PROCEEDINGS OF SCIENCE

Top physics results from CMS

Mintu Kumar[a,*]

[a] *Tata Institute of Fundamental Research, Homi Bhabha Road, Mumbai, India*

E-mail: mintu.kumar@cern.ch

Recent results on the top quark physics related to the test of various discrete symmetries of the standard model have been presented. These include the search for charge asymmetries and violation of charged lepton flavor and charge-parity symmetries, conducted using proton-proton collision data collected by the CMS experiment during 2015-2018. The obtained results are compared with predictions of the standard model towards constraining physics beyond it.

8th Symposium on Prospects in the Physics of Discrete Symmetries (DISCRETE 2022)
7-11 November 2022
Baden-Baden, Germany

*Speaker

PoS(DISCRETE2022)057

1. Introduction

The top quark is the heaviest fundamental particle in the standard model (SM). Consequently, it is the most significant contributor among all elementary particles to radiative corrections to the mass of the W, Z, and Higgs bosons. At LHC, the top quark pair ($t\bar{t}$) production is the dominant process where the top quarks are produced copiously via the strong interaction. The cross section for $t\bar{t}$ production is about 831pb at the center-of-mass energy $\sqrt{s}$ = 13TeV. The second dominant process is the *t*-channel followed by the tW-channel, where the single top quarks are produced via weak interaction with a cross-section of $216.99^{+9.04}_{-7.71}$ pb and 71.7 ± 3.8 pb, respectively. One can perform symmetry measurements like the charge asymmetry (A_C), charge parity violation (CPV), charge-parity-time reversal symmetry (CPT), and charged lepton flavor violation in SM as well as the search for the physics beyond the standard model (BSM) in the top quark sector. This document reports the latest results on all these symmetry measurements from the CMS [1] experiment at LHC.

2. Charge asymmetry

At the LHC, in the $Qq \rightarrow t\bar{t}$ process, the top quark is expected to be produced in the direction of the incoming quark while the top antiquark is produced centrally. The top quark has a broader pseudorapidity distribution compared to the top antiquark quarks. These processes are symmetric under the charge conjugation at leading order, but higher order contribution gives rise to the charge asymmetry. Since the relative contribution of the valance quarks increases at high momentum transfer, A_c is measured in the $t\bar{t}$ process with the lepton+jets final state using highly Lorentz-boosted top quarks [2]. Such top quarks make searches more stringent probes of the QCD, as well as sensitive to BSM physics. In the final state, there should be one isolated charged lepton (e/μ) and two reconstructed jets with a cone radius of 0.4. To suppress the QCD background, $p_{\mathrm{T}}^{\mathrm{miss}} > 50$ GeV and $p_{\mathrm{T}}^{\mathrm{miss}} + p_{\mathrm{T}}^{\mu} > 150$ GeV ($p_{\mathrm{T}}^{\mathrm{miss}} > 120$ GeV) in the μ + jets (e + jets) channel. The larger value of the e + jets $p_{\mathrm{T}}^{\mathrm{miss}}$ requirement efficiently reduces the larger QCD multijet background in this channel and obviates the need for a separate requirement on $p_{\mathrm{T}}^{\mathrm{miss}} + p_{\mathrm{T}}^{\mathrm{e}}$. This measurement is done in three categories: boosted, semi-resolved, and resolved based on the soft drop cleaning [3] of the jets and the jet subjetiness ratio in the final state. One top quark is tagged in the boosted category, and no W boson is tagged using the tagging algorithms. One W boson is tagged in the semi-resolved category, but no top quark is tagged. In the resolved category, neither the top quark nor the W boson is tagged. The top quark and the W boson tagging are applied only on the jets reconstructed with a cone radius of 0.8 with $p_{\mathrm{T}} > 400$ GeV. The A_C is defined as the rapidity difference between the top quark and antiquark ($\Delta|y|$) :

$$A_C = \frac{N(\Delta|y| > 0) - N(\Delta|y| < 0)}{N(\Delta|y| > 0) + N(\Delta|y| < 0)} \tag{1}$$

This measurement is further split into two categories based on the $t\bar{t}$ invariant mass ($M_{t\bar{t}}$), namely $750 < M_{t\bar{t}} < 900$ GeV and $M_{t\bar{t}} > 900$ GeV, which allows us to focus on the region of interest. A simultaneous maximum likelihood (ML) fit has been performed using the Δy distribution in 12 channels: two lepton flavors (μ + jets and e + jets), three years (2018, 2017, and 2016), and two mass regions ($750 < M_{t\bar{t}} < 900$ GeV and $M_{t\bar{t}} > 900$ GeV). The final fit distribution is shown in Fig.1.

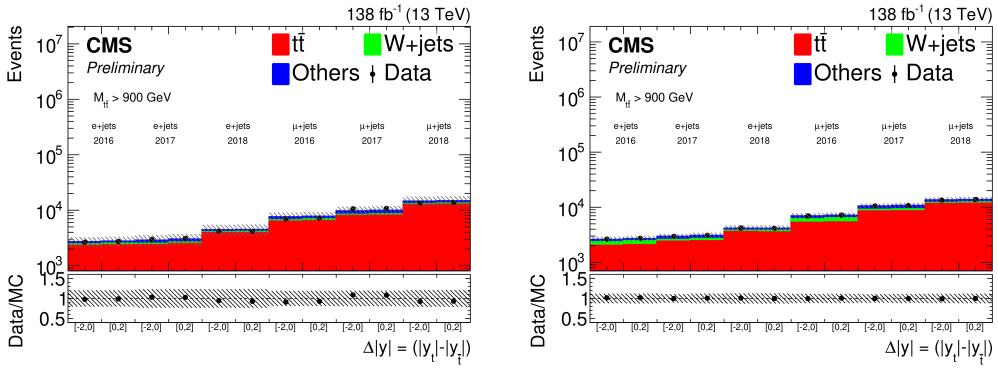

Figure 1: Postfit distributions of $\Delta|y|$ for the regions $750 < M_{t\bar{t}} < 900$ GeV (left) and $M_{t\bar{t}} > 900$ GeV (right) in the lepton+jets final state.

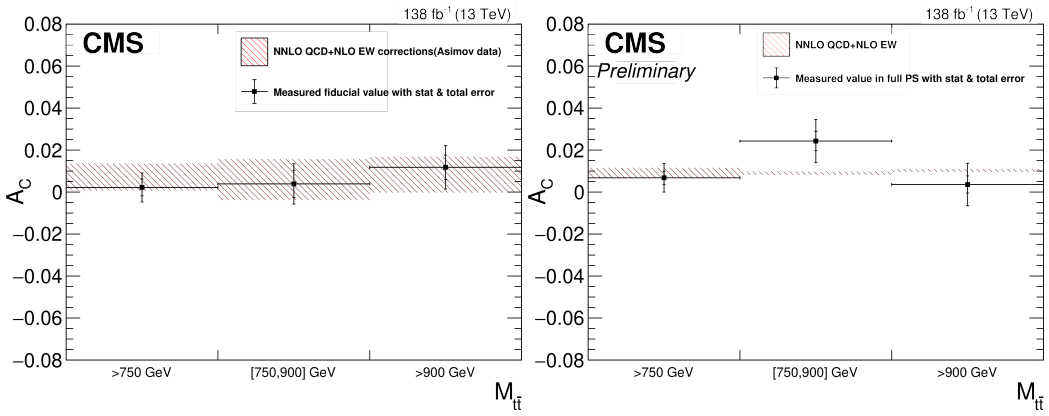

Figure 2: Measured A_C^{fid} (left) and A_C in the full phase space (right) presented in different mass regions after combining the μ + jets and e + jets channels. The vertical bars represent the total uncertainties, with the inner tick mark indicating the statistical uncertainty in the observed data.

The final A_C observable is measured in two ways: in the first method A_C^{fid}, the top quark charge asymmetry in the fiducial phase space is fitted for r_{pos} and r_{neg}, the signal strengths that scale the contribution of events with $\Delta|y| > 0$ and $\Delta|y| < 0$, respectively, and then use Eq.(1) to determine A_C^{fid}. Instead, we fit for r_{neg} and A_C^{fid} directly to ensure that the uncertainty in A_C^{fid} is correctly estimated. Figure2 summarises the results from both methods.

The measured inclusive asymmetry is $0.69^{+0.65}_{-0.69}$ %, which is consistent with the expected asymmetry of 1% at the LHC.

3. Charge-parity-time reversal symmetry (CPT)

According to the CPT theorem, the properties of the particles and their antiparticles, like mass and lifetime, must be the same. Therefore the mass difference and the ratio of the top quark and the antiquark masses are sensitive to the CPT symmetry. In single top production topologies, the masses of the top quark and antiquark are measured separately depending on the charge of the

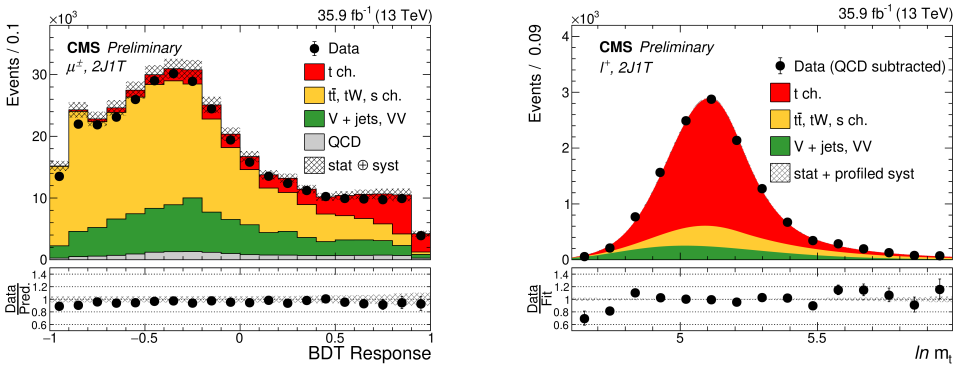

Figure 3: Prefit distribution of the BDT response (left [4]) and postfit distribution of $\ln m_t$ (right [4]).

lepton, where the W boson from the top quark decays leptonically. The mass difference and the ratio are obtained with the single top t-channel process using the data collected by the CMS detector during 2016 [4]. In this analysis, candidate events require at least one isolated lepton and a sizeable missing momentum due to an escaping neutrino. Exactly two jets are required, out of which one must be a b-tagged jet. The four-momentum of the top quark (and hence its mass) is reconstructed from its decay products: the charged lepton, the neutrino, and the b-tagged jet. Since QCD multijet production has a small acceptance in the phase space used in this analysis, QCD templates are derived from the sideband data. To suppress the QCD contribution, the transverse mass of the W boson (m_T) > 50 GeV is required for further analysis.

Several kinematic variables are combined into boosted decision trees (BDT) to separate single top quark events from backgrounds optimally. These variables are selected to have a significant power to distinguish signal and backgrounds and must have a low correction with the reconstructed top quark mass. A criterion on the BDT response > 0.8 has been applied (Fig. 3), which results in 64% (58%) signal purity and 20% efficiency for the muon (electron) final state. The distributions $y = (\ln m_t/1\ \text{GeV})$ obtained from the muon and electron final states are considered in a simultaneous ML fit. The QCD multijet contribution is subtracted from data before the fit, and a parametric model describes the remaining distribution as follows:

$$F(y; y_0, f_{t-ch}, f_{Top}, f_{EWK}) = f_{t-ch} \cdot F_{t-ch}(y; y_0) + f_{Top} \cdot F_{Top}(y; y_0) + f_{EWK} \cdot F_{EWK}(y)$$

Here y_0 is the parameter of interest which corresponds to the mean of the y distribution and the F_{t-ch}, F_{Top} and F_{EWK} are described by a sum of an asymmetric Gaussian core with a Landau tail, a Crystal ball function, a Novosibirsk function [5] to model the signal, top background, and electroweak background, respectively. The normalization scale factors f_{t-ch}, f_{Top}, and f_{EWK} are constrained using log-normal priors with 15, 6, and 10% based on their respective cross-section results [6] [7–9]. These constraints are included as nuisance parameters in the final fit, whereas other sources of systematic uncertainties are externalized. The top quark mass is obtained from the postfit $\ln(m_t/1\ \text{GeV})$ distribution, as shown in Fig. 3, by taking the exponential of the postfit value of the parameter of interest y_0. The mass of the top quark measured using t-channel single top quark events, inclusive of the lepton charge in the final state, is given by

$$m_t = 172.13 \pm 0.32\,(\text{stat} + \text{prof})\,^{+0.69}_{-0.70}\,(\text{syst})\ \text{GeV} = 172.13\,^{+0.76}_{-0.77}\ \text{GeV}, \tag{2}$$

reaching a sub-GeV precision for the first time in such a phase space.

The mass ratio and difference of the top antiquark to quark are determined to be $0.9952\,^{+0.0079}_{-0.0104}$ and $0.83\,^{+1.79}_{-1.35}$ GeV. The dominant source of systematic uncertainty is the jet energy scale, Parton shower scale, and b-quark hadronization model.

4. Charge parity violation (CPV)

CPV is manifested due to an irreducible phase in the Cabibbo–Kobayashi–Maskawa (CKM) matrix. The expected CPV in the SM can not fully account for the matter and antimatter asymmetry in the universe. But the BSM model can be the additional CPV source, manifested through a finite chromo-electric dipole moment (CEDM) in the top quark decay. BSM interaction modifies the tg coupling in the Lagrangian to accommodate the CEDM parameter d_{tG}, defined as:

$$d_{tG} = \frac{\sqrt{2}v}{\Lambda^2} Img(d_{tG}) + \iota\gamma_5 d_t^g) t G^a_{\mu\nu}, \tag{3}$$

where Λ is a high-mass scale of the BSM phenomena, and v is the vacuum expectation value for the Higgs boson field (≈ 246 GeV), $G^a_{\mu\nu}$ and d_t^g are the gluon field strength tensor and CP-odd CEDM. Higher d_{tG} values are expected to yield larger A_{CP} contributions. Such a BSM model introduces higher dimensional parameters. The CP observables are chosen to come from reconstructable final-state objects that can be well measured. The CP observables take the form $v_1.(v_2 \times v_3)$, where v_i are spin or momentum vectors and $i = 1 - 3$. The CPV asymmetry is defined as

$$A_{CP} = \frac{N(O_i > 0) - N(O_i < 0)}{N(O_i > 0) + N(O_i < 0)}, \tag{4}$$

where O_i is the A_{CP} sensitive operator of interest.

CPV measurement is performed in the $t\bar{t}$ process with lepton+jets [10], and fully leptonic [11] final states. For the lepton+jets final state, one lepton (e/μ) and at least four jets are required, out of which two jets must be b-tagged originating from the top decay. Four operators sensitive to CPV have been used, namely O_3, O_6, O_{12}, O_{14}. The contribution from the CEDM can be as large as 8 and 0.4% for $A_{CP}(O_3)$ and $A_{CP}(O_{12})$. The invariant mass of the lepton and b jet, $M_{lb} < 150$ GeV is required.

The $t\bar{t}$ production process is the dominating background in this analysis. A data-driven background method has been used to include contributions from other backgrounds. An ML fit has been performed to data using the M_{lb} distribution simultaneously in the e+jets and μ+jets channel to calculate the number of events for $O_i < 0$ and $O_i > 0$ to evaluate the A_{CP}. The final fit distribution after the fit is shown in Fig. 4.

Measured A'_{CP} is calibrated using the A_{CP} values at the generator level for each operator separately to estimate the actual value of the measured A'_{CP} as shown in Fig. 5. The final A_{CP} measurements are also summarised in Fig. 5.

We find $d_{tG} = 0.04 \pm 0.10(\text{stat}) \pm 0.07(\text{syst})$ determined from the measured A'_{CP} using the formula in equation Eq(6) where the parameters a, b, c, and d are taken from a χ^2 fit to the A_{CP}

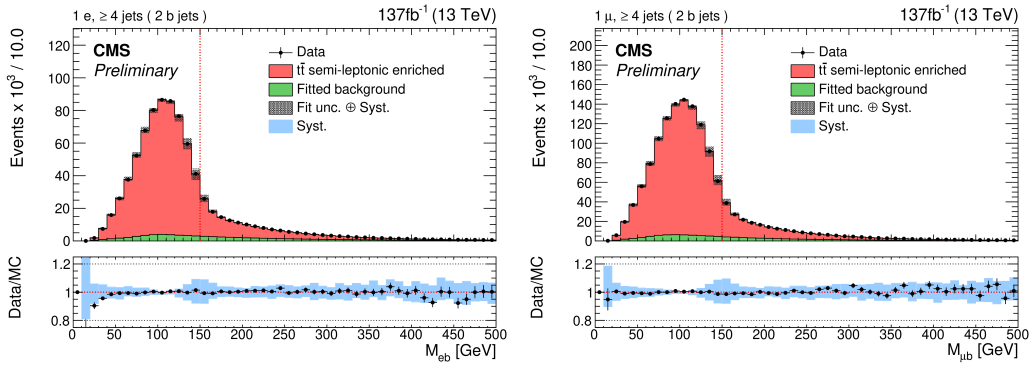

Figure 4: The distribution of the invariant mass M_{lb} distribution of the leptonically decaying top quark candidates in the electron (left) and muon (right) channels. The vertical bars on the data points in the lower panels indicate the statistical uncertainties in the data, and the hatched bands indicate the statistical uncertainties combined with the systematic uncertainties in the simulation. The blue bands represent the systematic uncertainties in the expected yield in the simulation, including all sources of systematic uncertainty, except for that due to changing the background template

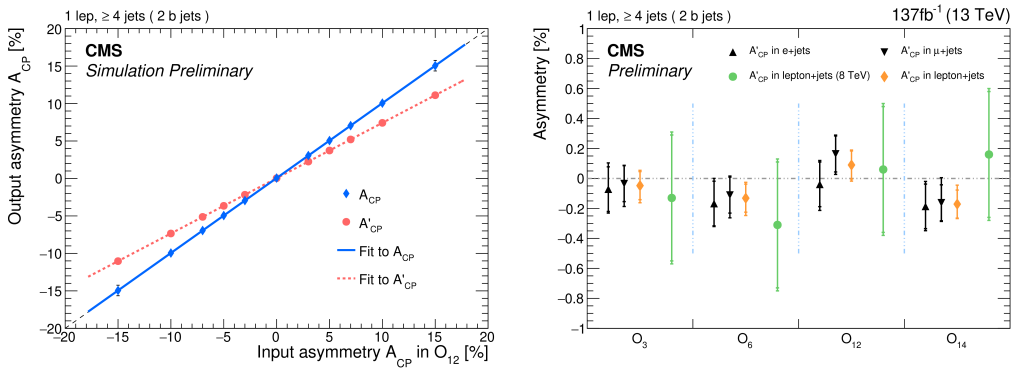

Figure 5: The left plot shows the effective asymmetries A'_{CP} for each observable shown for the separate electron and muon channels and the combined lepton+jets channel. The inner bars represent the statistical uncertainties, and the outer bars represent the combined statistical and systematic uncertainties added in quadrature. The right plot shows CPV asymmetries as a function of a varying input true asymmetry for each observable. The statistical uncertainties in the values of the asymmetries are smaller than the markers.

and d_{tG} values obtained from the simulated samples.

$$A_{CP} = \frac{d_{tg} + a}{b d_{tG}^2 + c d_{tG} + d} \tag{5}$$

Similarly, the analysis with the fully leptonic final state requires two leptons of opposite charge and at least two jets. One jet out of the selected jets has to be b-tagged. In this analysis, only the two operators, namely O_1 and O_3, are used, which are scalars under the Lorentz transformation. Three di-leptonic channels (e^+e^-, $\mu^+\mu^-$, $e^\pm\mu^\mp$) are considered with a requirement of lower bound on the di-lepton invariant mass $m_{\ell\ell} > 20$ GeV. The Z boson mass window ($76 < m_{ll} < 106$) GeV

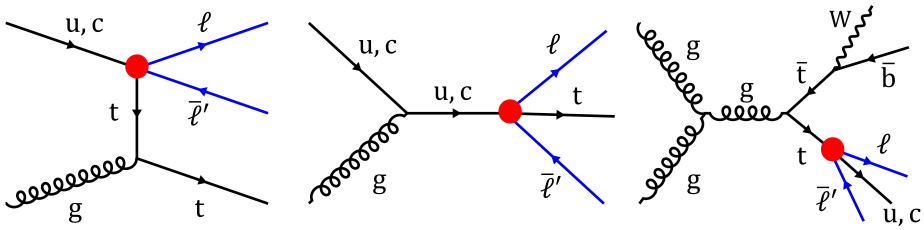

Figure 6: Feynman diagrams for single top quark production (left and middle) and top quark decays in SM $t\bar{t}$ events (right) via CLFV interactions. The CLFV vertex is marked as a filled circle.

is excluded for the same flavor lepton case. In this analysis, an ML fit has been performed using the O_1 and O_3 distributions to extract the $t\bar{t}$ cross-section. A_{CP} is measured using the number of events from the fit.

The A_{CP} measurement are combined using the best linear unbiased estimator method to take care of the statistical correlation ($\approx 46\%$) between O_1 and O_3.

The CEDM parameter Img(d_{tG}) is parameterized with the asymmetry using the Eq(6) where the coefficient a and b are extracted from the fitting the MC sample generated with different values of Img(d_{tg}) parameter values.

$$A_{CP} = a\text{Img}(d_{tg}) + b \tag{6}$$

The img(d_{tG}) derived from the measured asymmetries in O_1 and O_3 are 0.10 ± 0.12 (stat) ±0.12 (syst) and 0.00 ± 0.13 (stat) ±0.10 (syst), respectively. These results are consistent with the SM prediction.

5. Charged lepton flavour violation (CLFV)

CLFV is not allowed in the SM by construction, but it is clear from the discovery of neutrino oscillations that neutrinos have mass, and thus their flavors are not always conserved. BSM effects can be a possible source of CLFV in top quark decay, and the effective field theory (EFT) approach can probe such effects. Conventionally, an EFT Lagrangian consists of the six-dimensional Wilson coefficient as:

$$\mathcal{L} = \mathcal{L}_{SM} + \mathcal{L}_{eff} = \mathcal{L}_{SM} + \sum_x \frac{C_x}{\Lambda^2} O_x + \dots \tag{7}$$

This lagrangian introduces the CLFV interaction between up-type quarks, namely u,c with top and two opposite charged leptons, as shown in Fig. 6.

The search for the CLFV combines the search for "$e\mu tu$" and "$e\mu tc$" CLFV interactions in the top quark production which are generated separately [12]. The final state consists of two oppositely charged lepton of different flavors and one b-tagged jet. Leading lepton p_{T} and the invariant mass of the lepton-pair system has to be greater than 20 GeV to suppress the background. BDTs are used to extract the signal from the background. The training is done in the signal region and applied in the control region, which requires more than one b jets in the final state. A simultaneous ML fit has been performed using the BDT distribution in the signal and control regions to extract the signal strength as shown in Fig. 7.

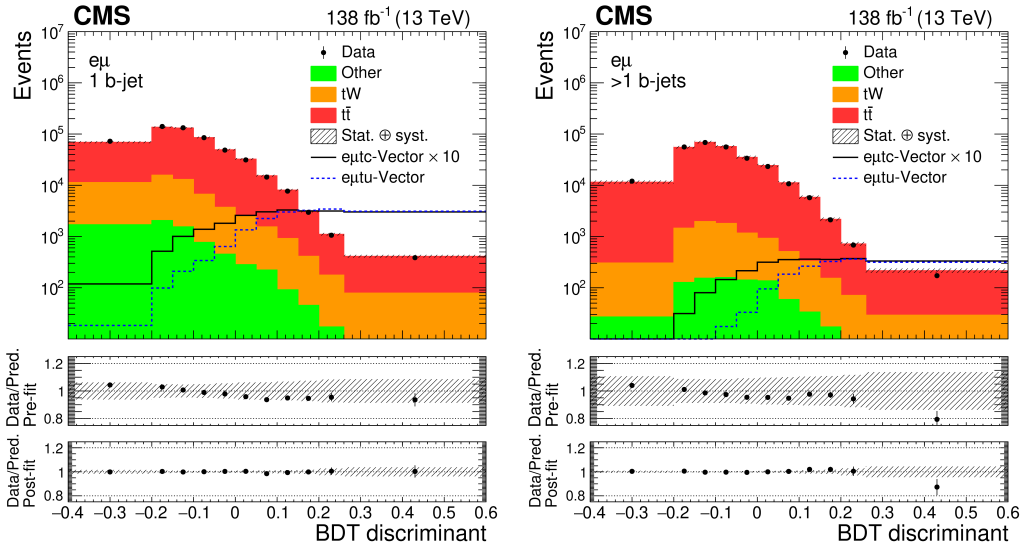

Figure 7: The BDT output distributions for data (points) and backgrounds (histograms) with the ratio of data to the total background yield, before (middle panel) and after (lower panel) the fit. Events with one or more b-tagged jets are shown in the left and right columns, respectively. The hatched bands indicate the total uncertainty (statistical and systematic added in quadrature) for the SM background predictions

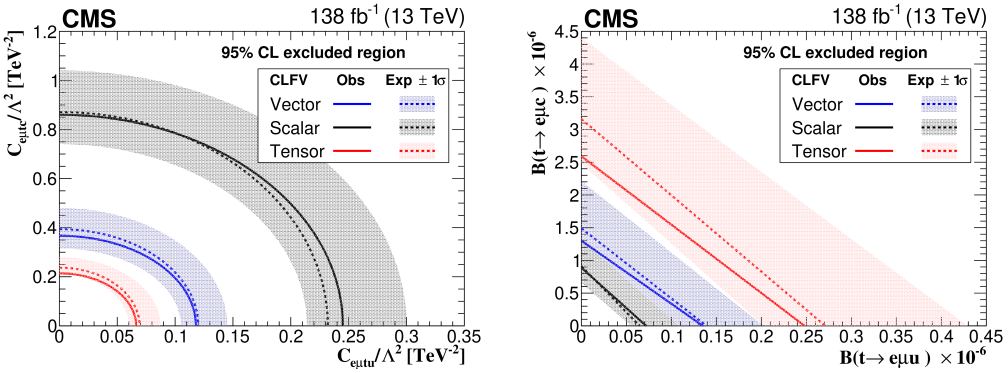

Figure 8: The observed 95% CL exclusion limits on the $e\mu tc$; of the $e\mu tu$ Wilson coefficient (left) and $B(t \to e\mu c)$ as a function of $B(t \to e\mu u)$ (right) for the vector-, scalar-, and tensor-like CLFV interactions. The hatched bands indicate the regions containing 68% of the distribution of limits expected under the background-only hypothesis.

Cross-section limits are calculated using the CLs method [13] with a 95% confidence level since no significant excess is observed over the SM expectation. The limits on the cross-section are first translated to limits on the Wilson coefficients since the cross-section is directly proportional to the square of the Wilson coefficient. Consequently, upper limits on the Wilson coefficients are translated to limits on the related top quark CLFV branching fractions as shown in Fig. 8.

Observed exclusion limits are set at 95% confidence levels on the respective branching fractions of a top quark to an eμ pair and an up (charm) the quark of $0.13 \times 10^{-6} (1.31 \times 10^{-6}), 0.07 \times$

$10^{-6}(0.89 \times 10^{-6})$, and $0.25 \times 10^{-6}(2.59 \times 10^{-6})$ for vector, scalar, and tensor CLFV interactions, respectively.

6. Summary

The latest results from CMS on tests of various asymmetry measurements like charge asymmetry (A_C), CP Violation, CPT symmetry, and charged lepton flavor violation in SM and the BSM for the top quark sector are summarized. There is no significant excess observed over SM expectations.

- Travel funding for this conference is sponsored by Infosys-TIFR Leading Edge Travel Grant.

References

[1] CMS Collaboration. The CMS experiment at the CERN LHC. *JINST*, 3 (2008) S08004. doi:10.1088/1748-0221/3/08/S08004.

[2] CMS Collaboration. Measurement of the $t\bar{t}$ charge asymmetry in events with highly Lorentz-boosted top quarks in pp collisions at $\sqrt{s} = 13$ TeV. *ArXiv e-prints*. arXiv:2208.02751v1.

[3] G. Soyez A. J. Larkoski, S. Marzani and J. Thaler. Soft drop. *ArXiv e-prints*, 05:146, 2014. arXiv:1402.2657, doi:10.1007/JHEP05(2014)146.

[4] CMS Collaboration. Measurement of the top quark mass using events with a single reconstructed top quark in pp collisions at $\sqrt{s} = 13$ TeV. *ArXiv e-prints*, 12:161, 2021. arXiv:2108.10407, doi:10.1007/JHEP12(2021)161.

[5] H. Ikeda et al. A detailed test of the CsI(Tl) calorimeter for BELLE with photon beams of energy between 20 MeV and 5.4 GeV. *Nucl. Instrum. Meth. A*, 441:401, (2000). doi:10.1016/S0168-9002(99)00992-4.

[6] M. Czakon, P. Fiedler, and A. Mitov. Total top-quark pair-production cross section at hadron colliders through $O(\alpha_S)$. *Phys. Rev. Lett.*, 110:252004, (2013). arXiv:1303.6254, doi:10.1103/PhysRevLett.110.252004.

[7] CMS Collaboration. Measurement of the differential cross sections for the associated production of a W boson and jets in proton-proton collisions at $\sqrt{s} = 13$ TeV. *Phys. Rev. D*, 96:072005, (2017). arXiv:1707.05979, doi:10.1103/PhysRevD.96.072005.

[8] CMS Collaboration. Measurement of differential cross sections for Z boson production in association with jets in proton-proton collisions at $\sqrt{s} = 13$ TeV. *Eur. Phys. J. C*, 78:965, (2018). arXiv:1804.05252, doi:10.1140/epjc/s10052-018-6373-0.

[9] CMS Collaboration. Measurement of the single top quark and antiquark production cross sections in the t channel and their ratio in proton-proton collisions at $\sqrt{s} = 13$ TeV. *Phys. Lett. B*, 800:135042, (2020). arXiv:1812.10514, doi:10.1016/j.physletb.2019.135042.

[10] CMS Collaboration. Search for CP violation using $t\bar{t}$ events in the lepton+jets channel in pp collisions at $\sqrt{s} = 13$ TeV. *ArXiv e-prints*. arXiv:2205.02314.

[11] CMS Collaboration. Search for CP violating top quark couplings in pp collisions at $\sqrt{s}$ = 13 TeV. *ArXiv e-prints*. arXiv:2205.07434.

[12] CMS Collaboration. Search for charged lepton flavor violation in top quark production and decay in proton-proton collisions at $\sqrt{s}$ = 13 TeV. *ArXiv e-prints*. arXiv:2201.07859.

[13] T. Junk. Confidence level computation for combining searches with small statistics. *ArXiv e-prints*, 434:435–443, 1999. arXiv:hep-ex/9902006, doi:10.1016/S0168-9002(99)00498-2.