SEARCH FOR CHARGED-LEPTON FLAVOR VIOLATION IN THE PRODUCTION AND DECAY OF TOP QUARKS AT $\sqrt{S} = 13$ TEV WITH THE CMS DETECTOR

JINGYAN LI

B.S. in Physics, Lanzhou University

A DISSERTATION SUBMITTED IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF DOCTOR OF PHILOSOPHY

DEPARTMENT OF PHYSICS
NORTHEASTERN UNIVERSITY

2023

Abstract

A search for charged-lepton flavor violation has been performed in the top quark sector through both top quark production and decay signal processes. The data were collected by the CMS experiment from proton-proton collisions at a center-of-mass energy of 13 TeV and correspond to an integrated luminosity of 138 fb^{-1}. The selected events are required to contain one opposite-sign electron-muon pair, a third charged lepton (electron or muon), at least one jet, and at most one jet associated with a bottom quark. The analysis utilizes boosted decision trees to separate background processes from a possible signal, exploiting differences in the kinematics of the final state particles. The data are found to be consistent with the standard model expectation. Upper limits at 95% confidence level are placed on the branching fractions involving up (charm) quarks, t$\to$ eμu (t$\to$ eμc), of 0.032×10^{-6} (0.498×10^{-6}), 0.022×10^{-6} (0.369×10^{-6}), and 0.012×10^{-6} (0.216×10^{-6}) for tensor, vector, and scalar interactions, respectively.

Contents

Contents

List of Figures

List of Tables

List of Acronyms

ATLAS A Toroidal LHC Apparatus

BDT Boosted Decision Tree

CLFV Charged-Lepton Flavor Violation

CMS Compact Muon Solenoid

EFT Effective Field Theory

HLT High Level Trigger

HL-LHC High Luminosity-LHC

LHC Large Hadron Collider

LO leading order

L1 Level-1

MC Monte-Carlo

ME matrix-element

NLO next-to-leading order

NN neural network

NNLO next-to-next-to-leading order

PD Primary Dataset

PDF Parton Distribution Function

PF particle flow

PU pile-up

PV primary vertex

QCD quantum chromodynamics

SM Standard Model

SV secondary vertex

WC Wilson Coefficient

CHAPTER 1

Introduction

Part I

Theoretical Framework

CHAPTER 2

Electroweak Standard Model

2.1 Gauge Theory

2.2 Higgs Mechanism

2.3 Flavor Sector

CHAPTER 3

Beyond the Standard Model

3.1 Overview of the BSM Theories

3.2 Leptoquark Model

CHAPTER 4

Effective Field Theory

4.1 Standard Model EFT

4.2 Low Energy EFT

Event Simulation for Colliders

5.1 Perturbative Quantum Chromodynamics

5.2 Nonperturbative Quantum Chromodynamics

Part II

Experimental Apparatus

CHAPTER 6

The Large Hadron Collider

CHAPTER 7

The Compact Muon Solenoid Detector

CHAPTER 8

Object Reconstruction in the CMS detector

8.1 Electron and Photon

8.2 Muon

8.3 Tau

8.4 Jets

8.5 Energy Sums

Run-3 Operation of the CMS detector

9.1   The CMS Run Organisation

9.2   Central Shifting Crew

9.3   Detector On-call Expert

CHAPTER 10

The Phase-2 Upgrade of the CMS Detector

10.1   Overview of the Upgrade

10.2   Leve-1 Track Finder

10.3   Leve-1 Electron Trigger Algorithm

Part III

Search for Flavor-Violating eµtq Interactions

CHAPTER 11

Previous searches

11.1 ATLAS

11.2 CMS

[1]

CHAPTER 12

Datasets, Simulated Samples and Triggers

This analysis is based on data collected by the Compact Muon Solenoid (CMS) experiment in 2016-2018 from pp collisions at a center-of-mass energy of 13 TeV corresponding to an integrated luminosity of 138 fb^{-1}. There were approximately 30 simultaneous pp collisions occurring per 25 ns. Based on online selection criteria, fully reconstructed collision data that contain high-level physics objects are divided into "Primary Datasets (PDs)". The PDs that make use of lepton information for selection include "DoubleEG", "DoubleMu", "MuonEG", "SingleElectron", and "SingleMuon" for 2016 and 2017 data taking era. In 2018, "SingleElectron", "DoubleEG" are replaced by "EGamma". The names of these PDs reflect the selection criteria. In addition to these PDs, MC samples are also generated to model both signal and background processes, which are described in section 12.1 and section 12.2, respectively. To account for the different data taking conditions across the years, all MC samples are generated separately for each year. High Level Trigger (HLT) triggers are used to select events offline, which is described in section 12.3.

12.1 Signal Samples

In this analysis, New Physics is described by Dimension-6 Effective Field Theory (EFT) operators,

$$\mathcal{L} = \mathcal{L}_{\text{SM}}^{(4)} + \frac{1}{\Lambda^2} \sum_a C_a^{(6)} O_a^{(6)} + O\left(\frac{1}{\Lambda^4}\right). \tag{12.1}$$

Among many of the Dimension-6 operators in Warsaw basis [2], a total of 6 operators are considered, which are summarized in Table 12.1. To reduce the number of free parameters, the permutations of fermion flavors are combined. Taking eμtu vertex as an example, the Wilson Coefficients (WCs) are parameterized in the following way:

$$C_{\text{lq}} = C_{\text{lq}}^{(1)1213} + C_{\text{lq}}^{(1)2113} + C_{\text{lq}}^{(1)1231} + C_{\text{lq}}^{(1)1213}, \tag{12.2}$$

$$C_{\text{lu}} = C_{\text{lu}}^{1213} + C_{\text{lu}}^{2113} + C_{\text{lu}}^{1231} + C_{\text{lu}}^{1213}, \tag{12.3}$$

$$C_{\text{eq}} = C_{\text{eq}}^{1213} + C_{\text{eq}}^{2113} + C_{\text{eq}}^{1231} + C_{\text{eq}}^{1213}, \tag{12.4}$$

$$C_{\text{eu}} = C_{\text{eu}}^{1213} + C_{\text{eu}}^{2113} + C_{\text{eu}}^{1231} + C_{\text{eu}}^{1213}, \tag{12.5}$$

Table 12.1: Summary of relevant dimension-6 operators considered in this analysis. Here, ε is the two dimensional Levi-Civita symbol, γ^μ the gamma matrix, and $\sigma^{\mu\nu} = \frac{i}{2}[\gamma^\mu, \gamma^\nu]$. The l and q denote left-handed doublets, whereas u and e denote right-handed singlets. The indices i and j are lepton flavor indices that run from 1 to 2 with $i \neq j$; m and n are quark flavor indices with the condition that one of them is 3 and the other one is 1 or 2.

Lorentz Structure	Operator		
vector	$O_{lq}^{(1)ijmn}$	$=$	$(\bar{l}_i\gamma^\mu l_j)(\bar{q}_m\gamma_\mu q_n)$
	O_{lu}^{ijmn}	$=$	$(\bar{l}_i\gamma^\mu l_j)(\bar{u}_m\gamma_\mu u_n)$
	O_{eq}^{ijmn}	$=$	$(\bar{e}_i\gamma^\mu e_j)(\bar{q}_m\gamma_\mu q_n)$
	O_{eu}^{ijmn}	$=$	$(\bar{e}_i\gamma^\mu e_j)(\bar{u}_m\gamma_\mu u_n)$
scalar	$O_{lequ}^{(1)ijmn}$	$=$	$(\bar{l}_i e_j)\, \varepsilon\, (\bar{q}_m u_n)$
tensor	$O_{lequ}^{(3)ijmn}$	$=$	$(\bar{l}_i \sigma^{\mu\nu} e_j)\, \varepsilon\, (\bar{q}_m \sigma_{\mu\nu} u_n)$

$$C_{lequ}^{(1)} = C_{lequ}^{(1)1213} + C_{lequ}^{(1)2113} + C_{lequ}^{(1)1231} + C_{lequ}^{(1)1213}, \tag{12.6}$$

$$C_{lequ}^{(3)} = C_{lequ}^{(3)1213} + C_{lequ}^{(3)2113} + C_{lequ}^{(3)1231} + C_{lequ}^{(3)1213}. \tag{12.7}$$

Additionally, all vector-like operators are combined,

$$O_{e\mu tu}^{vector} = O_{lq} + O_{lu} + O_{eq} + O_{eu}, \tag{12.8}$$

$$O_{e\mu tu}^{scalar} = O_{lequ}^{(1)} + h.c, \tag{12.9}$$

$$O_{e\mu tu}^{tensor} = O_{lequ}^{(3)} + h.c, \tag{12.10}$$

which results in 6 independent WCs: $C_{e\mu tu}^{vector}$, $C_{e\mu tu}^{scalar}$, $C_{e\mu tu}^{tensor}$, $C_{e\mu tc}^{vector}$, $C_{e\mu tc}^{scalar}$, $C_{e\mu tc}^{tensor}$.

To generate signal MC samples, the effective Lagrangian described above is implemented using the SmeftFR v2 [3] model, and saved in the "UFO" format [4]. Additionally, all the WCs are set to 1 with $\Lambda = 1$ TeV in the UFO, which then interfaces with the FEYNRULES [5] package to calculate Feynman diagrams. The output of the FEYNRULES is used in ME event generator MADGRAPH5_AMC@NLO v2.4.2 [6] to generate events at leading order (LO).

In general, the calculations done by the ME event generators are model-agnostic assuming the same EFT configurations. In other words, models like SmeftFR or SMEFTsim [7] are expected to give the same or very similar results in terms of cross sections and four-momenta of final-
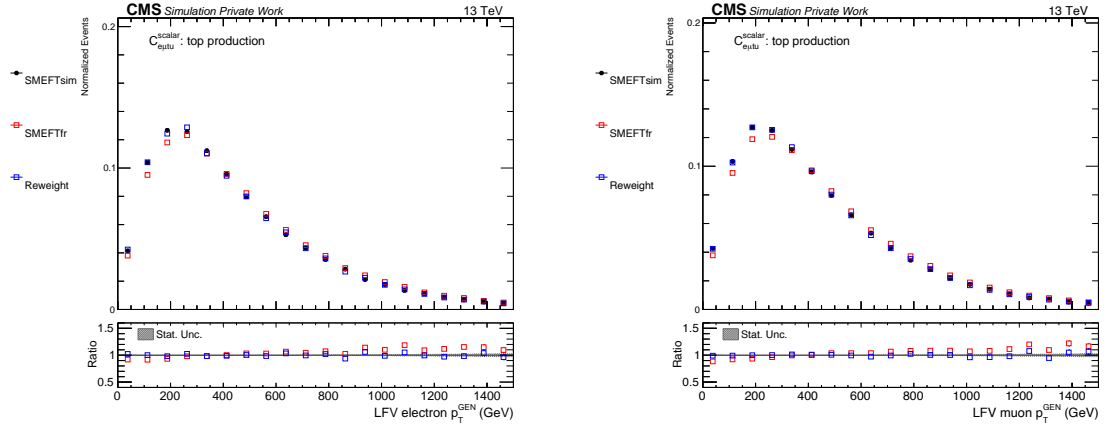
Figure 12.1: Comparison of kinematic distributions at ME-level produced by different models: LFV electron p_T (left), LFV muon p_T (right). The "SmeftFR" samples (shown in red curve) and "SMEFTsim" samples (shown in black curve) are statistically independent of each other. The "Reweight" (shown in blue curve) are produced by applying weights calculated by Equation 12.11 to "SmeftFR" samples.

state particles. Nevertheless, visible differences in kinematics have been observed and shown in Figure 12.1. Furthermore, the cross sections predicted by SmeftFR v2 also yield more than 20% difference relative to SMEFTsim due to a bug that was later fixed in SmeftFR v3. In the light of these differences, the CMS and A Toroidal LHC Apparatus (ATLAS) Collaborations agreed to adopt the SMEFTsim model as the common standard. To quantify the impact of the choice of models on kinematics, the following ratio is calculated for each event i,

$$R^i_{\mathrm{reweight}} = \frac{\omega^i_{\mathrm{SMEFTsim}}}{\omega^i_{\mathrm{SmeftFR}}}, \tag{12.11}$$

where ω^i_X is the per-event ME weight calculated by MADGRAPH5_AMC@NLO using model X. Since SMEFTsim was not used by CMS at the time when the signal samples were generated, R_{reweight} are used to "reweight" the original samples generated using SmeftFR.

The cross sections for top production signals are taken directly from MADGRAPH5_AMC@NLO with SMEFTsim UFO as input. The event generation for top decay signals at the ME-level take two steps: (i) production of the SM $t\bar{t}$, and (ii) CLFV decay of one of the top quarks. Therefore, the $t\bar{t}$ cross section at next-to-next-to-leading order (NNLO) precision [8] is used to normalize the top decay signals,

$$\sigma^{\mathrm{Top\ Decay}}_{\mathrm{CLFV}} = 2 \times \sigma^{\mathrm{NNLO}}_{t\bar{t}} \times \mathcal{B}(t \to e\mu q), \tag{12.12}$$

where q={u,c}, and $\mathcal{B}(t \to e\mu q)$ [9] can be expressed as,

$$\mathcal{B}(t \rightarrow e\mu q) = \begin{array}{cc} \dfrac{|C_{e\mu tq}^{vector}|^2}{\Lambda^4} & \dfrac{m_t^5}{384\pi^3\Gamma_t^{SM}} \\[2ex] \dfrac{|C_{e\mu tq}^{scalar}|^2}{\Lambda^4} & \dfrac{m_t^5}{3072\pi^3\Gamma_t^{SM}} \\[2ex] \dfrac{|C_{e\mu tq}^{tensor}|^2}{\Lambda^4} & \dfrac{m_t^5}{64\pi^3\Gamma_t^{SM}} \end{array} \tag{12.13}$$

where m_t and Γ_t^{SM} are taken to be 172.5 GeV and 1.33 GeV in this analysis, respectively. The choice of u or c quark in final states does not affect the cross sections of the top decay signals. The cross sections for all signal MC samples are summarized in Table 12.2.

Table 12.2: Theoretical cross sections for top production and decay for each CLFV coupling, calculated at $C/\Lambda^2 = 1$ TeV^{-2}. Uncertainties related to PDF and QCD scale in ME calculation are given ($\sigma_{-\text{scale}}^{+\text{scale}} \pm$ PDF).

Lorentz Structure	Samples	XS (fb)
vector	top production via u quark	$460^{+81}_{-64} \pm 6$
	top production via c quark	$33^{+5}_{-4} \pm 6$
	top decay via u/c quark	$32^{+0.8}_{-1.1} \pm 1.3$
scalar	top production via u quark	$97^{+18}_{-14} \pm 1$
	top production via c quark	$6.3^{+0.9}_{-0.8} \pm 1.4$
	top decay via u/c quark	$4.0^{+0.1}_{-0.1} \pm 0.2$
tensor	top production via u quark	$2143^{+368}_{-293} \pm 31$
	top production via c quark	$164^{+22}_{-18} \pm 27$
	top decay via u/c quark	$187^{+5}_{-6} \pm 8$

Steps other than the ME calculation concerning signal MC generation follow the CMS standard, which is described in the following section.

12.2 Background Samples

Besides tZq, tHq, tHW, and tWZ processes, the next-to-leading order (NLO) Parton Distribution Function (PDF) set from NNPDF3.0 [10] is used in 2016 to generate background MC samples. The NNLO PDF set from NNPDF3.1 [11] is used for tZq while the LO PDF set from NNPDF3.0 is used for tHq, tHW, and tWZ in 2016. In 2017 and 2018, the NNLO PDF set from NNPDF3.1 is used to generate all the samples.

The default choice of ME event generator is MADGRAPH5_AMC@NLO v2.4.2 (v2.2.2 for 2016), which is used to generate all but ZZ, t$\bar{t}$H, and t$\bar{t}$ samples. These three samples are

Table 12.3: Summary of the configurations of the MC samples. DYM50 (DYM10to50) denote DY sample with a dilepton invariant mass greater than 50 GeV (between 10 and 50 GeV). V includes W and Z bosons. The cross sections for samples without a citation are taken directly from their event generators.

Category	Process	Event Generator	Perturbative QCD	Tune	XS precision
prompt background	WZ	MadGraph	NLO	CUETP8M1(CP5)	NLO [16]
	ZZ	powheg	NLO	CUETP8M1(CP5)	NLO [16]
	VVV	MadGraph	NLO	CUETP8M1(CP5)	NLO
	$t\bar{t}W$, $t\bar{t}Z$	MadGraph	NLO	CUETP8M1(CP5)	NLO [17, 18]
	$t\bar{t}H$	powheg	NLO	CUETP8M1(CP5)	NLO [18]
	tZq	MadGraph	NLO	CP5	NLO
	tHq, tHW, tWZ	MadGraph	LO	CUETP8M1(CP5)	LO
nonprompt background	$t\bar{t}$	powheg	NLO	CUETP8M1(CP5)	NNLO [8]
	DYM50	MadGraph	NLO	CUETP8M1(CP5)	NNLO [19]
	DYM10to50	MadGraph	LO	CUETP8M1(CP5)	NLO [19]

generated with powheg v2 [12] instead. Samples with small contributions (tHq, tWZ, tHW, and low mass DY) are generated at LO while other samples are generated at NLO. Whenever possible and relevant, theoretical cross sections from high-order quantum chromodynamics (QCD) calculations are used. The references of these calculations are included in 12.3.

The pythia v8.2 [13] is used to model parton shower and hadronization. The CUETP8M1 [14] is used in 2016 for underlying event tuning while the CP5 [15] is used in 2017 and 2018. The configurations of the MC samples are summarized in Table 12.3. The background processes are divided into two categories: (i) processes with three or more *prompt* leptons in the final states are classified as "*prompt* background", and (ii) other processes are classified as "*nonprompt* background". The *nonprompt* backgrounds in this analysis are modeled with a data-driven technique, which is discussed in chapter 15. The MC samples listed in "*nonprompt*" category in Table 12.3 are therefore only used for the purpose of validations.

12.3 Triggers

The target final states of this analysis contain three prompt leptons, which make lepton triggers the most optical choice to select events. To achieve the maximum acceptance, a combination of single-lepton, di-lepton and tri-lepton triggers are used. These triggers are summarized in Appendix A. Events in simulated samples are required to fire at least one of the triggers listed in Table A.1-A.3. Since multple PDs are used to record data events and the orthogonality of these

PDs are not guaranteed by the online selection criteria, the following trigger logic is implemented to remove the overlap between different PDs:

- Events in SingleMuon datasets are required to fire at least one of the triggers listed under "SingleMuon".

- Events in DoubleMuon datasets are required to fire at least one of the triggers listed under "DoubleMu". Events are removed if they also fire at least one of the triggers listed under "SingleMuon".

- Events in MuonEG' datasets are required to fire at least one of the triggers listed under "MuonEG". Events are removed if they also fire at least one of the triggers listed under "SingleMuon" or "DoubleMu".

- Events in Single Electron datasets are required to fire at least one of the triggers listed under "SingleElectron". Events are removed if they also fire at least one of the triggers listed under "SingleMuon" or "DoubleMu" or "MuonEG".

- Events in DoubleEG datasets are required to fire at least one of the triggers listed under "DoubleEG". Events are removed if they also fire at least one of the triggers listed under "SingleMuon" or "DoubleMu" or "MuonEG" or "SingleElectron".

- Events in EGamma datasets are required to fire at least one of the triggers listed under "EGamma". Events are removed if they also fire at least one of the triggers listed under "SingleMuon" or "DoubleMu" or "MuonEG".

CHAPTER 13

Object Selection

Objects described in chapter 8, referred to as "candidates", are further selected with more stringent requirements with the goal of suppressing the contributions from background processes while maintaining a high signal acceptance. In particular, prompt electron and muon candidates are identified through a custom-trained Boosted Decision Tree (BDT) classifier, which is discussed in section 13.1. Two jet identification algorithms are deployed to select jet candidates originating from hard collisions, which is discussed in section 13.2. Furthermore, jet candidates that originate from b quarks are identified with a neural network (NN) based algorithm, which is discussed in section 13.3.

13.1 Lepton Selection

The target final states of this analysis feature exactly three leptons that originate either from decays of electroweak bosons or from the CLFV interaction, which in this analysis is a contact interaction that involves four fermions. These leptons, referred to as *prompt* leptons, typically appear to be isolated and not far away from the primary vertex (PV). In contrast, *nonprompt* leptons are leptons that originate from decays of hadrons, or from photon conversions, or misidentified leptons. They often travel a noticeable distance away from the PV and appear to be less isolated due to nearby activities. Due to the high multiplicity of leptons in our selection, backgrounds with at least one *nonprompt* lepton outnumber any other Standard Model (SM) processes that produce three or more *prompt* leptons. It is therefore crucial to exploit the differences between *nonprompt* and *prompt* leptons and bring the *nonprompt* background under control.

13.1.1 TOP LeptonMVA

The TOP LEPTONMVA is an offline lepton identification algorithm that is originally developed for tZq analyses [20, 21]. It is based on Gradient BDT implemented using the TMVA package [22]. A total of 13 features are used as input to the BDT. They can be categorised into four groups: (i) positions and momenta of the lepton candidates, (ii) isolation variables, (iii) variables associated to the closest jet, and (iv) a quality variable that is specific to electron or muon candidate.

The version of TOP LEPTONMVA used by this analysis is the same as [21], where a detailed description of all input features can be found.

Prompt leptons from t$\bar{\text{t}}$W, t$\bar{\text{t}}$Z, and tZq samples are used as signals in the BDT training while *nonprompt* leptons from t$\bar{\text{t}}$ samples are used as backgrounds. The trained BDT outputs a single score for each lepton candidate ranging from -1 to 1 with -1 (1) being the most background- (signal-) like. the tight working point with a threshold of ($>$) 0.9 is chosen as the selection criteria for both electron candidates and muon candidates, which corresponds to a signal(background) efficiency of 90%(1%). The strategy is to trade a small percentage ($<$10%) of signal efficiency for several factors of background rejection.

13.1.2 Full Selection

In addition to the TOP LEPTONMVA requirement, a set of common selection criteria are applied to both electron and muon candidates. The minimum p_T requirement is 38GeV, 20 GeV, and 20 GeV for the leading, sub-leading, and trailing lepton in p_T, respectively. This requirement is driven by the p_T thresholds of the HLT triggers to avoid inefficiency at turn-on. Electron and muon candidates are required to be in the pseudorapidity range $|\eta| <$2.4, which corresponds to the acceptance of CMS tracker and muon system in 2016-2018. The transverse (longitudinal) impact parameters with respect to the PV, denoted as d_{xy} (d_z), is required to be in the range $|d_{xy}| < 0.05$ cm ($|d_z| < 0.05$ cm). The significance of the 3-dimensional impact parameter, denoted as SIP_3, is defined as the 3-dimensional impact parameter divided by its uncertainty. It is required that $\text{SIP}_3 < 8$. The three cuts on impact parameters are added due to the difference in distributions of these parameters between *prompt* and *nonprompt* leptons. Also, they are part of the pre-selection requirement in the BDT training.

Furthermore, all lepton candidates are required to be isolated. This is achieved by first defining a cone with a distance parameter of ΔR around each lepton candidate, where $\Delta R = \sqrt{\Delta\eta^2 + \Delta\phi^2}$. Only particles within $\Delta R < R_{\text{max}}$ can contribute to the isolation variable, where R_{max} is referred to as the size of the cone. Secondly, a particle flow (PF) based isolation variable is defined as,

$$I_{\text{mini}}^{\text{rel}} = \frac{1}{p_T^{\ell}}\{ \sum_{\text{charged}} p_T + \max(0, \sum_{\text{neutral}} p_T - \rho\mathcal{A}[\frac{\Delta R}{0.3}]^2)\}, \tag{13.1}$$

where p_T^{ℓ} is the p_T of the lepton candidate, the first term inside the curly braces is the scalar sum of all charged particles associated with the PV while the second term evaluates the contribution from neutral particles. This is done by first scalar-summing over p_T of all neutral particles associated to the PV. A correction term, known as effective area correction [23], is then subtracted. This term is used to mitigate the impact of pile-up (PU) interactions. The size of the cone scales with p_T^{ℓ} as,

$$R_{\mathrm{max}} = \mathrm{max}(0.05, \mathrm{min}(0.2, \frac{10\mathrm{GeV}}{p_{\mathrm{T}}^{\ell}})). \tag{13.2}$$

.

This type of isolation variable is known as "mini" isolation, which maximises the signal efficiency at p_{T}^{ℓ} by reducing the cone size. It is required that lepton candidates to have $I_{\mathrm{mini}}^{\mathrm{rel}} < 0.12$.

For electron specifically, candidates are required to have a GSF track with one of less missing inner hit. Electron candidates reconstructed in the transition region between ECAL barrel and endcap (i.e. $1.44 < |\eta_{\mathrm{SC}}| < 1.57$) are removed from consideration. For muon specifically, candidates are required to be PF muons and pass the medium working point discussed in section 8.2.

Lepton candidates that pass all requirements stated above are referred to as "*tight*" leptons. Leptons selected with a separate set of criteria, known as "*loose*", is used in estimating the *nonprompt* background, which is discussed in chapter 15. Unless explicitly stated, all lepton objects presented in this search are *tight* leptons.

The energy of electron candidates are calibrated through. The moment scale of muon candidates are calibrated for muon candidates with $p_{\mathrm{T}} < 200$ GeV. Scale factors are applied to *tight* leptons to correct for the differences in reconstruction, isolation, and identification between data and MC. These scale factors are obtained using dilepton events in Z resonance window.

13.2 Jet Selection

Jet candidates are reconstructed from PF candidates using the anti-k_{T} algorithm described in section 13.2 with a cone size of 0.4. Charged hadrons that are not associated to the PV are removed. Jet candidates are required to have a minimum p_{T} of 30 GeV and in in the pseudorapidity range $|\eta| < 2.4$, where b-tagging are still effective. It is further required that all jet candidates to be isolated from *tight* leptons. A cone of the size 0.4 around each jet candidate is defined and candidates will be removed if any *tight* leptons are found within such a cone. This procedure is implemented to remove the overlap between leptons and jets.

The two primary sources of background are (i) detector noise, and (ii) jets from PU interactions. To suppress detector noise, a set of cut-based selections are applied to jet candidates. This algorithm utilizes information from PF candidates, including: (i) fraction of charged (neutral) hadrons energy, (ii) fraction of charged (neutral) EM energy, (iii) fraction of muon energy, and (iv) object multiplicity. The "tightLepVeto" working point is chosen to select jet candidates, which corresponds to 98-99% signal efficiency.

The second algorithm is designed to reject jet candidates that originate from PU interactions. This algorithm is based on a BDT that utilizes: (i) the trajectories of tracks associated to the jets, (ii) the topology of the jet shape, and (iii) object multiplicity. The *loose* working point is chosen to select jet candidates with $p_{\mathrm{T}} < 50$ GeV, which corresponds to 99% signal efficiency.

Applying this algorithm to jet candidates with $p_T > 50$ GeV is both ineffective and unnecessary as PU jets mostly reside in low p_T spectrum. The overall effect of this algorithm on this analysis is small as PU jets constitute only a small fraction of all jet candidates in the phase space of this analysis.

As discussed in section 8.4, the energy scale for all jet candidates (data and MC) are calibrated. One extra correction is applied to simulated jets to recreate the jet energy resolution as measured in data.

13.3   Identification of b jets

The DeepJet algorithm [24] is used to identify jets that originate from b quark. The core strategy of this algorithm is to minimise information. This is achieved by removing entirely the selection of jet constituents, which limits the number of jet constituents considered. Additionally, an effort is made to use as many low-level features as possible, which further further deepens the feature space. Approximately a total of 650 input features are used, which can be categorized into four groups: (i) global variables, (ii) charged PF candidate features, (iii) neutral PF candidate features, (iv) and secondary vertex (SV) features associated with the jet.

The DeepJet algorithm outputs a score ranging between 0 and 1, with 0 (1) being the most background- (signal-) like. The medium working point is chosen to tag b jet candidate, which corresponds to 90% signal efficiency. The shape of the DeepJet output distribution is corrected for the differences between data and MC in signal and background efficiencies. The per-event correction weight ω is defined as,

$$\omega = \prod_{i}^{N_{\text{jets}}} \text{SF}(p_{T_i}, \eta_i, F_i, D_i),\tag{13.3}$$

where SF is the ratio of efficiency in data to efficiency to MC parameterized as a function of p_T, η, (MC truth) flavor F, as well as DeepJet output D of each jet in the event. ω is applied to all MC samples, and additional scale factors are also applied to remove the normlization effect of ω. These scale factors are measured using MC in e$\mu\ell$ channel described in chapter 14. The effect of these scale factors are shown in Figure 13.1.
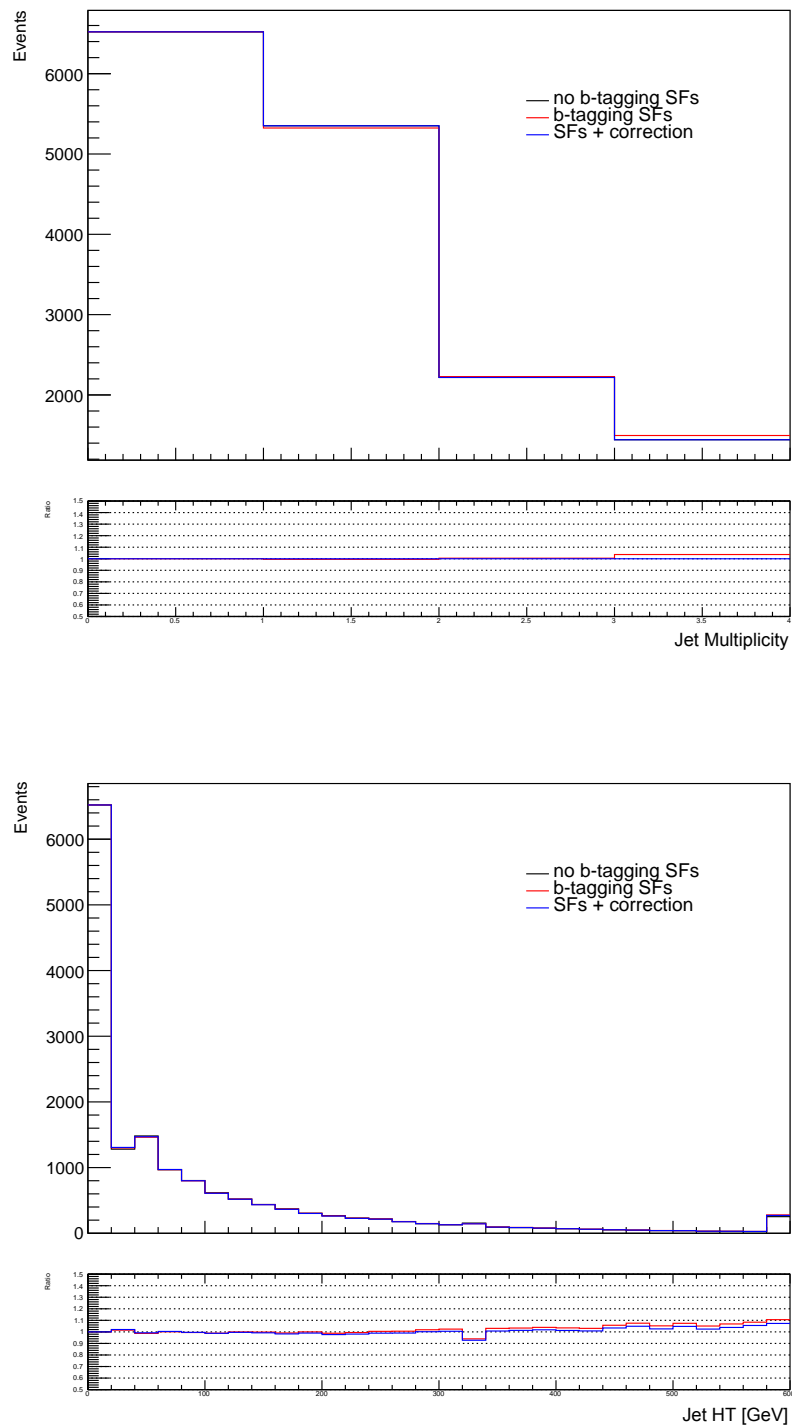
Figure 13.1: Simulated events in $e\mu l$ channel (run-II). Top: jet multiplicity. Bottom: H_T of jets.

CHAPTER 14

Event Selection

Events to contain exactly three *tight* leptons. Events with different lepton compositions are further categorized into different channels: eee, μμμ, eμℓ. In all three channels, the sum of the charges of three leptons are required to be 1 or -1. The leading leptons in all selected events are required to be matched with trigger objects via $\Delta R < 0.2$. Within each channel, different regions are defined to further understand signal and background.

eμℓ is the channel where the presence of the signal is expected. This channel is split up to signal-enriched signal regions and signal-depleted validation regions, which are discussed in section 14.1 and section 14.2, respectively. Events with the presence of the same flavor leptons are categorized into eee or μμμ are signal-depleted by definition. Regions in these two channels are used to study background processes, which is discussed in section 14.2.

14.1 Signal Region

In $e\mu l$ channel, we set up our signal region (SR) targeting LFV signal processes. Apart from the OSOF pair, we require the events to have big MET (MET > 20 GeV). Events with an Opposite-Sign-Same-Flavor (OSSF) lepton pair, the invariant mass of which is between 50 GeV and 106 GeV, are removed to suppress contribution from Z production. To further suppress background contribution, we require at least one jet and at most one b-tag jet.

Distributions of leading lepton p_T and invariant mass of Opposite-Sign $e\mu$ pair are shown in Figure 14.1. Please note: Although we intend to deploy a data-driven technique to estimate non-prompt background in the actual analysis, all backgrounds in Figure 14.1 are estimated using MC simulation only. This serves as a preliminary check to understand the components of different backgrounds in SR.

Using the invariant mass of the Opposite-Sign $e\mu$ pair, SR is further divided into two subsets to create ST and TT enriched regions:

- SR1, $M_{e\mu} <$ 150GeV: TT enriched.

- SR2, $M_{e\mu} >$ 150GeV: ST enriched.
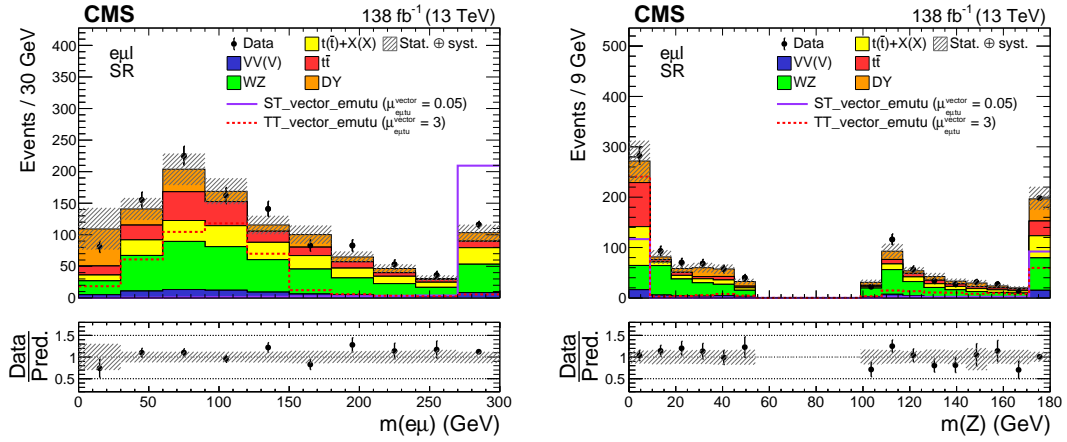
Figure 14.1: Distributions of different kinematic variables estimated in SR (full run II). Backgrounds are estimated using MC only. From left to right: invariant mass of the $e\mu$ pair, leading lepton p_T.

14.2 Validation Region

Events are required to pass trigger requirements listed in Table A.1-A.3. The leading lepton is required to have a $p_T >$38 GeV. The sum of the charges of the three selected leptons is required to be either 1 or -1. Selections applied to different regions are summarized in Table 14.1 and are illustrated in Figure 14.2.

- Acronyms

  - C_i: electric charge of the i-th lepton

  - OnZ: requiring the presence of at least one Opposite-Sign-Same-Flavor lepton pair with an invariant mass between 50 GeV and 106 GeV.

  - OffZ: veto events that pass OnZ

  - MET: missing transverse energy

  - njet: number of jets that pass jet selection

  - nbjet: number of b-tag jets that pass jet selection

  - SR: signal region

  - VR: validation region for data-driven method

  - WZ CR: WZ control region

There are two types of signal-depleted regions defined within eee/$\mu\mu\mu$ channel: validation region (VR) and WZ control region. The distinction between these two regions comes down to

Channel	Region	$\lvert\Sigma_i C_i\rvert=1$	OnZ	OffZ	MET $>$ 20 GeV	njet$>$=1	nbjet$<$=1
eee	VR	✓	-	-	-	-	-
	WZ VR	✓	✓	-	✓	✓	✓
eμℓ	SR	✓	-	✓	✓	✓	✓
	Nonprompt VR	✓	✓	-	-	-	-
	WZ VR	✓	✓	-	✓	✓	✓
μμμ	Nonprompt VR	✓	-	-	-	-	-
	WZ VR	✓	-	✓	✓	✓	✓

Table 14.1: Summary of the cuts applied to different regions
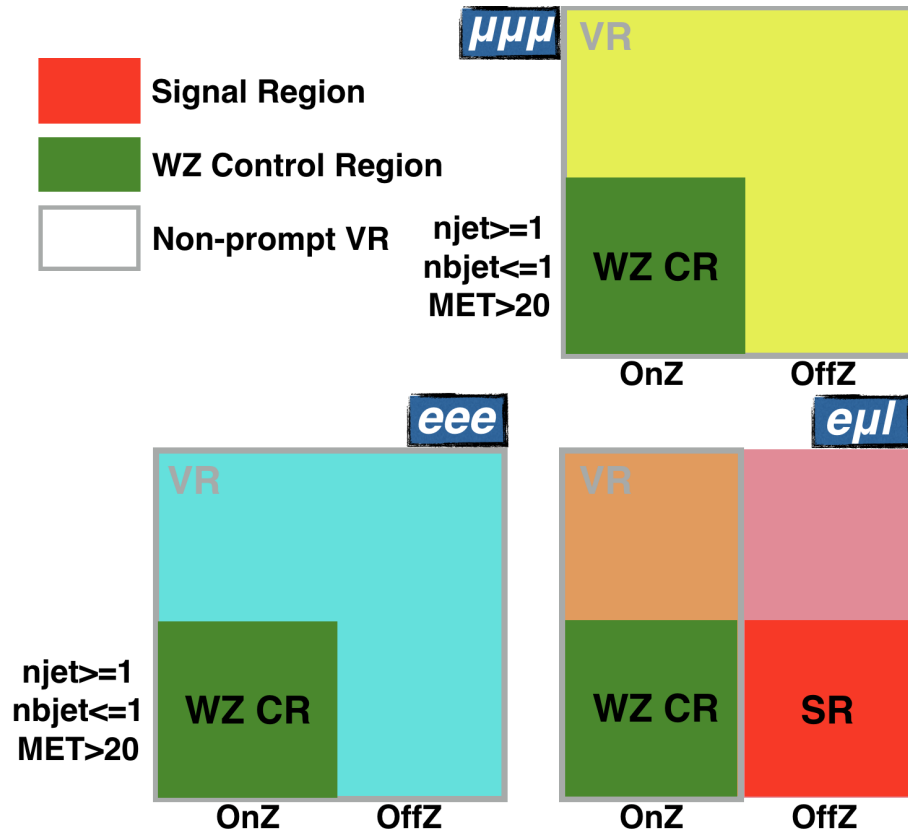
Figure 14.2: Illustration of event selections applied to different regions

background compositions. It is expected that the VR has a significant fraction of non-prompt backgrounds while WZ production is responsible for most of the backgrounds in the WZ control

region. Distributions of leading lepton p_T and leading lepton η in WZ control region can be found in Figure 14.3-14.5. Discussion about validation region can be found in chapter 15.
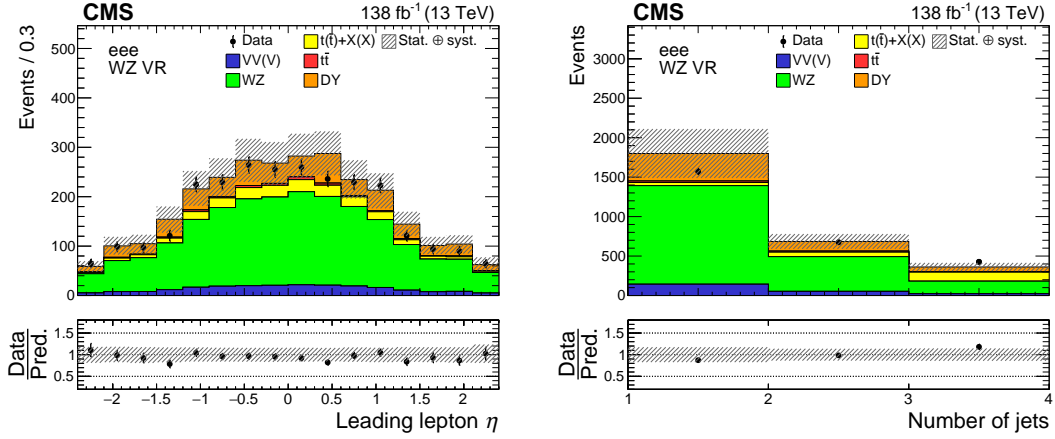
Figure 14.3: Distributions of different kinematic variables estimated in WZ control region (full run II). Backgrounds are estimated using MC only. From left to right: leading lepton η, jet multiplicity.
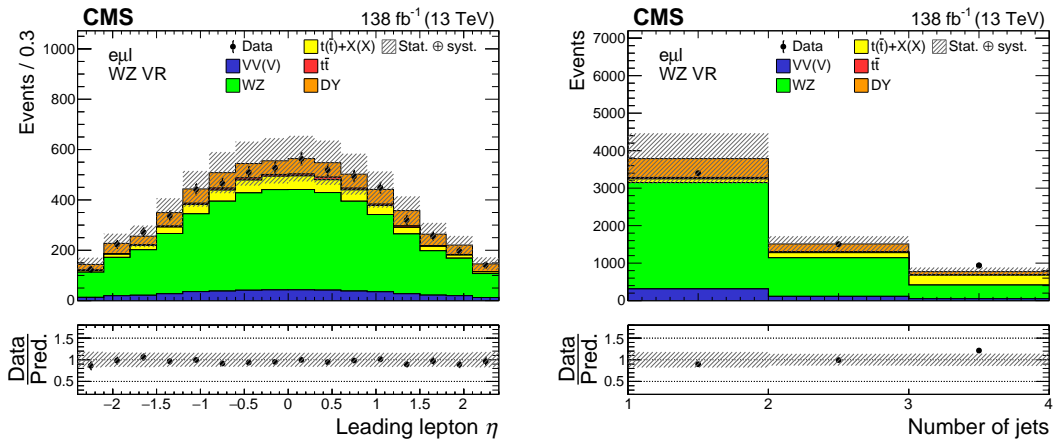
Figure 14.4: Distributions of different kinematic variables estimated in WZ control region (full run II). Backgrounds are estimated using MC only. From left to right: leading lepton η, jet multiplicity.

14.3 Kinematic Reconstruction

A kinematic reconstruction is performed in this channel. The OSOF lepton pair is assumed to be the product of cLFV interaction, while the third lepton (bachelor lepton) is assumed to be the product of the SM top-quark decay. Jet with the highest b-tagging score is assumed to originate from bottom-quark decay, therefore, it is combined with MET and the bachelor lepton to build the SM top-quark.
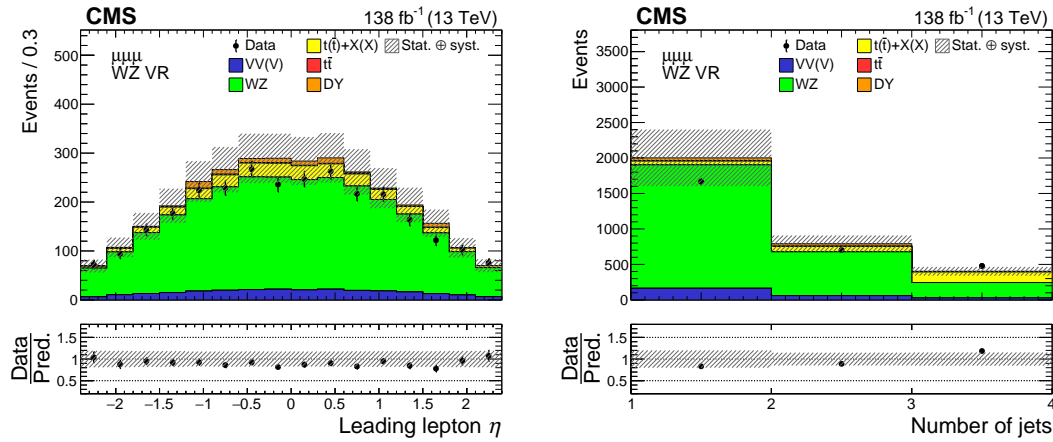
Figure 14.5: Distributions of different kinematic variables estimated in WZ control region (full run II). Backgrounds are estimated using MC only. From left to right: leading lepton η, jet multiplicity.

- We take the x and y component of MET as measurements of neutrino p_x and p_y.

- The z component of neutrino momentum is calculated by imposing the constraint that the invariant mass of the combined object (bachelor lepton+neutrino) must be equal to W boson mass.

- If there is no real solution, we take the real part of the complex solution.

- If there is more than one real solution, we chose the solution that is the closest to the p_z of the bachelor lepton.

In events where there is more than one candidate of bachelor lepton, the lepton that gives a top mass that is the closest to the the SM top-quark mass is chosen as the bachelor lepton.

Once the bachelor lepton has been determined, the OSOF pair is combined with jets One-By-One to reconstruct LFV top candidates. Jet with the highest b-tagging score is excluded from this reconstruction since it is assumed to be from the decay of the SM top-quark. Out of all the LFV top candidates, the candidate that gives a top mass that is the closest to the the SM top-quark mass is chosen.

Kinematic reconstruction of heavy objects like top-quark is carried out using selected leptons, jets, and MET.

- **Z boson candidate**

  - Z boson candidate is reconstructed by combining two Opposite-Sign-Same-Flavor leptons

  - Whenever possible, Z candidate with an invariant mass that is closest to Z boson

mass is chosen

- **Standard model top-quark candidate**

  - Standard model top-quark candidate is reconstructed by combining a jet, a lepton, and MET

  - Jet with the highest b-tagging score is always chosen

  - In $e\mu l$ channel, flavor-violating leptons are excluded from reconstruction

    * Whenever more than one lepton is available due to possible combinatorics of flavor-violating $e\mu$ pair, lepton that forms a top-quark candidate with an invariant mass that is closest to the top-quark mass is chosen

  - In $eee/\mu\mu\mu$ channel, leptons that form Z candidate are excluded

  - Lepton is combined with MET to reconstruct W candidate

    * Transverse component of the neutrino momentum is taken from MET

    * Neutrino p_z is obtained by W mass constrain set on the lepton-neutrino system

- **Flavor-violating top-quark candidate**

  - Flavor-violating top-quark candidate is reconstructed by combining a jet and an Opposite-Sign $e\mu$ lepton pair

  - Jet with the highest b-tagging score is excluded from reconstruction unless it fails b-tagging

  - Whenever possible, the jet that forms a flavor-violating top-quark candidate with an invariant mass that is the closest to the top-quark mass is chosen

  - Lepton that forms the standard model top-quark candidate is excluded

- **The first b-jet+lepton system**

  - Jet with the highest b-tagging score is always chosen

  - All three leptons can form this system with the chosen jet

    * Lepton that forms a jet-lepton system with the lowest invariant mass is chosen

- **The second b-jet+lepton system**

  - Jet with the highest b-tagging score is excluded unless it passes b-tagging

  - Whenever possible (after the previous step), jet with the highest b-tagging score is chosen

  - Lepton that forms the first jet-lepton system is excluded

– Leptons that have the same sign as the lepton that forms the first jet-lepton system is excluded

– Whenever possible, lepton that forms a jet-lepton system with the lowest invariant mass is chosen

CHAPTER 15

Nonprompt Background Estimate

In this analysis, the term *real* leptons refers to prompt leptons that originate from the primary interaction, like the decays of vector bosons that are produced via hard collisions. *Fake* leptons refer to non-prompt leptons that are mostly produced in the decays of b-quarks, c-quarks, pions, or from photon conversion. A relatively low ratio of *fake* leptons in a lepton selection is desired and can be achieved by applying stringent isolation requirements and/or an MVA-based lepton identification specifically trained against *fake* leptons.

Non-prompt backgrounds are defined to be backgrounds with at least one *fake* lepton passing the analysis selection, in this case generally dominated by Drell-Yan and $t\bar{t}$ production. An accurate estimation of non-prompt backgrounds is difficult to achieve through MC modeling. Therefore, a data-driven technique called the 3D matrix method is used to estimate the non-prompt backgrounds.

A brief description of the matrix method is given in this section followed by details on it's implementation. In autorefyou can find more information on the generalization of the matrix method to 2D and 3D. The appendices also contain a validation of the method and its performance is compared to the results of the fake factor method, another commonly used data-driven technique.

15.1   The Matrix Method

The matrix method citematrix2012 is a data driven technique used to estimate the fraction of *fake* leptons that pass a given lepton selection, referred to as "*tight*". The *tight* selection usually incorporates tight lepton identification and isolation requirements and should be similar if not exactly the same as lepton selection in the actual analysis. The *loose* selection is obtained by loosening the *tight* selection. The *loose* selection is used as a baseline such that any leptons that pass it will fall into one of the two exclusive categories: *tight* or *not tight*. The matrix method treats *real* and *fake* leptons separately. As a result, *real* and *fake* efficiencies are introduced (see Figure 15.1).

In a simplified scenario with only one lepton in the final state, the *real* efficiency r measures the likelihood of *real* leptons pass *tight* selection. It is treated as an observable that can be obtained
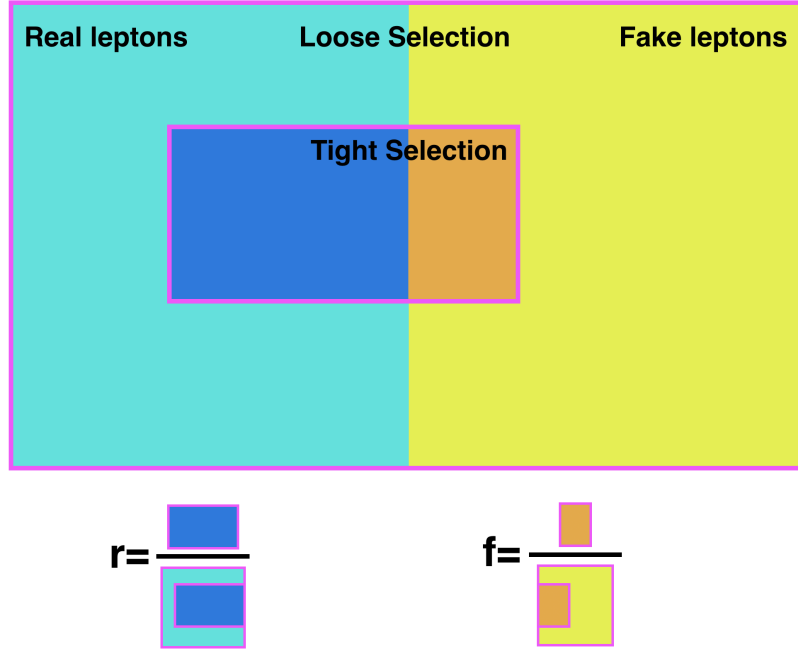
Figure 15.1: Illustration of the *real* efficiency *r* and the *fake* efficiency *f*.

through measurement:

$$r = \frac{n_R^T}{n_R^T + n_R^{\overline{T}}}, \tag{15.1}$$

in which $n_R^T / n_R^{\overline{T}}$ denotes the number of events with a *real* lepton that is *tight/not tight*.

Similarly, *fake* efficiency *f* can be expressed as

$$f = \frac{n_F^T}{n_F^T + n_F^{\overline{T}}}, \tag{15.2}$$

in which $n_F^T / n_F^{\overline{T}}$ denotes the number of events with a *fake* lepton that is *tight/not tight*.

The measurement of r/f is often performed in dedicated control regions, where high purity of *real/fake* leptons is expected. These regions are referred to as the measurement regions (MR). It is assumed that r/f is a universal property of *real/fake* leptons that is independent of physics processes. Therefore, r/f extracted from MR can be used to estimate the contamination of *fake* leptons in a different region (e.g. SR) even though these two regions are orthogonal to each other.

In our simplified scenario, the total number of events in the region of interest (e.g. SR/VR) with a *tight/not tight* lepton can be expressed in a system of equations:

$$N^T = N_R^T + N_F^T$$
$$N^{\overline{T}} = N_R^{\overline{T}} + N_F^{\overline{T}},$$

(15.3)

in which capital letter "N" is used to indicate that these numbers are referring to events in a region that is different from MR. $N_R^{\overline{T}}/N_F^{\overline{T}}$ can be expressed in terms of r/f and N_R^T/N_F^T according to Equation 15.1/15.2 and the assumption that r/f is transferable to other regions:

$$N^T = r\frac{N_R^T}{r} + f\frac{N_F^T}{f}$$
$$N^{\overline{T}} = (1-r)\frac{N_R^T}{r} + (1-f)\frac{N_F^T}{f}.$$

(15.4)

Equation 15.4 can also be expressed in the form of matrix:

$$\begin{pmatrix} N^T \\ N^{\overline{T}} \end{pmatrix} = \begin{pmatrix} r & f \\ 1-r & 1-f \end{pmatrix} \begin{pmatrix} N_R^T/r \\ N_F^T/f \end{pmatrix}.$$

(15.5)

Since those two numbers that appear on the lefthand side of the equation can be obtained through measurement, a matrix inversion can give us an estimation of *fake* leptons that pass *tight* selection, denoted by N_F^T.

15.2   Extraction of Prompt and Nonprompt Efficiencies

The description in previous section deals with a scenario where only one lepton is studied. This analysis uses a generalized version of the matrix method (i.e. 3D matrix method. See autorefsec:appendixnonprompt for more details), where all three leptons are considered to be possibly *fake*.

The measurement of both r and f are performed in dilepton MRs using the Tag-and-Probe approach. Datasets used to measure r and f are listed in Table. More information about MC samples listed in Table ??-. More can be found. Events are required to fire either single-electron or single-muon trigger listed in Table 15.1.

Both r and f are parameterized in bins of lepton p_T, $|\eta|$, and jet multiplicity. The bin range is optimized to retain sufficient statistics for each bin:

- Electron p_T bin range: {20.0, 24.6, 28.8, 33.0, 37.2, 41.4, 46.1, 52.1, 59.3, 68.3, 82.7, 110.6} GeV,

Channel	path	dataset	2016	2017	2018
Electron	HLT_Ele27_WPTight_Gsf	data & MC	✓	-	-
	HLT_Ele35_WPTight_Gsf	data & MC	-	✓	✓
Muon	HLT_IsoMu27	data & MC	✓	✓	✓

Table 15.1: Summary of the triggers used in the measurement of r and f.

- Muon p_T bin range: {20.0, 23.8, 27.7, 31.3, 35.0, 38.9, 42.8, 45.6, 50.7, 59.5, 72.9, 94.3} GeV,

- $|\eta|$ bin range: {0, 0.8, 1.6, 2.4},

- Jet multiplicity: {0 jet, at least 1 jet}.

The addition of jet multiplicity as a dependency is motivated by the unideal assumption that both r and f are independent of physics processes. We know that *fake* leptons of different origins are often reconstructed at a different rate. For instance, *fake* electrons coming from the decays of b-quarks have very different kinematics than *fake* electrons that originate from photon conversion. Therefore, the probabilities of these two types of *fake* electrons to pass any kinematic-dependent selections are going to differ. In other words, the proportion of $t\bar{t}$ events that constitutes the non-prompt backgrounds has a direct impact on r and f. To achieve a better representation of the target region (in terms of background composition), r and f are linked to the number of jets in events.

The *fake* efficiency is measured in a same-sign dilepton region, in which the leading p_T lepton, used as the *tag* (See Table 15.2 for definition), is required to be matched with trigger objects via $\Delta R <0.2$. The sub-leading lepton is required to pass the *loose* selection and is taken as the *probe*. Two orthogonal regions are defined within the dilepton MR, which account for the two bins of jet multiplicity:

- Events in dilepton MR with at 0 jets.

- Evens in dilepton MR with at least 1 jet, and at most 1 b-tag jet.

The selection of the second MR is motivated by the selection of the SR, in which at least one jet and at most one b-tag jet are required.

The contribution from prompt backgrounds, estimated from MC simulation, are subtracted from data (see Figure 15.2). Therefore, the fake efficiency f is calculated as:

$$f = \frac{n_{data}^{T} - n_{Prompt\ MC}^{T}}{n_{data}^{T+\overline{T}} - n_{Prompt\ MC}^{T+\overline{T}}}. \tag{15.6}$$

Lepton	Selection	*Loose*	Tag	*Tight*[b] (Passing-Probe)
Electron	p_T	> 20 GeV	> 38 GeV[a]	> 20 GeV
	I_{iso}^{mini}	<0.4	<0.1	<0.12
	TOP MVA ID	>-0.9	>0.95	>0.9
	Match with trigger objects	-	✓	-
Muon	p_T	> 20 GeV	> 30 GeV	> 20 GeV
	I_{iso}^{mini}	<0.4	<0.1	<0.12
	cut-Based ID	-	Medium WP	Medium WP
	TOP MVA ID	>0.5	>0.9	>0.9
	Match with trigger objects	-	✓	-

Table 15.2: Summary of the lepton selections needed for the measurement of r and f. Please note: a) the minimum p_T cut for *tag* electron in 2016 dataset is reduced to 30 GeV to adjust for the trigger threshold; b) the *tight* selection here is the same as the lepton selection listed in autorefsec:reconstruction.1.
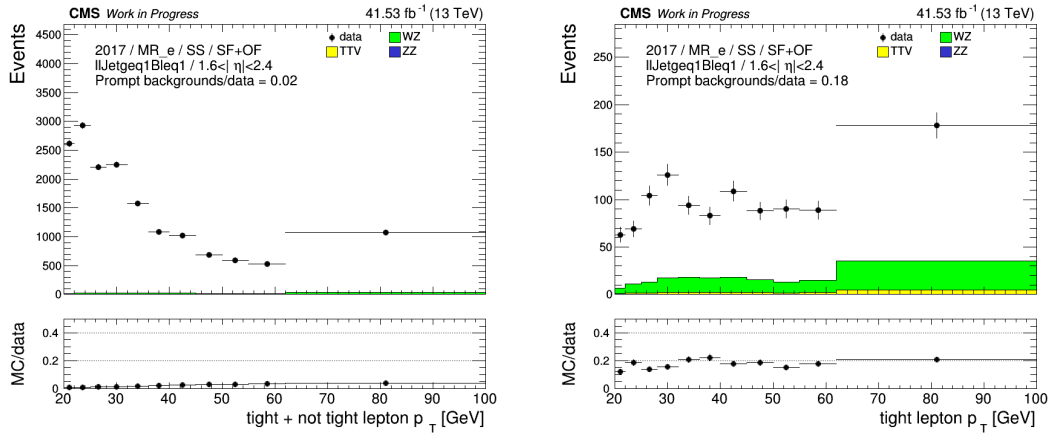
Figure 15.2: Distribution of lepton p_T in an electron *fake* efficiency MR (2017). In this particular example: both $e + e$ (Same-Sign-Same-Flavor) and $e + \mu$ (Same-Sign-Opposite-Flavor) pair are considered. At least one jet and at most one b-tag jet are required (the second jet multiplicity bin). *Probe* electron is required to have $1.6 < |\eta| < 2.4$ (the third η bin). Contamination from prompt backgrounds are estimated with MC simulation. From left to right: *tight+not tight* electron, only *tight* electron.

The dependency of f on background composition is also reflected in the difference between f measured with different flavor compositions. For electrons, the *fake* efficiency measured with $e + \mu$ selection is visibly different than the *fake* efficiency measured with $e + e$ selection, which
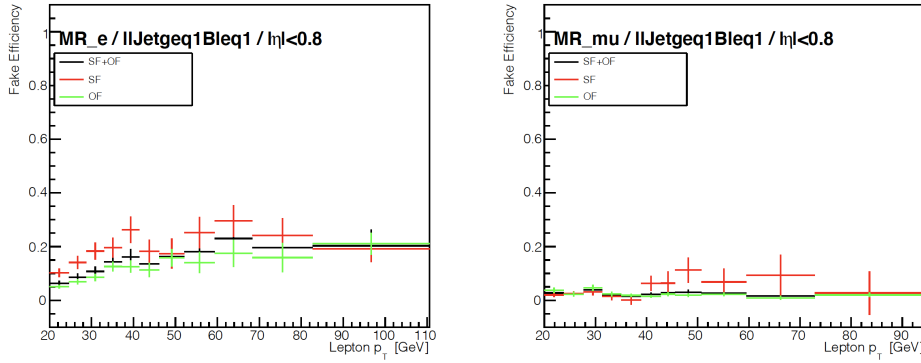
Figure 15.3: Fake efficiency measured in 2017 data. These plots correspond to the first $|\eta|$ bin ($|\eta| <0.8$) and the second jet multiplicity bin. Error bars displayed in these plots include statistical uncertainty only. From left to right: electron f, muon f.

could suggest:

- *Fake* electrons that originate from the decays of b-quarks have a lower probability of being selected by *tight* selection than *fake* electrons coming from Z+jets events.

- The Same-Sign $e + e$ selection is contaminated with *real* electrons and, to a less extent, the Same-Sign $e + \mu$ selection.

Events that have two same-sign electrons with an invariant mass between 76 GeV and 106 GeV are removed from MR to suppress the charge-flip effect of electrons. No such cut has been introduced to muon MR due to the negligible rate of charge-flip in muons.

Lepton selections listed in Table 15.2 are referred to as *baseline*, which has been optimized to provide a robust non-prompt estimate. Variation of the *basline* selection has also been studied.

There is a trade-off between muon r and f when changing the isolation requirement (See Figure 15.4-15.5). While it is ideal to reduce statistical uncertainties by loosening the *loose* definition, a low *real* efficiency should also be avoided. This change proves to have a minor effect on the overall estimate.

The *real* efficiency r is estimated with simulated $t\bar{t}$ events in an Opposite-Sign dilepton region. The same lepton selection (See Table 15.2) is used to perform Tag-and-Probe. The leading lepton in p_T is used as a *tag* while the sub-leading lepton (with opposite sign) is taken as a *probe*. The variation of r between $e + e$ and $e + \mu$ flavor combination is negligible (See Figure 15.6). Therefore, only $e + \mu$ events are used to measure *real* efficiency in order to minimize the contamination of *fake* leptons.

The selection criteria for r/f measurement region is summarized in Table 15.3. The latest version
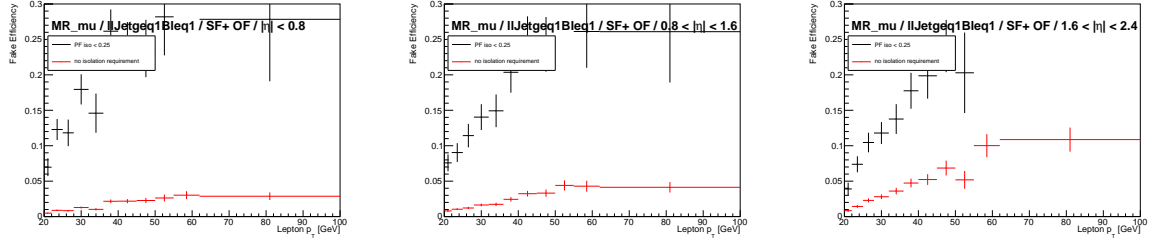
Figure 15.4: Comparison of the muon *fake* efficiencies measured with isolation requirement in *loose* definition (2017 datasets). From left to right: $|\eta| < 0.8$, $0.8 < |\eta| < 1.6$, $1.6 < |\eta| < 2.4$.
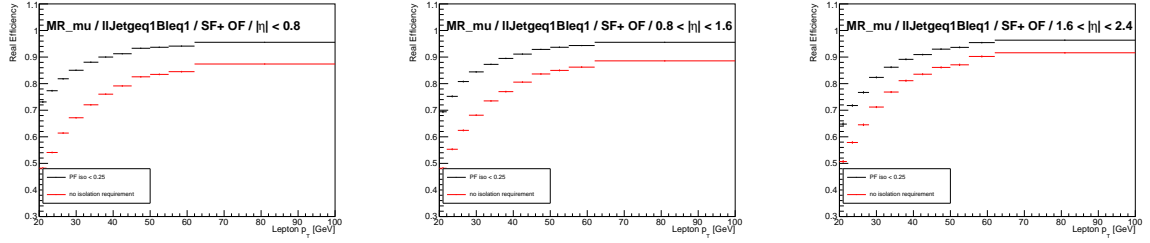
Figure 15.5: Comparison of the muon *real* efficiencies measured with different *loose* definition (2017 datasets). From left to right: $|\eta| < 0.8$, $0.8 < |\eta| < 1.6$, $1.6 < |\eta| < 2.4$.
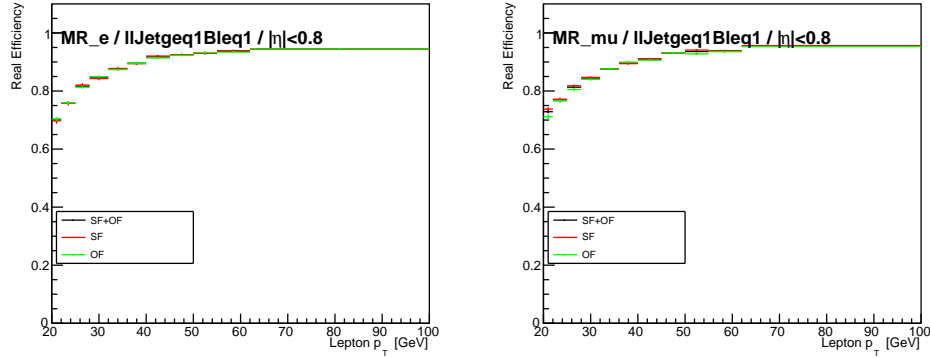
Figure 15.6: *Real* efficiency measured in simulated $t\bar{t}$ events (2017). These plots correspond to the first $|\eta|$ bin ($|\eta| < 0.8$) and the second jet multiplicity bin. Error bars displayed in these plots include statistical uncertainty only. From left to right: electron r, muon r.

of the *real* and *fake* efficiency is summarized in autorefsec:RandF.

Observable	jet bin	# of selected leptons	lepton flavor composite	$\lvert \Sigma_i C_i \rvert$	OffZ	njet>=1	nbjet<=1
f	0 jet	2	any	1	✓	-	-
	1 or more jet	2	any	1	✓	✓	✓
r	0 jet	2	eμ only	0	✓	-	-
	1 or more jet	2	eμ only	0	✓	✓	✓

Table 15.3: Summary of the cuts applied to the r/f measurement region.

15.3 Validation of the Matrix Method

The performance of the matrix method is validated using three regions that are tangential to the SR, referred to as VRs. In these VRs, prompt backgrounds are estimated with MC simulation while non-prompt backgrounds are estimated with the matrix method. A summary of the selections applied to these VRs is given in autorefsec:selection.4.

Distribution of lepton lepton η and jet multiplicity are shown in Figure 15.7-15.9. Distributions of other variables are included in autorefsec:VRmatrix
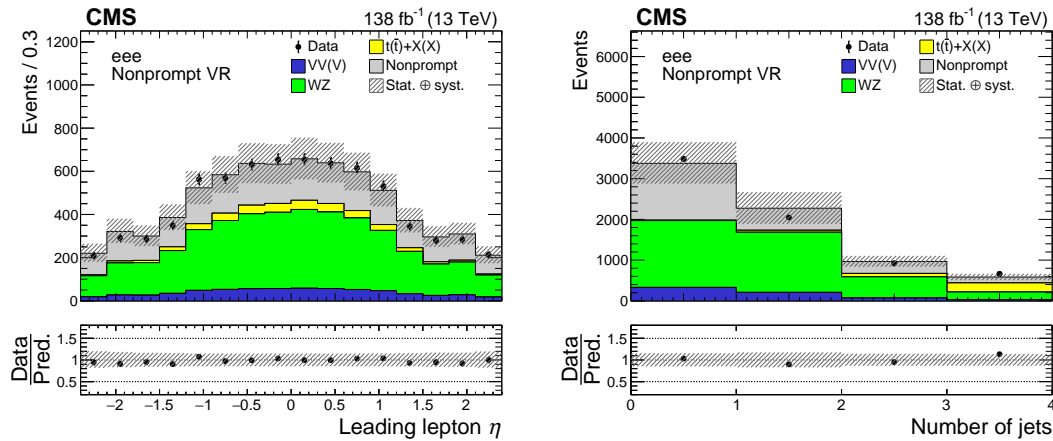
Figure 15.7: Distributions of different kinematic variables estimated in VR with three electrons. From left to right: leading lepton η, jet multiplicity.

15.4 Nonprompt estimate in SR

The matrix method is used to estimate non-prompt backgrounds in the signal region due to its superior performance over other alternative methods (e.g. fake factor method discussed in autorefsec:appendixnonprompt). Distributions of leading lepton η, jet multiplicity, n-jet
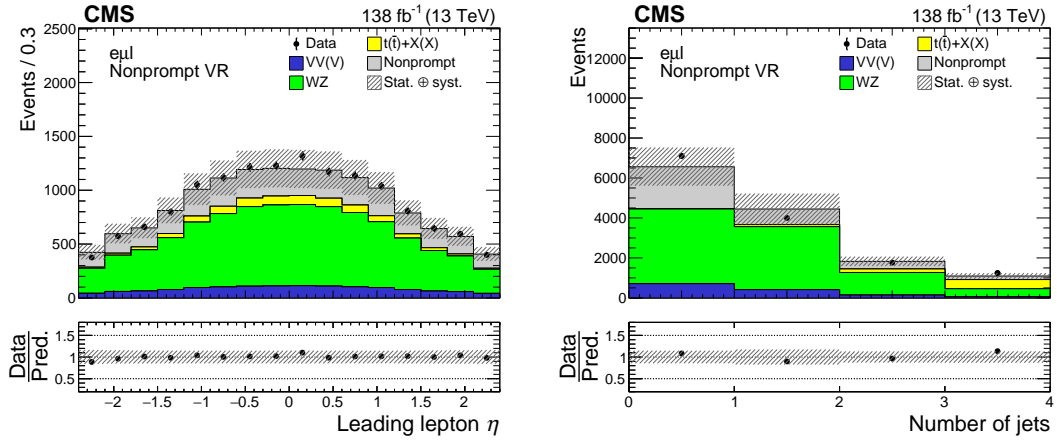
Figure 15.8: Distributions of different kinematic variables estimated in VR with electron, muon and a third light lepton. From left to right: leading lepton η, jet multiplicity.
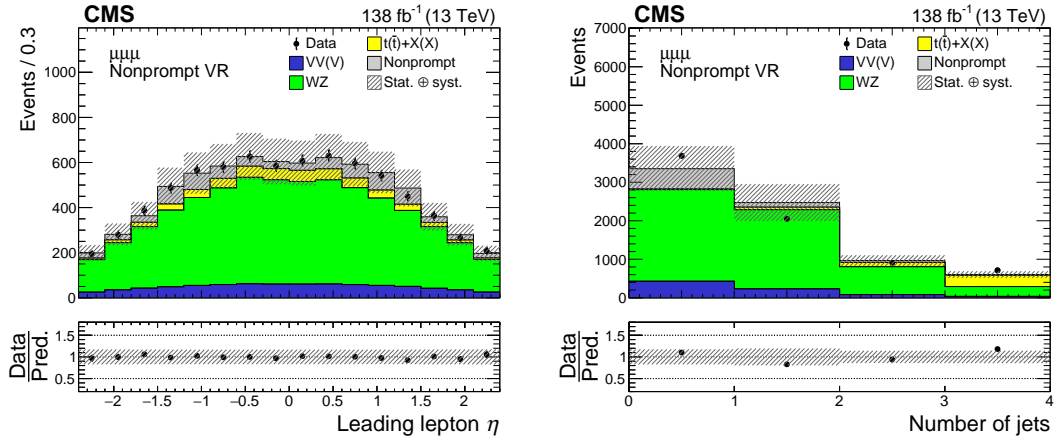
Figure 15.9: Distributions of different kinematic variables estimated in VR with three muons. From left to right: leading lepton η, jet multiplicity.

multiplicity, missing transverse energy, LFV-Top mass and Z mass are shown in Figure 15.10-15.11. Distributions of other variables are included in autorefsec:SRDataDriven.

The number of expected events from various kinds of backgrounds are shown in Table 15.4. Representative bin yields and their statistical uncertainties are summarized in citeStatError.
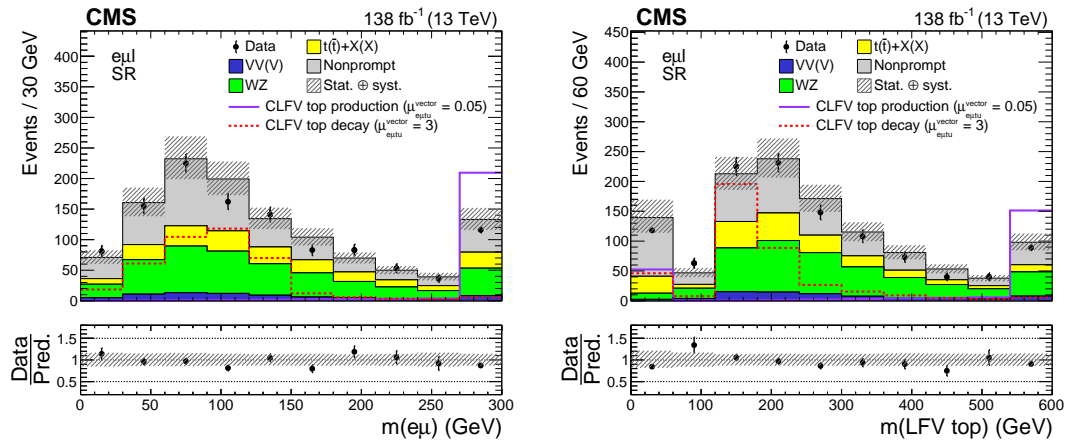
Figure 15.10: Distributions of different kinematic variables estimated in SR (full run II). From left to right: b-jet multiplicity, missing transverse energy.
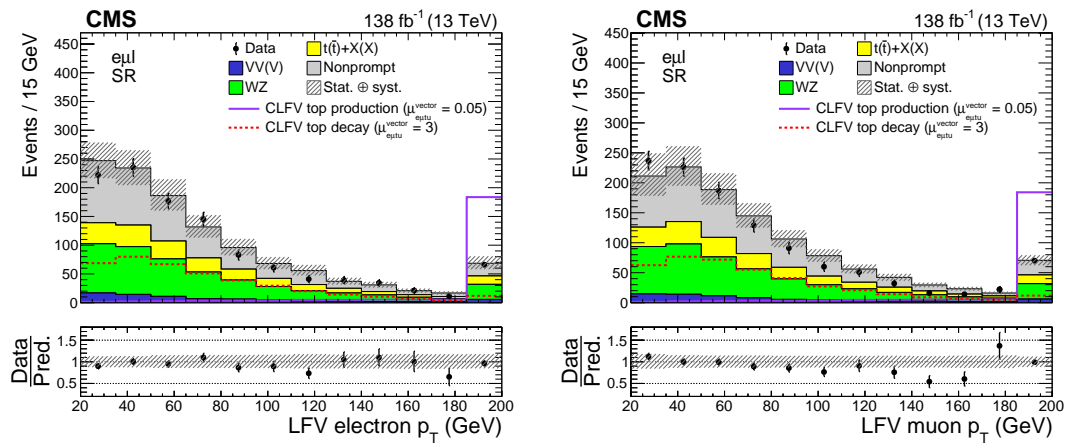
Figure 15.11: Distributions of different kinematic variables estimated in SR (full run II). From left to right: b-jet multiplicity, missing transverse energy.
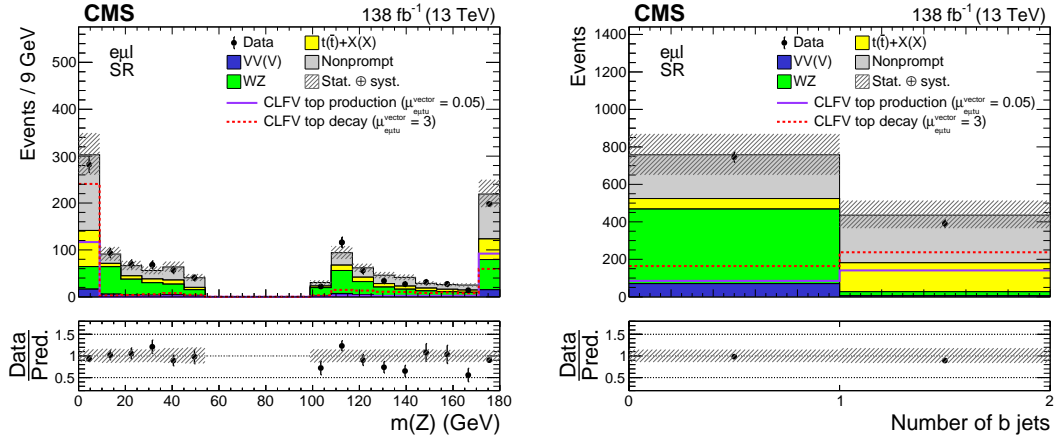
Figure 15.12: Distributions of different kinematic variables estimated in SR (full run II). From left to right: mass of the flavor-violating-top-quark candidate, mass of the Z boson candidate

Table 15.4: Expected background contributions and the number of events observed in data collected during 2016–2018. The statistical and systematic uncertainties are added in quadrature. The category "other backgrounds" includes smaller background contributions containing one or two top quarks plus a boson or quark. The CLFV signal, generated with $C_{e\mu tu}^{vector}/\Lambda^2 = 1 \mathrm{TeV}^{-2}$, is also listed for reference. The signal yields include contributions from both top production and decay modes.

Process	m(eμ)<150 GeV	m(eμ)>150 GeV
Nonprompt	351 ± 92	146 ± 38
WZ	275 ± 64	145 ± 35
ZZ	33.2 ± 6.5	13.1 ± 2.6
VVV	17.0 ± 8.5	12.0 ± 6.0
t$\bar{\text{t}}$W	47.6 ± 10.0	40.0 ± 9.1
t$\bar{\text{t}}$Z	39.1 ± 7.9	25.8 ± 5.4
t$\bar{\text{t}}$H	28.2 ± 4.5	10.0 ± 1.6
tZq	5.5 ± 1.1	2.5 ± 0.5
Other	7.3 ± 3.7	4.5 ± 2.3
Total expected	805 ± 123	398 ± 57
Data	783	378
CLFV	207 ± 15	4440 ± 215

16 | CHAPTER

Signal Extraction with Boosted Decision Trees

To further separate the LFV signal from the standard model backgrounds in our signal region, a Boosted Decision Tree (BDT) was employed using the xgBoost package citeChen:2016:XST:2939672.2939785.

16.1 BDT Setup

Since the kinematic distributions of final state particles in our two signal channels (single top production and top decay) differ significantly (see Figure 16.3-16.7), the SR is divided into two parts (SR1: $M_{e\mu} \leq$150GeV; SR2: $M_{e\mu} >$150GeV) in order to achieve the optimal sensitivity. Within each sub-region of the SR, a binary BDT is trained.

The signal datasets used for BDT training are MC samples that correspond to single top LFV production (ST) and top LFV decay (TT). The Lorentz structure and the flavor of the up-type quark involved in LFV interaction were shown to have a minor impact on the kinematics of final state particles (See Figure 16.1). Therefore, they were not further distinguished.
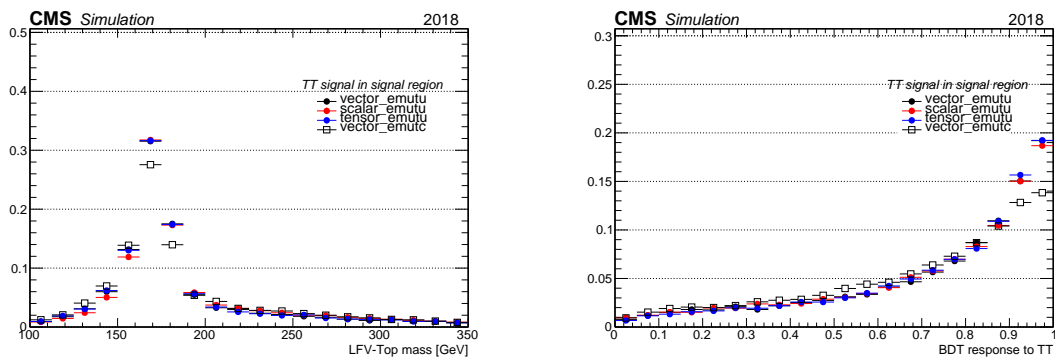
Figure 16.1: Normalized distribution in SR1. From left to right: LFV top mass, BDT shape

The background MC samples include all the samples listed in Table **??-??**. It was determined that splitting the background into separate classes (prompt vs non-prompt backgrounds) did not increase the signal discrimination power and caused overtraining which is why in the backgrounds are combined.

The MC dataset is divided into 5 subsets evenly. A 5-fold cross-validation scheme is deployed:
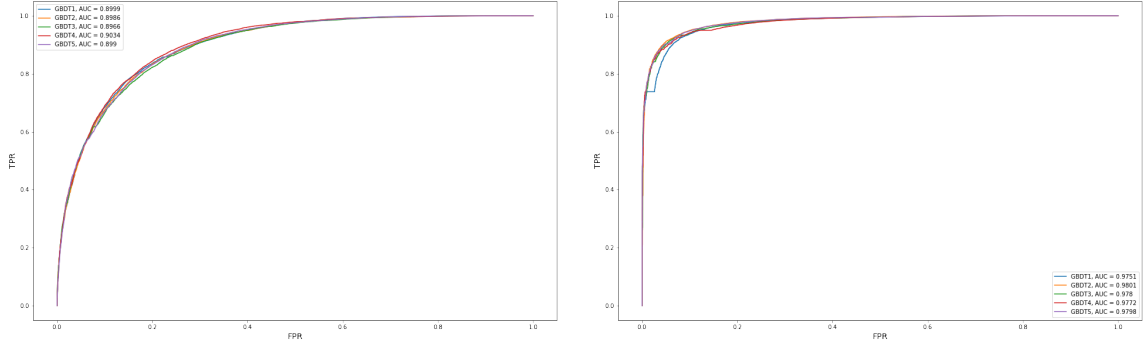
Figure 16.2: ROC curve with 5-fold cross validation. From left to right: BDT targeting TT (SR1), BDT targeting ST (SR2).

the training set consists of 3 out of the 5 subsets, while validation/testing set each consists of 1 subset. The number of estimators is set to 1000, max depth 5, and sub-sample 0.8 with the evaluation metric set to "mlogloss". The results of the cross-validation study are shown in Figure 16.2. The trained BDT model targeting TT(ST) signal is used to evaluate each event in SR1(SR2): a BDT weight is assigned to each event, giving the likelihood that it originated from TT(ST) signal.
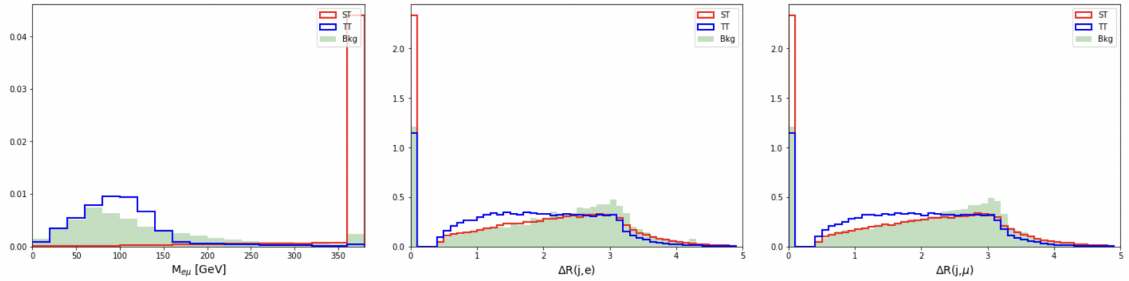
Figure 16.3: Normalized distribution of various features in SR. From to left to right: MVA_$M_{e\mu}$, MVA_JeDr, MVA_JeDr.

16.2 Input Variables

The observables used in training are referred to as "features" in this analysis. The following features are used for both ST and TT classifier:

- **MVA_$M_{e\mu}$: invariant mass of the Opposite-Sign $e\mu$ pair**
- **MVA_LFVePt:** p_T **of the flavor-violating electron**
- **MVA_LFVmuPt:** p_T **of the flavor-violating muon**
- **MVA_LFVTopmass: invariant mass of the flavor-violating top-quark candidate**

- **MVA_Zmass: invariant mass of Z boson candidate**

- **MVA_Jet2Btag: b-tagging score of the jet with the second highest b-tagging score**

- **MVA_Mbl2: invariant mass of the second b-jet+lepton system**

- **MVA_njet: number of jets**

- **MVA_nbjet: number of b-jets**

- **MVA_tM: transverse mass of the W boson candidate (from standard model top-quark)**

- **MVA_llDr: ΔR between flavor-violating electron and muon**

- **MVA_SSee_Zmass: invariant mass of a Same-Sign di-electron pair**

- **MVA_Topmass: invariant mass of the standard model top-quark candidate**

- **MVA_Met: missing transverse energy**

The following features are only used for TT classifier:

- **MVA_Ht: scalar sum of the p_T of all objects**

- **MVA_Mbl1: invariant mass of the second bjet+lepton system**

- **MVA_JeDr: ΔR between flavor-violating electron and a light jet (non-b-jet)**

- **MVA_JmuDr: ΔR between flavor-violating muon and a light jet (non-b-jet)**

The following features are used for ST classifier:

- **MVA_BaPt: p_T of the standard model lepton**

- **MVA_JetHt: scalar sum of the p_T of jets**

Distributions of selected features are shown in Figure 16.3-16.7.
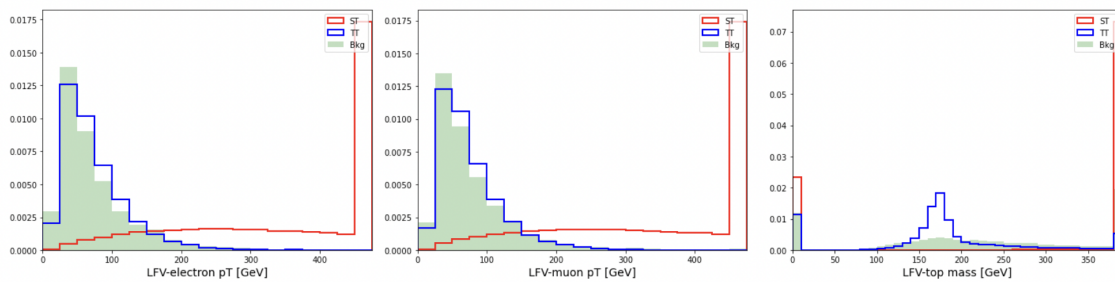
Figure 16.4: Normalized distribution of additional features in SR. From to left to right: MVA_LFVePt, MVA_LFVmuPt, MVA_LFVTopmass.
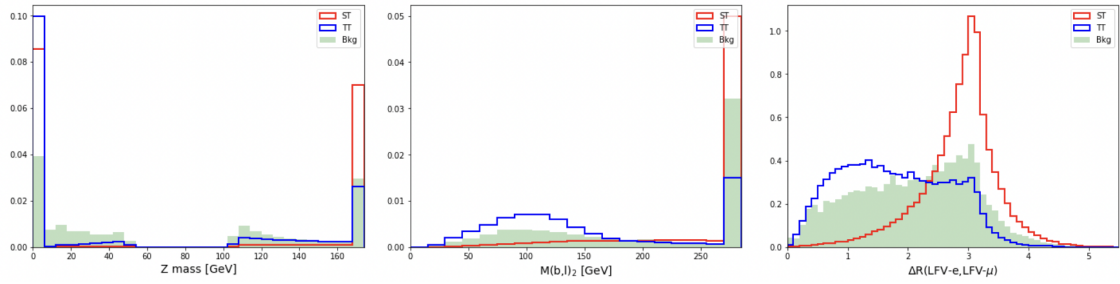
Figure 16.5: Normalized distribution of additional features in SR. From to left to right: MVA_Zmass, MVA_Mbl2, MVA_llDr.
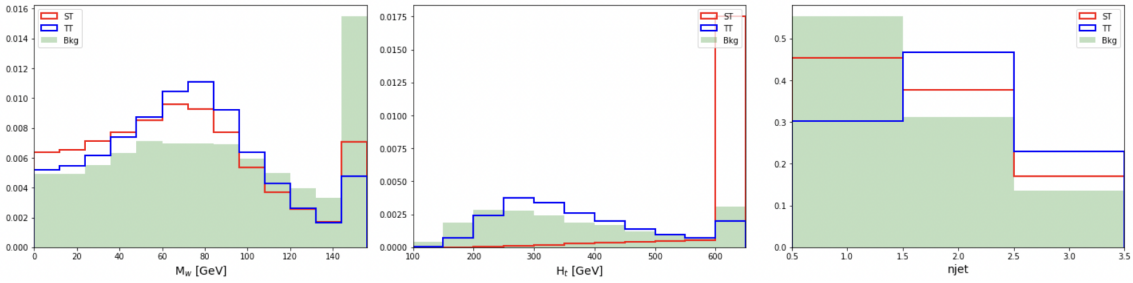
Figure 16.6: Normalized distribution of additional features in SR. From to left to right: MVA_tM, MVA_Ht, MVA_njet.
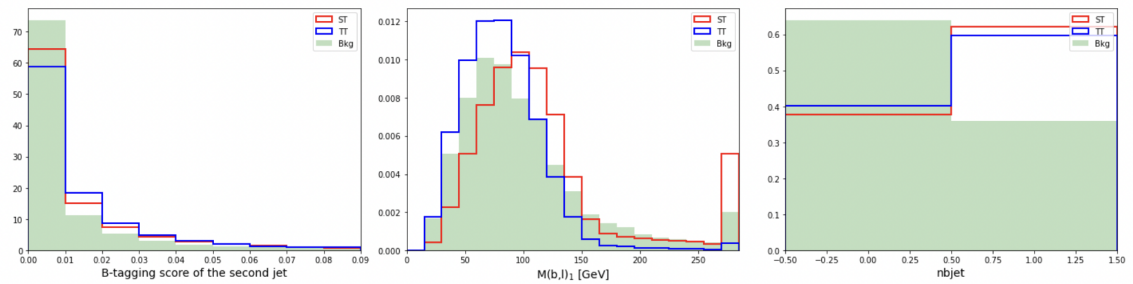
Figure 16.7: Normalized distribution of additional features in SR. From to left to right: MVA_Jet2Btag, MVA_Mbl1, MVA_nbjet.

16.2.1 Input ranking and correlation

To determine the relative importance of these features, feature importance is extracted using the "gain" method and is shown in Figure 16.10.
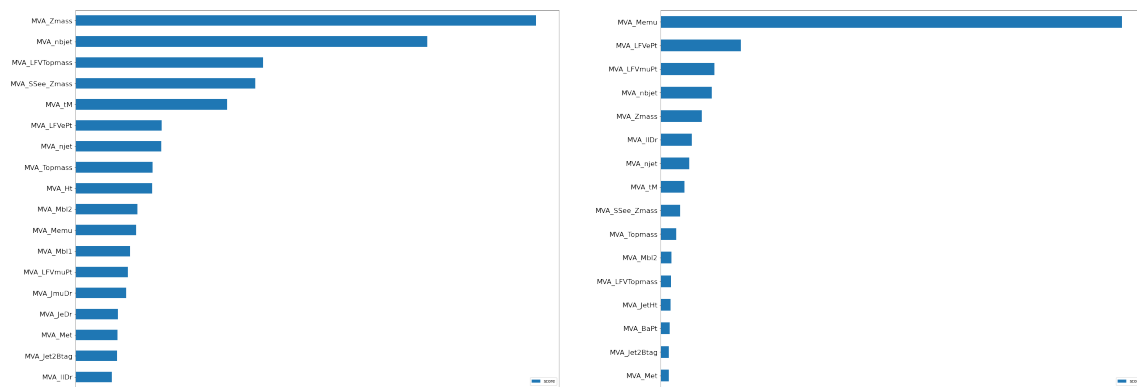
Figure 16.8: List of features ranked by their relative importance. From left to right: BDT targeting TT (SR1), BDT targeting ST (SR2)
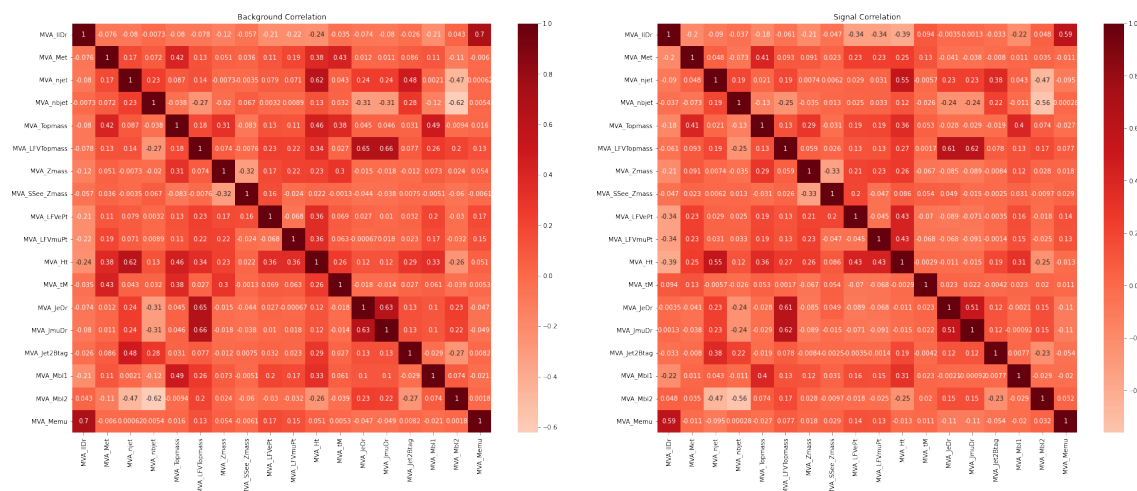
Figure 16.9: Correlation matrices (SR1), from left to right: background correlation, signal correlation.
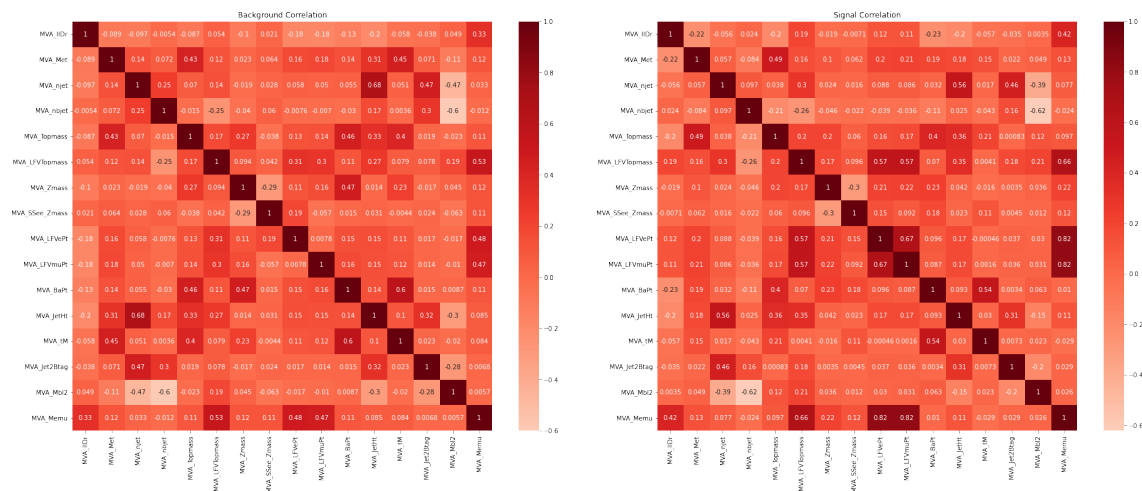
Figure 16.10: Correlation matrices (SR2), from left to right: background correlation, signal correlation.

16.2.2 Input validation

The background modeling of the input features are checked in SR. The input features are shown in Figure 16.11-16.17. Note: all input samples come from MC simulation.
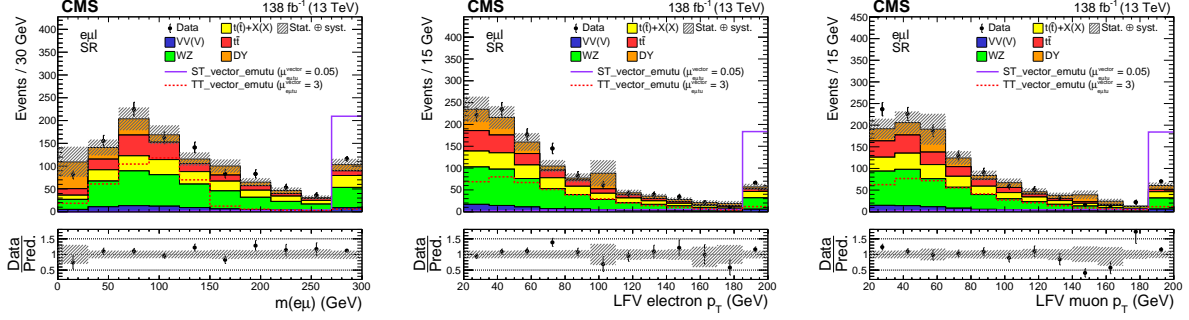
Figure 16.11: BDT input validation, from left to right: MVA_$M_{e\mu}$, MVA_LFVePt, MVA_LFVmuPt
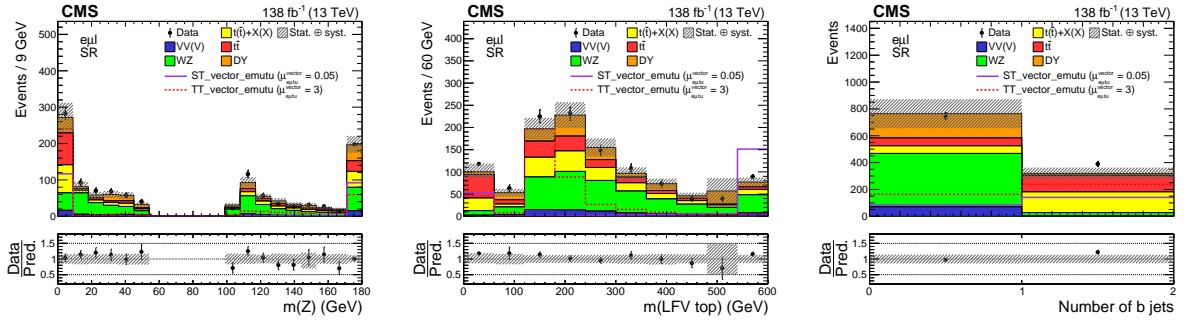
Figure 16.12: BDT input validation, from left to right: MVA_Zmass, MVA_LFVTopmass, MVA_nbjet
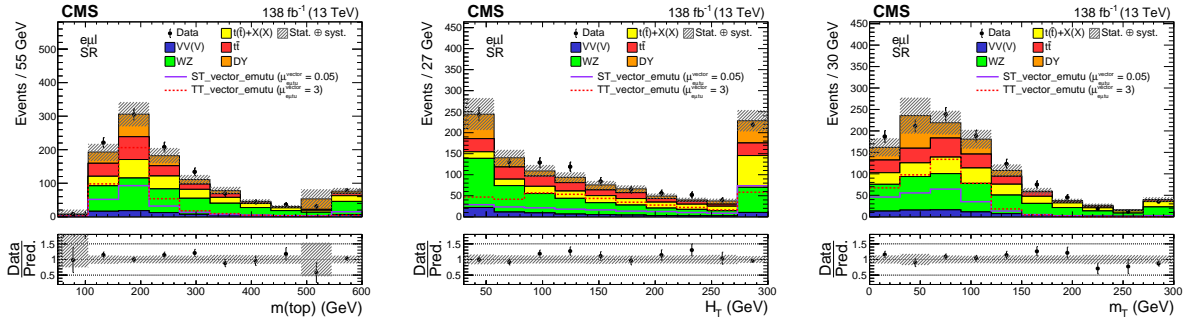
Figure 16.13: BDT input validation, from left to right: MVA_Topmass, MVA_JetHt, MVA_tM

Figure 16.14: BDT input validation, from left to right: MVA_njet, MVA_Ht, MVA_Met

Figure 16.15: BDT input validation, from left to right: MVA_SSZmass, MVA_JeDr, MVA_JmuDr

Figure 16.16: BDT input validation, from left to right: MVA_Mbl1, MVA_Mbl2, MVA_BaPt

16.3 BDT Output
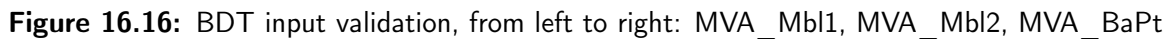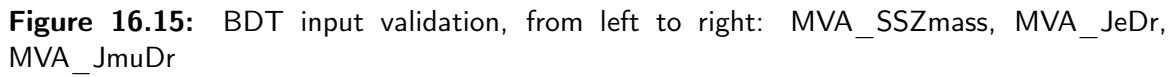
The output of the BDTs are shown in Figure 16.18. Note: all but the nonprompt backgrounds are estimated with MC simulation. The nonprompt background backgrounds are estimated with the matrix method.
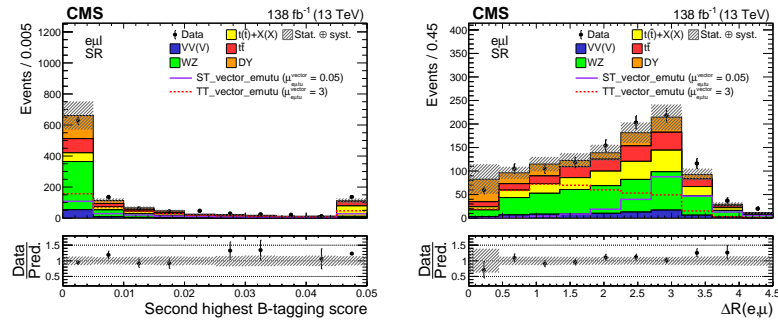
Figure 16.17: BDT input validation, from left to right: MVA_Jet2Btag, MVA_llDr
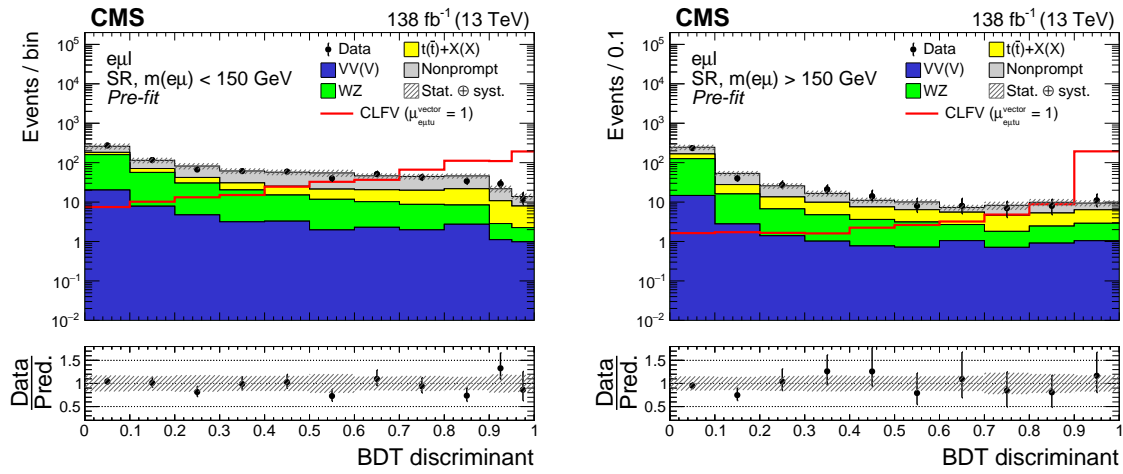
Figure 16.18: Distribution of BDT output, from left to right TT enriched SR (SR1), ST enriched SR (SR2)

CHAPTER 17

Systematic Uncertainties

For this analysis, we consider the following sources of systematic uncertainties:

17.1 Theoretical Uncertainties

- Prompt background cross-sections: These uncertainties arise from the uncertainties in the measured cross-sections of the processes which are used for scaling the MC to the luminosity of data. Systematic uncertainties associated to the prompt backgrounds WZ (as discussed in autorefsec:CRWZ), ZZ , ttW/ttZ/ttH, tZq are 6%, 6%, 15% and 20% respectively. Measurements performed on the WZ and ZZ processes currently have uncertainties of order 5-8% citeWZ2019 citeZZ2020. A 50 % normalization uncertainty is assigned to other smaller backgrounds.

- PDF: Using 100 the replicas of the NNPDF 3.1 set, the PDF uncertainties are obtained by taking the root mean square value of the variations. Only shape components of the variation is considered, this is done by normalizing LHE weights of the replicas such that the sum of the LHE wights (in SR) stays the same for each replica. This uncertainty treated as correlated. We consider this uncertainty for all the signals and major prompt backgrounds (i.e. WZ, TTZ, TTW and TTH).

- Renormalization and factorization scale: this uncertainty is evaluated by considering an envelope of 6 variation (μ_R 2, μ_R 0.5, μ_F 2, μ_F 0.5, μ_R and μ_F 2, μ_R and μ_F 0.5). This uncertainty is treated as correlated and only shape component is considered. We consider this uncertainty for all the signals and major prompt backgrounds (i.e. WZ, TTZ, TTW and TTH).

- ISR and FSR: initial state radiation and final state radiation uncertainty is evaluated by using the default set of variation (2 and 0.5). This uncertainty is treated as correlated and only shape component is considered. We consider this uncertainty for all the signals.

17.2 Nonprompt Uncertainties

We consider the following sources of systematic uncertainties (shape) associated with the *fake* efficiency.

- Prompt background correction

- Sample dependence

- Statistical uncertainty

An important systematic uncertainty comes from the estimate of prompt contamination in MR. As is discussed in **??**, prompt backgrounds (estimated with MC) are subtracted from total event yields measured in MR (data). A flat 20 % uncertainty (α in Equation 17.1) is assigned to the event yields of the prompt background and the resulting variation of f is taken as the uncertainty, which is estimated to be between 5 % and 15 %.

$$f = \frac{n_{data}^{T} - (1+\alpha)n_{Prompt\ MC}^{T}}{n_{data}^{T+\overline{T}} - (1+\alpha)n_{Prompt\ MC}^{T+\overline{T}}}. \tag{17.1}$$
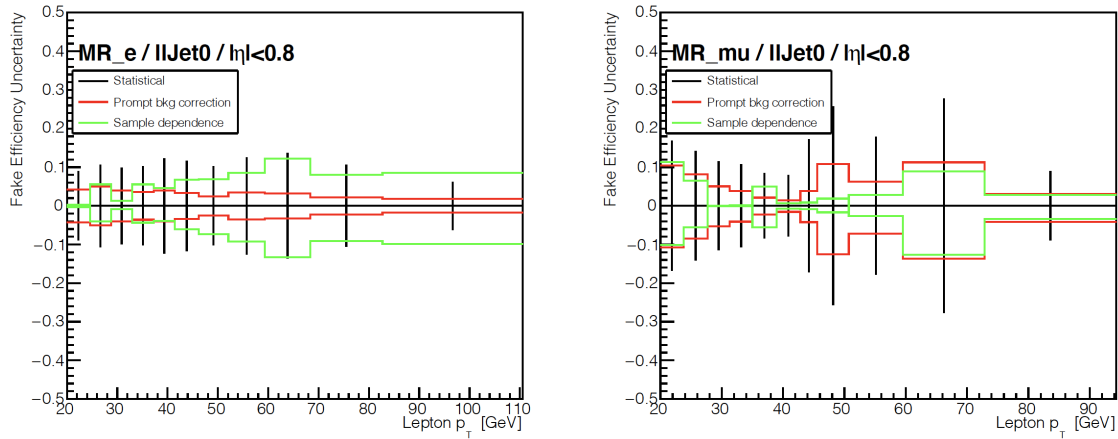
Figure 17.1: Comparison of different components of the uncertainties associated to the *fake* efficiency measured in 2017 dataset (njet=0 bin, $|\eta| <0.8$ bin). From left to right: electron f uncertainty, muon f uncertainty. Note: black uncertainty bars represent statistical uncertainties.

Sample dependence is one of the primary sources of systematic uncertainty and it is concerned with the possibility that the *fake* efficiency f measured in MR may not reflect characteristics of the non-prompt leptons in SR due to different compositions of backgrounds. This type of uncertainty is estimated by introducing a variation factor β between the proportions of Same-Flavor and Opposite-Flavor pairs in MR. For example, electron f can be calculated as (prompt background correction is ignored from the equation):

$$f_e = \frac{(1+\beta)n^T_{F,e+e} + (1-\beta)n^T_{F,e+\mu}}{(1+\beta)n^{T+\overline{T}}_{F,e+e} + (1-\beta)n^{T+\overline{T}}_{F,e+\mu}}. \tag{17.2}$$

A 20% variation (β) is assigned the resulting variation of f is taken as the uncertainty, which is estimated to be between 3 % and 15 %.

A comparison of different sources of systematic uncertainties (using 2017 datasets) are shown in Figure 17.1 and Figure 17.2. A complete summary of uncertainties associated to *real* and *fake* efficiency can be found in.

Figure 17.2: Comparison of different components of the uncertainties associated to the *fake* efficiency measured in 2017 dataset (njet>0 bin, $|\eta|$ <0.8 bin). From left to right: electron f uncertainty, muon f uncertainty. Note: black uncertainty bars represent statistical uncertainties.

In addition to statistical uncertainties, MC uncertainties are propagated to the *real* efficiency as sources of systematic uncertainties:

- Statistical uncertainty

- MC uncertainty

  – Electron reconstruction

  – b-tagging

  – JES/JER

  – Pile-up rewighting

  – ECAL prefiring

The uncertainties associated to the *real* efficiency are relatively small when compared to the *fake* efficiency uncertainties. A comparison of different sources of *real* efficiency uncertainties are shown in Figure 17.3.
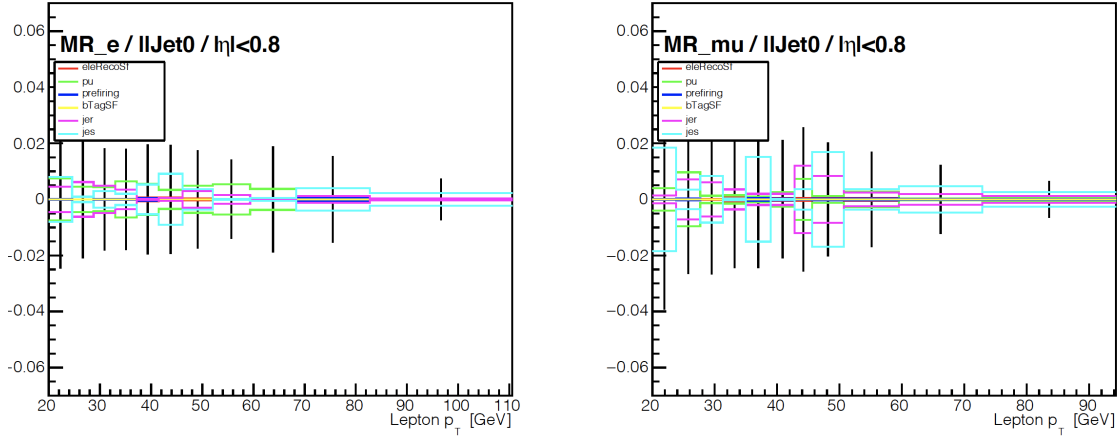
Figure 17.3: Comparison of different components of the uncertainties associated to the *real* efficiency measured in 2017 dataset (njet=0 bin, $|\eta| <$0.8 bin). From left to right: electron *r* uncertainty, muon *r* uncertainty. Note: black uncertainty bars represent statistical uncertainties.

In addition to the four components (i.e. r_e, r_μ, f_e, f_μ) of the shape uncertainties associated to the measurement of r and f, a fifth component is considered that accounts for the potential bias introduced when implementing 3D matrix method: Four out of the eight regions that appear on the lefthand side of the Equation ?? (i.e. $N^{\overline{T}TT}$, $N^{\overline{T}\overline{T}T}$, $N^{\overline{T}T\overline{T}}$, $N^{\overline{T}\overline{T}\overline{T}}$) are selected by requiring the leading lepton to fail the *tight* criteria (See Table 15.2). Effectively this means that the isolation requirement is reversed for leading lepton that enter these four regions. Selecting the leading lepton by a loose/reversed isolation requirement is not ideal since the leading lepton is required to match with iso-trigger. To account for this bias, a 50 % uncertainty is assigned to the f_1 (*fake* efficiency of the leading lepton) for events that enter these four regions. The variation of the non-prompt estimate due to trigger matching is largely covered by this uncertainty (See Figure 17.4).

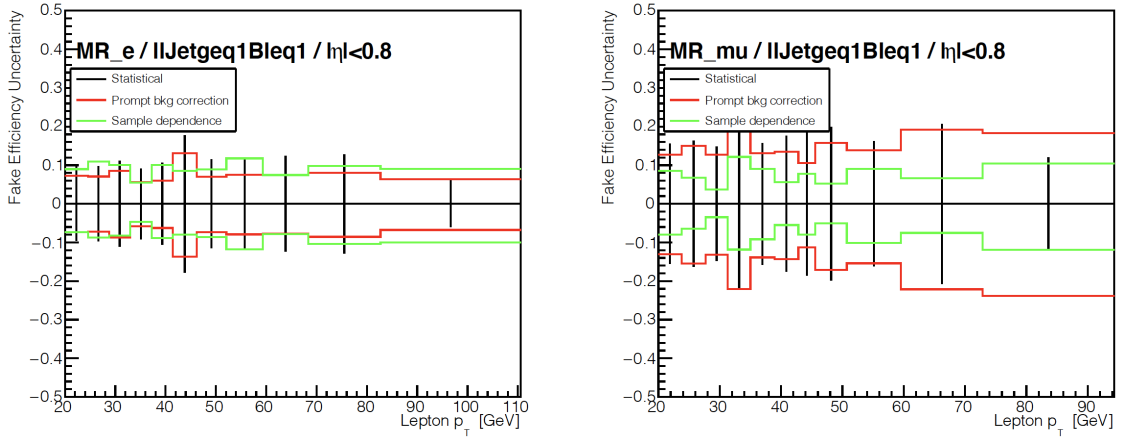The five components of the uncertainties discussed in this section are propagated through the matrix inversion. The resulting variations of the non-prompt estimates are taken as the shape uncertainties assigned to non-prompt backgrounds. These uncertainties are treated as correlated between the years. In addition to these shape uncertainties, an overall normalization uncertainty of 10% is assigned in order to cover any other potential bias.

17.3   Diboson Uncertainties

Mismodeling of the jet multiplicity is observed in WZ control region (See autorefsec:CRWZ). This is largely due to the fact that jets generated from diboson samples are coming from radiation (due to our trilepton selection). Therefore, the accuracy of the jet multiplicity distribution is suboptimal. To take this into account, a dedicated jet-dependent uncertainty is assigned to each
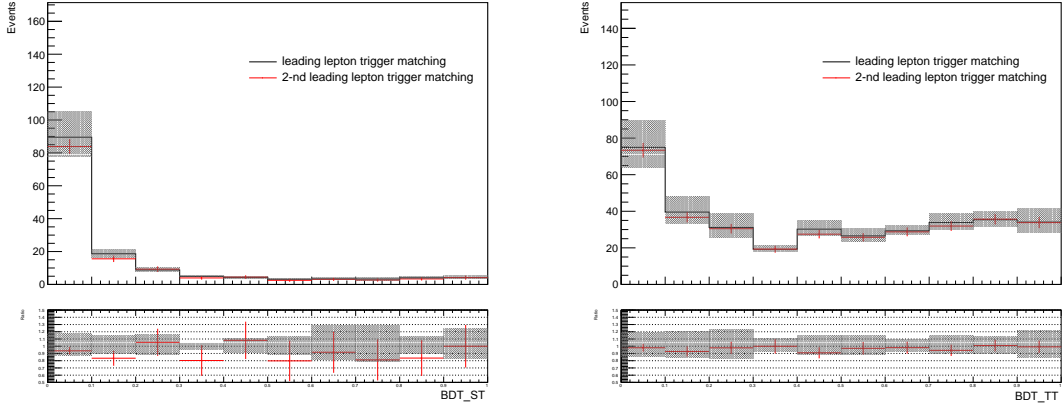
Figure 17.4: The impact of matching leptons to trigger objects on non-prompt estimate. From left to right: non-prompt estimate in ST enriched SR, non-prompt estimate in TT enriched SR. The nominal configuration of the matrix method is to match the leading lepton with trigger objects. Matching the 2nd-leading with the trigger objects is taken as an alternative to evaluate the robustness of the non-prompt estimate. The uncertainty band only covers the variation of the non-prompt estimate as a result of varying leading lepton f by 50 %. Uncertainty bars only include statistical uncertainties.

event. This uncertainty is determined using diboson control region:

- Exactly three leptons (any flavor composite),

- OnZ,

- MET$>$85GeV,

- Veto events with at least one b-tag jet ($p_T >$20GeV).

The jet multiplicity distributions are shown in Figure 17.5. For each year, a scale factor parameterized as bins of jet multiplicity is derived:

$$\epsilon = \frac{N_{data} - N_{VVV} - N_{TX} - N_{TTbar} - N_{others}}{N_{VV}}. \tag{17.3}$$

The scale factor ϵ is used to estimate the uncertainty, denoted by Δ:

$$\Delta = |1 - \epsilon| \tag{17.4}$$

This uncertainty is about 10-20%, as is shown in Figure 17.6.
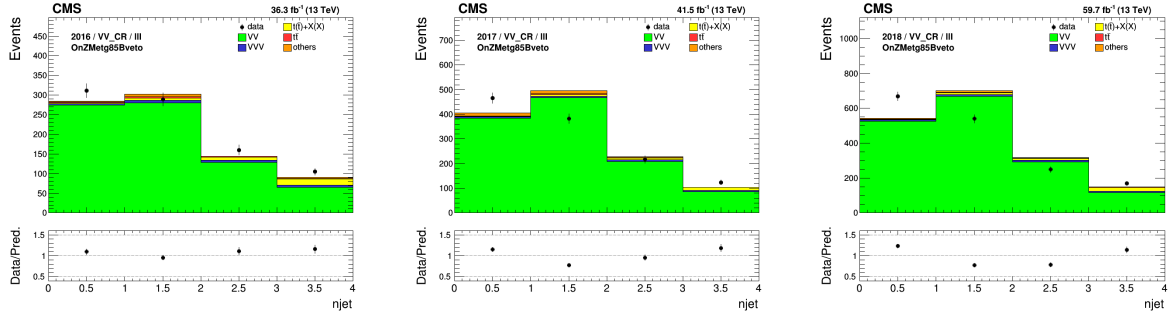
Figure 17.5: Diboson control region, from left to right: 2016, 2017 and 2018 datasets.
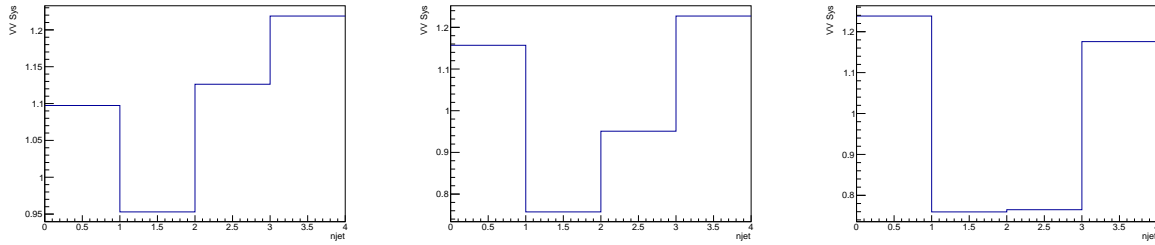
Figure 17.6: Scale factors derived from diboson control region, from left to right: 2016 and 2018 datasets.

17.4   Other Experimental Uncertainties

- Lepton reconstruction, identification, and isolation: electron/muon reconstruction uncertainties are provided by relevant POG. Uncertainties associated with lepton identification and isolation scale factors were evaluated by the authors of the mvaTOP ID citemvaTOP. The statistical components of these uncertainties are treated as uncorrelated while the other components are treated as fully correlated. For high p_T muons ($p_T >$ 200GeV), an additional uncertainty (denoted "muIDHighPt") is assigned and it increases linearly from 0 to 10% (200GeV-1000GeV) and is caped at 10% after 1000GeV.

- Muon scale uncertainties: uncertainties on muon momentum (denoted "MuonScale") are assigned using different method according to the muon p_T: Rochester algorithm is used for low p_T muons ($p_T <$ 200Gev), while GEciteGEmethod method is used for high p_T muons ($p_T >$ 200Gev).

- b-tagging: b-tagging efficiency and uncertainty is provided by the BTV group.

- trigger scale factor: trigger scale factors are set to 1 and a flat 2% uncertainty is assigned. This uncertainty is treated as uncorrelated.

- Luminosity: The uncertainties of 1.0%, 2.0% and 1.5% are assigned to the integrated luminosity for 2016, 2017, 2018, respectively. When processing the run 2 data the individual year uncertainties are treated as uncorrelated while there are two additional correlated luminosity uncertainties; firstly, 0.6%, 0.9% and 2.0% for each year respectively to account for the correlation between 2016, 2017 and 2018 data and secondly, 0.6%,0.2% for 2017 and 2018 respectively to account for the correlation between 2017 and 2018 data.

- jet energy scale and resolution: The jet energy scale, JES, and jet energy resolution, JER, are centrally provided by the JET/MET POG citeJEC. Variations of jet energy scale and resolution are propagated to the MET and b-tagging SF.

- pile-up reweighting: The measured minimum bias cross-section is varied up and down by 4.6%. This uncertainty is treated as correlated.

- MET unclustered missing energy is considered and treated as uncorrelated across the years.

- L1 ECAL prefiring: In the 2016 and 2017 datasets, L1 EGamma triggers fired early causing many uninteresting events to be recorded while the later interesting events were rejected. Since this effect is not present in the MC simulation, a tool citeECALPre provided by the L1 DPG is used to reweight MC events. This uncertainty is treated as correlated.

- HEM15/16 Issue: The HEM15/16 issue refers to two HCAL modules whose power supply died in the middle of the data taking (runs>=319077, i.e. last certified run of 2018B, and all of 2018C+D). The HEM issue is likely a very small effect but we still have to check it following the procedure in citeHEM

Uncertainties related to B-tagging scale factors are split into different sources. For b and udsg jets, we applied lf, hf, hfstats1/2, and lfstats1/2 uncertainties. For c jets, we applied cferr1/2 uncertainties. Correlations between different sources are specified in the Table ??.

Table 17.1: A hyphen ($-$) denotes that a source is not correlated between the different years.

Source	Correlated	Description
lf	✓	udsg+c jets in heavy flavor region
hf	✓	b+c jets in light flavor region
hfstats1	-	Linear fluctuations of c jets
hfstats2	-	Quadratic fluctuations of c jets
lfstats1	-	Linear fluctuations of udsg jets
lfstats2	-	Quadratic fluctuations of udsg jets
cferr1	✓	Linear fluctuations of c jets
cferr2	✓	Quadratic fluctuations of c jets

17.4.1 Jet energy scale uncertainties

Uncertainties associated with JES are split into their 27 components and properly treated the correlations of the split uncertainties by year as recommended by the JETMET POG and described in the Table 17.2.

Table 17.2: Summary of the sources of uncertainty related to the JECs. A hyphen (−) denotes that a source is not correlated between the different years.

Source	Correlated	Source	Correlated
AbsoluteStat	-	RelativePtHF	✓
AbsoluteScale	✓	RelativeBal	✓
AbsoluteMPFBias	✓	RelativeSample	-
Fragmentation	✓	RelativeFSR	-
SinglePionECAL	✓	RelativeStatFSR	✓
SinglePionHCAL	✓	RelativeStatEC	-
FlavorQCD	✓	RelativeStatHF	-
TimePtEta	-	PileUpDataMC	✓
RelativeJEREC1	-	PileUpPtRef	✓
RelativeJEREC2	-	PileUpPtBB	✓
RelativeJERHF	✓	PileUpPtEC1	✓
RelativePtBB	✓	PileUpPtEC2	✓
RelativePtEC1	-	PileUpPtHF	✓
RelativePtEC2	-		

B-tagging scale factors and MET vector are computed for each of the JES templates and treated as uncertainties that are fully correlated to the respective JES sources.

The overall systematic uncertainty on background is estimated to be about 15 % in SR.

A comparison of different sources of systematic uncertainties of the background estimates in SR is shown in Figure 17.7.

A comparison of different sources of systematic uncertainties of the signal estimates in SR is shown in Figure 17.8.

Representative range of systematic uncertainties on background and signal MC samples are summarized in citeSysError. These uncertainties are extracted from the last bins of the BDT distribution.
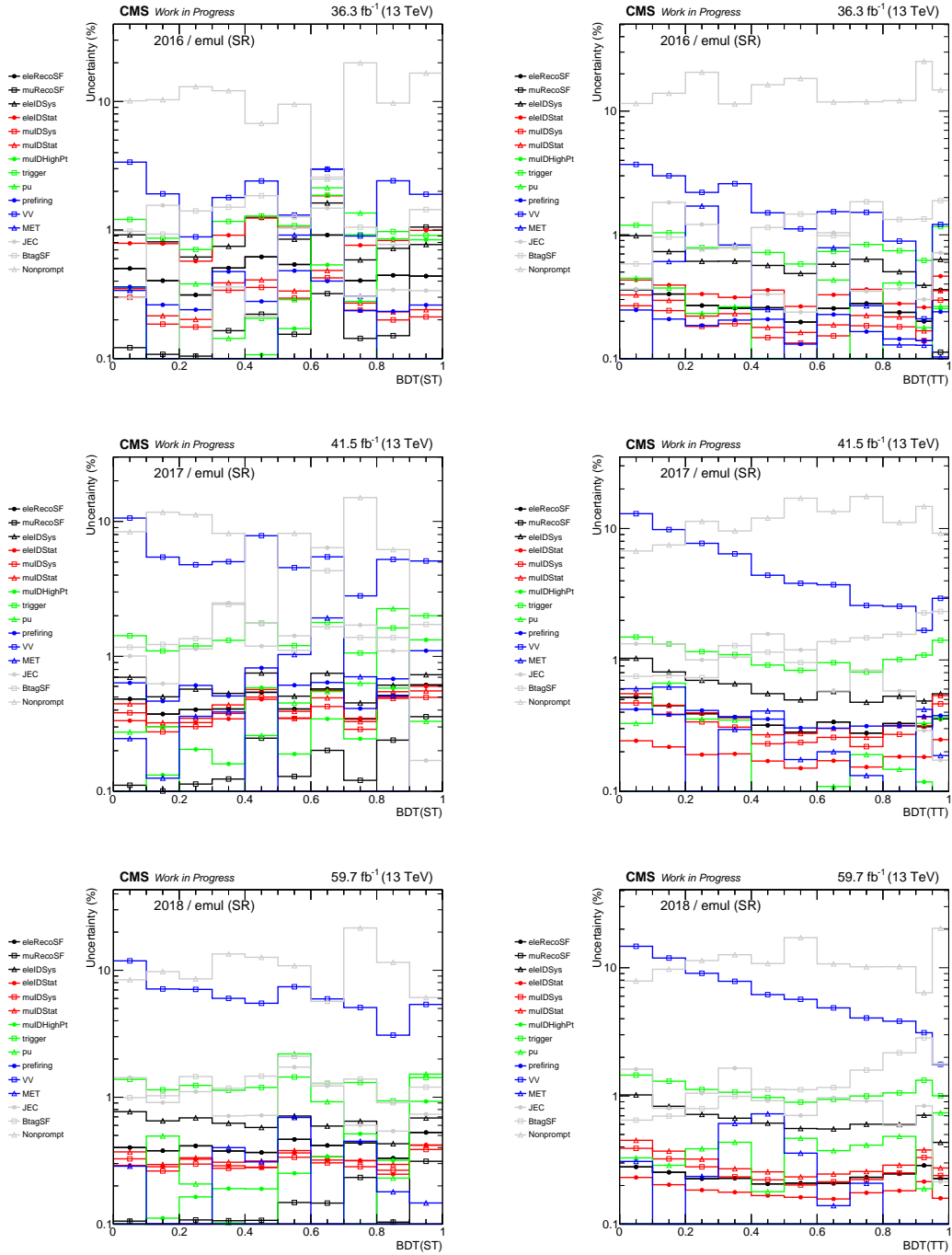
Figure 17.7: Distributions of relative uncertainties on total expected backgrounds as a function of BDT output in ST enriched SR (left), TT enriched SR (right). Luminosity and normalization uncertainties are not included in these plots. JES, JER and HEM are combined into "JEC". Sources of b-tagging uncertainties listed in Table ?? are combined into "BtagSF". From top to bottom: 2016, 2017 and 2018 datasets.
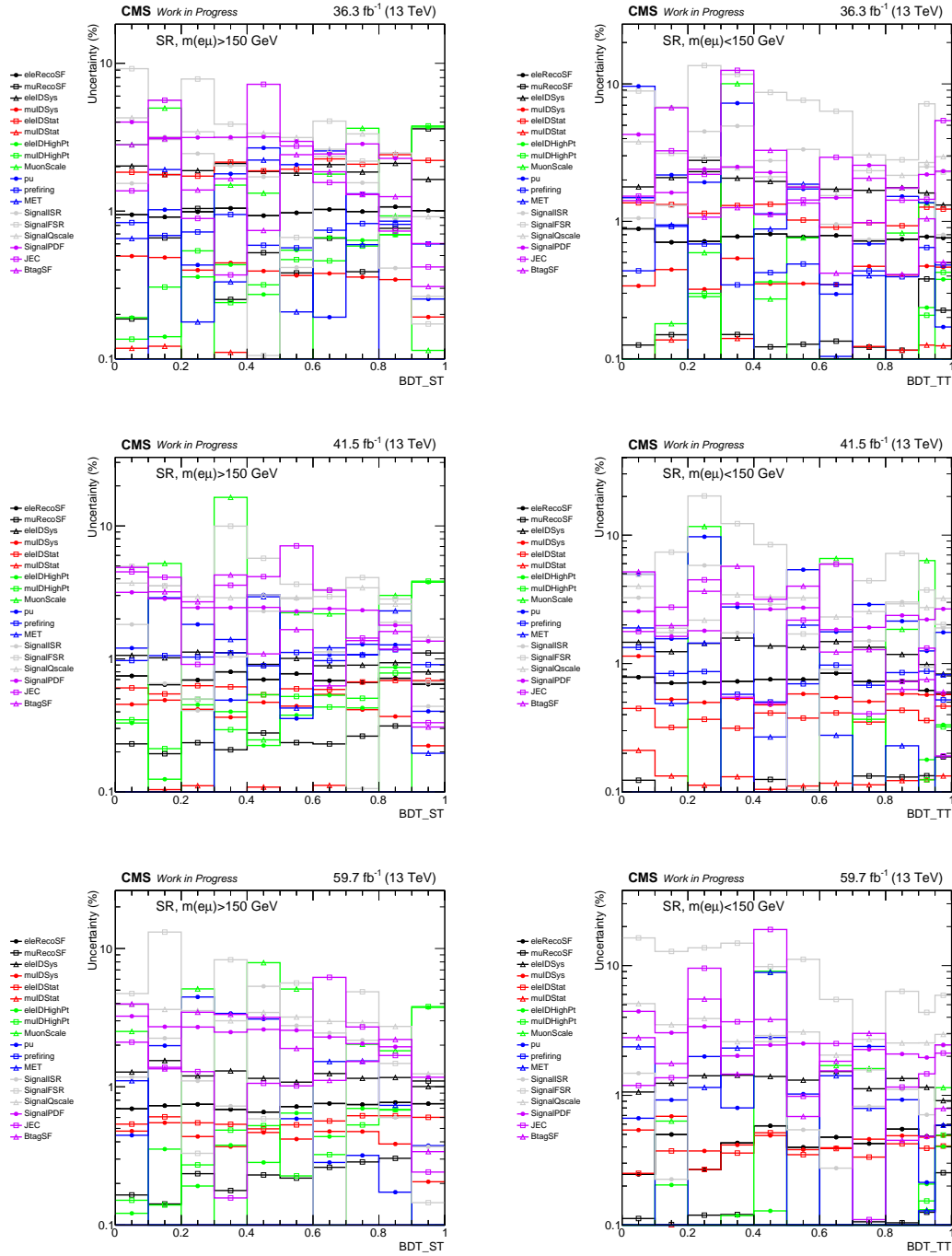
Figure 17.8: Distributions of relative uncertainties on signal ($C_{e\mu tu}^{vector}$ is used as an example) as a function of BDT output in ST enriched SR (left), TT enriched SR (right). Luminosity and normalization uncertainties are not included in these plots. JES, JER and HEM are combined into "JEC". Sources of b-tagging uncertainties listed in Table **??** are combined into "BtagSF". From top to bottom: 2016, 2017 and 2018 datasets.

CHAPTER 18

Statistical Analysis

18.1 Profile Likelihood Fit

A binned likelihood function $\mathcal{L}(\mu, \theta)$ is constructed to perform the statistical analysis on the BDT discriminator distributions. The parameter of interest (POI), denoted by μ, is the signal strength that governs the cross section of the ST and TT signals simultaneously. All the uncertainties are incorporated into the likelihood function as nuisance parameters, denoted by θ. The uncertainties that affect the shape of the BDT discriminator distributions are considered with Gaussian distributions while other uncertainties that only affect the normalizations are considered with log-normal distributions.

To control the systematic uncertainties, a profile likelihood fit is performed simultaneously in six regions (three data-taking years and two signal regions) by maximizing the likelihood function $\mathcal{L}(\hat{\mu}, \hat{\theta})$. Both $\hat{\mu}$ and $\hat{\theta}$ are the maximum likelihood estimators for the signal strength and nuisance parameters, respectively. The scaling of the signal for the purpose of a better visualization is optimized based on "VecU" signal, and it is kept the same for the other signals to compare their cross sections relative to each other. The post-fit distributions of the BDT discriminator are shown in Fig. 18.1-18.6. The most prominent uncertainties affecting the likelihood fit are the statistical uncertainties on MC samples.

18.2 Upper Limits

The results for the one-dimensional limits are summarized in Table 18.1. Assuming a linear relationship between $\mathcal{B}(t \rightarrow e\mu c)$ and $\mathcal{B}(t \rightarrow e\mu c)$ in the case of nonvanishing signals, the two-dimensional limits can also be obtained through interpolation (see Fig. 18.7). This analysis constitutes the most stringent limits on these processes to date.
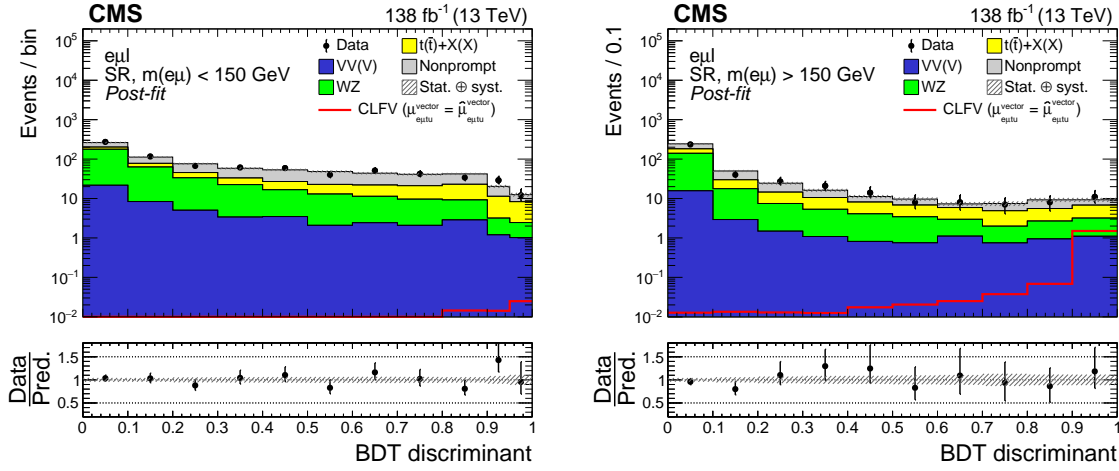
Figure 18.1: Distributions of the posterior BDT discriminator distributions for the TT-enriched SR (left) and the ST-enriched SR (right). Signals are generated with vector-like operator involving an Up quark. The three data-taking years are aggregated for better visualization.
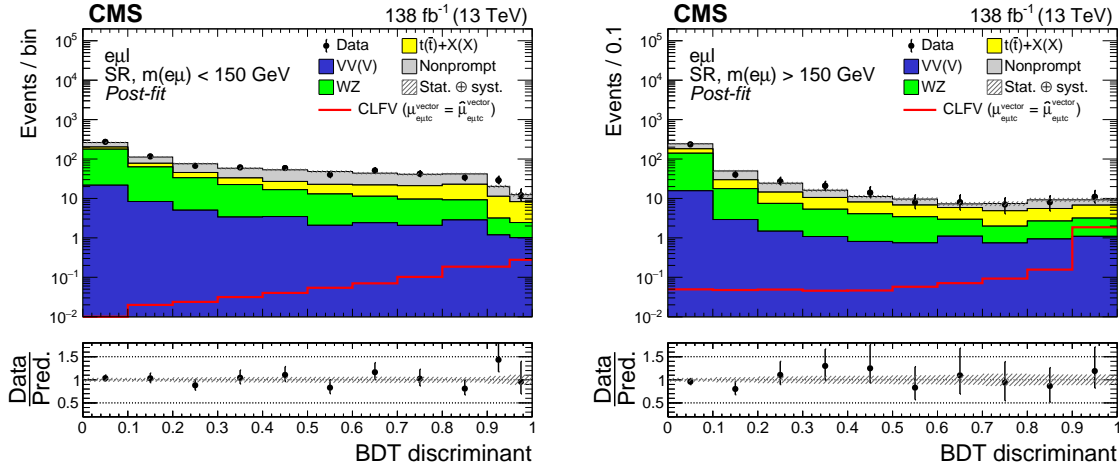
Figure 18.2: Distributions of the posterior BDT discriminator distributions for the TT-enriched SR (left) and the ST-enriched SR (right). Signals are generated with vector-like operator involving a Charm quark.
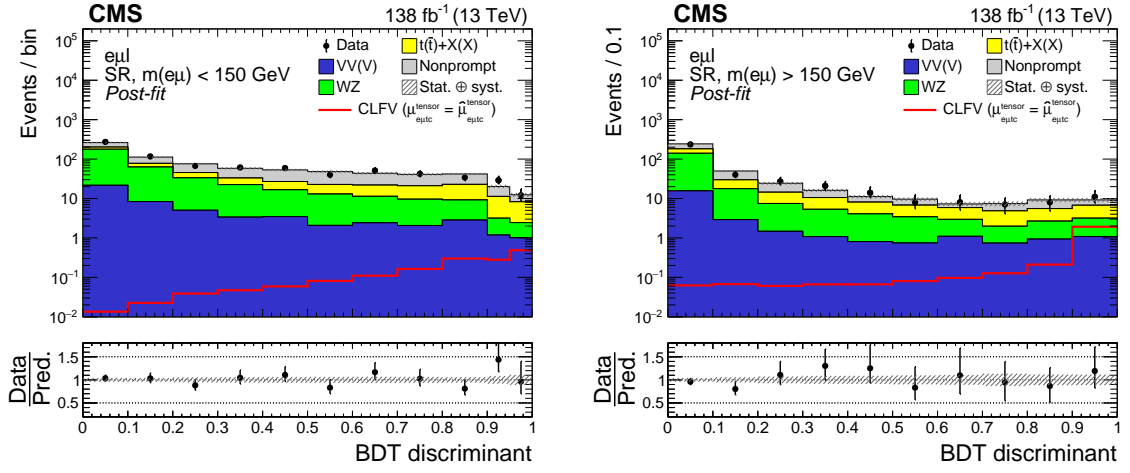
Figure 18.3: Distributions of the posterior BDT discriminator distributions for the TT-enriched SR (left) and the ST-enriched SR (right). Signals are generated with tensor-like operator involving an Up quark.
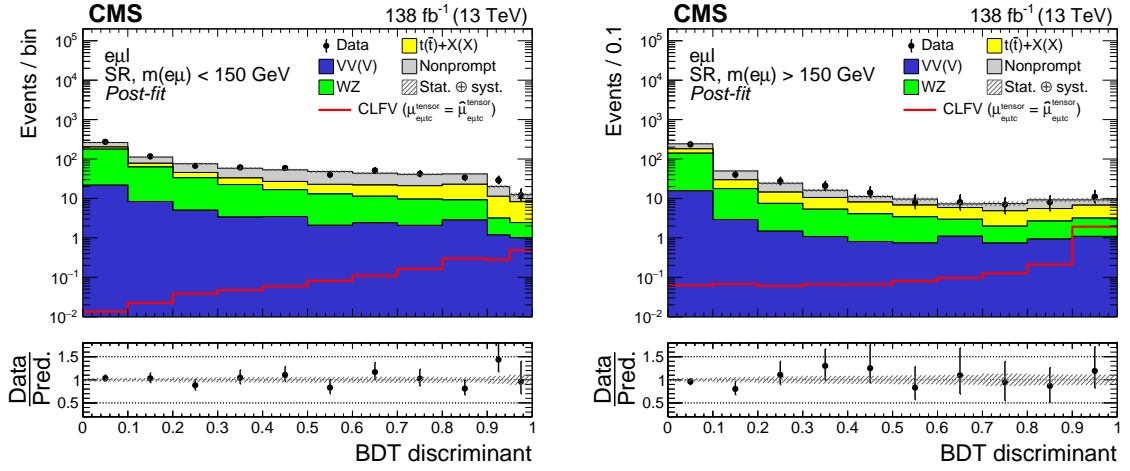
Figure 18.4: Distributions of the posterior BDT discriminator distributions for the TT-enriched SR (left) and the ST-enriched SR (right). Signals are generated with tensor-like operator involving a Charm quark.
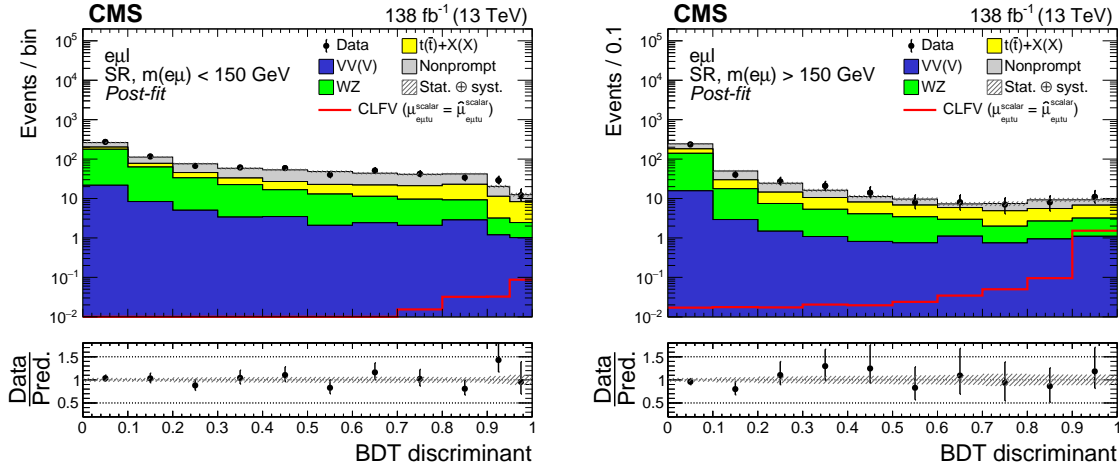
Figure 18.5: Distributions of the posterior BDT discriminator distributions for the TT-enriched SR (left) and the ST-enriched SR (right). Signals are generated with scalar-like operator involving an Up quark.
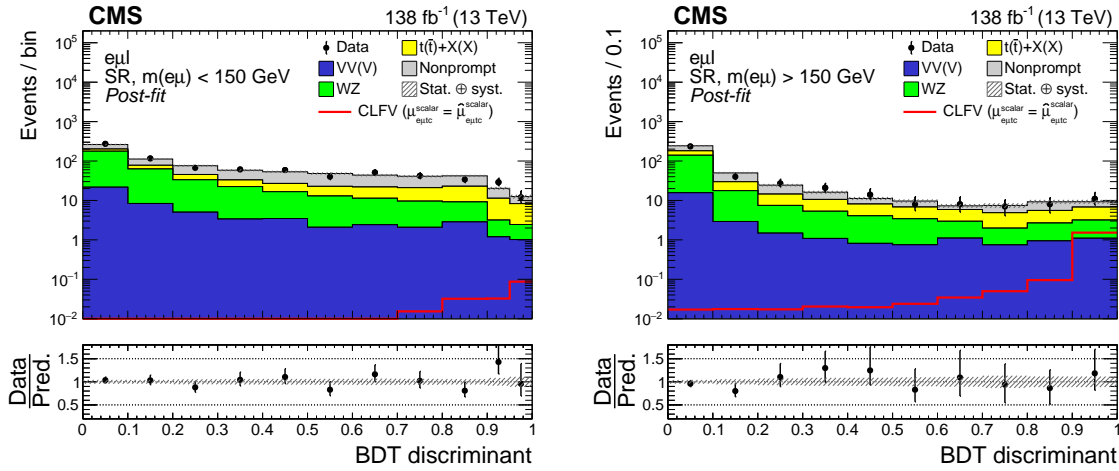
Figure 18.6: Distributions of the posterior BDT discriminator distributions for the TT-enriched SR (left) and the ST-enriched SR (right). Signals are generated with scalar-like operator involving a Charm quark.

CLFV coupling	Lorentz structure	$C_{e\mu tq}/\Lambda^2$ (TeV^{-2}) exp (68% range)	obs	$\mathcal{B}(t \to e\mu q) \times 10^{-6}$ exp (68% range)	obs
eμtu	tensor	0.022 (0.018–0.026)	**0.024**	0.027 (0.018–0.040)	**0.032**
	vector	0.044 (0.036–0.054)	**0.048**	0.019 (0.013–0.028)	**0.022**
	scalar	0.093 (0.077–0.114)	**0.101**	0.010 (0.007–0.016)	**0.012**
eμtc	tensor	0.084 (0.069–0.102)	**0.094**	0.396 (0.272–0.585)	**0.498**
	vector	0.175 (0.145–0.214)	**0.196**	0.296 (0.203–0.440)	**0.369**
	scalar	0.385 (0.318–0.471)	**0.424**	0.178 (0.122–0.266)	**0.216**

Table 18.1: Upper limit on the LFV signal using the full Run 2 data set.
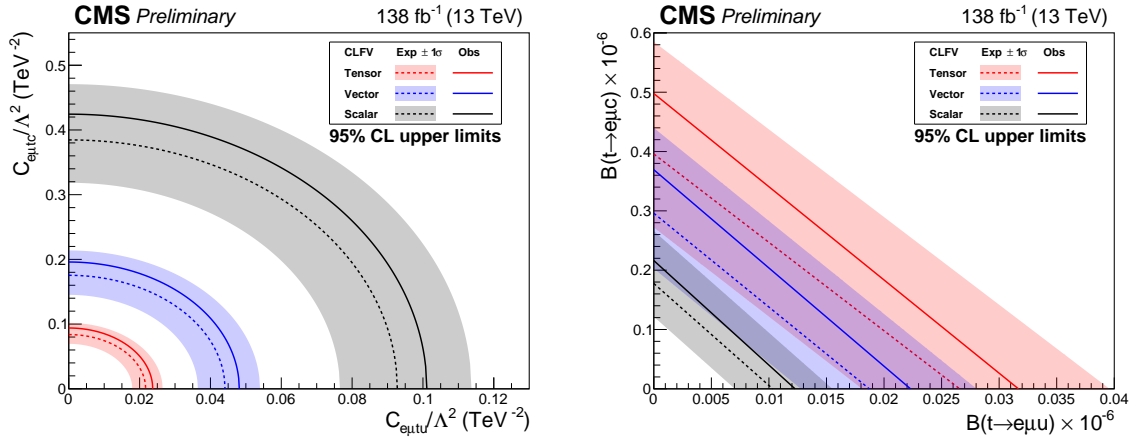
Figure 18.7: Two-dimensional upper limits on the Wilson coefficients (left) and the branching ratios (right).

Part IV

Outlook for CLFV Searches Using Top Quarks

CHAPTER 19

Inclusive Signal Generation

19.1   UFO Model

19.2   EFT Reweighting

CHAPTER 20

Object and Event Selection

20.1 Object Selection

20.2 Event Selection

Fake Tau Estimate

21.1 ABCD Method

21.2 Validation

CHAPTER 22

Expected Sensitivity

22.1   Asimov Fit

22.2   Expected Upper Limits

CHAPTER 23

Run-2 Combination

23.1   Correlation Scheme

23.2   Minimal Flavor Violation Interpretation

CHAPTER 24

Conclusion

References

[1] **CMS** Collaboration. *Search for charged-lepton flavor violation in top quark production and decay in pp collisions at $\sqrt{s} = 13$ TeV*. JHEP **06**, 082 (2022). `arXiv:2201.07859`, `doi:10.1007/JHEP06(2022)082`. (Cited on page 14.)

[2] D. Barducci et al. *Interpreting top-quark LHC measurements in the standard-model effective field theory* (2018). `arXiv:1802.07237`. (Cited on page 15.)

[3] A. Dedes, M. Paraskevas, J. Rosiek, K. Suxho, and L. Trifyllis. *SmeftFR − Feynman rules generator for the Standard Model Effective Field Theory*. Comput. Phys. Commun. **247**, 106931 (2020). `arXiv:1904.03204`, `doi:10.1016/j.cpc.2019.106931`. (Cited on page 16.)

[4] Celine Degrande, Claude Duhr, Benjamin Fuks, David Grellscheid, Olivier Mattelaer, and Thomas Reiter. *UFO - The Universal FeynRules Output*. Comput. Phys. Commun. **183**, 1201–1214 (2012). `arXiv:1108.2040`, `doi:10.1016/j.cpc.2012.01.022`. (Cited on page 16.)

[5] Neil D. Christensen and Claude Duhr. *FeynRules - Feynman rules made easy*. Comput. Phys. Commun. **180**, 1614–1641 (2009). `arXiv:0806.4194`, `doi:10.1016/j.cpc.2009.02.018`. (Cited on page 16.)

[6] J. Alwall, R. Frederix, S. Frixione, V. Hirschi, F. Maltoni, O. Mattelaer, H. S. Shao, T. Stelzer, P. Torrielli, and M. Zaro. *The automated computation of tree-level and next-to-leading order differential cross sections, and their matching to parton shower simulations*. JHEP **07**, 079 (2014). `arXiv:1405.0301`, `doi:10.1007/JHEP07(2014)079`. (Cited on page 16.)

[7] Ilaria Brivio, Yun Jiang, and Michael Trott. *The SMEFTsim package, theory and tools*. JHEP **12**, 070 (2017). `arXiv:1709.06492`, `doi:10.1007/JHEP12(2017)070`. (Cited on page 16.)

[8] Michal Czakon and Alexander Mitov. *Top++: A Program for the Calculation of the Top-Pair Cross-Section at Hadron Colliders*. Comput. Phys. Commun. **185**, 2930 (2014). `arXiv:1112.5675`, `doi:10.1016/j.cpc.2014.06.021`. (Cited on pages 17 and 19.)

[9] Jennifer Kile and Amarjit Soni. *Model-Independent Constraints on Lepton-Flavor-Violating Decays of the Top Quark*. Phys. Rev. D **78**, 094008 (2008). `arXiv:0807.4199`, `doi:10.1103/PhysRevD.78.094008`. (Cited on page 17.)

[10] **NNPDF** Collaboration. *Parton distributions for the LHC Run II*. JHEP **04**, 040 (2015). arXiv:1410.8849, doi:10.1007/JHEP04(2015)040. (Cited on page 18.)

[11] **NNPDF** Collaboration. *Parton distributions from high-precision collider data*. Eur. Phys. J. C **77(10)**, 663 (2017). arXiv:1706.00428, doi:10.1140/epjc/s10052-017-5199-5. (Cited on page 18.)

[12] Stefano Frixione, Paolo Nason, and Carlo Oleari. *Matching NLO QCD computations with Parton Shower simulations: the POWHEG method*. JHEP **11**, 070 (2007). arXiv:0709.2092, doi:10.1088/1126-6708/2007/11/070. (Cited on page 19.)

[13] Torbjörn Sjöstrand, Stefan Ask, Jesper R. Christiansen, Richard Corke, Nishita Desai, Philip Ilten, Stephen Mrenna, Stefan Prestel, Christine O. Rasmussen, and Peter Z. Skands. *An introduction to PYTHIA 8.2*. Comput. Phys. Commun. **191**, 159–177 (2015). arXiv:1410.3012, doi:10.1016/j.cpc.2015.01.024. (Cited on page 19.)

[14] **CMS** Colloboration. *Event generator tunes obtained from underlying event and multiparton scattering measurements*. Eur. Phys. J. C **76(3)**, 155 (2016). arXiv:1512.00815, doi:10.1140/epjc/s10052-016-3988-x. (Cited on page 19.)

[15] **CMS** Colloboration. *Extraction and validation of a new set of CMS PYTHIA8 tunes from underlying-event measurements*. Eur. Phys. J. C **80(1)**, 4 (2020). arXiv:1903.12179, doi:10.1140/epjc/s10052-019-7499-4. (Cited on page 19.)

[16] John M. Campbell, R. Keith Ellis, and Ciaran Williams. *Vector boson pair production at the LHC*. JHEP **07**, 018 (2011). arXiv:1105.0020, doi:10.1007/JHEP07(2011)018. (Cited on page 19.)

[17] Rikkert Frederix and Ioannis Tsinikos. *On improving NLO merging for $t\bar{t}W$ production*. JHEP **11**, 029 (2021). arXiv:2108.07826, doi:10.1007/JHEP11(2021)029. (Cited on page 19.)

[18] Anna Kulesza, Leszek Motyka, Daniel Schwartländer, Tomasz Stebel, and Vincent Theeuwes. *Associated top quark pair production with a heavy boson: differential cross sections at NLO+NNLL accuracy*. Eur. Phys. J. C **80(5)**, 428 (2020). arXiv:2001.03031, doi:10.1140/epjc/s10052-020-7987-6. (Cited on page 19.)

[19] Ye Li and Frank Petriello. *Combining QCD and electroweak corrections to dilepton production in FEWZ*. Phys. Rev. D **86**, 094034 (2012). arXiv:1208.5967, doi:10.1103/PhysRevD.86.094034. (Cited on page 19.)

[20] **CMS** Colloboration. *Observation of Single Top Quark Production in Association with a Z Boson in Proton-Proton Collisions at $\sqrt{s} =13$ TeV*. Phys. Rev. Lett. **122(13)**, 132003 (2019). arXiv:1812.05900, doi:10.1103/PhysRevLett.122.132003. (Cited on page 21.)

[21] **CMS** Colloboration. *Inclusive and differential cross section measurements of single top quark production in association with a Z boson in proton-proton collisions at $\sqrt{s} = 13$ TeV.* JHEP **02**, 107 (2022). `arXiv:2111.02860`, `doi:10.1007/JHEP02(2022)107`. (Cited on pages 21 and 22.)

[22] **TMVA** Colloboration. *TMVA - Toolkit for Multivariate Data Analysis* (2007). `arXiv: physics/0703039`. (Cited on page 21.)

[23] Matteo Cacciari and Gavin P. Salam. *Pileup subtraction using jet areas.* Phys. Lett. B **659**, 119–126 (2008). `arXiv:0707.1378`, `doi:10.1016/j.physletb.2007.09.077`. (Cited on page 22.)

[24] Emil Bols, Jan Kieseler, Mauro Verzetti, Markus Stoye, and Anna Stakia. *Jet Flavour Classification Using DeepJet.* JINST **15(12)**, P12012 (2020). `arXiv:2008.10519`, `doi: 10.1088/1748-0221/15/12/P12012`. (Cited on page 24.)

List of Trigger Paths

Table A.1: Triggers used to record events during data taking in 2016.

Dataset	Trigger Path
SingleMuon	HLT_IsoMu22_eta2p1, HLT_IsoTkMu22_eta2p1
	HLT_IsoMu24, HLT_IsoTkMu24
	HLT_Mu50, HLT_TkMu50, HLT_Mu45_eta2p1
SingleElectron	HLT_Ele25_eta2p1_WPTight_Gsf
	HLT_Ele27_WPTight_Gsf
	HLT_Ele27_eta2p1_WPTight_Gsf
	HLT_Ele32_eta2p1_WPTight_Gsf
	HLT_Ele105_CaloIdVT_GsfTrkIdT
	HLT_Ele115_CaloIdVT_GsfTrkIdT
DoubleMuon	HLT_Mu17_TrkIsoVVL_Mu8_TrkIsoVVL
	HLT_Mu17_TrkIsoVVL_TkMu8_TrkIsoVVL
	HLT_TkMu17_TrkIsoVVL_TkMu8_TrkIsoVVL
	HLT_Mu17_TrkIsoVVL_Mu8_TrkIsoVVL_DZ
	HLT_Mu17_TrkIsoVVL_TkMu8_TrkIsoVVL_DZ
	HLT_TkMu17_TrkIsoVVL_TkMu8_TrkIsoVVL_DZ
	HLT_Mu30_TkMu11, HLT_TripleMu_12_10_5
DoubleEG	HLT_Ele23_Ele12_CaloIdL_TrackIdL_IsoVL_DZ
	HLT_Ele16_Ele12_Ele8_CaloIdL_TrackIdL
	HLT_DoubleEle33_CaloIdL_MW
	HLT_DoubleEle33_CaloIdL_GsfTrkIdVL
	HLT_DoubleEle33_CaloIdL_GsfTrkIdVL_MW
MuonEG	HLT_Mu23_TrkIsoVVL_Ele8_CaloIdL_TrackIdL_IsoVL
	HLT_Mu23_TrkIsoVVL_Ele8_CaloIdL_TrackIdL_IsoVL_DZ
	HLT_Mu8_TrkIsoVVL_Ele23_CaloIdL_TrackIdL_IsoVL
	HLT_Mu30_Ele30_CaloIdL_GsfTrkIdVL
	HLT_Mu33_Ele33_CaloIdL_GsfTrkIdVL
	HLT_DiMu9_Ele9_CaloIdL_TrackIdL
	HLT_Mu8_DiEle12_CaloIdL_TrackIdL

Table A.2: Triggers used to record events during data taking in 2017.

Dataset	Trigger Path
SingleMuon	HLT_IsoMu24_eta2p1
	HLT_IsoMu27
	HLT_Mu50
	HLT_OldMu100
	HLT_TkMu100
SingleElectron	HLT_Ele32_WPTight_Gsf_L1DoubleEG
	HLT_Ele35_WPTight_Gsf
	HLT_Ele115_CaloIdVT_GsfTrkIdT
DoubleMuon	HLT_Mu17_TrkIsoVVL_Mu8_TrkIsoVVL
	HLT_Mu17_TrkIsoVVL_Mu8_TrkIsoVVL_DZ
	HLT_Mu17_TrkIsoVVL_Mu8_TrkIsoVVL_DZ_Mass8
	HLT_Mu17_TrkIsoVVL_Mu8_TrkIsoVVL_DZ_Mass3p8
	HLT_Mu19_TrkIsoVVL_Mu9_TrkIsoVVL_DZ_Mass3p8
	HLT_Mu37_TkMu27
	HLT_TripleMu_12_10_5
	HLT_TripleMu_10_5_5_DZ
	HLT_TripleMu_5_3_3_Mass3p8to60_DZ
DoubleEG	HLT_Ele23_Ele12_CaloIdL_TrackIdL_IsoVL
	HLT_Ele16_Ele12_Ele8_CaloIdL_TrackIdL
	HLT_DoubleEle25_CaloIdL_MW
	HLT_DoubleEle33_CaloIdL_MW
	HLT_DiEle27_WPTightCaloOnly_L1DoubleEG
MuonEG	HLT_Mu23_TrkIsoVVL_Ele12_CaloIdL_TrackIdL_IsoVL
	HLT_Mu23_TrkIsoVVL_Ele12_CaloIdL_TrackIdL_IsoVL_DZ
	HLT_Mu8_TrkIsoVVL_Ele23_CaloIdL_TrackIdL_IsoVL_DZ
	HLT_Mu12_TrkIsoVVL_Ele23_CaloIdL_TrackIdL_IsoVL_DZ
	HLT_Mu27_Ele37_CaloIdL_MV
	HLT_Mu37_Ele27_CaloIdL_MV
	HLT_DiMu9_Ele9_CaloIdL_TrackIdL
	HLT_DiMu9_Ele9_CaloIdL_TrackIdL_DZ
	HLT_Mu8_DiEle12_CaloIdL_TrackIdL

**Table A.3:** Triggers used to record events during data taking in 2018.

| Dataset	Trigger Path
SingleMuon	HLT_IsoMu24
	HLT_IsoMu27
	HLT_Mu50
	HLT_OldMu100
	HLT_TkMu100
EGamma	HLT_Ele32_WPTight_Gsf
	HLT_Ele115_CaloIdVT_GsfTrkIdT
	HLT_Ele23_Ele12_CaloIdL_TrackIdL_IsoVL
	HLT_Ele16_Ele12_Ele8_CaloIdL_TrackIdL
	HLT_DoubleEle25_CaloIdL_MW
	HLT_DiEle27_WPTightCaloOnly_L1DoubleEG
DoubleMuon	HLT_Mu17_TrkIsoVVL_Mu8_TrkIsoVVL_DZ_Mass3p8
	HLT_Mu37_TkMu27
	HLT_TripleMu_12_10_5
	HLT_TripleMu_10_5_5_DZ
	HLT_TripleMu_5_3_3_Mass3p8to60_DZ
MuonEG	HLT_Mu23_TrkIsoVVL_Ele12_CaloIdL_TrackIdL_IsoVL
	HLT_Mu23_TrkIsoVVL_Ele12_CaloIdL_TrackIdL_IsoVL_DZ
	HLT_Mu8_TrkIsoVVL_Ele23_CaloIdL_TrackIdL_IsoVL_DZ
	HLT_Mu12_TrkIsoVVL_Ele23_CaloIdL_TrackIdL_IsoVL_DZ
	HLT_Mu27_Ele37_CaloIdL_MW
	HLT_Mu37_Ele27_CaloIdL_MW
	HLT_DiMu9_Ele9_CaloIdL_TrackIdL_DZ
	HLT_Mu8_DiEle12_CaloIdL_TrackIdL

APPENDIX B

Nuisance Parameter Impact

The observed and expected ($\mu_{exp} = 0$) impacts of the nuisance parameters on the profile likelihood fit are shown in Figure B.1-B.2. They were computed using the following commands and plotted below for full run 2.

The expected ($\mu_{exp} = 1$) impacts of the nuisance parameters on the profile likelihood fit are shown in Figure B.3.
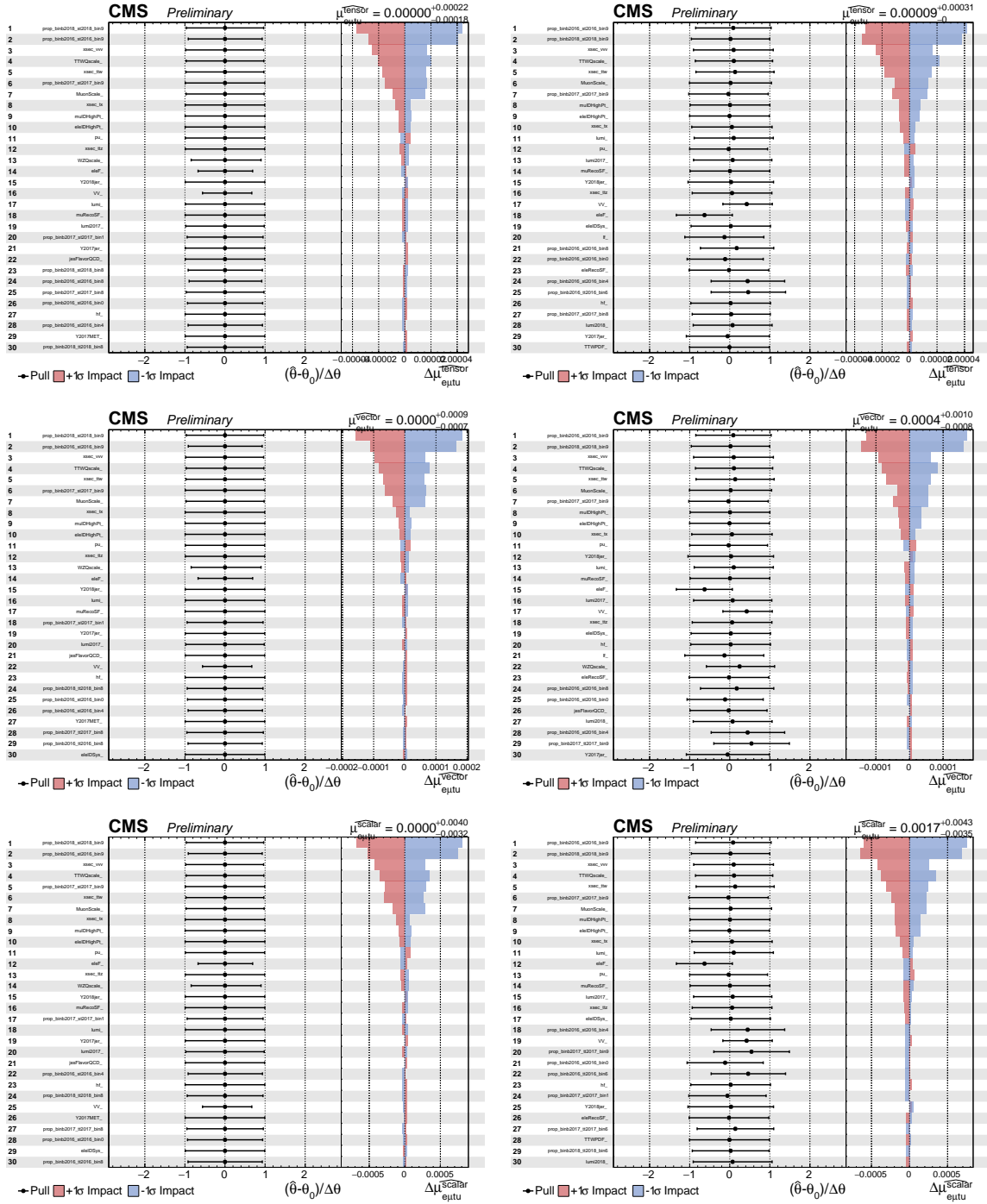
Figure B.1: Impacts of nuissance parameters for run II limit setting. From top to bottom: $e\mu tu$-tensor, $e\mu tu$-vector, $e\mu tu$-scalar. From left to right: expected impact (expected signal strength at 0), observed impact.

Figure B.2: Impacts of nuissance parameters for run II limit setting. From top to bottom: $e\mu tc$-tensor, $e\mu tc$-vector, $e\mu tc$-scalar. From left to right: expected impact (expected signal strength at 0), observed impact.
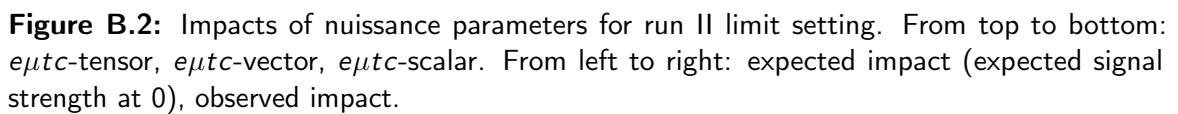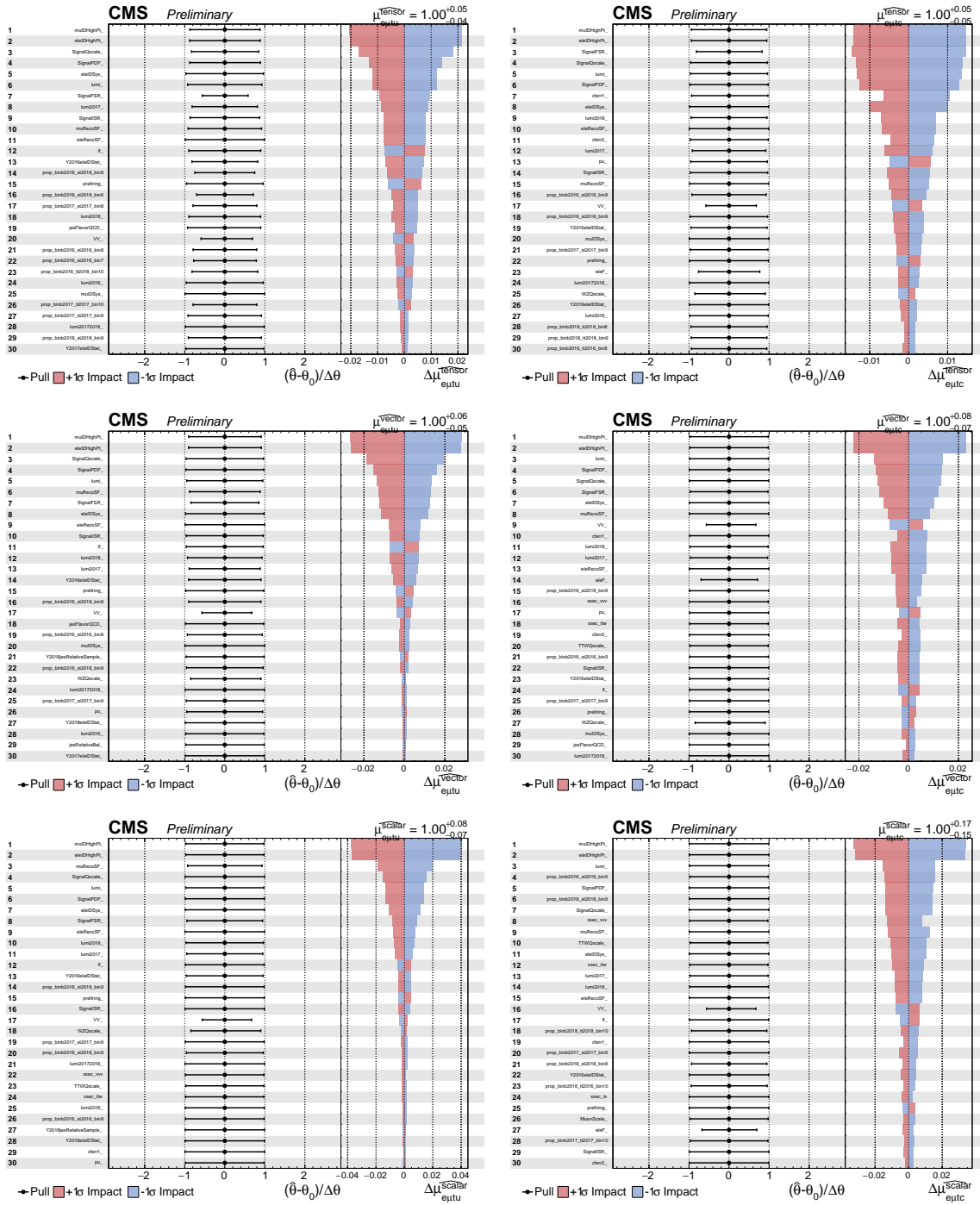
Figure B.3: Expected impact with an expected signal strength at 1. From top to bottom: tensor, vector, scalar. From left to right: $e\mu tu$, $e\mu tc$.

APPENDIX C

More materials