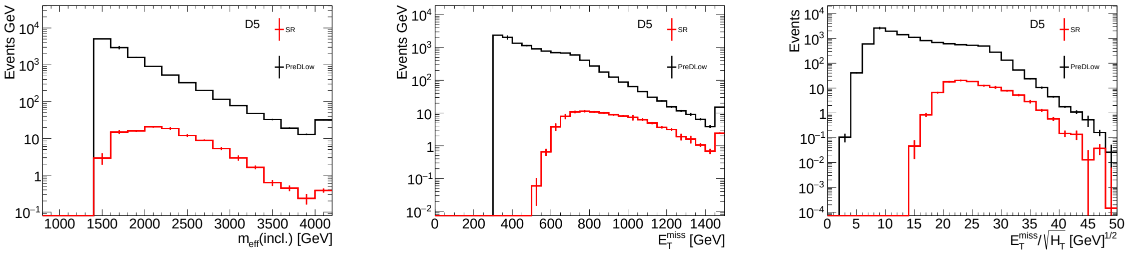


(g) 1st jet $|\eta|$ (h) 1st jet W_{trk} (i) 4th jet W_{trk}

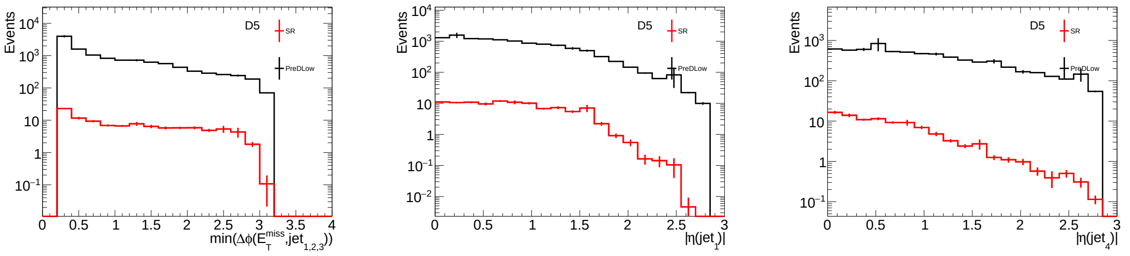


(j) 1st jet p_{T} (k) 4th jet p_{T}

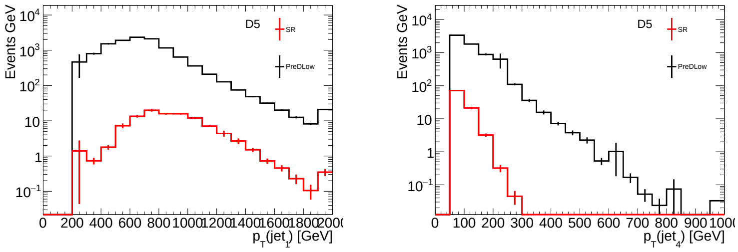
Figure B.18: Distributions of $m_{\mathrm{eff}}(\mathrm{inc})$, $E_{\mathrm{T}}^{\mathrm{miss}}$, $E_{\mathrm{T}}^{\mathrm{miss}}/m_{\mathrm{eff}}(4\mathrm{j})$, $\Delta\phi(\mathrm{j}_{1,2,(3)}, \vec{E}_{\mathrm{T}}^{miss})_{\mathrm{min}}$, $\Delta\phi(\mathrm{j}_{i>3}, \vec{E}_{\mathrm{T}}^{miss})_{\mathrm{min}}$, Aplanarity, 1st jet $|\eta|$, 1st jet W_{trk}, 4th jet W_{trk}, 1st jet p_{T}, and 4th jet p_{T} of **SRD4** in Monte Carlo simulation. The total number of events is normalized by cross-section.



(a) $m_{\mathrm{eff}}(\mathrm{inc})$ (b) $E_{\mathrm{T}}^{\mathrm{miss}}$ (c) $E_{\mathrm{T}}^{\mathrm{miss}}/\sqrt{H_{\mathrm{T}}}$

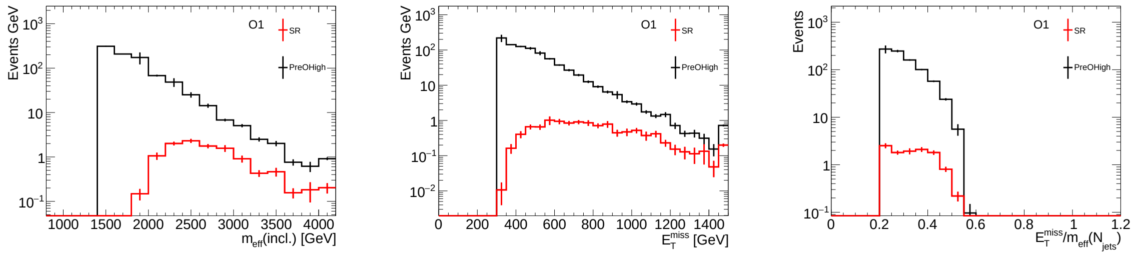


(d) $\Delta\phi(\mathrm{j}_{1,2,(3)}, \vec{E}_{\mathrm{T}}^{miss})_{\mathrm{min}}$ (e) 1st jet $|\eta|$ (f) 4th jet $|\eta|$

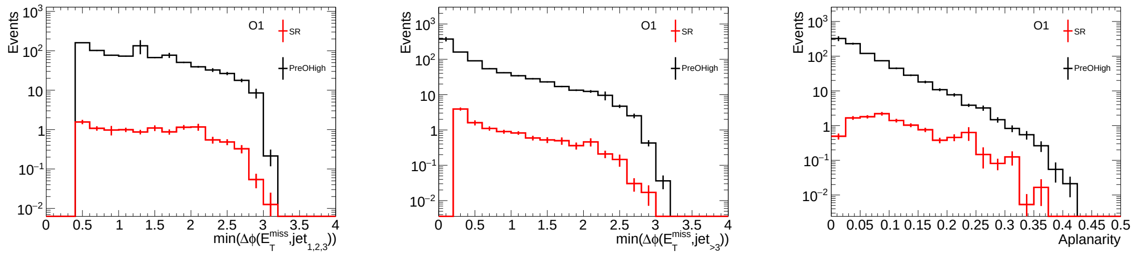


(g) 1st jet p_{T} (h) 4th jet p_{T}

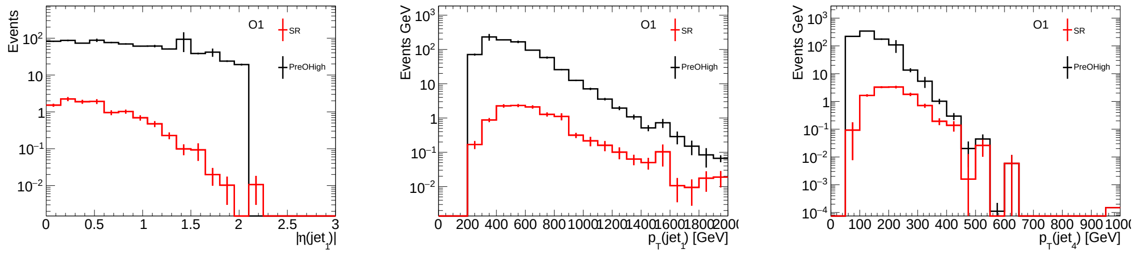
Figure B.19: Distributions of $m_{\mathrm{eff}}(\mathrm{inc})$, $E_{\mathrm{T}}^{\mathrm{miss}}$, $E_{\mathrm{T}}^{\mathrm{miss}}/\sqrt{H_{\mathrm{T}}}$, $\Delta\phi(\mathrm{j}_{1,2,(3)}, \vec{E}_{\mathrm{T}}^{miss})_{\mathrm{min}}$, 1st jet $|\eta|$, 4th jet $|\eta|$, 1st jet p_{T}, and 4th jet p_{T} of **SRD5** in Monte Carlo simulation. The total number of events is normalized by cross-section.



(a) $m_{\rm eff}({\rm inc})$ (b) $E_{\rm T}^{\rm miss}$ (c) $E_{\rm T}^{\rm miss}/m_{\rm eff}(4{\rm j})$

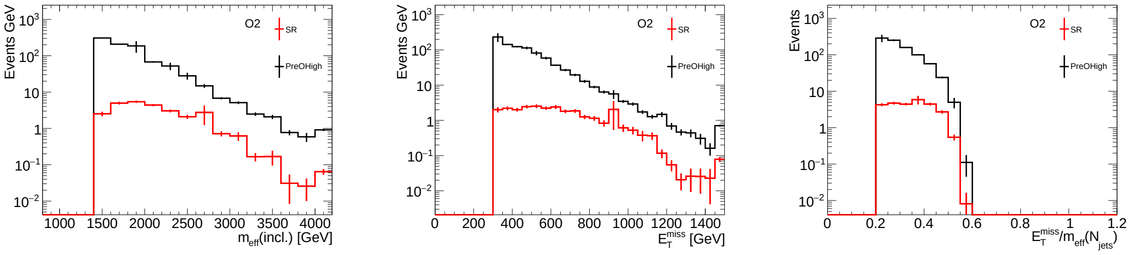


(d) $\Delta\phi({\rm j}_{1,2,(3)}, \vec{E}_{\rm T}^{miss})_{\rm min}$ (e) $\Delta\phi({\rm j}_{i>3}, \vec{E}_{\rm T}^{miss})_{\rm min}$ (f) Aplanarity



(g) 1st jet $|\eta|$ (h) 1st jet $p_{\rm T}$ (i) 4th jet $p_{\rm T}$

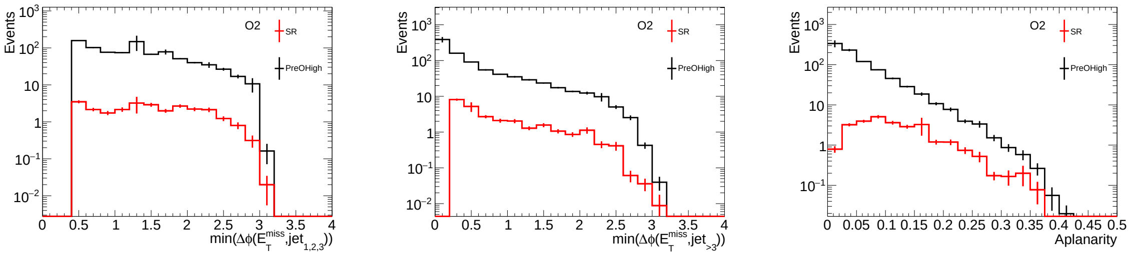
Figure B.20: Distributions of $m_{\rm eff}({\rm inc})$, $E_{\rm T}^{\rm miss}$, $E_{\rm T}^{\rm miss}/m_{\rm eff}(4{\rm j})$, $\Delta\phi({\rm j}_{1,2,(3)}, \vec{E}_{\rm T}^{miss})_{\rm min}$, $\Delta\phi({\rm j}_{i>3}, \vec{E}_{\rm T}^{miss})_{\rm min}$, Aplanarity, 1st jet $|\eta|$, 1st jet $p_{\rm T}$, and 4th jet $p_{\rm T}$ of **SRO1** in Monte Carlo simulation. The total number of events is normalized by cross-section.



(a) m_{eff}(inc)

(b) $E_{\mathrm{T}}^{\mathrm{miss}}$

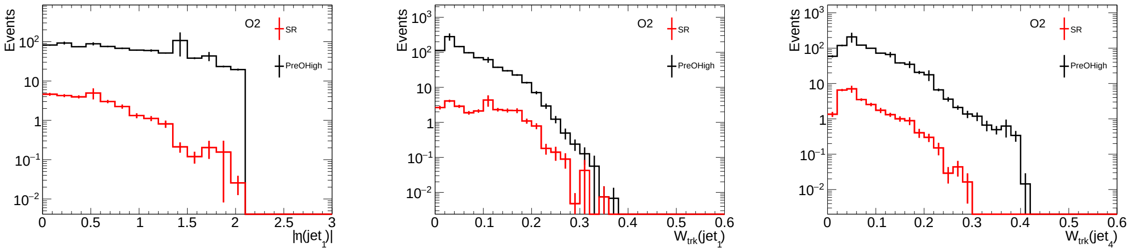
(c) $E_{\mathrm{T}}^{\mathrm{miss}}/m_{\mathrm{eff}}$(4j)



(d) $\Delta\phi(\mathrm{j}_{1,2,(3)}, \vec{E}_{\mathrm{T}}^{miss})_{\mathrm{min}}$

(e) $\Delta\phi(\mathrm{j}_{i>3}, \vec{E}_{\mathrm{T}}^{miss})_{\mathrm{min}}$

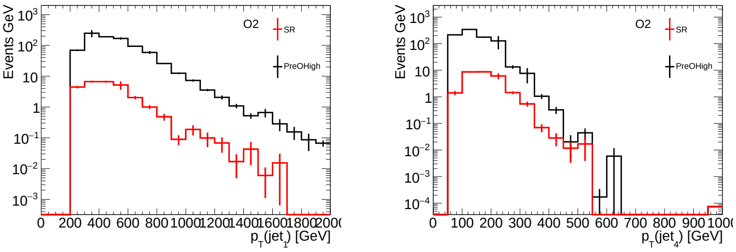
(f) Aplanarity



(g) 1st jet $|\eta|$

(h) 1st jet W_{trk}

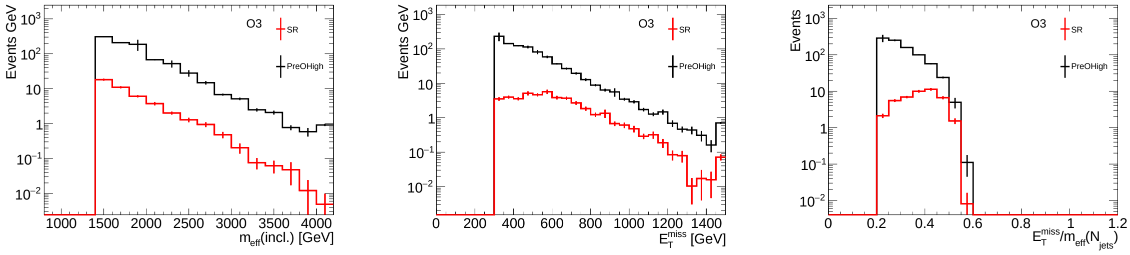
(i) 4th jet W_{trk}



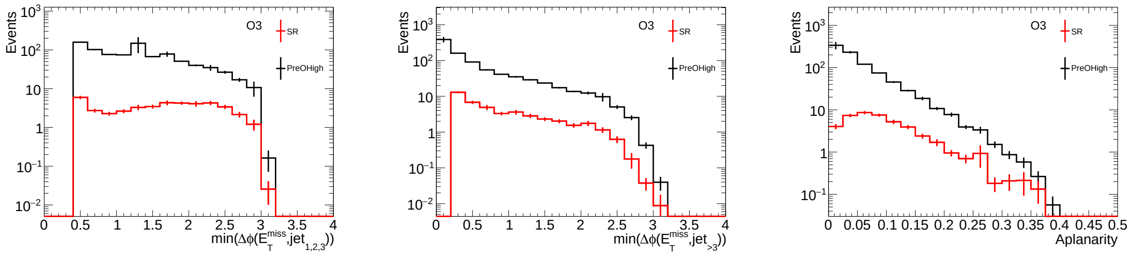
(j) 1st jet p_{T}

(k) 4th jet p_{T}

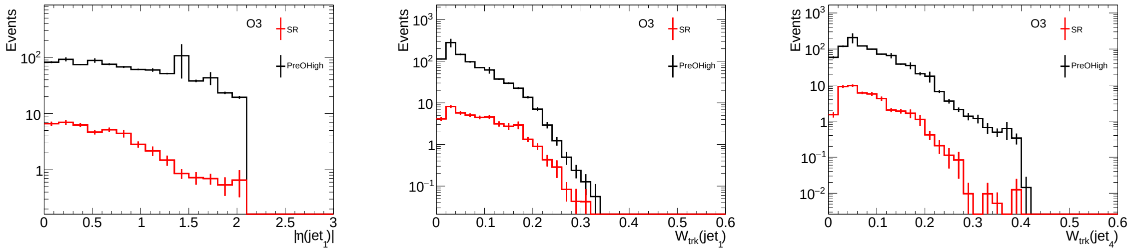
Figure B.21: Distributions of m_{eff}(inc), $E_{\mathrm{T}}^{\mathrm{miss}}$, $E_{\mathrm{T}}^{\mathrm{miss}}/m_{\mathrm{eff}}$(4j), $\Delta\phi(\mathrm{j}_{1,2,(3)}, \vec{E}_{\mathrm{T}}^{miss})_{\mathrm{min}}$, $\Delta\phi(\mathrm{j}_{i>3}, \vec{E}_{\mathrm{T}}^{miss})_{\mathrm{min}}$, Aplanarity, 1st jet $|\eta|$, 1st jet W_{trk}, 4th jet W_{trk}, 1st jet p_{T}, and 4th jet p_{T} of **SRO2** in Monte Carlo simulation. The total number of events is normalized by cross-section.



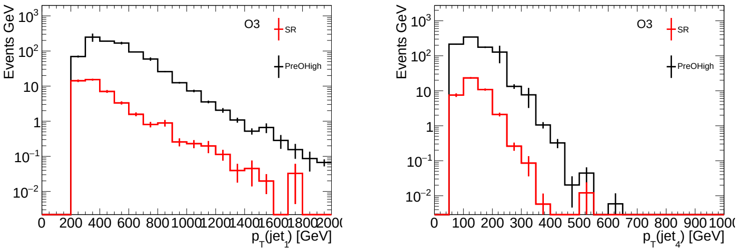
(a) $m_{\mathrm{eff}}(\mathrm{inc})$ (b) $E_{\mathrm{T}}^{\mathrm{miss}}$ (c) $E_{\mathrm{T}}^{\mathrm{miss}}/m_{\mathrm{eff}}(4\mathrm{j})$



(d) $\Delta\phi(\mathrm{j}_{1,2,(3)}, \vec{E}_{\mathrm{T}}^{miss})_{\mathrm{min}}$ (e) $\Delta\phi(\mathrm{j}_{i>3}, \vec{E}_{\mathrm{T}}^{miss})_{\mathrm{min}}$ (f) Aplanarity

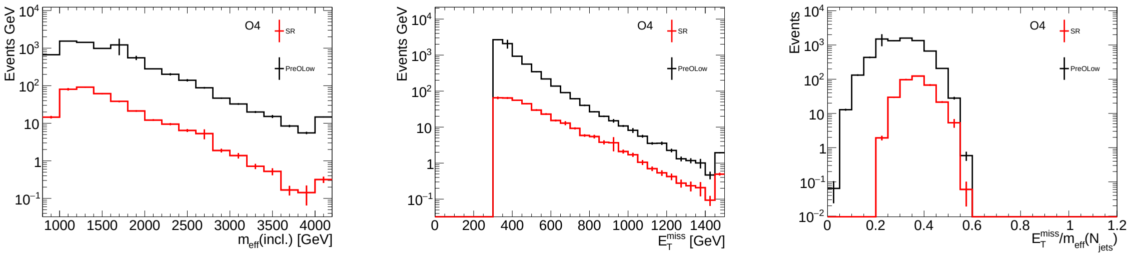


(g) 1st jet $|\eta|$ (h) 1st jet W_{trk} (i) 4th jet W_{trk}

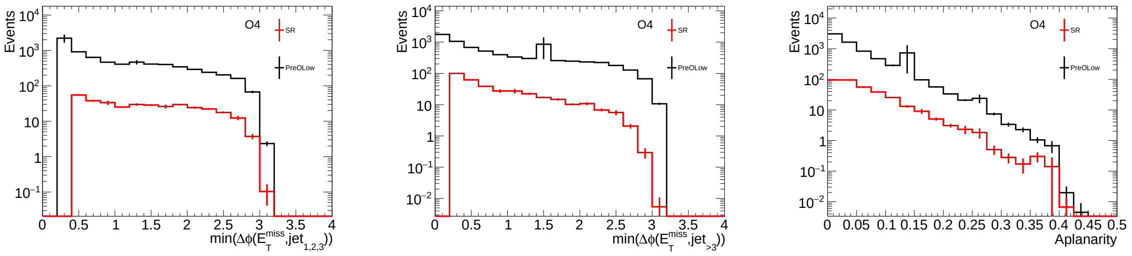


(j) 1st jet p_{T} (k) 4th jet p_{T}

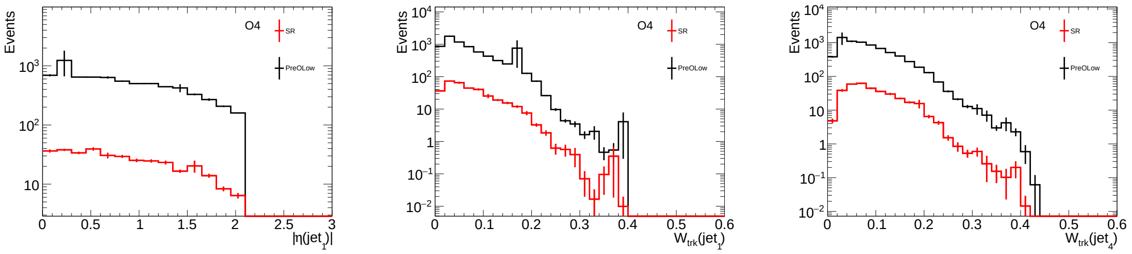
Figure B.22: Distributions of $m_{\mathrm{eff}}(\mathrm{inc})$, $E_{\mathrm{T}}^{\mathrm{miss}}$, $E_{\mathrm{T}}^{\mathrm{miss}}/m_{\mathrm{eff}}(4\mathrm{j})$, $\Delta\phi(\mathrm{j}_{1,2,(3)}, \vec{E}_{\mathrm{T}}^{miss})_{\mathrm{min}}$, $\Delta\phi(\mathrm{j}_{i>3}, \vec{E}_{\mathrm{T}}^{miss})_{\mathrm{min}}$, Aplanarity, 1st jet $|\eta|$, 1st jet W_{trk}, 4th jet W_{trk}, 1st jet p_{T}, and 4th jet p_{T} of **SRO3** in Monte Carlo simulation. The total number of events is normalized by cross-section.



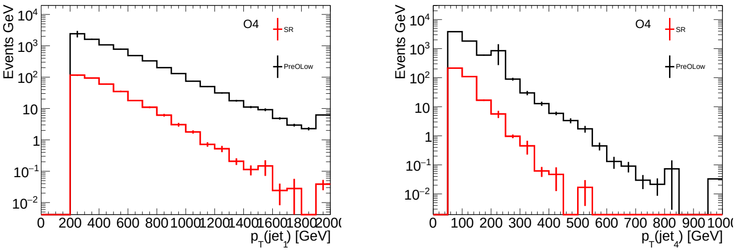
(a) $m_{\mathrm{eff}}(\mathrm{inc})$ (b) $E_{\mathrm{T}}^{\mathrm{miss}}$ (c) $E_{\mathrm{T}}^{\mathrm{miss}}/m_{\mathrm{eff}}(4\mathrm{j})$



(d) $\Delta\phi(\mathrm{j}_{1,2,(3)}, \vec{E}_{\mathrm{T}}^{miss})_{\mathrm{min}}$ (e) $\Delta\phi(\mathrm{j}_{i>3}, \vec{E}_{\mathrm{T}}^{miss})_{\mathrm{min}}$ (f) Aplanarity

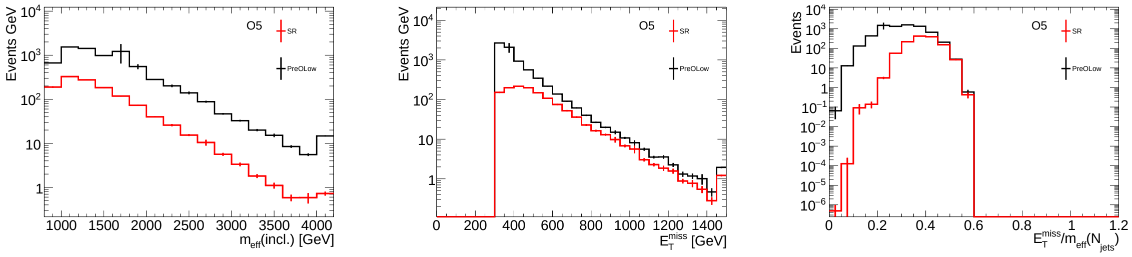


(g) 1st jet $|\eta|$ (h) 1st jet W_{trk} (i) 4th jet W_{trk}

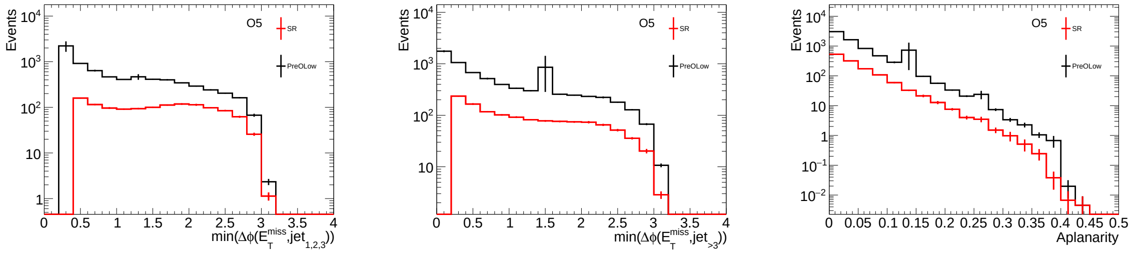


(j) 1st jet p_{T} (k) 4th jet p_{T}

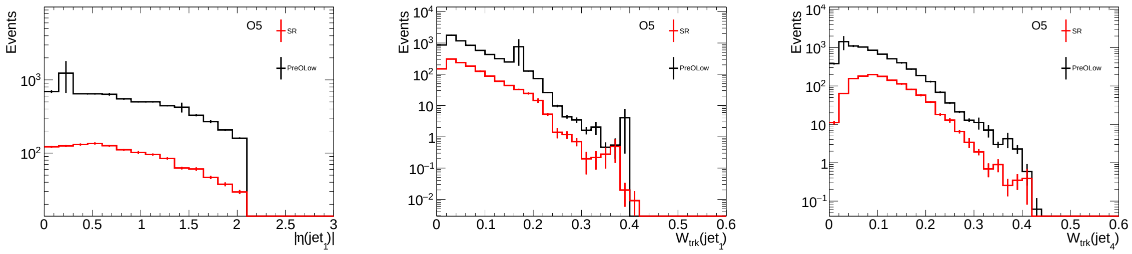
Figure B.23: Distributions of $m_{\mathrm{eff}}(\mathrm{inc})$, $E_{\mathrm{T}}^{\mathrm{miss}}$, $E_{\mathrm{T}}^{\mathrm{miss}}/m_{\mathrm{eff}}(4\mathrm{j})$, $\Delta\phi(\mathrm{j}_{1,2,(3)}, \vec{E}_{\mathrm{T}}^{miss})_{\mathrm{min}}$, $\Delta\phi(\mathrm{j}_{i>3}, \vec{E}_{\mathrm{T}}^{miss})_{\mathrm{min}}$, Aplanarity, 1st jet $|\eta|$, 1st jet W_{trk}, 4th jet W_{trk}, 1st jet p_{T}, and 4th jet p_{T} of **SRO4** in Monte Carlo simulation. The total number of events is normalized by cross-section.



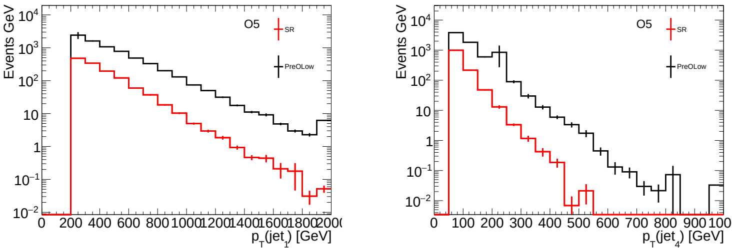
(a) $m_{\mathrm{eff}}(\mathrm{inc})$ (b) $E_{\mathrm{T}}^{\mathrm{miss}}$ (c) $E_{\mathrm{T}}^{\mathrm{miss}}/m_{\mathrm{eff}}(4\mathrm{j})$



(d) $\Delta\phi(\mathrm{j}_{1,2,(3)}, \vec{E}_{\mathrm{T}}^{miss})_{\mathrm{min}}$ (e) $\Delta\phi(\mathrm{j}_{i>3}, \vec{E}_{\mathrm{T}}^{miss})_{\mathrm{min}}$ (f) Aplanarity



(g) 1^{st} jet $|\eta|$ (h) 1^{st} jet W_{trk} (i) 4^{th} jet W_{trk}

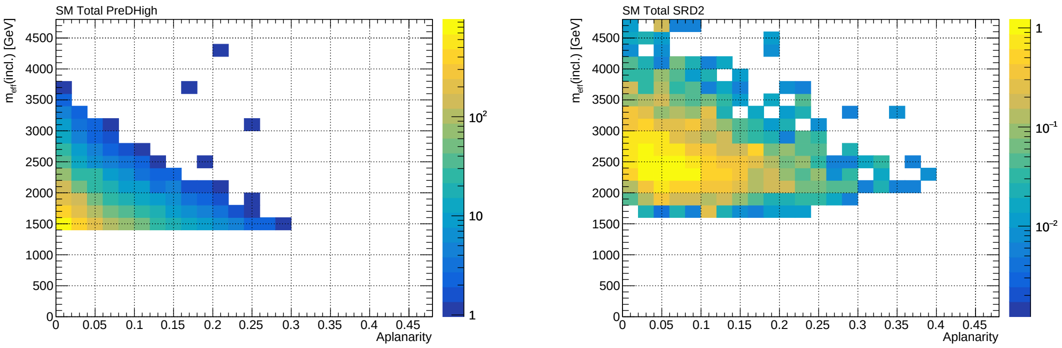


(j) 1^{st} jet p_{T} (k) 4^{th} jet p_{T}

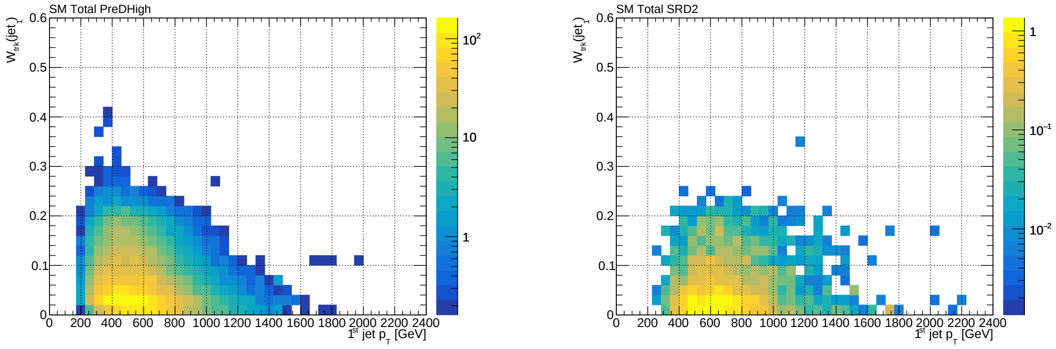
Figure B.24: Distributions of $m_{\mathrm{eff}}(\mathrm{inc})$, $E_{\mathrm{T}}^{\mathrm{miss}}$, $E_{\mathrm{T}}^{\mathrm{miss}}/m_{\mathrm{eff}}(4\mathrm{j})$, $\Delta\phi(\mathrm{j}_{1,2,(3)}, \vec{E}_{\mathrm{T}}^{miss})_{\mathrm{min}}$, $\Delta\phi(\mathrm{j}_{i>3}, \vec{E}_{\mathrm{T}}^{miss})_{\mathrm{min}}$, Aplanarity, 1^{st} jet $|\eta|$, 1^{st} jet W_{trk}, 4^{th} jet W_{trk}, 1^{st} jet p_{T}, and 4^{th} jet p_{T} of **SRO5** in Monte Carlo simulation. The total number of events is normalized by cross-section.

B.4.2 Two-dimensional distributions

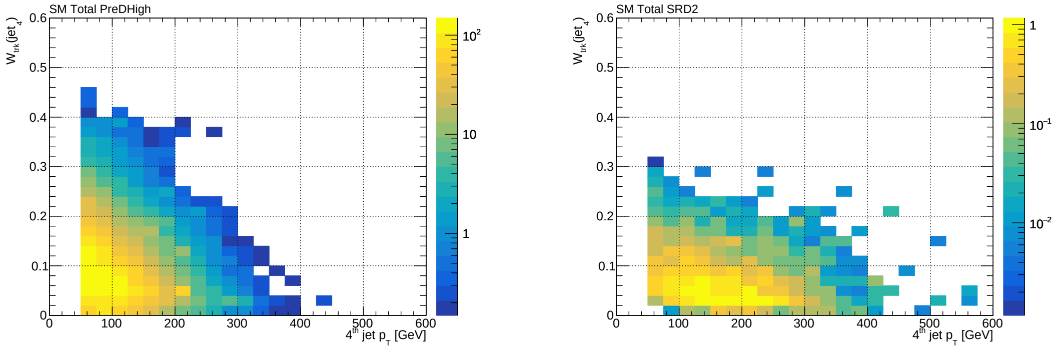
Figures B.25 show the two-dimensional distributions of the kinematic variables in the preselection PreDHigh and SRD2.



(a) Aplanarity v.s. m_{eff}(inc) (PreDHigh) (b) Aplanarity v.s. m_{eff}(inc) (SRD2)



(c) 1st jet p_{T} v.s. 1st jet W_{trk} (PreDHigh) (d) 1st jet p_{T} v.s. 1st jet W_{trk} (SRD2)



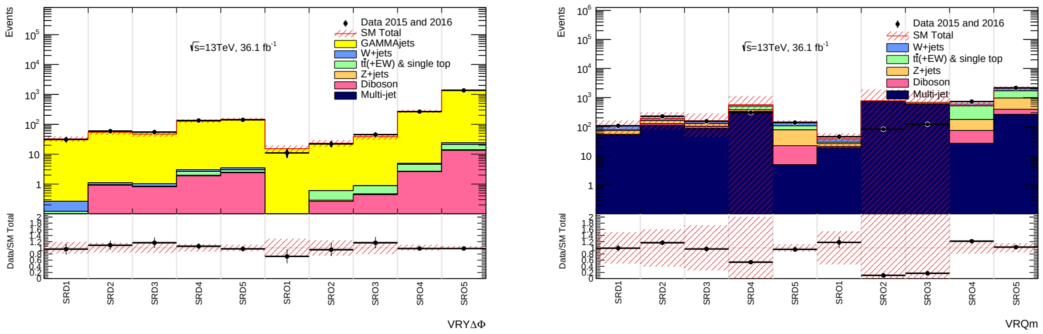
(e) 4th jet p_{T} v.s. 4th jet W_{trk} (PreDHigh) (f) 4th jet p_{T} v.s. 4th jet W_{trk} (SRD2)

Figure B.25: Two-dimensional distributions of Aplanarity v.s. m_{eff}(inc), 1st jet p_{T} v.s. 1st jet W_{trk}, and 4th jet p_{T} v.s. 4th jet W_{trk} in (left) **PreDHigh** and (right) **SRD2** of the SM background events, which are normalized to 36.1fb^{-1}by cross-section.

C Auxiliary material for result

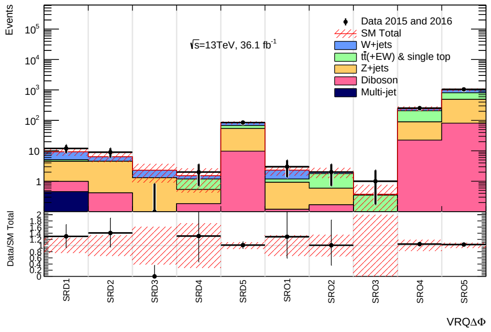
C.1 Validation regions

Figures C.1 show comparisons of the data and background prediction in VRYdPhi, VRQm, and VRQdPhi.



(a) VRYdPhi

(b) VRQm



(c) VRQdPhi

Figure C.1: Numbers of observed data and the SM background prediction in (a) VRYdPhi, (b) VRQm, and (d) VRQdPhi after the background-only fit. Each bin indicates an observed number of data (Black point) and a predicted number of the SM background (Color filled bars). Each color shows a different background process. The red shaded band to the SM total background represents its total uncertainties. The bottom panel shows the number of the data divided by the number of the total background prediction.

C.2 Yield tables

Tables C.1–C.10 show the yields in the SR and CRs for SRD1–D5 and SRO1–O5.

SRD1 channel	CRT	CRW	CRQ	CRY	SR
Observed events	5	12	124	41	23
Fitted bkg events	5.01 ± 1.60	11.99 ± 3.54	124.08 ± 11.56	40.99 ± 6.40	19.17 ± 5.52
Fitted Multijets events	0.00 ± 0.00	0.00 ± 0.00	103.37 ± 14.63	0.00 ± 0.00	$0.04^{+0.04}_{-0.04}$
Fitted Wjets events	3.65 ± 1.52	10.98 ± 3.78	10.21 ± 3.75	0.20 ± 0.08	7.09 ± 2.66
Fitted Zjets events	0.02 ± 0.01	0.11 ± 0.03	4.84 ± 2.04	0.02 ± 0.01	10.06 ± 4.94
Fitted GAMMAjets events	0.00 ± 0.00	0.00 ± 0.00	0.00 ± 0.00	40.65 ± 6.40	0.00 ± 0.00
Fitted Top events	$1.21^{+2.00}_{-1.21}$	$0.27^{+0.49}_{-0.27}$	$4.48^{+8.00}_{-4.48}$	$0.09^{+0.18}_{-0.09}$	$0.52^{+1.21}_{-0.52}$
Fitted Diboson events	0.12 ± 0.09	0.63 ± 0.05	1.19 ± 0.19	0.03 ± 0.00	1.46 ± 0.51
MC exp. SM events	8.74	10.55	118.14	47.45	21.08
MC exp. Multijets events	0.00	0.00	82.19	0.00	0.03
MC exp. Wjets events	2.84	8.52	7.94	0.15	5.51
MC exp. Zjets events	0.02	0.13	5.58	0.02	11.59
MC exp. GAMMAjets events	0.00	0.00	0.00	46.79	0.00
MC exp. Top events	5.76	1.27	21.25	0.45	2.49
MC exp. Diboson events	0.12	0.63	1.19	0.03	1.46

Table C.1: **S**RD1 : Background fit results for the CRT, CRW, CRQ, CRY and SR regions, for an integrated luminosity of 36.1 fb^{-1}. Nominal MC expectations (normalized to MC cross-sections) are given for comparison. The errors shown are the statistical plus systematic uncertainties. The errors shown for the signal region are systematic uncertainties only.

SRD2 channel	CRT	CRW	CRQ	CRY	SR
Observed events	3	12	210	89	29
Fitted bkg events	3.30 ± 3.05	11.75 ± 3.07	210.00 ± 23.32	89.02 ± 9.44	28.30 ± 5.42
Fitted Multijets events	0.00 ± 0.00	0.00 ± 0.00	189.80 ± 21.07	0.00 ± 0.00	$0.00^{+0.00}_{-0.00}$
Fitted Wjets events	3.01 ± 1.02	10.02 ± 3.12	10.10 ± 3.52	0.21 ± 0.09	5.57 ± 1.84
Fitted Zjets events	0.01 ± 0.01	0.14 ± 0.12	$8.19^{+28.16}_{-8.19}$	0.03 ± 0.01	17.91 ± 4.01
Fitted GAMMAjets events	0.00 ± 0.00	0.00 ± 0.00	0.00 ± 0.00	87.07 ± 9.51	0.00 ± 0.00
Fitted Top events	$0.00^{+2.91}_{-0.00}$	$0.00^{+0.58}_{-0.00}$	$0.00^{+11.92}_{-0.00}$	$0.00^{+0.33}_{-0.00}$	$0.00^{+1.04}_{-0.00}$
Fitted Diboson events	0.27 ± 0.11	1.59 ± 0.98	1.91 ± 0.77	1.70 ± 1.46	4.83 ± 2.60
MC exp. SM events	11.01	17.41	244.86	97.47	34.52
MC exp. Multijets events	0.00	0.00	193.93	0.00	0.00
MC exp. Wjets events	4.38	14.43	14.56	0.31	8.04
MC exp. Zjets events	0.02	0.15	8.57	0.04	19.42
MC exp. GAMMAjets events	0.00	0.00	0.00	94.73	0.00
MC exp. Top events	6.35	1.26	25.90	0.72	2.26
MC exp. Diboson events	0.27	1.57	1.89	1.68	4.81

Table C.2: **S**RD2 : Background fit results for the CRT, CRW, CRQ, CRY and SR regions, for an integrated luminosity of 36.1 fb^{-1}. Nominal MC expectations (normalized to MC cross-sections) are given for comparison. The errors shown are the statistical plus systematic uncertainties. The errors shown for the signal region are systematic uncertainties only.

SRD3 channel	CRT	CRW	CRQ	CRY	SR
Observed events	7	12	87	75	20
Fitted bkg events	$7.09^{+14.31}_{-7.09}$	11.98 ± 3.61	87.06 ± 17.68	74.97 ± 8.70	24.24 ± 7.82
Fitted Multijets events	0.00 ± 0.00	0.00 ± 0.00	79.93 ± 10.13	0.00 ± 0.00	$0.00^{+0.00}_{-0.00}$
Fitted Wjets events	3.95 ± 1.49	10.96 ± 3.42	4.13 ± 1.64	0.26 ± 0.09	5.73 ± 2.82
Fitted Zjets events	0.02 ± 0.01	0.10 ± 0.04	$2.09^{+4.94}_{-2.09}$	0.07 ± 0.02	14.10 ± 3.73
Fitted GAMMAjets events	0.00 ± 0.00	0.00 ± 0.00	0.00 ± 0.00	73.22 ± 8.73	0.00 ± 0.00
Fitted Top events	$0.03^{+15.26}_{-0.03}$	$0.01^{+2.56}_{-0.01}$	$0.04^{+17.28}_{-0.04}$	$0.00^{+0.92}_{-0.00}$	$0.01^{+5.85}_{-0.01}$
Fitted Diboson events	3.08 ± 1.25	0.91 ± 0.35	0.87 ± 0.47	$1.42^{+1.53}_{-1.42}$	4.39 ± 2.32
MC exp. SM events	15.51	15.30	83.57	93.36	31.64
MC exp. Multijets events	0.00	0.00	66.53	0.00	0.00
MC exp. Wjets events	4.70	13.00	4.90	0.31	6.80
MC exp. Zjets events	0.02	0.12	2.59	0.09	17.53
MC exp. GAMMAjets events	0.00	0.00	0.00	91.09	0.00
MC exp. Top events	7.68	1.28	8.68	0.46	2.93
MC exp. Diboson events	3.10	0.91	0.87	1.41	4.39

Table C.3: **S**RD3 : Background fit results for the CRT, CRW, CRQ, CRY and SR regions, for an integrated luminosity of 36.1 fb^{-1}. Nominal MC expectations (normalized to MC cross-sections) are given for comparison. The errors shown are the statistical plus systematic uncertainties. The errors shown for the signal region are systematic uncertainties only.

SRD4 channel	CRT	CRW	CRQ	CRY	SR
Observed events	59	38	25	153	76
Fitted bkg events	59.00 ± 7.68	38.00 ± 6.17	25.01 ± 5.34	153.00 ± 12.37	98.12 ± 33.87
Fitted Multijets events	0.00 ± 0.00	0.00 ± 0.00	17.21 ± 8.98	0.00 ± 0.00	0.00 ± 0.00
Fitted Wjets events	10.09 ± 3.67	28.85 ± 7.59	1.23 ± 0.64	0.47 ± 0.16	20.37 ± 5.87
Fitted Zjets events	0.10 ± 0.07	0.14 ± 0.04	0.56 ± 0.45	0.15 ± 0.04	35.92 ± 6.95
Fitted GAMMAjets events	0.00 ± 0.00	0.00 ± 0.00	0.00 ± 0.00	149.18 ± 12.39	0.00 ± 0.00
Fitted Top events	44.88 ± 9.16	5.59 ± 3.65	5.77 ± 5.59	0.92 ± 0.67	$32.05^{+33.22}_{-32.05}$
Fitted Diboson events	3.92 ± 1.53	3.42 ± 1.15	0.24 ± 0.15	2.28 ± 0.77	9.78 ± 4.27
MC exp. SM events	56.18	44.14	44.80	191.49	108.42
MC exp. Multijets events	0.00	0.00	37.20	0.00	0.00
MC exp. Wjets events	12.46	35.61	1.53	0.58	25.15
MC exp. Zjets events	0.13	0.18	0.71	0.19	45.18
MC exp. GAMMAjets events	0.00	0.00	0.00	187.62	0.00
MC exp. Top events	39.66	4.94	5.12	0.82	28.32
MC exp. Diboson events	3.92	3.42	0.24	2.28	9.78

Table C.4: **S**RD4 : Background fit results for the CRT, CRW, CRQ, CRY and SR regions, for an integrated luminosity of 36.1 fb^{-1}. Nominal MC expectations (normalized to MC cross-sections) are given for comparison. The errors shown are the statistical plus systematic uncertainties. The errors shown for the signal region are systematic uncertainties only.

SRD5 channel	CRT	CRW	CRY	SR
Observed events	37	39	161	85
Fitted bkg events	36.91 ± 6.07	39.11 ± 6.26	161.06 ± 12.69	83.85 ± 8.91
Fitted Multijets events	0.00 ± 0.00	0.00 ± 0.00	0.00 ± 0.00	0.00 ± 0.00
Fitted Wjets events	6.08 ± 1.78	27.87 ± 6.96	0.52 ± 0.16	16.16 ± 4.21
Fitted Zjets events	0.07 ± 0.02	0.32 ± 0.05	0.12 ± 0.04	44.26 ± 5.91
Fitted GAMMAjets events	0.00 ± 0.00	0.00 ± 0.00	157.16 ± 12.69	0.00 ± 0.00
Fitted Top events	29.64 ± 6.60	5.80 ± 2.28	0.57 ± 0.43	13.65 ± 5.40
Fitted Diboson events	1.12 ± 0.26	5.12 ± 0.68	2.68 ± 0.10	9.78 ± 3.05
MC exp. SM events	53.49	57.52	185.10	105.64
MC exp. Multijets events	0.00	0.00	0.00	0.00
MC exp. Wjets events	9.53	43.69	0.82	25.33
MC exp. Zjets events	0.07	0.36	0.14	50.87
MC exp. GAMMAjets events	0.00	0.00	180.63	0.00
MC exp. Top events	42.77	8.36	0.82	19.66
MC exp. Diboson events	1.12	5.11	2.68	9.78

Table C.5: **S**RD5 : Background fit results for the CRT, CRW, CRY and SR regions, for an integrated luminosity of 36.1 fb^{-1}. Nominal MC expectations (normalized to MC cross-sections) are given for comparison. The errors shown are the statistical plus systematic uncertainties. The errors shown for the signal region are systematic uncertainties only.

SRO1 channel	CRT	CRW	CRQ	CRY	SR
Observed events	4	6	53	19	6
Fitted bkg events	3.99 ± 1.82	6.04 ± 2.48	52.94 ± 7.28	18.97 ± 4.35	6.32 ± 2.27
Fitted Multijets events	0.00 ± 0.00	0.00 ± 0.00	48.96 ± 7.50	0.00 ± 0.00	$0.00^{+0.00}_{-0.00}$
Fitted Wjets events	1.91 ± 1.06	5.22 ± 2.71	1.50 ± 0.85	0.06 ± 0.03	2.29 ± 1.24
Fitted Zjets events	0.02 ± 0.01	0.09 ± 0.03	0.70 ± 0.42	0.01 ± 0.00	3.06 ± 1.97
Fitted GAMMAjets events	0.00 ± 0.00	0.00 ± 0.00	0.00 ± 0.00	18.87 ± 4.35	0.00 ± 0.00
Fitted Top events	$2.04^{+2.18}_{-2.04}$	$0.36^{+0.55}_{-0.36}$	$1.60^{+1.78}_{-1.60}$	$0.04^{+0.06}_{-0.04}$	$0.46^{+0.53}_{-0.46}$
Fitted Diboson events	0.02 ± 0.01	0.36 ± 0.09	0.19 ± 0.14	0.00 ± 0.00	0.50 ± 0.17
MC exp. SM events	13.51	8.73	46.24	33.75	11.20
MC exp. Multijets events	0.00	0.00	34.27	0.00	0.00
MC exp. Wjets events	2.27	6.21	1.78	0.07	2.72
MC exp. Zjets events	0.04	0.16	1.23	0.02	5.43
MC exp. GAMMAjets events	0.00	0.00	0.00	33.46	0.00
MC exp. Top events	11.19	2.00	8.78	0.21	2.55
MC exp. Diboson events	0.02	0.36	0.19	0.00	0.50

Table C.6: **SRO1** : Background fit results for the CRT, CRW, CRQ, CRY and SR regions, for an integrated luminosity of 36.1 fb^{-1}. Nominal MC expectations (normalized to MC cross-sections) are given for comparison. The errors shown are the statistical plus systematic uncertainties. The errors shown for the signal region are systematic uncertainties only.

SRO2 channel	CRT	CRW	CRQ	CRY	SR
Observed events	20	5	30	30	23
Fitted bkg events	20.00 ± 4.47	4.99 ± 2.23	29.97 ± 5.48	30.01 ± 5.48	17.38 ± 4.13
Fitted Multijets events	0.00 ± 0.00	0.00 ± 0.00	23.19 ± 8.24	0.00 ± 0.00	$0.00^{+0.00}_{-0.00}$
Fitted Wjets events	$1.26^{+1.29}_{-1.26}$	3.02 ± 2.99	$0.52^{+0.55}_{-0.52}$	$0.01^{+0.01}_{-0.01}$	$1.71^{+1.72}_{-1.71}$
Fitted Zjets events	0.01 ± 0.00	0.07 ± 0.02	0.30 ± 0.19	0.03 ± 0.01	4.88 ± 1.51
Fitted GAMMAjets events	0.00 ± 0.00	0.00 ± 0.00	0.00 ± 0.00	29.34 ± 5.50	0.00 ± 0.00
Fitted Top events	18.50 ± 4.81	$1.55^{+1.89}_{-1.55}$	5.94 ± 5.81	$0.37^{+0.45}_{-0.37}$	7.15 ± 3.19
Fitted Diboson events	0.23 ± 0.10	0.34 ± 0.10	0.02 ± 0.01	0.26 ± 0.16	3.63 ± 2.48
MC exp. SM events	26.32	11.51	28.19	58.59	27.01
MC exp. Multijets events	0.00	0.00	18.81	0.00	0.00
MC exp. Wjets events	3.83	9.19	1.58	0.03	5.21
MC exp. Zjets events	0.01	0.14	0.59	0.05	9.62
MC exp. GAMMAjets events	0.00	0.00	0.00	57.79	0.00
MC exp. Top events	22.24	1.85	7.20	0.45	8.57
MC exp. Diboson events	0.23	0.34	0.02	0.26	3.61

Table C.7: **SRO2** : Background fit results for the CRT, CRW, CRQ, CRY and SR regions, for an integrated luminosity of 36.1 fb^{-1}. Nominal MC expectations (normalized to MC cross-sections) are given for comparison. The errors shown are the statistical plus systematic uncertainties. The errors shown for the signal region are systematic uncertainties only.

SRO3 channel	CRT	CRW	CRY	SR
Observed events	52	10	49	37
Fitted bkg events	52.03 ± 7.31	10.07 ± 3.13	49.07 ± 7.01	34.53 ± 5.58
Fitted Multijets events	0.00 ± 0.00	0.00 ± 0.00	0.00 ± 0.00	0.00 ± 0.00
Fitted Wjets events	$1.22^{+2.01}_{-1.22}$	$3.58^{+5.80}_{-3.58}$	$0.02^{+0.04}_{-0.02}$	$2.28^{+3.72}_{-2.28}$
Fitted Zjets events	0.01 ± 0.01	0.06 ± 0.03	0.05 ± 0.01	9.54 ± 2.88
Fitted GAMMAjets events	0.00 ± 0.00	0.00 ± 0.00	47.81 ± 7.02	0.00 ± 0.00
Fitted Top events	50.20 ± 7.92	$4.58^{+7.45}_{-4.58}$	0.57 ± 0.44	20.05 ± 6.94
Fitted Diboson events	0.60 ± 0.33	1.85 ± 0.41	0.62 ± 0.13	2.65 ± 0.93
MC exp. SM events	47.13	20.06	79.85	44.09
MC exp. Multijets events	0.00	0.00	0.00	0.00
MC exp. Wjets events	4.85	14.23	0.08	9.06
MC exp. Zjets events	0.02	0.10	0.08	15.68
MC exp. GAMMAjets events	0.00	0.00	78.60	0.00
MC exp. Top events	41.66	3.89	0.47	16.70
MC exp. Diboson events	0.60	1.85	0.62	2.65

Table C.8: **SRO3** : Background fit results for the CRT, CRW, CRY and SR regions, for an integrated luminosity of 36.1 fb^{-1}. Nominal MC expectations (normalized to MC cross-sections) are given for comparison. The errors shown are the statistical plus systematic uncertainties. The errors shown for the signal region are systematic uncertainties only.

SRO4 channel	CRT	CRW	CRY	SR
Observed events	263	80	355	253
Fitted bkg events	262.93 ± 16.21	80.09 ± 8.96	354.99 ± 18.84	241.60 ± 42.99
Fitted Multijets events	0.00 ± 0.00	0.00 ± 0.00	0.00 ± 0.00	0.00 ± 0.00
Fitted Wjets events	14.19 ± 5.50	43.10 ± 13.60	0.42 ± 0.18	$33.69^{+35.73}_{-33.69}$
Fitted Zjets events	0.15 ± 0.04	0.56 ± 0.19	0.16 ± 0.04	66.46 ± 14.06
Fitted GAMMAjets events	0.00 ± 0.00	0.00 ± 0.00	347.97 ± 18.88	0.00 ± 0.00
Fitted Top events	246.01 ± 17.64	25.89 ± 8.79	2.68 ± 1.06	118.94 ± 20.75
Fitted Diboson events	2.59 ± 1.21	10.54 ± 4.20	3.76 ± 0.90	22.51 ± 8.76
MC exp. SM events	342.63	138.07	537.12	346.41
MC exp. Multijets events	0.00	0.00	0.00	0.00
MC exp. Wjets events	30.90	94.04	0.92	73.45
MC exp. Zjets events	0.23	0.85	0.24	101.02
MC exp. GAMMAjets events	0.00	0.00	528.84	0.00
MC exp. Top events	308.92	32.62	3.36	149.36
MC exp. Diboson events	2.59	10.55	3.77	22.57

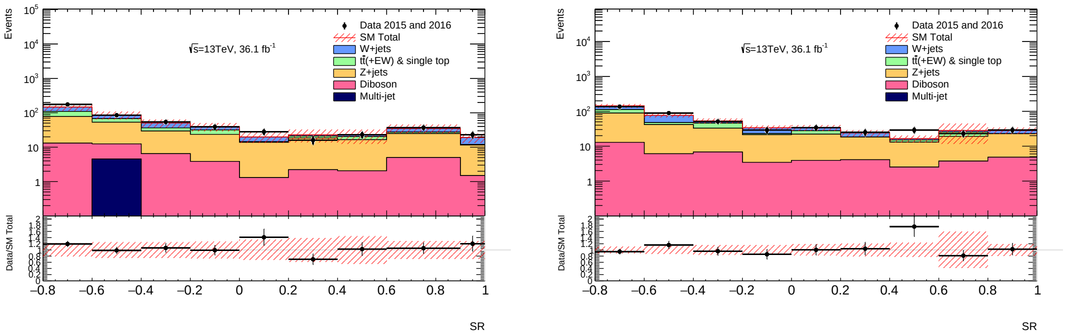
Table C.9: **SRO4** : Background fit results for the CRT, CRW, CRY and SR regions, for an integrated luminosity of 36.1 fb^{-1}. Nominal MC expectations (normalized to MC cross-sections) are given for comparison. The errors shown are the statistical plus systematic uncertainties. The errors shown for the signal region are systematic uncertainties only.

SRO5 channel	CRT	CRW	CRY	SR
Observed events	822	448	1694	1047
Fitted bkg events	821.93 ± 28.67	447.99 ± 21.17	1694.10 ± 41.16	1014.17 ± 80.67
Fitted Multijets events	0.00 ± 0.00	0.00 ± 0.00	0.00 ± 0.00	$0.00^{+0.00}_{-0.00}$
Fitted Wjets events	80.15 ± 18.45	298.20 ± 42.01	3.73 ± 0.74	205.93 ± 36.90
Fitted Zjets events	1.08 ± 0.34	3.05 ± 1.59	0.96 ± 0.17	409.43 ± 68.54
Fitted GAMMAjets events	0.00 ± 0.00	0.00 ± 0.00	1662.69 ± 41.50	0.00 ± 0.00
Fitted Top events	724.66 ± 35.76	101.26 ± 31.75	9.72 ± 3.35	317.57 ± 41.63
Fitted Diboson events	16.04 ± 4.81	45.48 ± 11.53	17.01 ± 4.77	81.24 ± 31.77
MC exp. SM events	996.87	620.67	2075.07	1271.28
MC exp. Multijets events	0.00	0.00	0.00	0.00
MC exp. Wjets events	121.06	451.14	5.65	311.52
MC exp. Zjets events	1.32	3.73	1.17	502.19
MC exp. GAMMAjets events	0.00	0.00	2039.74	0.00
MC exp. Top events	858.49	120.32	11.52	376.36
MC exp. Diboson events	16.01	45.48	16.99	81.22

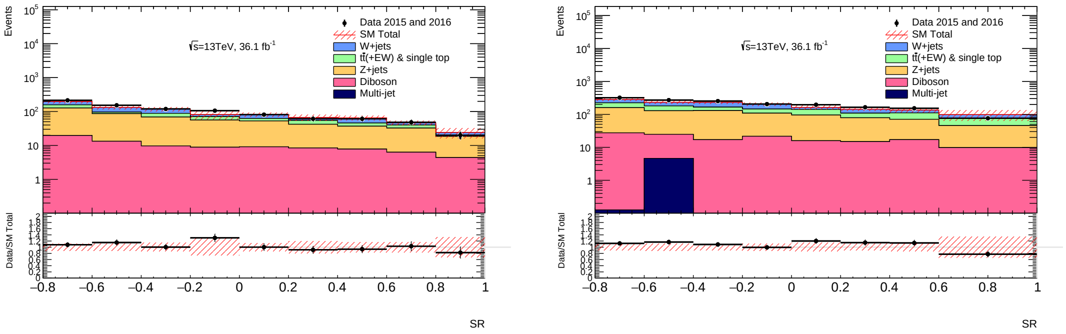
Table C.10: **S**RO5 : Background fit results for the CRT, CRW, CRY and SR regions, for an integrated luminosity of 36.1 fb^{-1}. Nominal MC expectations (normalized to MC cross-sections) are given for comparison. The errors shown are the statistical plus systematic uncertainties. The errors shown for the signal region are systematic uncertainties only.

C.3 Background-only fit in each BDT ranges

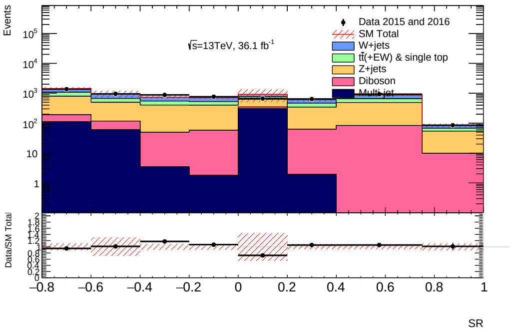
Figures C.2 and C.3 show comparisons of the data and background prediction in each BDT ranges of SR. The prediction is obtained by background-only fit in each BDT ranges. CRs in each fit is defined by the same BDT range as in SR.



(a) BDTD1 (b) BDTD2

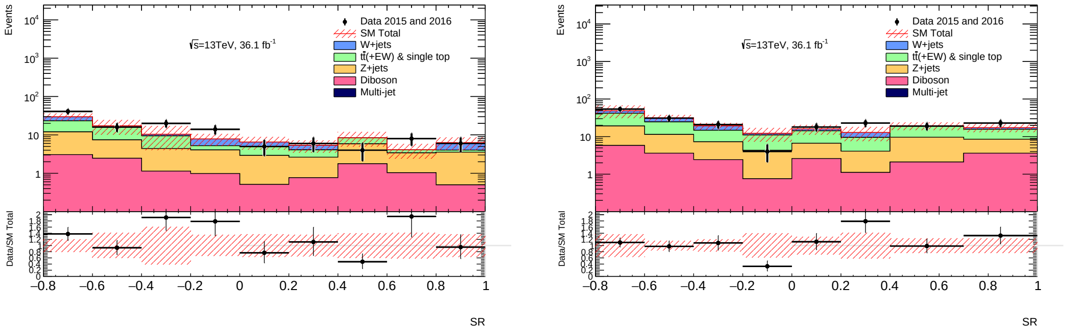


(c) BDTD3 (d) BDTD4

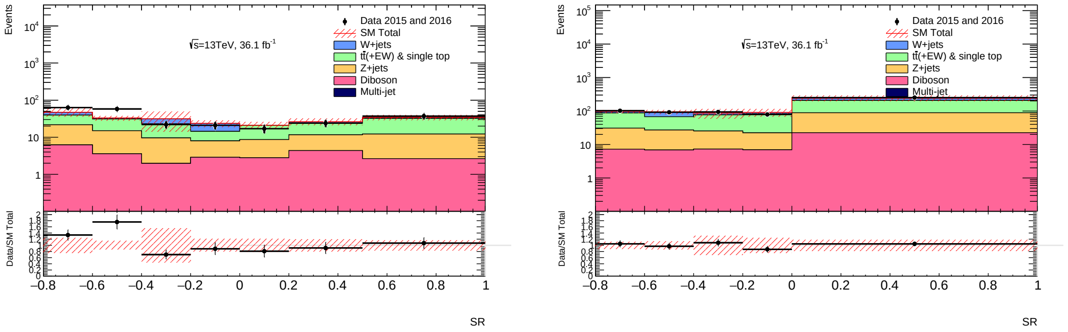


(e) BDTD5

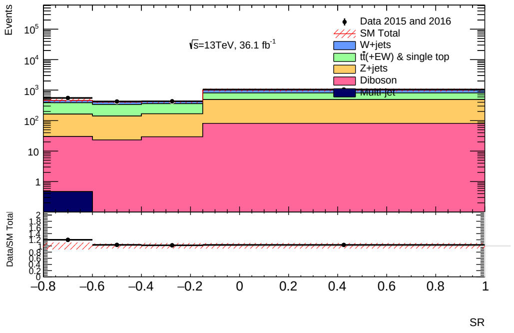
Figure C.2: Numbers of observed data and the SM background prediction in each BDT range of (a) D1, (b) D2, (c) D3, (d) D4, and (e) D5 BDT scores after the background-only fit in each BDT range. Each bin indicates an observed number of data (Black point) and a predicted number of the SM background (Color filled bars). Each color shows a different background process. The red shaded band to the SM total background represents its total uncertainties. The bottom panel shows the number of the data divided by the number of the total background prediction.



(a) BDTO1 (b) BDTO2



(c) BDTO3 (d) BDTO4



(e) BDTO5

Figure C.3: Numbers of observed data and the SM background prediction in each BDT range of (a) O1, (b) O2, (c) O3, (d) O4, and (e) O5 BDT scores after the background-only fit in each BDT range. Each bin indicates an observed number of data (Black point) and a predicted number of the SM background (Color filled bars). Each color shows a different background process. The red shaded band to the SM total background represents its total uncertainties. The bottom panel shows the number of the data divided by the number of the total background prediction.

Bibliography

[1] ATLAS Collaboration, "Summary of Standard Model cross section measurements compared to theoretical predictions", 2017, URL: `https://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/CombinedSummaryPlots/SM/index.html#ATLAS_a_SMSummary_TotalXsect`.

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

[3] I. J. R. Aitchison, "Supersymmetry and the MSSM: An Elementary Introduction", 2005, eprint: `arXiv:hep-ph/0505105`.

[4] U. Amaldi, W. de Boer, and H. Fürstenau, *Comparison of grand unified theories with electroweak and strong coupling constants measured at LEP*, Physics Letters B **260** (1991) p. 447, ISSN: 0370-2693, URL: `http://www.sciencedirect.com/science/article/pii/0370269391916418`.

[5] E. Corbelli and P. Salucci, *The extended rotation curve and the dark matter halo of M33*, Monthly Notices of the Royal Astronomical Society **311** (2000) p. 441, eprint: `/oup/backfile/content_public/journal/mnras/311/2/10.1046/j.1365-8711.2000.03075.x/2/311-2-441.pdf`, URL: `+%20http://dx.doi.org/10.1046/j.1365-8711.2000.03075.x`.

[6] Planck Collaboration et al., *Planck 2015 results - I. Overview of products and scientific results*, A&A **594** (2016) A1, URL: `https://doi.org/10.1051/0004-6361/201527101`.

[7] T. Han, Z. Liu, and A. Natarajan, *Dark matter and Higgs bosons in the MSSM*, Journal of High Energy Physics **2013** (2013) p. 8, ISSN: 1029-8479, URL: `https://doi.org/10.1007/JHEP11(2013)008`.

[8] ATLAS Collaboration, *Summary of the searches for squarks and gluinos using* $\sqrt{s} = 8$ *TeV pp collisions with the ATLAS experiment at the LHC*, Journal of High Energy Physics **2015** (2015) p. 54, ISSN: 1029-8479, URL: `https://doi.org/10.1007/JHEP10(2015)054`.

[9] P. Draper, G. Lee, and C. E. M. Wagner, *Precise estimates of the Higgs mass in heavy supersymmetry*, Phys. Rev. D **89** (5 2014) p. 055023, URL: `https://link.aps.org/doi/10.1103/PhysRevD.89.055023`.

[10] S. P. Martin, *A Supersymmetry primer*, (1997), [Adv. Ser. Direct. High Energy Phys.18,1(1998)], arXiv: `hep-ph/9709356 [hep-ph]`.

[11] A. M. Baldini et al., *Search for the lepton flavour violating decay $\mu^+ \rightarrow e^+\gamma$ with the full dataset of the MEG experiment*, The European Physical Journal C **76** (2016) p. 434, ISSN: 1434-6052, URL: `https://doi.org/10.1140/epjc/s10052-016-4271-x`.

[12] B. Fuks, M. Klasen, D. R. Lamprea, and M. Rothering, *Gaugino production in proton-proton collisions at a center-of-mass energy of 8 TeV*, JHEP **10** (2012) p. 081, arXiv: `1207.2159 [hep-ph]`.

[13] B. Fuks, M. Klasen, D. R. Lamprea, and M. Rothering, *Precision predictions for electroweak superpartner production at hadron colliders with Resummino*, Eur. Phys. J. C **73** (2013) p. 2480, arXiv: `1304.0790 [hep-ph]`.

[14] B. Fuks, M. Klasen, D. R. Lamprea, and M. Rothering, *Gaugino production in proton-proton collisions at a center-of-mass energy of 8 TeV*, JHEP **10** (2012) p. 081, arXiv: `1207.2159 [hep-ph]`.

[15] B. Fuks, M. Klasen, D. R. Lamprea, and M. Rothering, *Precision predictions for electroweak superpartner production at hadron colliders with Resummino*, Eur. Phys. J. C **73** (2013) p. 2480, arXiv: `1304.0790 [hep-ph]`.

[16] B. Fuks, M. Klasen, D. R. Lamprea, and M. Rothering, *Revisiting slepton pair production at the Large Hadron Collider*, JHEP **01** (2014) p. 168, arXiv: `1310.2621 [hep-ph]`.

[17] J. Alwall, M. Le, M. Lisanti and J. G. Wacker, *Searching for Directly Decaying Gluinos at the Tevatron*, Phys. Lett. B **666** (2008) p. 34, arXiv: `0803.0019 [hep-ph]`.

[18] J. Alwall, P. Schuster, and N. Toro, *Simplified Models for a First Characterization of New Physics at the LHC*, Phys. Rev. D **79** (2009) p. 075020, arXiv: `0810.3921 [hep-ph]`.

[19] D. Alves et al., *Simplified Models for LHC New Physics Searches*, J. Phys. G: Nucl. Part. Phys. **39** (2012) p. 105005, arXiv: `1105.2838 [hep-ph]`.

[20] L. Evans and P. Bryant, *LHC Machine*, Journal of Instrumentation **3** (2008) S08001, URL: `http://stacks.iop.org/1748-0221/3/i=08/a=S08001`.

[21] S. Myers, *The LEP Collider, from design to approval and commissioning*, John Adams' Lecture, Delivered at CERN, 26 Nov 1990, CERN, 1991, URL: `http://cds.cern.ch/record/226776`.

[22] E. Mobs, *The CERN accelerator complex. Complexe des accélérateurs du CERN*, (2016), General Photo, URL: `https://cds.cern.ch/record/2197559`.

[23] ATLAS Collaboration, *LHC Commissioning : Longer term LHC schedule*, 2017, URL: `http://lhc-commissioning.web.cern.ch/lhc-commissioning/schedule/LHC-long-term.htm`.

[24] ATLAS Collaboration, *The ATLAS Experiment at the CERN Large Hadron Collider*, Journal of Instrumentation **3** (2008) S08003, URL: `http://stacks.iop.org/1748-0221/3/i=08/a=S08003`.

[25] ATLAS Collaboration, *ATLAS magnet system: Technical Design Report, 1*, Technical Design Report ATLAS, CERN, 1997, URL: `https://cds.cern.ch/record/338080`.

[26] M. Capeans et al., "ATLAS Insertable B-Layer Technical Design Report", tech. rep. CERN-LHCC-2010-013. ATLAS-TDR-19, 2010, URL: `https://cds.cern.ch/record/1291633`.

[27] K. A. Olive et al., *Review of Particle Physics*, Chin. Phys. **C38** (2014) p. 090001.

[28] D. Banfi, M. Delmastro, and M. Fanti, *Cell response equalisation of the ATLAS electromagnetic calorimeter without the direct knowledge of the ionisation signals*, Journal of Instrumentation **1** (2006) P08001, URL: `http://stacks.iop.org/1748-0221/1/i=08/a=P08001`.

[29] ATLAS Collaboration, *Performance of the ATLAS trigger system in 2015*, The European Physical Journal C **77** (2017) p. 317, ISSN: 1434-6052, URL: `https://doi.org/10.1140/epjc/s10052-017-4852-3`.

[30] ATLAS Collaboration, *Luminosity Public Results Run2*, 2017, URL: `https://twiki.cern.ch/twiki/bin/view/AtlasPublic/LuminosityPublicResultsRun2`.

[31] A. Salzburger, "The ATLAS Track Extrapolation Package", tech. rep. ATL-SOFT-PUB-2007-005. ATL-COM-SOFT-2007-010, CERN, 2007, URL: `https://cds.cern.ch/record/1038100`.

[32] "Track Reconstruction Performance of the ATLAS Inner Detector at $\sqrt{s} = 13$ TeV", tech. rep. ATL-PHYS-PUB-2015-018, CERN, 2015, URL: `https://cds.cern.ch/record/2037683`.

[33] R. Frühwirth, *Application of Kalman filtering to track and vertex fitting*, Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment **262** (1987) p. 444, ISSN: 0168-9002, URL: `http://www.sciencedirect.com/science/article/pii/0168900287908874`.

[34] "Performance of primary vertex reconstruction in proton-proton collisions at $\sqrt{s}$ =7 TeV in the ATLAS experiment", tech. rep. ATLAS-CONF-2010-069, CERN, 2010, URL: `https://cds.cern.ch/record/1281344`.

[35] ATLAS Collaboration, *Topological cell clustering in the ATLAS calorimeters and its performance in LHC Run 1*, The European Physical Journal C **77** (2017) p. 490, ISSN: 1434-6052, URL: `https://doi.org/10.1140/epjc/s10052-017-5004-5`.

[36] M. Cacciari, G. P. Salam, and G. Soyez, *The anti- k t jet clustering algorithm*, Journal of High Energy Physics **2008** (2008) p. 063, URL: `http://stacks.iop.org/1126-6708/2008/i=04/a=063`.

[37] M. Cacciari and G. P. Salam, *Pileup subtraction using jet areas*, Physics Letters B **659** (2008) p. 119, ISSN: 0370-2693, URL: `http://www.sciencedirect.com/science/article/pii/S0370269307011094`.

[38] S. D. Ellis and D. E. Soper, *Successive combination jet algorithm for hadron collisions*, Phys. Rev. D **48** (7 1993) p. 3160, URL: `https://link.aps.org/doi/10.1103/PhysRevD.48.3160`.

[39] ATLAS Collaboration, *Jet energy scale measurements and their systematic uncertainties in proton-proton collisions at* $\sqrt{s} = 13$ TeV *with the ATLAS detector*, Phys. Rev. D **96** (7 2017) p. 072002, URL: `https://link.aps.org/doi/10.1103/PhysRevD.96.072002`.

[40] ATLAS Collaboration, *Missing Transverse Momentum Distribution and Performance in 2016 data*, 2016, URL: `https://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/PLOTS/JETM-2016-008/`.

[41] "Expected performance of missing transverse momentum reconstruction for the ATLAS detector at $\sqrt{s} = 13$ TeV", tech. rep. ATL-PHYS-PUB-2015-023, CERN, 2015, URL: `http://cds.cern.ch/record/2037700`.

[42] ATLAS Collaboration, *2015-2016 TST Systematics and Forward Pileup Suppression in MET*, 2017, URL: `https://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/PLOTS/JETM-2017-001/`.

[43] S. Moch, *Expectations at LHC from hard QCD*, Journal of Physics G: Nuclear and Particle Physics **35** (2008) p. 073001, URL: `http://stacks.iop.org/0954-3899/35/i=7/a=073001`.

[44] S. Dulat et al., *New parton distribution functions from a global analysis of quantum chromodynamics*, (2015), eprint: `arXiv:1506.07443`.

[45] A. Collaboration, "Search for squarks and gluinos in final states with jets and missing transverse momentum using 36 fb^{-1} of $\sqrt{s}$=13 TeV *pp* collision data with the ATLAS detector", 2017, eprint: `arXiv:1712.02332`.

[46] "Discrimination of Light Quark and Gluon Jets in *pp* collisions at $\sqrt{s} = 8$ TeV with the ATLAS Detector", tech. rep. ATLAS-CONF-2016-034, CERN, 2016, URL: `https://cds.cern.ch/record/2200202`.

[47] ATLAS Collaboration, *Light-quark and gluon jet discrimination in pp collisions at* $\sqrt{s} = 7$ TeV *with the ATLAS detector*, The European Physical Journal C **74** (2014) p. 3023, ISSN: 1434-6052, URL: `https://doi.org/10.1140/epjc/s10052-014-3023-z`.

[48] "Quark versus Gluon Jet Tagging Using Charged Particle Multiplicity with the ATLAS Detector", tech. rep. ATL-PHYS-PUB-2017-009, CERN, 2017, URL: `https://cds.cern.ch/record/2263679`.

[49] "Performance of quark/gluon discrimination in 8 TeV pp data", tech. rep. CMS-PAS-JME-13-002, CERN, 2013, URL: `https://cds.cern.ch/record/1599732`.

[50] H.-J. Yang, B. P. Roe, and J. Zhu, *Studies of boosted decision trees for MiniBooNE particle identification*, Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment **555** (2005) p. 370, ISSN: 0168-9002, URL: `http://www.sciencedirect.com/science/article/pii/S0168900205018322`.

[51] A. Collaboration, "Evidence for the $H \to b\bar{b}$ decay with the ATLAS detector", 2017, eprint: `arXiv:1708.03299`.

[52] A. Buckley et al., *LHAPDF6: parton density access in the LHC precision era*, The European Physical Journal C **75** (2015) p. 132, ISSN: 1434-6052, URL: `https://doi.org/10.1140/epjc/s10052-015-3318-8`.

[53] R. D. Ball et al., *Parton distributions for the LHC run II*, Journal of High Energy Physics **2015** (2015) p. 40, ISSN: 1029-8479, URL: `https://doi.org/10.1007/JHEP04(2015)040`.

[54] H.-L. Lai et al., *New parton distributions for collider physics*, Phys. Rev. D **82** (7 2010) p. 074024, URL: `https://link.aps.org/doi/10.1103/PhysRevD.82.074024`.

[55] L. A. Harland-Lang, A. D. Martin, P. Motylinski, and R. S. Thorne, *Parton distributions in the LHC era: MMHT 2014 PDFs*, The European Physical Journal C **75** (2015) p. 204, ISSN: 1434-6052, URL: `https://doi.org/10.1140/epjc/s10052-015-3397-6`.

[56] A. Collaboration, *Measurement of event shapes at large momentum transfer with the ATLAS detector in pp collisions at* $\sqrt{s} = 7\text{TeV}$, The European Physical Journal C **72** (2012) p. 2211, ISSN: 1434-6052, URL: `https://doi.org/10.1140/epjc/s10052-012-2211-y`.

[57] A. Collaboration, *Search for squarks and gluinos in final states with jets and missing transverse momentum using 36 fb*$^{-1}$ *of* $\sqrt{s} = 13$ *TeV pp collision data with the ATLAS detector*, (2017), ATLAS-CONF-2017-022, URL: `https://cds.cern.ch/record/2258145`.

[58] J. T. Linnemann, *Measures of Significance in HEP and Astrophysics*, (2003), eprint: `arXiv:physics/0312059`.

[59] K. Cranmer, *Statistical Challenges for Searches for New Physics at the LHC*, (2005), eprint: `arXiv:physics/0511028`.

[60] R. Brun and F. Rademakers, *ROOT — An object oriented data analysis framework*, Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment **389** (1997) p. 81, New Computing Techniques in Physics Research V, ISSN: 0168-9002, URL: `http://www.sciencedirect.com/science/article/pii/S016890029700048X`.

[61] ATLAS Collaboration, *Public Egamma Trigger Plots for Collision Data*, 2017, URL: `https://twiki.cern.ch/twiki/bin/view/AtlasPublic/EgammaTriggerPublicResults#Electron_and_photon_trigger_AN2`.

[62] G. Cowan, K. Cranmer, E. Gross, and O. Vitells,
Asymptotic formulae for likelihood-based tests of new physics,
The European Physical Journal C **71** (2011) p. 1554, ISSN: 1434-6052,
URL: `https://doi.org/10.1140/epjc/s10052-011-1554-0`.

[63] A. L. Read, *Presentation of search results: the CL s technique*,
Journal of Physics G: Nuclear and Particle Physics **28** (2002) p. 2693,
URL: `http://stacks.iop.org/0954-3899/28/i=10/a=313`.

[64] ATLAS Collaboration,
b-tagging efficiency calibration from 2015+2016 data on lepton+jets tt events, 2017,
URL: `http://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/PLOTS/FTAG-2017-003/`.

[65] E. Gerwick, S. Schumann, S. Höche, and S. Marzani, *Soft evolution of multi-jet final states*,
Journal of High Energy Physics **2015** (2015) p. 106, ISSN: 1029-8479,
URL: `https://doi.org/10.1007/JHEP02(2015)106`.

[66] S. Catani, F. Krauss, B. R. Webber, and R. Kuhn, *QCD Matrix Elements + Parton Showers*,
Journal of High Energy Physics **2001** (2001) p. 063,
URL: `http://stacks.iop.org/1126-6708/2001/i=11/a=063`.

[67] "Identification of boosted, hadronically-decaying W and Z bosons in $\sqrt{s}$ = 13 TeV Monte Carlo Simulations for ATLAS", tech. rep. ATL-PHYS-PUB-2015-033, CERN, 2015,
URL: `http://cds.cern.ch/record/2041461`.

[68] ATLAS Collaboration,
"Commissioning of the ATLAS b-tagging algorithms using $t\bar{t}$ events in early Run-2 data",
tech. rep. ATL-PHYS-PUB-2015-039, CERN, 2015,
URL: `https://cds.cern.ch/record/2047871`.

[69] "Secondary vertex finding for jet flavour identification with the ATLAS detector",
tech. rep. ATL-PHYS-PUB-2017-011, CERN, 2017,
URL: `https://cds.cern.ch/record/2270366`.

[70] "Expected performance of the ATLAS b-tagging algorithms in Run-2",
tech. rep. ATL-PHYS-PUB-2015-022, CERN, 2015,
URL: `https://cds.cern.ch/record/2037697`.

[71] G. Piacquadio and C. Weiser,
A new inclusive secondary vertex algorithm for b-jet tagging in ATLAS,
Journal of Physics: Conference Series **119** (2008) p. 032032,
URL: `http://stacks.iop.org/1742-6596/119/i=3/a=032032`.

[72] ATLAS Collaboration,
b-tagging performance plots in a ttbar-dominated sample from early 2016 ATLAS data,
2016, URL: `http://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/PLOTS/FTAG-2016-001/`.

[73] W. Lampl et al., "Calorimeter Clustering Algorithms: Description and Performance",
tech. rep. ATL-LARG-PUB-2008-002. ATL-COM-LARG-2008-003, CERN, 2008,
URL: `https://cds.cern.ch/record/1099735`.

[74] "Improved electron reconstruction in ATLAS using the Gaussian Sum Filter-based model for bremsstrahlung", tech. rep. ATLAS-CONF-2012-047, CERN, 2012, URL: `https://cds.cern.ch/record/1449796`.

[75] "Electron efficiency measurements with the ATLAS detector using the 2015 LHC proton-proton collision data", tech. rep. ATLAS-CONF-2016-024, CERN, 2016, URL: `https://cds.cern.ch/record/2157687`.

[76] "Photon identification in 2015 ATLAS data", tech. rep. ATL-PHYS-PUB-2016-014, CERN, 2016, URL: `https://cds.cern.ch/record/2203125`.

[77] ATLAS Collaboration, *Electron and photon energy calibration with the ATLAS detector using LHC Run 1 data*, The European Physical Journal C **74** (2014) p. 3071, ISSN: 1434-6052, URL: `https://doi.org/10.1140/epjc/s10052-014-3071-4`.

[78] J. Illingworth and J. Kittler, *A survey of the hough transform*, Computer Vision, Graphics, and Image Processing **44** (1988) p. 87, ISSN: 0734-189X, URL: `http://www.sciencedirect.com/science/article/pii/S0734189X88800331`.

[79] ATLAS Collaboration, *Muon reconstruction performance of the ATLAS detector in proton–proton collision data at* $\sqrt{s}$ *=13 TeV*, The European Physical Journal C **76** (2016) p. 292, ISSN: 1434-6052, URL: `https://doi.org/10.1140/epjc/s10052-016-4120-y`.