UNIVERSITY OF CALIFORNIA
RIVERSIDE

Search for Standard Model Production of Four Top Quarks in Proton-Proton Collisions at 13 TeV in the Opposite-Sign Dilepton Channel Using CMS Data From 2017 and 2018

A Dissertation submitted in partial satisfaction
of the requirements for the degree of

Doctor of Philosophy

in

Physics

by

Nicholas James Manganelli

September 2022

Dissertation Committee:

Professor Robert Clare, Co-Chairperson
Professor Stephen Wimpenny, Co-Chairperson
Professor Owen Long

The Dissertation of Nicholas James Manganelli is approved:

Committee Co-Chairperson

Committee Co-Chairperson

University of California, Riverside

Acknowledgments

I am grateful to my advisors for their support and guidance, and my collaborators within the CMS Experiment for their assistance and wisdom.

To my father, my brother, and the friends who have graced my life.

ABSTRACT OF THE DISSERTATION

Search for Standard Model Production of Four Top Quarks in Proton-Proton Collisions at 13 TeV in the Opposite-Sign Dilepton Channel Using CMS Data From 2017 and 2018

by

Nicholas James Manganelli

Doctor of Philosophy, Graduate Program in Physics
University of California, Riverside, September 2022
Professor Robert Clare, Co-Chairperson
Professor Stephen Wimpenny, Co-Chairperson

A search for a rare Standard Model Process, four top quark production in the di-lepton plus jets final state (e e, μ μ, and e μ), is presented. The analysis utilizes data equivalent to 101.5 fb^{-1} of luminosity at $\sqrt{s} = 13$ TeV recorded by the CMS experiment during 2017 and 2018. Following a baseline selection to choose events with a pair of opposite sign charged leptons, four or more hadronic jets, of which at least 2 jets are b-tagged, the data are categorized in jet and b-tag multiplicity (4, 5, 6, 7, and 8 or more jets; 2, 3, and 4 or more b-tags). Events with at least 7 jets and 3 b-tagged jets are the most sensitive to the $\mathrm{t\bar{t}t\bar{t}}$ signal. The H_{T} distributions for the 15 categories per year and decay channel are fitted to simulation using Maximum Likelihood Estimation to obtain an upper limit on the cross section $\sigma_{\mathrm{t\bar{t}t\bar{t}}}$. An a-priori expected limit of $2.9^{+2.7}_{-1.4} \times \sigma^{SM}_{\mathrm{t\bar{t}t\bar{t}}}$ (35^{+32}_{-16}fb) is obtained at the 95% CL, with an a-priori expected signal significance of 0.64σ. The a-posteriori asymptotic limits are $3.0^{+2.9}_{-1.4} \times \sigma^{SM}_{\mathrm{t\bar{t}t\bar{t}}}$ (36^{+35}_{-17}fb). The asymptotic observed limits are $6.2 \times \sigma^{SM}_{\mathrm{t\bar{t}t\bar{t}}}$ (74fb) with a significance of 1.9σ.

Contents

List of Figures

List of Tables

Chapter 1

Four Top Production in the Standard Model

In the Standard Model, SM, the production of four top quarks, $t\bar{t}t\bar{t}$, results from gluon fusion or quark-antiquark annihilation. At the LHC, the gluon fusion diagram dominates and is shown in Fig 1.1. The cross section for this process is predicted to be extremely small with $\sigma^{SM}_{t\bar{t}t\bar{t}} \approx 12$ fb [5] at $\sqrt{s} = 13$ TeV. Next to leading order corrections (QCD + Electroweak) are predicted to be positive and on the order of 25%. Initial studies with 8 TeV and 13 TeV data yielded limits on $\sigma^{SM}_{t\bar{t}t\bar{t}}$ but no observation [6, 7, 8, 9, 10].

Recent studies by the ATLAS Collaboration [11, 12] based on 139 fb^{-1} of data at 13 TeV suggest a possible cross section of around 24 fb which is approximately a factor of two larger than the predicted SM cross section of $12.6^{+5.8}_{-5.2}$ fb [5]. If proven to be correct, this could provide an indication new physics, such as enhanced $t\bar{t}t\bar{t}$ production via gluon pair production from supersymmetric theories [13, 14], pair production of scalar gluons [15, 16],

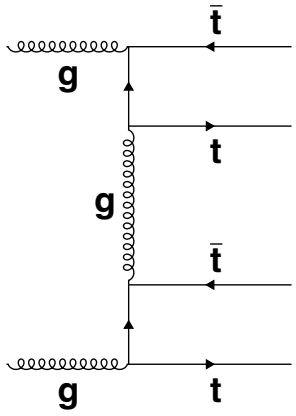

Figure 1.1: The dominant SM diagram for $t\bar{t}t\bar{t}$ production at the LHC is shown.

or by the production of a heavy scalar or pseudoscalar boson in association with a $t\bar{t}$ pair in type II two-Higgs-doublet models [17, 18, 19].

The analysis presented in this dissertation describes the opposite-sign dilepton analysis performed using the dataset recorded by CMS during 2017 and 2018 with luminosity = 101.5 fb^{-1}. It is one of three new analyses, which in combination with the result from [10] will probe the ATLAS results. This analysis employs a cut-based strategy in conjunction with improved Monte Carlo simulation and more efficient b-tagging algorithms to achieve similar performance per unit integrated luminosity as the Multi-Variate technique used in 2016 for the same channel (for this reason, 2016 data are not re-analyzed with a cut-based approach).

Chapter 2

The Standard Model

The Standard Model has proven extraordinarily successful at categorizing the most fundamental constituents of matter and their interactions. The theory accurately predicts the rates of decay and production of both fundamental and composite particles, at multiple scales. Fundamental particles are classified as the spin-0 (Scalar) and spin 1 (Vector) Bosons and spin $\frac{1}{2}$ Fermions, which are described in this section. The underlying Quantum Field Theory constructed to fit experimental measurements is briefly over-viewed in Section 2.2. The theory is not yet complete, as evidenced by the failure to explain several features of the Universe, namely the matter-antimatter asymmetry, neutrino masses, dark matter, dark energy, and unification with gravity, further discussed in Section 2.4.

2.1 Particles

Bosons

The Vector Bosons are force exchange mediators; the photon (γ), gluon (g), and $W^{\pm}$ and Z^0 Bosons are outlined in Table 2.1, along with their associated forces. The Higgs Boson (H) is the particle associated with the Higgs Field, which was discovered in 2012 by the CMS and ATLAS experiments at the Large Hadron Collider. It is the only fundamental scalar particle in the theory. Gluons (g) are the mediators for the Strong Nuclear Force, the residual effect of which is seen as the nuclear binding force keeping the protons and neutrons of the nucleus together. The $W^{\pm}$ and Z^0 were discovered in 1983 by the UA1 and UA2 experiments at the CERN Super Proton Synchrotron. These particles' quantum numbers are displayed in the bottom part of Fig. 2.1.

Table 2.1: The force mediators, their mass, relative strength and range.

Force	Mediator	Mass [GeV]	Relative Strength	Range
Strong Force	g	0	10^2	10^{-15}m
Electromagnetism	γ	0	1	∞
Weak Force (Neutral)	Z^0	91.2	10^{-11}	10^{-18}m
Weak Force (Charged)	$W^{\pm}$	80.4	10^{-11}	10^{-18}m

The Bosons collectively obey Bose-Einstein statistics. The Klein-Gordon Equation (2.1a) is appropriate to scalar fields (with $\eta^{\mu\nu}$ being the Minkowski Metric with positive-timelike element), while vector fields satisfy the Proca Equation (2.1b). The photon and Z^0 are their own anti-particles, but the charged $W^{\pm}$ particles are a particle-antiparticle pair.

$$(i\eta^{\mu\nu}\partial_\mu\partial_\nu + m^2)\Phi(x) = 0 \tag{2.1a}$$

$$\partial_\mu(\partial^\mu X^\nu - \partial^\nu X^\mu) + m^2 X^\nu = 0 \tag{2.1b}$$

$$X = \text{4-component field, i.e. Electromagnetic field } A \tag{2.1c}$$

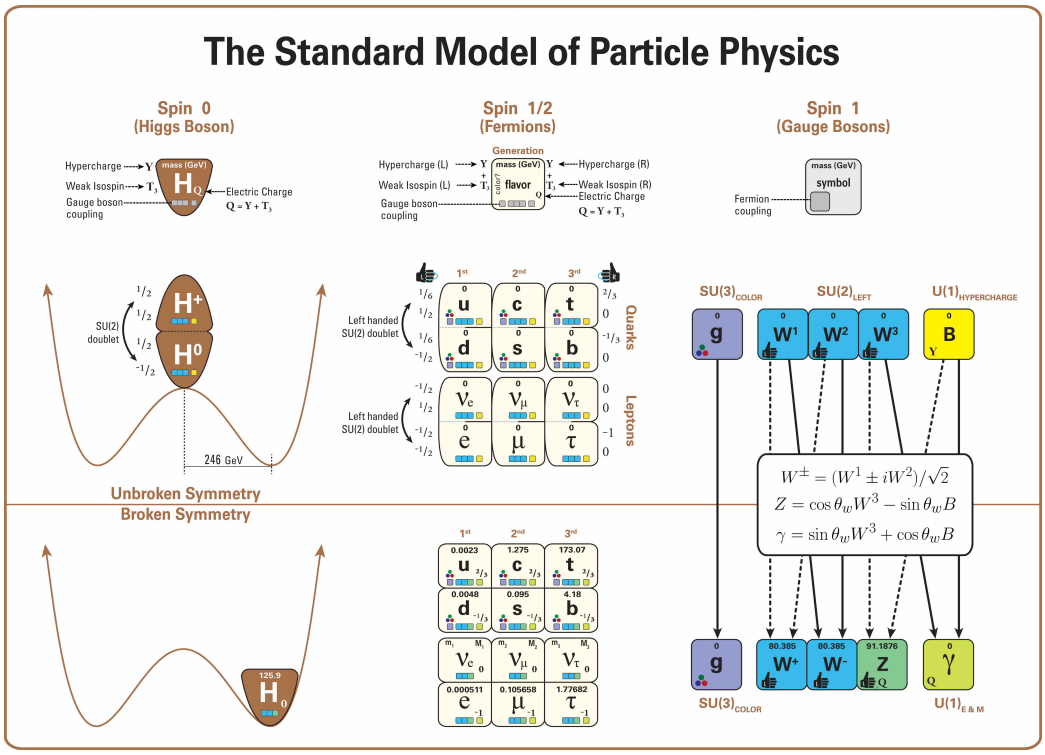

Figure 2.1: The particles (quanta) of the Standard Model fields both before and after Electro-Weak Symmetry Breaking.

Fermions

Fermions are the constituents of the Baryonic matter we observe in the Universe, carrying an intrinsic spin of $\frac{1}{2}$, thereby obeying Fermi-Dirac statistics and satisfying the Dirac Equation 2.2a.

$$(i\gamma^\mu \partial_\mu - m)\Psi(x) = 0 \tag{2.2a}$$

$$\Psi(x) = \text{field for a Dirac Spinor, 4-component and complex-valued} \tag{2.2b}$$

They are divided into two types: leptons and quarks. Each of these types has 3 generations, which behave very similarly except for differences in their mass. For the quarks, the lightest set of the 3 generations is the u (up) and d (down), which form a left-handed (or Chiral) doublet and 2 right-handed singlets. The weak force mediators demonstrate an asymmetry in their interaction with left and right Chiral particles, with the weak charged currents ($W^{\pm}$ mediated) demonstrating no coupling to right-handed particles. In the group theory view, these Bosons can "rotate" the upper and lower elements in the left-handed doublet (in the plane of hypercharge and 3rd component of weak isospin), such that $u_L \longrightarrow d_L$ and vice-a-versa, but cannot transform the right-handed $u_R \longrightarrow d_R$. The up-type quarks have a fractional $+\frac{2}{3}$ electric charge, the down-type $-\frac{1}{3}$. Amongst the leptons, the electron, muon, and tau (e, μ, τ) all have electric charge -1, while their isospin-doublet partners are neutrinos with no electric charge. The masses of the neutrinos are not (all) zero, but have yet to be measured. Each of the fermions has an anti-matter partner.

Higgs Boson

The discovery of the Higgs Boson, as announced on July 4, 2012, completed one of the greatest chapters yet in the history of Particle Physics. The Higgs is the only known fundamental scalar particle in the Standard Model, an excitation of one of the Brout-Englert-Higgs (BEH) Fields which gives rise to mass (in a gauge-invariant manner) for most of the SM particles. Quarks, leptons, and the vector bosons couple directly to the Higgs boson, but photons and gluons do so through indirectly via loops [20]. Its observation was essential in verifying a coherent Standard Model, even if some mysteries remain.

2.2 Quantum Field Theory

Noether's Theorem

Emmy Noether was invited by David Hilbert and Felix Klein to the University of Göttingen due to her expertise in invariants, to work on the apparent problem of non-conservation of energy in Albert Einstein's Theory of General Relativity. In 1915, she proved the theorem that now bears her name. She went further than proving the energy was conserved, as she showed that any continuous symmetry of nature is invariably linked to a conserved quantity[21]. Noether's Theorem (Eq. 2.3a), and several extensions of it to discrete symmetries (c.f. Slavnov-Taylor, Ward-Takahashi identities), form a foundational pillar upon which modern Field Theories are based. The generalized currents (Eq. 2.3b) that are derived from its application are ubiquitous in Quantum Field Theory, and the importance of the theorem is underscored by Albert Einstein's remark that "The old guard at Göttingen should take some lessons from Miss Noether! She seems to know her stuff." after receiving her paper on invariants.

$$\Sigma \partial_\mu j_i^\mu(x) = 0 \tag{2.3a}$$

$$j_i^\mu(x) = \frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi_i)} \Delta\phi_i - \mathcal{J}_i^\mu \tag{2.3b}$$

Symmetries

The Standard Model encodes several observed symmetries of nature. Translational invariance, in combination with Noether's Theorem, leads to the conservation of momentum; rotational invariance to conservation of angular momentum. In QFT, this is the sum of the

orbital (L) and spin (S) angular momenta; the quantities from these two operators are not generally conserved, but it can be shown that the sum of squares of their quanta is [22]. There are a number of other continuous and discrete symmetries which are also fully or approximately obeyed by the Universe. Under Charge Conjugation (C), a particle of the opposite charge should behave the same; and under Parity (P), where the coordinates and linear momenta are mirrored ($\vec{r} \longrightarrow -\vec{r}$, but angular momentum $J = \vec{r} \times \vec{p}$ is not), similarly so. Under perfect Time Reversal Symmetry (T), all interactions would behave the same with the time coordinate reversed (i.e. a $2 \longrightarrow 1$ interaction should have a corresponding $1 \longrightarrow 2$ interaction at the same rate). Z^0 exchange respects CP symmetry but violates C and P separately, and W boson interactions maximally violate C and P. [21].

The Standard Model Lagrangian

Before Electroweak Symmetry Breaking

Prior to Electroweak Symmetry Breaking (EWSB), the SM Lagrangian density encapsulates our understanding of the fields in the observable universe, the excitations of which are the standard particles we can interact with and measure. This is encapsulated by Eq. 2.4, where the fermion fields Ψ^f and the gauge fields X^j (and associated field strength tensors) are summed over all the types (Ψ, X) and components (j) of the SU(3) $\otimes$ SU(2) $\otimes$ U(1) symmetry groups.

$$\mathcal{L}_{SM} = \mathcal{L}_V + \mathcal{L}_f + \mathcal{L}_H + \mathcal{L}_{\text{Yuk}} \tag{2.4}$$

Equation 2.5b describes the Field Strength Tensor for a given field X (Eq. 2.1c), with some constant c and structure constant f^{abc} (Eq. 2.5c) which is governed by the commutator

relations of the Group's generators. The first term, expanded in Eq. 2.5a encompasses the 3 sets of Gauge fields. There are 8 generators T^a and fields G^a corresponding to the SU(3) gauge group (which is non-Abelian, hence a non-vanishing structure constant, and with associated charge "color"); 3 generators and fields W^i for the SU(2) group, similarly non-Abelian (associated charge: "isospin"); and 1 generator Y and field B for the Abelian U(1) group (associated charge: "hypercharge").

$$\mathcal{L}_V = -\frac{1}{4}B_{\mu\nu}B^{\mu\nu} - \frac{1}{4}W^i_{\mu\nu}W^{\mu\nu,i} - \frac{1}{4}G^a_{\mu\nu}G^{\mu\nu,a} \tag{2.5a}$$

$$X^{\mu\nu}_a = \partial_\mu X^\nu_a - \partial_\nu X^\mu_a - c f_{abc} X^b_\mu X^c_\nu \tag{2.5b}$$

$$[T^a, T^b] = i f^{abc} T^c \tag{2.5c}$$

The second component of the Lagrangian, Eq 2.6a, describes the spin-$\frac{1}{2}$ fermionic fields and their interaction with the gauge fields. The covariant derivative included is expanded in Eq. 2.6b, where the left-handed quark fields may couple to all three fields, the right handed to G and B, the Higgs to W and B, and so on according to their (L or R) isospin, hypercharge, and color.

$$\mathcal{L}_f = \sum_{\Psi \in Q_L, u_R, d_R, e_R, L} \bar{\Psi} i\gamma^\mu D^{\mathrm{YIC}}_\mu \Psi \tag{2.6a}$$

$$D^{\mathrm{YIC}}_\mu = (\partial_\mu - ig' B_\mu Y - igW^i_\mu T^i - ig_S G^a_\mu T^a) \tag{2.6b}$$

The third component of the SM Lagrangian (Eq. 2.7a) groups together subterms involving ϕ and the gauge-field interactions. Equation 2.7a) separates out the ϕ-fermion interactions. The most notable feature of Eq. 2.4 is the absence of mass terms for the fermions and vector Bosons. Explicit fermion and gauge-field mass terms would break the

gauge invariance of the theory. It is through the interactions with the Higgs fields that these two classes of particles will gain "mass."

$$\mathcal{L}_H = |D_\mu^{\text{YIC}}\phi|^2 + \mu^2\phi^\dagger\phi - \lambda(\phi^\dagger\phi)^2 \tag{2.7a}$$

$$\mathcal{L}_{\text{Yuk}} = y_{ij}\bar{\Psi}_i\phi\Psi_j + \text{hermitian conjugate} \tag{2.7b}$$

Electro-Weak Symmetry Breaking and the Emergence of Mass

"Our hope of finding a gauge theory of weak interactions with massive gauge bosons looks forlorn. It appears that we shall also have unwanted (unobserved) massless scalar particles to worry about. Nevertheless, let us proceed from a global to a local gauge theory. A miracle is about to happen."[23, p.325].

A triumph of particle physics is the development of (and numerous experimental evidence for) the Electro-Weak Symmetry Breaking mechanism [21]. The field in Eq. 2.7a describes the kinematics for an SU(2) doublet field (two complex values). It's expected for most of the quantum fields to typically be in a 0-like ground state (space is mostly empty, vacuum excitations excepted), but in the event that $\lambda > 0$ and $\mu^2 > 0$, then in the unitary gauge one can find $\langle\phi\rangle = \binom{0}{v}$ with $v = \sqrt{\frac{\mu^2}{\lambda}}$ a real quantity. Then ϕ may be expanded about this ground state expectation, and this parameter v, called the vacuum expectation value (vev) leads to mass-terms for the Bosons ($M_{\text{W}} = \frac{1}{2}vg$ and $M_{\text{Z}} = \frac{1}{2}v\sqrt{g^2 + g'^2}$), where these bosons are linear combinations of the W and B fields coupling to the Higgs (see Sec. 2.3 for more on the mechanism). The Yukawa terms can be reduced to $m_f\bar{\Psi}_f\Psi_f$ with m_f proportional to $y_f v$ in the flavor eigenbasis, becoming the the mass parameter of the Dirac equation of motion. As such, it is said that mass emerges as a post-EWSB

feature from interactions with the scalar Higgs field. While the non-zero vev is responsible for both, the vector Bosons additionally gain an extra degree of freedom in polarization, which is elucidated in Section 2.3 on the Goldstone Boson Equivalence Theorem. Of note, no right-handed neutrinos being observed in nature leads to the question of how they attain mass.

After EWSB

With the Higgs field having the given vacuum expectation value, the Lagrangian can be rearranged to better match the observed currents and eigenstates. The observed Vector Bosons are Electromagnetic photons (field given by Eq. 2.8a), Weak Charged Bosons (Eq.2.8c), and Weak Neutral Boson (Eq. 2.8b), which are all eigenstates of charge $Q = T^3 + Y$. The covariant derivative may be re-written as in Eq. 2.9, where the charged and neutral weak currents are clear, and the electromagnetic field is made manifest.

$$A = \sin\theta_{\mathrm{W}}\mathrm{W}_3 + \cos\theta_{\mathrm{W}}B \tag{2.8a}$$

$$Z = \cos\theta_{\mathrm{W}}\mathrm{W}_3 - \sin\theta_{\mathrm{W}}B \tag{2.8b}$$

$$\mathrm{W}^{\pm} = \frac{1}{\sqrt{2}}(\mathrm{W}_1 \mp i\mathrm{W}_2) \tag{2.8c}$$

$$\cos\theta_{\mathrm{W}} = \frac{M_W}{M_Z} = \frac{g}{\sqrt{g^2 + g'^2}} \tag{2.8d}$$

$$
\begin{aligned}
D_\mu =& \partial_\mu - i\frac{g}{\sqrt{2}}(\mathrm{W}^+{}_\mu T^+ + \mathrm{W}^-{}_\mu T^-) \\
& - i\frac{1}{\sqrt{g^2+g'^2}}\mathrm{Z}^0{}_\mu (g^2 T^3 - g'^2 Y) \\
& - i\frac{gg'}{\sqrt{g^2+g'^2}} A_\mu (T^3 + Y) \\
& - ig_S G^a_\mu T^a
\end{aligned}
\tag{2.9}
$$

2.3 Goldstone Boson Equivalence Theorem

The weak gauge bosons ($\mathrm{W}^\pm$, Z^0) are massive, and accordingly have the short range that gives them their name as the weak force mediators. The presence of 'naturally massive' gauge fields, however, presents a problem in a Yang-Mills theory; when no longer gauge invariant, renormalizability becomes an impossibility [24] [25]. Through the efforts of Goldstone [26]; Glashow [27]; Salam, and Weinberg [28]; Brout, Englert [29]; Higgs [30]; and Guralnik, Hagen and Kibble [31], it was shown that a potential solution is spontaneously breaking the symmetry (SSB) of a group $SU(2)_L \times U(1)_Y$ to $U(1)_{EM}$, which then *dynamically* generates mass for both gauge and spinor fields.

Under the SSB induced by the Higgs mechanism (involving a scalar field with two complex components and a non-zero vacuum expectation value for one of the four components), 3 massless Goldstone bosons and the massive Higgs boson become part of the scheme. The Goldstone bosons are absorbed by 3 of the 4 (linear combination) vector boson fields from the original $SU(2)_L \times U(1)_Y$ gauge groups, giving mass to 3 of their quanta ($\mathrm{W}^\pm$, Z^0) and leaving the 4th massless (γ). While massless, these bosons have 2 possible physical transverse polarization (TP) states, and upon becoming massive, gain a third longitudinal

polarization (LP). The polarization vectors collectively satisfy $k_\mu \epsilon^\mu = 0$ and $\epsilon^2 = -1$, and the states are quite disparate under Lorentz boosts:

$$\epsilon_L = (0,0,0,1) \rightarrow (k/m_W, 0, 0, E_k/m_W) \tag{2.10}$$

$$\epsilon_T = \frac{1}{\sqrt{2}}(0,1,\pm i,0) \rightarrow \frac{1}{\sqrt{2}}(0,1,\pm i,0) \tag{2.11}$$

This fundamentally distinguishes the transverse and longitudinal modes. Summing over polarizations, they manifest (in the Unitary gauge) as the first and second terms of Eq. 2.12, respetively.

$$\sum_{spin} \epsilon_\mu(k)\epsilon^*_\nu(k) = -g_{\mu\nu} + \frac{k^\mu k^\nu}{m_{\mathrm{W}}^2} \tag{2.12}$$

The Goldstone Boson Equivalence Theorem, first proven by Cornwall, Levin, and Tiktopoulos [25], relies upon the Ward identities, $0 = \langle \partial_\mu J^\mu \rangle = k_\mu \langle J^\mu \rangle$. The essential idea is that a gauge current can couple directly with some one-Particle-Irreducible (1PI) vertex, or it may first couple to a scalar boson in the Higgs sector which mediates the coupling into the same 1PI vertex. As such,

$$\langle J^\mu \rangle = \Gamma^\mu(k) + (igFk^\mu)\frac{i}{k^2}\Gamma(k) \tag{2.13}$$

where the former is for the direct coupling, and the latter parameterizes the scalar mediated process ($gF = m_{\mathrm{W}}$).

$$\begin{aligned} 0 = k_\mu \langle J^\mu \rangle &= k_\mu [\Gamma^\mu(k) - k^\mu \frac{m_{\mathrm{W}}}{k^2}\Gamma(k)] \\ &= k_\mu \Gamma^\mu(k) - m_{\mathrm{W}} \frac{k^2}{k^2}\Gamma(k) \end{aligned} \tag{2.14}$$

$$k_\mu \Gamma^\mu(k) = m_{\mathrm{W}}\Gamma(k) \tag{2.15}$$

Returning to $\epsilon_L = (k/m_{\mathrm{W}}, 0, 0, E_k/m_{\mathrm{W}})$, in the high energy limit, this is $\frac{k^\mu}{m_{\mathrm{W}}} + O(m_{\mathrm{W}}/E_k)$, as the numerator is essentially the four momentum with components swapped.

Then, replacing k^μ, dividing out m_{W}, and dropping the error term of $O(m_{\mathrm{W}}^2/E_k)$ gives the final result.

$$\epsilon_{L\mu}\Gamma^\mu(k) = \Gamma(k) \tag{2.16}$$

The physical interpretation of this statement is that the emission or absorption of a longitudinally polarized gauge boson is equal to that for a Goldstone boson in the high energy limit.

By itself, the theorem provides a powerful tool for alternative calculations involving transversely polarized gauge bosons. It's clear the equivalent scalar boson vertex can simplify calculations, despite the requirement of technically adding more diagrams to the given process.

An illuminating application of the GBET is the partial decay width of the top quark (momentum p) to a W boson (momentum k) and b quark (momentum q), for which the amplitude and subsequent partial width (making use of Eq. 2.12 when summing outgoing and averaging incoming spins) are:

$$i\mathcal{M} = \frac{ig}{\sqrt{2}}\bar{u}(q)\gamma^\mu\left(\frac{1-\gamma^5}{2}\right)u(p)\epsilon_\mu^*(k) \tag{2.17}$$

$$\Gamma = \frac{g^2}{64\pi}\frac{m_{\mathrm{t}}^3}{m_{\mathrm{W}}^2}\left(1-\frac{m_{\mathrm{W}}^2}{m_{\mathrm{t}}^2}\right)^2\left(1+2\frac{m_{\mathrm{W}}^2}{m_{\mathrm{t}}^2}\right) \tag{2.18}$$

Alternatively, to compute the decay to just a longitudinally polarized W, one can use the GBET and the Yukawa coupling of the top to the Higgs (λ_{t}), finding:

$$i\mathcal{M} = i\lambda_{\mathrm{t}}\bar{u}(q)\gamma^\mu\left(\frac{1+\gamma^5}{2}\right)u(p) \tag{2.19}$$

$$\Gamma = \frac{g^2}{64\pi}\frac{m_{\mathrm{t}}^3}{m_{\mathrm{W}}^2}\ (m_{\mathrm{t}} \gg m_{\mathrm{W}}\ \textit{approx.}) \tag{2.20}$$

Equation 2.20 reproduces the leading order term of Eq. 2.18 (See [21] for more details). What has been glossed over is that the leading order term does not vary as $\alpha_{\mathrm{W}} = \frac{g^2}{4\pi}$, but has an additional enhancement in the form of $\frac{m_{\mathrm{t}}^2}{m_{\mathrm{W}}^2}$, which can be recast as $2\frac{\lambda_{\mathrm{t}}^2}{g^2}$. In other words,

$$\Gamma = \frac{\lambda_{\mathrm{t}}^2}{32\pi} m_{\mathrm{t}} \tag{2.21}$$

and the decay width in the leading term scales with the top-Yukawa coupling [32] to the Higgs sector, rather than the weak coupling!

The Equivalence Theorem connects the Goldstone bosons to the weak bosons' tertiary modes after symmetry breaking, serving as a signature for the Higgs mechanism in action It is a useful tool for simplifying calculations in the high energy regime where longitudinal polarizations of the vector bosons appear.

2.4 Incompleteness of the Standard Model

Baryon-Antibaryon Asymmetry

While our observable Universe's energy density appears to be composed of just 4% visible baryonic matter, virtually none of it is anti-matter. A symmetric initial state for the creation of baryonic matter is presumed in the very early moments of the Universe. The Sakharov conditions [33] lay out the conditions in which such an evolution may occur, requiring that not only is there Charge-Parity (CP) violation prior to reaching thermal equilibrium, but also Charge (C) symmetry violation and non-conservation of Baryon (B) Number. A non-anthropomorphic argument for this imbalance in the visible universe does not yet exist. Charge-Parity (CP) violation has been observed in experi-

ments [34], including BaBar [35] Belle [36] and more recently LHCb [37], which goes part-way to explaining this conundrum. However, the magnitude of the violation, encoded in the Cabibbo–Kobayashi–Maskawa (CKM) matrix, is far too small to explain the observed imbalance.

Neutrino Masses

There is no mechanism in the core theory explains the non-zero masses of neutrinos. In a curious relation to the matter-antimatter asymmetry, the weak current $J^{\mu} = \bar{\Psi^{e}}\gamma^{\mu}\frac{1}{2}(1-\gamma^{5})\Psi_{\nu}$ relating the charged electron and neutral neutrino contains the Vector$-$Axial current $(V-A)$ that projects out left-handed neutrinos (right-handed anti-neutrinos). In other words, the charged leptons would only interact with these neutrino types if they were massless. With non-zero neutrino mases, which have been experimentally verified through neutrino-oscillations [38, 39], the handedness may swap under a Lorentz transformation, and so we should experimentally find right-handed neutrinos (left-handed anti-neutrinos). However, so far no evidence for this exists. Two mechanisms have been postulated for how neutrinos gain mass in a way consistent with the known SM. If neutrinos are their own anti-particle (Majorana neutrinos), the Majorana Mechanism can explain their non-zero mass. If this is not the case, and they are Dirac neutrinos, then a sterile neutrino which only interacts with the neutral component of the Higgs field could be a missing component.

Gravity

A particle carrying the gravitational force has not been discovered (and may not exist; the incompatibility of General Relativity and Quantum Field Theory continues to

frustrate Theoretical Physicists). The symmetric rank-2 tensor encoding the gravitational field ($g_{\mu\nu}$) indicates that a quantization of it would be a spin-2 particle [40].

Hierarchy Problem

The Mass Hierarchy/Naturalness problem has been present since the discovery of the Higgs Boson in 2012. The large discrepancy between the effective scales at which different forces act, such as gravity and the weak force (10^{35} separation in strength), is not technically in violation of any SM principle. However, it means there is a massive, unexplained gap. Similarly, the Higgs mass is extremely light compared to the predictions that would produce a "Natural" model, which is one in which model parameters do not have to be "fine-tuned" to arbitrary scales. Various extensions of the Standard Model propose mechanisms that can protect or naturally lead to such a small mass. One of them is Supersymmetry (SUSY), which balances out quantum loop corrections in the renormalization that would scale with the square of a Λ_{UV}, with oppositely signed counter-terms. Specifically, fermionic particles such as the top quark would have bosonic partners (top squark). No evidence of SUSY has yet been found, and large swathes of parameter space have ruled out. Most remaining and still-plausible SUSY models would require similar fine-tuning as the SM. A second set of models involve multiple Higgs doublets from a larger set of fields than described by a minimal BEH mechanism. In the minimal model, the Higgs field obeys an SU(2) symmetry with 2 neutral components, and 2 electrically charged components, which leads to one massive scalar particle, "the" Higgs Boson. Instead of a single Higgs, there could be two or more Higgs Doublets. Various permutations of allowed couplings and mass hierarchies lead to Higgs particles with charge, CP-even and CP-odd counterparts, etc.

Dark Matter

A candidate for the Dark Matter which is known to compose 1/5th of the observable Universe has not been found. From observations of the Bullet Cluster [41], galactic rotation curves [42], gravitational lensing (c.f. [43]) and other cosmological observations, there is a known discrepancy between observable matter distributions and measurable distributions in the typical Einstein Gravity. Accordingly, this type of matter which is unobserved directly must account for 21% of the Universe's energy density, compared to just 4% for typical Baryonic matter.

Dark Energy

Sixth, the SM gives no explanation for the effect currently attributed to Dark Energy, which forms much of the energy density of the universe and drives its expansion. The expansion of the universe is well established, most notably by the discovery of the Cosmic Microwave Background (CMB; by Penzias and Wilson in 1965), in which photons drop below the threshold to ionize the matter of the Universe in an epoch called Recombination, and subsequently cool through further expansion from approximately 1000K to the 3K currently observed [23, p. 352]. In subsequent years, evidence has emerged that the Universe is not only expanding, but it is doing so at an increasingly greater rate [44, 45]. This acceleration is contrary to the long-range Gravitational force, necessitating that $\frac{3}{4}$ of the energy density of the Universe be attributed to what we now call Dark Energy, and that its properties drive the acceleration.

Chapter 3

The Experimental Apparatus

3.1 The Large Hadron Collider

The Large Hadron Collider (LHC) is 50m – 150m underground on the border of France and Geneva, Switzerland, with a circumference of 27km. It is the most energetic human-made particle collider yet operated, accelerating bunches of protons to 6.5 TeV during collisions fromf 2015 to 2018 (Run II), giving a Center of Mass (CoM) energy of 13 TeV. Radio-Frequency (RF) cavities – complex wave-guides which "kick" proton packets to add half an MeV per turn – and magnetic dipoles – to bend the proton beams – are the main manipulators employed. The machine exceeds twice the design luminosity(Eq. 3.1) of $10^{34}\text{cm}^2\text{s}^{-1}$; approximately 1 billion proton-proton interactions occur per-second, even while intentionally limiting the rate of collisions in a process called "luminosity leveling" during Run II and beyond, c.f. [46]. A massive liquid-Helium cryogenic system is required to keep the magnets at superconducting temperatures. The collider was designed to discover

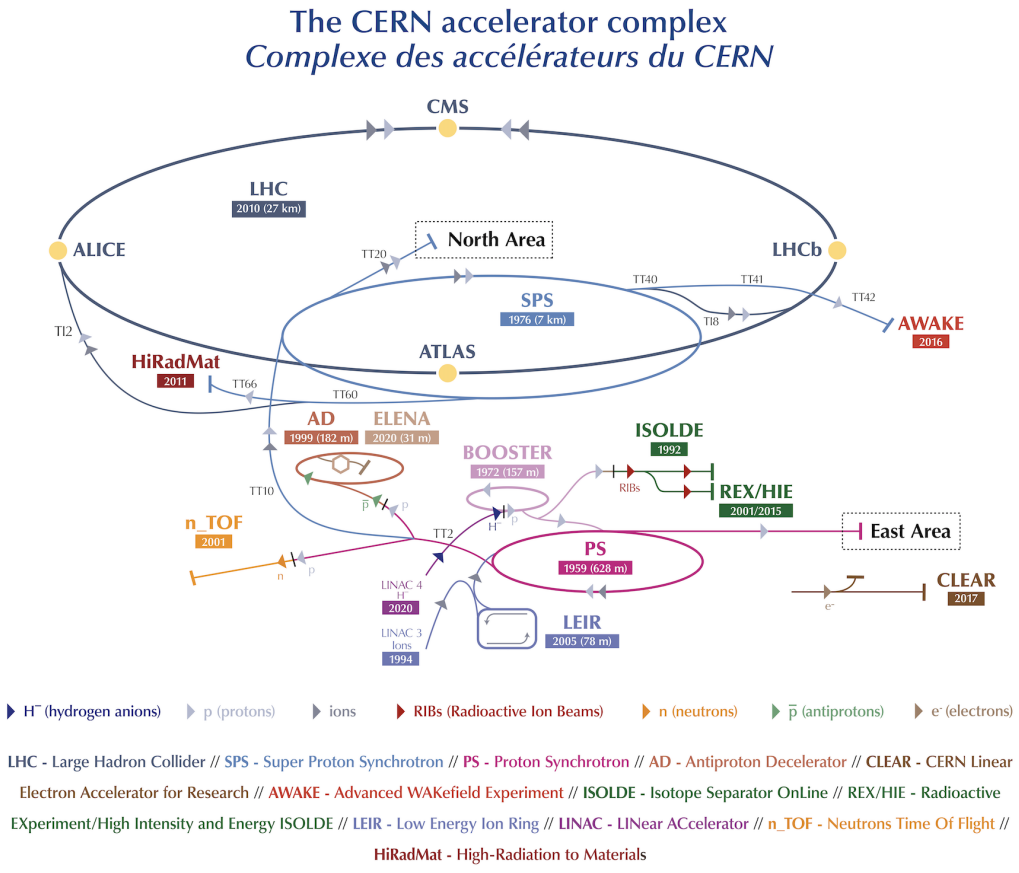

Figure 3.1: The LHC Accelerator Complex. Various accelerators are used to ramp up the energy of protons to 13 TeV, with 4 collision points on the final LHC ring for experiments.

and measure the properties of the Higgs Boson, a success story announced to the world in July of 2012, but its versatility has deepened our understanding of electroweak and strong physics as well. The LHC produces abundant W and Z bosons, as well as top quarks.

LHC Optics

The Large Hadron Collider employs a number of different magnets as beam optics. The three principle ones are magnetic dipoles, quadrapoles, and sextupoles. The 1232 superconducting dipoles are responsible for bending the proton beams around the circumference of the LHC. Quadrapoles serve the same purpose as focusing lenses for optical light. Along one axis perpendicular to the beam, they focus the proton packet. Along the other, they defocus the packet. By alternating the orientation 90 degrees, the quadrapoles focus and defocus the protons, keeping the packet within an acceptable envelope. Sextupole magnets (688 in total) serve for cleaning the beam packets, due to a magnetic field shape which becomes increasingly stronger the further from the center of the beam. This Lorentz force on the protons furthest from the center is greatest, which tends to drive protons straying too far from the beam axis out of the packet.

3.2 The Compact Muon Solenoid Detector

The Compact Muon Solenoid experiment is one of four large particle physics experiments constructed at an Interaction Point (IP) on the Large Hadron Collider, specifically at Point 5 (P5). Along with ATLAS at Point 1 (P1), it follows a general-purpose design for studying proton-proton collisions, and during the 2015 – 2018 timeframe ("Run II") the

LHC ran at a nominal center-of-mass (CoM) energy of 13 TeV. CMS and ATLAS share the ring with LHCb and ALICE, whose designs are optimized for B-physics [47] and Heavy Ion physics (such as the study of the quark-gluon plasma) [48], respectively. The detector is designed to be as hermetic as possible, with a silicon tracker forming the heart of the system, surrounded by calorimeters to measure particle energy, a Solenoid capable of generating a 3.8 Tesla magnetic field (to induce a Lorentz force on charged particles, central to momentum determination), and muon detectors within the iron return yolk. The detector has a mass of 12.5 kilotons, and is roughly cylindrical, with a length of 21.6m and height of 14.6m. As its name suggests, identification of muons is central to the CMS design, and a system of 3 gaseous detector technologies were employed up through the end of 2018 (expanded to 4 for collisions beginning in 2022).

Pixel Tracker

The innermost detector in CMS is a Silicon Pixel Tracker ("Pixels"), whose primary job is efficiently reconstructing tracks and measuring the origin of the particles producing them. The Pixels have 4 cylindrical layers 3.0 cm – 16.0 cm from the beamline, approximately 0.55 m long (Barrel Pixels or BPIX), and 3 disks forming a cap structure at either end (Forward Pixels or FPIX), covering pseudo-rapidity -2.5 to 2.5. This represents an upgrade from the original design with 3 barrel layers and 2 endcap disks the subsystem had prior to the 2016 Year End Technical Stop (YETS), which notably increased the performance. As an example, the efficiency of a b-tagging algorithm improved from approximately 50% to 65% at the same false-tagging rate [49]. The name of this subsystem

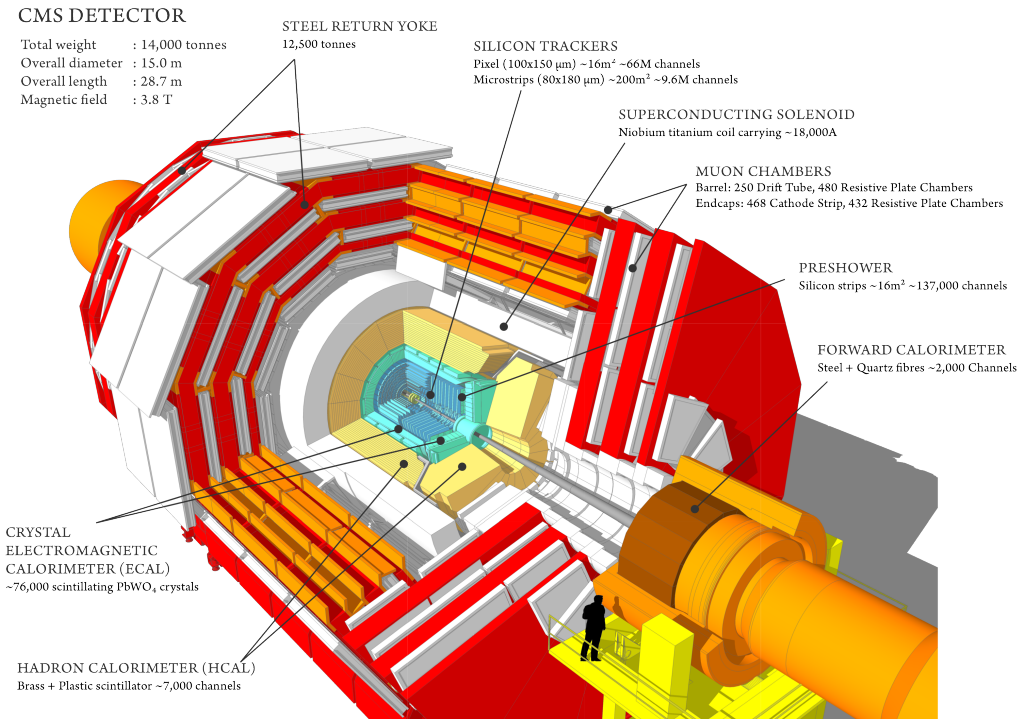

Figure 3.2: The CMS detector design prior to Long Shutdown 1. Upgrades and expansions to subdetectors not included, such as 72 chambers for CSCs (ME4/2) [1].

comes from the general structure employed: silicon sensors with a rectangular (100 by 150 μm^2) shape form the active detector material, and are read out through dedicated chips, connected by fiber-optic cables to the rest of the system. The subsystem must be kept at 251 Kelvin during operation by the CO_2 cooling system.

Strip Tracker

Outside the Pixels, another silicon detector is the Strip Tracker, enclosing the Pixels in a cylindrical volume (r = 1.3m, l = 5.8m). As its name suggests, the active area is formed of silicon microstrips, which are rotated 90 degrees in each layer such that the smaller resolution dimension alternates between η and ϕ. 10 layers in total provide the necessary lever-arms to measure the bending and hence the momentum of tracks from charged particles, including electrons and muons. The number of layers also permit the tracker to handle the high multiplicity of tracks the CMS detector can expect from LHC and High-Luminosity LHC collisions, with pileup (PU; additional proton-proton collisions besides the hardest one in a bunch-crossing) potentially reaching 200 at times. While there are dedicated triggering and tracking detectors for muons outside the solenoid, the Strip (+Pixel) Tracker still provides the bulk of the momentum resolution for muons (up to several hundred GeV). CO_2 cooling keeps the subsystem at the operating temperature of 253 Kelvin.

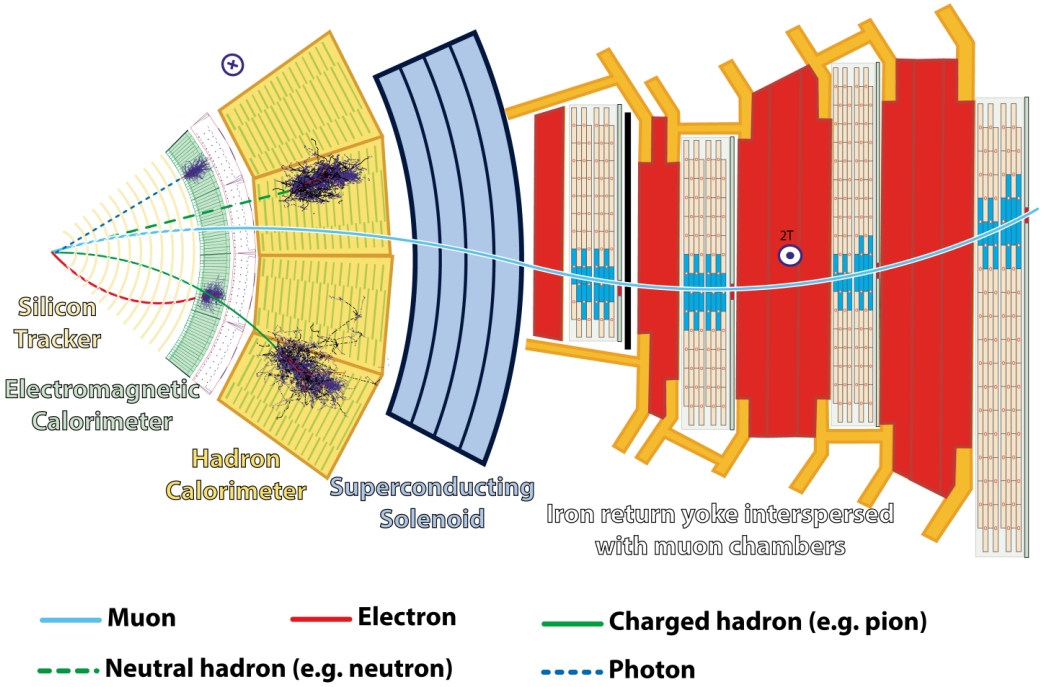

Figure 3.3: Visual description of how different particles are detected by CMS. Diagram is an x-y slice from the barrel section that includes DTs and RPCs for muon tracking and triggering [2].

Electromagnetic Calorimeter

The Electromagnetic Calorimeter (ECAL) is situated just outside the Strip and Pixel Trackers. The system consists of lead tungstate ($PbWO_4$) crystals up to pseudorapidity $|\eta| < 3.0$, which scintillate upon interaction with particles. The light generated is read out via photodiodes (Avalanche type – APDs – in the barrel and Vacuum type – VPTs – in the endcap). In front of the ECAL endcap is a preshower system, whose purpose is to enable π^0 rejection [46]. The ECAL subsystem is designed to detect the electromagnetic (EM) interaction of electrons and photons (through induced pair-production of electrons and positrons; it represents approximately 25 radiation lengths (X_0)), as shown in Fig. 3.3. Muons tend to deposit very little energy in the ECAL; neutral hadrons may deposit some energy, but typically escape to either be captured in the Hadronic Calorimeter or leave the detector entirely (virtually all neutrinos); electrons and photons are usually captured entirely within the system.

Hadron Calorimeter

The Hadronic Calorimeter (HCAL) is crucial for determining the energy of neutral particles, especially those generated by hadronization (formation of jets from colored particles), and plays an important role in particle identification. The CMS detector is effectively transparent for neutrinos, but together ECAL and HCAL can estimate the transverse component of missing momentum that is central to many analyses. The HCAL utilizes a sampling design, in which layers of brass and plastic scintillator are interleaved (iron return yolk with quartz fibers in the forward detector, where signals are detected via Cerenkov

light), covering $-5.0 < \eta < 5.0$. Over that η range, the subsystem spans 7 – 11 interaction lengths (λ_I). Wavelength-shifting fibers convert the scintillation light for readout by Photomultipliers (PMTs). Originally these were Hybrid PMTs, chosen to tolerate the high magnetic field inside the solenoid, but during Long Shutdown 1, these were upgraded. In the Hadron Forward (HF) calorimeter, multi-anode PMTs were used; for the rest of the HCAL, Silicon Photomultipliers (SiPMs) were chosen to improve measurement of the shower development and permit depth segmentation. A general improvement in noise reduction was also achieved, via reduced sensitivity to spurious signals and better pulse-time measurement in the HF [50]. HCAL captures most of the energy of neutral particles besides neutrinos (since they only interact through the Weak force, the detector has insufficient interaction lengths to estimate their energy). CMS uses the estimates of the EM and hadronic energy from the two calorimeters (in combination with tracking information from the tracker and muon system) to perform particle identification.

Drift Tubes

Drift Tubes (DTs) are used for muon tracking in the barrel region, $|\eta| < 1.4$, where the magnetic field from the solenoid is weaker (down from 3.8T to 2T and below) and the neutron-induced background and rate of muons are both low. A rectangular aluminum cathode-tube contains an anode-wire, and is filled with ionizable gas, and a chamber is constructed from multiple DTs assembled together into r-ϕ or z measurement planes. The 250 chambers are divided among the 5 wheels of the barrel, with 3 inner layers composed of 12 DT chambers, and a 4th outer-most layer with 14 chambers. Operating in the amplification region of the gas-gain curve, DTs are a cost-effective technology for reconstructing

the position of muons passing through, utilizing the drift times of the electron cascade to the anode and ion-drift to the cathode to determine where in the active volume the muon deposited charge, and hence calculate the muon position vector to better than 100 μm and 1 milliradian.[46]

The Cathode Strip Chambers are the principal muon triggering detectors for the high-η region, $1.4 \leq |\eta| \leq 2.4$, and complement the silicon inner tracker for measuring muon momentum. The 540 chambers are separated into 18 or 36 chambers per ring, 2 or 3 rings per disk, 4 disks per endcap, and 2 endcaps. Each chamber consists of gold-plated tungsten wires sandwiched between a solid copper sheet and azimuthally-segmented copper strips (forming one interior wall of each of the 6 gas gaps; the other is solid copper) in each chamber. The copper strips are etched into the surface of G-10 Printed Circuit Board (PCB) material. The wire is not perfectly perpendicular, to accommodate the drift direction of charges in the magnetic field. In total, there are 266,112 cathode strips and 210,816 wire-groups (WG). The WGs are read out by Anode Front End Boards (AFEBs), which are installed at the edges of the chambers and connect to sixteen WGs across two layers.

Cathode Strip Chambers

The AFEBs send their data to the Anode Local Charged Track (ALCT) Base-board, which houses an ALCT Mezzanine that contains the main FPGA and logic for readout and forming ALCT to send to the trigger and data readout electronics. For the strip readout, Cathode Front End Boards are used, for which there are 4 to 7 boards, each

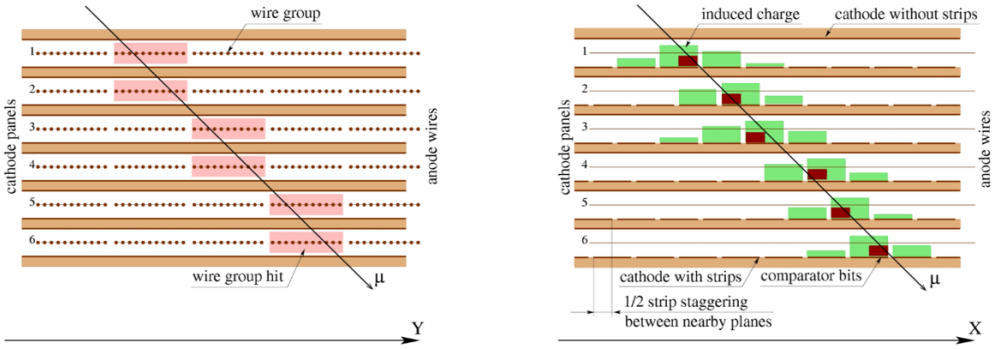

Figure 3.4: (left) The pattern of wiregroup hits formed by a muon passing through a CSC chamber. (right) The cathode strips are read out individually, with the detailed level of charge in 3 neighboring strips used to determine the per-layer hit location to better than half-strip resolution.

reading out all 6 layers of strips, 16 strips per layer. 40% argon, 50% CO_2, and 10% CF_4 (being reduced to 5% to meet CERN obligations to cut greenhouse-gas use) forms the working gas, where argon is the main component being ionized by muons and participating in the rest of the electron cascade and ion drift induced by the 2.9–3.6 kV working potential applied between anodes and cathodes, see Fig. 3.4. CO_2 acts as a quenching gas, and so-called magic gas (CF_4) protects against the worst aging effects. CSCs are used in this region of the detector due to their robustness, capable of sustaining relatively high rates of neutron-induced background and muon tracks passing through, and their tolerance for high and varying magnetic fields as necessarily exist in the endcap region. They have good timing resolution, which complements an excellent spatial resolution, allowing them to operate independently of any other muon detector and provide both triggering and tracking. More details of the system may be found in Appendix C.2.

Resistive Plate Chambers

To complement the excellent position resolution of the the DTs in the barrel region and CSCs in the endcap, Resistive Plate Chambers (RPCs) have been constructed and installed. In the barrel, the inner 3 layers of DT chambers are mated to 1 or 2 RPCs, and the outer rings of the CSC stations have matching RPCs. RPCs are another gaseous muon detector which operates in avalanche mode, which gives them a characteristically excellent time resolution and quick response (with weaker position resolution than DTs and CSCs), allowing for detailed identification of the originating bunch-crossing collision of the muons. Each chamber has a gas gap of 2mm, with bakelite walls (coated in linseed oil). Unlike the streamer mode typically used with RPCs, in avalanche mode, a muon ionizing the gas between the cathode and anode induces an electron cascade which must be amplified by readout electronics. While this has disadvantages, it improves the recovery time, allowing the chambers to withstand particle rates over 10kHz. Together with DTs and CSCs, the RPCs contribute to the Level-1 Trigger System. RPCs cannot sustain the highest particle rate regions covered by the innermost rings of CSCs, but an improved (iRPC) chamber type will be installed before HL-LHC collisions to provide redundancy and coverage in the ME3/1 and ME4/1 rings.

Precision Proton Spectrometer

The PPS subdetector was previously a separate but coupled-to-CMS experiment, TOTEM, which utilizes Roman Pots (RPs) approximately 200m upstream/downstream of the CMS collision point. The RPs contain diamond detectors for measuring the momentum

of intact Protons from (semi-)elastic p-p collisions. The PPS is not used in many LHC analyses as it's rare for both or even one of the interacting protons to remain intact for the hard interactions typically analyzed. However, the PPS provides crucial information about momentum along the z direction that's typically missing in those kinds of analyses; it is being utilized to do specialized fully-inclusive physics measurements and searches that are otherwise impossible.

BRIL

The Beam Radiation, Instrumentation, and Luminosity (BRIL) group is responsible for dozens of luminometers in and around the CMS detector. The luminometers make use of both dedicated detectors and information from the main subdetectors described here to measure the instantaneous (and therefore the integrated luminosity) at CMS. Eq. 3.1 shows the factorization of parameters determining the instantaneous value. The Lorentz factor (γ), number of bunches (k_B), number of protons in a bunch (N_p), revolution frequency (f), reduction factor (F; due to the crossing angle), normalized transverse emittance (ϵ_n; design value is 3.75 μm), and Betatron function (β^*) all contribute [46].

$$\mathcal{L} = \frac{\gamma f k_B N_p^2}{4\pi\epsilon_n \beta^*} F \tag{3.1}$$

This is a crucial responsibility to enable physics measurements at CMS, since the instantaneous and integrated luminosity set predictions for how many events from a given process we may expect, as well as how much PU (c.f. Sec. 4.5) can be expected on-average in the events analyzed. Many downstream effects in terms of reconstruction and performance that crucially depend on characterizing the "typical" interactions in given blocks of time

(Luminosity Block or Section) therefore depend on the BRIL group's work. BRIL produces summaries characterizing the final luminosity uncertainties and correlations between years for analyses to apply in their fits to simulation.

Chapter 4

Four Top Analysis

4.1 Data Sample and Event Simulation

Monte Carlo simulation from the 2017 ReReco campaign, and the 2018 ReReco (Run Periods A, B, C) and 2018 Prompt (Run Period D) are used. These samples are from the NanoAODv7 data tier/campaign. Data from the NanoAODv7 data tier/campaign are used. Events collected during periods when the detector is not fully ready are filtered out via the standard Golden JSON files for 2017 and 2018:

- Cert_294927-306462_13TeV_EOY2017ReReco_Collisions17_JSON_v1
- Cert_314472-325175_13TeV_17SeptEarlyReReco2018ABC_PromptEraD_Collisions18_JSON

Background samples are grouped into several split/aggregate categories for the final fit. The EWK category includes Drell-Yan, W + Jets, and diboson (WW, WZ, ZZ). $t\bar{t}$ is split into the $t\bar{t}b\bar{b}$ and $t\bar{t}b$ (which includes $t\bar{t}\bar{b}$) subprocesses and non-$t\bar{t}b(\bar{b})$, with single top ($\bar{t}W$ and tW) being merged with the latter. $t\bar{t}V$ + Jets includes $t\bar{t}W$ + Jets

and $t\bar{t}Z$ + Jets. $t\bar{t}H$ is kept as a single category. Finally, ttultrarare groups together all remaining processes listed with $t\bar{t}$ and associated production ($t\bar{t}HH$, $t\bar{t}ZH$, $t\bar{t}ZZ$, $t\bar{t}WH$, $t\bar{t}WZ$, $t\bar{t}WW$, $t\bar{t}tW$, $t\bar{t}tJ$). The $t\bar{t}W$ and $t\bar{t}Z$ samples used in this analyis are chosen over alternatives in other generators due to the latter having very low simulation statistics and high fraction of negative weights relative to the ones listed in this section. Those samples therefore produced unreliable background templates. Overall, these backgrounds are less important than the ttultrarare and $t\bar{t}H$ process groups in the signal rich categories.

For splitting $t\bar{t}$ into $t\bar{t}b\bar{b}/t\bar{t}b$ and non-$t\bar{t}b(\bar{b})$, a Ghost-Hadron Clustering algorithm is employed. The method reclusters reconstructed jets in simulation while including particles from Generator level (with their momenta scaled down to infinitesimal values, such that the kinematics are not affected). When a jet clusters a B or C hadron, the algorithm tags this jet, and if it does not come from the decay of the main $t\bar{t}$ pair, it is classified as an additional b-jet or c-jet, respectively. For this analysis, no distinction is made between the case when a jet includes two B-hadrons and a single B-hadron, but only events including one or two additional jets, each with at least one B-hadron clustered within, are classified as $t\bar{t}b\bar{b}/t\bar{t}b$. All other events are classified as non-$t\bar{t}b(\bar{b})$.

The signal sample for 2017 integrates a bug-fix relating to the number of partons used in the Born approximation, indicated in the dataset name. Ultimately, this fix was found in other samples to not have a substantial effect on observables; nevertheless the version with the error is not used. The 2018 signal sample was not affected by this error, and uses the nominal naming convention, therefore matching the listed 2017 sample here.

Table 4.1: Simulation samples used in the analysis. "DL" stands for the dilepton channel, "AH" for all hadronic, "SL" for semileptonic, "UE" for underlying event and "GF" for the filtered generator events.

Process	σ (pb)	Data Set
$t\bar{t}$ (AH)	380.11	/TTToHadronic_TuneCP5_13TeV-powheg-pythia8
$\rightarrow h_{\text{damp}}$ (down/up)	-	/TTToHadronic_hdamp(DOWN/UP)_TuneCP5_13TeV-powheg-pythia8
$\rightarrow$ UE (down/up)	-	/TTToHadronic_TuneCP5(down/up)_13TeV-powheg-pythia8
$t\bar{t}$ (DL)	87.33	/TTTo2L2Nu_TuneCP5_13TeV-powheg-pythia8
$\rightarrow h_{\text{damp}}$ (down/up)	-	/TTTo2L2Nu_hdamp(DOWN/UP)_TuneCP5_13TeV-powheg-pythia8
$\rightarrow$ UE (down/up)	-	/TTTo2L2Nu_TuneCP5(down/up)_13TeV-powheg-pythia8
$t\bar{t}$(DL) (GF)	1.456	/TTTo2L2Nu_HT500Njet7_TuneCP5_13TeV-powheg-pythia8
$\rightarrow h_{\text{damp}}$ (down/up)	-	/TTTo2L2Nu_HT500Njet7_hdamp(DOWN/UP)_TuneCP5_13TeV-powheg-pythia8
$\rightarrow$ UE (down/up)	-	/TTTo2L2Nu_HT500Njet7_TuneCP5(down/up)_13TeV-powheg-pythia8
$t\bar{t}$ (SL)	364.31	/TTToSemiLeptonic_TuneCP5_13TeV-powheg-pythia8
$\rightarrow h_{\text{damp}}$ (down/up)	-	/TTToSemiLeptonic_hdamp(DOWN/UP)_TuneCP5_13TeV-powheg-pythia8
$\rightarrow$ UE (down/up)	-	/TTToSemiLeptonic_TuneCP5(down/up)_13TeV-powheg-pythia8
$t\bar{t}$ (SL) (GF)	2.268	/TTToSemiLepton_HT500Njet9_TuneCP5_13TeV-powheg-pythia8
$\rightarrow h_{\text{damp}}$ (down/up)	-	/TTToSemiLepton_HT500Njet9_hdamp(DOWN/UP)_TuneCP5_13TeV-powheg-pythia8
$\rightarrow$ UE (down/up)	-	/TTToSemiLepton_HT500Njet9_TuneCP5(down/up)_13TeV-powheg-pythia8

Table 4.2: Simulation samples used in the analysis. "DL" stands for the dilepton channel, "AH" for all hadronic, "SL" for semileptonic, "UE" for underlying event and "GF" for the filtered generator events.

Process	σ (pb)	Data Set
tW	19.553	/ST_tW_top_5f_NoFullyHadronicDecays_TuneCP5_13TeV-powheg-pythia8
$\bar{t}$W	19.553	/ST_tW_antitop_5f_NoFullyHadronicDecays_TuneCP5_13TeV-powheg-pythia8
t$\bar{t}$H (SL, bb)	0.1285	/ttHTobb_ttToSemiLep_M125_TuneCP5_13TeV-powheg-pythia8
t$\bar{t}$H (DL, bb)	0.0308	/ttHTobb_ttTo2L2Nu_M125_TuneCP5_13TeV-powheg-pythia8
t$\bar{t}$W+Jets	0.611	/ttWJets_TuneCP5_13TeV_madgraphMLM_pythia8
t$\bar{t}$Z+Jets	0.783	/ttZJets_TuneCP5_13TeV_madgraphMLM_pythia8
t$\bar{t}$HH	0.000741	/TTHH_TuneCP5_13TeV-madgraph-pythia8
t$\bar{t}$WH	0.00157	/TTWH_TuneCP5_13TeV-madgraph-pythia8
t$\bar{t}$WW	0.00788	/TTWW_TuneCP5_13TeV-madgraph-pythia8
t$\bar{t}$WZ	0.00297	/TTWZ_TuneCP5_13TeV-madgraph-pythia8
t$\bar{t}$ZH	0.00125	/TTZH_TuneCP5_13TeV-madgraph-pythia8
t$\bar{t}$ZZ	0.00157	/TTZZ_TuneCP5_13TeV-madgraph-pythia8
t$\bar{t}$tJ	0.000474	/TTTJ_TuneCP5_13TeV-madgraph-pythia8
t$\bar{t}$tW	0.00788	/TTTW_TuneCP5_13TeV-madgraph-pythia8
t$\bar{t}$t$\bar{t}$	0.012	/TTTT_TuneCP5_PSweights_13TeV-amcatnlo-pythia8_correctnPartonsInBorn (2017)
		/TTTT_TuneCP5_13TeV-amcatnlo-pythia8 (2018)

Table 4.3: Simulation samples used in the analysis. "DL" stands for the dilepton channel, "AH" for all hadronic, "SL" for semileptonic, "UE" for underlying event and "GF" for the filtered generator events.

Process	σ (pb)	Data Set
D-Y (DL, HT100)	181.92	/DYJetsToLL_M-50_HT-100to200_TuneCP5_PSweights_13TeV-madgraphMLM-pythia8
D-Y (DL, HT200)	54.933	/DYJetsToLL_M-50_HT-200to400_TuneCP5_PSweights_13TeV-madgraphMLM-pythia8
D-Y (DL, HT400)	7.8581	/DYJetsToLL_M-50_HT-400to600_TuneCP5_PSweights_13TeV-madgraphMLM-pythia8
D-Y (DL, HT600)	1.9477	/DYJetsToLL_M-50_HT-600to800_TuneCP5_PSweights_13TeV-madgraphMLM-pythia8
D-Y (DL, HT800)	0.8587	/DYJetsToLL_M-50_HT-800to1200_TuneCP5_PSweights_13TeV-madgraphMLM-pythia8
D-Y (DL, HT1200)	0.2028	/DYJetsToLL_M-50_HT-1200to2500_TuneCP5_PSweights_13TeV-madgraphMLM-pythia8
D-Y (DL, HT2500)	0.003219	/DYJetsToLL_M-50_HT-2500toInf_TuneCP5_PSweights_13TeV-madgraphMLM-pythia8
WJets (SL, HT100)	1627.45	/WJetsToLNu_HT-100To200_TuneCP5_13TeV-madgraphMLM-pythia8
WJets (SL, HT200)	435.237	/WJetsToLNu_HT-200To400_TuneCP5_13TeV-madgraphMLM-pythia8
WJets (SL, HT400)	59.181	/WJetsToLNu_HT-400To600_TuneCP5_13TeV-madgraphMLM-pythia8
WJets (SL, HT600)	16.3296	/WJetsToLNu_HT-600To800_TuneCP5_13TeV-madgraphMLM-pythia8
WJets (SL, HT800)	8.0005	/WJetsToLNu_HT-800To1200_TuneCP5_13TeV-madgraphMLM-pythia8
WJets (SL, HT1200)	2.1419	/WJetsToLNu_HT-1200To2500_TuneCP5_13TeV-madgraphMLM-pythia8
WJets (SL, HT2500)	0.1634	/WJetsToLNu_HT-2500ToInf_TuneCP5_13TeV-madgraphMLM-pythia8
WW	118.7	/WW_TuneCP5_13TeV-pythia8
WZ	47.13	/WZ_TuneCP5_13TeV-pythia8
ZZ	16.523	/ZZ_TuneCP5_13TeV-pythia8

Table 4.4: 2017 and 2018 Data, indicating the primary decay channel targeted.

Chan.	Run Period	Data Set
ee	2017B	/DoubleEG/Run2017B-02Apr2020-v1/NANOAOD
	2017C	/DoubleEG/Run2017C-02Apr2020-v1/NANOAOD
	2017D	/DoubleEG/Run2017D-02Apr2020-v1/NANOAOD
	2017E	/DoubleEG/Run2017E-02Apr2020-v1/NANOAOD
	2017F	/DoubleEG/Run2017F-02Apr2020-v1/NANOAOD
	2018A	/EGamma/Run2018A-02Apr2020-v1/NANOAOD
	2018B	/EGamma/Run2018B-02Apr2020-v1/NANOAOD
	2018C	/EGamma/Run2018C-02Apr2020-v1/NANOAOD
	2018D	/EGamma/Run2018D-02Apr2020-v1/NANOAOD
eμ	2017B	/MuonEG/Run2017B-02Apr2020-v1/NANOAOD
	2017C	/MuonEG/Run2017C-02Apr2020-v1/NANOAOD
	2017D	/MuonEG/Run2017D-02Apr2020-v1/NANOAOD
	2017E	/MuonEG/Run2017E-02Apr2020-v1/NANOAOD
	2017F	/MuonEG/Run2017F-02Apr2020-v1/NANOAOD
	2018A	/MuonEG/Run2018A-02Apr2020-v1/NANOAOD
	2018B	/MuonEG/Run2018B-02Apr2020-v1/NANOAOD
	2018C	/MuonEG/Run2018C-02Apr2020-v1/NANOAOD
	2018D	/MuonEG/Run2018D-02Apr2020-v1/NANOAOD
μμ	2017B	/DoubleMuon/Run2017B-02Apr2020-v1/NANOAOD
	2017C	/DoubleMuon/Run2017C-02Apr2020-v1/NANOAOD
	2017D	/DoubleMuon/Run2017D-02Apr2020-v1/NANOAOD
	2017E	/DoubleMuon/Run2017E-02Apr2020-v1/NANOAOD
	2017F	/DoubleMuon/Run2017F-02Apr2020-v1/NANOAOD
	2018A	/DoubleMuon/Run2018A-02Apr2020-v1/NANOAOD
	2018B	/DoubleMuon/Run2018B-02Apr2020-v1/NANOAOD
	2018C	/DoubleMuon/Run2018C-02Apr2020-v1/NANOAOD
	2018D	/DoubleMuon/Run2018D-02Apr2020-v1/NANOAOD

4.2 Triggers Used for Data Collection

Events are selected using the dilepton triggers listed in Tables 4.5 and 4.6. These sets of HLT triggers are not prescaled triggers for the data from either year. Because the input datastreams are non-exclusive, a priority cascade is applied to avoid duplicate event selection. Events from the eμ datastream which fire the eμ triggers are selected first. Next, events from the μμ datastream are selected if they fire the μμ triggers, but do not fire the eμ triggers. Similarly, events from the ee datastream are selected if and only if they fire the ee triggers but not the eμ or μμ triggers. Simulated events are required to meet the same requirements to enter each of the analyses. An explicit check has been performed after the baseline analysis (Section 4.4) to ensure that the event selection is exclusive for the three channels in both data and simulation.

Table 4.5: List of eμ, μμ and ee triggers used for the 2017 opposite sign dilepton analyses

Channel	Run period(s)	HLT trigger name
eμ	BCDEF	HLT_Mu23_TrkIsoVVL_Ele12_CaloIdL_TrackIdL_IsoVL_DZ
	BCDEF	HLT_Mu12_TrkIsoVVL_Ele23_CaloIdL_TrackIdL_IsoVL_DZ
μμ	B	HLT_Mu17_TrkIsoVVL_Mu8_TrkIsoVVL_DZ
	CDEF	HLT_Mu17_TrkIsoVVL_Mu8_TrkIsoVVL_DZ_Mass3p8
ee	BCDEF	HLT_Ele23_Ele12_CaloIdL_TrackIdL_IsoVL

Table 4.6: List ofeμ, μμ and ee triggers used for the 2018 opposite sign dilepton analyses

Channel	Run period(s)	HLT trigger name
eμ	ABCD	HLT_Mu12_TrkIsoVVL_Ele23_CaloIdL_TrackIdL_IsoVL_DZ
	ABCD	HLT_Mu23_TrkIsoVVL_Ele12_CaloIdL_TrackIdL_IsoVL_DZ
μμ	ABCD	HLT_Mu17_TrkIsoVVL_Mu8_TrkIsoVVL_DZ_Mass3p8
ee	ABCD	HLT_Ele23_Ele12_CaloIdL_TrackIdL_IsoVL

4.3 Particle and Jet Reconstruction

Electrons, muons, and jets are reconstructed using the Particle Flow (PF) algorithm, which combines information from the tracker, electromagnetic and hadronic calorimeters, and the muon systems.

Electrons are identified using the 94X-V2-loose selection criteria, [51], which includes isolation criteria. Muons are identified using the loose ID criteria[52]. As the muon ID does not include any isolation requirements, we also apply the tight PF requirement[52].

Jets are clustered with the anti-k_{T} algorithm using a radius parameter of 0.4, and the Charged Hadron Subtraction method is used in the clustering to mitigate the effects of pileup from additional pp collisions. This method subtracts charged particles with tracks not coming from the primary collision, and uses these to estimate the energy fraction from the corresponding neutral particles. That estimated neutral energy is then removed from the overall energy of the jet.

The energy scales of the reconstructed jets are adjusted using a factorized approach. The level 1 corrections are derived and applied separately for data and simulation to remove the dependence on the additional pp collisions in the events. This reduces the bias for the remaining corrections. Corrections are also performed to account for the difference between particle level and reconstructed jet p_{T}. This adjusts for the momentum that escapes undetected via neutrinos. Residual corrections derived from the momentum balance in Z + jets events are used to improve the data-to-simulation agreement to the percent level. The jet energy resolution in the simulation has been seen to be better than in the data. This is corrected by adding an additional smearing to the energy of simulated jets

so that the resolution of simulated jets matches that of data jets. Table 4.7 lists the tags for corrections used in the analyses presented here. These were applied via the nanoAOD-tools package.

Table 4.7: The jet energy and resolution correction tags used in this analysis

Jet energy scale corrections		
	2017	2018
MC JES	Fall17_17Nov2017_V32_MC	Autumn18_V19_MC
Data JES RunA	N/A	Autumn18_RunA_V19_DATA
Data JES RunB	Fall17_17Nov2017B_V32_DATA	Autumn18_RunB_V19_DATA
Data JES RunC	Fall17_17Nov2017C_V32_DATA	Autumn18_RunC_V19_DATA
Data JES RunD	Fall17_17Nov2017DE_V32_DATA	Autumn18_RunD_V19_DATA
Data JES RunE	Fall17_17Nov2017DE_V32_DATA	N/A
Data JES RunF	Fall17_17Nov2017F_V32_DATA	N/A
Jet energy resolution corrections		
	2017	2018
MC JER	Fall17_V3_MC	Autumn18_V7_MC

The efficient identification of jets that originate from the fragmentation of b quarks is essential in the search for $t\bar{t}t\bar{t}$ events. For this analysis the tagging[53] of b jets is performed using the DeepJet algorithm[54], whose performance on data is superior to the previous algorithms, including DeepCSV[55]. This algorithm uses deep neural networks to define a discriminant which characterizes the quality of the tag probability of a particular jet. This analysis uses the medium working point[56].

4.4 Baseline Event Selection

The baseline criteria for event selection are detailed below. Events passing these criteria are used in all control and signal categories in the analysis. Two isolated, oppo-

sitely charged leptons (e, μ) are selected with p_{T} above the corresponding trigger threshold outlined in section 4.2. For events with same flavor leptons, two cuts on the invariant mass are employed to remove low-mass resonances and events with leptonically decaying Z bosons. The invariant mass must not be below 20 GeV or within 15 GeV of the Z boson mass, 91 GeV, in these two channels.

4.4.1 Primary Vertex and MET Selection

The starting point is to impose the standard CMS criteria to define the Primary Vertex (PV) and the Missing Transverse Momentum (MET). These are defined below. Standard Primary Vertex and MET filter requirements are imposed for all events thereafter.

The Primary Vertex is defined as the collision vertex, reconstructed from charged particle tracks in the detector (alternatives algorithms exist for photon-enriched event signatures), with the most momentum and energy associated with it (i.e. most likely to be from the hardest proton-proton interaction in the collision bunches). Events passing the following criteria are kept for further analysis.

- number of degrees of freedom > 4 - corresponds to the number of good tracks associated to the reconstructed vertex
- $|z_{PV}| < 24.0$ - distance from the center of the detector along the z-axis
- $\rho_{PV} < 2.0$ - distance from the center of the beamline, the radial coordinate

The initial proton-proton interaction has approximately zero transverse momentum, and therefore through conservation laws, the MET may be calculated as the balance of transverse momentum that zeros the vectorial sum of measurable particles in the detector.

This serves as the proxy for the transverse component of neutrino momentum/momenta in the event. The MET must satisfy several filters designed to remove bad reconstruction, noise, and other erroneous effects.

- MET FIlter: globalSuperTightHalo2016Filter
- MET FIlter: goodVertices
- MET FIlter: HBHENoiseFilter
- MET FIlter: HBHENoiseIsoFilter
- MET FIlter: EcalDeadCellTriggerPrimitiveFilter
- MET FIlter: BadPFMuonFilter
- MET FIlter: ecalBadCalibFilterV2 (2017, 2018)

4.4.2 Jet and Common Lepton Selection

Criteria for jet selection and hadronic activity are applied to all events. Analyzing the AK4 jets (reconstructed according to the description in Sec. 4.3), the following criteria must be met:

- Four or more jets, with $p_{\mathrm{T}} > 30\,\mathrm{GeV}$ and $|\eta| < 2.5$, $\Delta\ R > 0.4$ cross-cleaned against selected leptons, and passing Jet PileUp ID when $30\,\mathrm{GeV} \leq p_{\mathrm{T}} < 50\,\mathrm{GeV}$ (Loose working point, 94X and 102X training in 2017 and 2018, respectively)
- Events with any number of b tagged jets are kept to facilitate shape calibration; control and signal regions only contain events with two or more medium DeepJet tagged jets.

- $H_{\mathrm{T}} > 500\,\mathrm{GeV}$

In addition to the jet criteria, a number of common lepton selection criteria are applied to enumerate the number of isolated muons and electrons in the event. In order to select prompt leptons from $\mathrm{W}^{\pm}$ decays (equally valid for Z^{0}), the following cuts on distance from the PV and 3D Impact Parameters are applied to the reconstructed lepton candidates.

- Electrons must satisfy $|dz| < 0.10$ and $|3dImpactParameter| < 0.05$ in the barrel region or $|dz| < 0.20$ and $|3dImpactParameter| < 0.10$ in the endcap region.

- Muons must satisfy $|dz| < 0.2$ and $|dxy| <= 0.10$

4.4.3 Channel Specific Selection

The three dilepton decay channels have similar criteria, differing mostly in the μ and e specific variables. The two same-flavor channels invoke invariant mass cuts to both match trigger selection criteria and to remove the large background from Z^{0} and γ^{*} decays, and other pair production from sub-20 GeV. The details are as follows. μμ channel:

- Pass non-prescaled di-muon trigger

- Exactly two loose cut-based ID muons with tight Particle Flow isolation, with leading muon $p_{\mathrm{T}} > 25\,\mathrm{GeV}$, subleading $p_{\mathrm{T}} > 15\,\mathrm{GeV}$, with $|\eta| < 2.4$

- No additional loose cut-based ID/tight isolated muons or electrons, with the same p_{T} and η restrictions as the subleading selected leptons

- $M_{\mu\mu} > 20\,\mathrm{GeV}$ and $|M_{\mu\mu} - 91GeV| > 15\,\mathrm{GeV}$

eμ channel:

- Pass non-prescaled muon-electron trigger
- Exactly one loose cut-based ID muon with tight Particle Flow isolation and loose cut-based ID electron, with leading lepton $p_{\mathrm{T}} > 25\,\mathrm{GeV}$, subleading $p_{\mathrm{T}} > 15\,\mathrm{GeV}$, and $|\eta| < 2.4$ (muons) or 2.5 (electrons)
- No additional loose cut-based ID/tight isolated muons or electrons, with the same p_{T} and η restrictions as the subleading selected leptons

ee channel:

- Pass non-prescaled di-electron trigger
- Exactly two loose cut-based ID electrons, with leading electron $p_{\mathrm{T}} > 25\,\mathrm{GeV}$, subleading $p_{\mathrm{T}} > 15\,\mathrm{GeV}$, and $|\eta| < 2.5$
- No additional loose cut-based ID/tight isolated muons or electrons, with the same p_{T} and η restrictions as the subleading selected leptons
- $M_{\mathrm{ee}} > 20\,\mathrm{GeV}$ and $|M_{\mathrm{ee}} - 91\,\mathrm{GeV}| > 15\,\mathrm{GeV}$

4.5 Corrections Applied to MC Simulation

4.5.1 Leptons

The muons and electron simulations are corrected to account for discrepancies between the baseline simulation and the data. For electrons, dedicated measurements from CMS are used for the reconstruction and identification efficiencies, using the selected ID and working point. For muons, both the ID and isolation efficiencies are accounted for,

again with the selected ID and working point, based on official measurements from the Muon POG. The corrections are parameterized in terms of lepton transverse momentum (p_{T}) and pseudorapidity (η), where the absolute value for the latter is used to match the derived corrections.

4.5.2 b-jet Tagging

As originally applied, there is a small shape mismatch in the DeepJet results between the data and the simulation. This is corrected by applying reshaping corrections to the simulation to make it match the data. These are taken from the BTV POG[57] which has measured discriminant scale factors using an iterative fit approach, which parameterizes the corrections as a function of the jet p_{T}, η, discriminant score, and jet type (i.e. b quark, c quark, lighter quark or gluon). The reshaping is accomplished by calculating the product of all selected jets' scale factors and applying it as a multiplicative factor to the event weight. This process by itself does not inherently preserve the normalization of the simulation and produces changes in the cross section. To account for this, a renormalization factor is calculated for each Monte Carlo process. However, because the scale factors are derived in Z + Jets and $\mathrm{t\bar{t}}$ events, they do not necessarily account for all relevant parameters when applied to the analysis phase space, which requires significantly more jets. The renormalization factor is parameterized in terms of jet multiplicity and the scalar sum of jet p_{T} (H_{T}) to account for additional dependencies not captured by the iterative fit derivation, and is computed separately for each sample and shape-varying systematic variation, with all analysis cuts applied except for b tag multiplicity (events with any number of b tags enter the numerator and denominator for the renormalization factor calculation). This si-

multaneously renormalizes the simulation and accounts for the discrepancy that would exist when extrapolating from the derivation phase space to the analysis phase spaces. Examples of renormalization maps are available in Appendix D.11.

4.5.3 Pileup

As the Monte Carlo used in these analyses were produced prior to completion of corresponding data taking, the presumed distribution of additional pp collisions does not match, and so simulation is reweighted using a shape-only correction to better match the real data. The MC distribution of true (Poisson mean) pileup, from which the per-event additional pp interactions is sampled, is reweighted to match the minimum bias cross section measured in data, corresponding to 69.2 mb[58]. More details and cross-checks are in Appendix D.8.

4.5.4 ECAL Prefiring Correction (2017 data)

During the 2016 and 2017 data taking period, the ECal prefiring issue permitted certain events to self-veto, due to a shift in the timing not being propagated to the Level 1 system. Because of this, objects reconstructed in the ECal could be improperly associated with the previous collission, and due to Level 1 trigger rules disallowing consecutive bunch crossings to both fire, an event can effectively veto itself from passing to the High Level Trigger stage. The EGamma POG has produced a recipe for calculating the probability of a simulation event failing to pass L1, and this is applied as an event level scalefactor for all Monte Carlo prior to 2018, where the issue was fully mitigated at trigger level. In later 2021, it was additionally discovered that a similar but smaller effect existed for muon

triggering. A compatible correction and uncertainty for muons were not available for this analysis, but the larger ECal prefiring is O(1%) and the muon prefiring is thus negligible in this analysis.

4.5.5 HLT z Vertex Position Inefficiency (2017 data)

Due to imprecise measurements of the online beamspot z position prior to run 301046, there is an inefficiency for electron triggers. A cylinder is defined around the presumed beamspot to use as the active tracking volume, and electrons' whose z_{vtx} lay outside it have very low efficiency. The effect was partially mitigated at the beginning of Run C, and fully after run 301046, approximately halfway through Run C. The NanoAOD data tier does not include required beamspot information to help reproduce the cuts at trigger level, so the data were studied in different channels in 3 periods, Run B, early Run C, and post-mitigation. Within statistical error, no measurable discrepancies were found between these three eras, and so a correction factor of 0.991, a weighted average of efficiencies over all of the 2017 Run periods derived in the EGamma POG for single electron triggers, is applied.

4.5.6 PileUp Jet ID

In keeping with the chosen btagging algorithm and calibration method, the loose working point of the pileup jet ID algorithm (94X and 102X trainings) is applied. This is a tagger trained to discriminate between jets originating from pileup and those from the hardscatter event. Compared to not using the ID, the dominant simulated background $t\bar{t}$ yields are reduced $4.1 - 7.9\%$ in the higher jet multiplicity control regions (6 to 8 jets) while

signal is reduced only $0.1 - 2.2\%$. The data see a corresponding reduction matching the main backgrounds ($4.6 - 6.8\%$).

4.5.7 Heavy Flavor Re-weighting

The measured and predicted cross sections for the $t\bar{t}b\bar{b}$ and $t\bar{t}b$ subprocesses in $t\bar{t}$ events do not agree, with the measurements being systematically above the predictions. However, the predictions for the cross sections for the $t\bar{t}c\bar{c}$ and the $t\bar{t}jj$ subprocesses are consistent with the latest measurements. To correct for this, a re-scaling of the simulated $t\bar{t}b\bar{b}$ and $t\bar{t}b$ cross section (which is applied to the $t\bar{t}b\bar{b}/t\bar{t}b$ subset of events in the principal $t\bar{t}$ samples used) is performed using the ratio of the data to predicted cross sections in POWHEG + Pythia8. Using the latest CMS measurement[59], this gives a correction factor of 36%, with an uncertainty of 8% for $t\bar{t}b\bar{b}/t\bar{t}b$. The cross section of the remaining events are within 1 standard deviation, and thus are not re-scaled based on this measurement.

4.5.8 Top p_{T} Modelling

QCD calculations to Next-to-Leading Order (NLO) precision do not accurately describe the shape of the top quark p_{T} distribution. In the 2016 CMS $t\bar{t}t\bar{t}$ searches in these channels an event level re-weighting correction was applied to the simulation to correct for this discrepancy. The size of the effect in the current POWHEG simulations is somewhat reduced compared to that of the simulations used in the analysis of the 2016 data. We have studied the effect of making a similar correction for this analysis, and we find that the correction has a negligible effect on the level of agreement between the data and simulation in the $t\bar{t}t\bar{t}$ sensitive regions. Moreover, the analysis sensitivity to making or not making

the correction is insignificant. We therefore choose not to make a correction for this effect. The results of this study are summarized in Appendix D.3 of this note.

4.5.9 Large H_{T} Corrections

The POWHEG + Pythia8 simulation shows growing disparity with the data as event activity increases in regions of phase space where events have high transverse mass, high-H_{T}, and high top (anti)quark p_{T}. While there is only a small normalization discrepancy when looking at events inclusive of low-H_{T} ($\geq 250\,\mathrm{GeV}$), this grows to the order of 20% once the 500 GeV cut to reduce backgrounds is introduced.

The simulation appears to model the kinematics (p_{T}, η) of the two hardest jets reasonably well, and ratio of data to simulation is consistent with flat beyond an overall normalization factor in both 2017 and 2018. These jets will be predominantly from the b quark hadronization in the leptonically-decaying top quarks, as $\mathrm{t\bar{t}}$ constitutes the bulk of the background. However, the remaining jets show an additional shape discrepancy, and these addional jets in data are softer in p_{T}. The 1 b-tag control region is used to correct for the normalization difference between the data and simulation. This is determined from the combined 2017 and 2018 datasets, so as to minimize the statistical fluctuations in the data. The correction is derived from the non-$\mathrm{t\bar{t}b(\bar{b})}$ component of $\mathrm{t\bar{t}}$. The result is a SF of 0.79 ± 0.054 applied to the non-$\mathrm{t\bar{t}b(\bar{b})}$ components of $\mathrm{t\bar{t}}$ in simulation, where the uncertainty is derived from the difference between the correction and the two individual years' 1 b-tag control categories. The residual after SF application is well within systematic uncertainties (for the analysis phase space). Weighting by jet multiplicity (as was done for the 2016 OSDL $\mathrm{t\bar{t}t\bar{t}}$ analysis) was also studied. The results, shown in Appendix D.2

indicate that this is not needed for the 2017 and 2018 simulation, which uses an updated TuneCP5 that demonstrates better agreement with the data in the analysis phase space.

4.6 Control distributions

This section contains pre-fit distributions in various event categories. In subsection 4.6.1, for each year and dilepton decay channel, variables using events with exactly 2 b-tagged jets and 4 or more jets are plotted. These plots serve as checks of the data to simulation agreement in a background-dominated phase space. Similarly, in subsection 4.6.2, there are corresponding plots using events with 3 or more b-tagged jets. These serve as more signal-like control regions, dominated by the $t\bar{t}b\bar{b}$ background.

The errors shown in the top panels are total shape systematic and simulation statistical uncertainties added in quadrature using their pre-fit values. These shape uncertainties may or may not have a rate-effect on the total inclusive cross-section of a given simulation sample, but usually have a rate-effect in each category (b-tagged jet multiplicity or jet multiplicity). In the ratio plots, the uncertainties are separated into the systematic and statistical components. Normalization-only uncertainties (log normal for Combine fitting) are not displayed here.

4.6.1 Background Regions

The following plots show pre-fit control distributions in the background-dominant regions for the three channels in the 2017 and 2018 data. The plots are made for data and simulation events which pass the baseline selection defined in Section 4.4 of which exactly 2 jets are medium b tagged using the DeepJet algorithm. Additional plots with 3 or more b tagged jets shown in the next subsection 4.6.2. In general, the distributions in this section show good agreement between data and simulation, with many distributions having agreement within data statistical uncertainties alone. The systematic uncertainty bands do not include some pure normalization uncertainties (across all categories) to help isolate shape compatibility.

- Jet p_{T}: The p_{T} distributions for the four highest p_{T} jets - Figs. 4.1, 4.2, 4.3, 4.4.
- Jet η: The η distributions for the the two leading jets - Figs. 4.5, 4.6.
- Jet multiplicity: Fig. 4.7
- b tagged jet multiplicity: Fig. 4.8
- Deep Jet b-tag discriminants, p_{T} ranking: The discriminant distributions for the four leading jets - Figs. 4.9, 4.10, 4.11, 4.12,
- Deep Jet b-tag discriminants, b-tag ranking: The discriminant distributions for the four highest tagged jets - Figs. 4.13, 4.14, 4.15, 4.16
- dR^{bb}: ΔR between the two leading b-tagged jets in the event - Fig. 4.17.
- H: Scalar sum of $|\vec{p_T}|$ of the selected jets - Fig. 4.18.

- H_T: Scalar sum of p_{T} of the selected jets - Fig. 4.19.

- H_T^{b}: H_T for the subset of selected jets passing the medium b tagging working point - Fig. 4.20.

- Muon p_{T}, η, PF Iso: The p_{T}, η, and total PF isolation distribution of the leading muon - Figs. 4.21, 4.22, 4.23

- Electron p_{T}, η, PF Iso: The p_{T}, η, and total PF isolation distribution of the leading electron - Fig. 4.24, 4.25, 4.26

- Subleading lepton p_{T}: Subleading muon and electron in the same-flavor channels - Fig. 4.27

- dR^{ll}: ΔR between the two leptons in the event - Fig. 4.28.

- Muon vertexing: The dz and 3D impact parameter of the muons for the e μ channel are shown in Fig. 4.29 and the distributions for the μ μ channel are shown in Fig. 4.30.

- Electron vertexing: The dz and 3D impact parameter of the electrons for the e μ channel are shown in Fig. 4.31 and the distributions for the e e channel are shown in Fig. 4.32.

- $M_T(l, MET)$: The transverse mass distributions for the ee and μ μ channels are shown in Fig. 4.33 and the distributions for the e μ channel are shown in Fig. 4.34, where each possible combination of lepton and p_{T}^{Miss} is calculated.

- Missing transverse momentum (p_{T}^{Miss}): The missing transverse momentum magnitude is shown in Fig. 4.35 and the phi distribution is in Fig. 4.36.

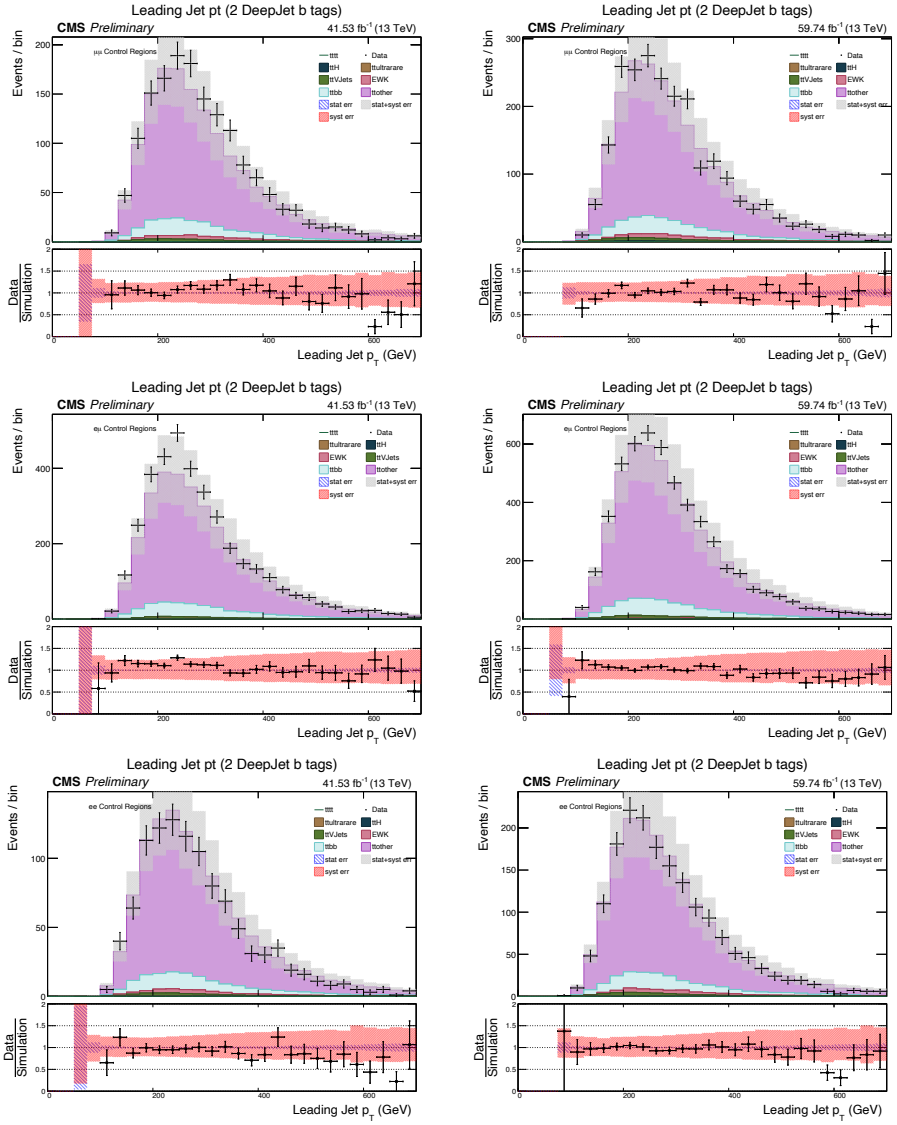

Figure 4.1: Leading Jet p_{T} for the 2017 and 2018 datasets for the the e e channels. The left panels are for 2017 and the right panels are for 2018. The μ μ channel is shown in the top row, the e μ channel is in the middle row and the e e channel is in the bottom row.

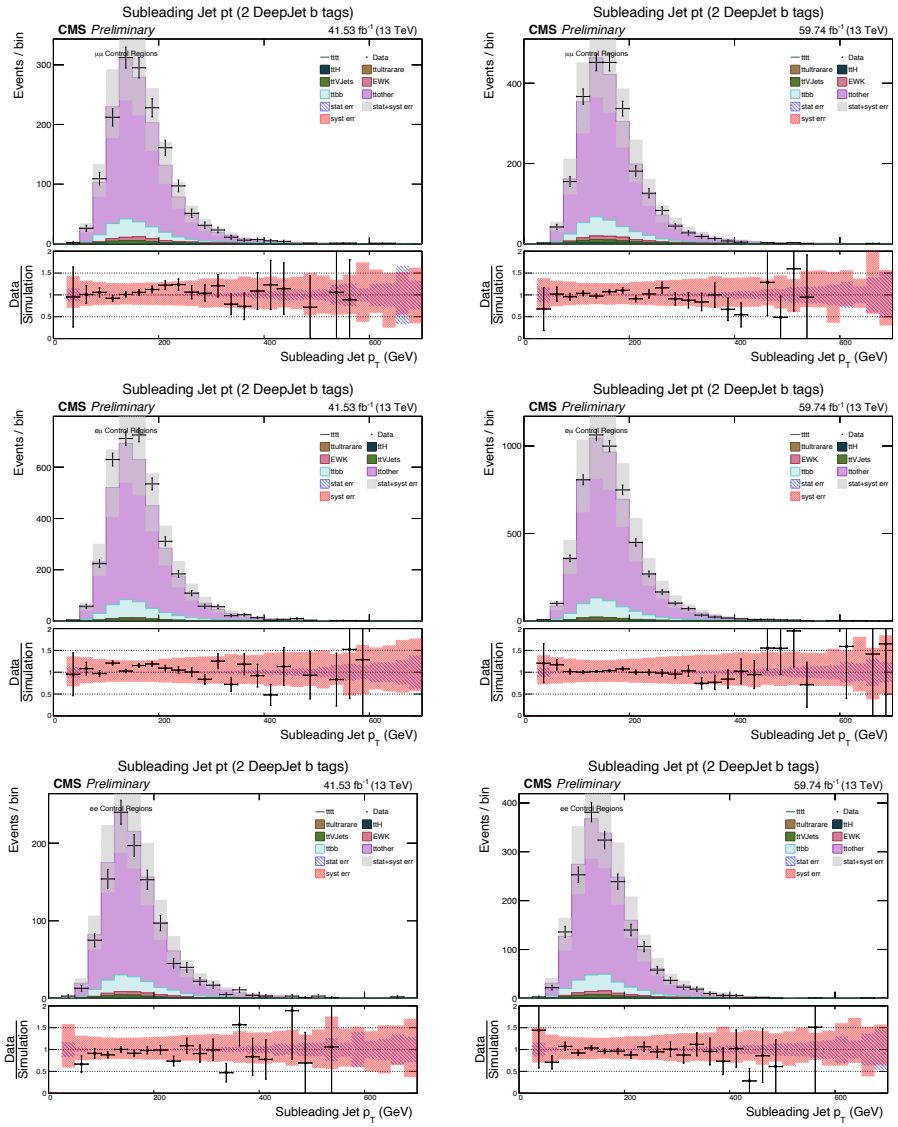

Figure 4.2: Sub-leading Jet p_{T} for the 2017 and 2018 datasets for the three channels. The left panels are for 2017 and the right panels are for 2018. The μ μ channel is shown in the top row, the e μ channel is in the middle row and the e e channel is in the bottom row.

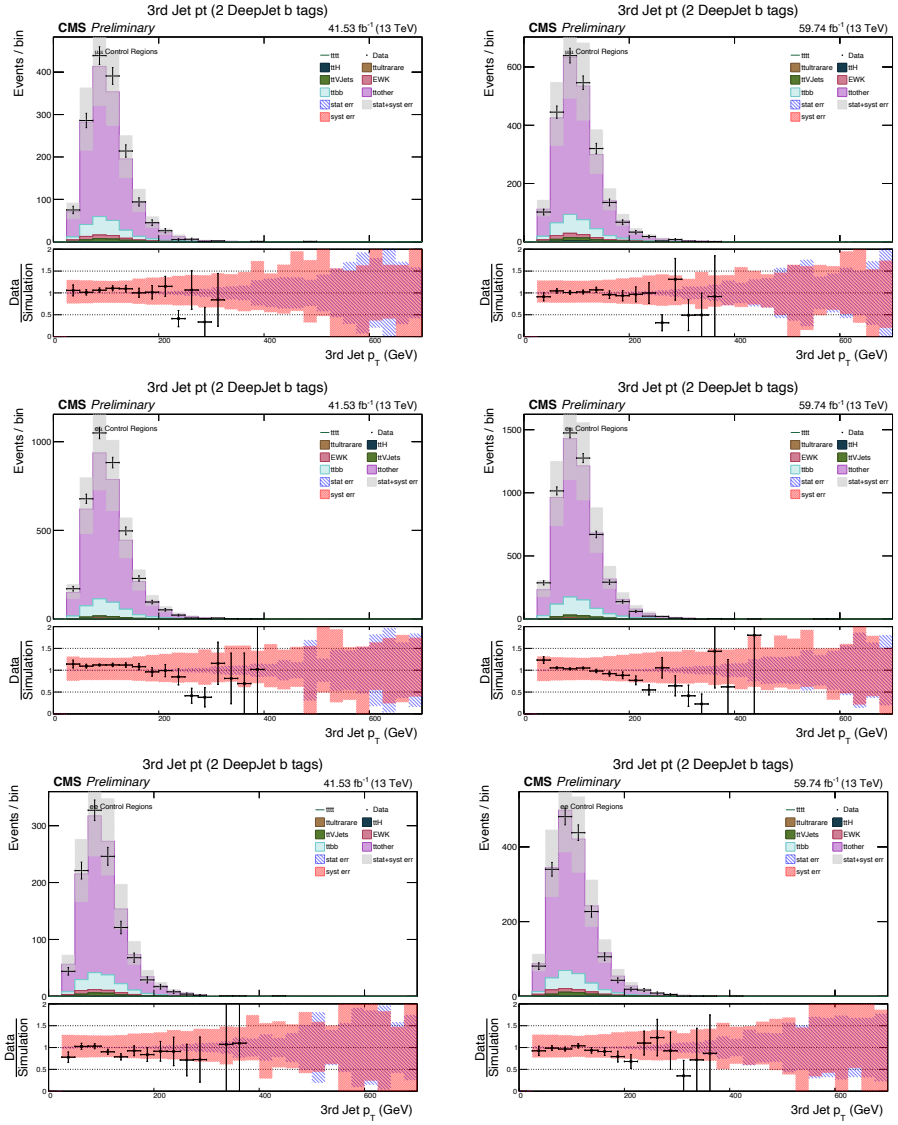

Figure 4.3: p_{T} of Jet 3 for the 2017 and 2018 datasets for the three channels. The left panels are for 2017 and the right panels are for 2018. The μ μ channel is shown in the top row, the e μ channel is in the middle row and the e e channel is in the bottom row.

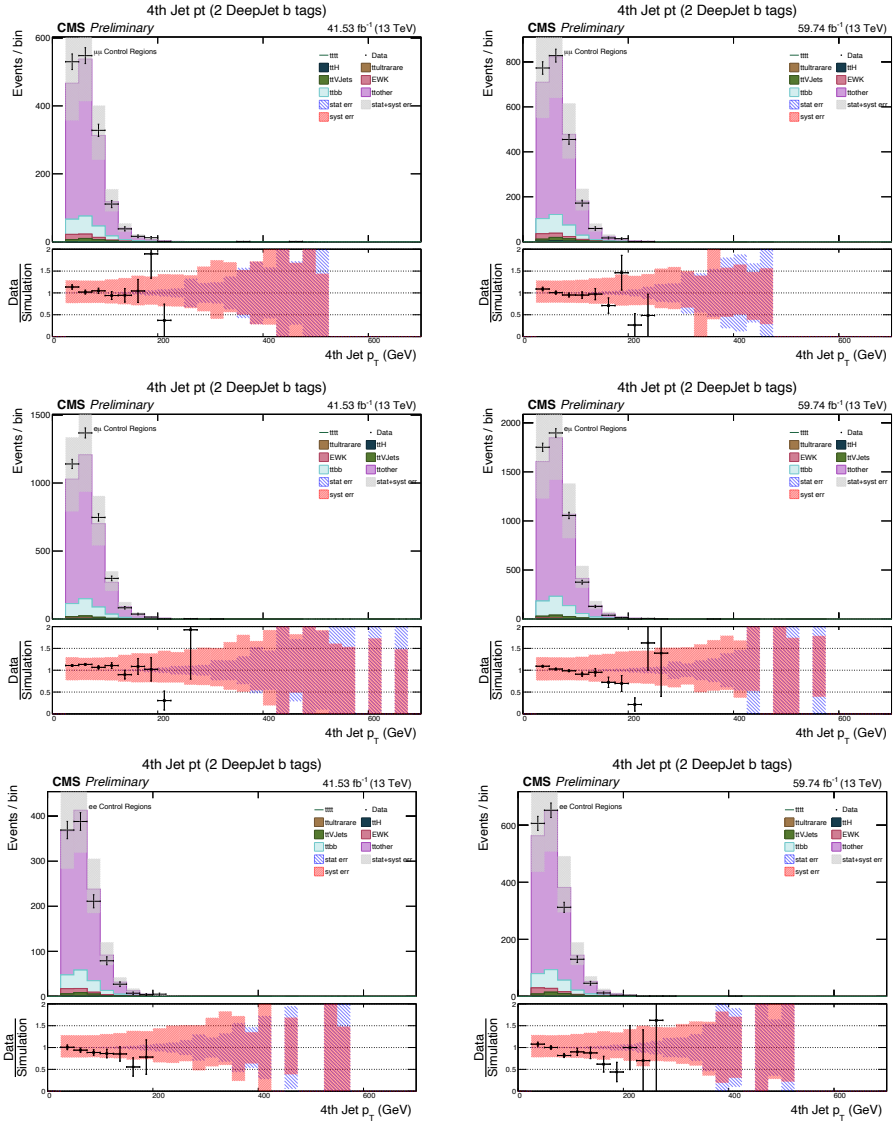

Figure 4.4: p_{T} of Jet 4 for the 2017 and 2018 datasets for the three channels. The left panels are for 2017 and the right panels are for 2018. The μ μ channel is shown in the top row, the e μ channel is in the middle row and the e e channel is in the bottom row.

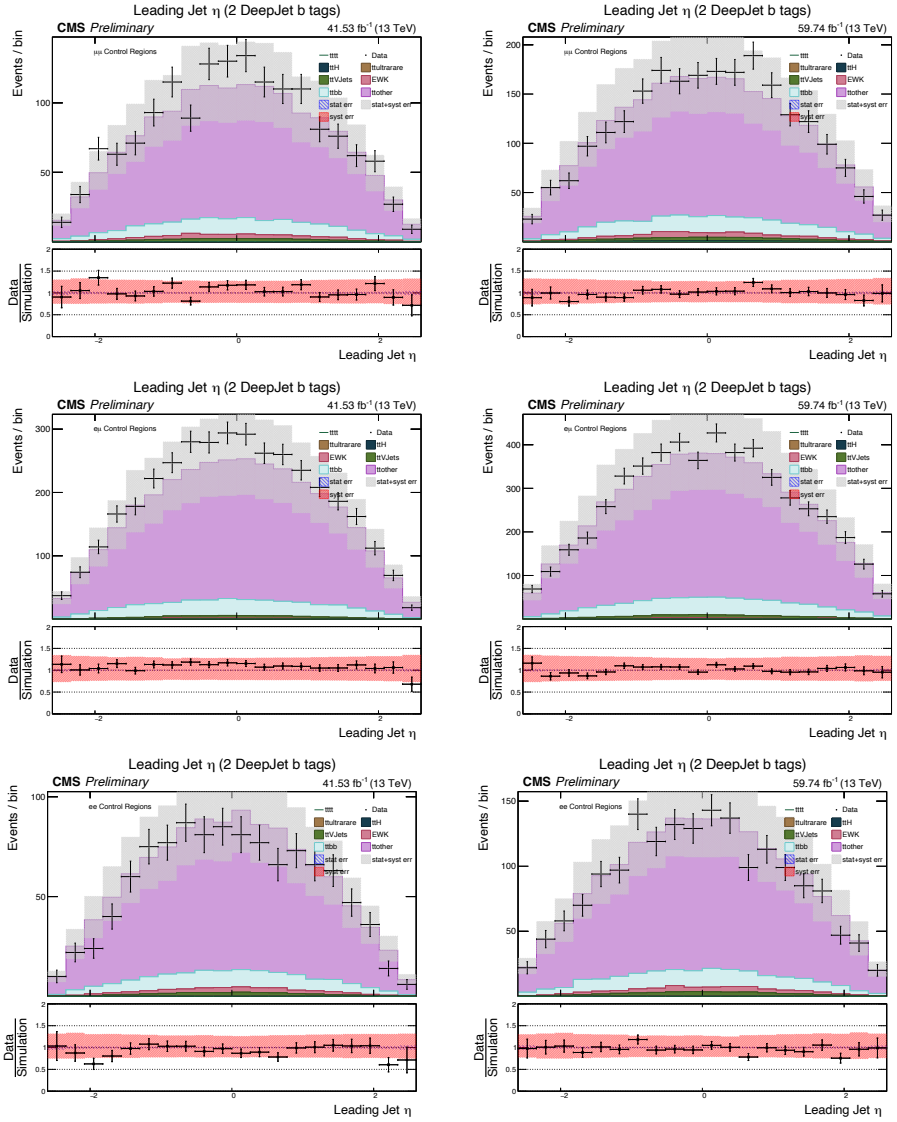

Figure 4.5: Leading Jet η for the 2017 and 2018 datasets for the three channels. The left panels are for 2017 and the right panels are for 2018. The μ μ channel is shown in the top row, the e μ channel is in the middle row and the e e channel is in the bottom row.

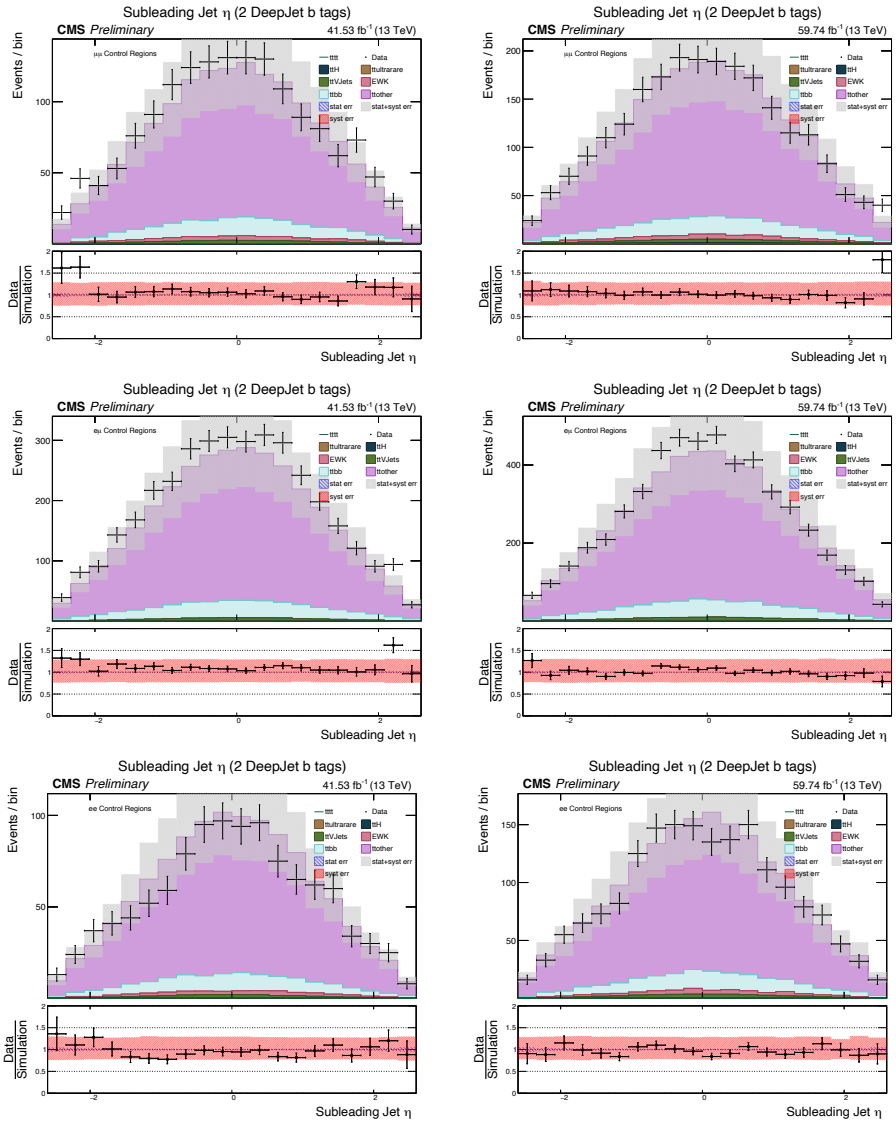

Figure 4.6: Sub-leading Jet η for the 2017 and 2018 datasets for the three channels. The left panels are for 2017 and the right panels are for 2018. The μ μ channel is shown in the top row, the e μ channel is in the middle row and the e e channel is in the bottom row.

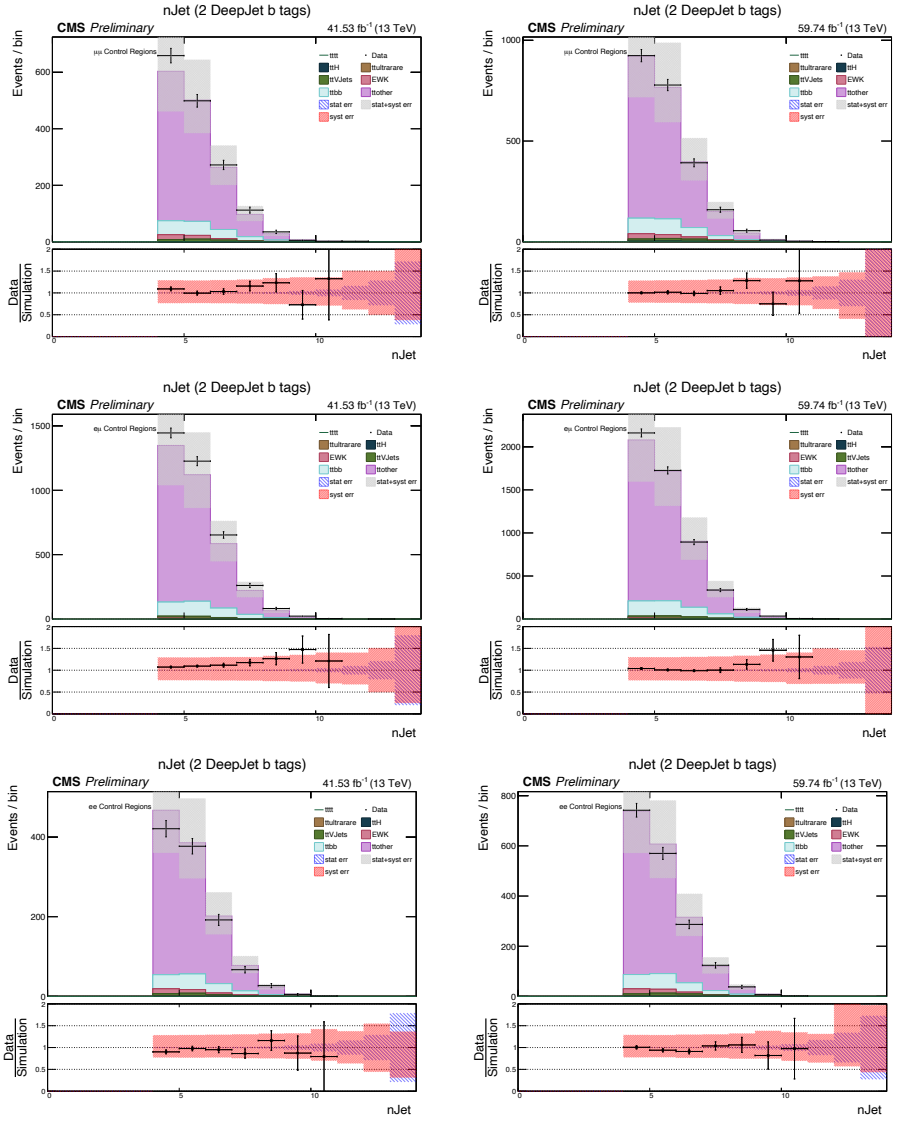

Figure 4.7: Jet multiplicity for the 2017 and 2018 datasets for the three channels. The left panels are for 2017 and the right panels are for 2018. The μ μ channel is shown in the top row, the e μ channel is in the middle row and the e e channel is in the bottom row.

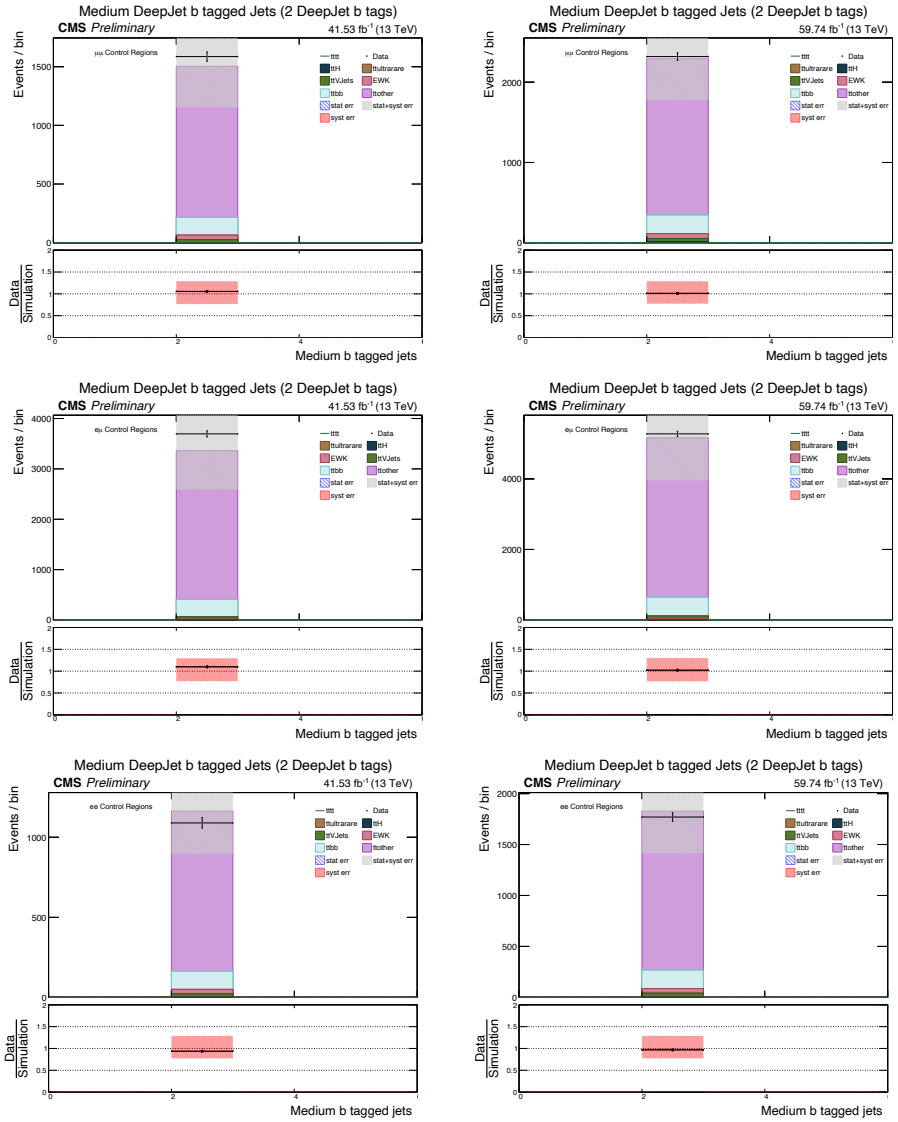

Figure 4.8: Medium b tagged jet multiplicity for the 2017 and 2018 datasets for the three channels. The left panels are for 2017 and the right panels are for 2018. The μ μ channel is shown in the top row, the e μ channel is in the middle row and the e e channel is in the bottom row.

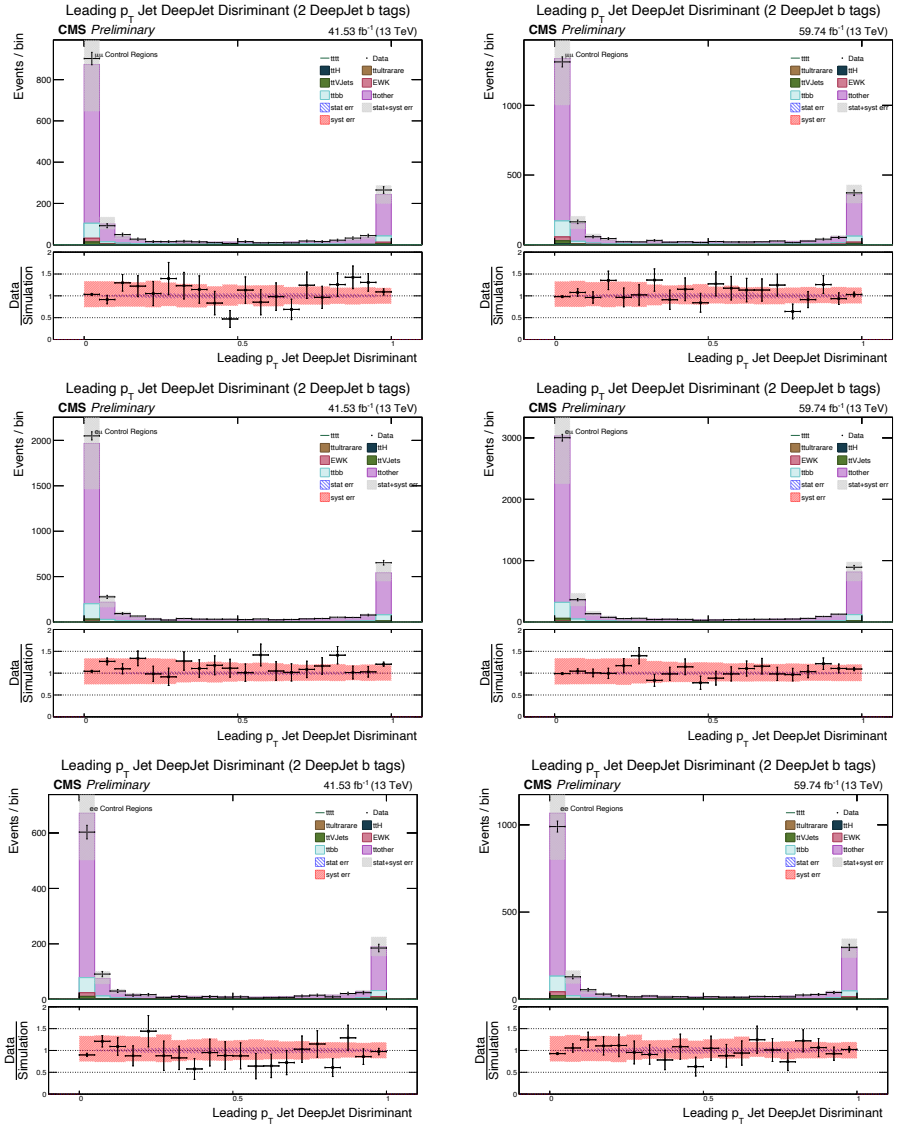

Figure 4.9: Leading Jet b-tag disciminant for the 2017 and 2018 datasets for the three channels. The left panels are for 2017 and the right panels are for 2018. The μ μ channel is shown in the top row, the e μ channel is in the middle row and the e e channel is in the bottom row.

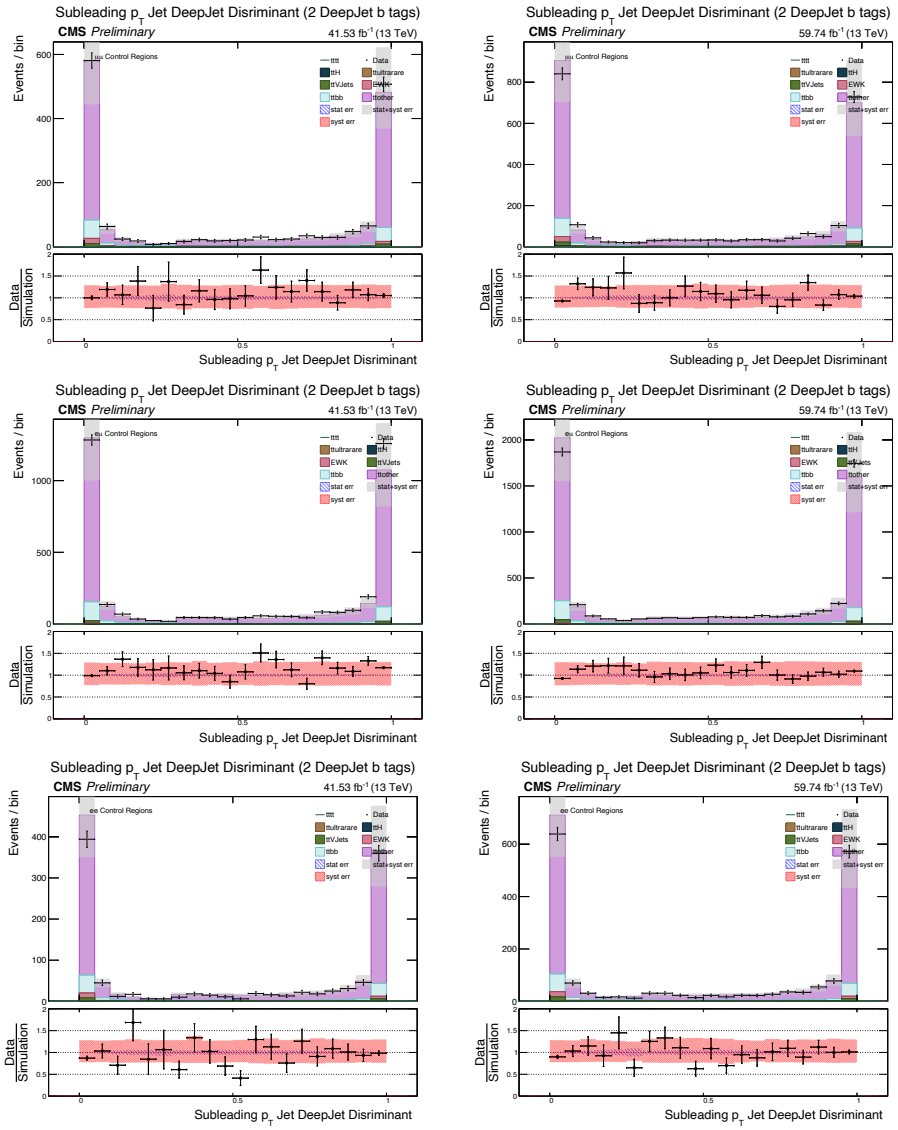

Figure 4.10: Sub-leading Jet b-tag disciminant for the 2017 and 2018 datasets for the three channels. The left panels are for 2017 and the right panels are for 2018. The μ μ channel is shown in the top row, the e μ channel is in the middle row and the e e channel is in the bottom row.

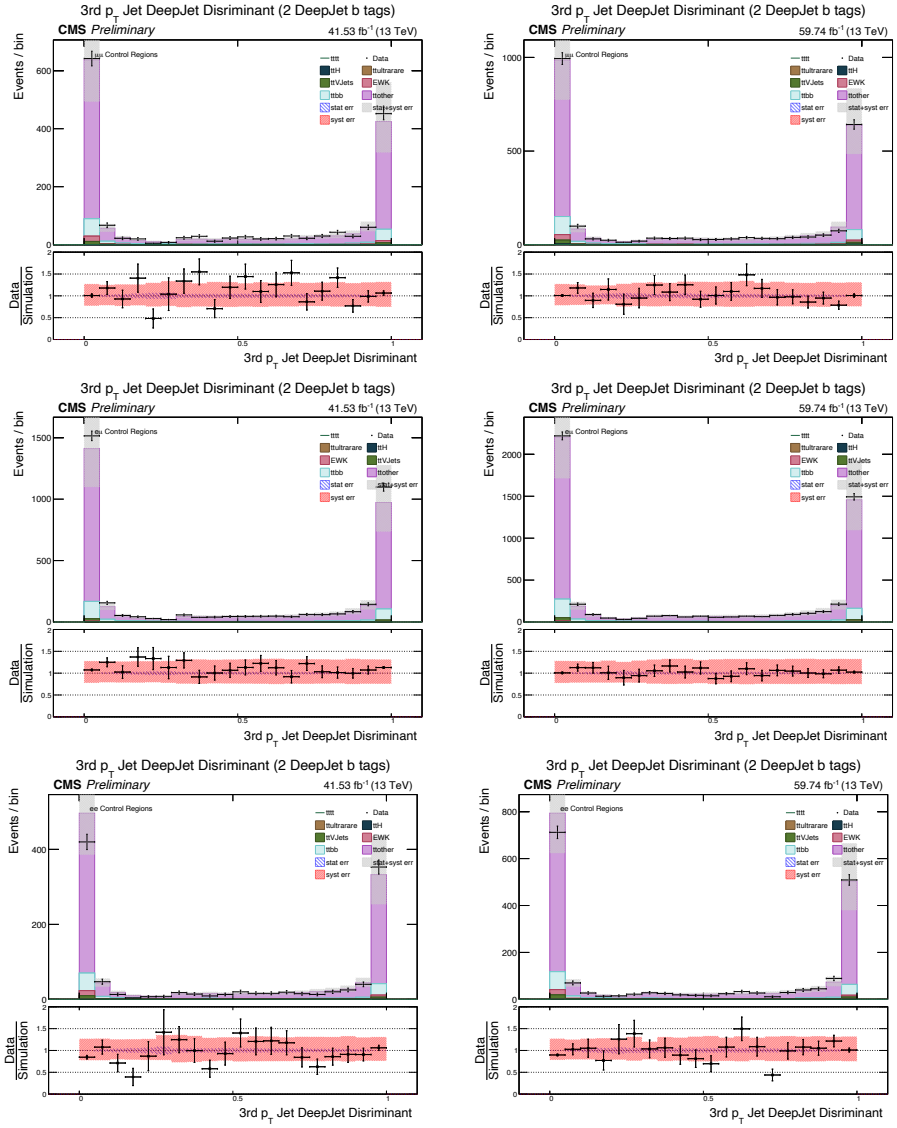

Figure 4.11: Third highest p_{T} jet b-tag disciminant for the 2017 and 2018 datasets for the three channels. The left panels are for 2017 and the right panels are for 2018. The μ μ channel is shown in the top row, the e μ channel is in the middle row and the e e channel is in the bottom row.

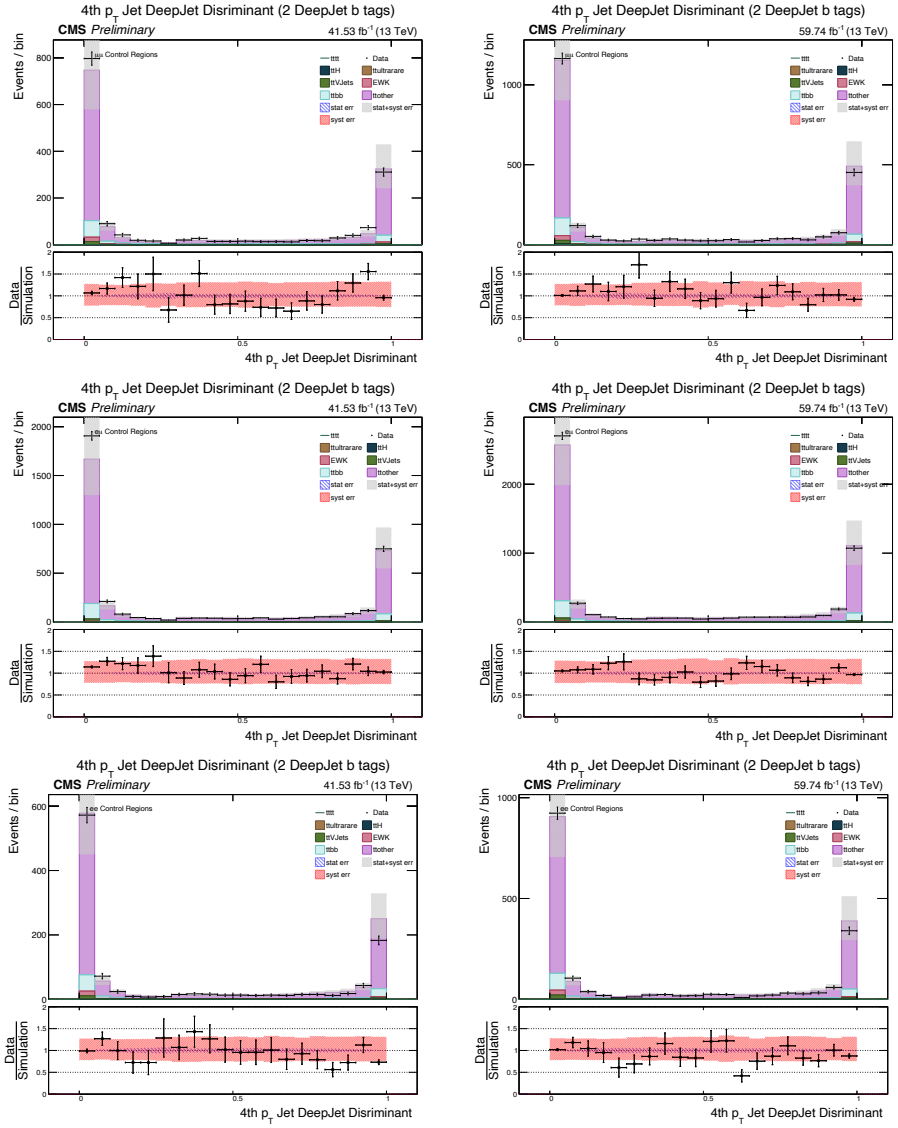

Figure 4.12: Fourth highest p_{T} jet b-tag disciminant for the 2017 and 2018 datasets for the three channels. The left panels are for 2017 and the right panels are for 2018. The μ μ channel is shown in the top row, the e μ channel is in the middle row and the e e channel is in the bottom row.

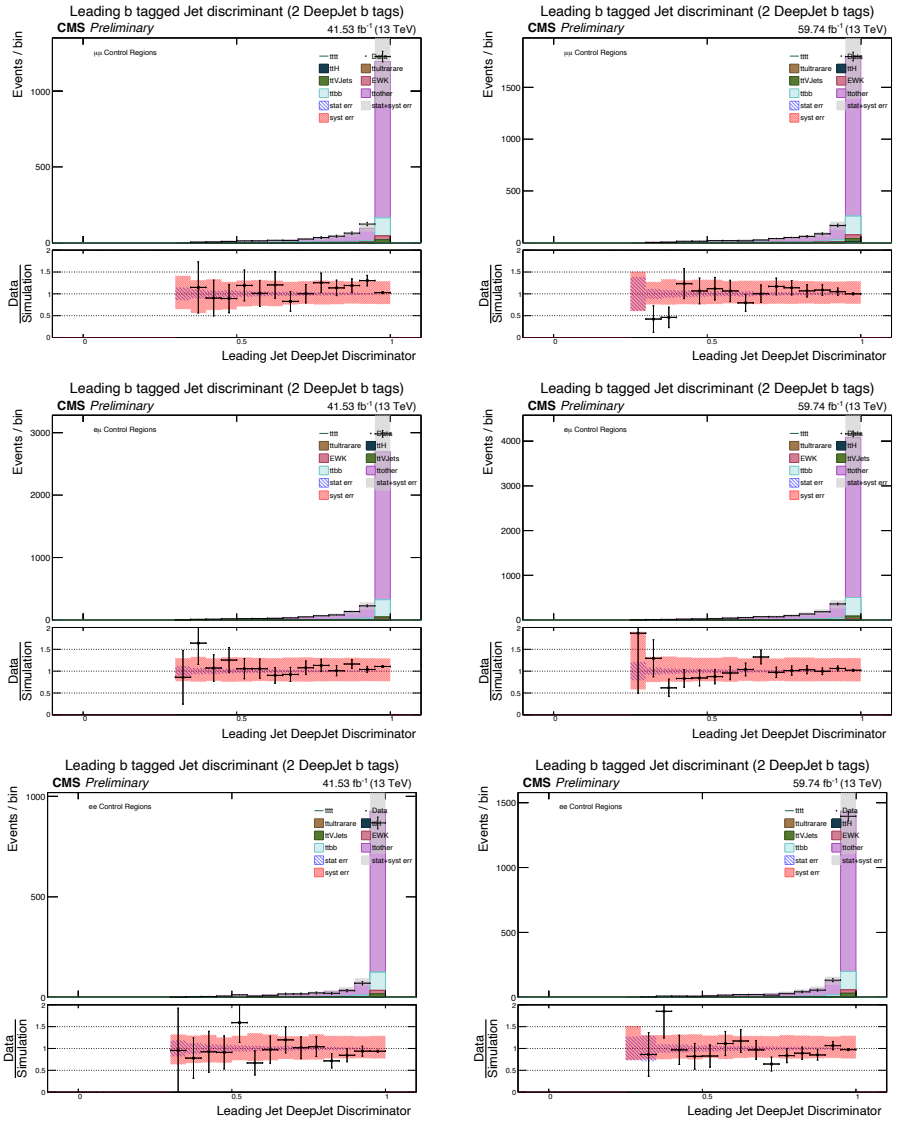

Figure 4.13: Highest b tagged jet disciminant for the 2017 and 2018 datasets for the three channels. The left panels are for 2017 and the right panels are for 2018. The μ μ channel is shown in the top row, the e μ channel is in the middle row and the e e channel is in the bottom row.

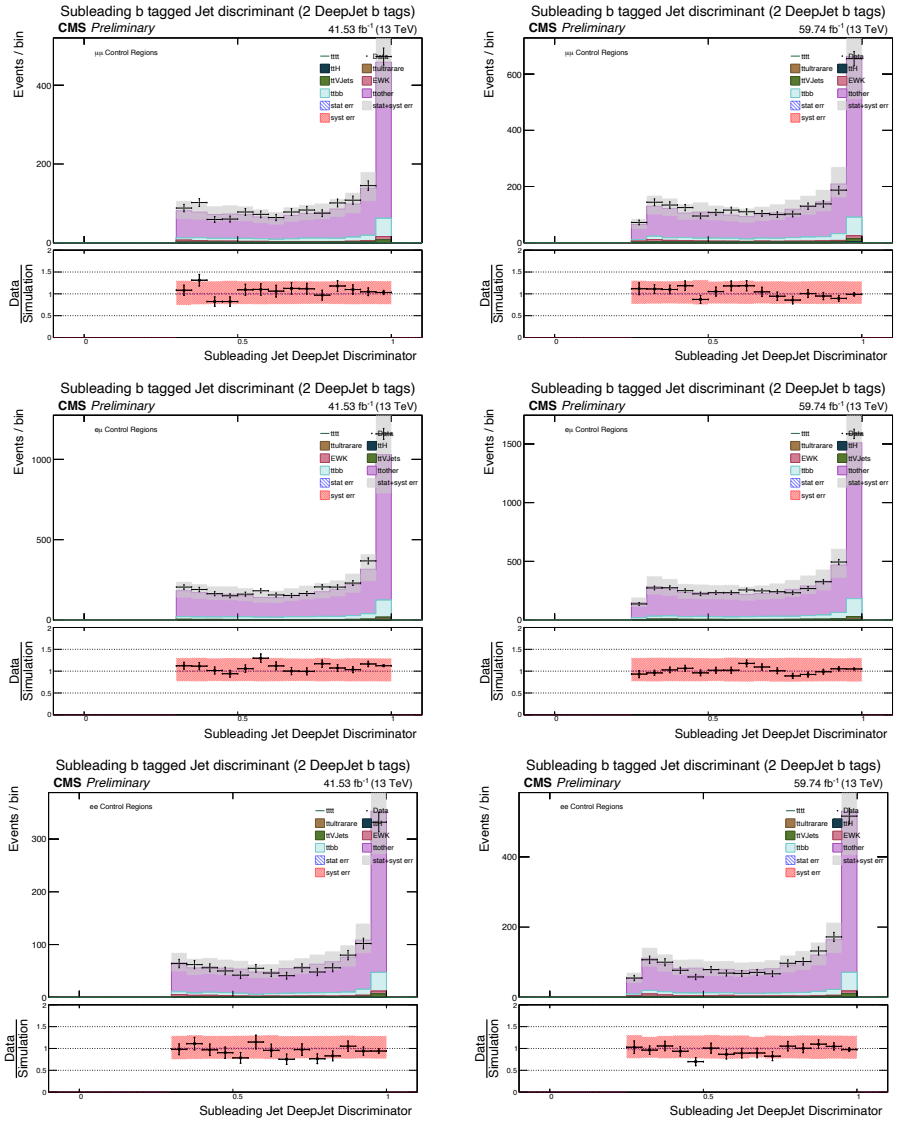

Figure 4.14: Second highest b tagged jet disciminant for the 2017 and 2018 datasets for the three channels. The left panels are for 2017 and the right panels are for 2018. The μ μ channel is shown in the top row, the e μ channel is in the middle row and the e e channel is in the bottom row.

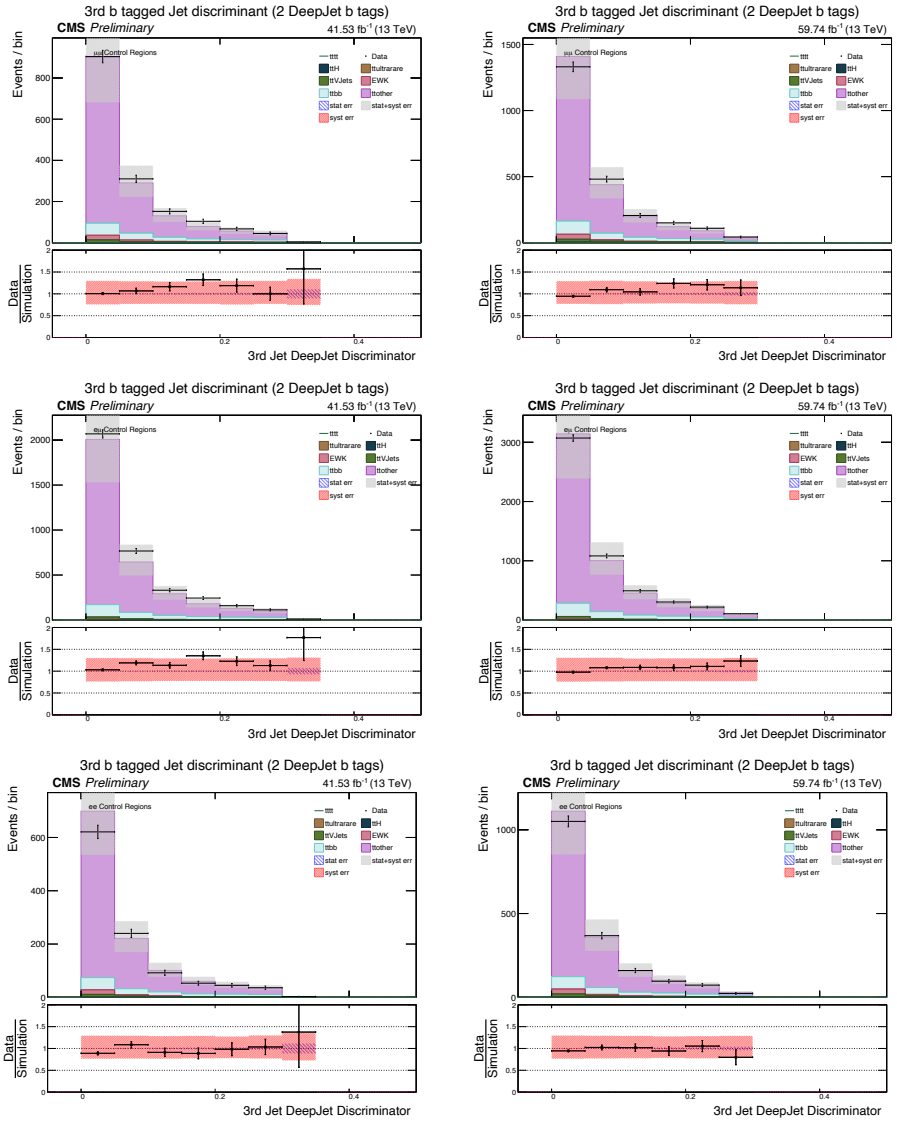

Figure 4.15: Third highest b tagged jet disciminant for the 2017 and 2018 datasets for the three channels, covering only the range of jets untagged at the medium working point. The left panels are for 2017 and the right panels are for 2018. The μ μ channel is shown in the top row, the e μ channel is in the middle row and the e e channel is in the bottom row.

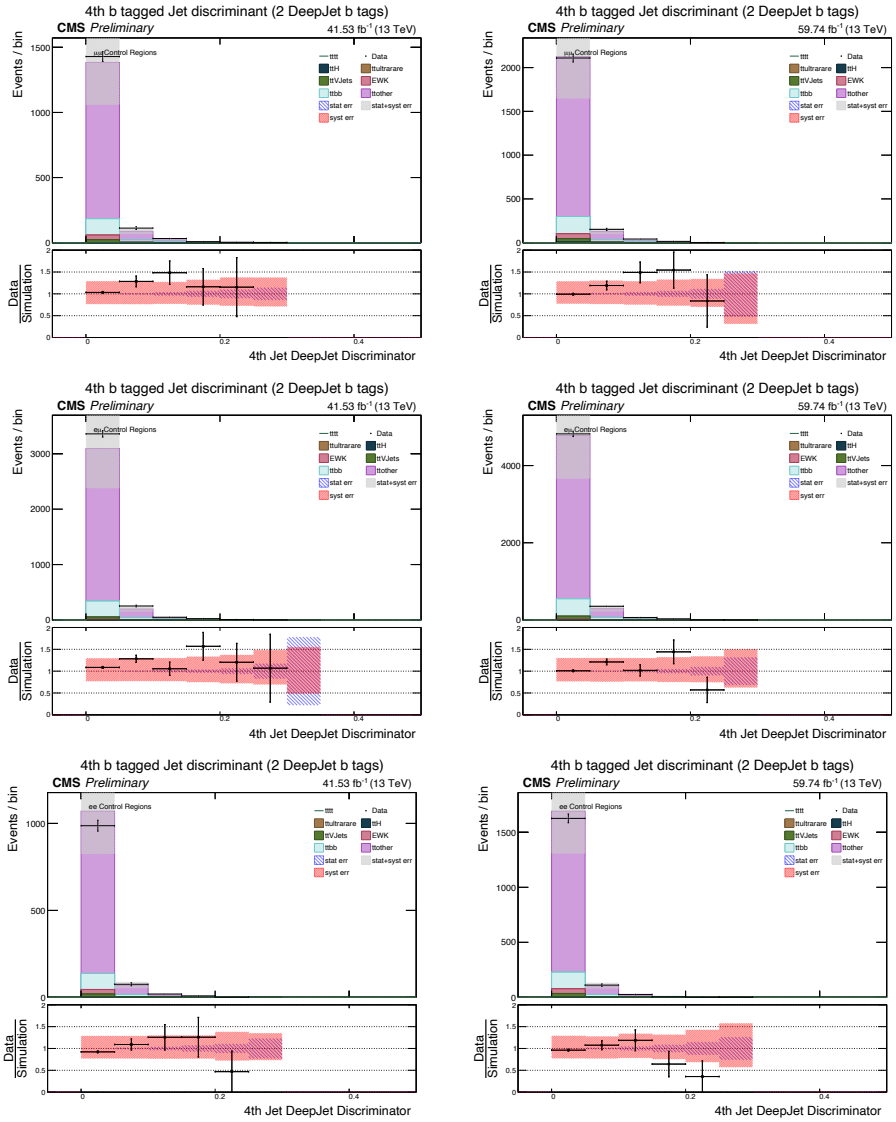

Figure 4.16: Fourth highest b tagged jet disciminant for the 2017 and 2018 datasets for the three channels, covering only the range of jets untagged at the medium working point. The left panels are for 2017 and the right panels are for 2018. The μ μ channel is shown in the top row, the e μ channel is in the middle row and the e e channel is in the bottom row.

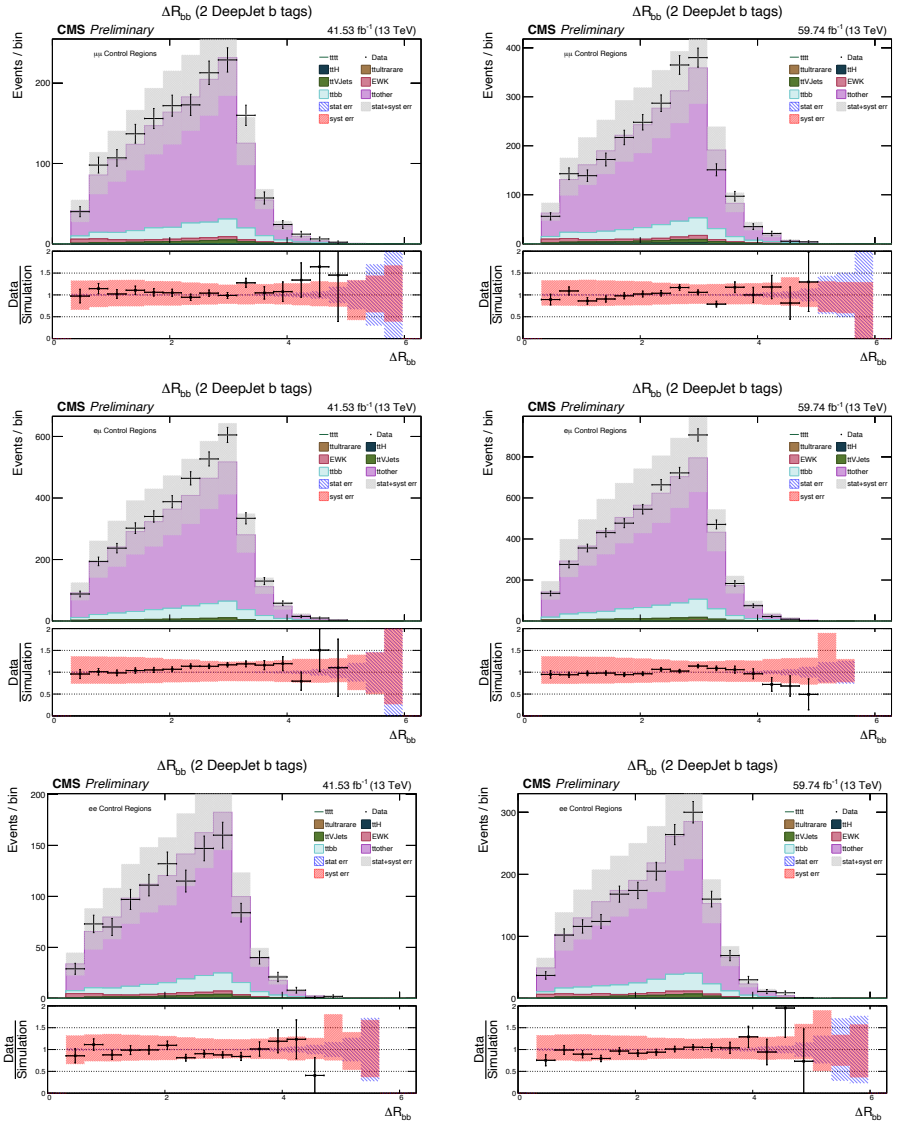

Figure 4.17: dR^{bb} for the 2017 and 2018 datasets for the three channels. The left panels are for 2017 and the right panels are for 2018. The μ μ channel is shown in the top row, the e μ channel is in the middle row and the e e channel is in the bottom row.

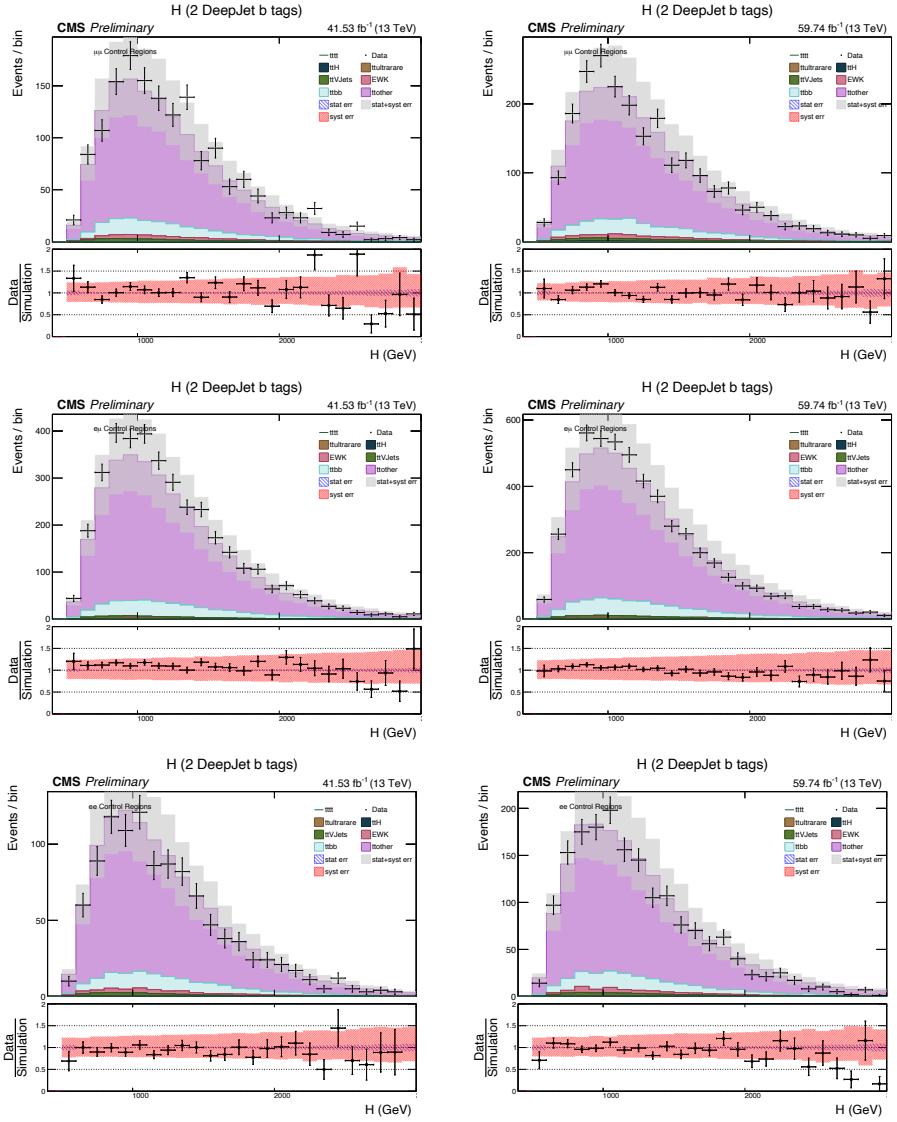

Figure 4.18: H for the 2017 and 2018 datasets for the three channels. The left panels are for 2017 and the right panels are for 2018. The μ μ channel is shown in the top row, the e μ channel is in the middle row and the e e channel is in the bottom row.

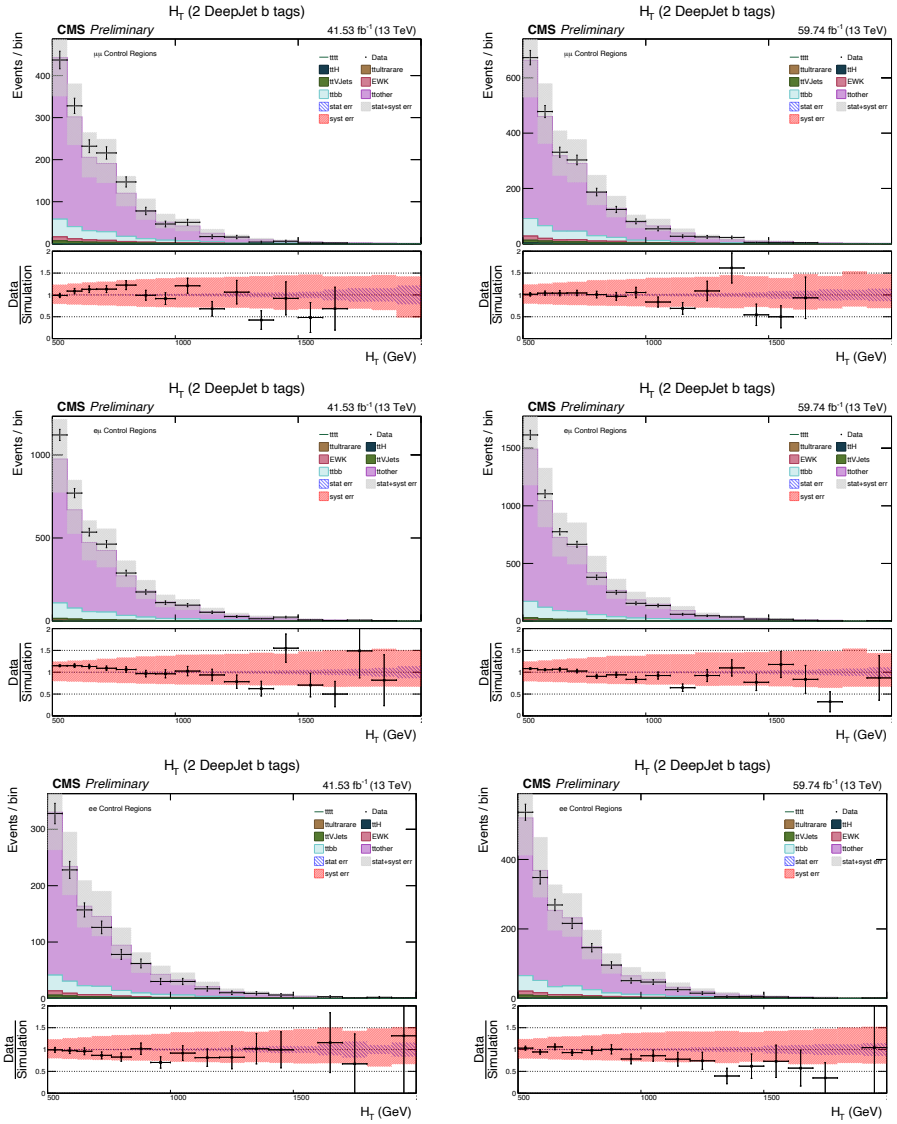

Figure 4.19: H_T for the 2017 and 2018 datasets for the three channels. The left panels are for 2017 and the right panels are for 2018. The μ μ channel is shown in the top row, the e μ channel is in the middle row and the e e channel is in the bottom row.

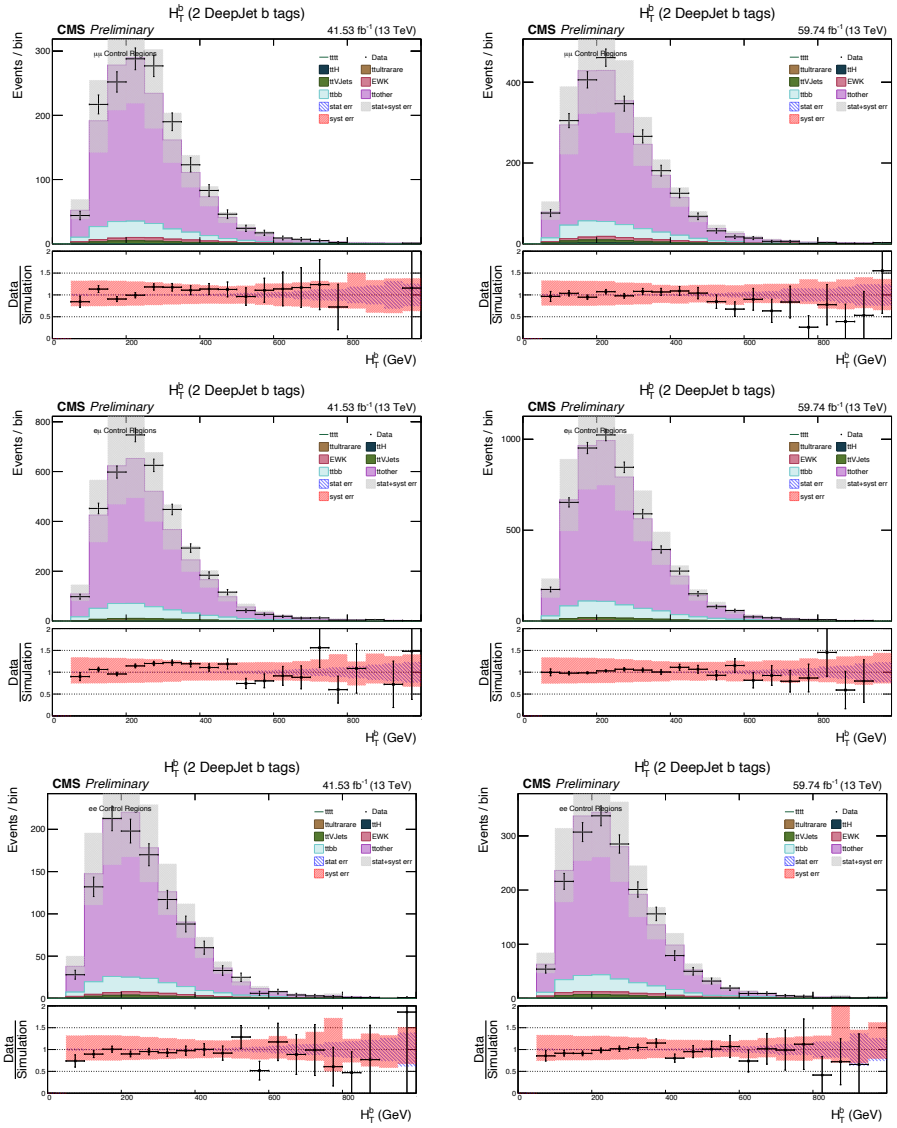

Figure 4.20: H_T^{b} for the 2017 and 2018 datasets for the three channels. The left panels are for 2017 and the right panels are for 2018. The μ μ channel is shown in the top row, the e μ channel is in the middle row and the e e channel is in the bottom row.

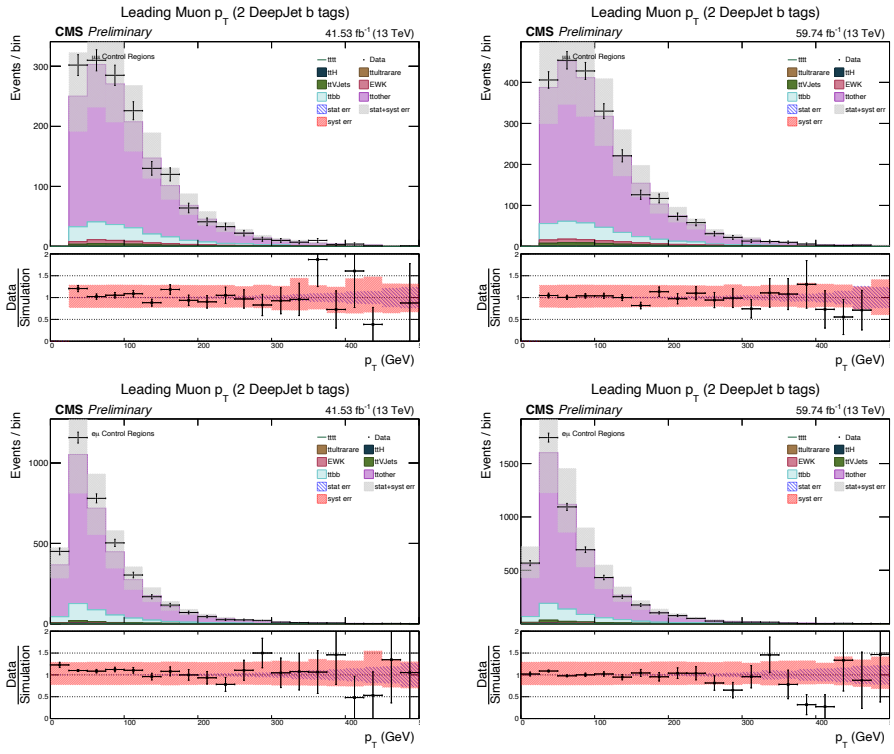

Figure 4.21: Leading muon p_{T} for the 2017 and 2018 datasets. The left panels are for 2017 and the right panels are for 2018. The μ μ channel is shown in the top row, the e μ channel is in the bottom row

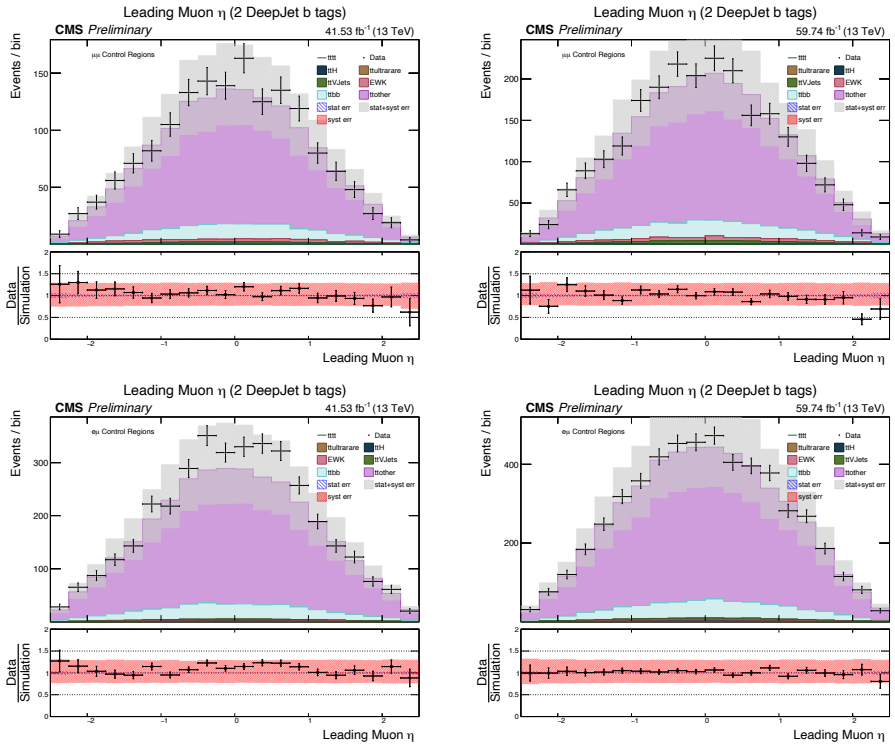

Figure 4.22: Leading muon η for the 2017 and 2018 datasets. The left panels are for 2017 and the right panels are for 2018. The μ μ channel is shown in the top row, the e μ channel is in the bottom row

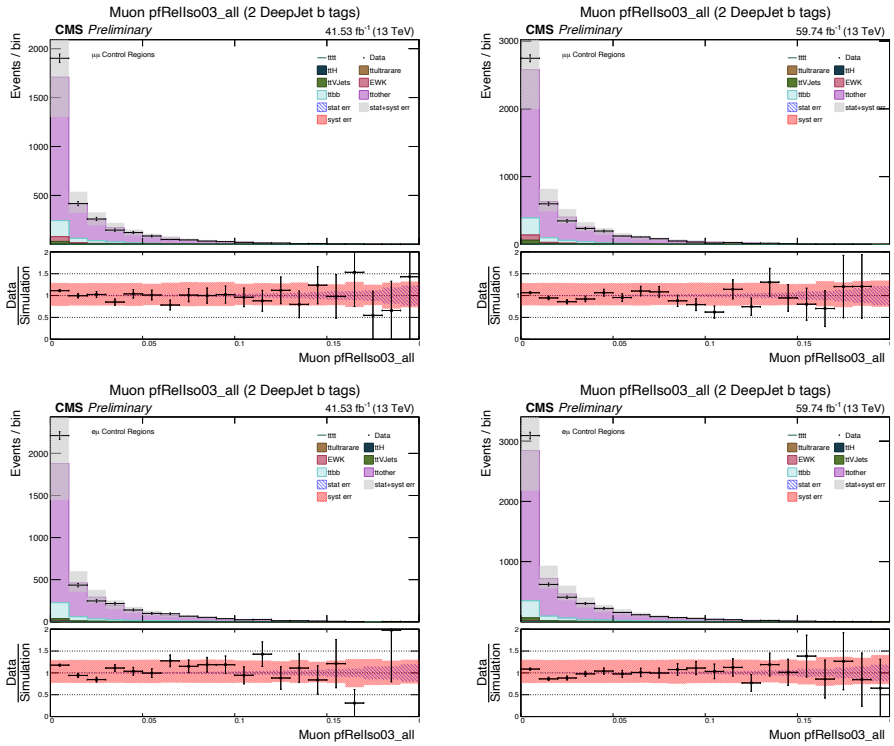

Figure 4.23: Muon PF Relative Isolation (R=0.03) for the 2017 and 2018 datasets. The left panels are for 2017 and the right panels are for 2018. The muons of the μ μ channel are shown in the top row, the muons of the e μ channel are in the bottom row

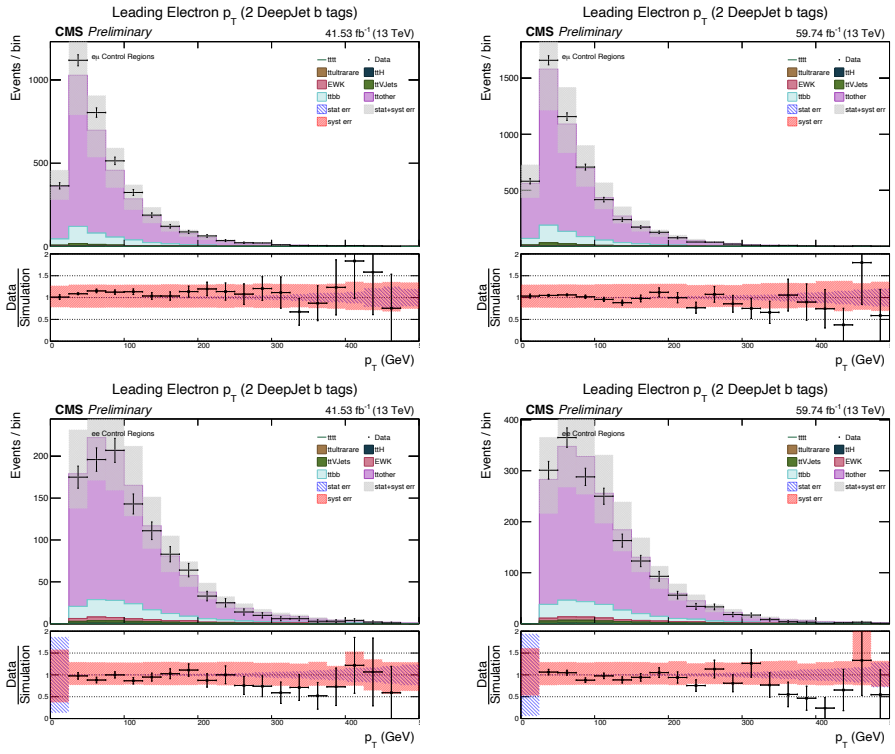

Figure 4.24: Leading electron p_{T} for the 2017 and 2018 datasets. The left panels are for 2017 and the right panels are for 2018. The e μ channel is shown in the top row, the e e channel is in the bottom row

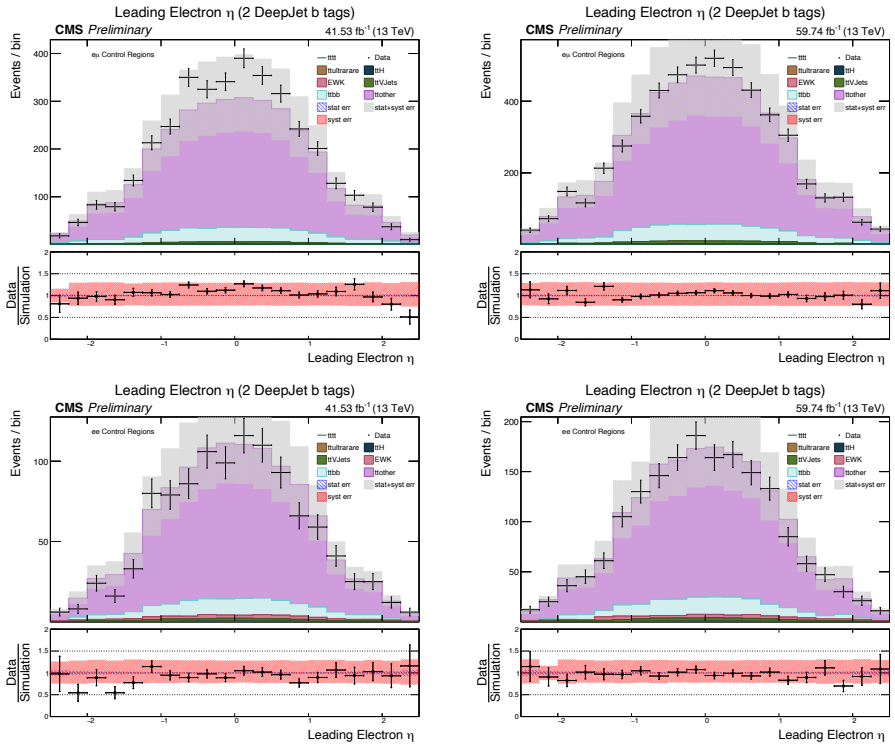

Figure 4.25: Leading electron η for the 2017 and 2018 datasets. The left panels are for 2017 and the right panels are for 2018. The e μ channel is shown in the top row, the e e channel is in the bottom row

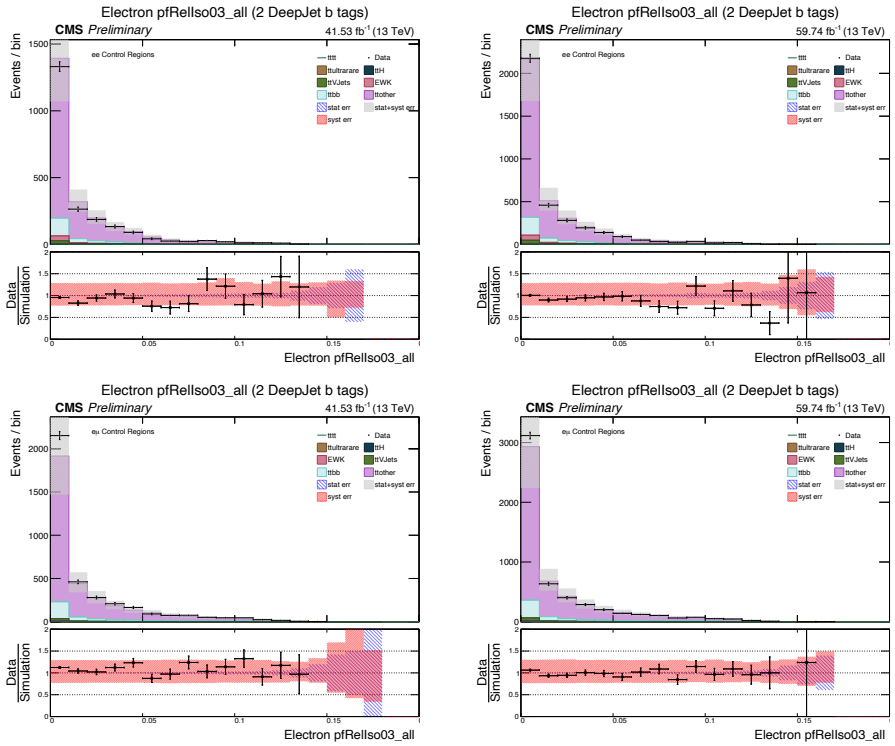

Figure 4.26: Electron PF Relative Isolation (R=0.03) for the 2017 and 2018 datasets. The left panels are for 2017 and the right panels are for 2018. The electrons of the e e channel are shown in the top row, the electrons of the e μ channel are in the bottom row

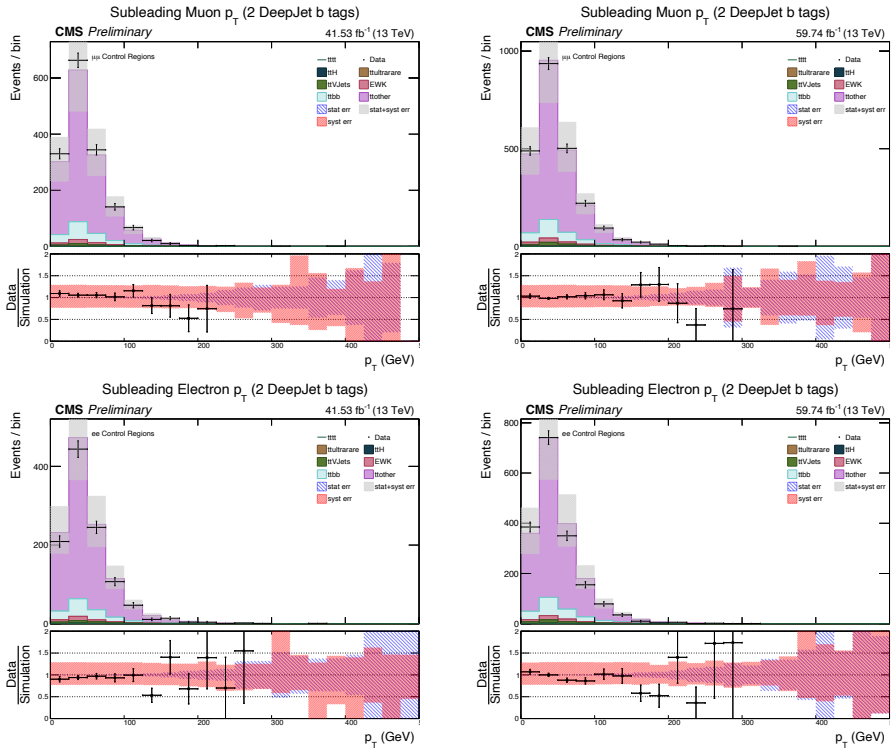

Figure 4.27: Subleading lepton p_{T} for the 2017 and 2018 datasets. The left panels are for 2017 and the right panels are for 2018. The muons of the μ μ channel are shown in the top row, the electrons of the e μ channel are in the bottom row

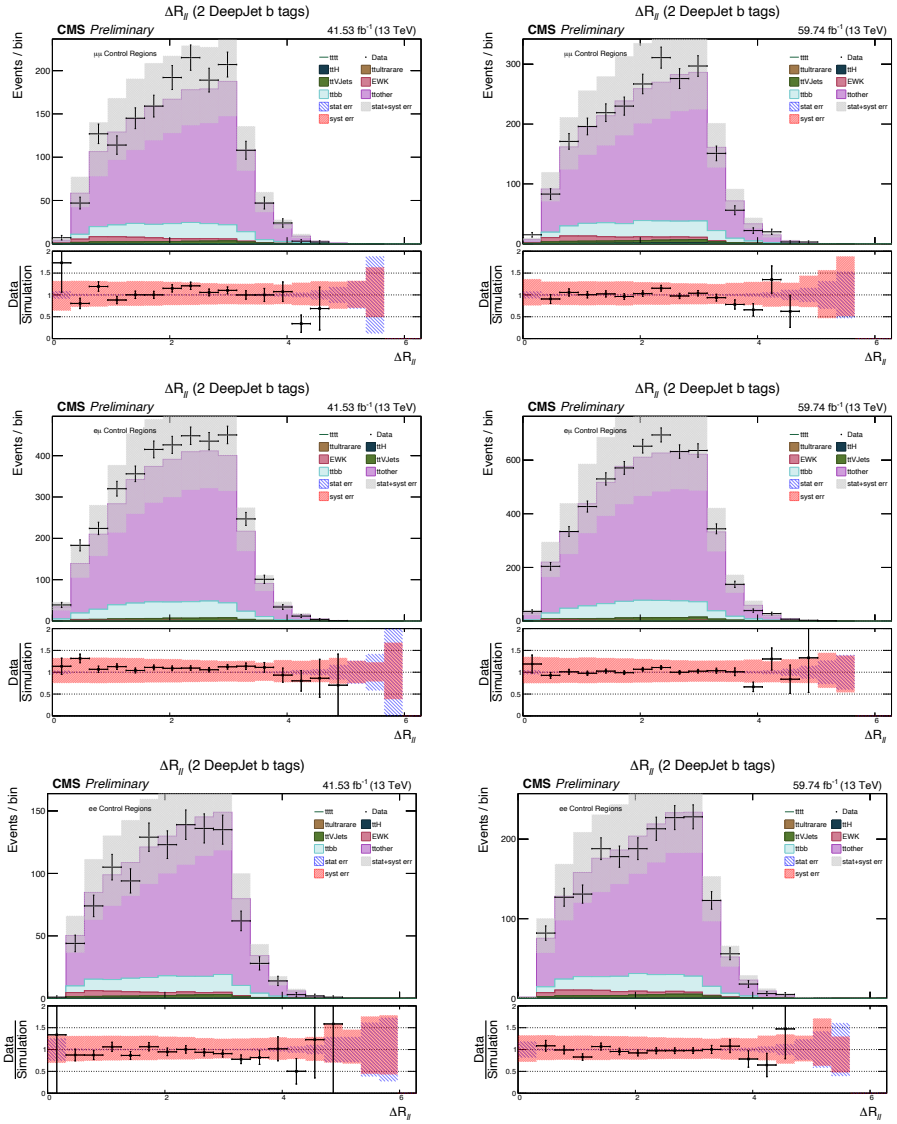

Figure 4.28: dR^{bb} for the 2017 and 2018 datasets for the three channels. The left panels are for 2017 and the right panels are for 2018. The μ μ channel is shown in the top row, the e μ channel is in the middle row and the e e channel is in the bottom row.

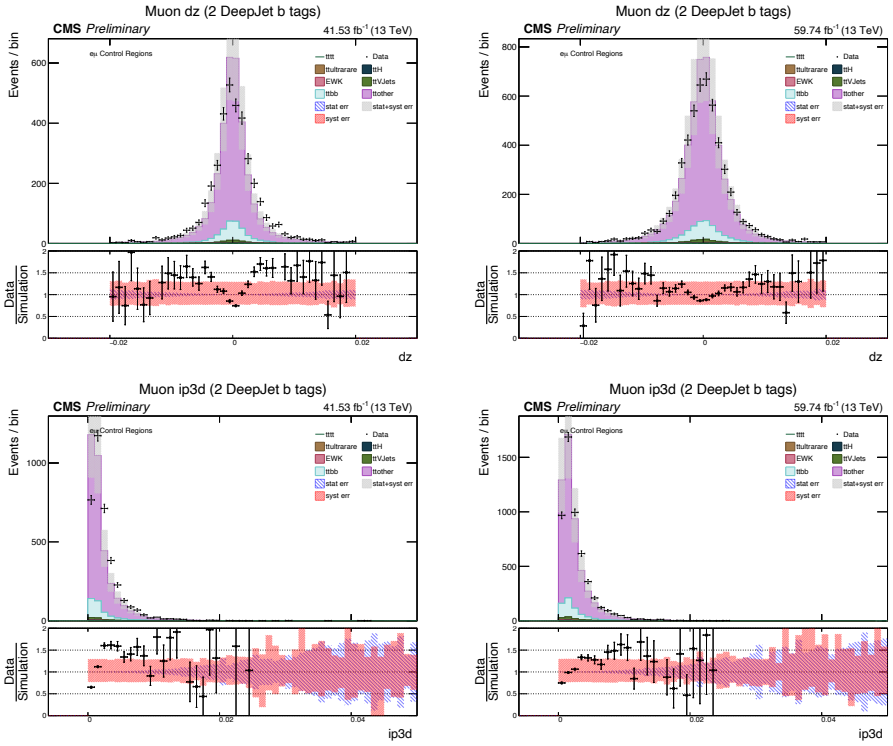

Figure 4.29: e μ muon dz and 3D impact parameter distributions, with 2017 on the left and 2018 on the right, dz on top and IP3D on bottom

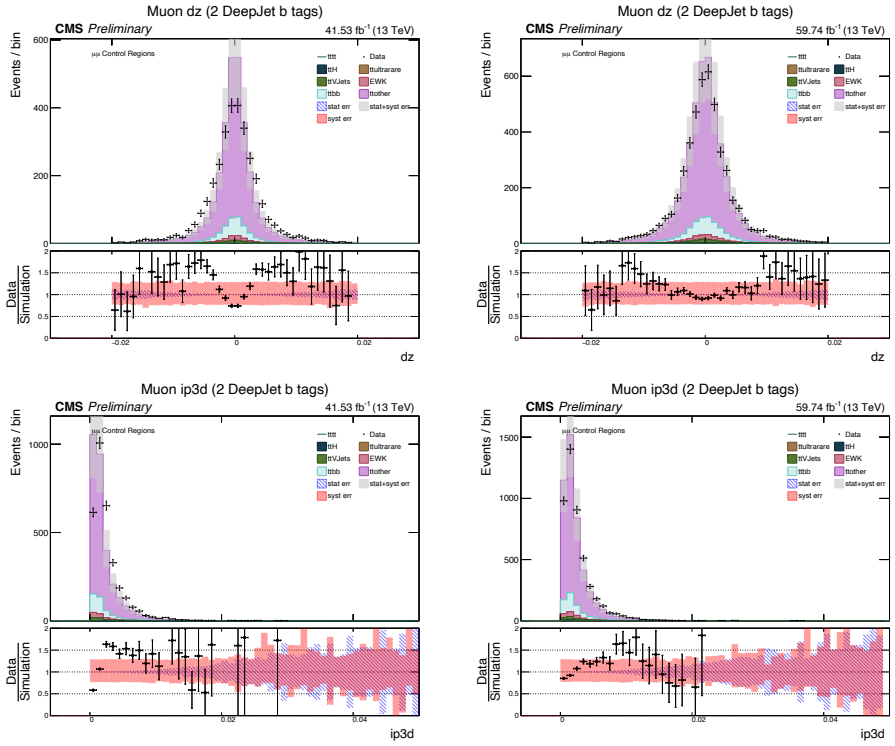

Figure 4.30: μ μ muon dz and 3D impact parameter distributions, with 2017 on the left and 2018 on the right, dz on top and IP3D on bottom

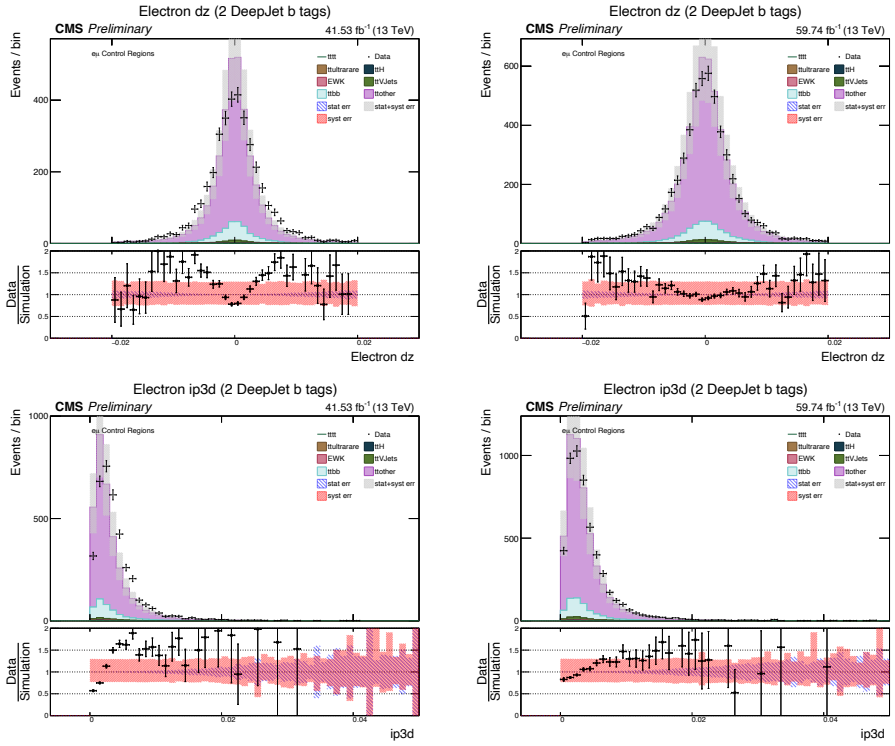

Figure 4.31: e μ electron dz and 3D impact parameter distributions, with 2017 on the left and 2018 on the right, dz on top and IP3D on bottom

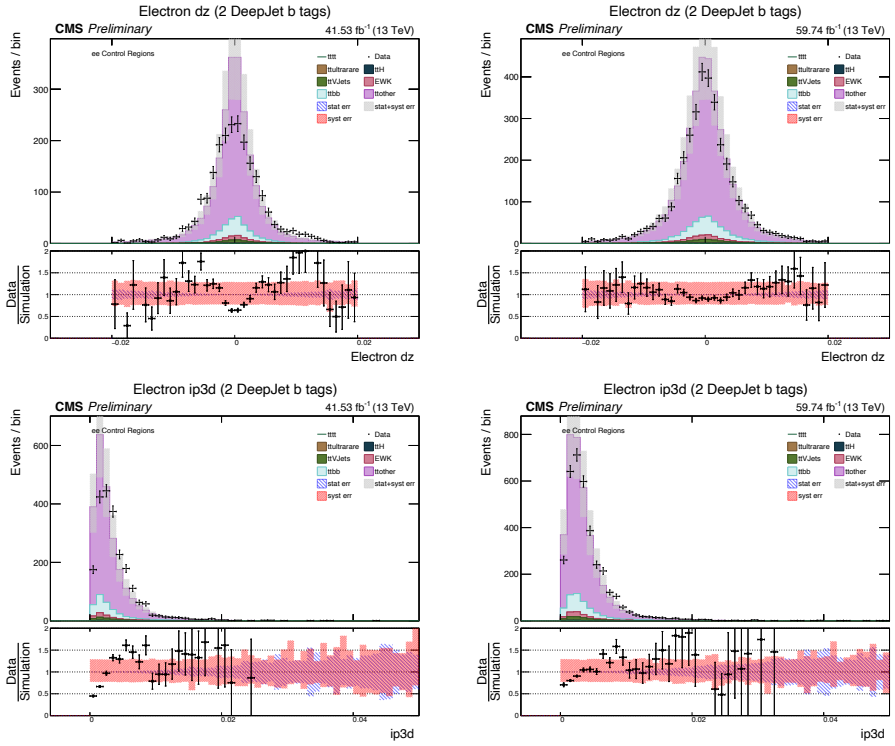

Figure 4.32: ee electron dz and 3D impact parameter distributions, with 2017 on the left and 2018 on the right, dz on top and IP3D on bottom

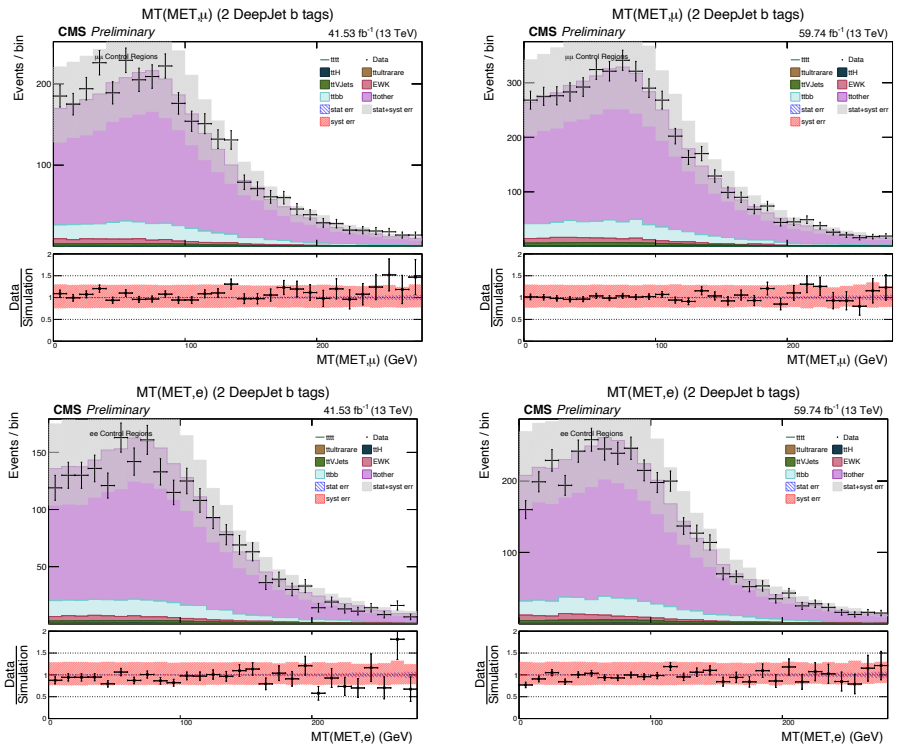

Figure 4.33: M_T for the 2017 and 2018 datasets for the ee and μ μ channels. The left panels are for 2017 and the right panels are for 2018. The μ μ channel is shown in the top row, and the e e channel is in the bottom row.

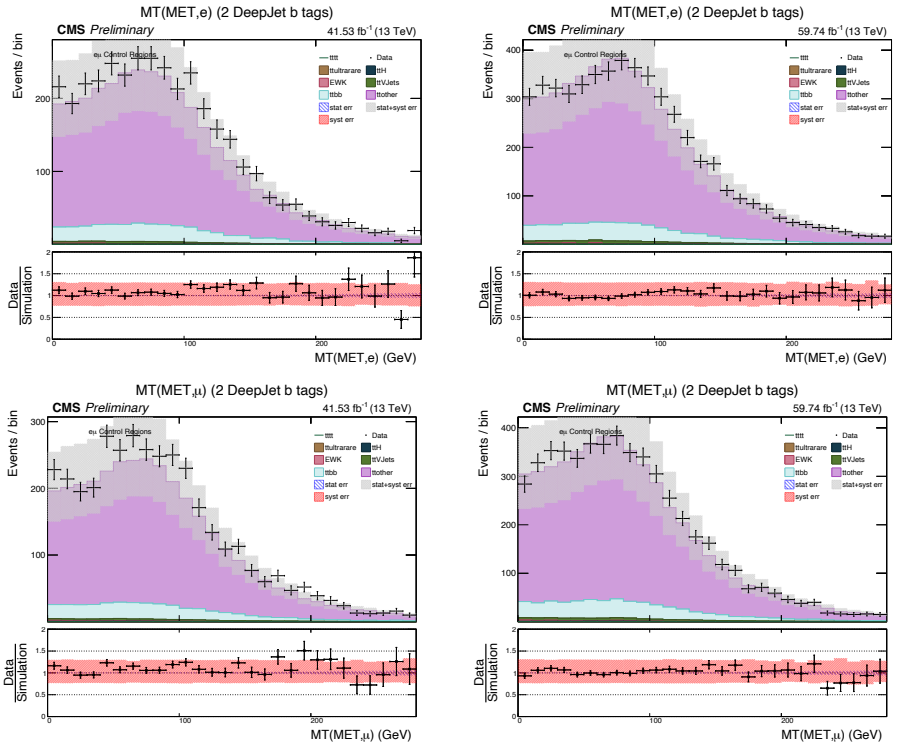

Figure 4.34: M_T for the 2017 and 2018 datasets for the e μ channel. The left panels are for 2017 and the right panels are for 2018. The combination of the e and MET is shown in the top row, and the μ and MET is in the bottom row.

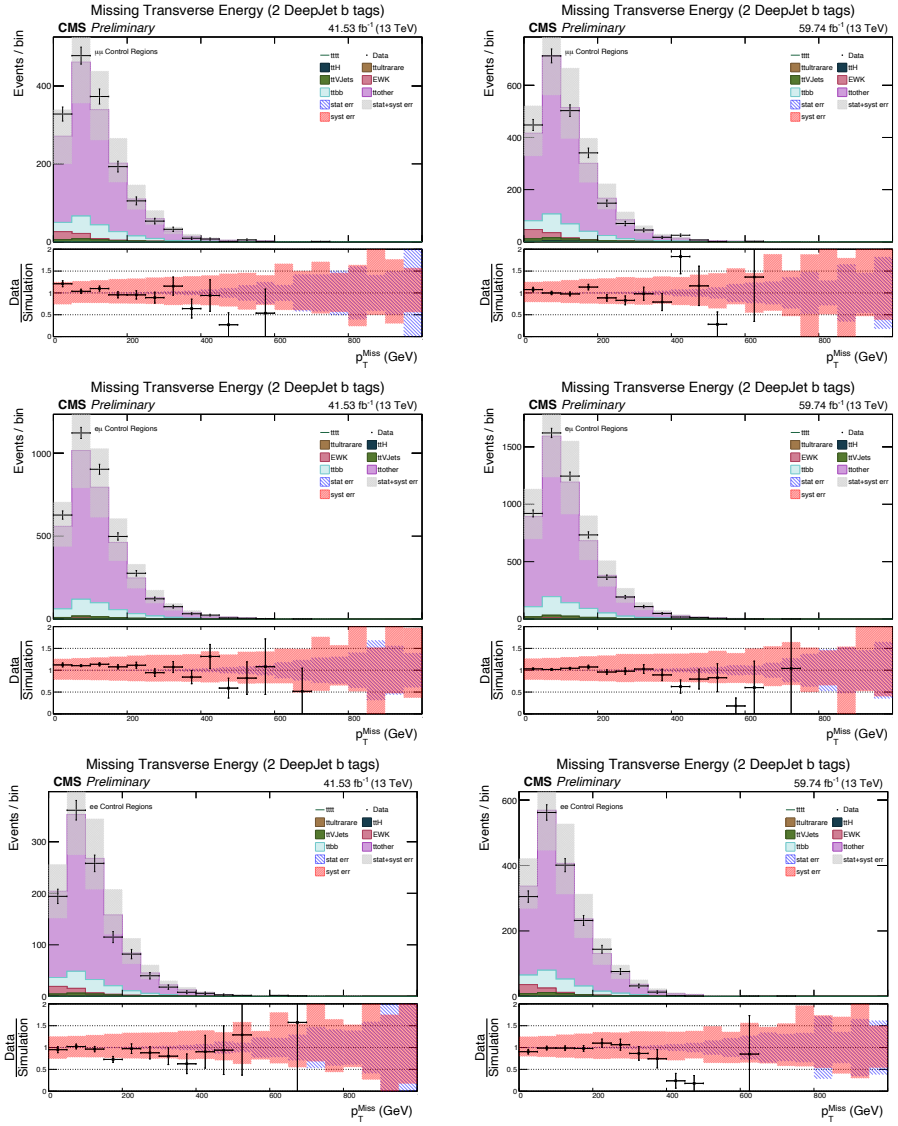

Figure 4.35: p_{T}^{Miss} for the 2017 and 2018 datasets for the three channels. The left panels are for 2017 and the right panels are for 2018. The μ μ channel is shown in the top row, the e μ channel is in the middle row and the e e channel is in the bottom row.

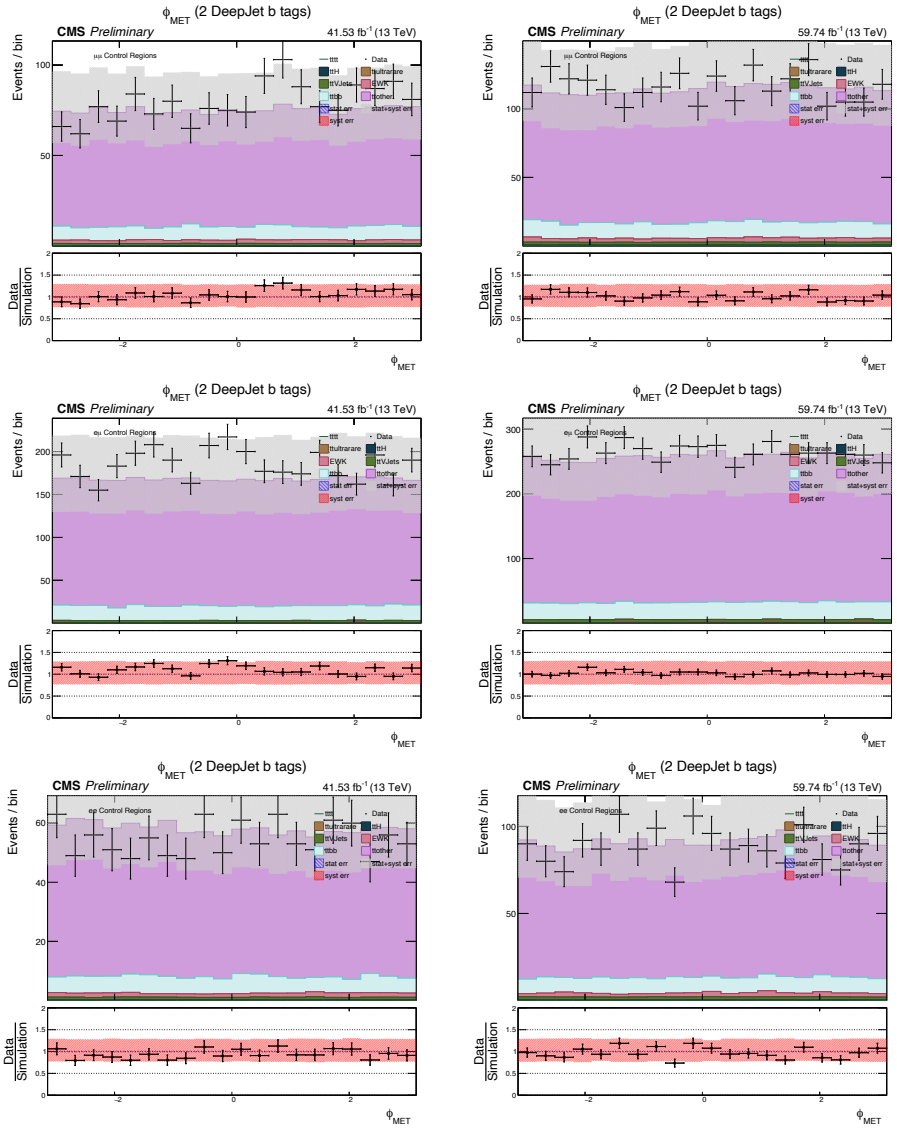

Figure 4.36: ϕ_{MET} for the 2017 and 2018 datasets for the three channels. The left panels are for 2017 and the right panels are for 2018. The μ μ channel is shown in the top row, the e μ channel is in the middle row and the e e channel is in the bottom row.

4.6.2 Signal Enriched Regions

The following plots show pre-fit control distributions for signal-enriched regions. The plots are made for data and simulation events which pass the baseline selection defined in Section 4.4 of which 3 or more jets are medium b tagged. Plots are aggregates of all jet and btag multiplicity categories within this subset, and thus are signal-enriched compared to the 2 b tagged regions (but not exclusively signal regions, due to the presence of lower jet multiplicities). In general, these distributions show that after the $\mathrm{t\bar{t}b\bar{b}}$ normalization correction, we have similar agreement between data and simulation in this section of phase space as the 2 b tag control regions.

- Jet p_{T}: The p_{T} distributions for the four highest p_{T} jets - Figs. 4.37, 4.38, 4.39, 4.40.
- Jet η: The η distributions for the the two leading jets - Figs. 4.41, 4.42.
- Jet multiplicity: Fig. 4.43
- b tagged jet multiplicity: Fig. 4.44
- Deep Jet b-tag discriminants, p_{T} ranking: The discriminant distributions for the four leading jets - Figs. 4.45, 4.46, 4.47, 4.48,
- Deep Jet b-tag discriminants, b-tag ranking: The discriminant distributions for the four highest tagged jets - Figs. 4.49, 4.50, 4.51, 4.52
- dR^{bb}: ΔR between the two leading b-tagged jets in the event - Fig. 4.53.
- H: Scalar sum of $|\vec{p_T}|$ of the selected jets - Fig. 4.54.

- H_T: Scalar sum of p_{T} of the selected jets - Fig. 4.55.

- H_T^{b}: H_T for the subset of selected jets passing the medium b tagging working point - Fig. 4.56.

- Muon p_{T}, η, PF Iso: The p_{T}, η, and total PF isolation distribution of the leading muon - Figs. 4.57, 4.58, 4.59

- Electron p_{T}, η, PF Iso: The p_{T}, η, and total PF isolation distribution of the leading electron - Fig. 4.60, 4.61, 4.62

- Subleading lepton p_{T}: Subleading muon and electron in the same-flavor channels - Fig. 4.63

- dR^{ll}: ΔR between the two leptons in the event - Fig. 4.64.

- Muon vertexing: The dz and 3D impact parameter of the muons for the e μ channel are shown in Fig. 4.65 and the distributions for the μ μ channel are shown in Fig. 4.66.

- Electron vertexing: The dz and 3D impact parameter of the electrons for the e μ channel are shown in Fig. 4.67 and the distributions for the e e channel are shown in Fig. 4.68.

- $M_T(l, MET)$: The transverse mass distributions for the ee and μ μ channels are shown in Fig. 4.69 and the distributions for the e μ channel are shown in Fig. 4.70, where each possible combination of lepton and p_{T}^{Miss} is calculated.

- Missing transverse momentum (p_{T}^{Miss}): The missing transverse momentum magnitude is shown in Fig. 4.71 and the phi distribution is in Fig. 4.72.

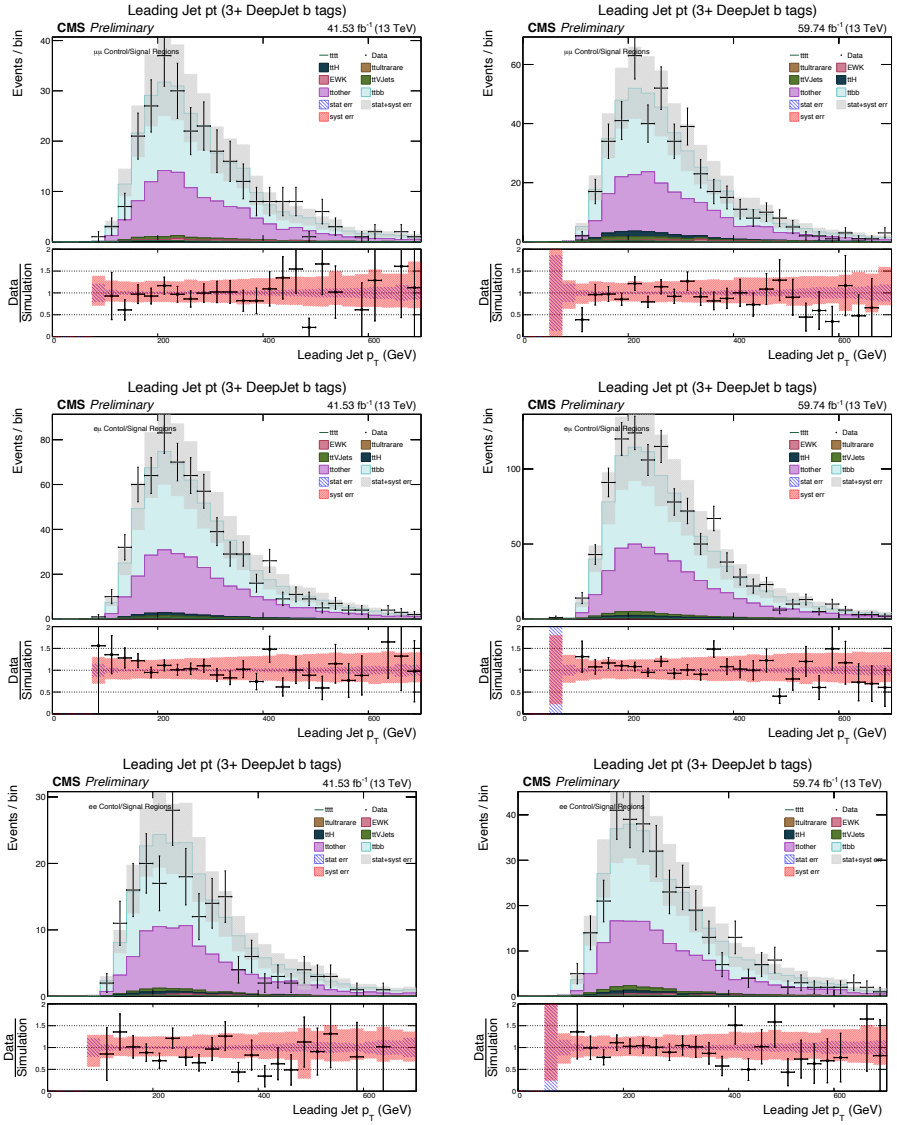

Figure 4.37: Leading Jet p_{T} for the 2017 and 2018 datasets for the three channels. The left panels are for 2017 and the right panels are for 2018. The μ μ channel is shown in the top row, the e μ channel is in the middle row and the e e channel is in the bottom row.

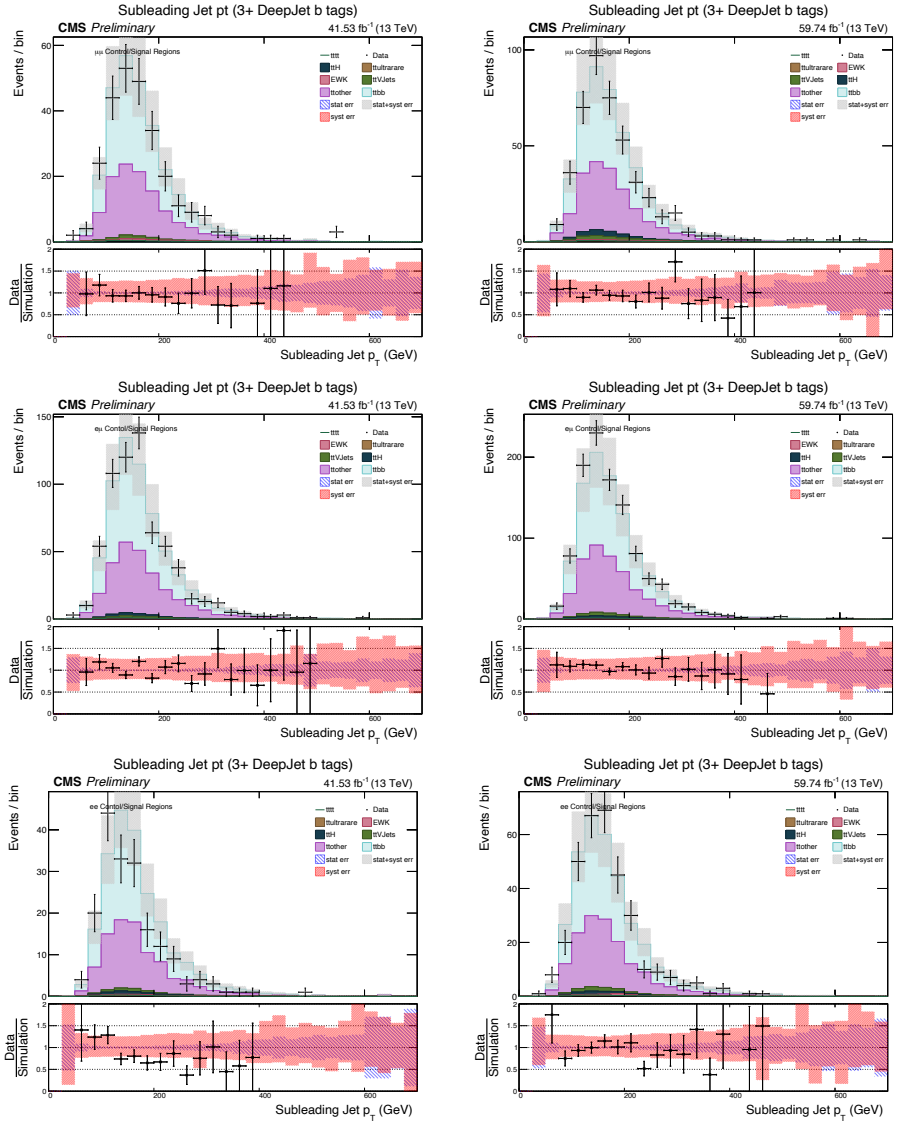

Figure 4.38: Sub-leading Jet p_{T} for the 2017 and 2018 datasets for the three channels. The left panels are for 2017 and the right panels are for 2018. The μ μ channel is shown in the top row, the e μ channel is in the middle row and the e e channel is in the bottom row.

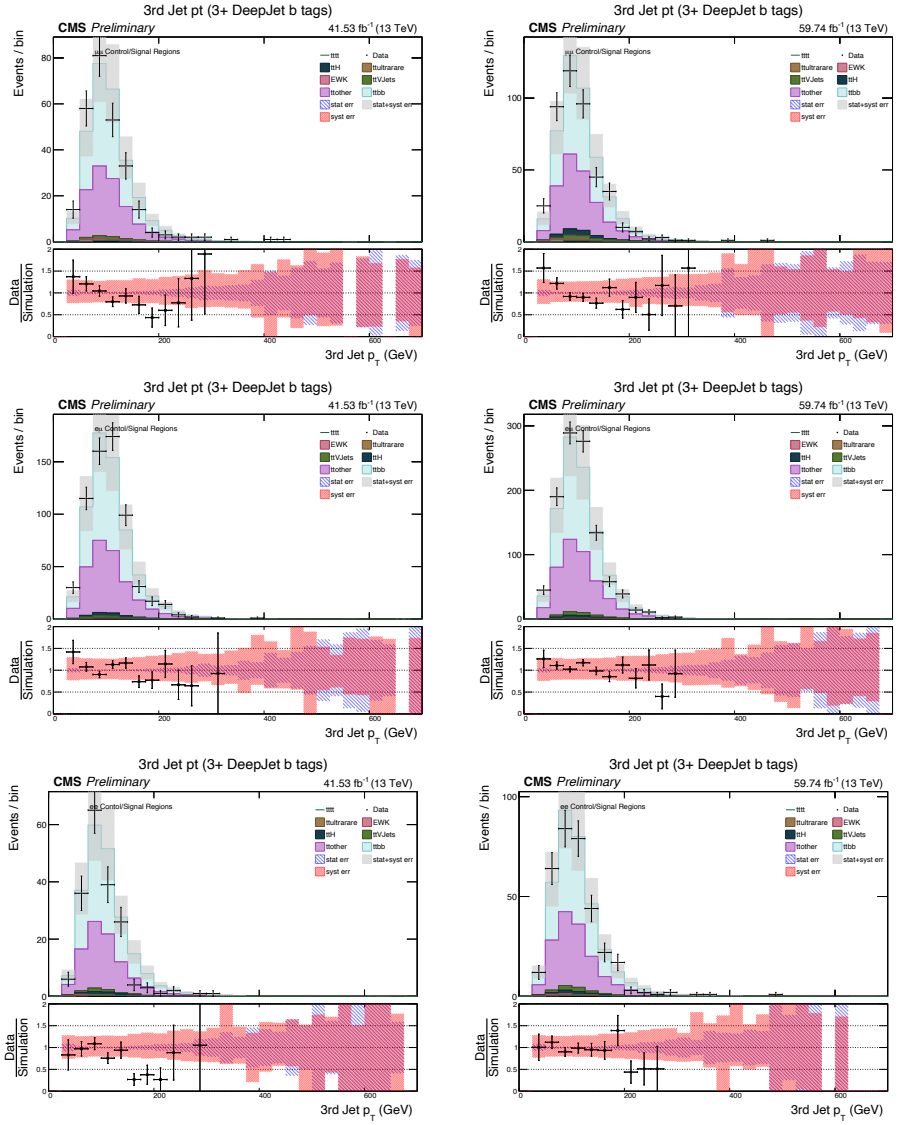

Figure 4.39: p_{T} of Jet 3 for the 2017 and 2018 datasets for the three channels. The left panels are for 2017 and the right panels are for 2018. The μ μ channel is shown in the top row, the e μ channel is in the middle row and the e e channel is in the bottom row.

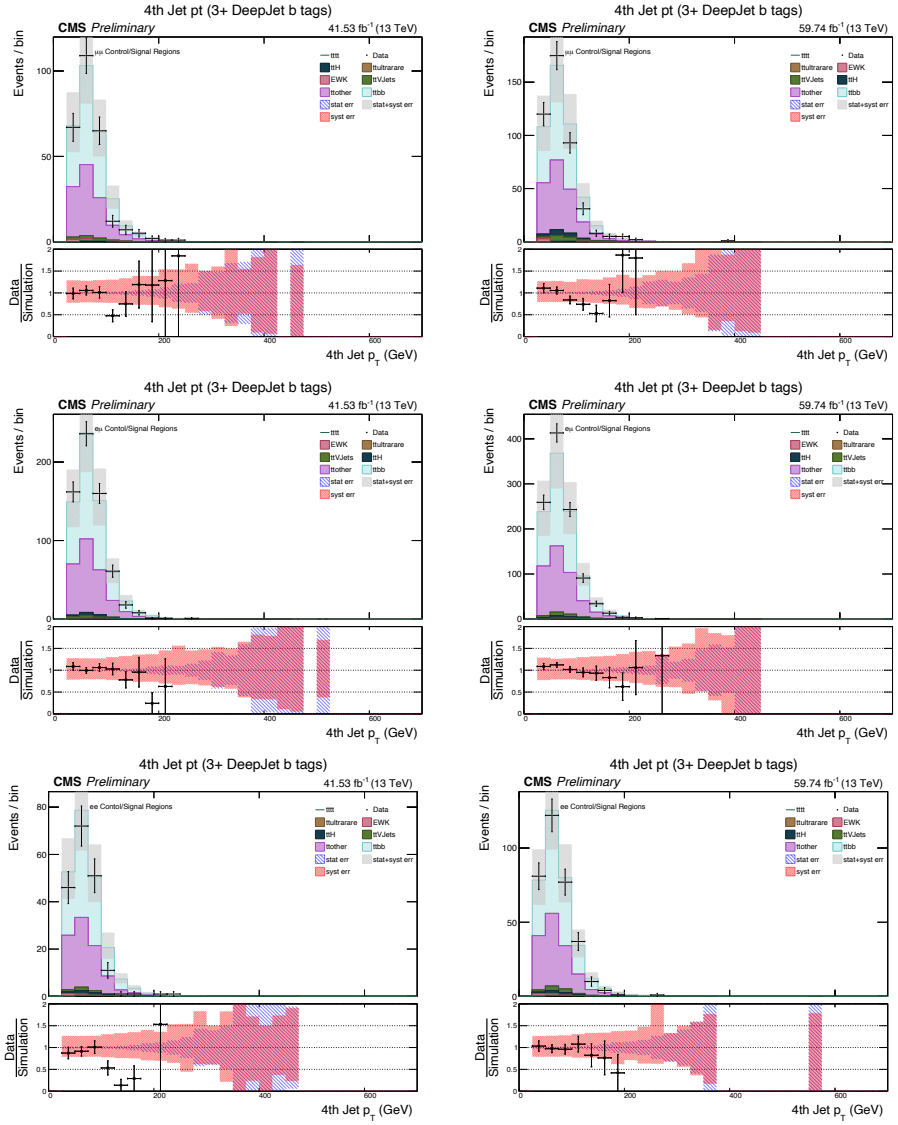

Figure 4.40: p_{T} of Jet 4 for the 2017 and 2018 datasets for the three channels. The left panels are for 2017 and the right panels are for 2018. The μ μ channel is shown in the top row, the e μ channel is in the middle row and the e e channel is in the bottom row.

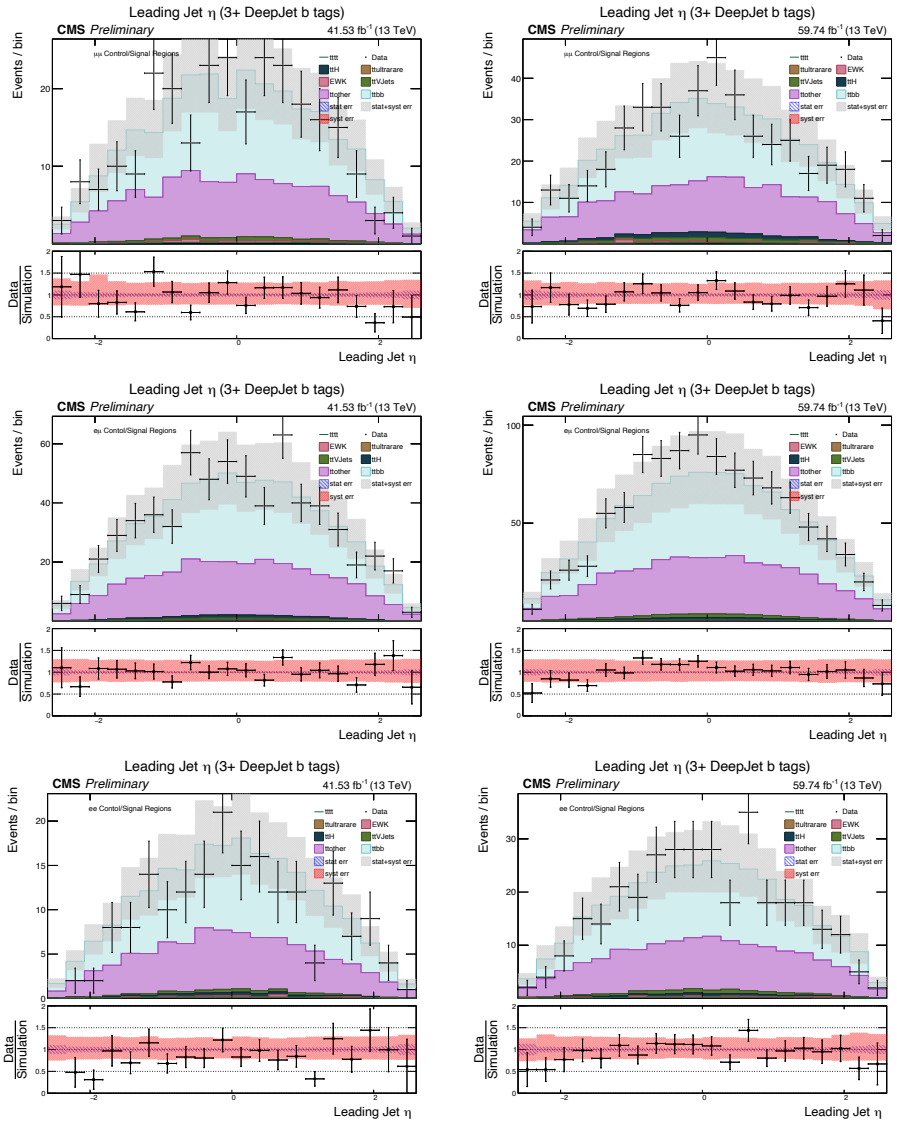

Figure 4.41: Leading Jet η for the 2017 and 2018 datasets for the three channels. The left panels are for 2017 and the right panels are for 2018. The μ μ channel is shown in the top row, the e μ channel is in the middle row and the e e channel is in the bottom row.

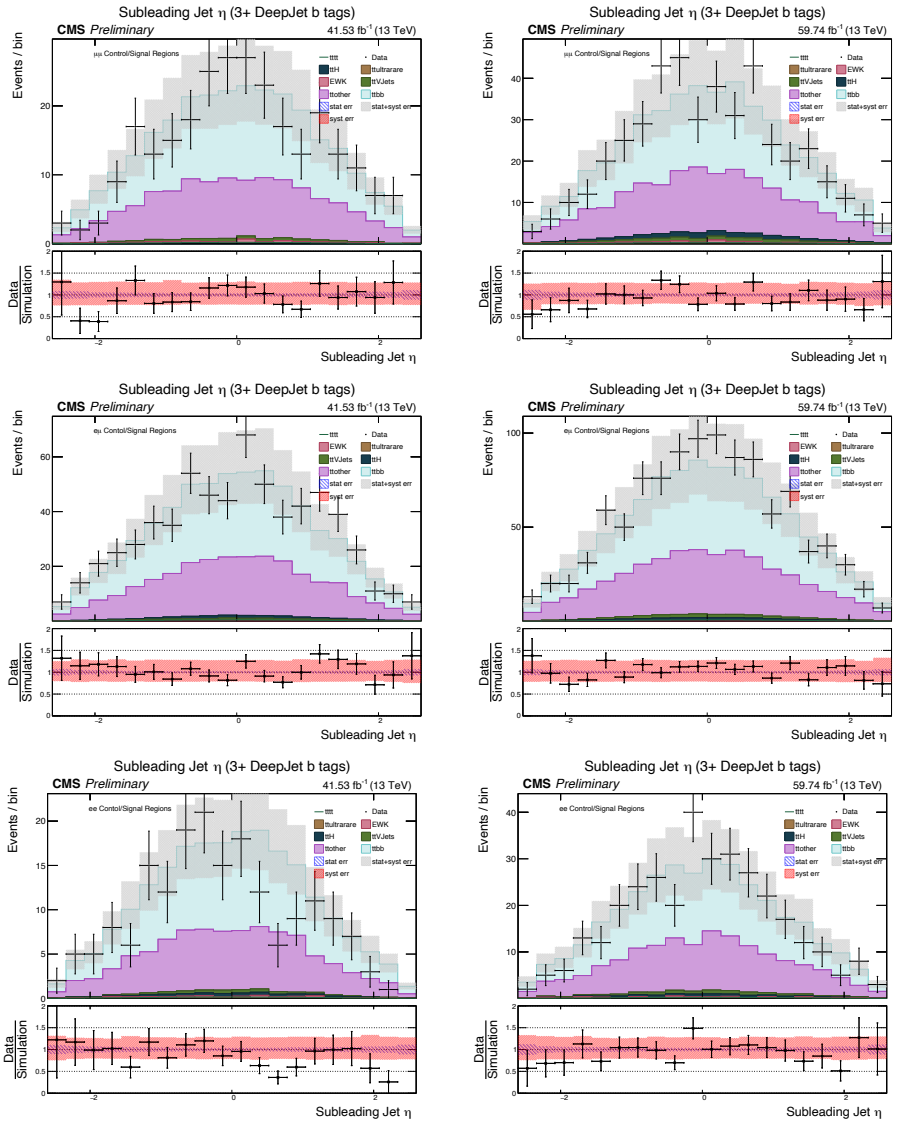

Figure 4.42: Sub-leading Jet η for the 2017 and 2018 datasets for the three channels. The left panels are for 2017 and the right panels are for 2018. The μ μ channel is shown in the top row, the e μ channel is in the middle row and the e e channel is in the bottom row.

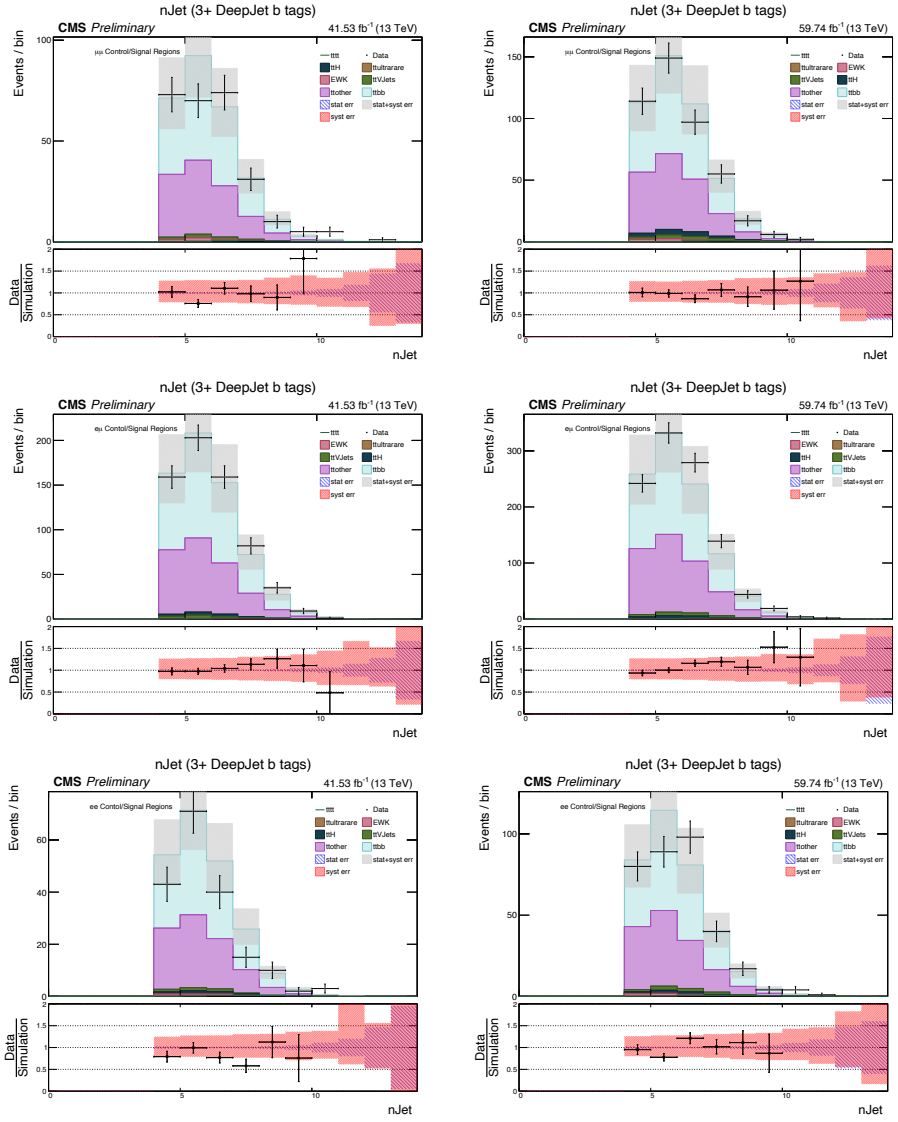

Figure 4.43: Jet multiplicity for the 2017 and 2018 datasets for the three channels. The left panels are for 2017 and the right panels are for 2018. The μ μ channel is shown in the top row, the e μ channel is in the middle row and the e e channel is in the bottom row.

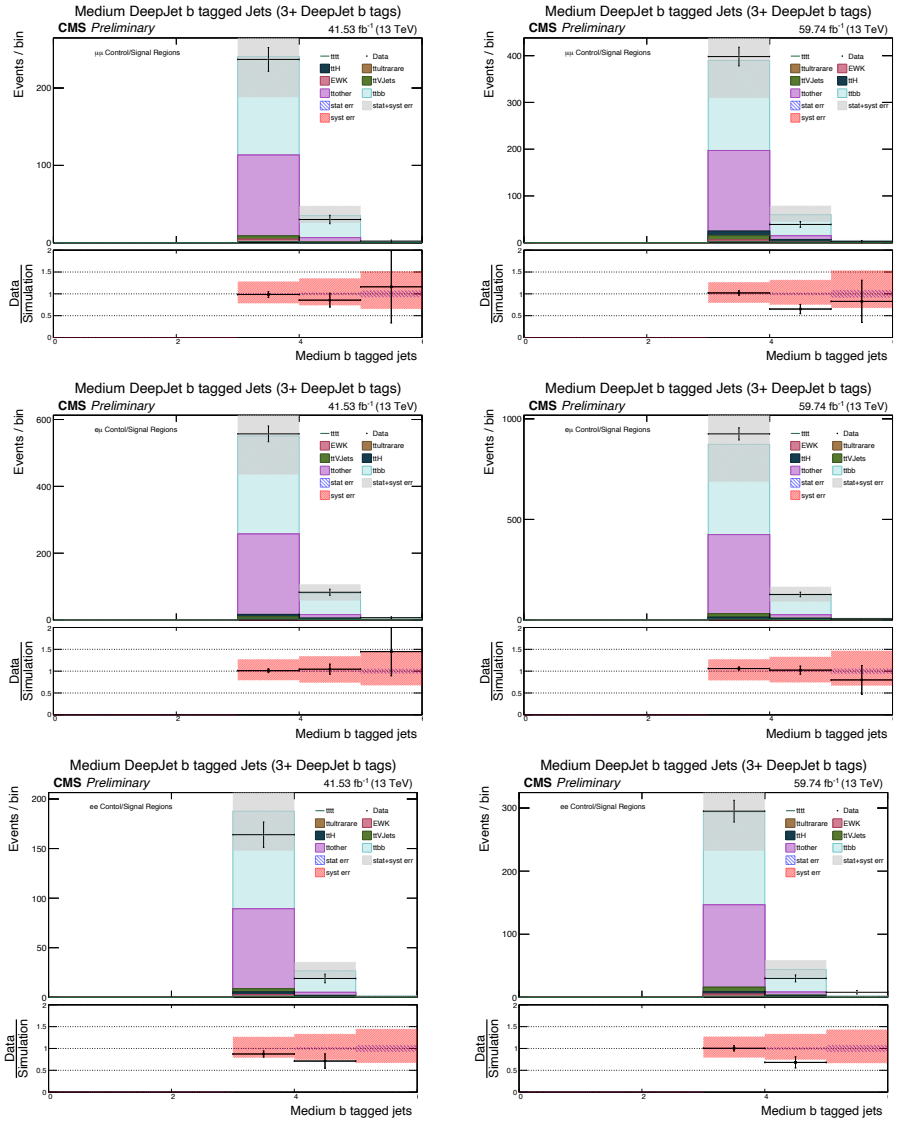

Figure 4.44: Medium b tagged jet multiplicity for the 2017 and 2018 datasets for the three channels. The left panels are for 2017 and the right panels are for 2018. The μ μ channel is shown in the top row, the e μ channel is in the middle row and the e e channel is in the bottom row.

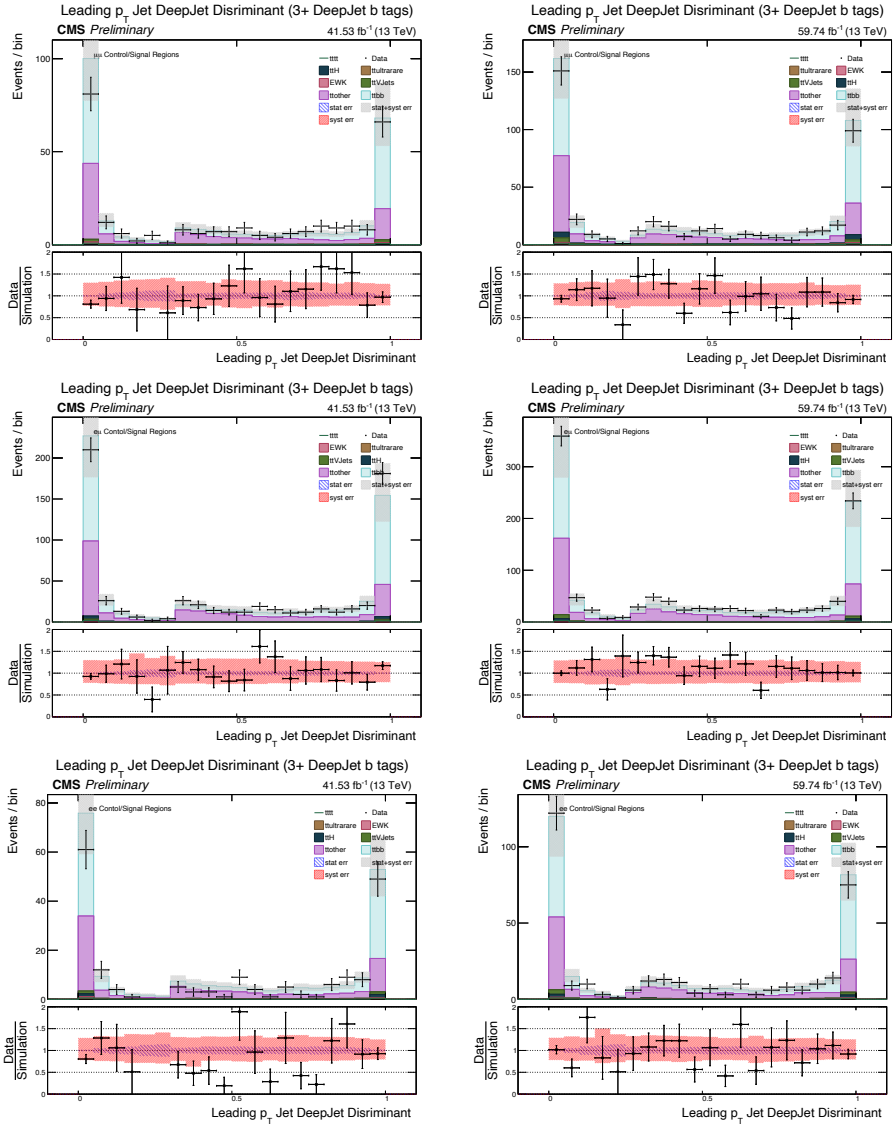

Figure 4.45: Leading Jet b-tag disciminant for the 2017 and 2018 datasets for the three channels. The left panels are for 2017 and the right panels are for 2018. The μ μ channel is shown in the top row, the e μ channel is in the middle row and the e e channel is in the bottom row.

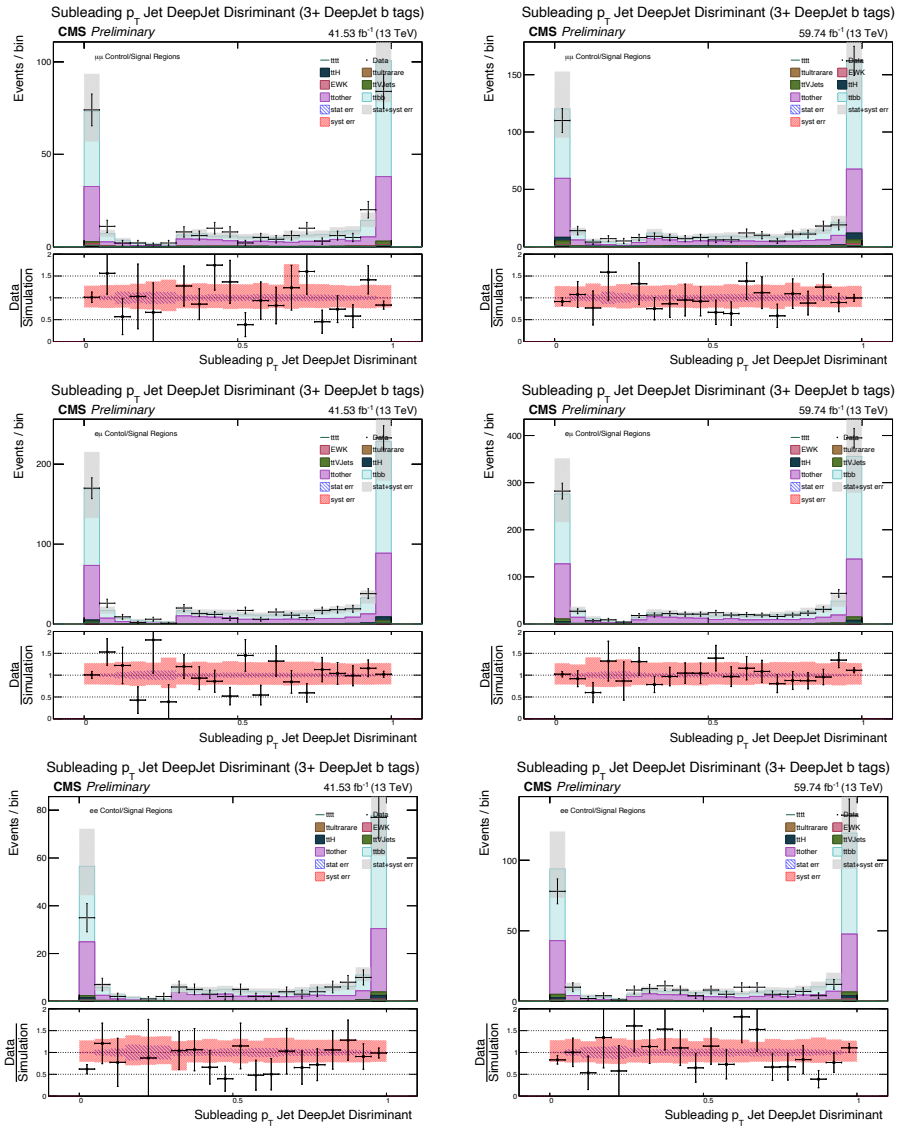

Figure 4.46: Sub-leading Jet b-tag disciminant for the 2017 and 2018 datasets for the three channels. The left panels are for 2017 and the right panels are for 2018. The μ μ channel is shown in the top row, the e μ channel is in the middle row and the e e channel is in the bottom row.

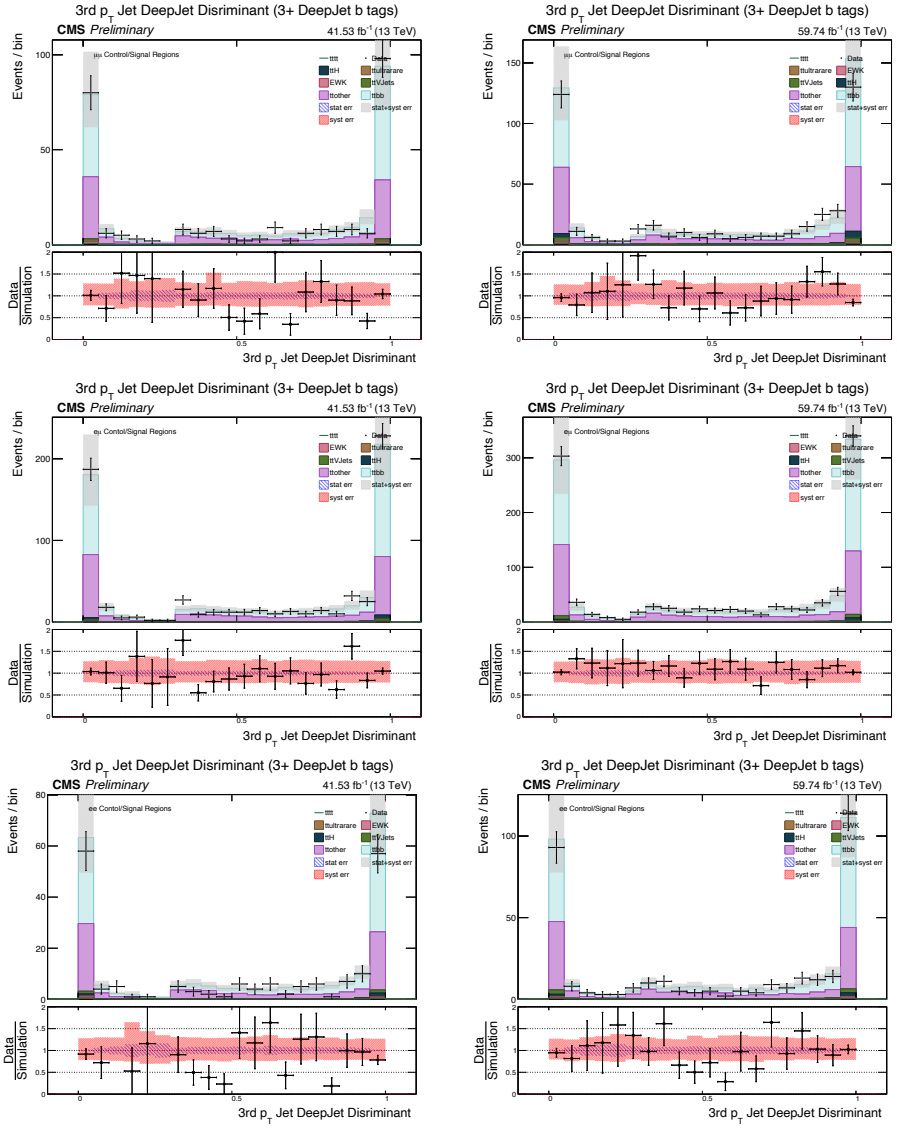

Figure 4.47: Third highest p_{T} jet b-tag disciminant for the 2017 and 2018 datasets for the three channels. The left panels are for 2017 and the right panels are for 2018. The μ μ channel is shown in the top row, the e μ channel is in the middle row and the e e channel is in the bottom row.

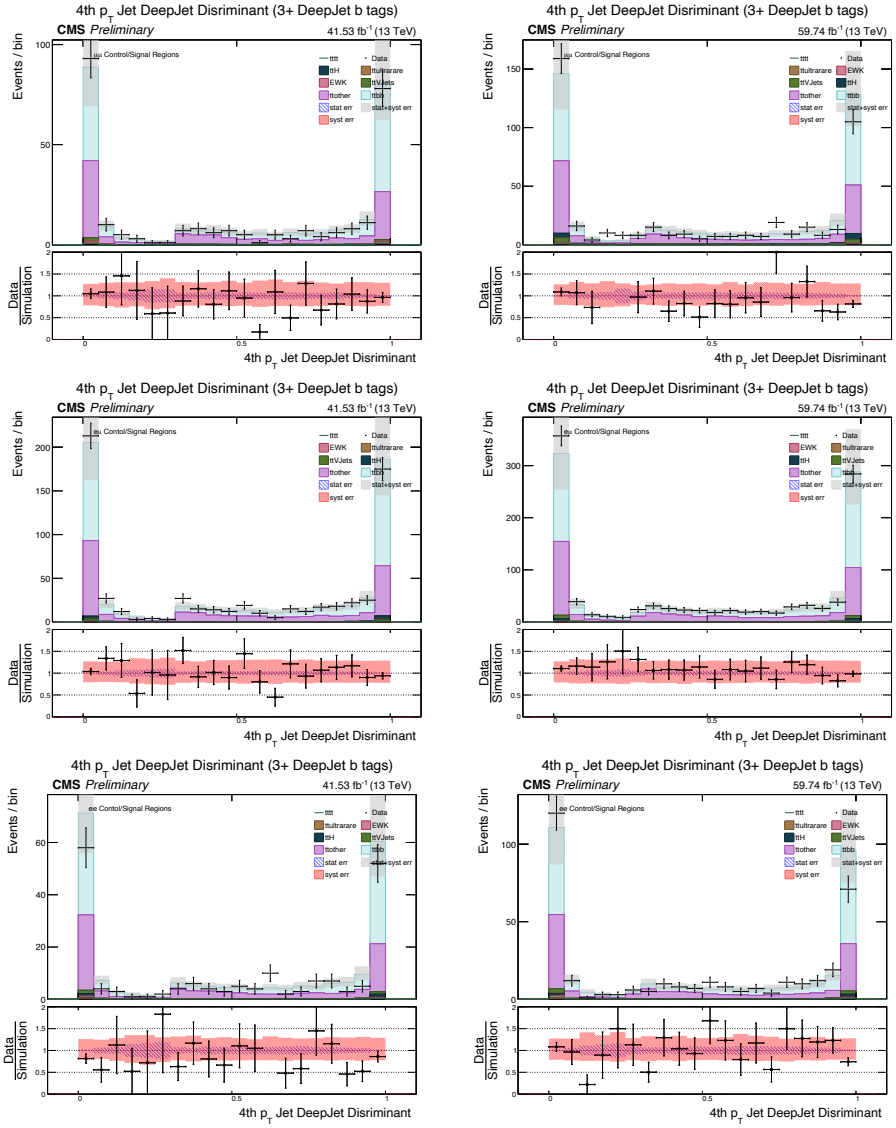

Figure 4.48: Fourth highest p_{T} jet b-tag disciminant for the 2017 and 2018 datasets for the three channels. The left panels are for 2017 and the right panels are for 2018. The μ μ channel is shown in the top row, the e μ channel is in the middle row and the e e channel is in the bottom row.

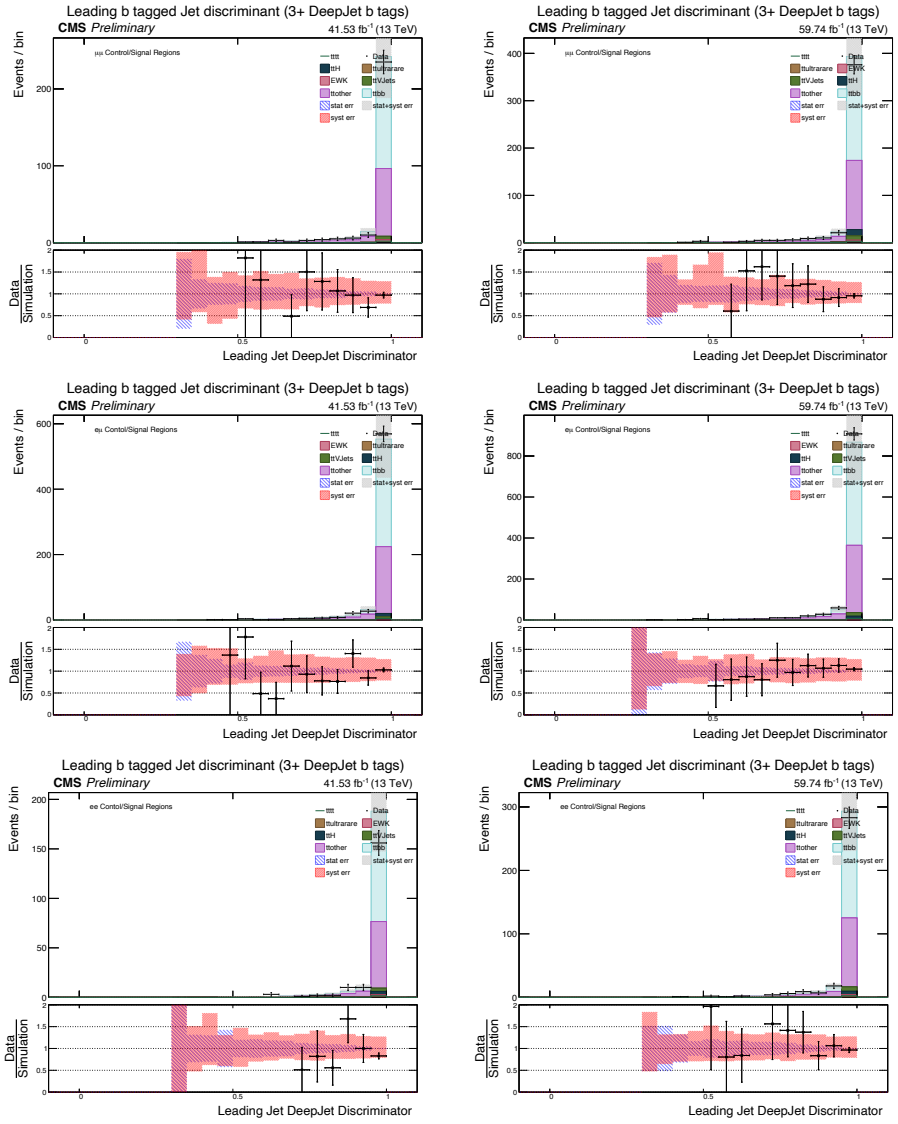

Figure 4.49: Highest b tagged jet disciminant for the 2017 and 2018 datasets for the three channels. The left panels are for 2017 and the right panels are for 2018. The μ μ channel is shown in the top row, the e μ channel is in the middle row and the e e channel is in the bottom row.

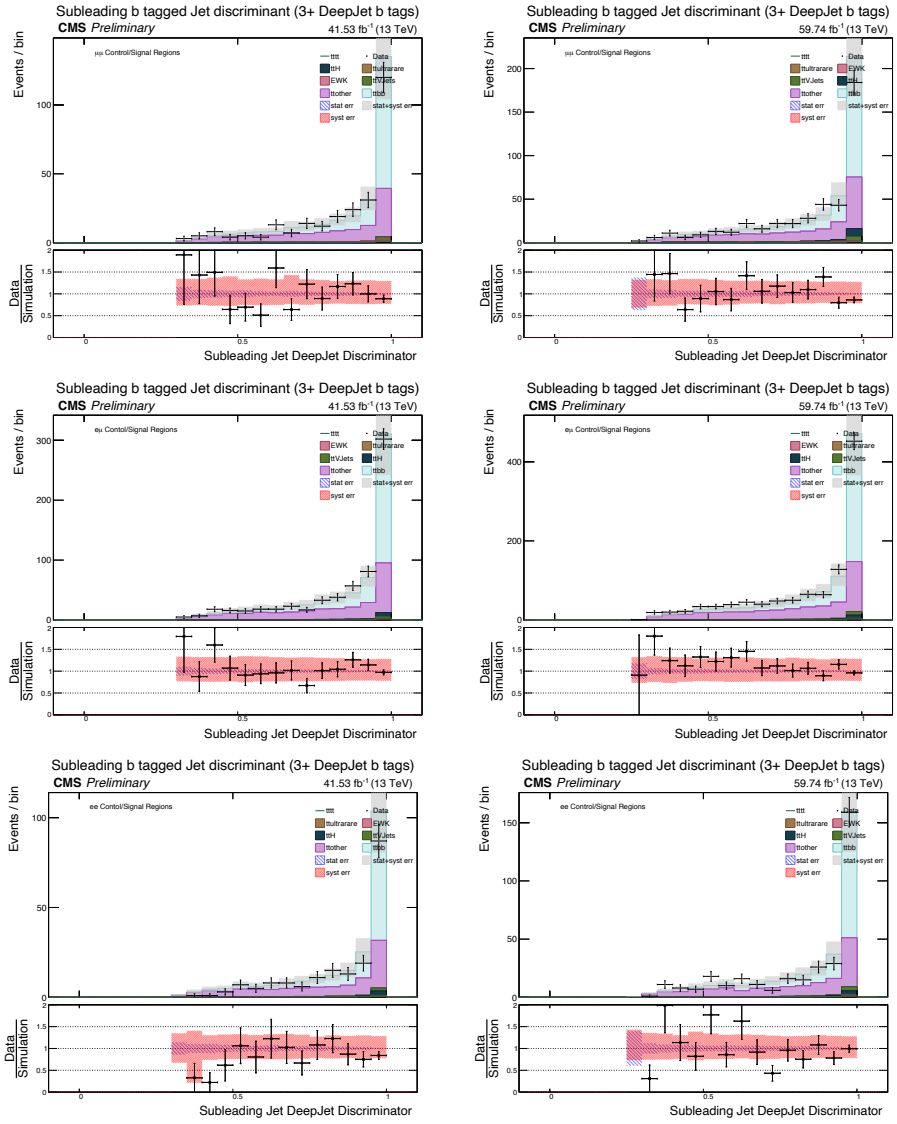

Figure 4.50: Second highest b tagged jet disciminant for the 2017 and 2018 datasets for the three channels. The left panels are for 2017 and the right panels are for 2018. The μ μ channel is shown in the top row, the e μ channel is in the middle row and the e e channel is in the bottom row.

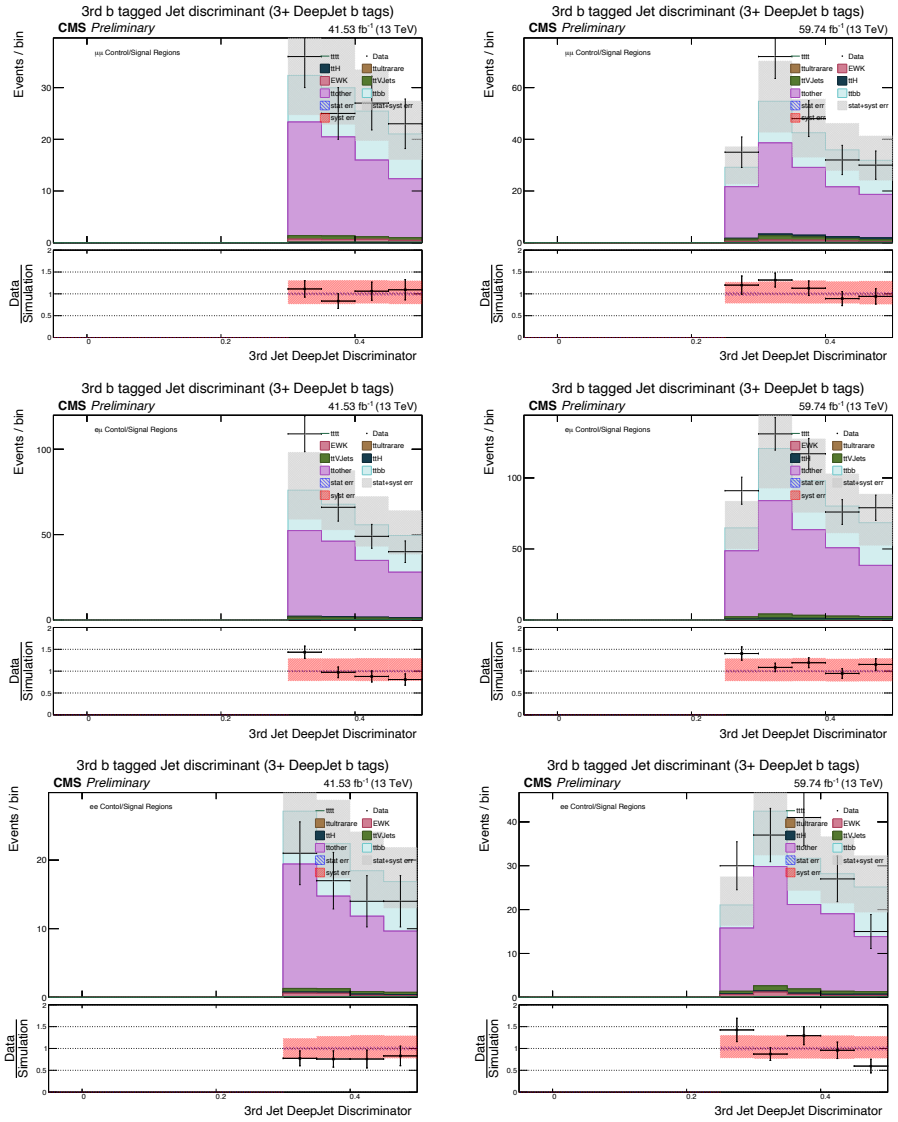

Figure 4.51: Third highest b tagged jet disciminant for the 2017 and 2018 datasets for the three channels, covering only the range of jets untagged at the medium working point. The left panels are for 2017 and the right panels are for 2018. The μ μ channel is shown in the top row, the e μ channel is in the middle row and the e e channel is in the bottom row.

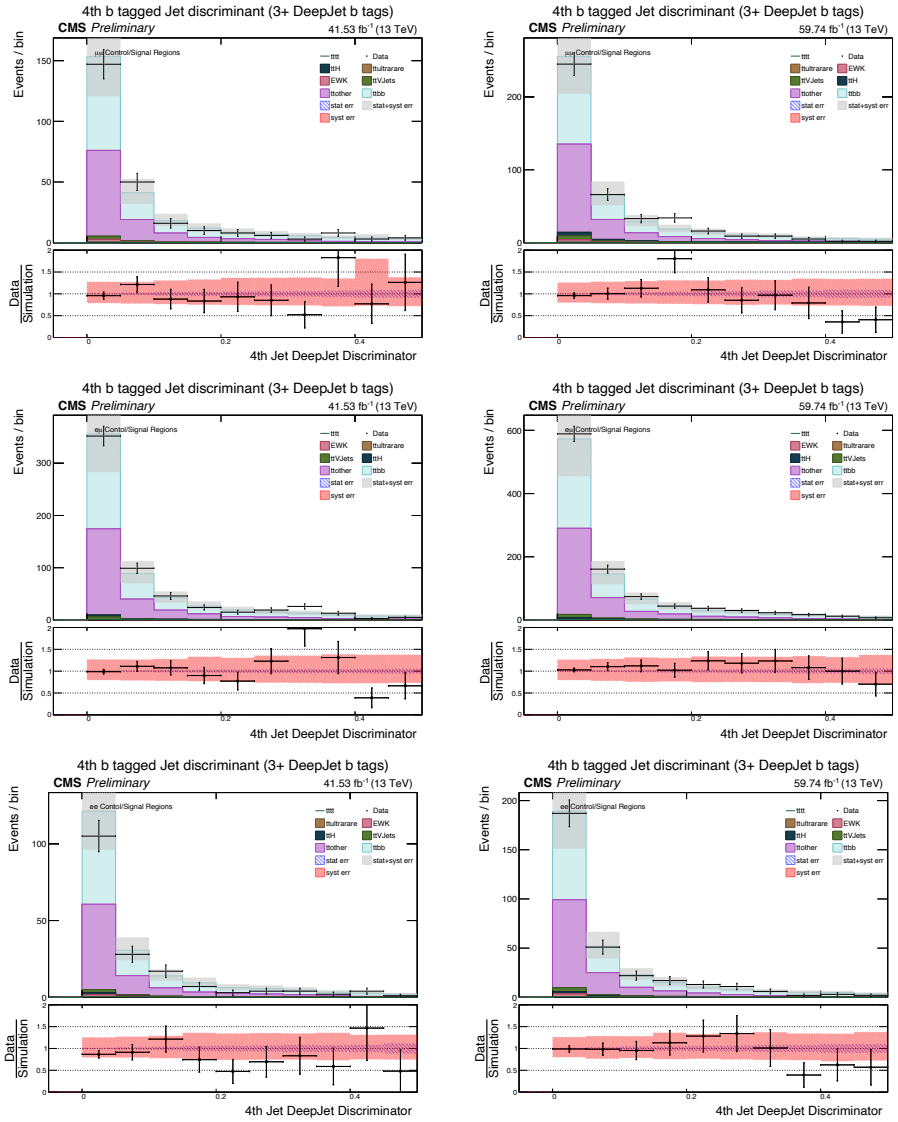

Figure 4.52: Fourth highest b tagged jet disciminant for the 2017 and 2018 datasets for the three channels, covering only the range of jets untagged at the medium working point. The left panels are for 2017 and the right panels are for 2018. The μ μ channel is shown in the top row, the e μ channel is in the middle row and the e e channel is in the bottom row.

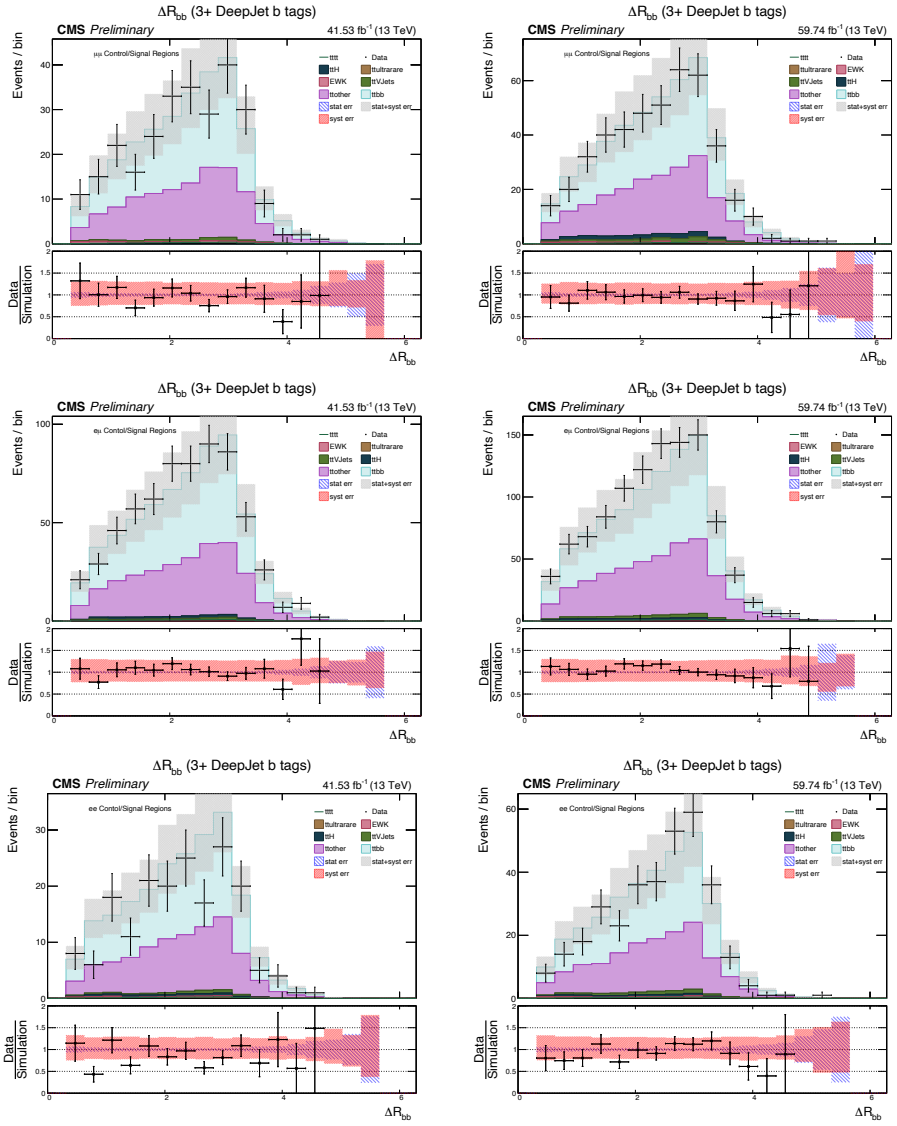

Figure 4.53: dR^{bb} for the 2017 and 2018 datasets for the three channels. The left panels are for 2017 and the right panels are for 2018. The μ μ channel is shown in the top row, the e μ channel is in the middle row and the e e channel is in the bottom row.

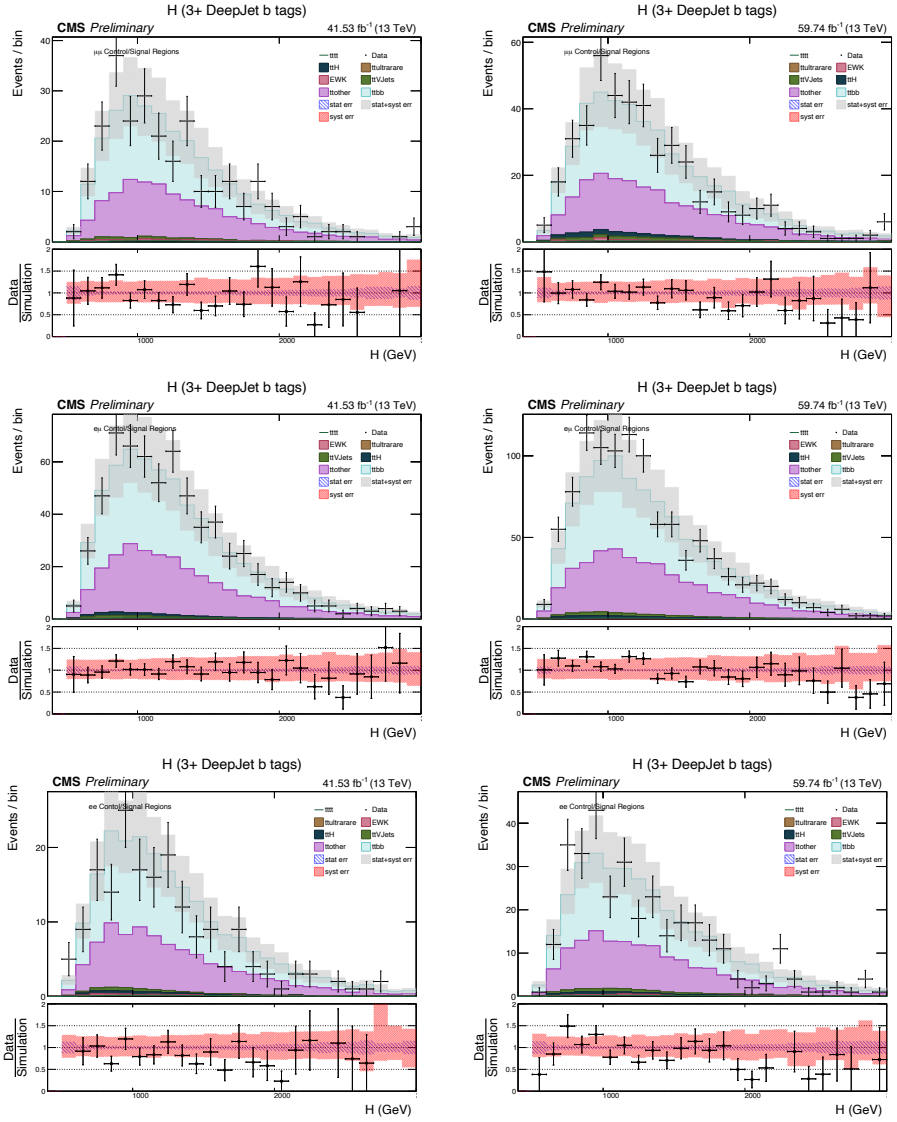

Figure 4.54: H for the 2017 and 2018 datasets for the three channels. The left panels are for 2017 and the right panels are for 2018. The μ μ channel is shown in the top row, the e μ channel is in the middle row and the e e channel is in the bottom row.

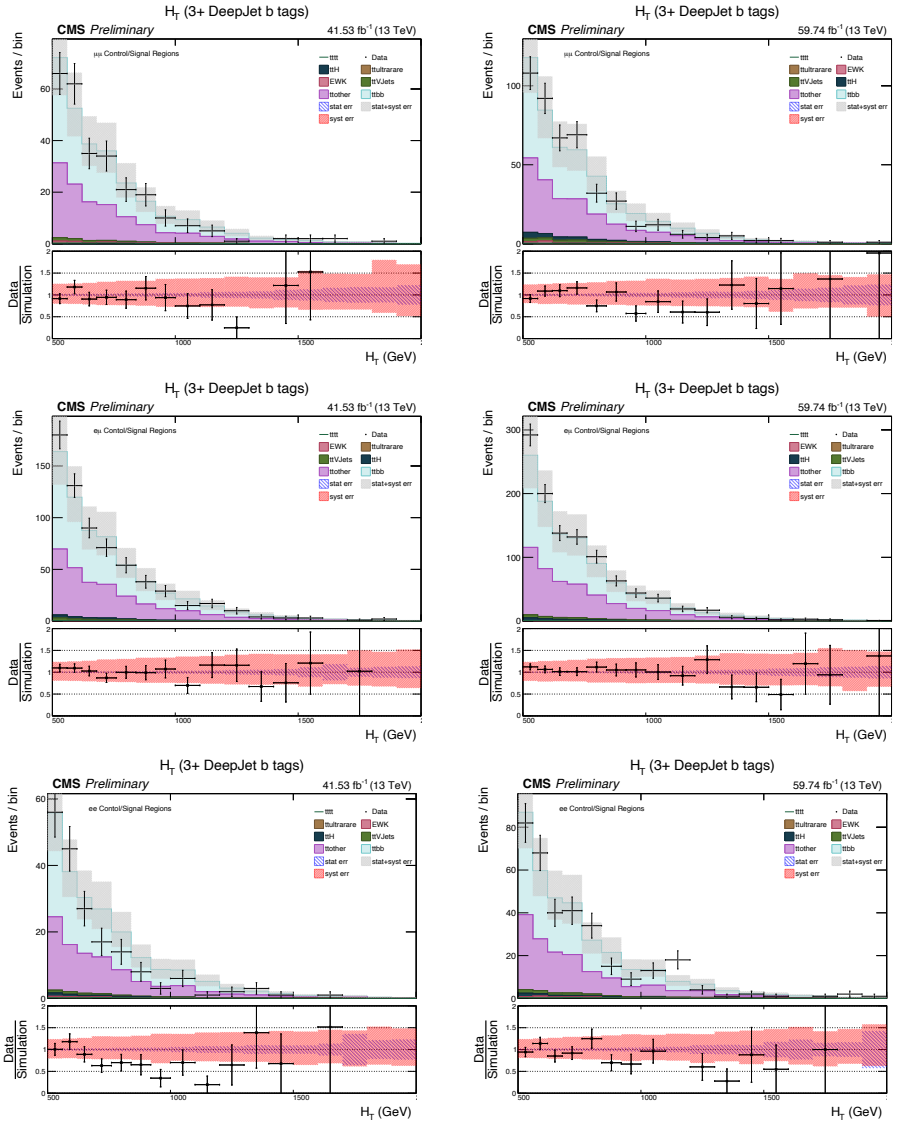

Figure 4.55: H_T for the 2017 and 2018 datasets for the three channels. The left panels are for 2017 and the right panels are for 2018. The μ μ channel is shown in the top row, the e μ channel is in the middle row and the e e channel is in the bottom row.

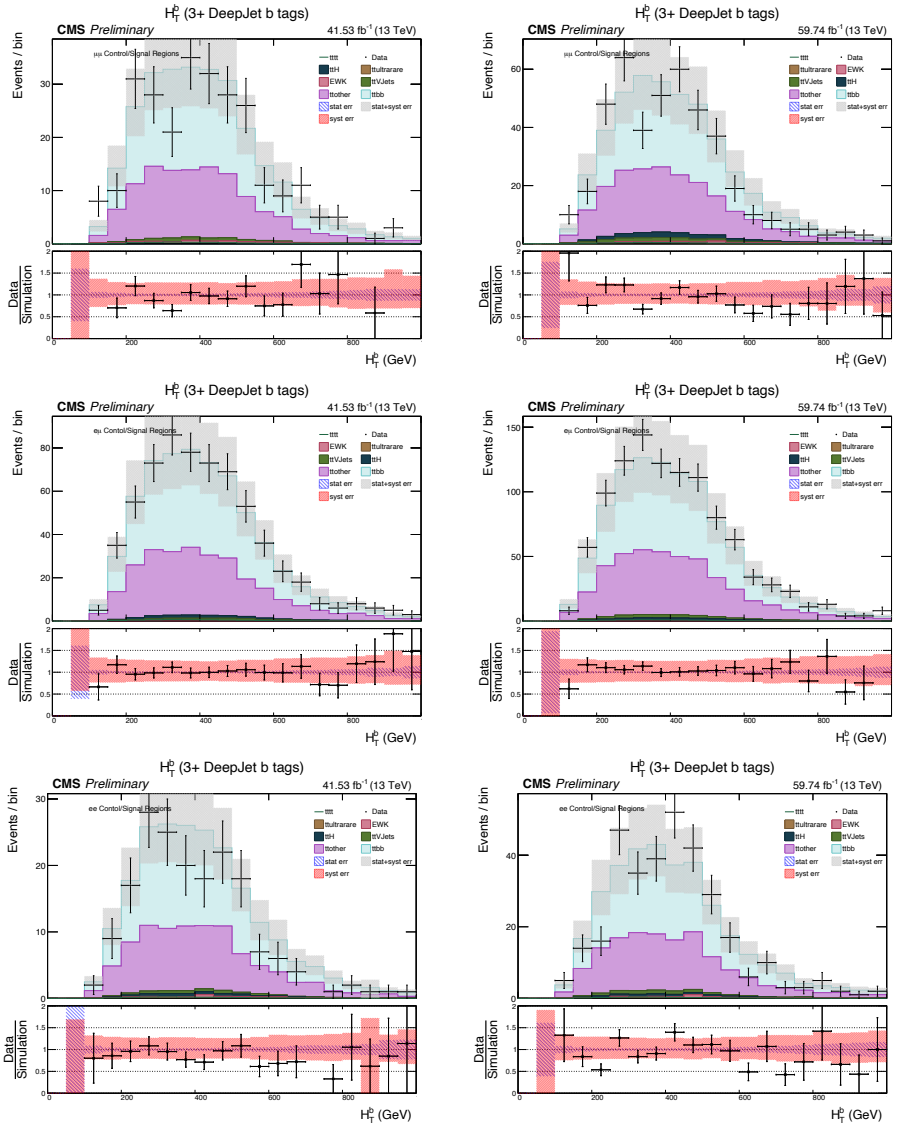

Figure 4.56: H_T^{b} for the 2017 and 2018 datasets for the three channels. The left panels are for 2017 and the right panels are for 2018. The μ μ channel is shown in the top row, the e μ channel is in the middle row and the e e channel is in the bottom row.

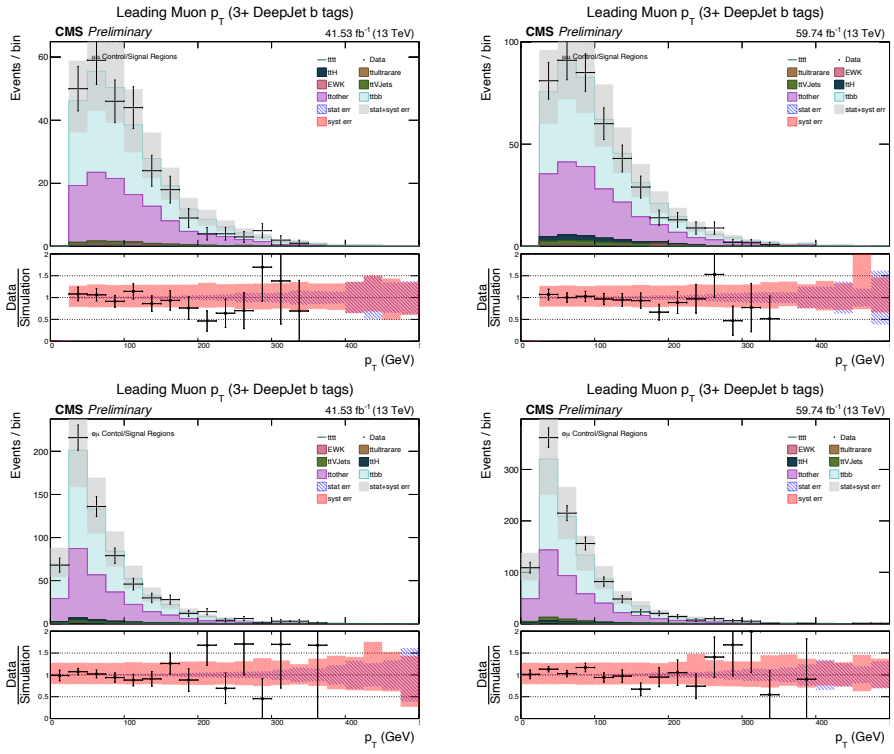

Figure 4.57: Leading muon p_{T} for the 2017 and 2018 datasets. The left panels are for 2017 and the right panels are for 2018. The μ μ channel is shown in the top row, the e μ channel is in the bottom row

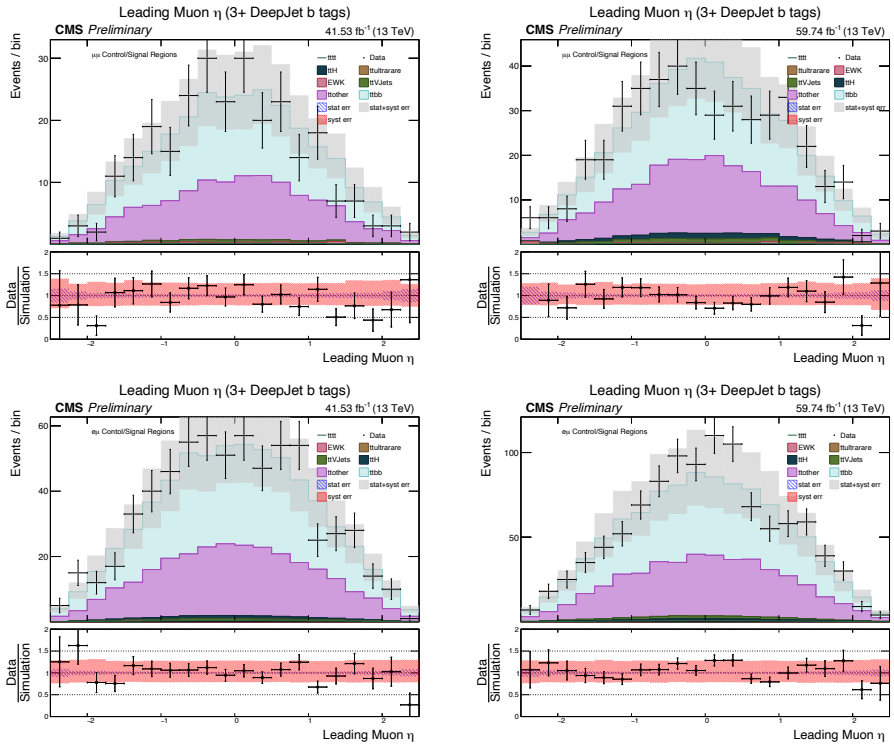

Figure 4.58: Leading muon η for the 2017 and 2018 datasets. The left panels are for 2017 and the right panels are for 2018. The μ μ channel is shown in the top row, the e μ channel is in the bottom row

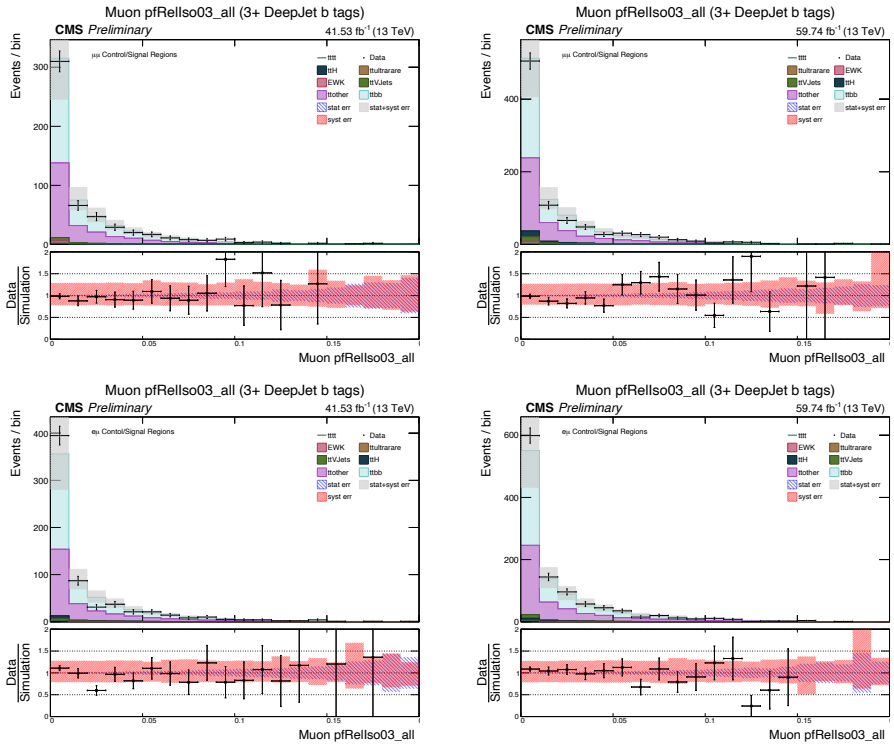

Figure 4.59: Muon PF Relative Isolation (R=0.03) for the 2017 and 2018 datasets. The left panels are for 2017 and the right panels are for 2018. The muons of the μ μ channel are shown in the top row, the muons of the e μ channel are in the bottom row

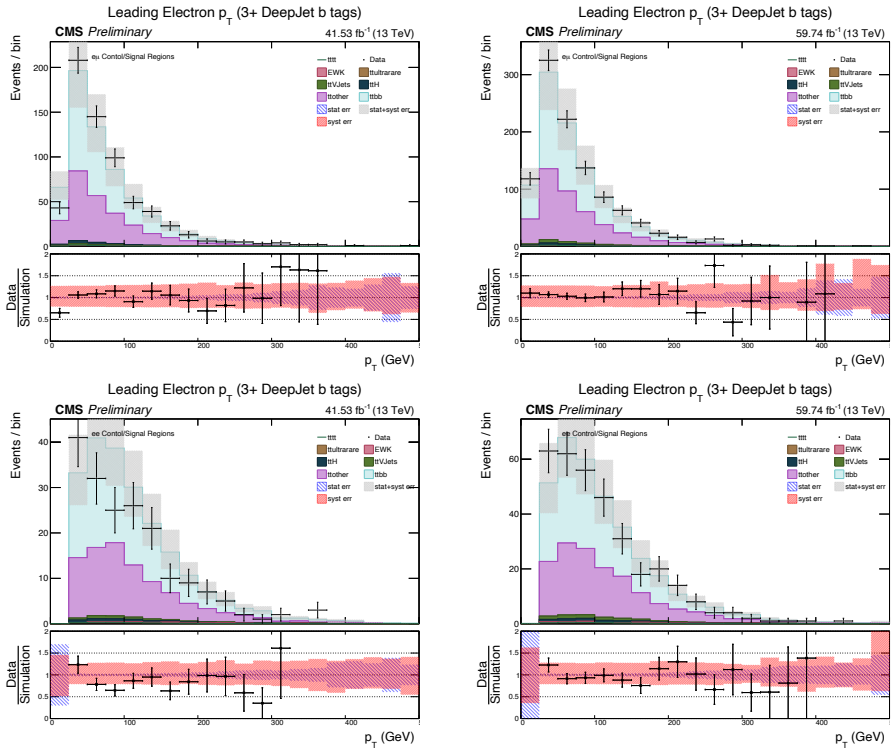

Figure 4.60: Leading electron p_{T} for the 2017 and 2018 datasets. The left panels are for 2017 and the right panels are for 2018. The e μ channel is shown in the top row, the e e channel is in the bottom row

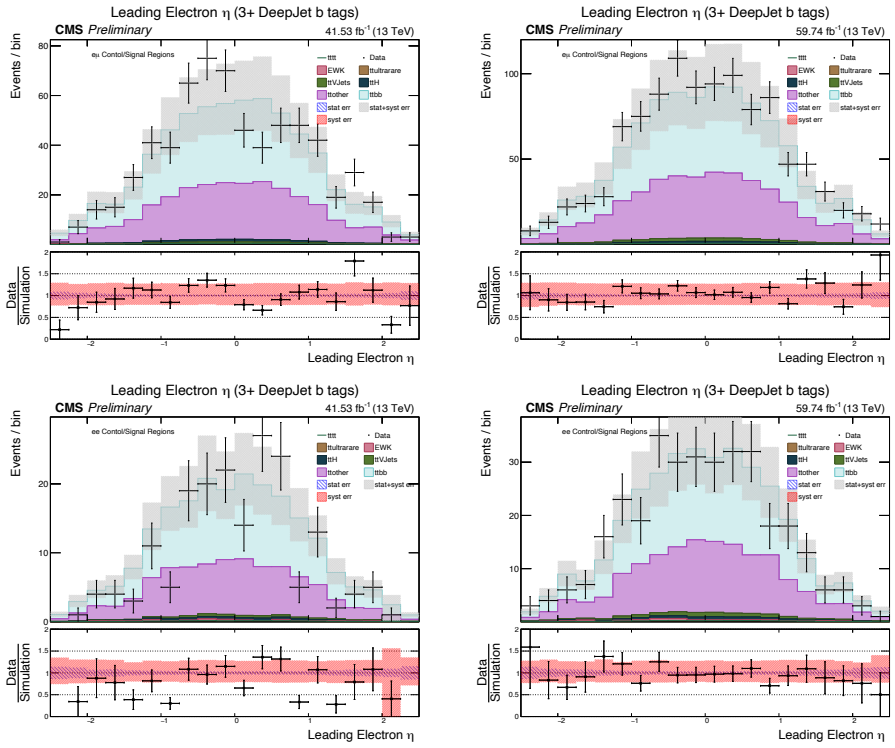

Figure 4.61: Leading electron η for the 2017 and 2018 datasets. The left panels are for 2017 and the right panels are for 2018. The e μ channel is shown in the top row, the e e channel is in the bottom row

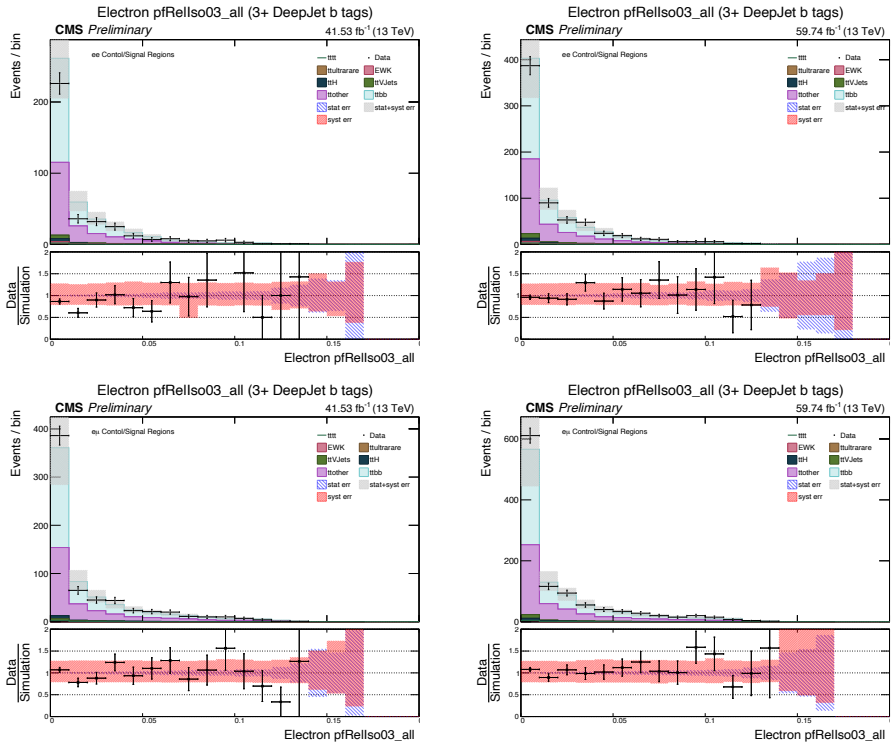

Figure 4.62: Electron PF Relative Isolation (R=0.03) for the 2017 and 2018 datasets. The left panels are for 2017 and the right panels are for 2018. The electrons of the e e channel are shown in the top row, the electrons of the e μ channel are in the bottom row

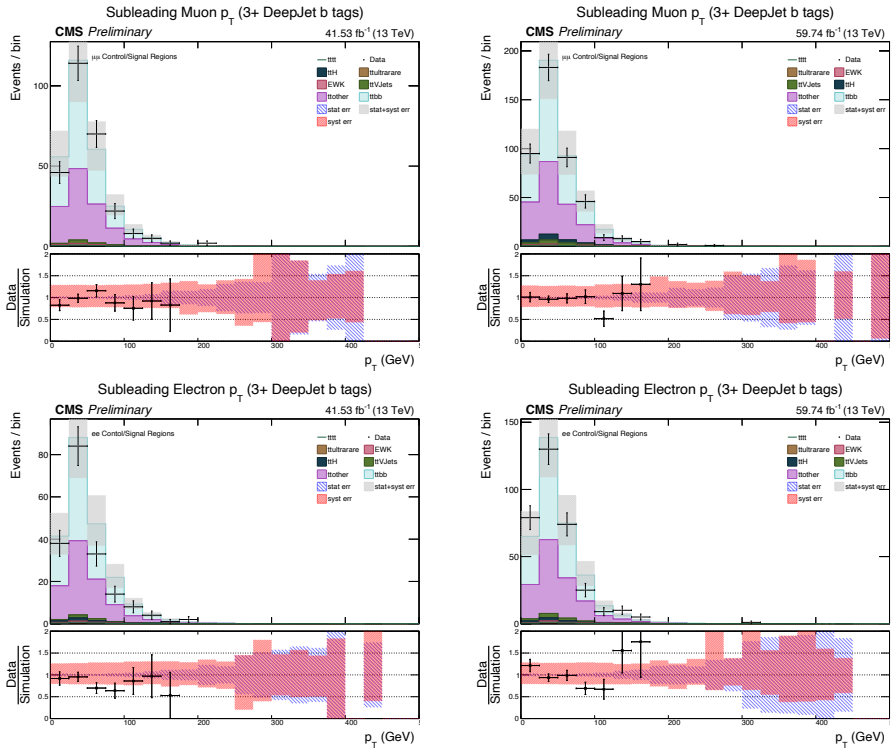

Figure 4.63: Subleading lepton p_{T} for the 2017 and 2018 datasets. The left panels are for 2017 and the right panels are for 2018. The muons of the μ μ channel are shown in the top row, the electrons of the e μ channel are in the bottom row

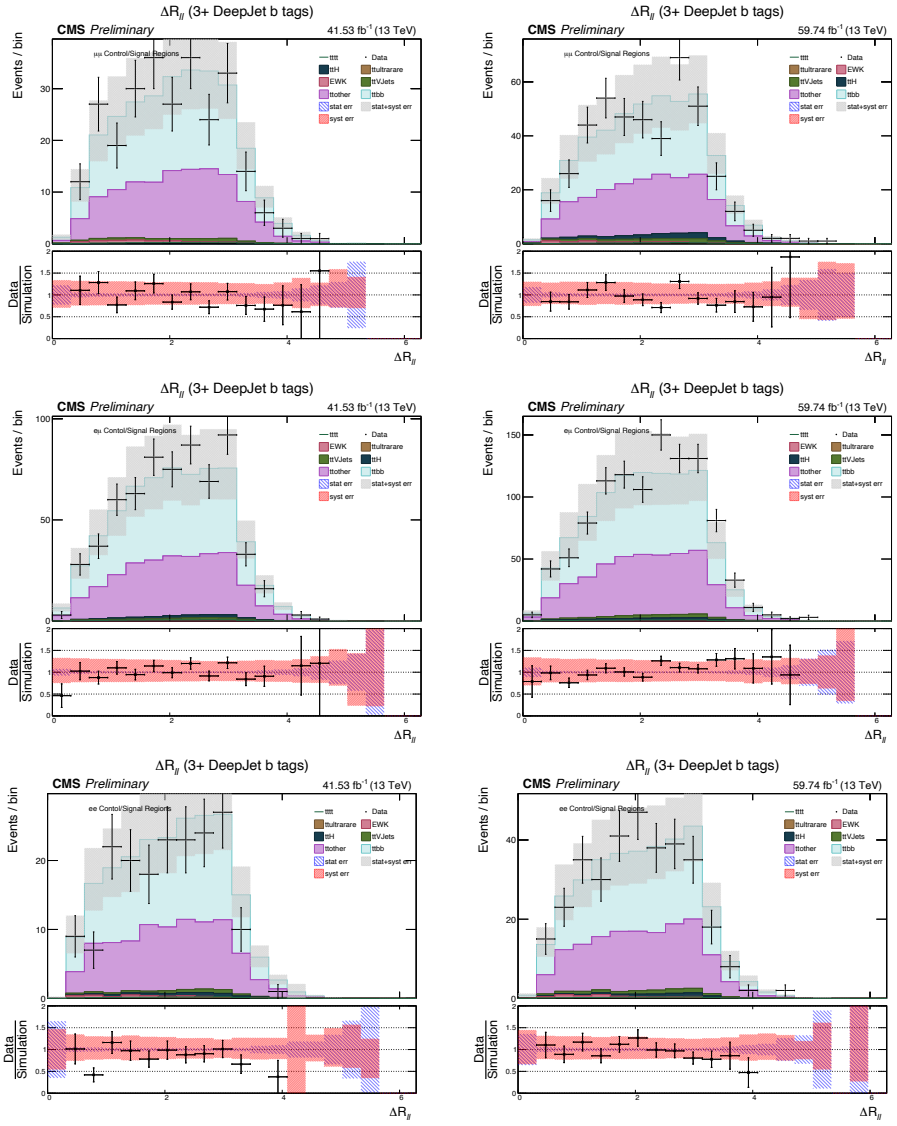

Figure 4.64: dR^{bb} for the 2017 and 2018 datasets for the three channels. The left panels are for 2017 and the right panels are for 2018. The μ μ channel is shown in the top row, the e μ channel is in the middle row and the e e channel is in the bottom row.

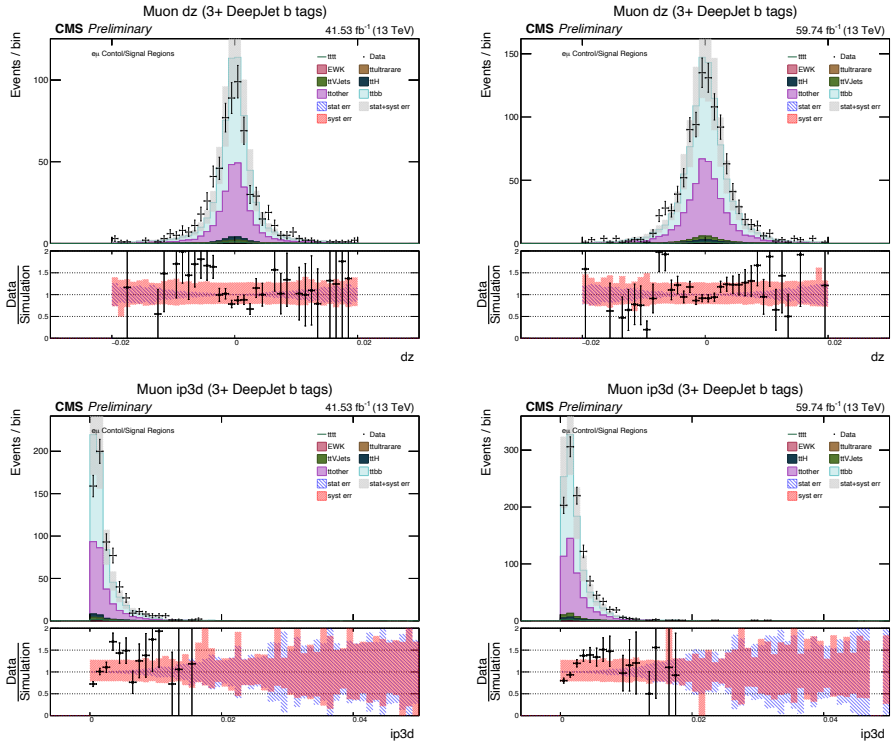

Figure 4.65: e μ muon dz and 3D impact parameter distributions, with 2017 on the left and 2018 on the right, dz on top and IP3D on bottom

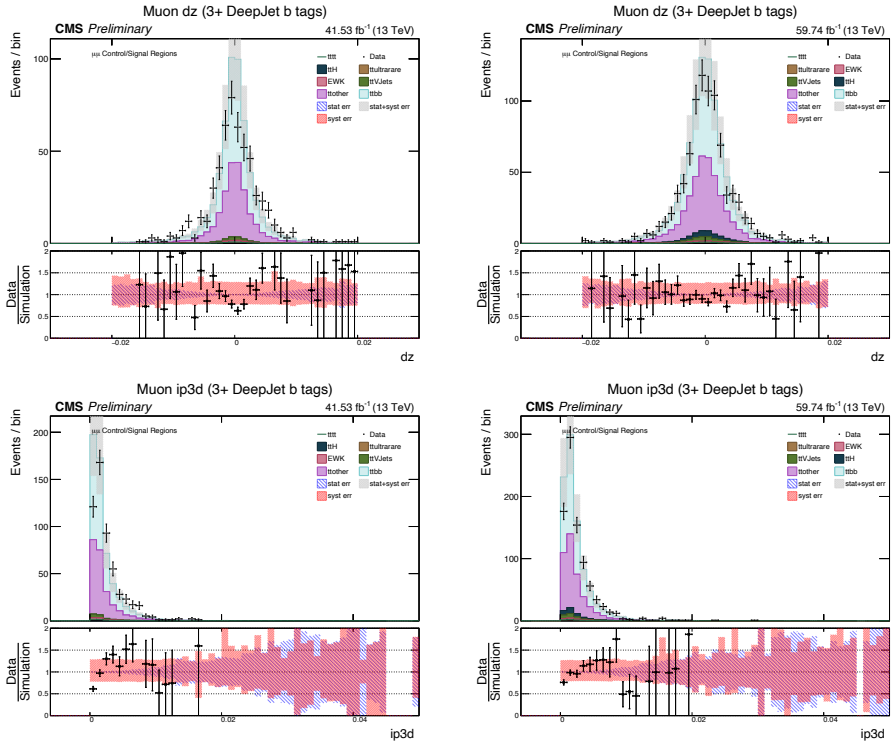

Figure 4.66: μ μ muon dz and 3D impact parameter distributions, with 2017 on the left and 2018 on the right, dz on top and IP3D on bottom

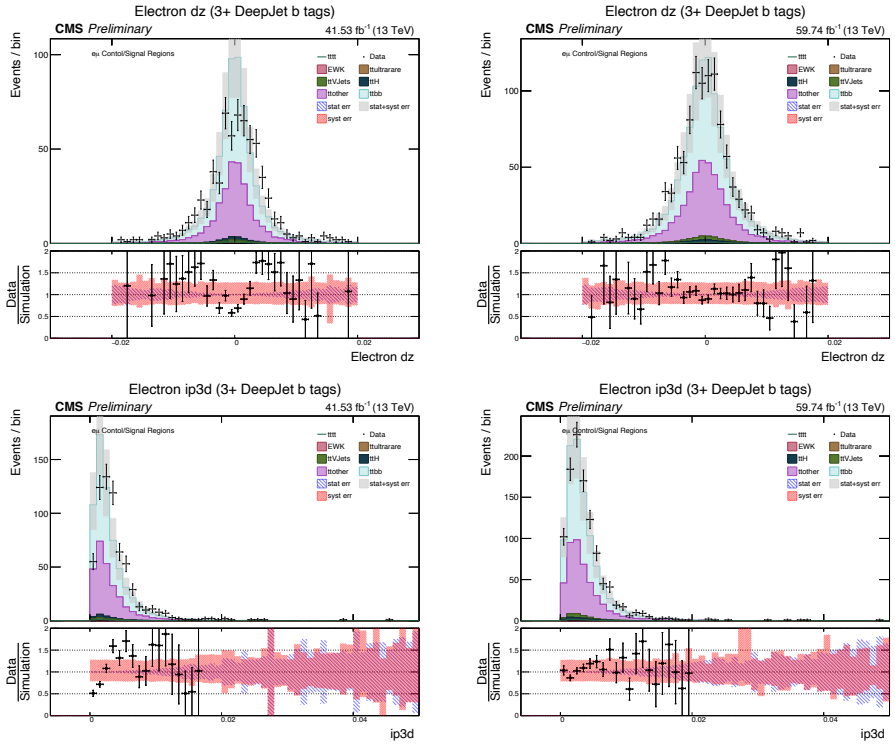

Figure 4.67: e μ electron dz and 3D impact parameter distributions, with 2017 on the left and 2018 on the right, dz on top and IP3D on bottom

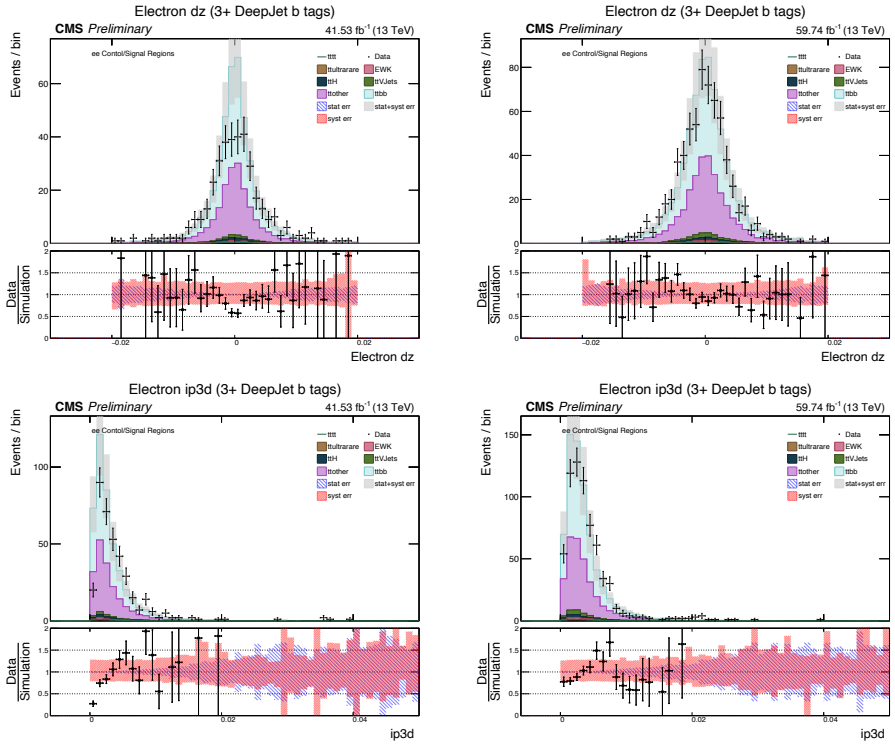

Figure 4.68: ee electron dz and 3D impact parameter distributions, with 2017 on the left and 2018 on the right, dz on top and IP3D on bottom

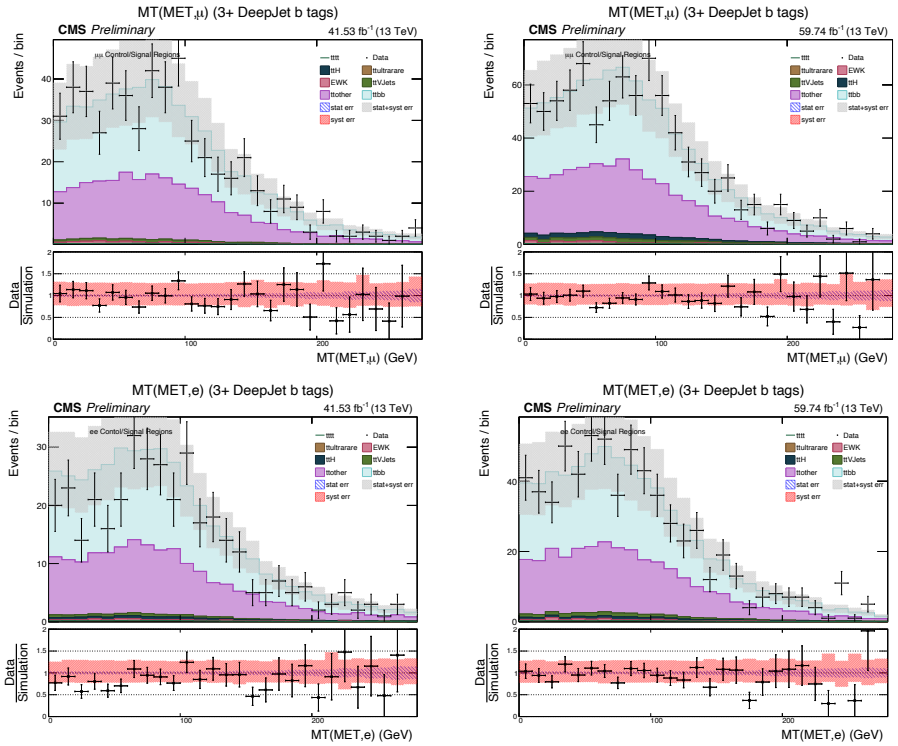

Figure 4.69: M_T for the 2017 and 2018 datasets for the ee and μ μ channels. The left panels are for 2017 and the right panels are for 2018. The μ μ channel is shown in the top row, and the e e channel is in the bottom row.

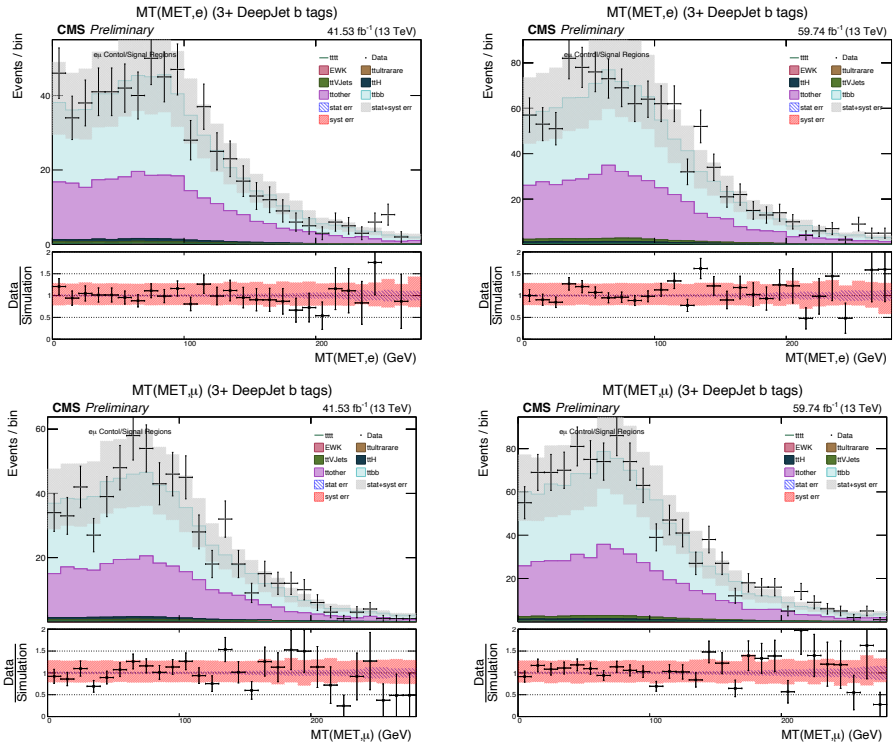

Figure 4.70: M_T for the 2017 and 2018 datasets for the e μ channel. The left panels are for 2017 and the right panels are for 2018. The combination of the e and MET is shown in the top row, and the μ and MET is in the bottom row.

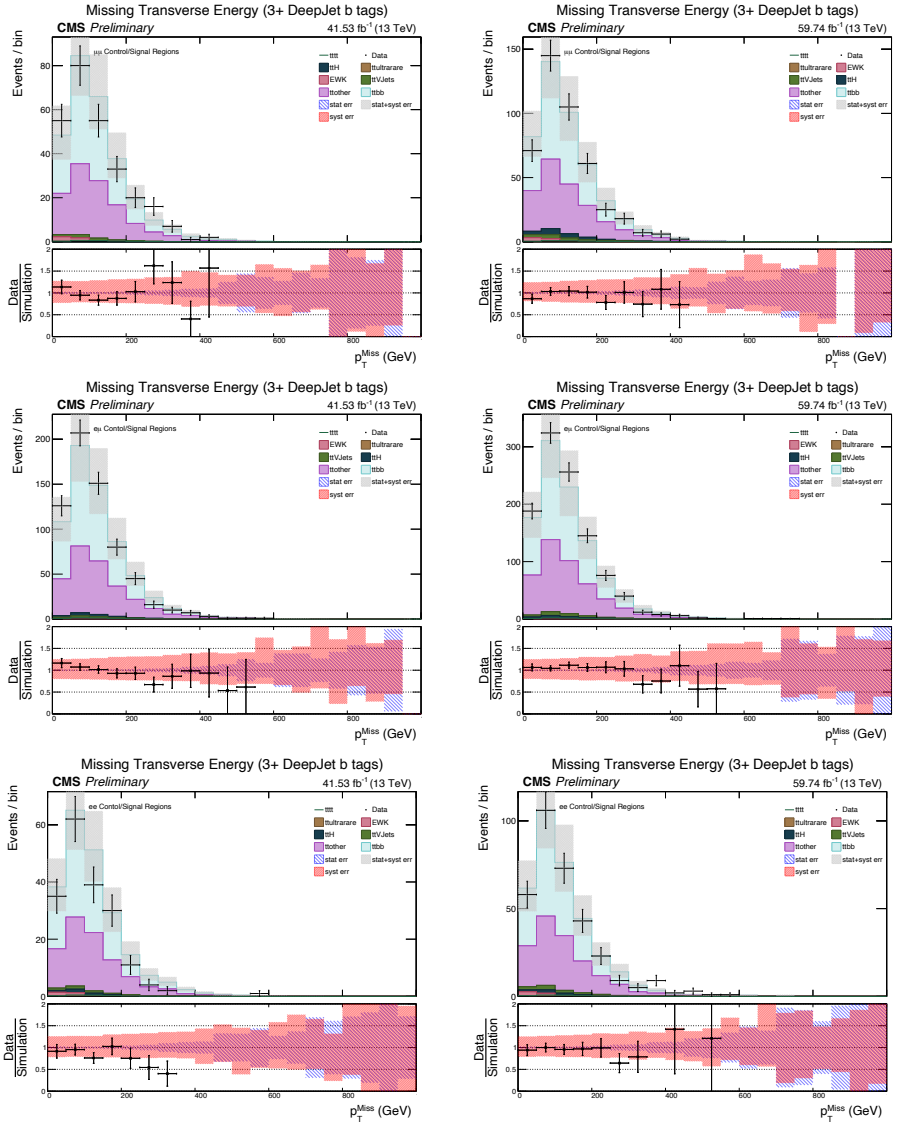

Figure 4.71: p_{T}^{Miss} for the 2017 and 2018 datasets for the three channels. The left panels are for 2017 and the right panels are for 2018. The μ μ channel is shown in the top row, the e μ channel is in the middle row and the e e channel is in the bottom row.

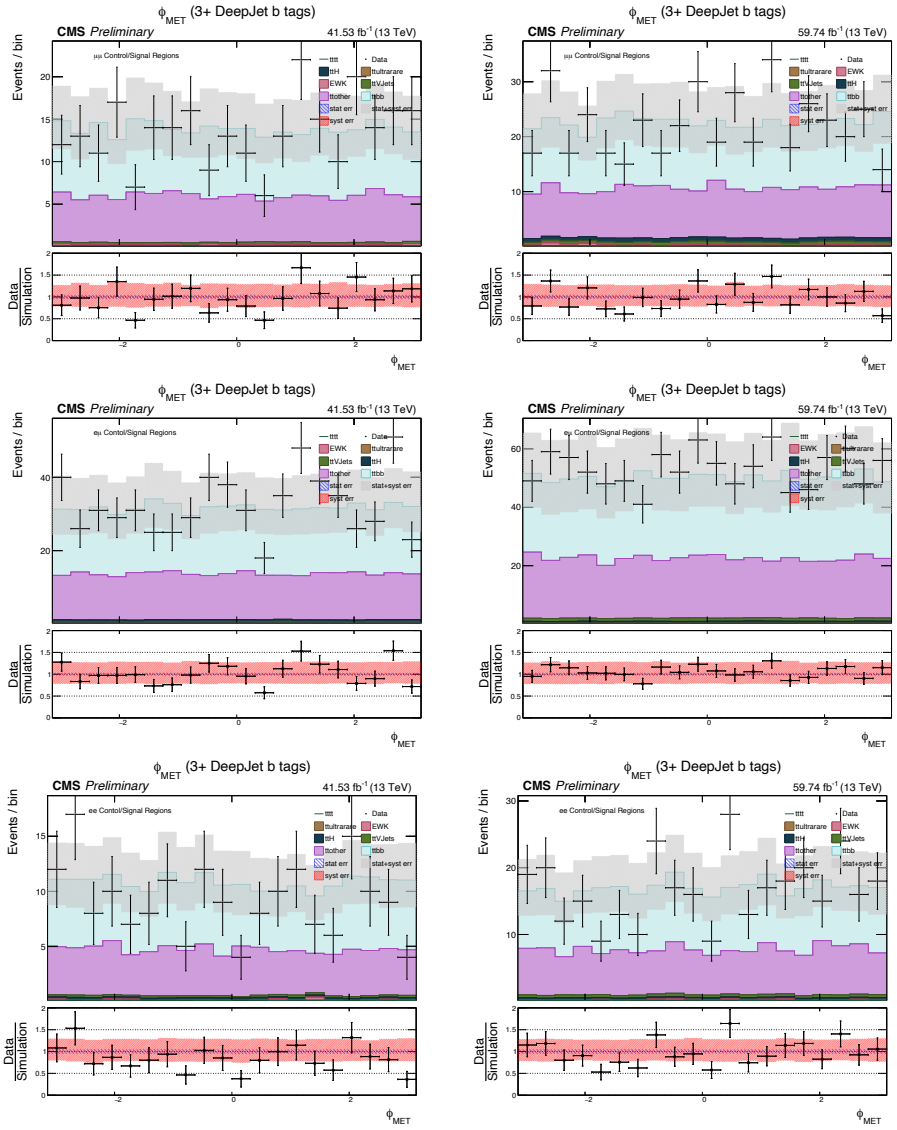

Figure 4.72: ϕ_{MET} for the 2017 and 2018 datasets for the three channels. The left panels are for 2017 and the right panels are for 2018. The μ μ channel is shown in the top row, the e μ channel is in the middle row and the e e channel is in the bottom row.

4.7 Discriminating Between the Four-Top Signal and the SM Background

The dominant background to the search for $t\bar{t}t\bar{t}$ production is $t\bar{t}$, which has a cross-section 4 orders of magnitude larger than the signal. Distinguishing characteristics of the signal versus the background are the large number of b jets (4 in the signal *vs.* 2 in the background) and the large overall number of jets (8 in the signal *vs.* 2 in the background). Of course, there are additional jets in the background due to ISR and FSR radiation, as well as from other processes such as $t\bar{t}b\bar{b}$. For these reasons, we categorize the data based on the number of jets as well as the number of b jets. The signal-rich regions are high jet, high b jet, while the background-rich regions are low jet and low b jet.

The events passing the baseline selection are categorized according to the number of jets passing the selection and the number of b tagged jets. The event categories used are 4, 5, 6, 7, and 8 or more jets and 2, 3 and 4 or more b tags, giving 15 categories per channel and year. The dominant background process in all categories is $t\bar{t}$, with the $t\bar{t}b\bar{b}$ subprocess becoming the principal background in the regions with four or more b tags. The non-$t\bar{t}$ backgrounds are very small in the low jet multiplicity regions, and the high-H_{T} cut eliminates almost all Drell-Yan and much of the single top processes (with the Z invariant mass veto doing the bulk of the removal for the former). At higher multiplicities the $t\bar{t}$ +X backgrounds are still small but have significant contributions. Of the 15 categories, the 7 and 8+ jet categories with 3 and 4+ b tags are the most sensitive to the $t\bar{t}t\bar{t}$ signal, with some additional contribution coming from the 8+ jets, 2 b tag categories. Discrimination between the $t\bar{t}$ background and the $t\bar{t}t\bar{t}$ signal is done using H_{T}.

The H_{T} distributions for each of the 15 categories are fitted to obtain a limit on the $\mathrm{t\bar{t}t\bar{t}}$ cross section. For the fits, the H_{T} distributions are plotted separately for each year and for each of the three lepton decay channels using a binning scheme which keeps the MC statistical uncertainty (using unweighted events) for the $\mathrm{t\bar{t}}$ (including $\mathrm{t\bar{t}b\bar{b}}$) background below 30% in all bins. Additionally, the bin widths are required to be close to the H_{T} resolution determined from the $\mathrm{t\bar{t}t\bar{t}}$ simulation. The details of this are in Appendix D.9. The distributions are then fitted simultaneously using the Higgs Combine tool to calculate the a-priori and a-posteriori expected limit, associated 95% confidence interval, the a-priori and a-posteriori expected significance, and the impact of all the experimental and theoretical nuisance parameters detailed in Section 4.8, using the Asimov dataset. These are provided in full to show the relative expected contributions for all lepton channels and years. For the final result, the full LHC-style test statistics are used, as not all the requirements of the Asymptotic methods are fulfilled for the analysis. The 2 b tagged jets distributions are shown in Fig. 4.73, the 3 b tagged jets distributions in Fig. 4.74, and the categories with 4+ b tags in in Fig. 4.75. The templates passed to the Higgs Combine tool are in *Events/bin* (variable width), shown in Appendix D.1, where these figures show *Events*/ GeV using a generally coarser binning that does not satisfy the template criteria in all bins, notably the far tails.

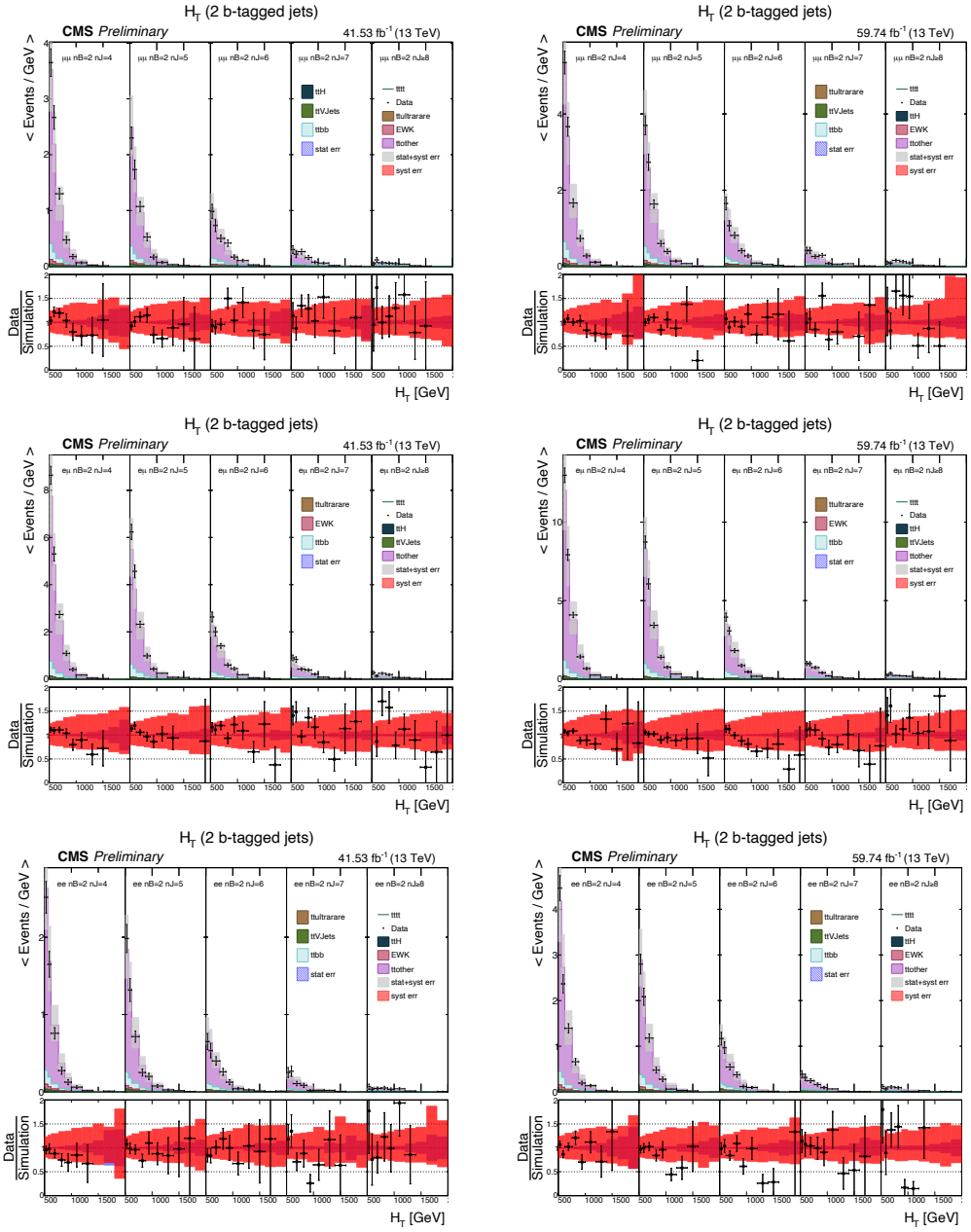

Figure 4.73: Differential, coarsely binned versions of the 2 b tags H_{T} templates passed to combine for the 2017 and 2018 datasets for the three channels. The left panels are for 2017 and the right panels are for 2018. The μ μ channel is shown in the top row, the e μ channel is in the middle row and the e e channel is in the bottom row. The templates passed to combine are in Figs. D.1, D.2, D.3

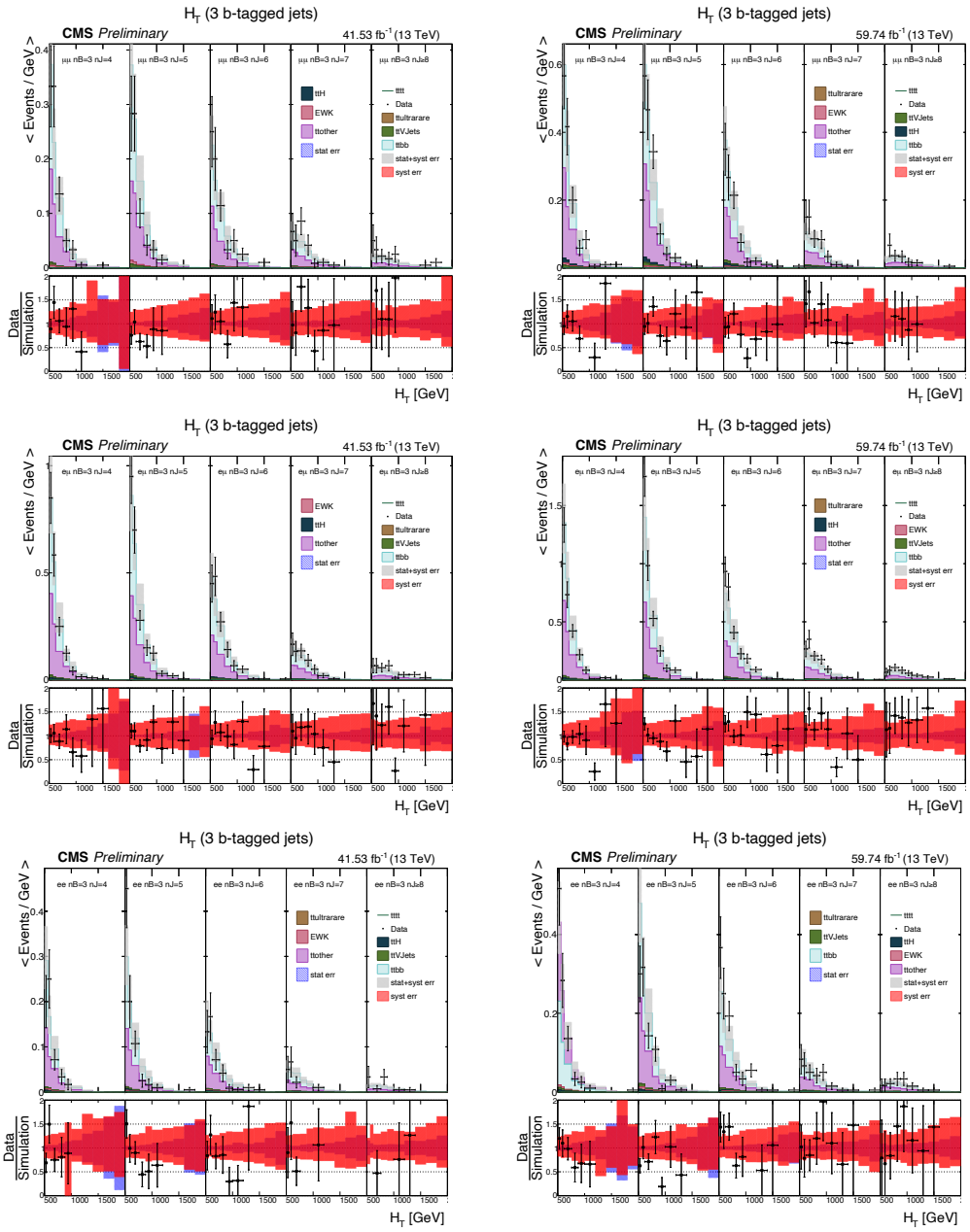

Figure 4.74: Differential, coarsely binned versions of the 3 b tags H_{T} templates passed to combine for the 2017 and 2018 datasets for the three channels. The left panels are for 2017 and the right panels are for 2018. The μ μ channel is shown in the top row, the e μ channel is in the middle row and the e e channel is in the bottom row. The templates passed to combine are in Figs. D.4, D.5, D.6

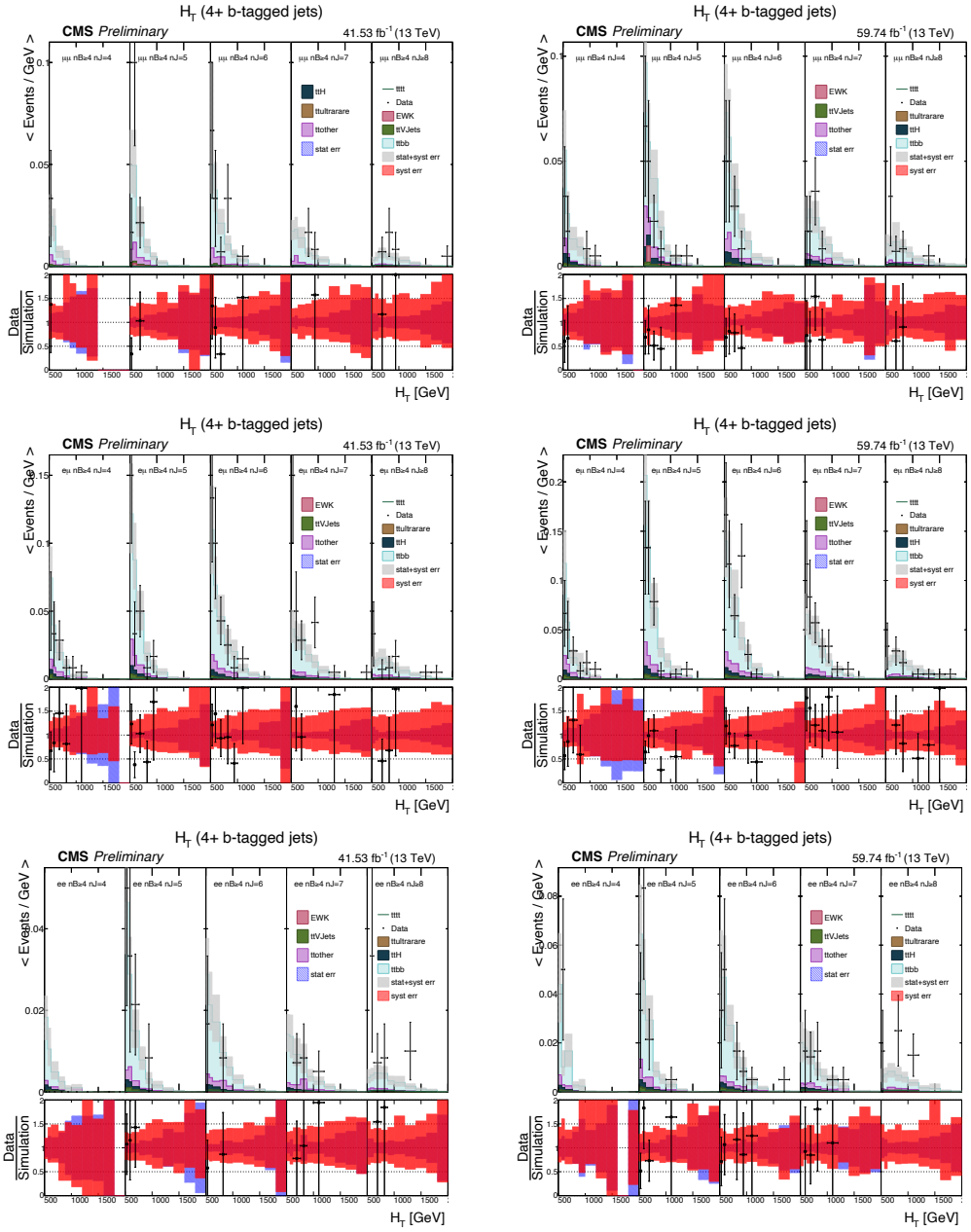

Figure 4.75: Differential, coarsely binned versions of the 4+ b tags H_{T} templates passed to combine for the 2017 and 2018 datasets for the three channels. The left panels are for 2017 and the right panels are for 2018. The μ μ channel is shown in the top row, the e μ channel is in the middle row and the e e channel is in the bottom row. The templates passed to combine are in Figs. D.7, D.8, D.9

4.8 Systematic Uncertainties

The systematic uncertainties listed in this section have been assessed on the Asimov dataset only, excluding Monte Carlo statistical uncertainties. The full list of systematic uncertainties considered is shown in Table 4.8 grouped according to the categories experimental uncertainties, cross sections, and other modeling uncertainties. A brief summary of treatment of each of these is given below. Table 4.8 also shows the type of uncertainty used in the limit setting (log normal, shape, and shape + normalization), the year of the datasets for which they are included, and the correlations which are assumed between the three decay channels and the two data years. Log-normal types are used for luminosity, process cross sections, lepton scale factors, the high-H_{T} correction nuisance, h_{damp}, and the heavy flavor uncertainty (the $\mathrm{t\bar{t}b\bar{b}}$ cross section uncertainty). The impact of most of these uncertainties is low, except the $\mathrm{t\bar{t}}$ heavy flavor uncertainty of 8%, as this subprocess constitutes the major background in the signal-rich phase-space. Due to the fit being unable to distinguish between the high-H_{T} correction nuisance, h_{damp}, and $\mathrm{t\bar{t}}$ cross section uncertainty, these have all been merged into the $\mathrm{t\bar{t}}$ cross section uncertainty. No change in the limit or significance is observed. See appendix D.5 for further discussion.

4.8.1 Experimental Effects

Jet energy scale and jet energy resolution are assessed by calculating all quantities after shifting all jet p_{T} values in the simulation up or down one standard deviation, using a reduced breakdown into 11 sources, with 6 being correlated between years and 5 uncorrelated. More details are available in Ref. [60].

For the b tagging shape calibration, 9 sources of uncertainty are considered. The effect of contamination in the heavy flavor derivation regions by non-b (u, d, s, c quarks; gluons) hadronization, and contamination in the light flavor derivation regions by b and c quark hadronization is varied by ±20% and propagated through. Linear and quadratic statistical fluctuations in data and simulation are accounted for, similarly, in the light flavor and heavy flavor scale factors. Two uncertainties for the charm jets are derived from the total systematic and statistical uncertainties on the b jets as in Ref. [61]. Finally, the systematic uncertainties on b tagging coming from the Jet Energy Scales are fully correlated with the JES uncertainties, including one for the HEM issue in 2018 (as regards the effect on jet kinematics). The b tagging uncertainties are correlated according to recommendations from the B Tagging and Vertexing (BTV) POG, and are externally correlated (depending on b tagging algorithm and correction method) with the other decay channels searching for the $t\bar{t}t\bar{t}$ process.

For electrons and muons, a separate 3% uncertainty is applied if the event contains two electrons or muons (with both being applied in the electron-muon channel at half the value). This accounts for the measured uncertainties from the POGs, plus any potential discrepancies arising from the extrapolation to $t\bar{t}t\bar{t}$ or $t\bar{t}$ events in the analysis phase space.

For pileup, the reweighting is shifted up and down according to the 4.6% uncertainty [58] on the measured distribution in data, and the resulting shape templates serve as the inputs to the fit.

A more conservative estimate of events triggering the prefire conditions, as well as the scenario in which this probability is set to 0, form the up and down systematic variation.

The luminosity uncertainty is split into correlated and uncorrelated parts according to the latest recommendations of the TOP PAG and Lumi POG, shown in table 4.9.

Due to the HEM15/16 issue, in which two HCAL modules lost power suplies after an emergency shutdown in the middle of 2018 data taking (affecting runs $\geq$ 319077, i.e. C, D, and the last certified in Run B), the jet energy measurements are impacted. To account for this, a systematic is implemented in which the energy of any jet within $-1.57 < \phi < -0.87$ and $-2.5 < \eta < -1.3$ is shifted down by 20%. This is considered the Down shift, and the Up shift is set equal to the nominal expectation producing a one-sided shape-varying uncertainty.

4.8.2 Modeling Uncertainties

The cross section uncertainties correspond to either the latest CMS measurement uncertainties or the uncertainties from the QCD calculations of the cross sections for the rare processes. The acceptance effects on the signal are accounted for by converting any shape and normalization systematic uncertainties into shape-only (meaning they are rescaled to preserve the inclusive-sample cross-section, prior to any acceptance effects), allowing differing acceptances on those variations to change the yields of the templates going into the fit.

The Matrix Element renormalization and factorization uncertainties pertain to the effect of missing higher order corrections in the ME calculation. The former is related to the effective UV cutoff, and the latter to the IR cutoff [62]. For each of the 8 simulation background groups, the renormalization, factorization, and correlated (μ_F and μ_R varied

simultaneously in the same direction) uncertainties are calculated, and contain both shape and normalization effects for the fit.

Initial/Final State Radiation (ISR/FSR) are calculated separately for the 8 process groups, and by construction from the hadronizer are shape uncertainties that do not alter the cross section prior to event selection and categorization. The process groups are defined as $t\bar{t}b\bar{b}$ and non-$t\bar{t}b(\bar{b})$ events from the $t\bar{t}$ simulation (with rarer single-top events included with the latter), electroweak (Drell-Yan), $t\bar{t}$ H, $t\bar{t}$W+Jets together with $t\bar{t}$Z+Jets, ttultrarare ($t\bar{t}$ with associated diboson production and three-top), and $t\bar{t}t\bar{t}$. ISR and FSR are the principle mechanisms by which lower jet multiplicity proceses like $t\bar{t}$ produce higher event activity and numbers of jets, akin to the $t\bar{t}t\bar{t}$ signature. The default recommendation for varying the Parton Showering up and down by 2.0 and 0.5 respectively are used to estimate the theory uncertainty. ISR and FSR are treated separately, and additionally, background processes from different mass regimes are treated separately, as ISR will not affect higher mass systems such as $t\bar{t}t\bar{t}$, $t\bar{t}$Z+Jets the same. These systematics are implemented as asymmetric log-normal uncertainties calculated individually per b-tag and jet multiplicity.

The high-H_{T} correction nuisance is set to the maximal difference between the normalization correction (common to 2017 and 2018 data taking) and either individual year, which is 4.0% (relative), and is formed into a symmetric up and down uncertainty.

The impact of h_{damp} and underlying event (UE) have been assessed and they manifest as principally normalization uncertainties. Due to the very limited size of the simulation samples available for these systematic uncertainties the uncertainties have been

estimated from inclusive simulations for ≥ 4 jets and ≥ 2 b-tags, using the combined results from the 3 decay channels. These give ratios of up/nominal and down/nominal of 1.10 and 0.93 for h_{damp} and 1.01 and 0.99 for UE, respectively where the distributions are consistent with being flat over the full H_{T} range. Due to the negligible impact of the latter, it has been removed from the fit, and the former has been merged with the $\text{t}\bar{\text{t}}$ cross section uncertainty (along with the high-H_{T} correction nuisance), as these uncertainties cannot be distinguished from eachother in the analysis phase space.

The Parton Distribution Function uncertainties are stored in NanoAOD from the PDF4LHC15_nnlo_30_pdfas set, a 30-element hessian breakdown of sources from different PDF sets, including α_s variations [63]. This PDF set combines NNPDF3.0, MMHT2014, and CT14; all set the central value of α_s to 0.118 (NLO, NNLO), and the two PDF scale variations shift α_s up and down by 0.0015. The α_s uncertainty is constructed via symmetrization of the up and down variations, and a PDF uncertainty is calculated using the prescription from the PDF4LHC15 recommendations.

The Barlow-Beeston Lite [64],[65] approach is used to account for bin-by-bin statistical uncertainties coming from the simulation (via the Higgs Combine auto-MC stat setting). In this construction, a single statistical uncertainty is assigned per bin for all of the backgrounds.

4.8.3 Systematic Uncertainty Impacts

The impacts for 2017 and 2018 combined are shown in Figs. 4.76 and 4.77. $\text{t}\bar{\text{t}}$ FSR and ISR are both highly ranked. The uncertainty on the $\text{t}\bar{\text{t}}\text{b}\bar{\text{b}}/\text{t}\bar{\text{t}}\text{b}$ subprocess cross-section is similarly important, due to the fact that this is the major background of the

signal-rich regions where 3 or more jets are b tagged. Several process factorization and renormalization scale uncertainties, jet energy scale, btag shape calibration, and inclusive process cross sections follow.

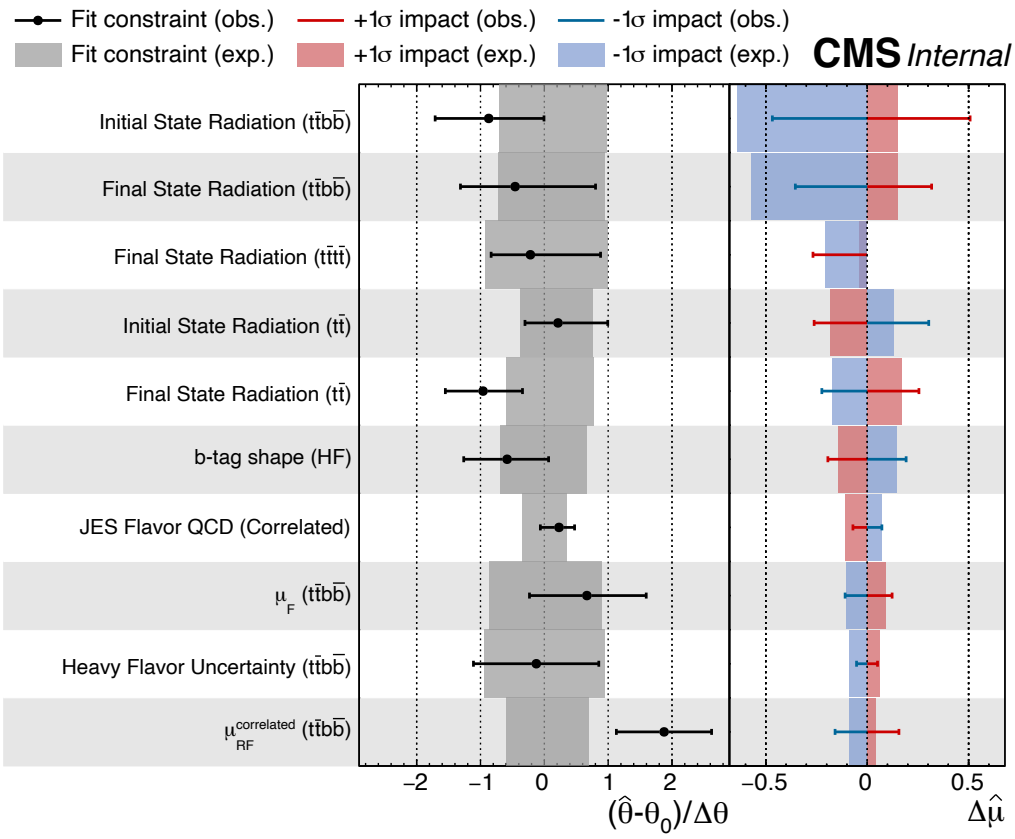

Figure 4.76: Highest impacts for the combination of 2017 and 2018 Asimov Dataset and Observed

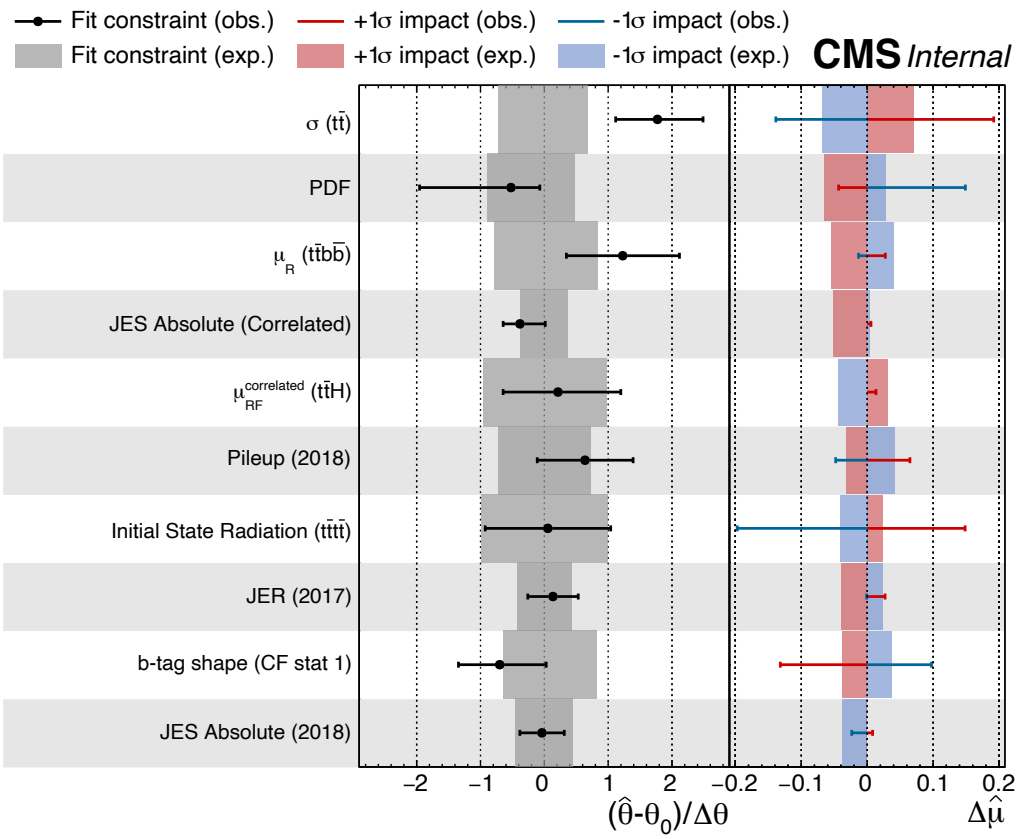

Figure 4.77: Subleading impacts for the combination of 2017 and 2018 Asimov Dataset and Observed

Table 4.8: Systematic Uncertainties for the combination of 2017 and 2018. Shape (S) and Normalization (N) specify uncertainty types; Correlated (C), Uncorrelated (U), Partially-correlated (P), e e-e μ (EEM), and μ μ-e μ (MME) indicate correlations.

Uncertainty	Type	2017	2018	Inter-year Correlation	Inter-chan. correlation
Experimental Effects					
Electron reconstruction/ID	N	3%	3%	U	EEM
Muon ID/Isolation	N	3%	3%	U	MME
Jet Energy Scale (6, C)	S	Yes	Yes	C	✓
Jet Energy Scale (5, U)	S	Yes	Yes	U	✓
Jet Energy Resolution	S	Yes	Yes	U	✓
b tagging Purity (HF, LF)	S	Yes	Yes	✓	✓
b tagging Charm (CFErr x2)	S	Yes	Yes	✓	✓
b tagging Stat. (HF x2, LF x2)	S	Yes	Yes	U	✓
Pileup	S	Yes	Yes	U	✓
Prefire	S	Yes	N/A	N/A	✓
Luminosity	N	2.3%	2.5%	P	✓
HEM	S	N/A	Yes	N/A	✓
Cross sections					
$t\bar{t}$ Activity Correction	N	5.4%	5.4%	✓	✓
$t\bar{t}$ ($+h_{\mathrm{damp}}$)	N	$^{+12.1\%}_{-10.1\%}$	$^{+12.1\%}_{-10.1\%}$	✓	✓
$t\bar{t}H$, $t\bar{t}V$	N	20%	20%	✓	✓
$t\bar{t}$ + rare	N	50%	50%	✓	✓
electroweak	N	3.8%	3.8%	✓	✓
Other modeling uncertainties					
$t\bar{t}$ heavy flavor ($t\bar{t}b\bar{b}$)	N	8.0%	8.0%	✓	✓
$t\bar{t}$, tW, $\bar{t}$W $\mu_{R,F,RF}$	S	Yes	Yes	✓	✓
$t\bar{t}$, tW, $\bar{t}$W ISR, FSR	S	Yes	Yes	✓	✓
$t\bar{t}b\bar{b}$ $\mu_{R,F,RF}$	S	Yes	Yes	✓	✓
$t\bar{t}b\bar{b}$ ISR, FSR	S	Yes	Yes	✓	✓
$t\bar{t}$ + rare $\mu_{R,F,RF}$	S	Yes	Yes	✓	✓
$t\bar{t}$ + rare ISR, FSR	S	Yes	Yes	✓	✓
electroweak $\mu_{R,F,RF}$	S	Yes	Yes	✓	✓
electroweak ISR, FSR	S	Yes	Yes	✓	✓
$t\bar{t}H$ $\mu_{R,F,RF}$	S	Yes	Yes	✓	✓
$t\bar{t}H$ ISR, FSR	S	Yes	Yes	✓	✓
$t\bar{t}V$ $\mu_{R,F,RF}$	S	Yes	Yes	✓	✓
$t\bar{t}V$ ISR, FSR	S	Yes	Yes	✓	✓
$t\bar{t}t\bar{t}$ $\mu_{R,F,RF}$	S	Yes	Yes	✓	✓
$t\bar{t}t\bar{t}$ ISR, FSR	S	Yes	Yes	✓	✓
PDF	S	Yes	Yes	✓	✓

Table 4.9: Luminosity correlation scheme (%)

Correlation group	2016	2017	2018
Uncorrelated 2016	1.0	0.0	0.0
Uncorrelated 2017	0.0	2.0	0.0
Uncorrelated 2018	0.0	0.0	1.5
Correlated 2016-2017-2018	0.6	0.9	2.0
Correlated 2017-2018	0.0	0.6	0.2

4.9 Limits on the Four Top Production Cross Section

Using the Asimov dataset, expected upper limits and significance are extracted for each year using the combined information from the three decay channels, with 95% CL intervals. The asymptotic results for the limits in terms of the signal strength ($\times\sigma^{SM}_{t\bar{t}t\bar{t}}$) are shown in Table 4.10, the same translated to fb in Table 4.11, and finally the significance in Table 4.12. The limits on the signal-strength are extracted with rate-effects on the signal, while for the cross-section limit, the rate-effects on the inclusive $t\bar{t}t\bar{t}$ cross-section are removed so that the conversion from signal-strength to cross-section is reduced to multiplication by the presumed 12fb. Limits show the apriori and aposteriori (after doing the background-fit to the data) expected results, in addition to the observed. For significance, apriori and observed results are included in the table. The 2017 and 2018 limits are comparable, as is expected based on the relative luminosities, with the eμ channel dominating. The limits and significances using the full LHC test statistics are in AN-2020/198 [4]. Table 4.13 shows the results from the published 2016 BDT analysis. The expected limits from the new analyses all show significant improvement over the expected limit from the 2016 data. It should also be noted here that the 2016 analysis used an older prediction of 9 fb for the $t\bar{t}t\bar{t}$ cross section, compared to the improved calculation of 12 fb which is used for all of the current analyses. The final result with the signal rescaled to 12fb is shown in the last row of the table.

4.9.1 Expected and Observed Limits

Table 4.10: Summary of asymptotic cross section limits for $t\bar{t}t\bar{t}$ production using the RunII dataset.

Era	Channel	Apriori limit $[\times\sigma^{SM}_{t\bar{t}t\bar{t}}]$	Aposteriori limit $[\times\sigma^{SM}_{t\bar{t}t\bar{t}}]$	Observed Limit $[\times\sigma^{SM}_{t\bar{t}t\bar{t}}]$
2017	μμ	$8.8^{+9.4}_{-4.3}$	$9.0^{+9.6}_{-4.4}$	14
2017	eμ	$6.3^{+6.4}_{-3.0}$	$6.4^{+6.4}_{-3.1}$	8.9
2017	ee	$10.9^{+12.1}_{-5.4}$	$9.5^{+10.8}_{-4.8}$	12
2017	Combined	$4.6^{+4.5}_{-2.2}$	$4.7^{+4.6}_{-2.2}$	8.2
2018	μμ	$7.9^{+8.3}_{-3.8}$	$7.9^{+8.3}_{-3.8}$	6.5
2018	eμ	$5.1^{+5.2}_{-2.4}$	$5.5^{+5.6}_{-2.6}$	7.6
2018	ee	$8.6^{+8.8}_{-4.2}$	$7.8^{+8.2}_{-3.8}$	13
2018	Combined	$3.8^{+3.6}_{-1.8}$	$3.9^{+3.8}_{-1.8}$	6.4
RunII	μμ	$5.8^{+5.8}_{-2.8}$	$5.7^{+5.9}_{-2.8}$	7.1
RunII	eμ	$3.9^{+3.8}_{-1.9}$	$4.2^{+4.2}_{-2.0}$	6.3
RunII	ee	$6.6^{+6.6}_{-3.2}$	$5.9^{+6.0}_{-2.9}$	10
RunII	Combined	$2.9^{+2.7}_{-1.4}$	$3.0^{+2.9}_{-1.4}$	6.2

Table 4.11: Summary of asymptotic cross section limits for $t\bar{t}t\bar{t}$ production using the RunII dataset.

Era	Channel	Apriori limit [fb]	Aposteriori limit [fb]	Observed Limit [fb]
2017	μμ	106^{+112}_{-52}	108^{+115}_{-53}	17
2017	eμ	75^{+76}_{-36}	77^{+76}_{-37}	110
2017	ee	131^{+145}_{-65}	114^{+130}_{-57}	140
2017	Combined	55^{+54}_{-26}	56^{+55}_{-27}	98
2018	μμ	95^{+99}_{-46}	95^{+99}_{-46}	78
2018	eμ	61^{+63}_{-29}	65^{+67}_{-32}	91
2018	ee	103^{+106}_{-50}	94^{+98}_{-46}	160
2018	Combined	45^{+44}_{-21}	46^{+46}_{-22}	77
RunII	μμ	69^{+70}_{-33}	68^{+70}_{-33}	85
RunII	eμ	47^{+46}_{-22}	51^{+51}_{-24}	76
RunII	ee	80^{+79}_{-38}	71^{+72}_{-34}	120
RunII	Combined	35^{+32}_{-16}	36^{+35}_{-17}	74

4.9.2 Expected and Observed Significance

Table 4.12: Summary of asymptotic significances for $t\bar{t}t\bar{t}$ production using the RunII dataset.

Era	Channel	Apriori significance [Std. Dev.]	Aposteriori significance [Std. Dev.]	Observed significance [Std. Dev.]
2017	μμ	0.25σ	0.25σ	1.3σ
2017	eμ	0.33σ	0.34σ	0.82σ
2017	ee	0.20σ	0.25σ	0.71σ
2017	Combined	0.43σ	0.46σ	1.7σ
2018	μμ	0.27σ	0.29σ	0.0σ
2018	eμ	0.41σ	0.40σ	0.74σ
2018	ee	0.25σ	0.28σ	1.2σ
2018	Combined	0.51σ	0.55σ	1.1σ
RunII	μμ	0.35σ	0.38σ	0.43σ
RunII	eμ	0.50σ	0.50σ	0.91σ
RunII	ee	0.31σ	0.37σ	1.6σ
RunII	Combined	0.64σ	0.69σ	1.9σ

4.9.3 2016 Expected Limits and Significance

Table 4.13: Summary of expected limits, expected cross section and signal significance of $t\bar{t}t\bar{t}$ production adding bin-to-bin statistical uncertainty on MC predictions of the 2016 OSDL analysis, using $t\bar{t}t\bar{t}$ cross section of 9 fb (the last result scales this to 12fb [4]).

Channel	Exp. limit $[\times\sigma^{SM}_{t\bar{t}t\bar{t}}]$	Expected limit [fb]	Expected signal significance [Std. Dev.]
μμ	$15.8^{+9.6}_{-5.4}$	144^{+87}_{-49}	0.16σ
eμ	$9.3^{+5.9}_{-3.3}$	85^{+54}_{-30}	0.28σ
ee	17^{+11}_{-6}	151^{+99}_{-54}	0.16σ
Combined	$7.2^{+4.4}_{-2.5}$	66^{+40}_{-23}	0.34σ
Combined ($t\bar{t}t\bar{t}$ 12fb)	$7.3^{+4.5}_{-2.5}$	66^{+40}_{-23}	0.45σ

4.10 Post-Fit Distributions

After performing the fit to data, histograms are produced sampling from the covariance matrix. The resulting histograms represent the distribution of background and signal simulation accounting for the nuisance parameters and signal-strength from the fit. In Fig. 4.78 are shown the distribution of jet multiplicities broken down by number of b-tags. The data and simulation from 2017 and 2018 across all 3 dilepton decay modes are aggregated together.

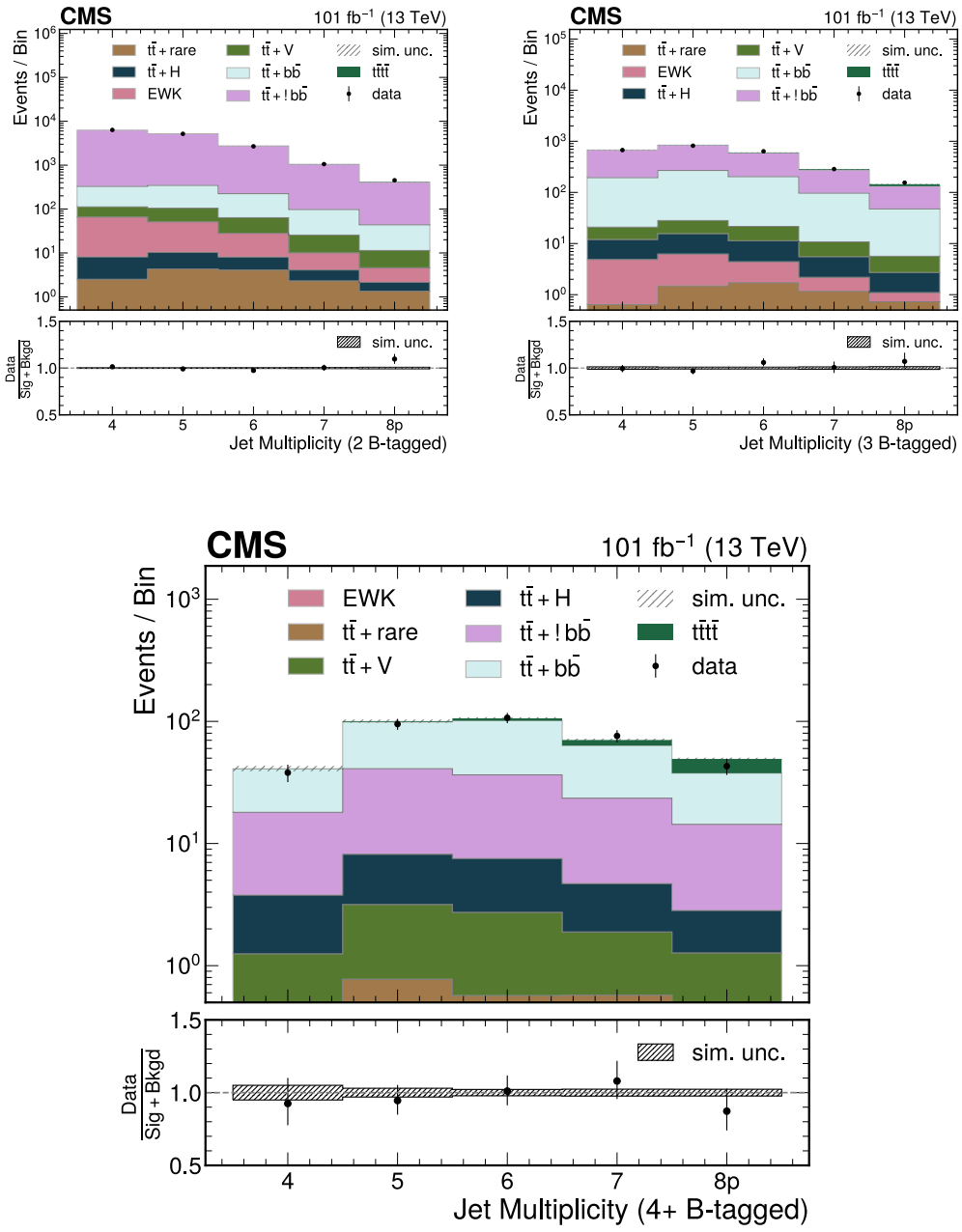

Figure 4.78: Postfit plots of jet multiplicity for the 2 b-tag (top), 3 b-tag (middle) and 4+ b-tag categories. Plots show the simulation and data stacked from 2017 and 2018 in all 3 dilepton decay channels.

Chapter 5

Summary and Conclusions

The analysis utilizes data equivalent to 101.5 fb^{-1} of luminosity at $\sqrt{s} = 13$ TeV recorded by the CMS experiment during 2017 and 2018. Following a baseline selection to select events with a pair of opposite sign leptons, four or more hadronic jets, of which at least 2 jets are b-tagged, the data are classified in bins of jet multiplicity (4, 5, 6, 7, and 8 or more jets), and b-tag multiplicity (2, 3, and 4 or more b-tagged jets). Of the 15 categories, the 7 and 8 or more jet bins, with 3 and 4 or more b-tags are the most sensitive to the $\text{t}\bar{\text{t}}\text{t}\bar{\text{t}}$ signal. The H_{T} distributions for the 15 categories are fitted to the sum of signal and background to obtain an upper limit on the the cross section $\sigma_{\text{t}\bar{\text{t}}\text{t}\bar{\text{t}}}$. An apriori expected upper limit on the four top cross section of $2.9^{+2.7}_{-1.4} \times \sigma^{SM}_{\text{t}\bar{\text{t}}\text{t}\bar{\text{t}}}$ (35^{+32}_{-16}fb) is obtained at the 95% CL with a signal significance of 0.64σ. The a-posteriori asymptotic limits, using all regions of data, are $3.0^{+2.9}_{-1.4} \times \sigma^{SM}_{\text{t}\bar{\text{t}}\text{t}\bar{\text{t}}}$ (36^{+35}_{-17}fb). The asymptotic observed limits are $6.2 \times \sigma^{SM}_{\text{t}\bar{\text{t}}\text{t}\bar{\text{t}}}$ (74fb) with observed significance 1.9σ. Finally, the full LHC statistics results are available in AN-2020/198 [4].

Bibliography

[1] CMS Collaboration, "CMS Detector Design". https://cms.cern/news/cms-detector-design. Accessed: 2022-06-06.

[2] D. Barney, "CMS Detector Slice", (Jan, 2016). CMS Collection.

[3] CMS Collaboration, "The Phase-2 Upgrade of the CMS Muon Detectors", technical report, CERN, Geneva, Sep, 2017. This is the final version, approved by the LHCC.

[4] CMS Collaboration, "Combination of four top quark production searches from 2016-2018 data.", Technical Report CMS-AN-2016/120, CERN, Geneva, 2016.

[5] R. Frederix, D. Pagani, and M. Zaro, "Large NLO corrections in $t\bar{t}W^{\pm}$ and $t\bar{t}t\bar{t}$ hadroproduction from supposedly subleading EW contributions", *JHEP* **02** (2018) 031, doi:10.1007/JHEP02(2018)031, arXiv:1711.02116.

[6] CMS Collaboration, "Search for standard model production of four top quarks in the lepton + jets channel in pp collisions at $\sqrt{s} = 8$ TeV", *JHEP* **11** (2014) 154, doi:10.1007/jhep11(2014)154, arXiv:1409.7339.

[7] CMS Collaboration, "Search for standard model production of four top quarks in proton–proton collisions at $\sqrt{s} = 13$ TeV", *Physics Letters B* **772** (2017) 336, doi:10.1016/j.physletb.2017.06.064, arXiv:1702.06164.

[8] CMS Collaboration, "Search for new physics in same-sign dilepton events in proton-proton collisions at $\sqrt{s} = 13$ TeV", *Eur. Phys. J. C.* **76** (2016) 439, doi:10.1140/epjc/s10052-016-4261-z, arXiv:1605.03171.

[9] CMS Collaboration, "Search for the production of four top quarks in the single-lepton and opposite-sign dilepton final states in proton-proton collisions at $\sqrt{s}$ = 13 TeV", *JHEP* **11** (2019) 082, doi:10.1007/JHEP11(2019)082, arXiv:1906.02805.

[10] CMS Collaboration, "Search for production of four top quarks in final states with same-sign or multiple leptons in proton-proton collisions at $\sqrt{s} = 13$ TeV", *Eur. Phys. J. C* **80** (2020) 75, doi:10.1140/epjc/s10052-019-7593-7, arXiv:1908.06463.

[11] ATLAS Collaboration, "Evidence for $t\bar{t}t\bar{t}$ production in the multilepton final state in proton-proton collisions at sqrt(s)=13 TeV with the ATLAS detector", *Eur. Phys. J.* **C80** (2020) 1085, doi:10.1140/epjc/s10052-020-08509-3.

[12] ATLAS Collaboration, "Measurement of the $t\bar{t}t\bar{t}$ production cross section in pp collisions at sqrt(s)=13 TeV with the ATLAS detector", Technical Report ATLAS-CONF-2021-013, 2021.

[13] H. P. Nilles, "Supersymmetry, Supergravity and Particle Physics", *Phys. Rept.* **110** (1984) 1.

[14] G. Farrar and P. Fayet, "Phenomenology of the Production, Decay and Detection of New Hadrionic States Associated with Supersymmery", *Phys. Lett. B* **76** (1978) 575.

[15] T. Plehn and T. M. P. Tait, "Seeking sgluons", Technical Report arXiv:0810.3919. ANL-HEP-PR-08-65. EDINBURGH 2008-43. NU-HEP-TH-08-08, Apr, 2009. Comments: published version, numerical bug removed, phenomenological results unchanged.

[16] S. Calvet, B. Fuks, P. Gris, and L. Valery, "Searching for sgluons in multitop events at a center-of-mass energy of 8 TeV", *JHEP* **04** (2013) 043, arXiv:1212.3360.

[17] D. Dicus, A. Stange, and S. Willenbrock, "Higgs decay to top quarks at hadron colliders", *Phys. Lett. B* **333** (1994) 126, arXiv:9404359.

[18] N. Craig et al., "The Hunt for the Rest of the Higgs Bosons", *JHEP* **06** (2015) 137, arXiv:1504.04630.

[19] N. Craig, J. Hajer, Y.-Y. Li, and H. Zhang, "Heavy Higgs Bosons at low tan β: from the LHC to 100 TeV", *JHEP* **01** (2017) 018, arXiv:1605.08744.

[20] J. Campbell, J. Huston, and F. Krauss, "The Black Book of Quantum Chromodynamics: A Primer for the LHC Era". Oxford University Press, 12, 2017. ISBN 978-0-19-965274-7.

[21] M. E. Peskin and D. V. Schroeder, "An Introduction to Quantum Field Theory". Westview Press, 1995. Reading, USA: Addison-Wesley (1995) 842 p.

[22] B. Martin and G. Shaw, "Particle Physics". Manchester Physics Series. Wiley, 2013. ISBN 9781118681664.

[23] F. Halzen, A. Martin, J. Wiley, and Sons, "Quarks and Leptons: An Introductory Course in Modern Particle Physics". Wiley, 1984. ISBN 9780471887416.

[24] G. 't Hooft, "Renormalization and gauge invariance", 2009.

[25] J. M. Cornwall, D. N. Levin, and G. Tiktopoulos, "Derivation of gauge invariance from high-energy unitarity bounds on the s matrix*", *Physical Review D. Vol 10, Issue 4* (15 August 1974).

[26] J. Goldstone, "Field theories with "superconductor" solutions", *Nuovo Cimento* **19** (Aug, 1960) 154–164, doi:10.1007/BF02812722.

[27] S. L. Glashow, "Partial-symmetries of weak interactions", *Nuclear Physics* **22** (1961), no. 4, 579–588, doi:https://doi.org/10.1016/0029-5582(61)90469-2.

[28] J. Goldstone, A. Salam, and S. Weinberg, "Broken symmetries", *Phys. Rev.* **127** (Aug, 1962) 965–970, doi:10.1103/PhysRev.127.965.

[29] F. Englert and R. Brout, "Broken Symmetry and the Mass of Gauge Vector Mesons", *Phys. Rev. Lett.* **13** (August, 1964) 321–323, doi:10.1103/PhysRevLett.13.321.

[30] P. W. Higgs, "Broken symmetries and the masses of gauge bosons", *Phys. Rev. Lett.* **13** (Oct, 1964) 508–509, doi:10.1103/PhysRevLett.13.508.

[31] G. S. Guralnik, C. R. Hagen, and T. W. B. Kibble, "Global conservation laws and massless particles", *Phys. Rev. Lett.* **13** (Nov, 1964) 585–587, doi:10.1103/PhysRevLett.13.585.

[32] "The standard model - theory: The goldstone boson equivalence theorem", 2015. www.perimeterinstitute.ca/videos/standard-model-theory-goldstone-boson-equivalence-theorem.

[33] A. D. Sakharov, "Violation of CP invariance, C asymmetry, and baryon asymmetry of the universe", *Soviet Physics Uspekhi* **34** (may, 1991) 392–393, doi:10.1070/pu1991v034n05abeh002497.

[34] J. H. Christenson, J. W. Cronin, V. L. Fitch, and R. Turlay, "Evidence for the 2π decay of the k_2^0 meson", *Phys. Rev. Lett.* **13** (Jul, 1964) 138–140, doi:10.1103/PhysRevLett.13.138.

[35] BABAR Collaboration, "Measurement of CP-violating asymmetries in b^0 decays to CP eigenstates", *Phys. Rev. Lett.* **86** (Mar, 2001) 2515–2522, doi:10.1103/PhysRevLett.86.2515.

[36] Belle Collaboration, "Observation of large CP violation in the neutral B meson system", *Phys. Rev. Lett.* **87** (Aug, 2001) 091802, doi:10.1103/PhysRevLett.87.091802.

[37] LHCb Collaboration, "Observation of cp violation in charm decays", *Phys. Rev. Lett.* **122** (May, 2019) 211803, doi:10.1103/PhysRevLett.122.211803.

[38] Q. R. Ahmad et al., "Direct evidence for neutrino flavor transformation from neutral-current interactions in the sudbury neutrino observatory", *Physical review letters* **89** (2002), no. 1, 011301.

[39] Y. Fukuda et al., "Evidence for oscillation of atmospheric neutrinos", *Physical review letters* **81** (1998), no. 8, 1562.

[40] R. Wald, "General Relativity". University of Chicago Press, 2010. ISBN 9780226870373.

[41] D. Clowe et al., "A Direct Empirical Proof of the Existence of Dark Matter", *APJL* **648** (September, 2006) L109–L113, doi:10.1086/508162, arXiv:astro-ph/0608407.

[42] V. C. Rubin, "The rotation of spiral galaxies", *Science* **220** (1983), no. 4604, 1339–1344, doi:10.1126/science.220.4604.1339, arXiv:https://www.science.org/doi/pdf/10.1126/science.220.4604.1339.

[43] R. Massey, T. Kitching, and J. Richard, "The dark matter of gravitational lensing", *Reports on Progress in Physics* **73** (jul, 2010) 086901, doi:10.1088/0034-4885/73/8/086901.

[44] P. J. E. Peebles and B. Ratra, "The cosmological constant and dark energy", *Rev. Mod. Phys.* **75** (Apr, 2003) 559–606, doi:10.1103/RevModPhys.75.559.

[45] S. Perlmutter et al., "Measurements of Ω and Λ from 42 High-Redshift Supernovae", *APJ* **517** (June, 1999) 565–586, doi:10.1086/307221, arXiv:astro-ph/9812133.

[46] CMS CollaborationG. L. Bayatian, et al., "CMS Physics: Technical Design Report Volume 1: Detector Performance and Software". Technical design report. CMS. CERN, Geneva, 2006. There is an error on cover due to a technical problem for some items.

[47] LHCb Collaboration, "The LHCb Detector". https://lhcb-outreach.web.cern.ch/detector/. Accessed: 2022-06-21.

[48] ALICE Collaboration, "ALICE". https://home.cern/science/experiments/alice. Accessed: 2022-06-21.

[49] A. Dominguezet al., "CMS Technical Design Report for the Pixel Detector Upgrade", technical report, Sep, 2012. Additional contacts: Jeffrey Spalding, Fermilab, Jeffrey.Spalding@cern.ch Didier Contardo, Universite Claude Bernard-Lyon I, didier.claude.contardo@cern.ch.

[50] J. Manset al., "CMS Technical Design Report for the Phase 1 Upgrade of the Hadron Calorimeter", technical report, Sep, 2012. Additional contact persons: Jeffrey Spalding, Fermilab, spalding@cern.ch, Didier Contardo, Universite Claude Bernard-Lyon I, contardo@cern.ch.

[51] CMS EGM POG, "Cut Based Electron ID for Run 2", 2020. https://twiki.cern.ch/twiki/bin/viewauth/CMS/CutBasedElectronIdentificationRun2.

[52] CMS Muon POG, "Baseline muon selections for Run-II", 2019. https://twiki.cern.ch/twiki/bin/view/CMS/SWGuideMuonIdRun2.

[53] CMS Collaboration, "Identification of heavy-flavour jets with the CMS detector in pp collisions at 13 TeV", *JINST* **13** (2018) P05011, doi:10.1088/1748-0221/13/05/P05011, arXiv:1712.07158.

[54] E. Bols et al., "Jet flavour classification using DeepJet", *JINST* **15** (Dec, 2020) P12012, doi:10.1088/1748-0221/15/12/p12012, arXiv:2008.10519.

[55] CMS Collaboration, "Performance of the DeepJet b tagging algorithm using 41.9/fb of data from proton-proton collisions at 13 TeV with Phase 1 CMS detector", CMS Detector Performance Note CMS-DP-2018-058, 2018.

[56] CMS Btag POG, "Recommendation for Using b-tag Objects in Physics Analyses", 2021. https://twiki.cern.ch/twiki/bin/viewauth/CMS/BtagRecommendation.

[57] CMS Collaboration, Btag POG, "Methods to apply b-tagging efficiency scale factors", 2020. https://twiki.cern.ch/twiki/bin/view/CMS/BTagSFMethods.

[58] CMS, "Utilities for accessing pileup information for data", Technical Report , 2015.

[59] CMS Collaboration, "Measurement of the cross section of top quark pair production with additional charm jets using the dileptonic final state in pp collisions at $\sqrt{s}$ = 13 TeV", Technical Report CMS-PAS-TOP-20-003, CERN, Geneva, 2020.

[60] CMS Collaboration, "Jet energy scale uncertainty sources", Technical Report , 2020.

[61] CMS Btag POG, "B-tagging discriminant shape calibration using event weights with a tag-and-probe method ", 2021. https://twiki.cern.ch/twiki/bin/viewauth/CMS/BTagShapeCalibration.

[62] J. Chýla, "Renormalization and factorization scale analysis of production in antiproton-proton collisions", *Journal of High Energy Physics* **2003** (Mar, 2003) 042–042, doi:10.1088/1126-6708/2003/03/042.

[63] J. Butterworth et al., "Pdf4lhc recommendations for lhc run ii", *Journal of Physics G: Nuclear and Particle Physics* **43** (Jan, 2016) 023001, doi:10.1088/0954-3899/43/2/023001.

[64] R. Barlow and C. Beeston, "Fitting using finite monte carlo samples", *Computer Physics Communications* **77** (1993), no. 2, 219–228, doi:https://doi.org/10.1016/0010-4655(93)90005-W.

[65] J. S. Conway, "Incorporating nuisance parameters in likelihoods for multisource spectra", 2011. doi:10.48550/ARXIV.1103.0354, https://arxiv.org/abs/1103.0354.

[66] .

[67] CMS Collaboration, "Commissioning of the particle-flow event reconstruction with the first LHC collisions recorded in the CMS detector", CMS Physics Analysis Summary CMS-PAS-PFT-10-001, 2010.

[68] CMS, "Top pag: Top reference selections for analyses using 2012 data", Technical Report , 2013.

[69] CMS Collaboration, "Search for standard model production of four top quarks with same-sign and multilepton final states in proton-proton collisions at $\sqrt{s} = 13$ TeV", arXiv:1710.10614. Submitted to *Eur. Phys. J. C*.

[70] CMS Collaboration, "Search for physics beyond the standard model in events with two leptons of same sign, missing transverse momentum, and jets in proton-proton collisions at sqrt(s)=13 TeV", *Eur. Phys. J.* **C77** (2017), no. 9, 578, doi:10.1140/epjc/s10052-017-5079-z, arXiv:1704.07323.

[71] CMS Collaboration, "Measurement of jet activity in ttbar events using l+jets final states at $\sqrt{s} = 13$ TeV",.

[72] CMS Collaboration, "Measurements of $t\bar{t}$ cross sections in association with b jets and inclusive jets and their ratio using dilepton final states in pp collisions at $\sqrt{s} =$ 13 TeV",.

[73] R. J. Barlow and C. Beeston, "Fitting using finite monte carlo samples", *Comput. Phys. Commun* **77** (1993) 219–228, doi:10.1016/0010-4655(93)90005-W.

[74] CMS Collaboration, "Measurement of the differential cross section for $t\bar{t}$ production in the dilepton final state at $\sqrt{s} = 13$ TeV", CMS Physics Analysis Summary CMS-PAS-TOP-16-011, 2016.

[75] Till Arndt, Carmen Diez Pardos, Andreas Meyer, "Measurement of the 2016 Trigger Efficiencies for a dilepton selection for a ttbar analysis",. CMS AN-16-392.

[76] W. Buchmüller and D. Wyler, "Constraints on SU(5)-type leptoquarks", *Phys. Lett. B* **177** (1986) 377, doi:10.1016/0370-2693(86)90771-9.

[77] B. Grzadkowski et al., "Dimension-Six Terms in the Standard Model Lagrangian", *JHEP* **10** (2010) 085, doi:10.1007/JHEP10(2010)085, arXiv:1008.4884.

[78] C. Zhang, "Constraining *qqtt* operators from four-top production: a case for enhanced EFT sensitivity", arXiv:1708.05928.

[79] CMS Collaboration, "Measurement of the inclusive and differential tt production cross sections in lepton + jets final states at 13 TeV", Technical Report CMS-PAS-TOP-16-008, CERN, Geneva, 2016.

[80] C. M. POG, "Cms muon pog run2", Technical Report , CMS, 2017.

[81] CMS Collaboration, "Calibration of the Combined Secondary Vertex b-Tagging discriminant using dileptonic ttbar and Drell-Yan events", Technical Report CMS-NOTE-2013-130, CERN, Geneva, 2013.

[82] CMS Collaboration, "Search for $t\bar{t}H$ production in the $H \to b\bar{b}$ decay channel with $\sqrt{s} = 13$ TeV pp collisions at the CMS experiment", Technical Report CMS-PAS-HIG-16-004, CERN, Geneva, 2016.

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

[84] M. Aldaya, "Search for tth, h to bb decays using the full 2016 data sample", CMS Note 2017/063, 2017.

[85] L. Ceard, "Measurement of the production cross sections of z bosons in association with b jets in pp collisions atsqrt(s) = 7 tev with the cms detector", CMS Note 2013/130, 2013.

[86] J. Andrea et al., "Search for standard model production of a single top associated with a z boson and a quark in three-lepton events at the lhc at $\sqrt{s} = 13$ tev", CMS Note 2016/276, 2017.

[87] M. Aldayaet al., "Discovery potential and search strategy for the standard model Higgs boson in the $H \to ZZ^{\star} \to 4\mu$ decay channel using a mass-independent analysis", CMS Note 2006/106, 2006.

[88] CMS Collaboration, "First measurement of the cross section ratio $\sigma_{ttbb}/sigma_{ttjj}$ using the dilepton final states in pp collisions at $\sqrt{s} = 13$ TeV",.

[89] "Cms luminosity based on pixel cluster counting - summer 2012 update", Technical Report CMS-PAS-LUM-12-001, CERN, Geneva, 2012.

[90] P. Biallasset al., "Parton distribution uncertainty determination within cmssw", CMS Note 2009/048, 2009.

[91] CMS Collaboration, "CMS technical design report, volume II: Physics performance", *J. Phys. G* **34** (2007) 995–1579, doi:10.1088/0954-3899/34/6/S01.

[92] CMS Collaboration, "Jet performance in pp collisions at $\sqrt{s}$=7 TeV", CMS Physics Analysis Summary CMS-PAS-JME-10-003, 2010.

[93] CMS Collaboration, "Search for susy in same-sign dilepton events at 13 tev", CMS Physics Analysis Summary CMS-PAS-SUS-16-020, 2016.

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

[95] Czakon, Michal and Fiedler, Paul and Mitov, Alexander, "The total top quark pair production cross-section at hadron colliders through o(α_s^4)", (2013). arXiv:1303.6254.

[96] G. Bevilacqua, M. Worek, "Constraining BSM Physics at the LHC: Four top final states with NLO accuracy in perturbative QCD", (2012). arXiv:1206.3064.

[97] Vernon Barger, Wai-Yee Keung, Brian Yencho, "Triple-top signal of new physics at the lhc", (2010). arXiv:1001.0221.

[98] M. Toharia and J. D. Wells, "Gluino decays with heavier scalar superpartners", *JHEP* **0602** (2006) 015, doi:10.1088/1126-6708/2006/02/015, arXiv:hep-ph/0503175.

[99] N. Bartosik, J. Hauk, and A. Geiser, "Identification of the origin of heavy flavour jets at generator level in monte carlo simulations at cms", Technical Report AN-14-093, DESY, 2014.

[100] "Jet energy resolution". Acessed: 2017-05-20. https://twiki.cern.ch/twiki/bin/view/CMS/JetResolution#Smearing_procedure.

[101] ALEPH, CDF, D0, DELPHI, L3, OPAL, SLD Collaborations, the LEP Electroweak Working Group, the Tevatron Electroweak Working Group, and the SLD Electroweak and Heavy Flavour Groups, "Precision electroweak measurements and constraints on the Standard Model", (2010). arXiv:1012.2367.

[102] I. Bertramet al., "A recipe for the construction of confidence limits", Technical Report TM-2104, Fermilab, 2000.

[103] CMS Top group, "Top systematic uncertainties wiki", Technical Report , 2013.

[104] M. Goerner, "Top p_{T} reweighing wiki", Technical Report , 2013.

[105] CMS JME POG, "Instructions for applying jet id", Technical Report , 2017.

[106] CMS Muon POG, "Cms muon reference efficiencies", Technical Report , 2015.

[107] CMS, "Cms miniaod documentation", Technical Report , 2015.

[108] P. Biallass et al., "Parton distribution uncertainty determination within cmssw", Technical Report AN-2009-48, CMS, 2009.

[109] R. D. Cousins, "Generalization of chisquare goodness-of-fit test for binned data using saturated models with application to histograms", technical report, 2013.

[110] A. H. et. al, "Tmva - toolkit for multivariate data analysis", Technical Report arXiv:physics/0703039, CERN, 2009.

[111] CMS Collaboration, "Methods to apply b-tagging efficiency scale factors", Technical Report , 2013.

[112] "Inclusive search for a sequential fourth generation of quarks", Technical Report CMS-PAS-EXO-11-098, CERN, Geneva, 2012.

[113] L. Monetaet al., "The RooStats Project", in *13^{th} International Workshop on Advanced Computing and Analysis Techniques in Physics Research (ACAT2010)*. SISSA, 2010. arXiv:1009.1003. PoS(ACAT2010)057.

[114] CMS Collaboration, "Search for pair production of first-generation scalar leptoquarks in pp collisions at $\sqrt{s} = 7$ TeV", (2010). arXiv:1012.4031. Submitted to *Phys. Rev. Lett.*

[115] K. Cranmeret al., "Histfactory: A tool for creating statistical models for use with roofit and roostats", Technical Report 2012-016, CERN, 2012.

[116] M. Cacciari, G. P. Salam, and G. Soyez, "Fitting using finite monte carlo samples", *Comput.Phys.Commun* **77** (2008) 219–228, doi:doi:10.1016/0010-4655(93)90005-W.

[117] CMS Collaboration, "Performance of cms muon reconstruction in pp collision events at $\sqrt{s} = 7\,\text{TeV}$", (2012). arXiv:1206.4071. Submitted to *J. Inst.*

[118] ATLAS Collaboration, "Search for the Higgs boson in the $H \rightarrow \text{WW}(*) \rightarrow \ell^+\nu\ell^-\bar{\nu}$ decay channel in pp collisions at $\sqrt{s} = 7$ TeV with the ATLAS detector", (2011). arXiv:1112.2577. Submitted to *Phys. Rev. Lett.*

[119] M. Cacciari, G. P. Salam, and G. Soyez, "The anti-k_t jet clustering algorithm", *JHEP* **04** (2008) 063, doi:10.1088/1126-6708/2008/04/063, arXiv:0802.1189.

[120] G. Bevilacqua et al., "Assault on the nlo wishlist: pp $\rightarrow$ tt bb", *JHEP* **0909** (2009) doi:10.1088/1126-6708/2009/09/109, arXiv:arXiv/0907.4723.

[121] A. Bredenstein, A. Denner, S. Dittmaier, and S. Pozzorini, "Nlo qcd corrections to top anti-top bottom anti-bottom production at the lhc: 2. full hadronic results", *JHEP* **1003** (2010) arXiv:arXiv/1001.4006.

[122] M. Worek, "On the next-to-leading order qcd k-factor for top anti-top bottom anti-bottom production at the tevatron", *JHEP* **1202** (2011) doi:10.1007/JHEP02(2012)043, arXiv:arXiv/1112.4325.

[123] A. Bredenstein, A. Denner, S. Dittmaier, and S. Pozzorini, "Next-to-leading order qcd corrections to $pp \rightarrow t\bar{t}b\bar{b} + x$ at the lhc", *Phys. Rev. Lett.* **103** (Jul, 2009) 012002, doi:10.1103/PhysRevLett.103.012002.

[124] A. Bredenstein, A. Denner, S. Dittmaier, and S. Pozzorini, "Next-to-leading order qcd corrections to $pp \rightarrow t\bar{t}b\bar{b} + x$ at the lhc", *JHEP* **108** (2008) doi:10.1088/1126-6708/2008/08/108.

[125] M. Worek, "Next-to-leading order qcd corrections to $t\bar{t}b\bar{b}$ production at the lhc", *Acta Phys.Polon.B* **40** (2009) 2937, doi:10.1088/1126-6708/2008/08/108.

[126] G. Bevilacqua, M. Czakon, C. G. Papadopoulos, and M. Worek, "Hadronic top-quark pair production in association with two jets at next-to-leading order qcd", *Phys. Rev. D* **84** (Dec, 2011) 114017, doi:10.1103/PhysRevD.84.114017.

[127] J. Aguilar-Saavedra et al., "Interpreting top-quark LHC measurements in the standard-model effective field theory", arXiv:1802.07237.

[128] A. Alloul et al., "FeynRules 2.0 - A complete toolbox for tree-level phenomenology", *Comput. Phys. Commun.* **185** (2014) 2250–2300, doi:10.1016/j.cpc.2014.04.012, arXiv:1310.1921.

[129] NNPDF Collaboration, "Parton distributions for the LHC Run II", *JHEP* **04** (2015) 040, doi:10.1007/JHEP04(2015)040, arXiv:1410.8849.

[130] R. Estalella and F. Sala, "On the projections of an m-dimensional ellipsoid", electronic note, 2016.

[131] J. Aguilar-Saavedra and J. Santiago, "Four tops and the t tbar forward-backward asymmetry", *Phys.Rev.* **D85** (2012) 034021, doi:10.1103/PhysRevD.85.034021, arXiv:1112.3778.

[132] W. Buchmuller and D. Wyler, "Effective Lagrangian Analysis of New Interactions and Flavor Conservation", *Nucl. Phys.* **B268** (1986) 621–653, doi:10.1016/0550-3213(86)90262-2.

[133] D. Goncalves-Netto et al., "Sgluon Pair Production to Next-to-Leading Order", *Phys.Rev.* **D85** (2012) 114024, doi:10.1103/PhysRevD.85.114024, arXiv:1203.6358.

[134] The ATLAS Collaboration, The CMS Collaboration, The LHC Higgs Combination Group Collaboration, "Procedure for the LHC Higgs boson search combination in Summer 2011", Technical Report CMS-NOTE-2011-005. ATL-PHYS-PUB-2011-11, CERN, Geneva, Aug, 2011.

[135] S. Chatrchyan et al., "Measurement of the production cross section in pp collisions at with lepton + jets final states", *Physics Letters B* **720** (2013), no. 1?3, 83 – 104, doi:http://dx.doi.org/10.1016/j.physletb.2013.02.021.

[136] CMS Collaboration, "Measurement of the $t\bar{t}$ production cross section in pp collisions at $\sqrt{s} = 7$ TeV with lepton + jets final states", *Phys. Lett. B* **720** (Dec, 2012) 83–104. 31 p.

[137] S. Chatrchyan et al., "Search for heavy quarks decaying into a top quark and a w or z boson using lepton + jets events in pp collisions at $\sqrt{s} = 7\,\text{TeV}$", *Journal of High Energy Physics* **2013** (2013), no. 1, 1–30, doi:10.1007/JHEP01(2013)154.

[138] CMS Collaboration, "Search for heavy quarks decaying into a top quark and a W or Z boson using lepton + jets events in pp collisions at $\sqrt{s} = 7$ TeV", *J. High Energy Phys.* **01** (Oct, 2012) 154. 29 p. Comments: Submitted to the Journal of High Energy Physics.

[139] CMS Collaboration, "Search for standard model four top quark production at 8 TeV in the lepton + jets channel", Technical Report CMS-PAS-TOP-13-012, CERN, Geneva, 2013.

[140] CMS Collaboration, "Search for SUSY in same-sign dilepton events at 13 TeV", Technical Report CMS-PAS-SUS-15-008, CERN, Geneva, 2015.

[141] CMS Collaboration, "Performance of b-tagging algorithms at 13 TeV", Technical Report CMS-AN-2016/036, CERN, Geneva, 2016.

[142] CMS Collaboration, "Tuning the Amount of QCD Radiation in Powheg+Pythia8 in the ttbar Process", Technical Report CMS-AN-2016/120, CERN, Geneva, 2016.

[143] M. Cacciari, G. P. Salam, and G. Soyez, "The anti- k t jet clustering algorithm", *Journal of High Energy Physics* **2008** (2008), no. 04, 063.

[144] CMS, "Jet energy scale and resolution in the cms experiment in pp collisions at 8 tev", doi:10.1088/1748-0221/12/02/P02014, arXiv:1607.03663.

[145] J. Pivarski et al., "Awkward Array", 10, 2018. doi:10.5281/zenodo.4341376.

[146] N. Manganelli, "FourTopNAOD", 11, 2019. https://github.com/NJManganelli/FourTopNAOD.

[147] N. Smith and L. Gray, "coffea", 2019. https://github.com/CoffeaTeam/coffea.

[148] "nanoAOD-tools", 2018. https://github.com/cms-nanoAOD/nanoAOD-tools.

[149] "HiggsCombine". http://cms-analysis.github.io/HiggsAnalysis-CombinedLimit.

[150] R. Brun et al., "root-project/root: v6.24/02", August, 2019. doi:10.5281/zenodo.3895860, https://doi.org/10.5281/zenodo.3895860.

[151] CMS Collaboration Collaboration, "CMS Achievement Awards 2019". https://cms.cern/content/achievement-awards-2019. Accessed: 2022-08-14.

[152] Piparo, Danilo et al., "Rdataframe: Easy parallel root analysis at 100 threads", *EPJ Web Conf.* **214** (2019) 06029, doi:10.1051/epjconf/201921406029.

[153] N. Manganelli, "12345: Lessons Learned building an Analysis Framework around RDataFrame and CMS NanoAOD". https://indico.fnal.gov/event/23628/contributions/240753/.

[154] N. Manganelli, "Upgrades to the CMS cathode strip chambers for the HL-LHC", *Journal of Instrumentation* **15** (mar, 2020) C03047–C03047, doi:10.1088/1748-0221/15/03/c03047.

[155] V. C. Rubin, J. Ford, W. K., and N. Thonnard, "Rotational properties of 21 SC galaxies with a large range of luminosities and radii, from NGC 4605 (R=4kpc) to UGC 2885 (R=122kpc).", *APJ* **238** (June, 1980) 471–487, doi:10.1086/158003.

Appendix A

Analysis Software

A significant amount of software supports this analysis. The most notable of these are the Compact Muon Solenoid Software (CMSSW), the ROOT toolkit[150], and the analysis code developed specifically for the search for four top production presented in this dissertation [146]. Within the ROOT ecosystem, a new subpackage called RDataFrame [152], developed since 2018, forms the basis for the analysis code.

A.1 RDataFrame

Declarative Analysis Tool

RDataFrame (RDF) adopts a functional, declarative approach to analysis. This means that traditional per-event control-flow like for-loops, case statements, and if-else blocks are discouraged in favor of declarations that form a computation graph. The computation graph starts from a data source, such as a root file in CMS's NanoAOD data

format. Using actions such as *Define*, a node of the computation graph is taken as input (with all associated variables defined up to that node), and new variables may be *Define*'d, producing a new graph node. In the new node, the previously-available variables as well as the one just *Define*'d will be available for further computation or as input to summary statistics. The action *Filter* has a similar structure, but has orthogonal functionality, in the sense that any event going through the action must pass a condition specified in the *Filter* call, otherwise that event becomes inaccessible in downstream nodes. Together, these two operations form the bulk of manipulations on the computation graph, respectively "adding columns" and "removing rows" in the columnar-data perspective.

Histograms and Statistics

Another class of transformation is used to aggregate results. *Sum* and *Stats* will take as input a node, name(s) of a variable or variables within that node, and then produce the intra-event sum or statistical summary. This can be used to compute means and variances for a quantity, such as the scalar sum of jet p_{T} or all selected muon p_{T}'s. More frequently, the action *HistoND* (where N is 1, 2, or 3) is called to take all events passing through a node and fill a corresponding 1, 2, or 3-dimensional histogram with the data. The result of this action forms the inputs for the maximum likelihood estimation in Sec. 4.8, as well as the summaries used to check agreement between data and simulation and perform most analysis studies.

Full Event Export

RDF has multiple ways of exporting results. One method is to use *Snapshot* to write to disk all variables for events passing through the filters up to the node specified. The operation *Snapshot* writes out data in root format, and can be configured to do so lazily and in parallel with other files. This permits some samples, such as $t\bar{t}$, to be split into $t\bar{t}$ +*light* and $t\bar{t}$ +*heavy* components, neatly separating contributions. Additionally, the ability to convert event data to NumPy arrays with the *AsNumpy* function, in which scalar primitives, such as floats and integers, are translated to NumPy arrays of the same type. Each event corresponds to an indexed position in the 1-dimensional array. Jagged data, such as a variable-length vector of floats per event, or non-primitive data such as structs and objects, are all exported as NumPy arrays of python objects. This incurs a performance penalty, and additionally prevents many typical NumPy operations from being used on that data, encouraging the conversion of any such data to single per-event, primitive types.

A.2 Workflow

Corrections and Systematic Variations

The general workflow for analysis starts with adding corrections for systematic effects and uncertainties. This is accomplished using the NanoAOD-tools package centrally maintained by CMS Cross-POG [148]. My code configures this to re-compute Jet Energy Corrections, and add the systematic variations on those, as well as add scale-factors and uncertainty variations for b-tagging and lepton identification, isolation, and tracking. An

additional part of my four top analysis code is used to add basic trigger, lepton, jet, and event selection criteria, mostly encoded as bitmasks, and perform basic event selection to reduce the amount of data written back to disk. The code is executed using the Worldwide LHC Computing Grid (WLCG) to run over the billions of simulated and real data events necessary in the analysis, and the output is stored as approximately 4 Terabytes worth of root files for further processing.

Selection, Categorization, Reduction

The main step of the analysis uses the output from the NanoAOD-tools. At this stage, leptons are selected based upon the bitmasks encoded in the previous step. Then, for each variation of the Jet Energy Scale and Resolution, jets are selected according to the criteria described in Sec. 4.4, and those identified as overlaps with the lepton collections are removed. The totality of event selections are applied. The analysis quantities not already used in selection are then computed per event and systematic variation and histograms are produced. Once all Monte Carlo simulation and data has been saved as histograms, another stage is used to aggregate histograms into one of the final process groups, also reducing the binning for each jet and b-tag multiplicity category according to the preferred MC statistical uncertainty threshold. JSON files containing the event yields per process group and systematic variation are written to disk, and this info is used to fill CMS Higgs Combine datacard templates.

Statistical Inference

The final step of the nominal analysis chain is to then use the Combine package to run statistical inference. The software is documented on the web [149]. Expected and final cross-section limit setting, significance calculations, and nuisance parameter profiling are amongst the most important modes employed.

A.3 RDataFrame Evaluation

Competing Solutions

At the time of development, RDataFrame featured a number of pros and cons versus the competing solutions. Easy multithreading and fast execution speed are strong benefits of using RDataFrame. On the other hand, systematic variations require complicated overhead in the analysis code. Additionally, competing frameworks such as *coffea* [147], born out of the IRIS-HEP initiative (and based upon the Awkward-Array library [145]) provide better integration with machine learning libraries, as they naturally convert data to and from NumPy arrays for batched evaluation.

Declarative Syntax

RDF just-in-time (JIT) compiles C++ code contained in standalone files or encoded as strings in python, which brings with it comparable speed to pre-compiled code. As part of the ROOT ecosystem, it naturally integrates statistical tools and other components already available in the large software ecosystem. RDF's computation graph model allows computing systematic variations in the same event-processing loop as nominal variations,

dramatically lowering compute-requirements versus traditional frameworks that process the event loop for every variation. Multi-threading may be enabled with a single command, and the bulk of central computation graph actions (i.e. *Define*, *Filter*, *Snapshot*, *HistoND*) are compatible. This enables scaling an analysis from initial prototype to production relatively easily.

Efficient Systematic Variations

At the time of development, RDF did not have a mechanism for handling systematic variations. My code developed one of several implementations in the community for this purpose; this was accomplished by tracking and matching alternative variable definitions, downstream filters and corresponding weights, such as jet p_{T}, H_{T} cuts, and b-tag weights (since these depend on exactly which jets are selected). Each final result, i.e. a histogram, that should vary depending on systematic uncertainty needed to be booked on the appropriate computation graph node, with the aforementioned matched arguments. This leads to a significant number of python for-loops and heavy use of string-replacement functions, which can balloon the setup time for the computation graph. An in-development RDF function called *Vary* will enable a single call to manage this kind of complexity implicitly. Internally, RDF will use the directions to store variations in how event quantities are computed. Then, all downstream variables that depend on those functions will implicitly have matching variations. When results are requested, such as histograms, compiled C++ code will handle the matching of inputs and book the histograms. In this way, future RDF users will benefit from significantly faster computation graph-building, and much less code will be necessary to manage variable dependencies and matching. As one of the earliest adopters

of RDF, the feedback from this analysis informed the development of this function by the ROOT development team [153].

Implicit Data Structure

RDF does not have a native way of bundling multiple arrays together. As an example, when selecting quantities for muons, the most straightforward way is to construct an array with boolean values, with each index position corresponding to a muon in the event. This *mask* can be the combination of multiple sub-selections, such as requiring muons to have p_{T} over a threshold, η within the muon tracker acceptance, and so on. To then create a subset of muon variables, ROOT's *RVec* functionality can be used, similar to NumPy fancy-indexing. However, this must be repeated for every variable individually, such that one might have code blocks as in Code A.3. A future feature to simplify user code could enable this via a single function call, which is still a proposal as of Summer 2022.

RDF works best on implicitly structured data. While structs and classes can be instantiated with the arrays of floats, integers, and other primitive types for each object in an event, this disrupts workflows by disabling optimizations in RDF by requiring that all requested branches in the root files be read and decompressed into RAM. To produce histograms for a given quantity, an array of primitive type must be extracted from such an array of structs or class instances, and this additional translation step is cumbersome. Competitors like *coffea* build python proxy objects that behave like structs or classes. Operations on these proxies are translated to manipulations on the underlying data, allowing for python-like idioms and syntax but with code execution comparable to compiled C/++

code. This results in much more concise code, with less spent keeping array manipulations synchronized. Building a proxy system like this for RDF presents challenges, but my feedback to the ROOT team, as in the talk for the ROOT User's Workshop [153], emphasized the importance of this feature for improving the User Experience (UX), and it is one of the chief development priorities in 2022.

RVec Fancy Indexing Example

```
#include "ROOT/RVec.hxx"
ROOT::VecOps::RVec<Float_t> mask;
mask = (Muon_pt > 30) && abs(Muon_eta) < 2.4 && Muon_looseId;
//mask will be std::vector-like containiner of booleans
// {true, false, false, true, false}
IsoMuon_pt = Muon_pt[mask];
//Using the mask in square brackets will select a subset
//Muon_pt {37.0f, 35.6f, 33.2f, 30.8f, 15.0f}
//IsoMuon_pt {37.0f, 30.8f}
IsoMuon_eta = Muon_eta[mask];
IsoMuon_phi = Muon_phi[mask];
IsoMuon_mass = Muon_mass[mask];
```

Appendix B

Data Taking, CSC Run Coordination, and DOC Training

B.1 Detector On Call Shifts for Data Taking

For six weeks in 2018, I served as Detector On Call (DOC) for the Cathode Strip Chamber subsystem. The Large Hadron Collider provided proton-proton collisions to the four large LHC experiments from 2015 to 2018, collectively referred to as "Run II." To support data-collection for the CMS experiment, a large group of physicists must serve as systems experts, one role being DOC. DOC shifters are on call 24 hours per day for a week at a time. They are required to do daily checks of system health, checking firmware for electronics and re-loading it; handling intermittent problems in high- and low-voltage power systems; and ensuring basic trigger and data-readout are working as expected via monitoring tools. In addition to being on-call for the entirety of the shift, DOCs serve as

the liaisons between the experts for a sub-detector and the rest of CMS. Problems in other subsystems must be communicated to the experts in CSC so that an appropriate response can be crafted, and tests in-between *fills* (i.e. interludes between periods when protons are circulating in the LHC and collisions are possible) must be coordinated to satisfy all the inter-dependencies in the experiment.

B.2 CSC Deputy Run Coordination

I was invited by the CSC Subsystem management team to apply for the position of CSC Deputy Run Coordinator (DRC), and was subsequently chosen to serve as such between September 2019 and October 2021; this made me the first graduate student to hold a CSC management position. While in this role, there were three principal tasks to accomplish, commissioning the upgraded chambers, coordinating activity with CMS, and ensuring preparedness for the Run III start in mid-2022.

Commissioning

First, the subsystem needed to be recommissioned once the 180 CSC chambers ("MEX/1" types, where X=1 – 4, that we extensively upgraded on the surface, see App. C) were reinstalled, and around 300 more non-inner-ring chambers also needed upgrades in-situ (most "MEX/2", "ME1/3" types). I was responsible for entering the experiment cavern and using cherry-pickers to replace ALCT mezzanine cards on chambers , some up to 15m above the cavern floor, as well as re-connecting re-installed chambers. Commissioning the system required updating databases with new parameters from measurements and scans,

deploying several iterations of new firmware for various electronics, and exercising their capability. For the hardware, alongside the CSC Run Coordinator (RC), I was running basic connectivity tests, and tested the system via data collection using the cosmic ray muons that make it through the 100-plus meters of rock above the detector. Replacing peripheral-crate electronics, re-seating connections between them and chambers, and checking fiber connections were routine parts of the job.

Coordination Responsibilities

Coordination with the rest of the experiment was another crucial aspect. The individual subsystems in CMS have a degree of independence, but there are a number of inter-dependencies for each individual subsystem to function correctly. Technical Coordinators were responsible for making sure that work in the experiment cavern was properly handled, such as ensuring detectors were not powered when people would be nearby. Subsystem (Deputy) Run Coordinators work together to ensure that joint tests succeed, such as in the case of the commissioning for the Gas Electron Multiplier (GEM) detectors installed near the ME1/1 chambers, where CSCs must be triggering such that coincidences in muon tracks can be analyzed and used for timing adjustments. As the GEM subsystem was installed in CMS during LS2, I had to plan joint tests with their operations experts throughout 2021. Similarly, CMS Mid-Week Global Runs (MWGRs) serve as some of the most important tests of the overall detector readiness. All available subsystems join the global run, jointly collecting cosmic ray data in the best approximation of collisions we have without LHC beams. Since the CSC subsystem is the only detector that can provide triggers for muons only passing through the endcap, having high availability of the system

for these runs was paramount, and this meant scheduling upgrades of the CSC subsystem appropriately.

Training and Documentation

The last subcategory of work involved other preparations for Run III and beyond. First, documentation for use by CSC operations needed to be updated, and numerous corrections to match the state of the post-LS2 configuration were added to twiki pages. Second, since normal data-taking had ended, most people qualified as DOCs had moved to new positions. I oversaw the training of several new shifters for both MWGRs and the start of Run III collisions, planned for mid-2022. Once new DOCs started signing up for shifts, I stepped back into a more typical supervisory role, working with the other CSC experts to plan activities while delegating the detector operations to the DOCs. As DRC, I needed to be available to communicate with the DOCs should any questions arise or change in plans be needed. I closed 2021 by serving as acting Run Coordinator from late August through October, and trained a new PostDoctoral researcher to become the incoming Deputy Run Coordinator.

Appendix C

Upgrades to the Cathode Strip Chambers

C.1 Cathode Strip Chambers

During the shutdown of the Large Hadron Collider between 2018 and 2022, internally labelled Long Shutdown 2 (LS2), the Cathode Strip Chambers (CSCs) were upgraded in anticipation of the High-Luminosity LHC project. The CMS detector needed to be adapted to handle the increased luminosity, which is expected to exceed 5×10^{34} cm^{-2}s^{-1}, more than a factor 5 over the original design value [154]. The CSCs are one of the three original sub-detectors in CMS for dedicated muon detection (the fourth subdetector, Gas Electron Multipliers (GEMs), being installed in two parts broken up between LS2 and a subsequent YETS). The CSCs have been the backbone of muon tracking and triggering for the $0.9 < |\eta| < 2.4$ region since the start of LHC collisions, performing admirably through

the discovery of the Higgs Boson and many other analyses. With the higher intensity expected in the HL-LHC, however, the readout electronics for many chambers close to the beam pipe had to be upgraded, as otherwise valuable data would be lost due to insufficient buffering and bandwidth. Upgrades to the chamber electronics were thus planned during Run II, then executed in LS2, owing to the number of project upgrades that could only be done in LS3 (expected to occur between 2026 and 2028). Of the 540 chambers installed, 72 ME1/1 chambers required upgrades to cathode and anode readout, 108 ME234/1 chambers additionally needed upgraded low-voltage distribution boards, and most other chambers in the system (excepting the ME4/2 chambers) required anode readout upgrades.

Upgrade Process

Since the CSCs are crucial to both triggering and reconstruction of muons in the end-cap region, it is paramount that they maintain high efficiency. Heading into the HL-LHC era, the rates from the higher instantaneous luminosity would lead to significant event losses in the more forward region, simply due to lack of buffering in the original electronics. The CMS detector is inherently a deeply pipe-lined system; raw data for events must be stored on the readout electronics in buffers while trigger information is propagated throughout the system. By the time a decision on whether to accept the event has been made, many events worth of information could be stored in the buffers, and that would have been overwhelming with the HL-LHC expected luminosity. Figure C.4 shows the fraction of events lost with old electronics in red, green, and blue for the 3 non-ME1/1 inner rings; with upgrades, all those chamber types will follow the purple curve.

Figure C.1: (left) The SX5 Surface Lab with mostly empty chamber tables, in Green. Chambers follow a zig-zag pattern, being refurbished near the middle-right before proceeding to the FAST test stands to the left, and then the Long Term Test stand (LTT) under the fume hood in the top of the image and Fig. C.2. (right) For most of the production, the lab was filled, with chambers being actively refurbished, tested, or awaiting delivery to the Experiment Cavern on tables.

The upgrade process consisted of a few stages, itself assembly-line-like, as shown in Figs. C.1. First, extraction of the chambers from the disks occurred using specialized rigs and trained personnel, precariously close to the detector, and they were transported to the surface lab. The second step was refurbishment, which involved stripping all the electronics attached to the cooling plate of each chamber (which primarily cools the electronics, the active volume of the chambers only needs to maintain a stable temperature and humidity), replacing that cooling system with a new joint-less one (zero welds) to reduce leak risks (ME1/1 types), installing improved electronics, and re-connecting all the fiber-optic and copper cables. Third, rigorous testing of all the electronics, both artificial tests involving generated pulses and noise checks, as well as those using real muons from cosmic rays; the testing aspect was my principal work between summer 2018 and October 2019. Fourth, re-installation of the chambers. Fifth, once chambers were re-installed, re-commissioning is performed: connectivity tests, software and firmware upgrades, performance testing,

calibration and timing procedures for collisions, and any ancillary activities needed to ensure readiness make up this stage. Some of this is further detailed in Sec. B.2.

Interfacing Hardware and Software for the CMS Critical Path

In June of 2018, after arriving at CERN, I started as the right-hand-man to the staff scientist, Karoly Banicz, in charge of preparing the SX5 lab for testing (a surface area situated next to the main access shaft to the CMS Underground Experiment Cavern (UXC) at LHC Point 5 in France). Our two-person operation was the nexus of establishing the test stands for all upgrade activity within CSC for LS2.

First, I set up requisitioned servers with fresh installs of the latest CentOS operating system for the experiment, HDD's in mirror-RAID configuration, Intel Network Interface Cards (NIC) for control and Data Aquisition (DAQ). Due to the custom ethernet-packets used in the system, corresponding custom kernel-drivers for these NICs needed to be updated to match changes in the OS. A full suite of Peripheral Crate electronics were procured and installed in the Versa Module Eurocard (VME) crate, including a VME Crate Controller (VCC) for communication and control of the other electronics; a Clock and Control Board (CCB) to distribute the on-board 40MHz clock to all electronics in the system (syncing); a Detector Dependent Unit (DDU); an Optical and regular DAQ Motherboard (ODMB, DMB), Optical and regular Trigger Motherboard (OTMB, TMB) which are detailed in Chapter C.2. Compared to a full system, this 'vertical-slice' lacked only the Muon Port Card (MPC) for sending triggers to the Endcap Muon Track Finder (EMTF) and more copies of the per-chamber (O)DMB+(O)TMB. Fibers were labeled and run between

all the components needed: connections for the chambers, to the ODMB and OTMB, from the ODMB to the server. Fig. C.2 shows the FAST Test Stands.

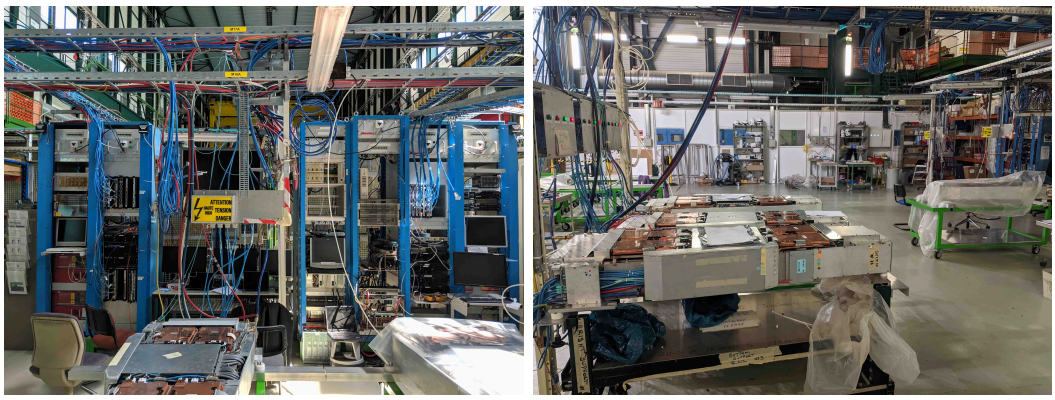

Figure C.2: (left) The FAST Test Stands, with servers (black) installed inside the blue racks. Fibers connect to the Peripheral Crate boards above and below. (right) The Long Term Test stand houses multiple chambers below the fume hood, with the electronics on the far right of the left image.

What followed was the more challenging stage: performing the first full integration tests of chambers with new electronics, FPGA firmware, and detector software. Most things had only been tested in pairwise fashion, such as firmware with one board, one on-chamber board with corresponding off-chamber board, etc. There was a significant amount of work in getting everything working together, simultaneously including all new board types with updated firmware and new software. Through detailed communication and coordination with other CSC experts, these elements were brought to a deliverable stage on time, taking several months of intense testing and experimentation. In Figure C.3 is the oscilloscope readout from when we were debugging a test pulse generator that was not configuring as expected from the commands sent by the software, confirming the hardware and software

were in disagreement on the state. This was just one among dozens of incompatibilities and bugs that we worked through. Within a few weeks of commissioning the test stands, the actual production testing began.

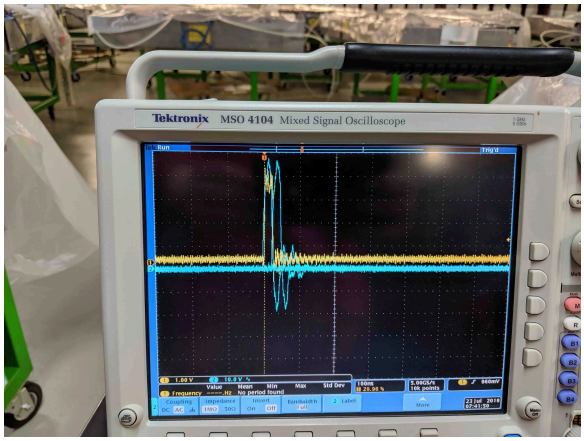

Figure C.3: The test pulses as measured by an oscilloscope, during the debugging of firmware and software for DCFEBs paired with DMBs.

Test Stand Expert

Since Karoly and I were the only ones intimately familiar with the test stand setup, we were mainline support for all testing efforts. As the more junior member of the collaboration, I thus took care of all the easy- and intermediate-level problems. A prominent one was recovering servers from power cuts (effectively a bi-monthly occurrence, since our setup was distributed and could be perturbed by (un)intentional power cuts at both the Point 5 SX5 lab and those at Building 904 on the Prévessin site). The role also included debugging problems with our network drivers and the DAQ system, hardware problems with the test stands such as poor or swapped connections, faulty boards, and doing anything to make the

testing team's work easier. As an example of the latter, I wrote a number of shell scripts to automate tasks for the testing team, including resetting drivers and a partial-recovery from power-cuts, intended to save crucial moments and make the strict deadlines doable. I fulfilled this duty from late 2018 to early 2020, as I transitioned more fully into the Deputy Run Coordinator role B.2.

Testing Phase II Muon Chambers

Simultaneous to supporting the test stands, I was embedded with one of the the two testing teams at the start of the upgrade campaign, leveraging my familiarity with the hardware and software. Each of the two groups, about half a dozen people, were scheduled for 1 week rotations. This was a very intense period, usually 6 days a week, 7am – 8pm, to meet deadlines for our LS2 schedule. After chambers were extracted from CMS and refurbished with new electronics, my team ensured their readiness for data taking and re-installation. I applied my problem-solving skills to identify issues with electronics, interconnects, and other hardware failures.

Testing required a full suite of diagnostics on the chambers, the most basic of which ensured that all connections internal to the chamber (and by necessity, from the chamber to the external electronics in the VME crate) worked as expected. Other tests that we could run with our test software checked that FPGA's configured correctly, that the noise read out from cathodes and anodes were below acceptable thresholds, and that injected test pulses (triggered via induced charge from test strips/wires in the chamber) could be appropriately

read back out in a loop-back style. Once the chamber could successfully pass all these tests, which frequently meant replacing a board or re-seating copper or fiber connections to it, it moved to the final stages. In this, we applied a high-voltage to the inner chamber layers, 2.9 – 3.6 kV (depending on chamber type), and ran the chambers through a suite of tests measuring the characteristics of real muon detection, such as gas-gain, resolution, and triggering. Since the chambers were oriented flat, with a face towards the sky, they would detect several 10s of Hz of real muons from cosmic rays.

An additional stage of testing consisted of connecting the chambers to a long-term test stand, which monitored voltages, temperatures, and currents over 48 – 72 hours while they were powered. This stage was designed to catch so-called "infant-mortality," electronics that could fail very early in the lifecycle. The restrictions on the experiment preclude re-extracting chambers and replacing or repairing them until LS3 or a Year End Technical Shutdown, so it was critical to catch as many such defective electronics as possible. Once this "burn-in" phase was finished, a select subset of tests from the initial testing phase were repeated to ensure the chambers' characteristics remained stable.

The bulk of this work involved deducing novel failure modes for the chambers. Many of these were easily-resolved by fixing connections or replacing electronics boards. Some cases involved investigation of the internal elements of the chamber, and in two cases spare chambers were swapped in, since repairing chambers is a prohibitive exercise. Among the more notable problems I solved was a case where chambers began mysteriously failing.

Aluminum covers to shield the electronics arrived late and were installed after all burn-in tests were completed, necessitating a third round of re-testing. A small fraction of chambers inconsistently failed. I discovered that a last-minute change to the upgraded mezzanines left mounting screws a few millimeters too high, touching the covers installed over them. The slight pressure this exerted on the mezzanines transferred to the PCB baseboard below, disrupting communication and damaging the boards. This failure actually cracked traces in the PCBs, which meant that our store of preciously-few spare ALCT baseboards was depleted in addressing the problem. My discovery enabled the CSC group to implement a fix and prevent future damage to valuable electronics.

Team Leader

Approximately halfway through my service on the team, I was promoted to team lead for my group. I became responsible for oversight and coordinating the efforts of the permanent members and several summer students or postdoctoral researchers visiting CERN. Carefully balancing the needs of the team members with the needs of the experiment took priority, which in practice meant ensuring work was being done safely despite the incredible time pressure placed on our testing team, with too little person-power assigned. Petitioning for more students and postdocs to help became one of my more important jobs. Additionally, I contributed substantially to updating our testing procedures, alongside the other team leader, Katerina Kuznetsova. Providing guidance on debugging chambers to my team formed the third pillar of my efforts. This overall work, lasting up to November 2019, was largely what earned me a CMS achievement award for 2019 [151].

C.2 CSC Details

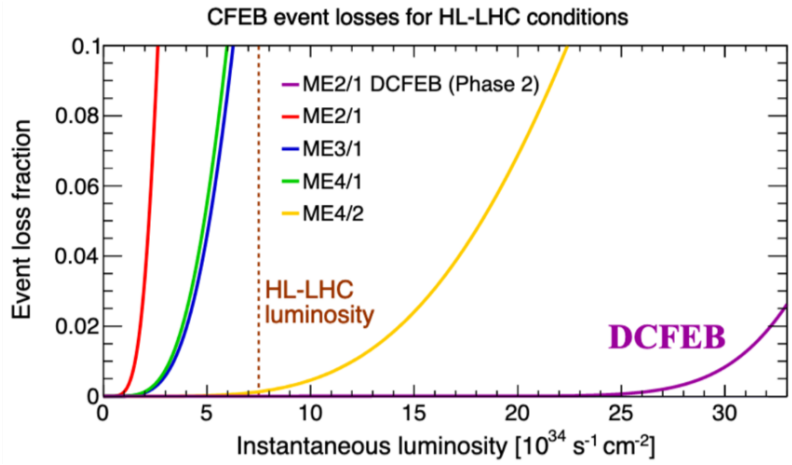

Figure C.4: (left) The event losses for ME2/1, ME3/1, and ME4/1 chambers with CFEBs installed become drastic at the proposed HL-LHC luminosity. However, replacing CFEBs with upgraded Digital CFEBs as seen on the rightmost curve entirely mitigates the issue [3].

(x)DCFEB

At the heart of the Digital Cathode Frton End Board (DCFEB), responsible for the readout of 6 layers of cathodes, 16 strips each, is a Virtex-6 FPGA paired with flash ADCs. Compared to the original CFEB installed on CSC chambers, which used Switch Capacitor Arrays to store 8 times samples of charge from all the strips (in a ring-buffer format), the digital pipeline of the DCFEB can keep 700 events worth of induced charge information. The DCFEB was used to upgrade the ME1/1 chambers in LS1, and during LS2 their Firefly optical tranceivers were replaced with the more radiation-hard Versatile

Twin Transmitter (VTTx) for the trigger and data paths, then installed on ME2/1, ME3/1, and ME4/1 chambers. In order to ensure operations throughout the HL-LHC lifetime, it was decided that supporting PROM-less programming, in which the on-board EEPROM is bypassed to program the FPGA directly from the control system via the Gigabit Transceiver (GBTx), was a necessity in the event EEPROMs suffered too much radiation damage. The DCFEBv2, colloquially called the xDCFEB, thus integrated the bidirectional Versatile Transceiver (VTRx) as well as a VTTx for trigger and data transmission off-chamber. Since the original DCFEB production was intended to furnish 72 chambers with 7 boards each, and this was insufficient for 108 ME234/1 chambers needing 5 boards each, xDCFEBs were installed on 9 ME2/1 chambers (3 Peripheral Crates with 3 ME2/1 chambers each; this kept some homogeneity in programming time and firmware types for operations and database convenience).

ALCT

The ALCT mezzanines of the inner-ring chambers were all upgraded to Spartan-6 FPGAs, which have 9x–12x the Block RAM of the Virtex-E from the original designs. By switching from copper-readout to optical data transmission, the upper-bound on transmission to off-chamber electronics is increased 8x–12x. Smaller chambers like ME1/1 receive an ALCT-LX100, sufficient for the 288–382 WGs readout (and including a bi-directional VTRx to receive PROM-less programming code from the ODMB7); larger chambers need an ALCT-LX150T design with more resources to handle the 576–672 WG readout. These latter mezzanines use a VTTx to provide upwards of 6.4Gbps bandwidth to the ODMB5 that will be installed, but during Run III the copper readout will be used for the data path.

OTMB

Optical Trigger Motherboards (OTMBs) are required for chambers upgraded with (x)DCFEBs, although the ALCT trigger path will always use the copper connection through the backplane. While the original OTMBs for the ME1/1 chambers were moved to ME3/1 and ME4/1 chambers, newer iterations of the OTMB needed to be produced for the ME2/1 and ME1/1 chamber types. The latter required support for PROM-less programming, and the former required support of connections with the GE2/1 GEM chambers to be installed in a YETS before LS3; additionally, not enough of the old OTMB design existed to support all chambers with (x)DCFEBs following LS2. The design is little-changed, mostly replacing the optical tranceivers with Samtec Firefly types. The additional input for ME2/1 type boards for GEM trigger information permits increased efficiency (by allowing substitution of GEM hits for missing CSC hits) and coincidence formation for new and novel triggers, like those for displaced triggering in the muon sub-system.

ODMB

For HL-LHC operations, chambers with (x)DCFEBs will be connected to new ODMB7 and ODMB5 boards with optical interlinks. For interim operations between LS2 and LS3, the non-ME1/1 chambers that now have (x)DCFEBs (but still must use DMBs) provide data through the backwards-compatible copper interconnects. Kintex Ultrascale FPGAs will form the logic core for the two new designs, and the ODMB7 for ME1/1 chambers will include bidirectional links to permit PROM-less programming through the GBTx of the new ALCT-LX100 mezzanines and xDCFEB boards. The ODMB5 boards,

paired with non-ME1/1 inner-ring chambers, will support 2–3 links to the new FED system, while 4 are used on the ODMB7. As seen in figure C.5, there will be an adequate safety margin for HL-LHC operations.

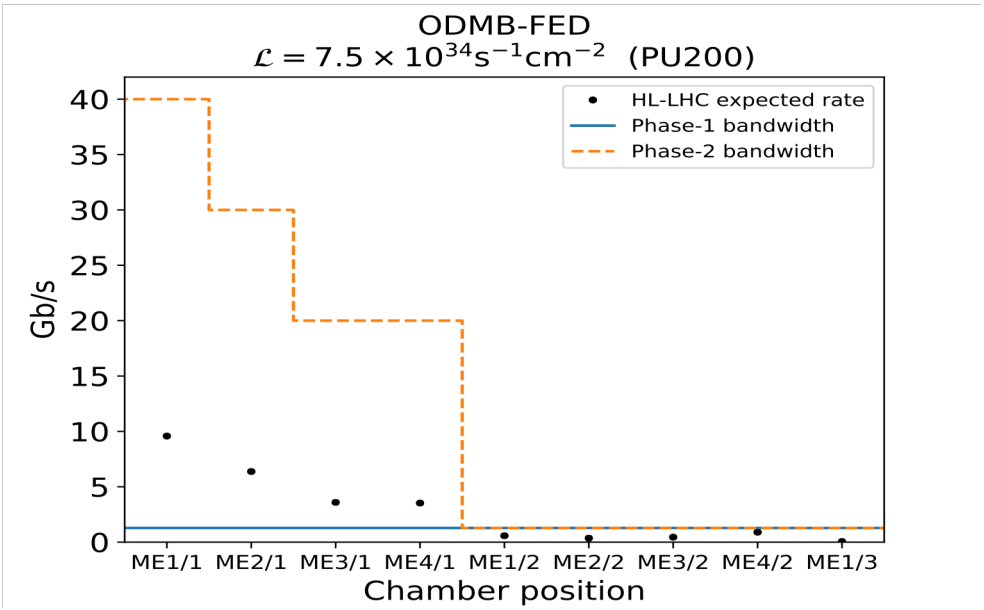

Figure C.5: The Phase-2 bandwidth will exceed the projected HL-LHC needs by a small margin, with the increased bandwidth of ODMB7s and ODMB5s apparent versus the Phase-1 bandwidth. Four Links are available for ME1/1 ODMB7s, while ODMB5s are configured to use 2 or 3 links to the FED based on the ring for which they are installed. The figure is produced with an expected PileUp of 200 pp collisions per bunch crossing (PU200) [from personal correspondence with Jaebak Kim, Oct. 2, 2019].

FED

Coinciding with upgrades to the ODMBs will be an update to the FED system, based on the Advanced Telecommunications Computing Architecture (ATCA) industry standard being employed in many upgrade projects throughout CMS, including the GEM and ME0 sub-systems. Part of the X2O Project, the GEM and CSC subsystems will

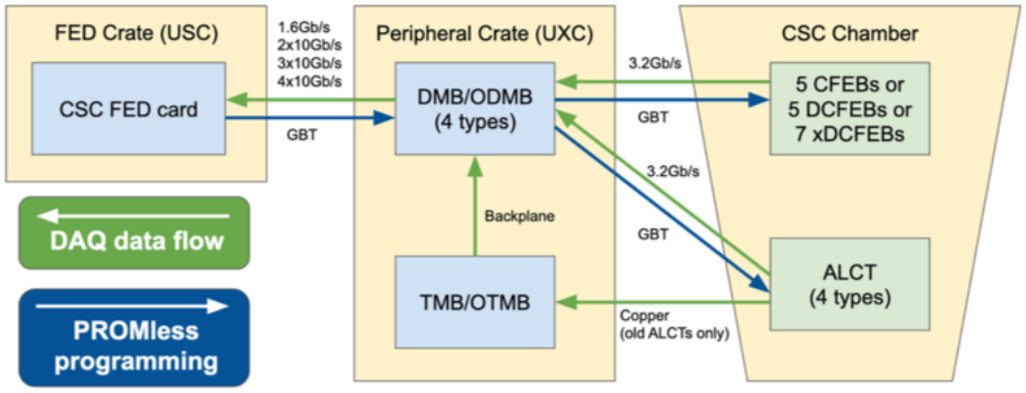

Figure C.6: Layout of the system after electronics upgrades

make use of the same processing module, differing only in some interfaces and firmware, built around the Xilinx Virtex Ultrascale VU13P and supporting 900 optical links for communication with the (O)DMBs. The total CSC Data Acquisition rate will be 2.8Tbps, aggregating the input of 72 ME1/1 chambers with 4 12.5Gbps links, 108 MEX/1 chambers with 2–3 similar links, and 360 chambers with 1.6Gbps links.

Appendix D

Four Top Analysis Appendices

D.1 Higgs Combine Templates

In this section, the templates as binned for input to Combine are presented. The templates begin with 100 equal-width bins between 400 and 2000 GeV, and bins are merged to satisfy the criteria that the total $t\bar{t}$ unweighted statistical uncertainty is less than the 30% threshold. These criteria are met per year and lepton decay channel.

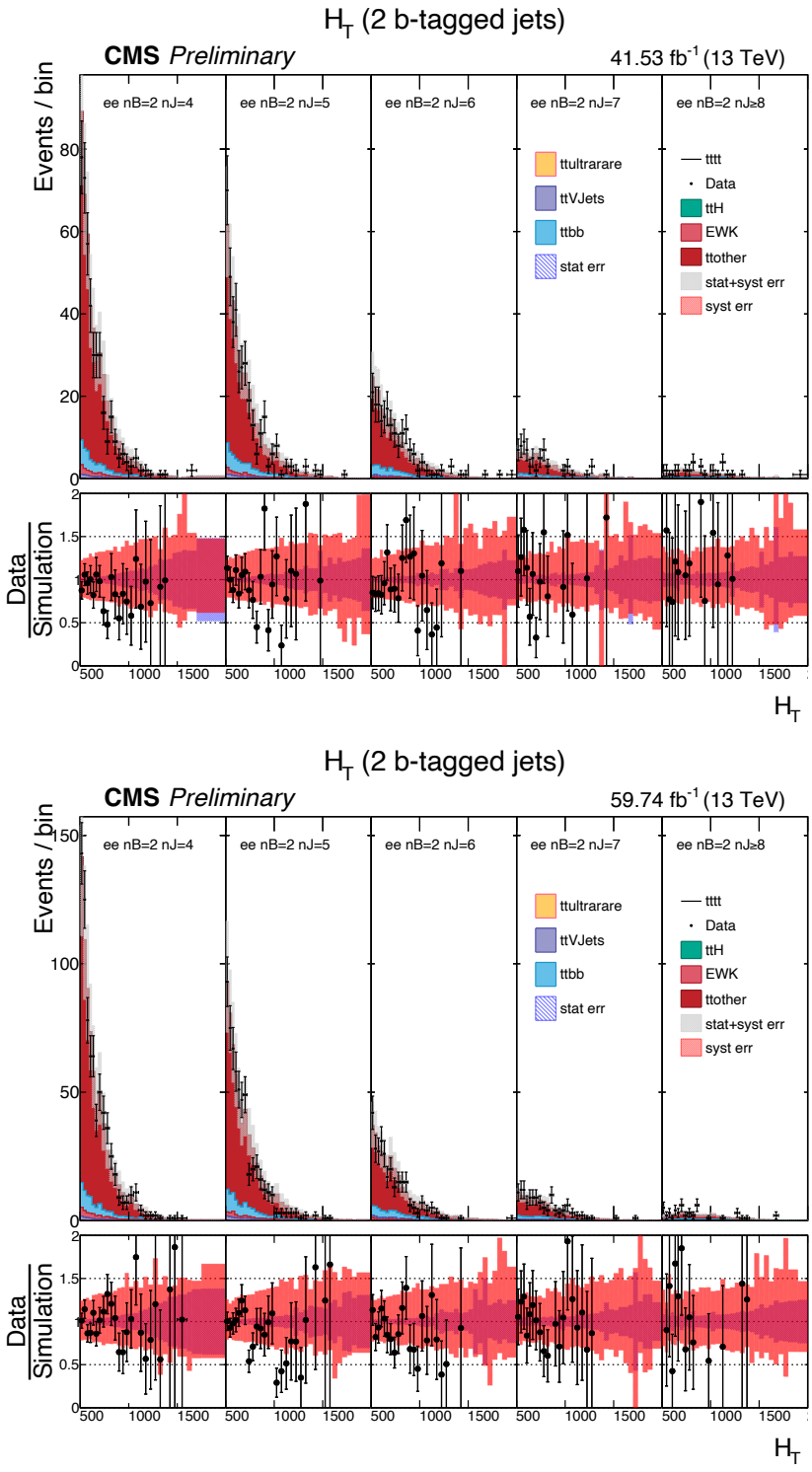

Figure D.1: H_{T} templates for the ee channel in the 2 btag category, 2017 top and 2018 bottom

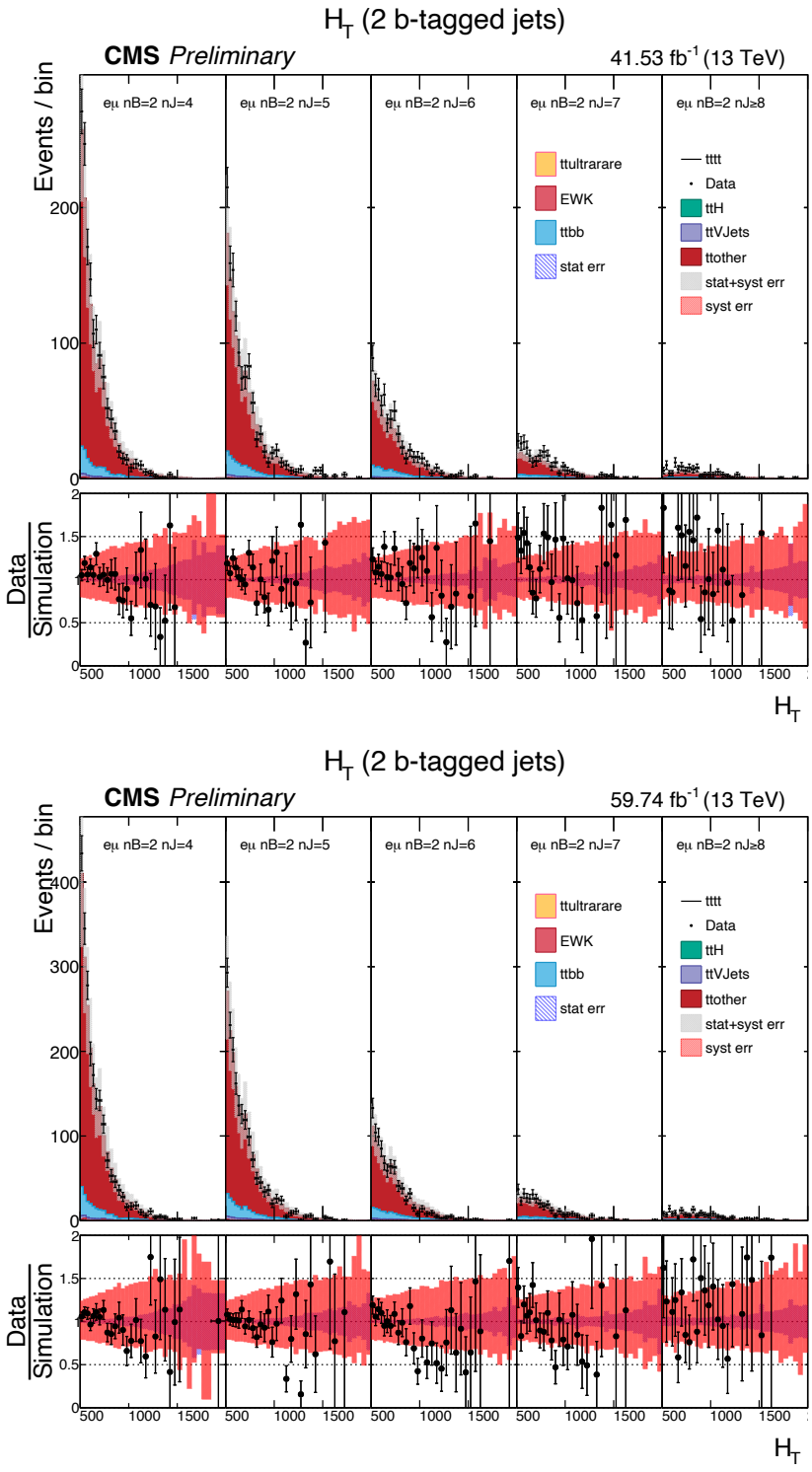

Figure D.2: H_{T} templates for the eμ channel in the 2 btag category, 2017 top and 2018 bottom

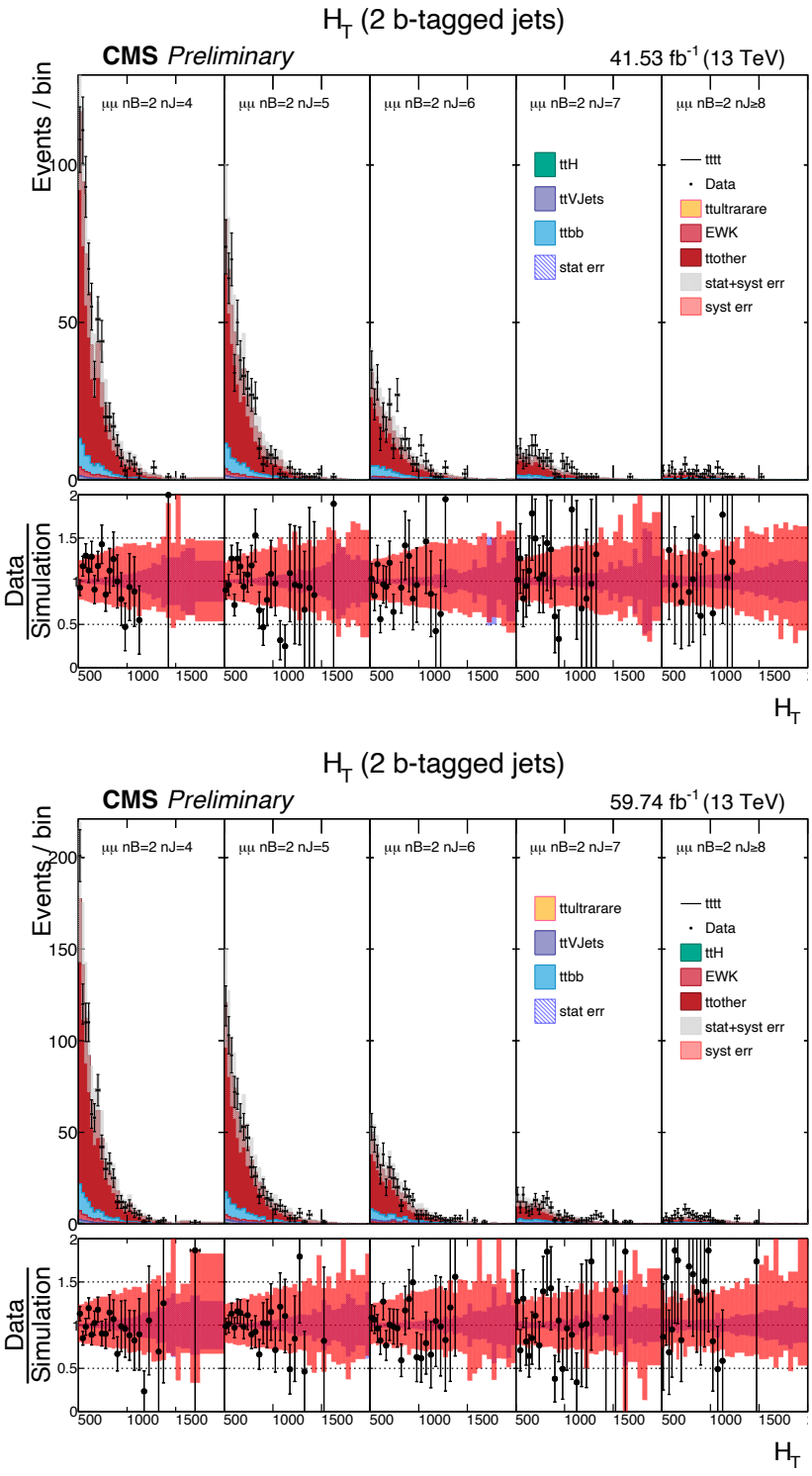

Figure D.3: H_{T} templates for the $\mu\mu$ channel in the 2 btag category, 2017 top and 2018 bottom

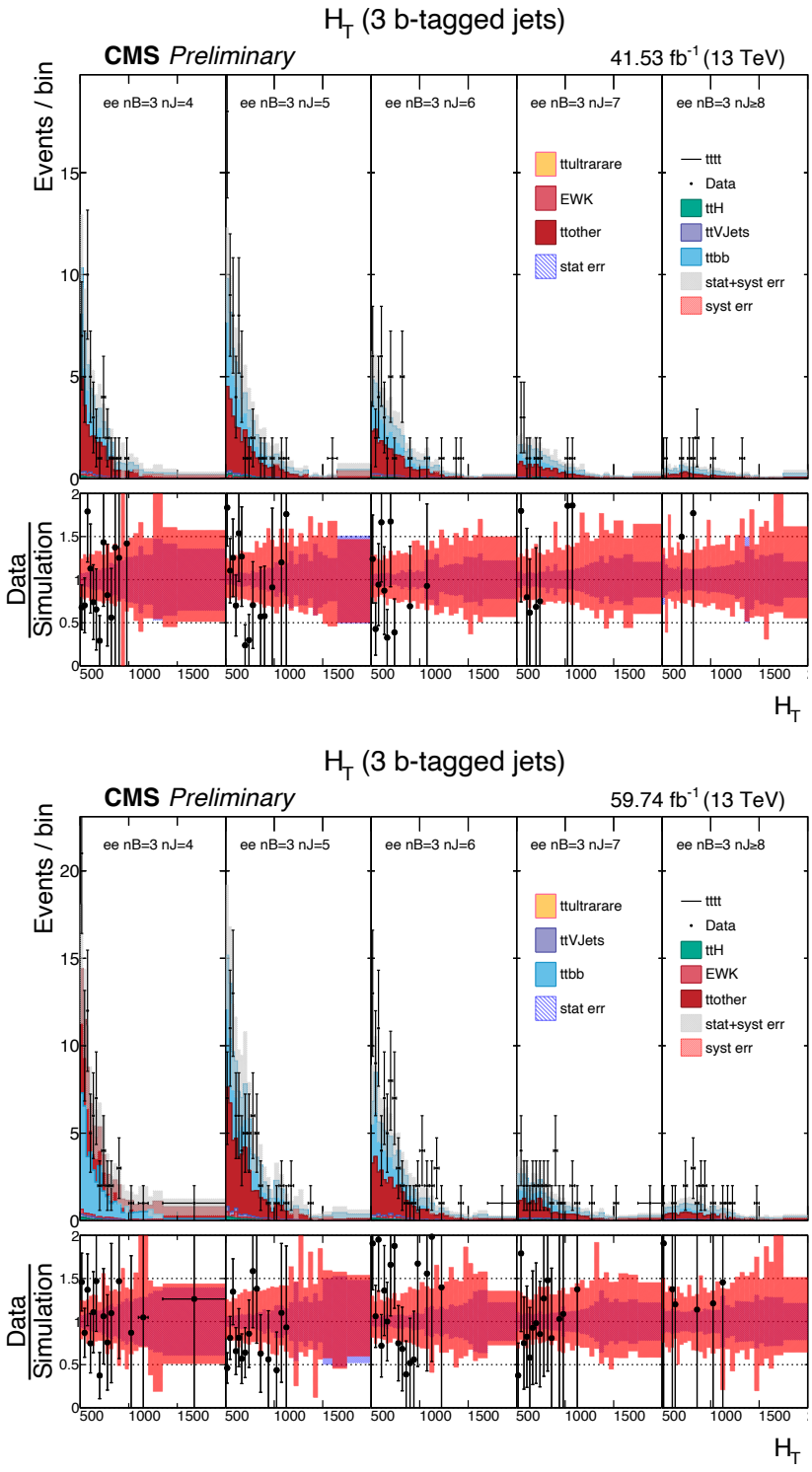

Figure D.4: H_{T} templates for the ee channel in the 3 btag category, 2017 top and 2018 bottom

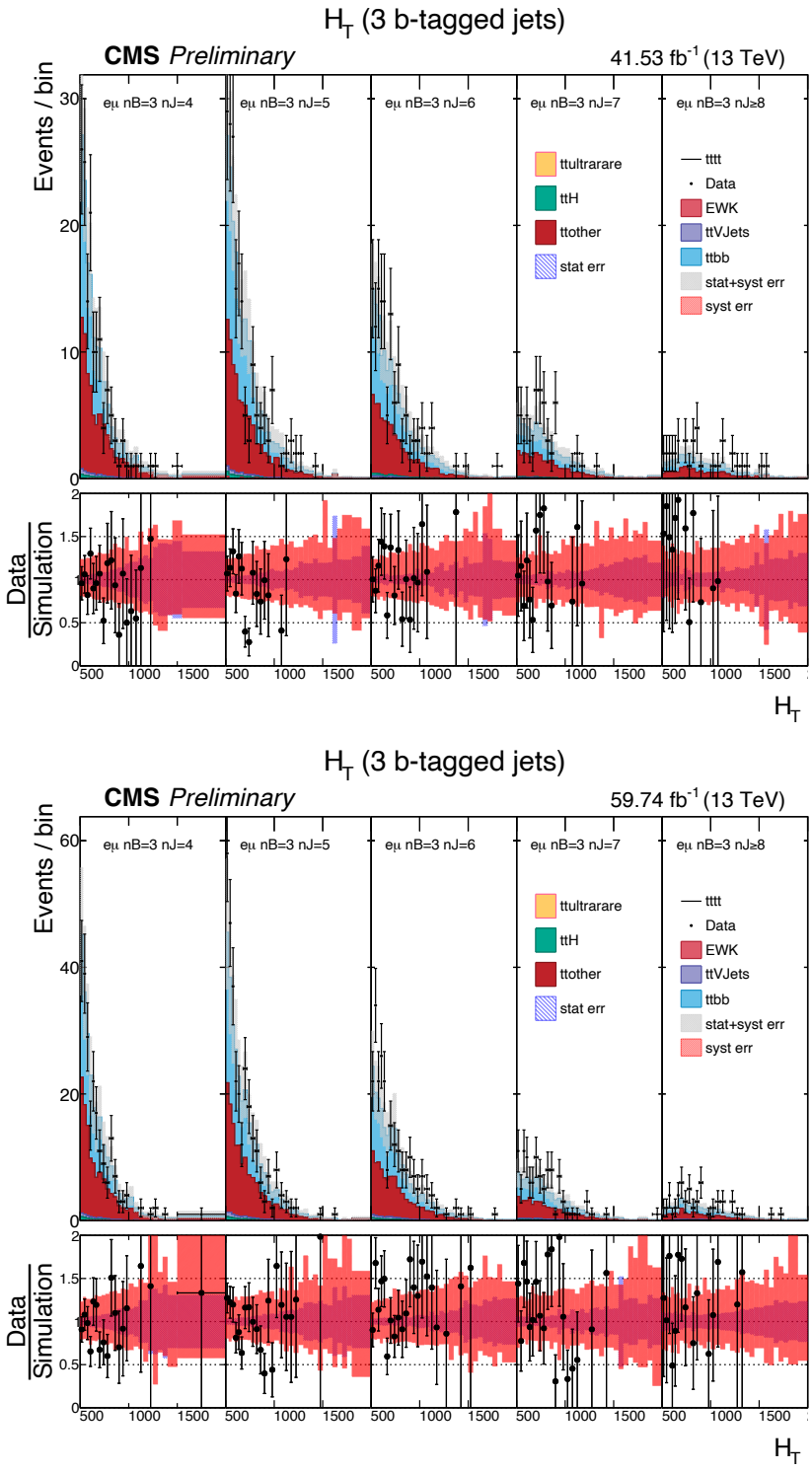

Figure D.5: H_{T} templates for the eμ channel in the 3 btag category, 2017 top and 2018 bottom

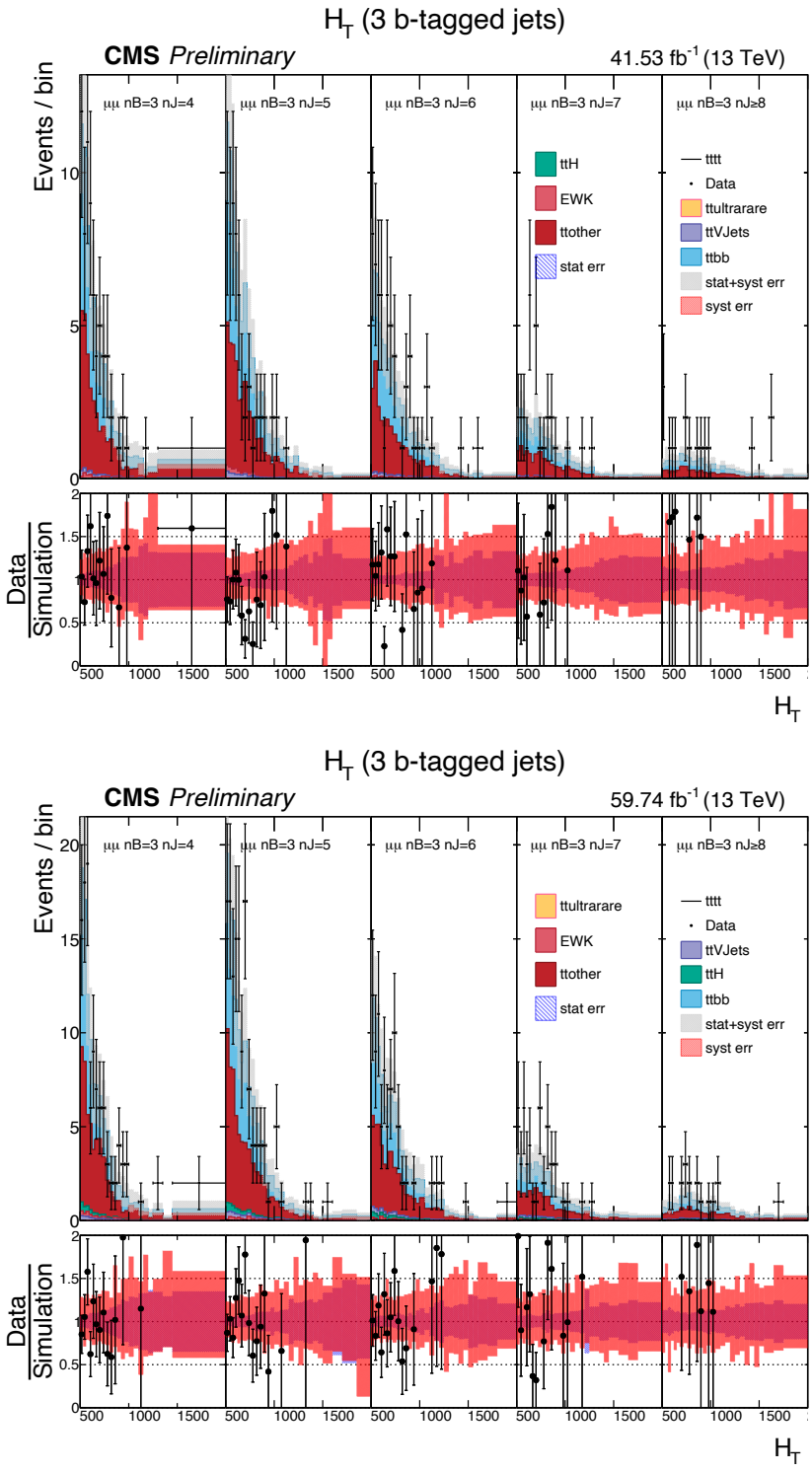

Figure D.6: H_{T} templates for the $\mu\mu$ channel in the 3 btag category, 2017 top and 2018 bottom

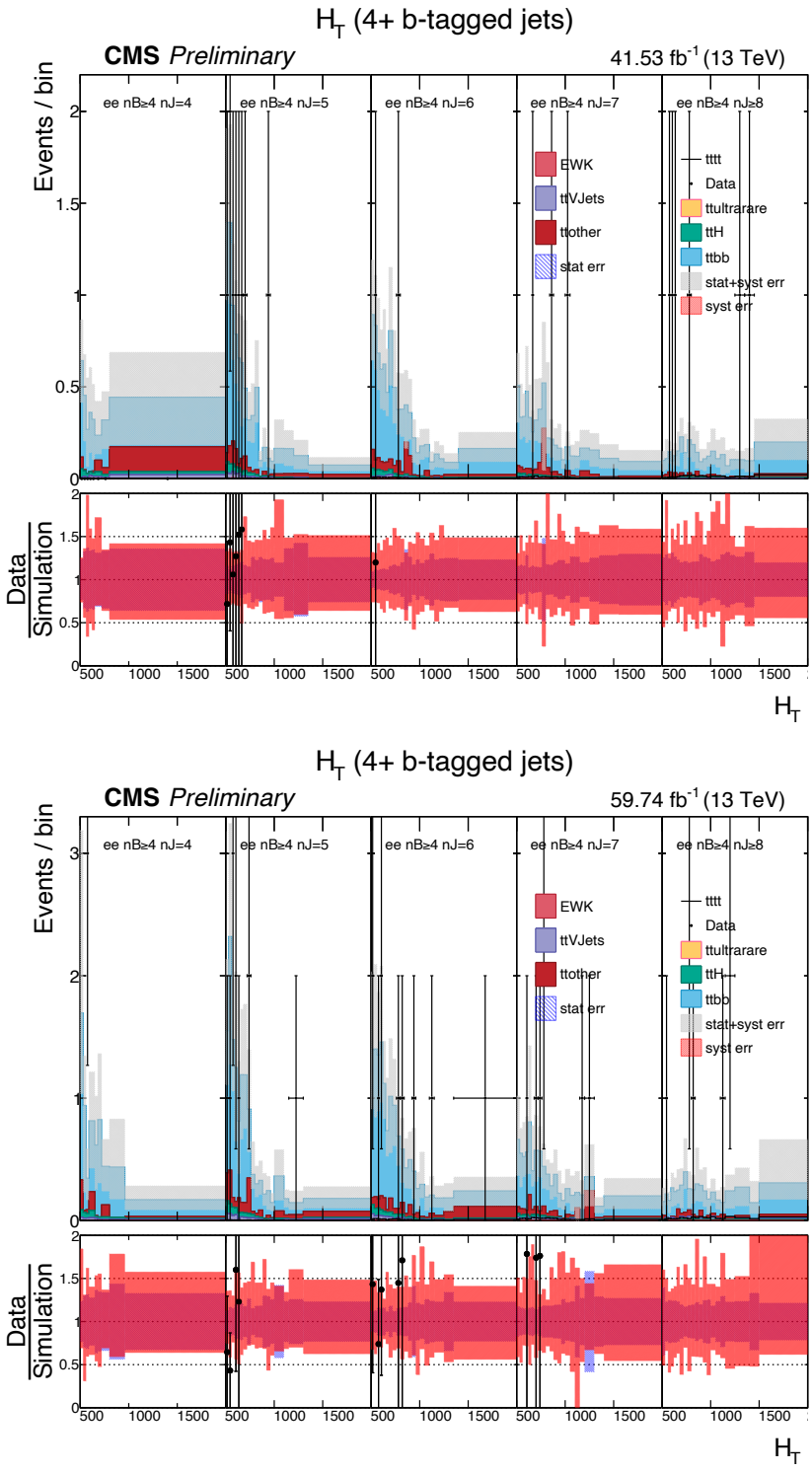

Figure D.7: H_{T} templates for the ee channel in the 4+ btag category, 2017 top and 2018 bottom

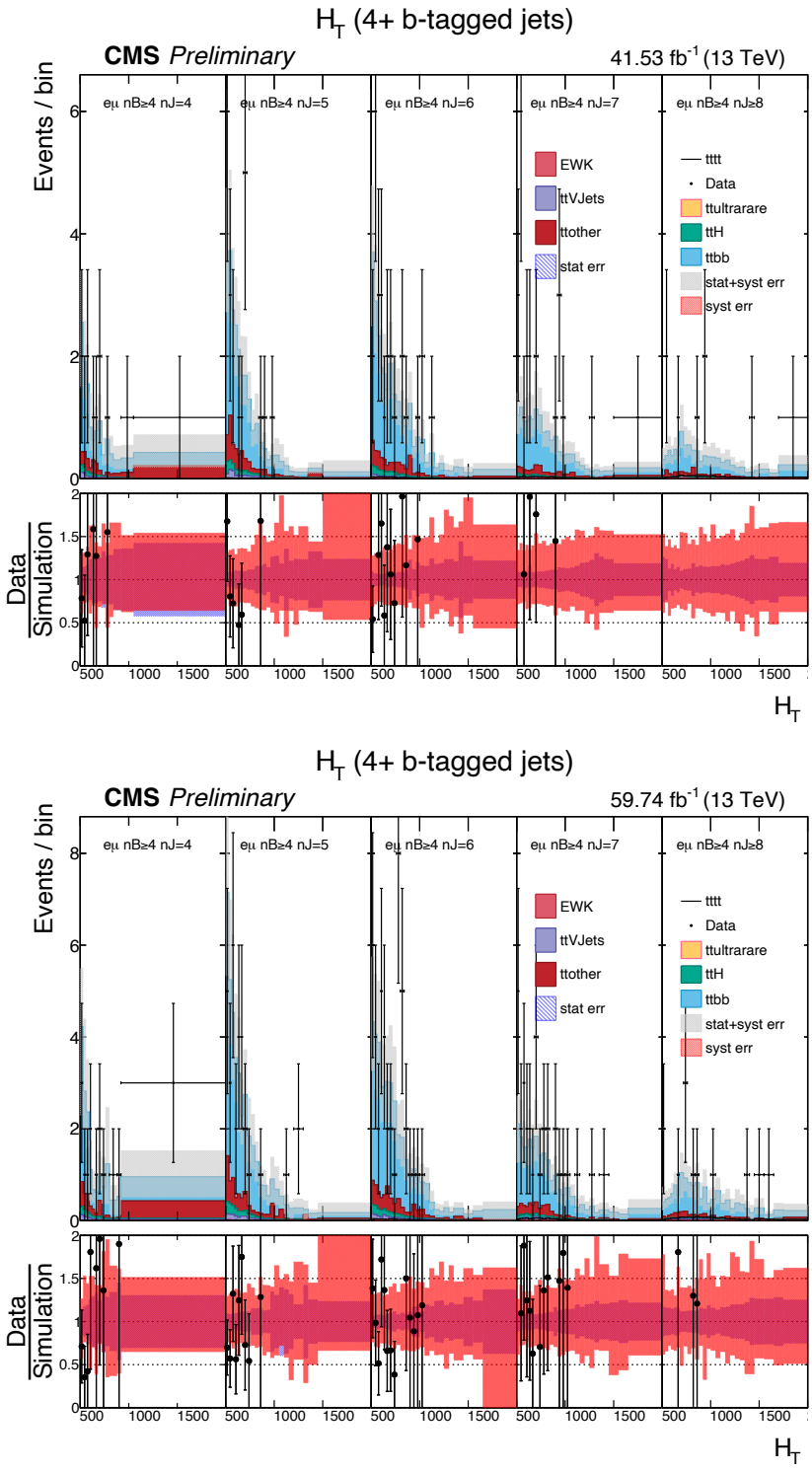

Figure D.8: H_{T} templates for the eμ channel in the 4+ btag category, 2017 top and 2018 bottom

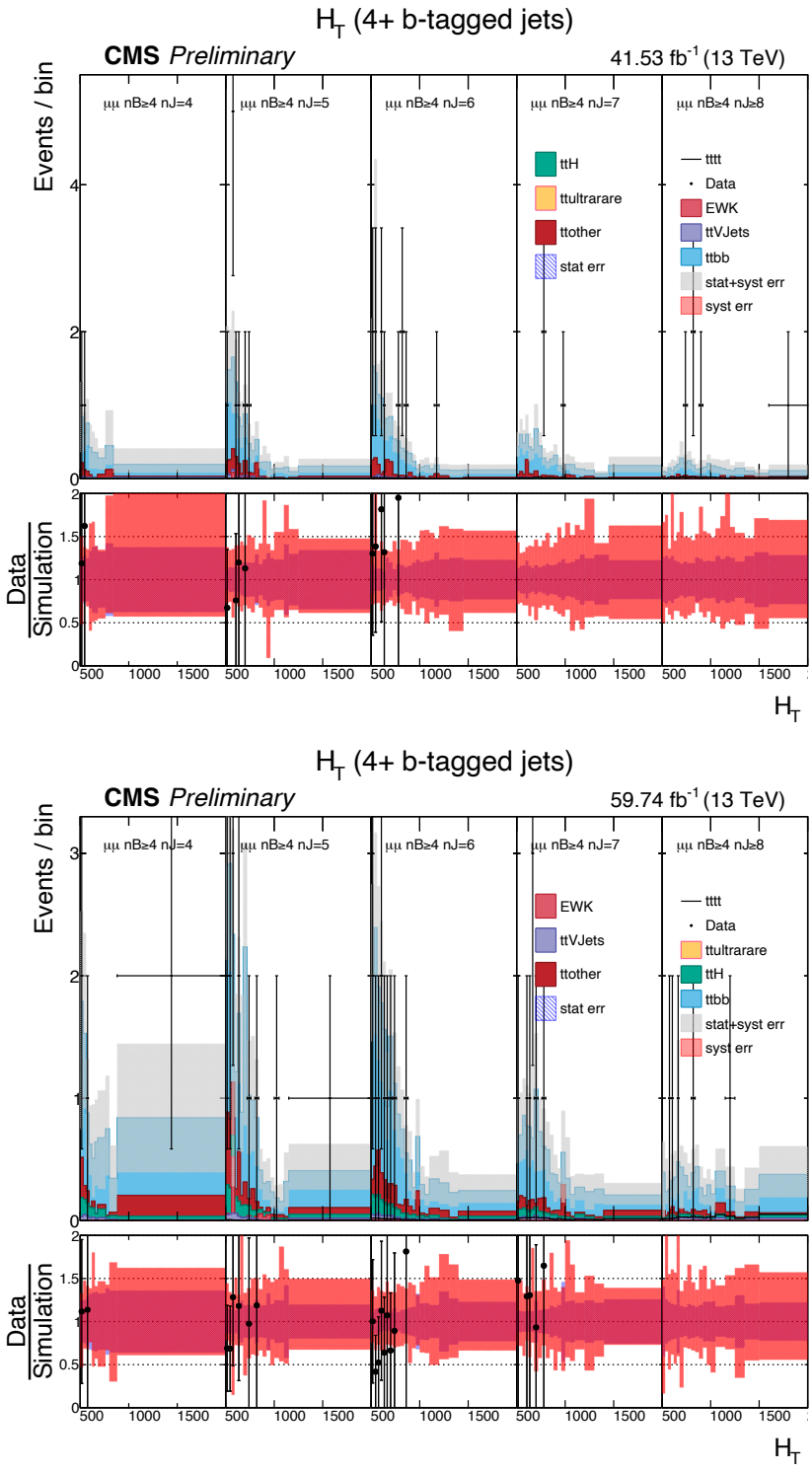

Figure D.9: H_{T} templates for the $\mu\mu$ channel in the 4+ btag category, 2017 top and 2018 bottom

D.2 Study of Jet Multiplicity Modelling and Reweighting

Previous analyses have observed discrepancies between the predicted and observed multiplicities for the additional jets in $t\bar{t}$ production. As the $t\bar{t}t\bar{t}$ signal that is being searched for is a high jet multiplicity environment, an error in the simulated jet multiplicities is potentially problematic. To address this concern we have studied multiplicities to potentially re-weight the simulated events in the large jet multiplicity region. These are determined in control regions which are insensitive to the $t\bar{t}t\bar{t}$ signal. The analysis has been performed in jet multiplicity bins of 4, 5, 6, 7, and 8+ jets for events with exactly 2 b-tagged jets. A relative jet multiplicity is extracted via the ratio of the data in a jet multiplicity bin to the data across all 5 bins, and the ratio of simulation in a jet multiplicity bin to that across all bins. By construction, the overall data to simulation normalization is ignored.

$$SF_{relative}^{nJet} = \frac{Data^{nJet}}{Data^{\geq 4jets}} \frac{MC^{\geq 4jets}}{MC^{nJet}} \tag{D.1}$$

The results are shown in Table D.1 for the 2017 and 2018 data separately and for the combined 2017+2018 dataset. The results for 2017 and 2018 are in very good agreement, and are consistent with unity where the statistical uncertainty is small. Based on this study, as the outlier for 8+ jets is covered by other systematic uncertainties, we choose not to apply a jet mutiplicity correction on the simulation.

Table D.1: Data and simulation Scale Factors ($SF^{nJet}_{relative}$) derived from the 2017, 2018 and combined 2017+2018 datasets, where the data-only statistical uncertainty is noted.

Jet multiplicity	2017	2018	2017 + 2018
4	0.98 ± 0.03	1.01 ± 0.02	1.00 ± 0.02
5	0.99 ± 0.03	0.99 ± 0.02	0.99 ± 0.02
6	1.01 ± 0.04	0.96 ± 0.03	0.98 ± 0.02
7	1.09 ± 0.06	1.03 ± 0.05	1.05 ± 0.04
8+	1.21 ± 0.09	1.16 ± 0.08	1.18 ± 0.06

D.3 Study of Limits and Impacts Under Various Alternative Scenarios

In this section, the sensitivity of the limits to a variety of alternative scenarios is presented, using the 2018 simulation and data. These cover the application of top quark p_{T} re-weighting, switching to the less efficient DeepCSV b-tagging algorithm, using tight lepton identification criteria (with corresponding ID/ISO SFs), switching from loose to tight jet PU ID, and rebinning templates with lower unweighted $\mathrm{t\bar{t}}$ statistical uncertainties (30% in the baseline scenario). The results are summarized in Table D.2 and they are discussed below in the context of performance relative to the baseline analysis. This uses the selection criteria discussed in section 4.4 together with the loose jet pileup ID corrections, no top quark p_{T} re-weighting, the Deep Jet b tagging algorithm, and loose lepton ID and associated scale factors. All comparisons should be taken as relative to the listed baseline in this section, as these were completed prior to finalization of the H_{T}-based analysis.

For comparison with the alternative scenarios, Figs. D.10 and D.11 show the H_{T} templates for the eμ channel data from 2018 and the impacts for the 3 channels combined for the baseline analysis, respectively. For the H_{T} templates the eμ channel is representative of

Table D.2: 2018 Data: Sensitivity of expected cross section limit and significance to choice of event selection/correction criteria

Change w.r.t Baseline	Limit [95% CI](fb)	Significance (s.d.)
–	44.81 [23.46, 87.95]	0.547
Apply top quark p_T reweighting	44.25 [23.34, 86.84]	0.553
Apply tight jet PU ID	45.56 [23.85, 89.42]	0.542
Rebin with 25% statistical uncertainty	45.19 [23.65, 88.40]	0.544
Rebin with 20% statistical uncertainty	45.94 [24.22, 89.57]	0.531
Rebin with 15% statistical uncertainty	46.50 [24.70, 90.67]	0.523
Rebin with 10% statistical uncertainty	47.63 [25.30, 92.74]	0.510
Use DeepCSV b tagging categorization	48.56 [25.61, 95.43]	0.506
Apply tight lepton ID (w/ matching ID/ISO SFs)	49.69 [26.01, 98.27]	0.501

the overall behavior and the 2017 and 2018 data agree with each other within the statistical uncertainty on data expected from the other channels.

D.3.1 Corrections Which Have Negligible or a Negative Effect

Switching from loose-loose to tight-tight ID for the leptons leptons produces a significant loss in data and gives a worse limit and significance. Additionally, as seen in Fig. D.26, the data to MC agreement does not improve. For completeness, the impact plot for this scenario is shown in Fig. D.27. The analysis is statistically limited, and the use of more stringent requirements on the leptons only serves to de-sensitize the searches. We conclude that this change is not justified for the current analyses.

We have also looked at using the older Deep CSV b tagging algorithm in place of the baseline DeepJet algorithm. The results for Deep CSV are shown in Figs. D.24 and D.25. Using the DeepCSV algorithm gives a significant degradation in the results for the limit and significance. Therefore, for this dissertation work the DeepJet algorithm is used exclusively, both for the main analysis as well as all further studies conducted.

Switching to the tighter jet PU ID working point modestly reduces the sensitivity of the analysis on the Asimov dataset.

More conservative binning schemes, where the unweighted $t\bar{t}$ background uncertainty is required to be less than 25%, 20%, 15%, or 10%, reduces the sensitivity as well. The loss spans the range from the Baseline scenario down to the DeepCSV and tight lepton ID scenarios.

The H_{T} templates and impacts for the baseline analysis with the addition of of top quark p_{T} re-weighting corrections are shown in Figs. D.12 and D.13. We find that the correction makes no significant change in the limit, significance or to the level of agreement between the data and simulation. We therefore choose not to apply this correction or add it as a nuisance parameter, since the other uncertainties sufficiently cover the effect.

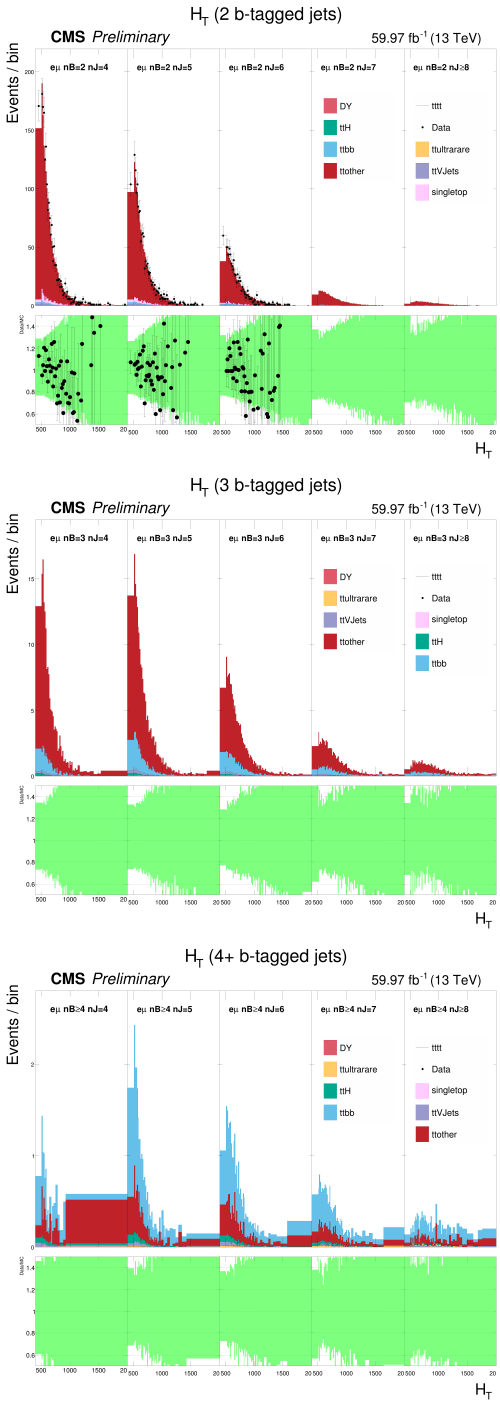

Figure D.10: H_{T} templates for the Baseline scenario

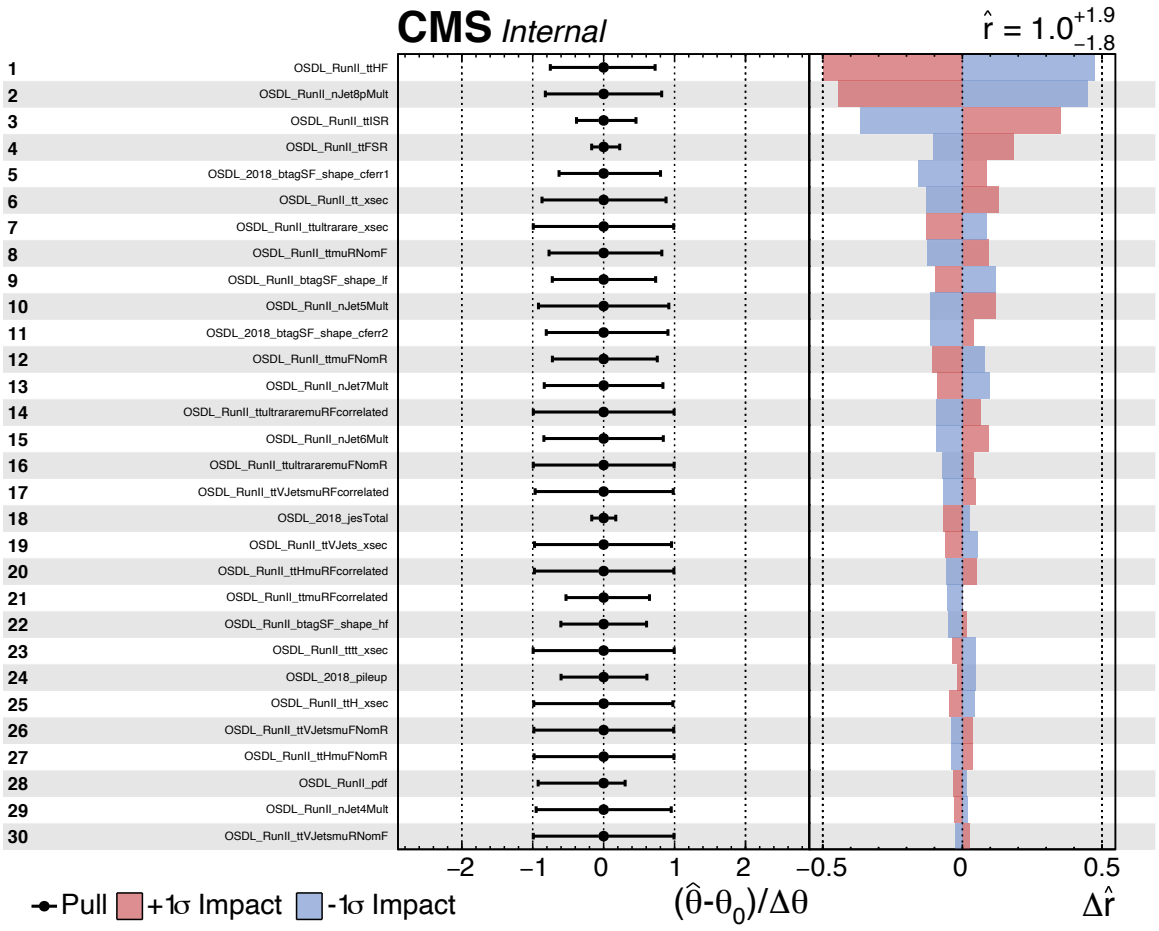

Figure D.11: Impacts for the Baseline scenario

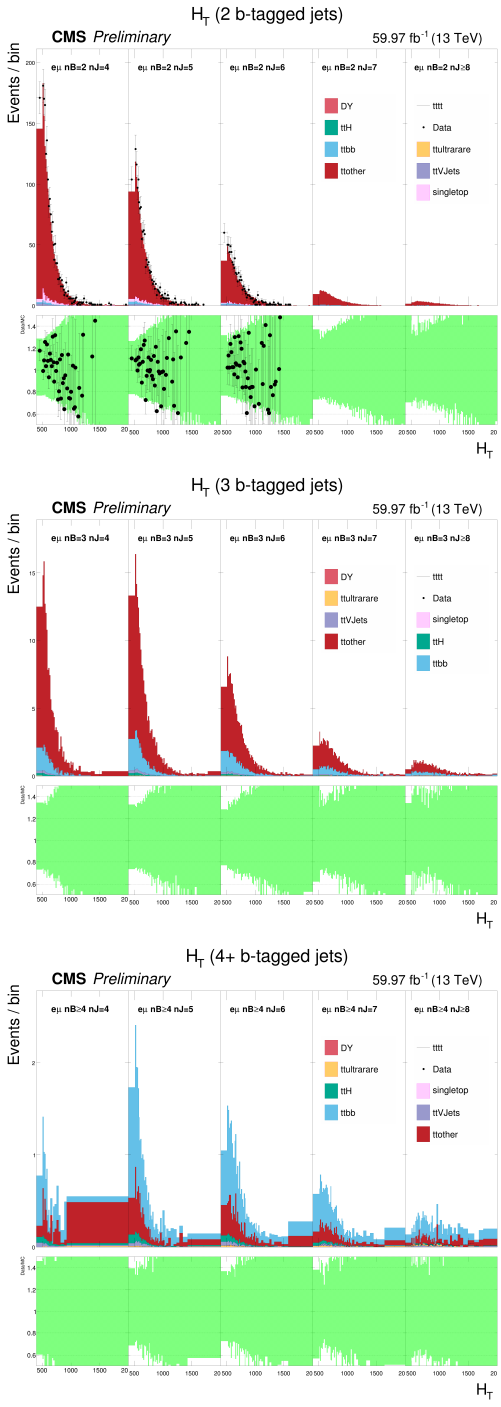

Figure D.12: H_{T} templates for the TopPtReweighting scenario

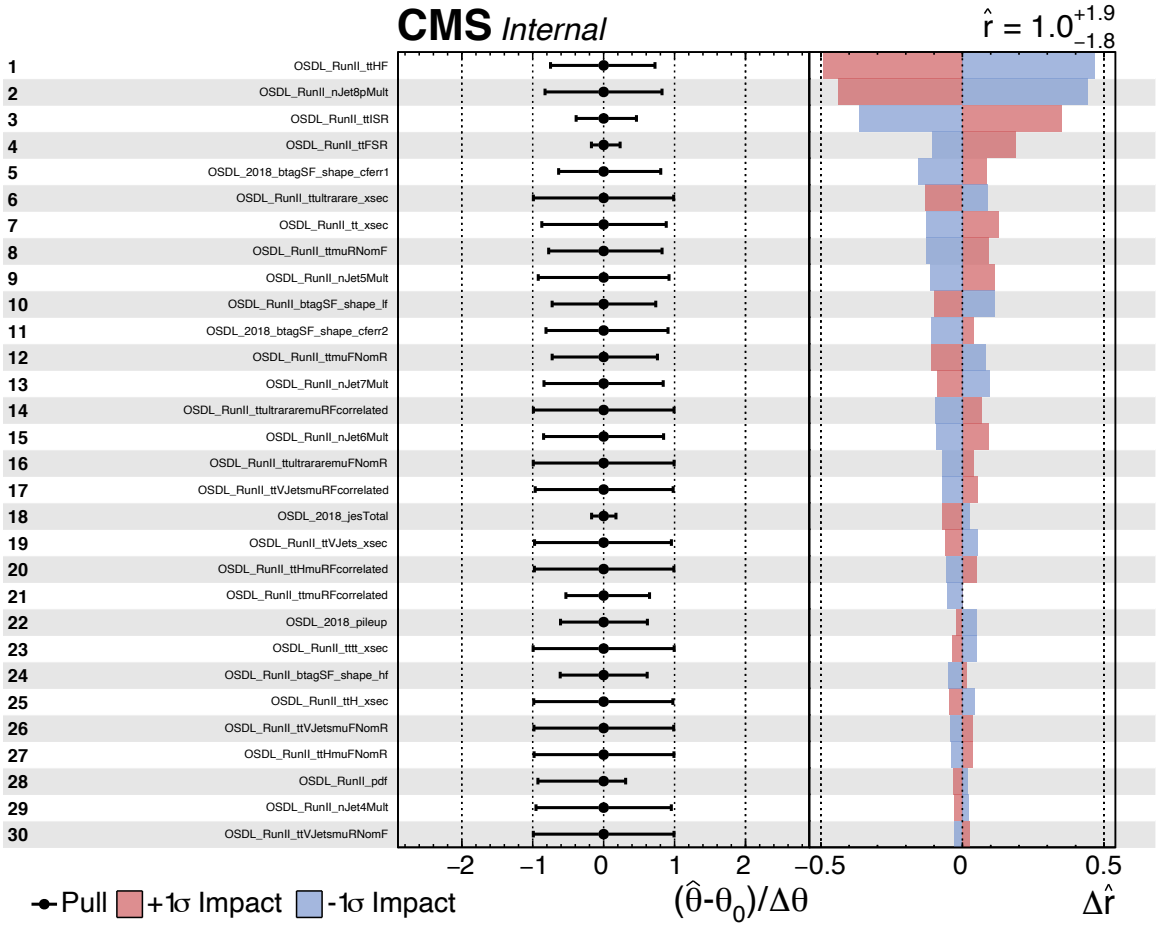

Figure D.13: Impacts for the TopPtReweighting scenario

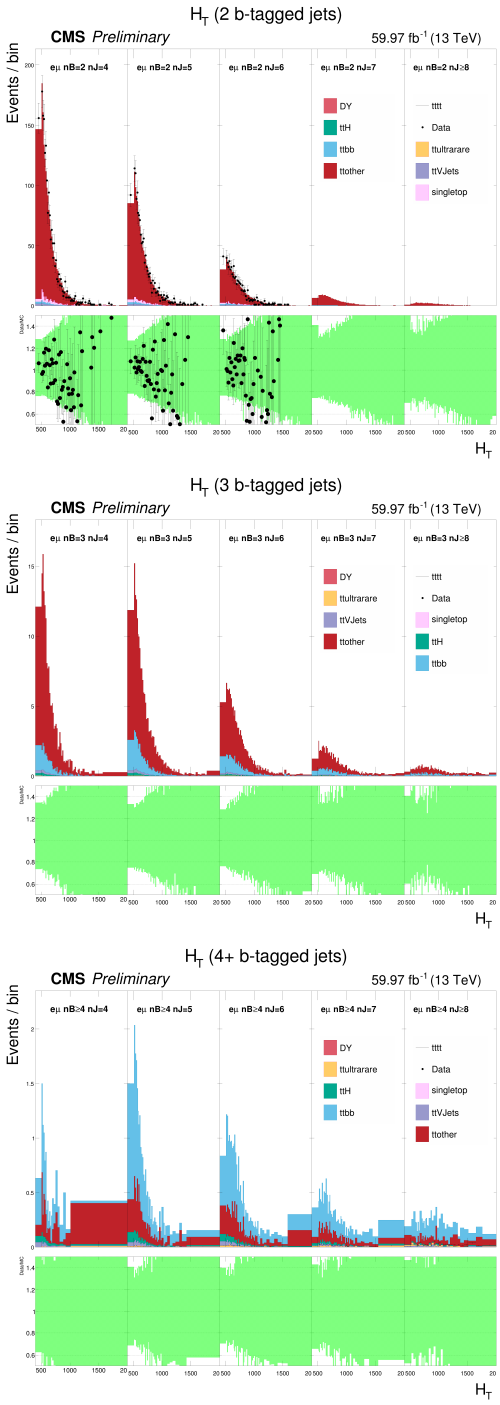

Figure D.14: H_{T} [GeV] templates for the TightPUID scenario

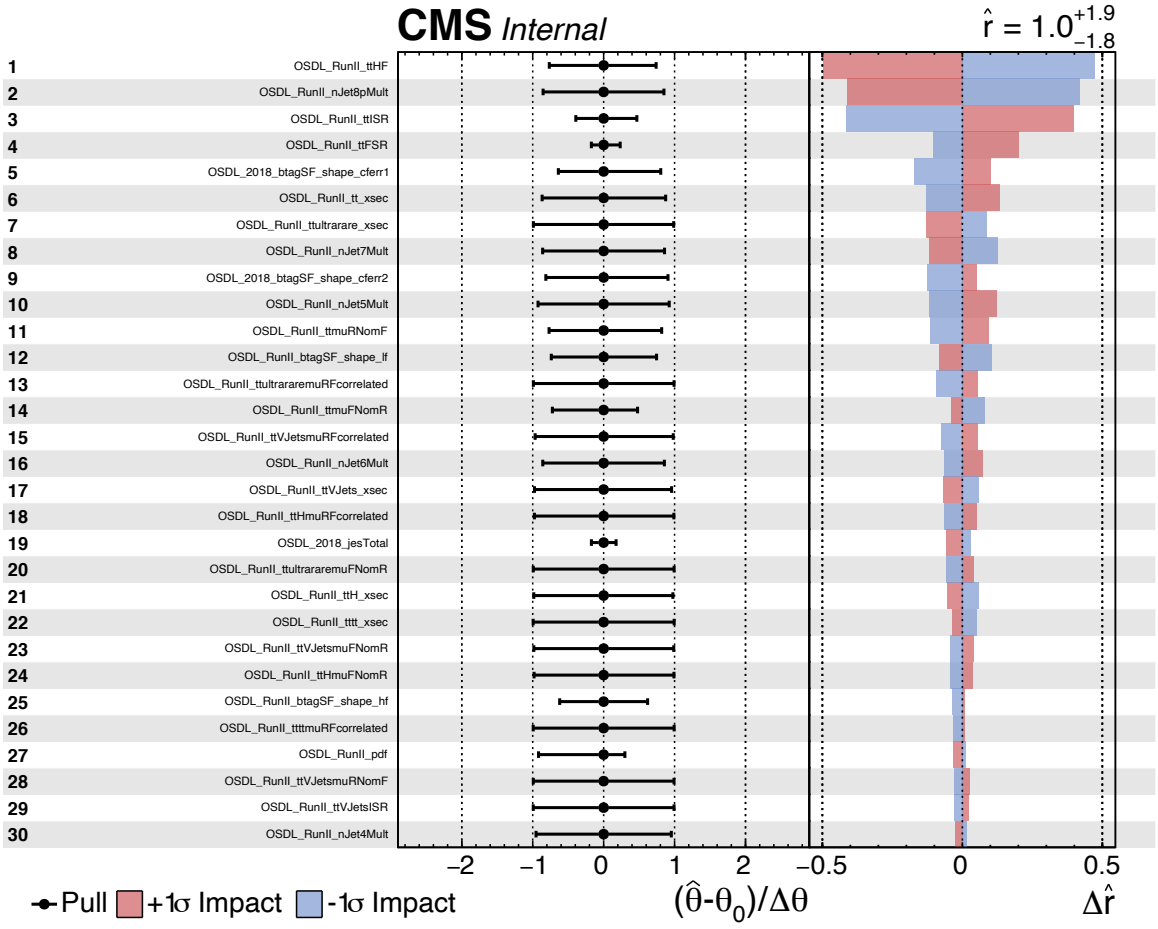

Figure D.15: Impacts for the TightPUID scenario

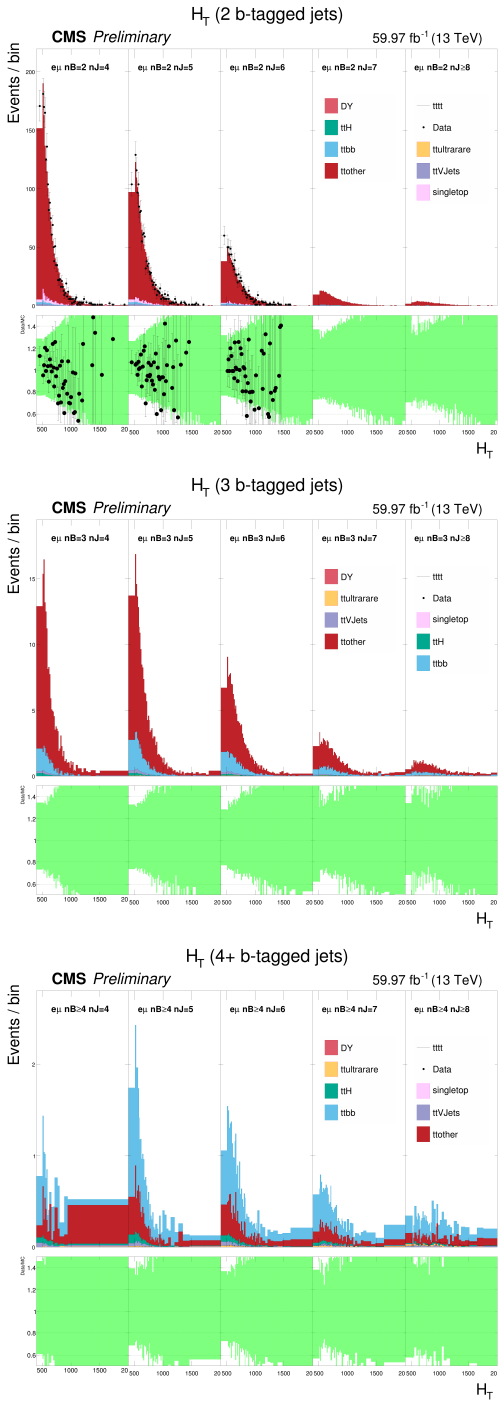

Figure D.16: H_{T} [GeV] templates for the 25 Percent Uncertainty scenario

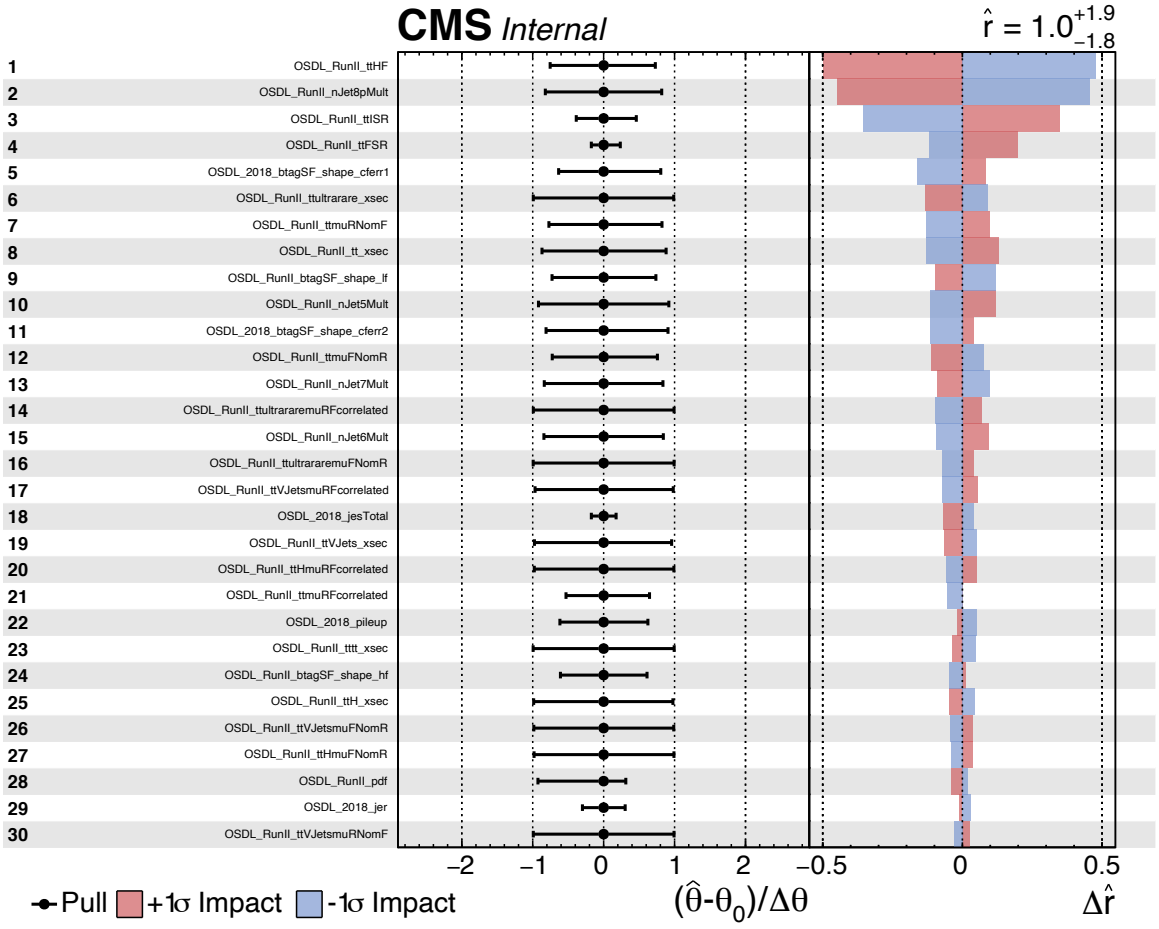

Figure D.17: Impacts for the 25 Percent Uncertainty scenario

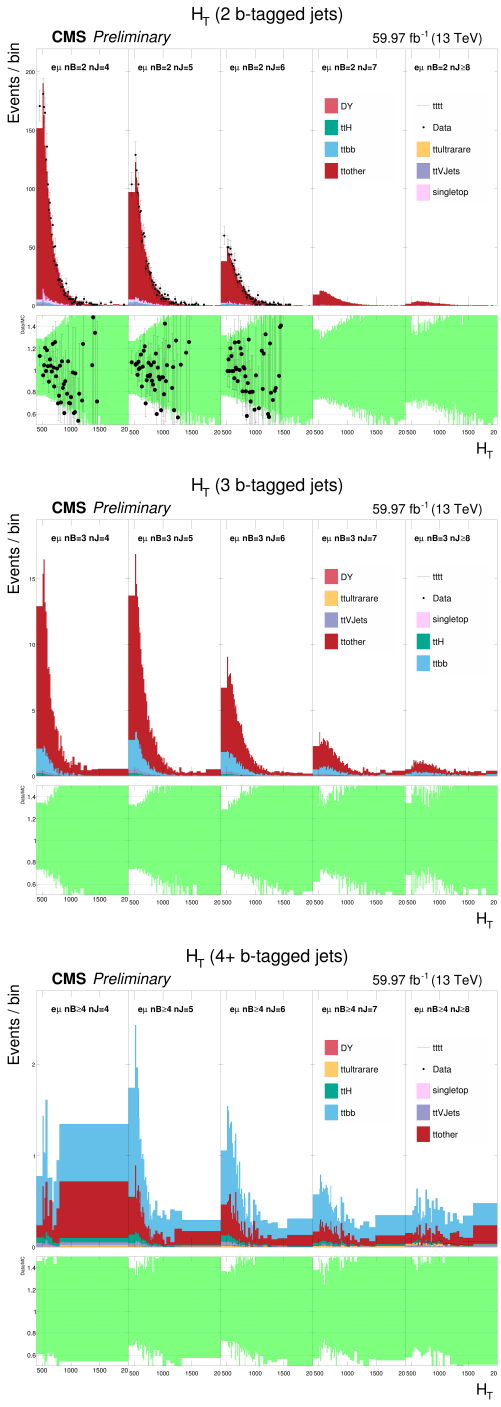

Figure D.18: H_{T} [GeV] templates for the 20 Percent Uncertainty scenario

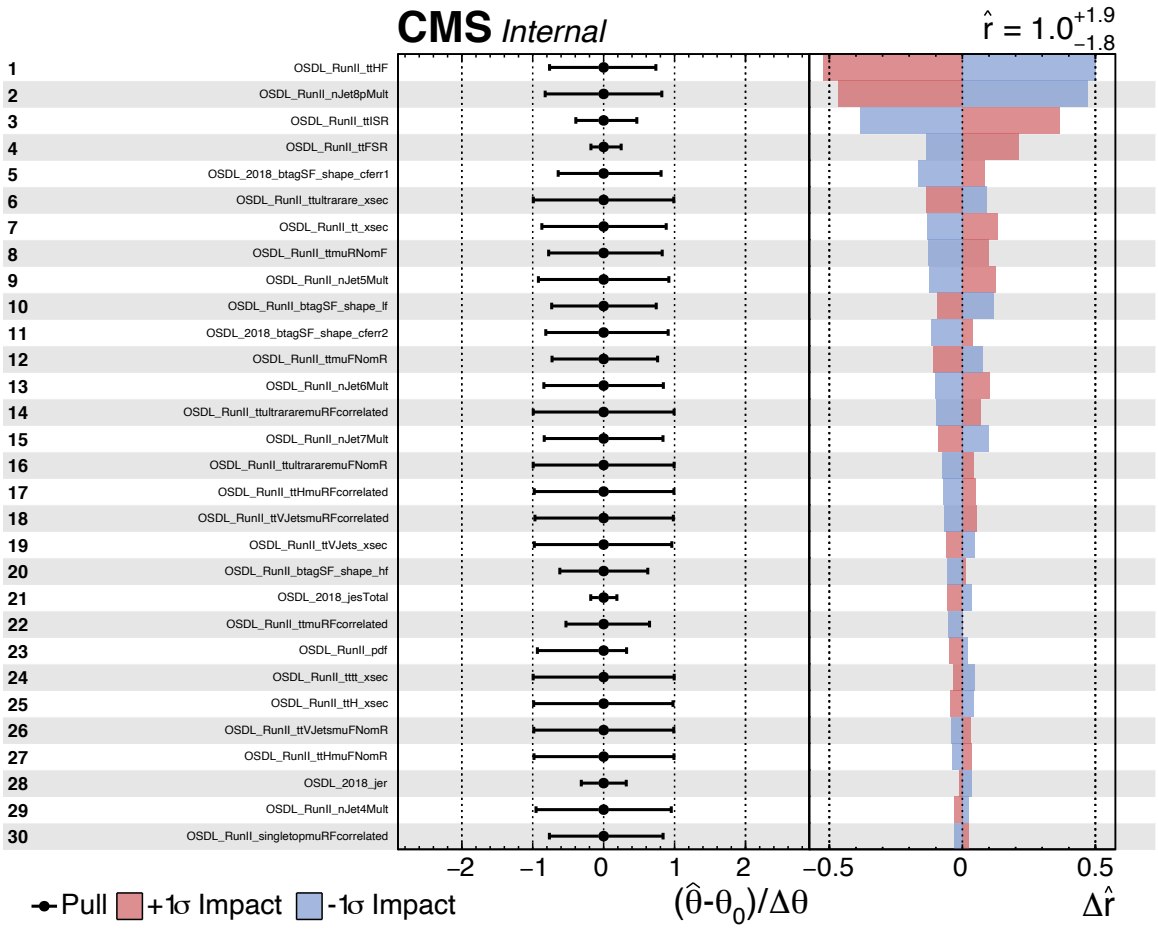

Figure D.19: Impacts for the 20 Percent Uncertainty scenario

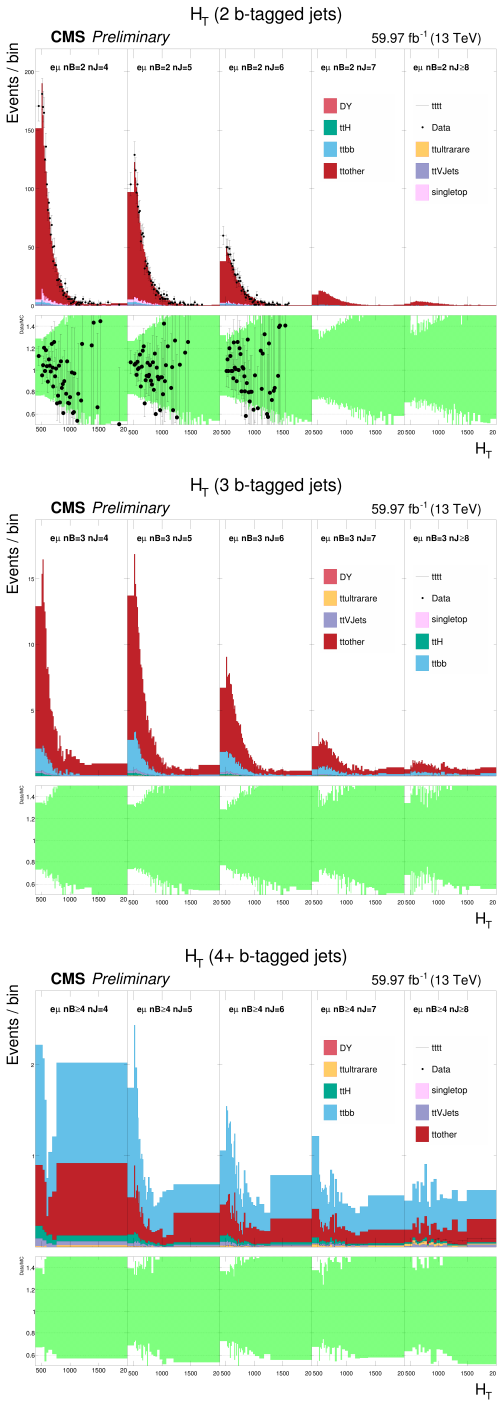

Figure D.20: H_{T} [GeV] templates for the 15 Percent Uncertainty scenario

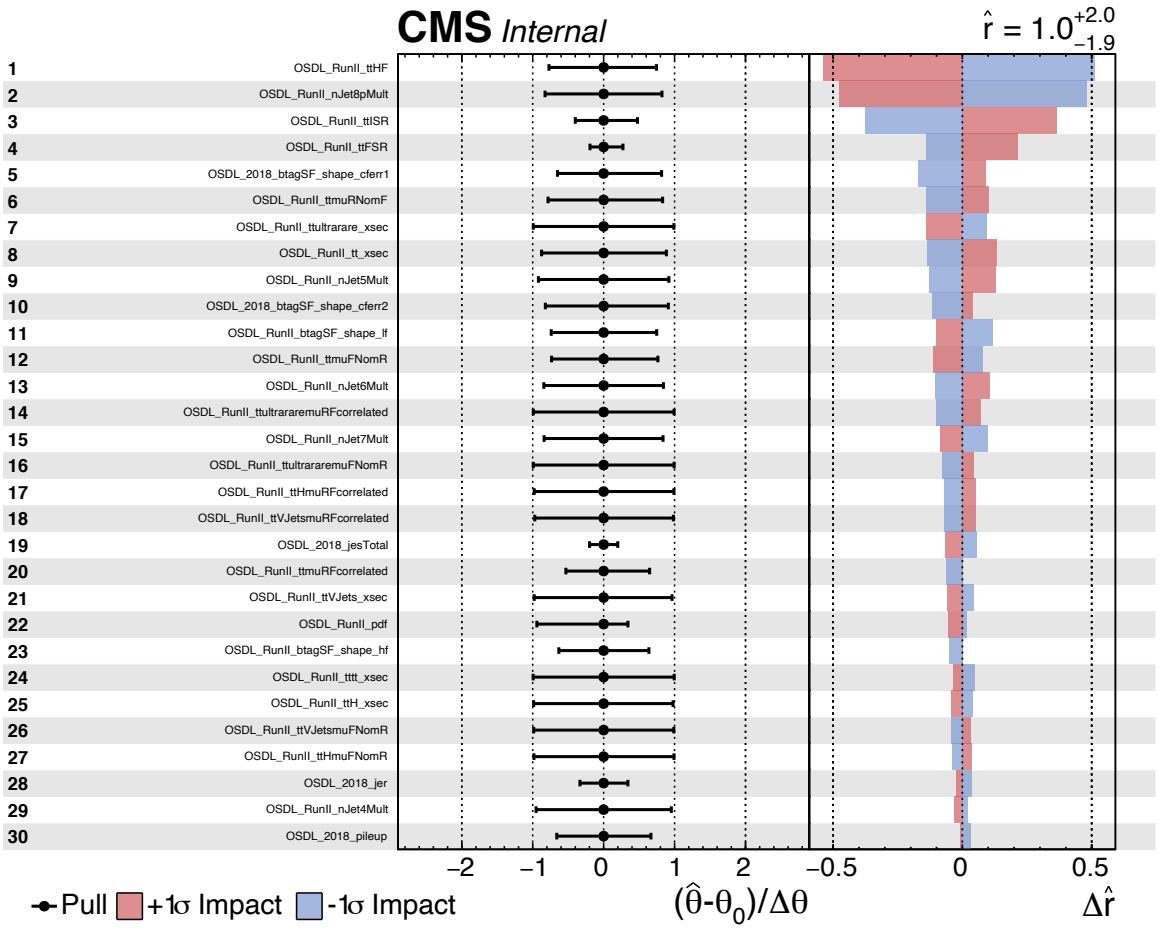

Figure D.21: Impacts for the 15 Percent Uncertainty scenario

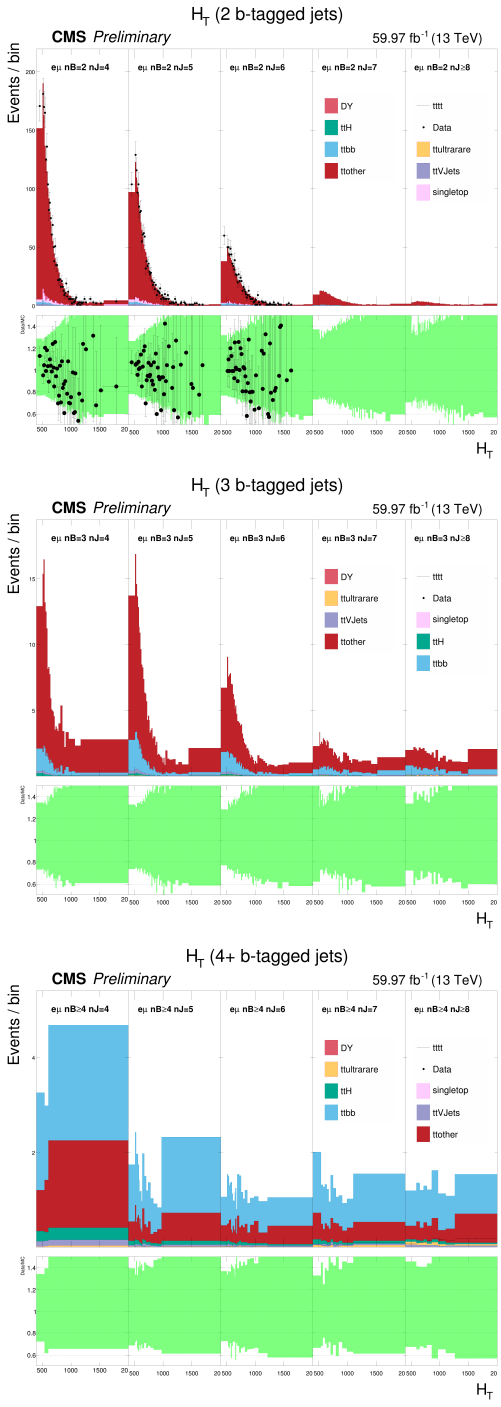

Figure D.22: H_{T} [GeV] templates for the 10 Percent Uncertainty scenario

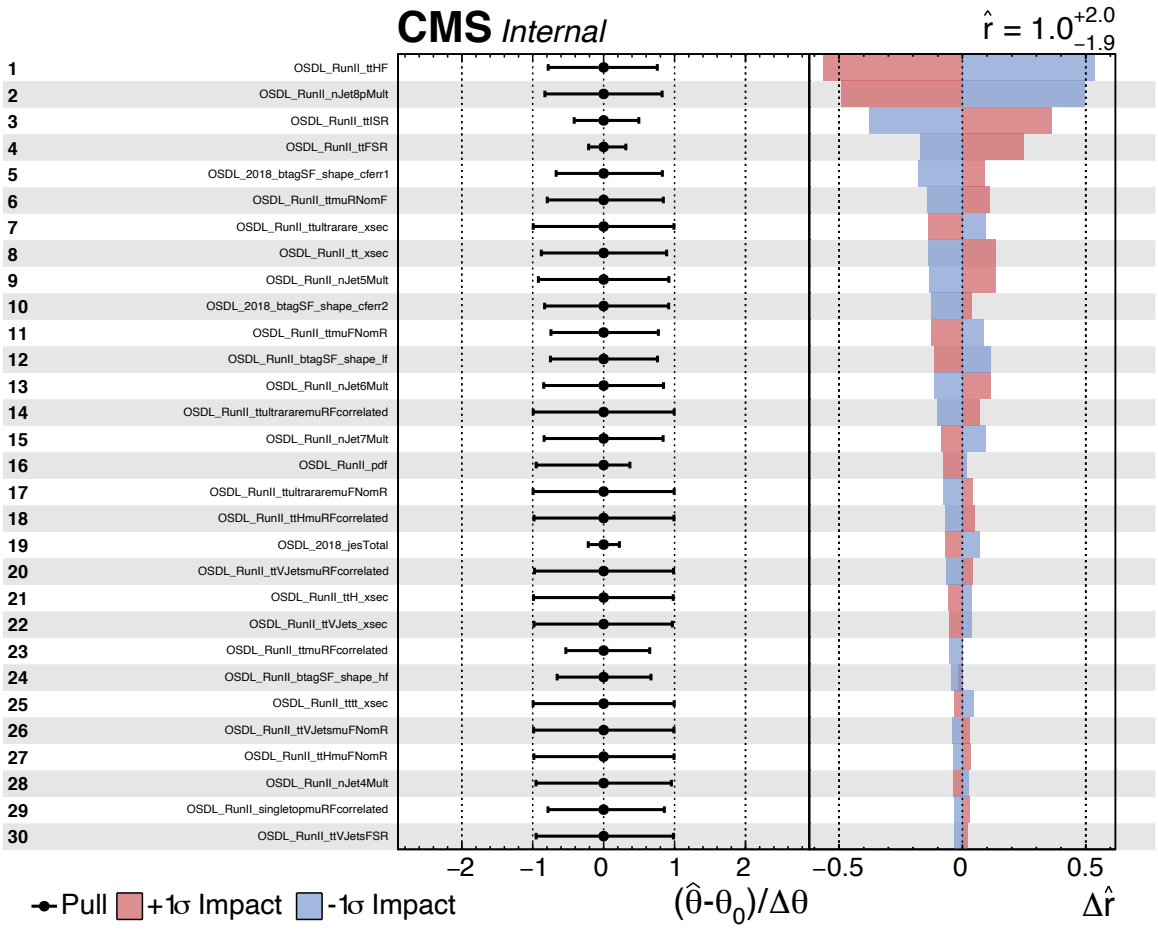

Figure D.23: Impacts for the 10 Percent Uncertainty scenario

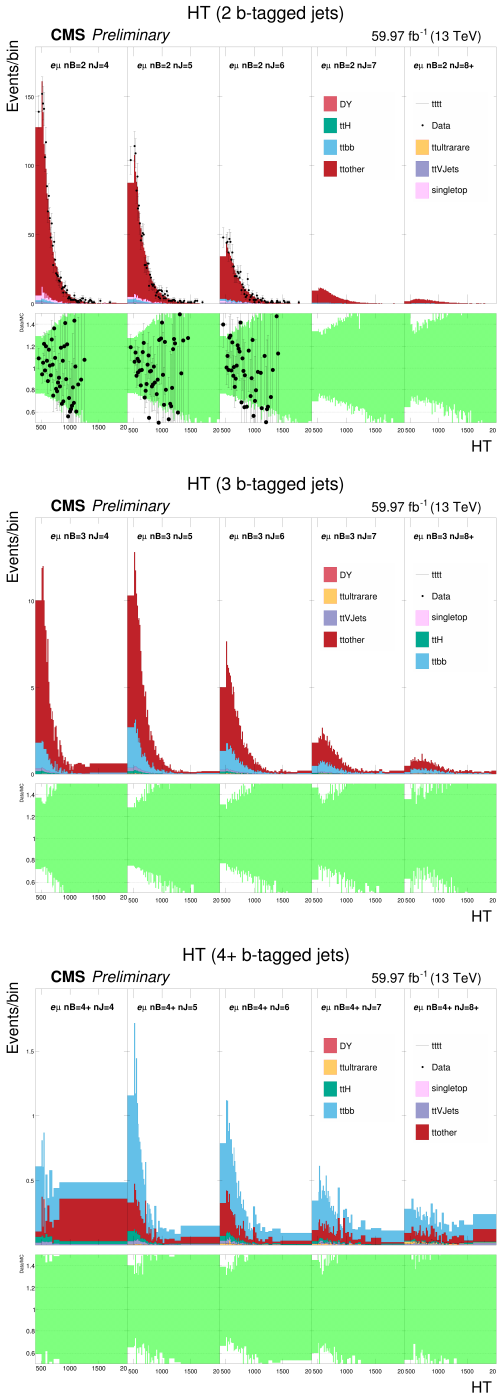

Figure D.24: H_{T} [GeV] templates for the DeepCSV scenario

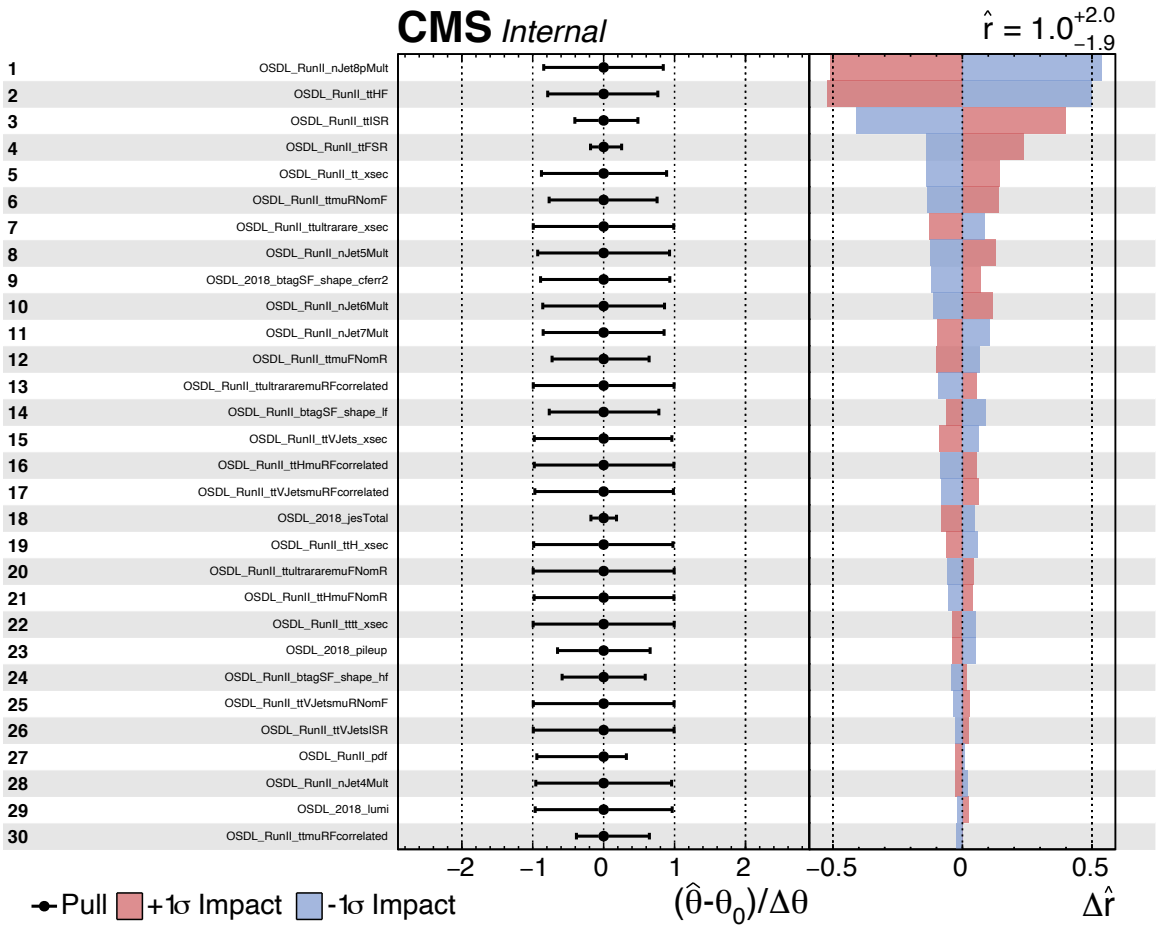

Figure D.25: Impacts for the DeepCSV scenario

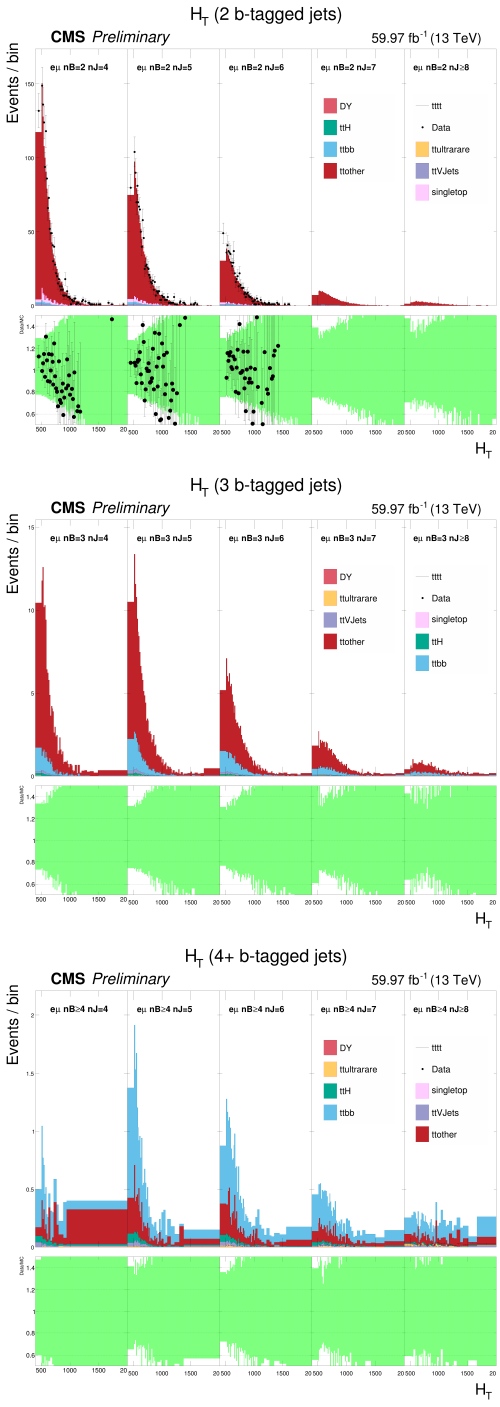

Figure D.26: H_{T} [GeV] templates for the Tight Lepton ID scenario

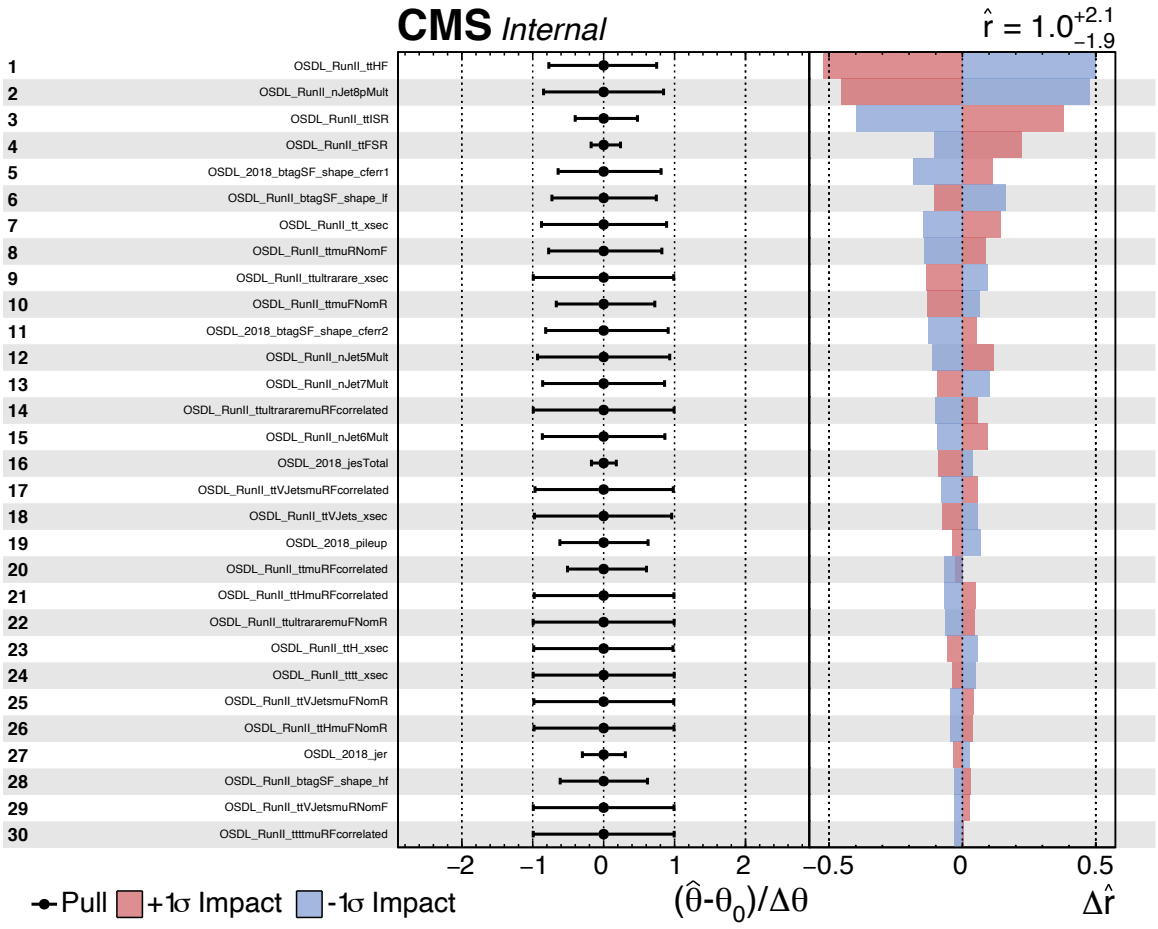

Figure D.27: Impacts for the Tight Lepton ID scenario

D.4 Stitching of Nominal and High-H_{T}, High Jet Multiplicity Samples

Because this analysis is dependent on an accurate simulation of $\mathrm{t\bar{t}}$ production wtih large numbers of jets and b-tags, the nominal $\mathrm{t\bar{t}}$ Monte Carlo does is insufficient. To correct for this, as in the 2016 OSDL analysis, special generator filtered (GF) samples have been produced for dilepton and semileptonic $\mathrm{t\bar{t}}$ channels. The generator-level cuts used to produce these samples are as follows. For the semileptonic generator: exactly one tau lepton, muon, or electron (particle or antiparticle) at LHE level; at least 9 generator jets with $p_{\mathrm{T}} > 30\,\mathrm{GeV}$; and $\Sigma p_{\mathrm{T}}^{\text{gen jet}} > 500$ using generator jets with $p_{\mathrm{T}} > 30\,\mathrm{GeV}$ and $|\eta| < 2.4$. For the dilepton generator: exactly two tau leptons, muons, or electrons (particle or antiparticle, any combination) at LHE level; at least 7 generator jets with $p_{\mathrm{T}} > 30\,\mathrm{GeV}$; and $\Sigma p_{\mathrm{T}}^{\text{gen jet}} > 500$ using generator jets with $p_{\mathrm{T}} > 30\,\mathrm{GeV}$ and $|\eta| < 2.4$.

In the matching mainline sample for each GF sample, the same cuts are applied to determine the total normalization of the filtered region. Additionally, for the nominal and GF samples, the number of entries $N = N_{\text{events}}^{\text{positive}} - N_{\text{events}}^{\text{negative}}$, where positive and negative denote the sign of the event generator weight, is calculated to assign a fractional contribution for the filtered region events from the nominal sample and the filtered region events from the GF sample, which by this point have been split into $\mathrm{t\bar{t}}$ categories using the GenHFMatcher algorithm ($\mathrm{t\bar{t}b\bar{b}}$ and non-$\mathrm{t\bar{t}b\bar{b}}$ subprocesses). The fractional contribution ($N^{\text{sample}}/\Sigma_{\text{sample}} N^{\text{sample}}$) is employed in the overall event weight for any Monte Carlo simulation from the filtered region of the nominal and filtered samples, thereby employing all of the MC simulated events, and keeping the final event weight distributions the same

regardless of whether events originated in the nominal or filtered sample. This increases the number of simulation events in this region by approximately 10%.

To verify the smoothness of the stitching process and the effect of the $t\bar{t}b\bar{b}$ renormalization, generator and reconstruction-level distributions are checked. In the figures below, the nominal $t\bar{t}$ nominal sample (without any $t\bar{t}b\bar{b}$ or fractional contribution reweighting) is shown in black. The equivalent distributions using a simple cut and replace strategy, such that events from the $t\bar{t}$ nominal sample are unused if they fall within the filtered region, are used as the denominator of the ratio plot. All events from the GF sample are used as replacement, with the same process normalization/cross section as the events cut out of the nominal sample. Finally, a third distribution shown with blue markers, combining all events from both samples is produced, using fractional contribution weighting and $t\bar{t}b\bar{b}$ rescaling. Good agreement is found, demonstrating that the stitching procedure works well at both the generator and the reconstruction levels. Additionally, it can be seen that the $t\bar{t}b\bar{b}$ rescaling effect is negligible until the high jet multiplicity region is reached. The same procedure is used for the h_{damp} Up and Down samples, as well as the Underlying Event (TuneCP5) systematic samples. The h_{damp} especially are found to have significantly different filtering efficiencies, indicating a shape morphing. The filtering efficiencies for the stitching procedure are in Table D.3.

Table D.3: Filter efficiencies calculated for semileptonic and dilepton $t\bar{t}$ samples, 2017 and 2018 EOY samples

Variation	2017 DL	2018 DL	2017 SL	2018 SL
Nominal	0.016679	0.016684	0.0061793	0.0061652
HDAMPup	0.018089	0.018131	0.0067113	0.0067297
HDAMPdown	0.015327	0.015305	0.0056451	0.0056740
TuneCP5up	0.016790	0.016813	0.0062286	0.0062316
TuneCP5down	0.016608	0.016598	0.0061086	0.0061255

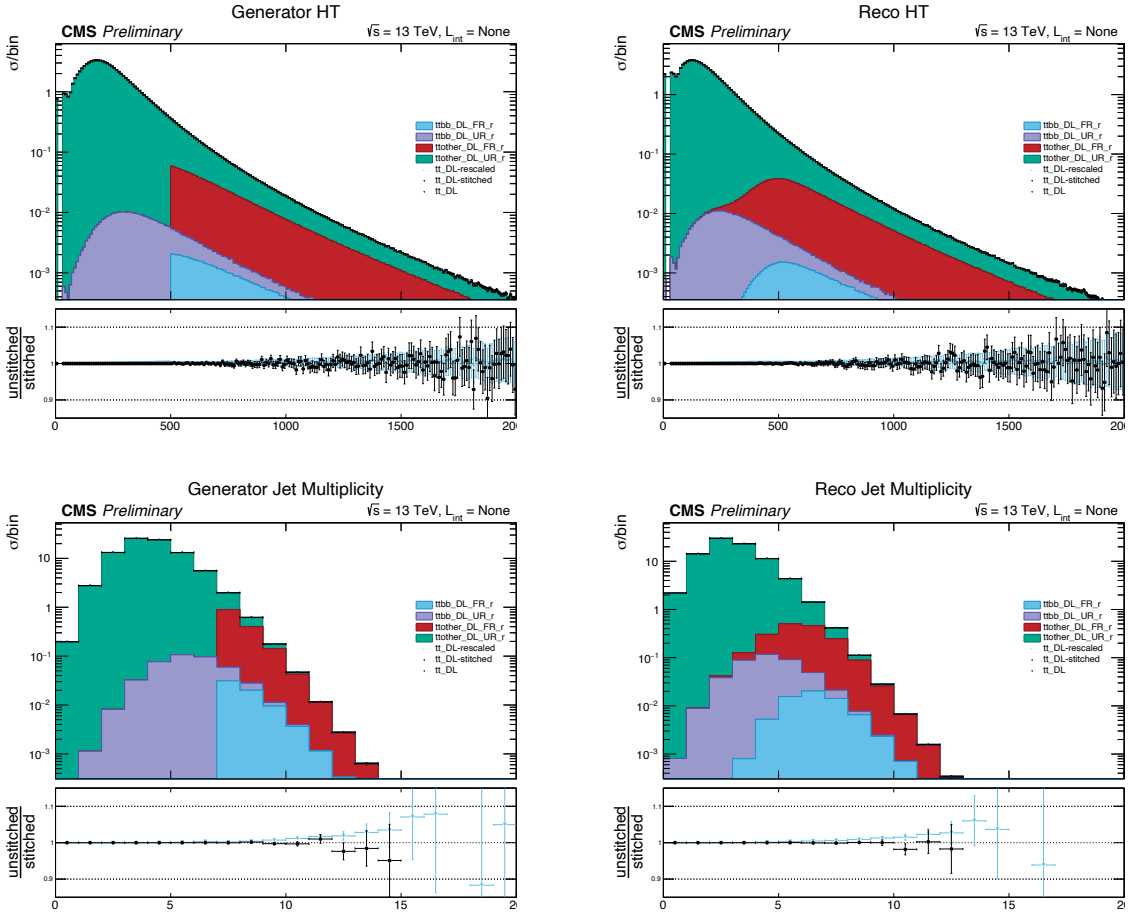

Figure D.28: Dilepton $t\bar{t}$ stitching, filter variables: Black markers denote the unstitched sample, blue markers include the $t\bar{t}b\bar{b}$ correction, both are divided by the stitched sample. Filled histograms show the rescaled process from the $t\bar{t}b\bar{b}$ and non-$t\bar{t}b\bar{b}$ subprocesses, divided into the Filtered Region (FR) and Unfiltered Region (UR).

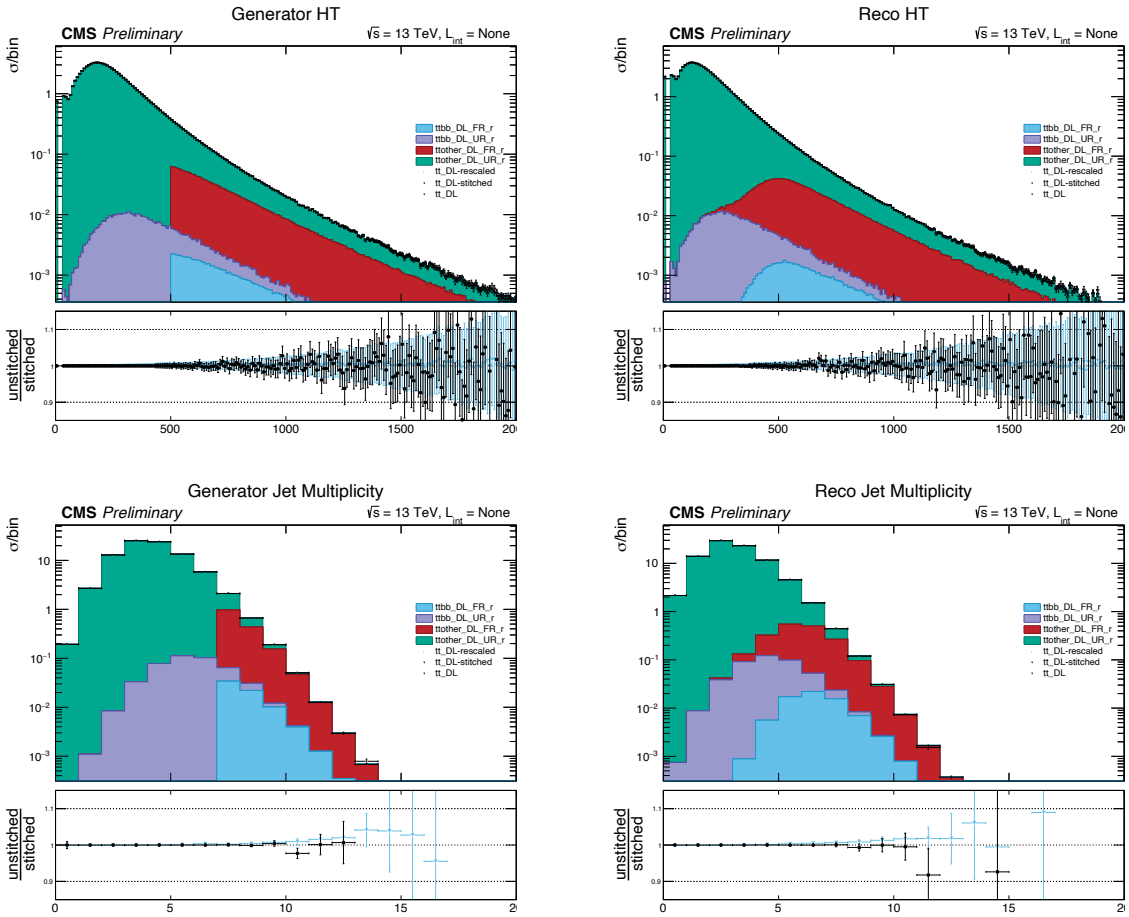

Figure D.29: Dilepton $\mathrm{t\bar{t}}$ h_{damp}-Up stitching, filter variables: Black markers denote the unstitched sample, blue markers include the $\mathrm{t\bar{t}b\bar{b}}$ correction, both are divided by the stitched sample. Filled histograms show the rescaled process from the $\mathrm{t\bar{t}b\bar{b}}$ and non-$\mathrm{t\bar{t}b\bar{b}}$ subprocesses, divided into the Filtered Region (FR) and Unfiltered Region (UR).

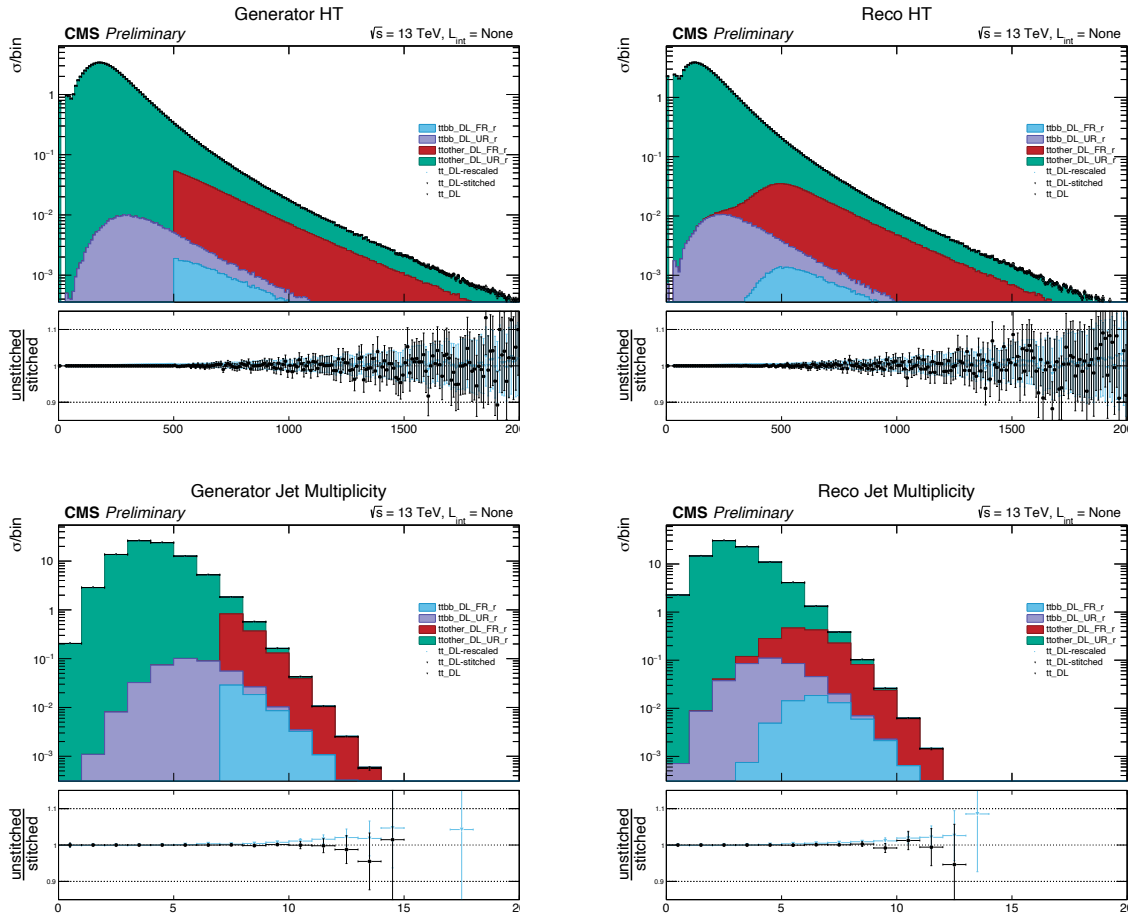

Figure D.30: Dilepton $\mathrm{t\bar{t}}$ h_{damp}-Down stitching, filter variables: Black markers denote the unstitched sample, blue markers include the $\mathrm{t\bar{t}b\bar{b}}$ correction, both are divided by the stitched sample. Filled histograms show the rescaled process from the $\mathrm{t\bar{t}b\bar{b}}$ and non-$\mathrm{t\bar{t}b\bar{b}}$ subprocesses, divided into the Filtered Region (FR) and Unfiltered Region (UR).

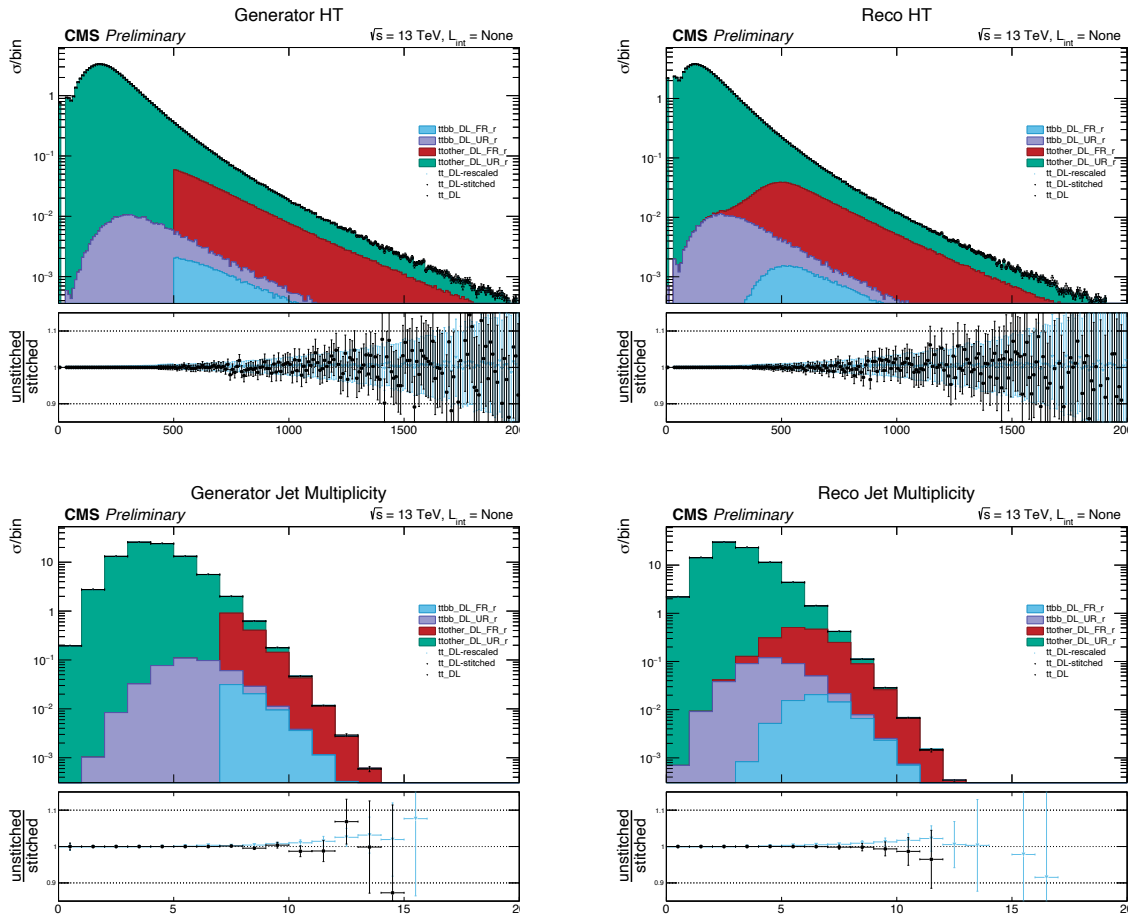

Figure D.31: Dilepton $t\bar{t}$ UE-Up stitching, filter variables: Black markers denote the unstitched sample, blue markers include the $t\bar{t}b\bar{b}$ correction, both are divided by the stitched sample. Filled histograms show the rescaled process from the $t\bar{t}b\bar{b}$ and non-$t\bar{t}b\bar{b}$ subprocesses, divided into the Filtered Region (FR) and Unfiltered Region (UR).

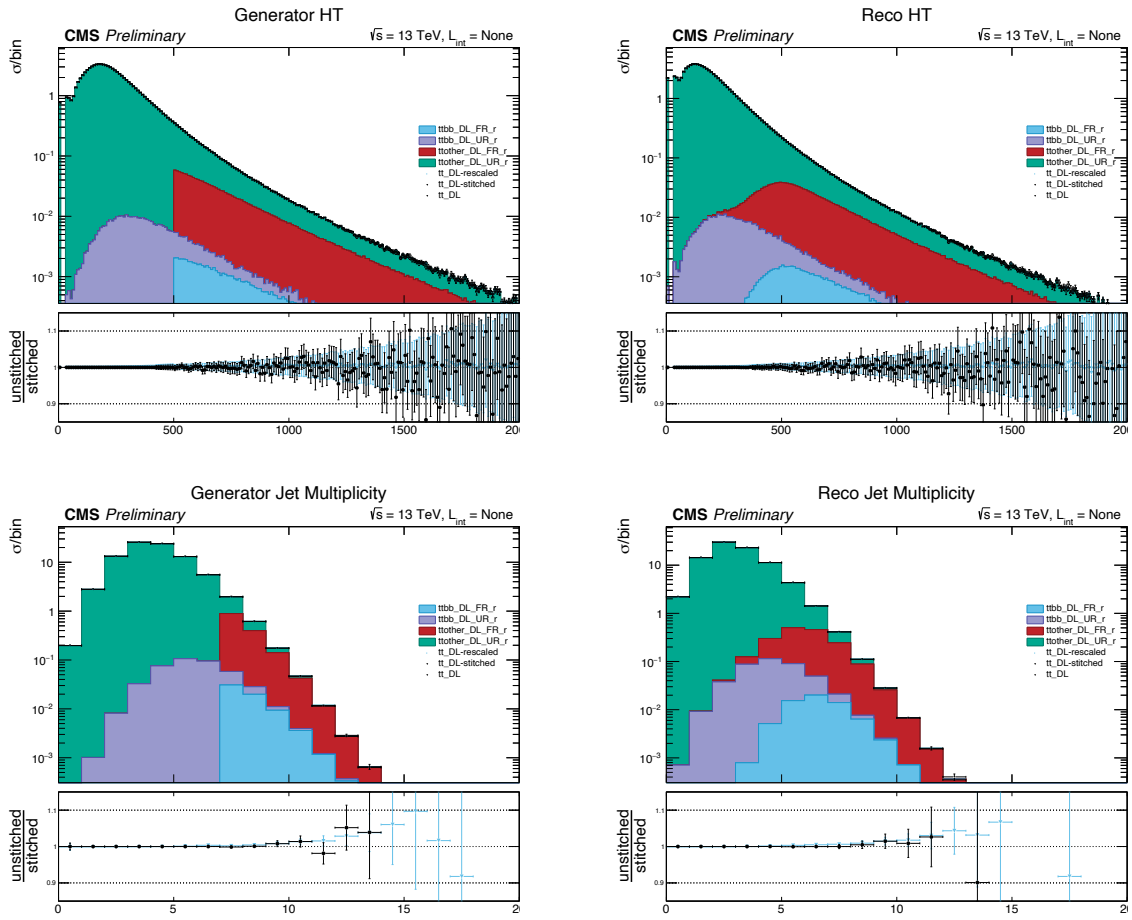

Figure D.32: Dilepton $t\bar{t}$ UE-Down stitching, filter variables: Black markers denote the unstitched sample, blue markers include the $t\bar{t}b\bar{b}$ correction, both are divided by the stitched sample. Filled histograms show the rescaled process from the $t\bar{t}b\bar{b}$ and non-$t\bar{t}b\bar{b}$ subprocesses, divided into the Filtered Region (FR) and Unfiltered Region (UR).

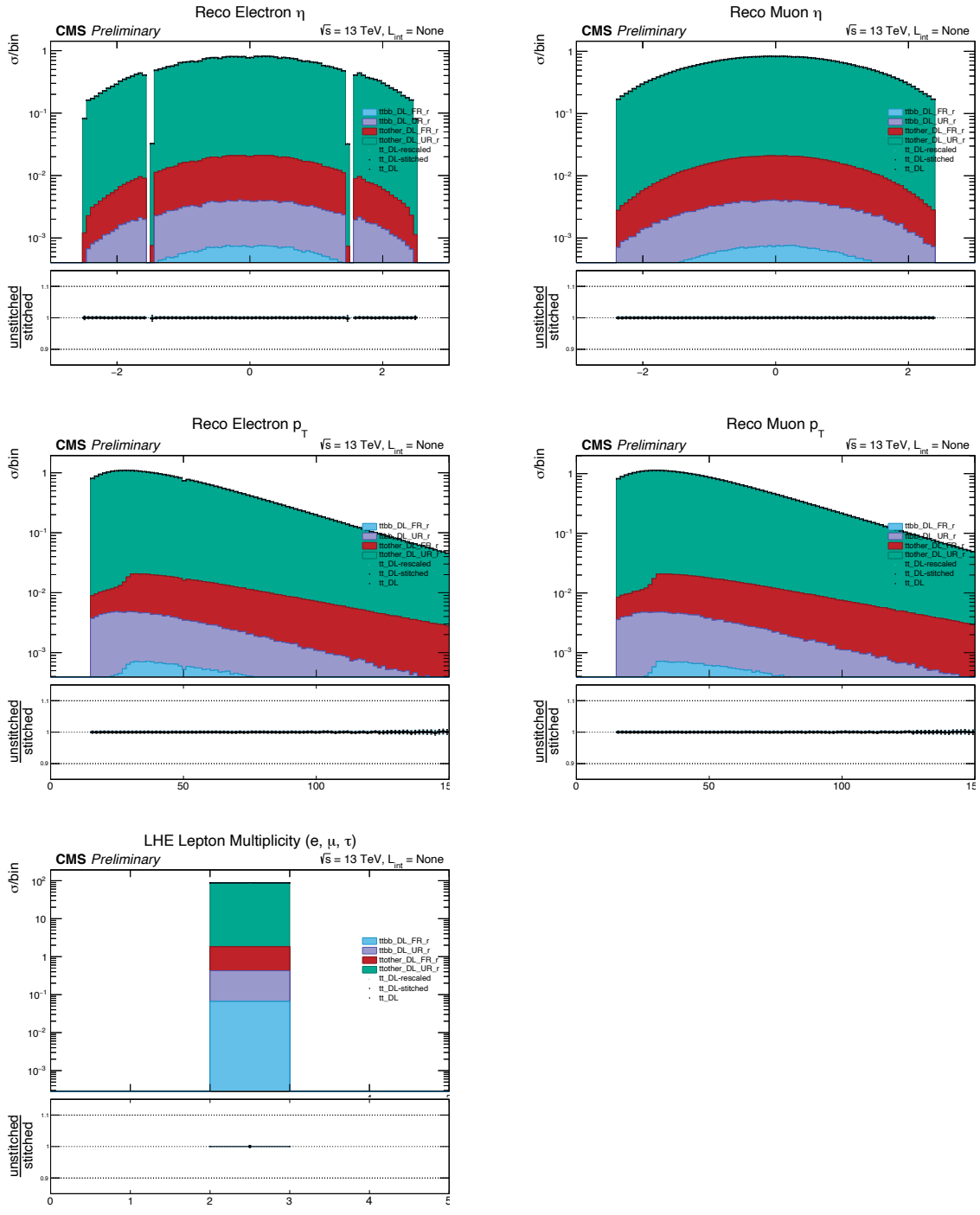

Figure D.33: Dilepton $t\bar{t}$ stitching, leptons

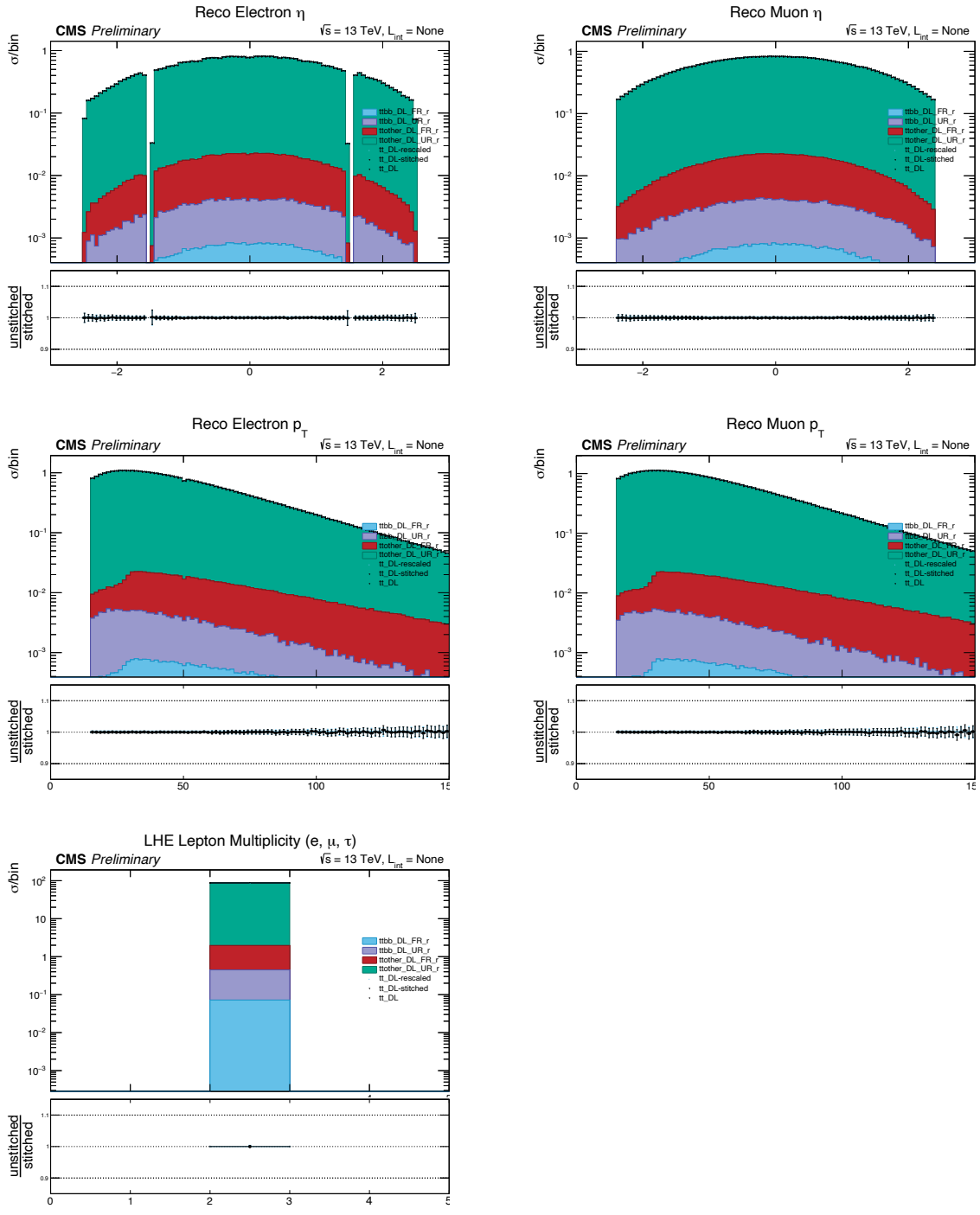

Figure D.34: Dilepton $\mathrm{t\bar{t}}$ h_{damp}-Up stitching, leptons

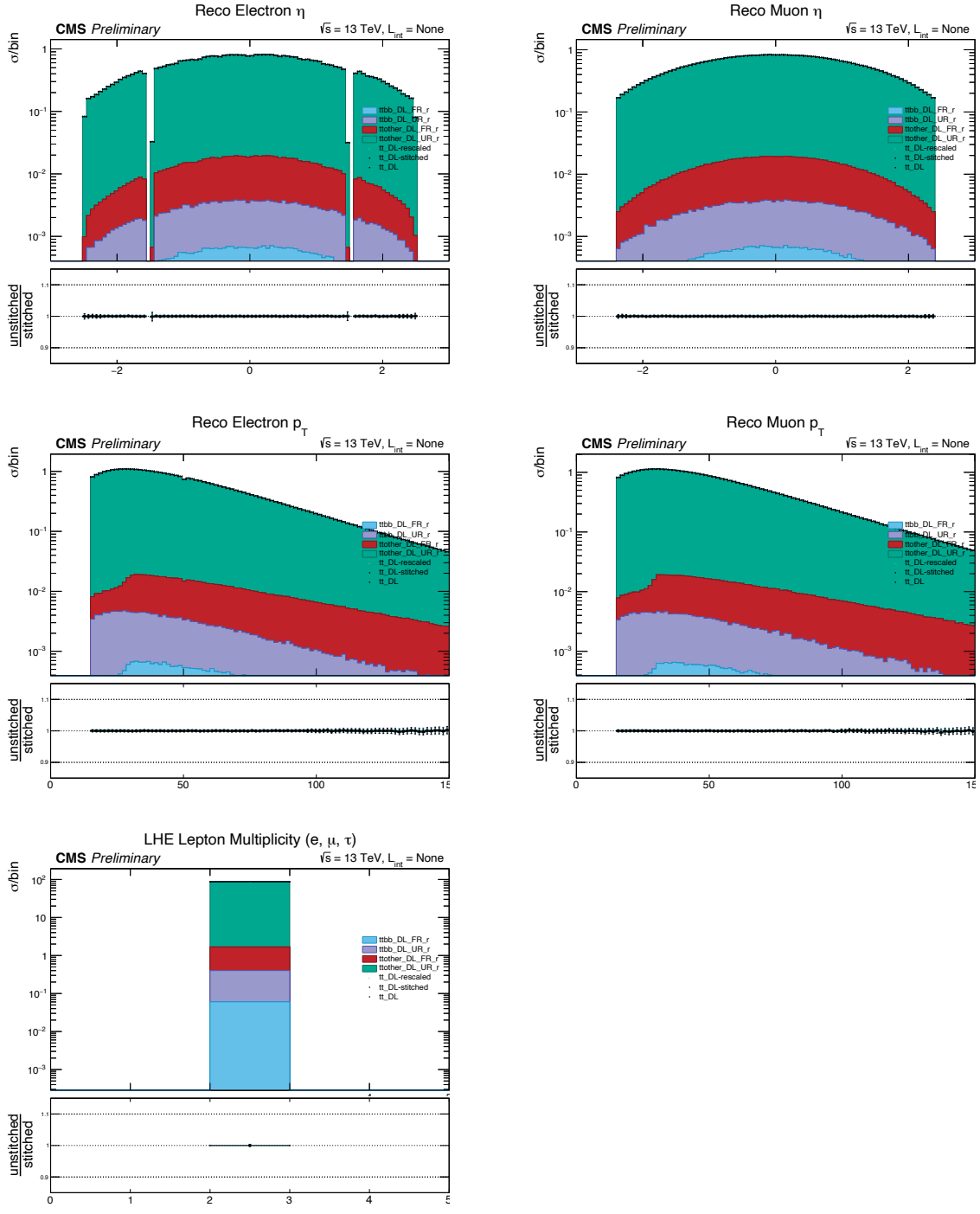

Figure D.35: Dilepton $t\bar{t}$ h_{damp}-Down stitching, leptons

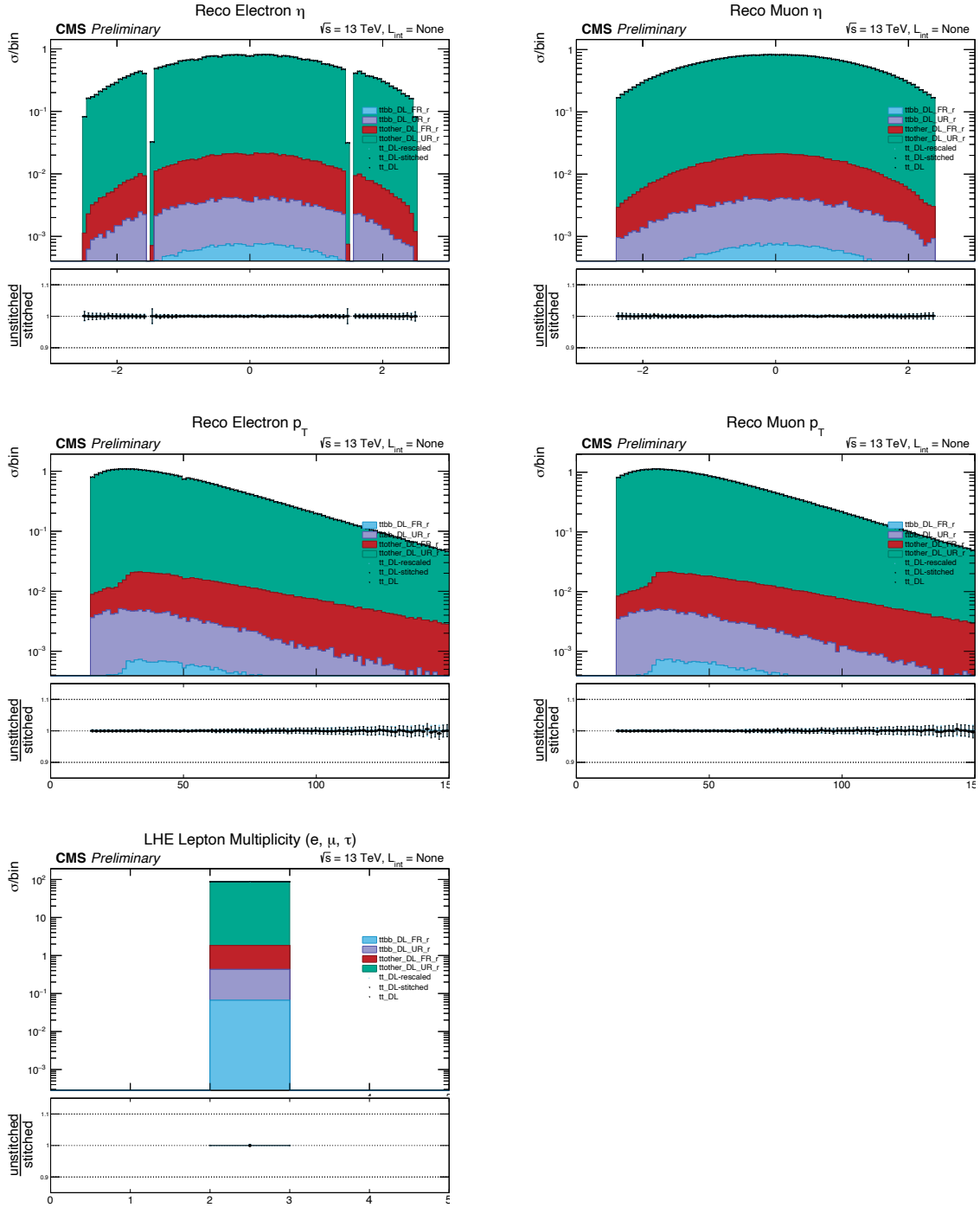

Figure D.36: Dilepton $t\bar{t}$ UE-Up stitching, leptons

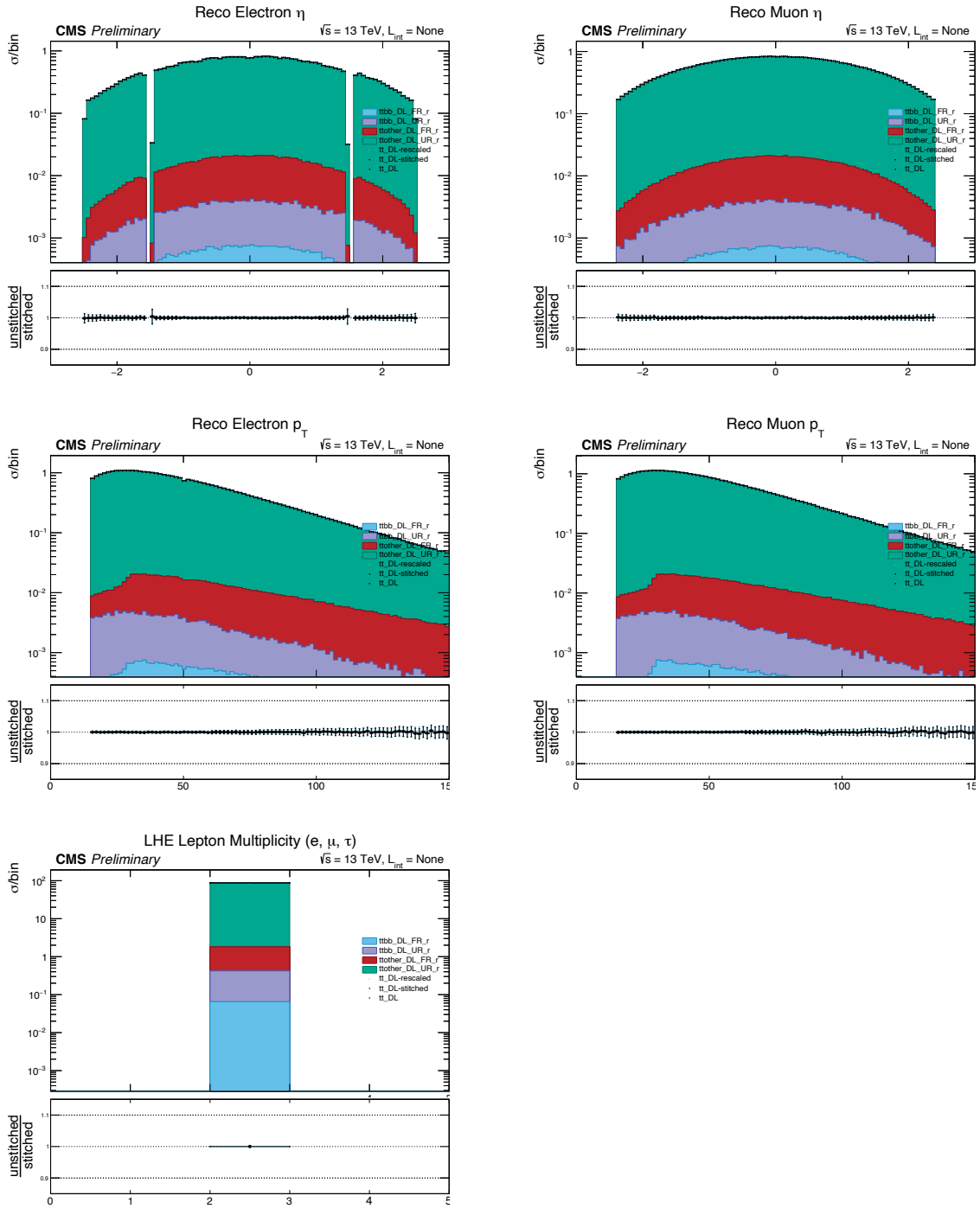

Figure D.37: Dilepton $t\bar{t}$ UE-Down stitching, leptons

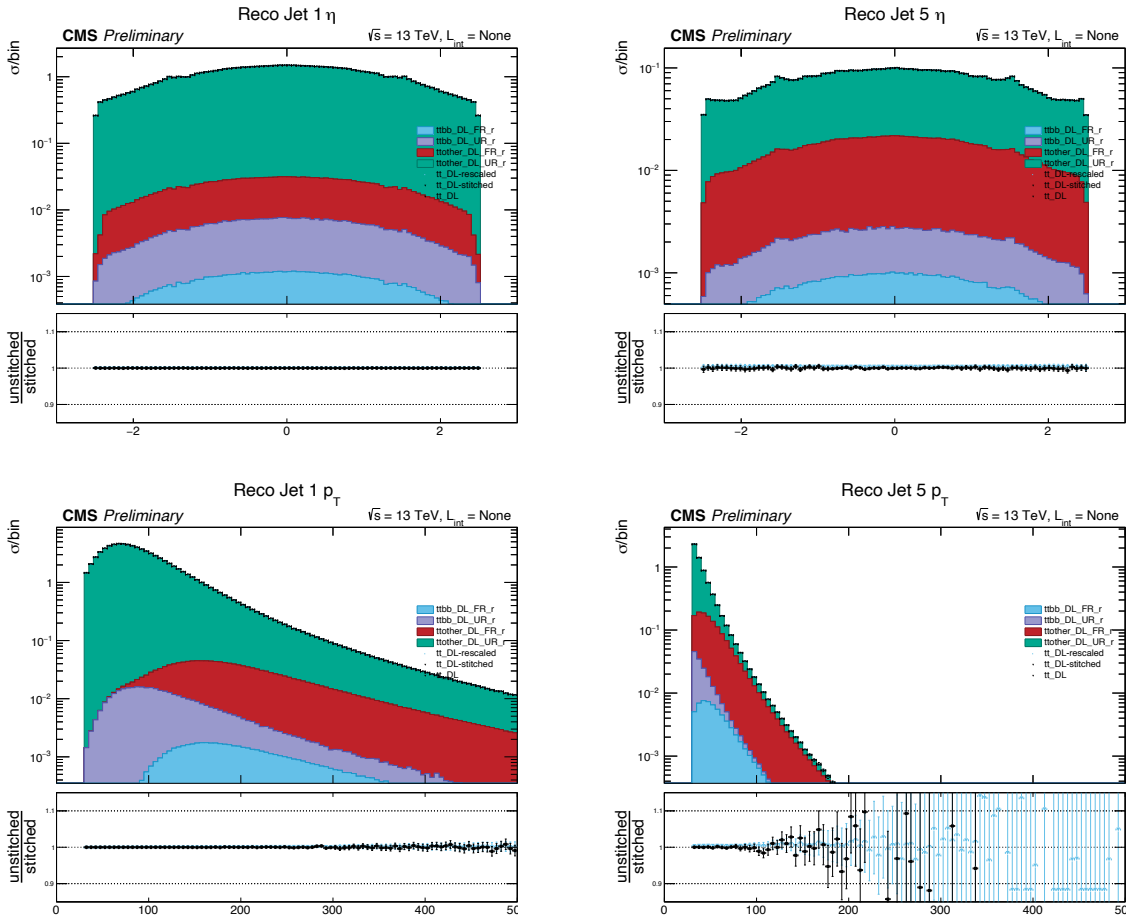

Figure D.38: Dilepton $t\bar{t}$ stitching, jets

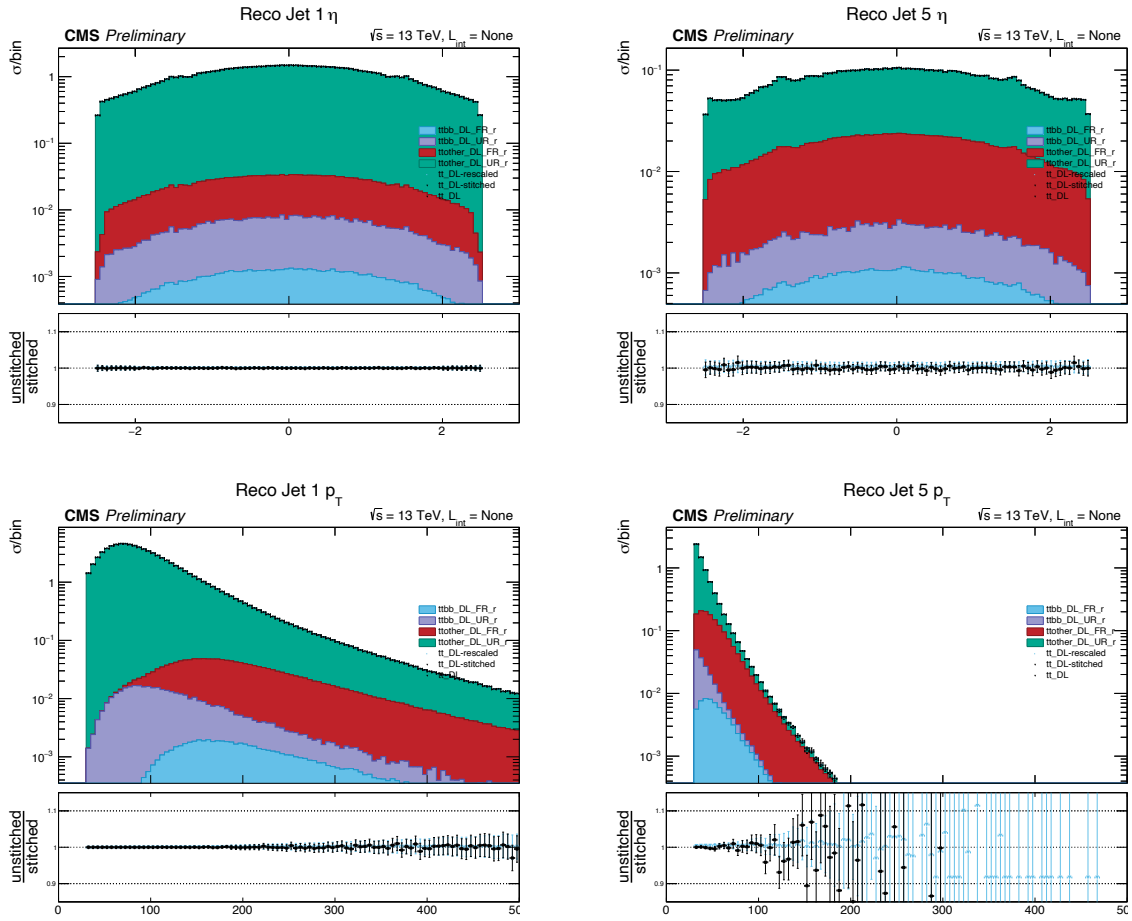

Figure D.39: Dilepton $t\bar{t}$ h_{damp}-Up stitching, jets

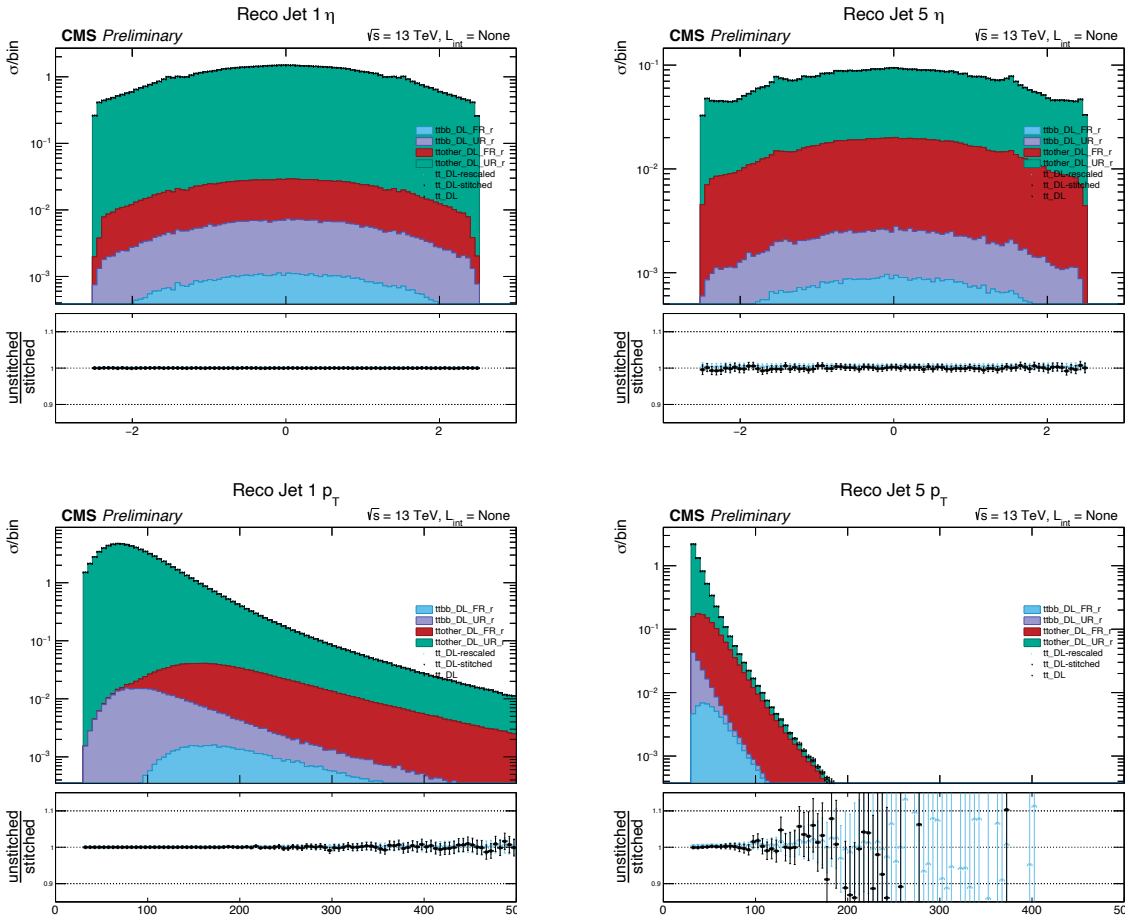

Figure D.40: Dilepton $t\bar{t}$ h_{damp}-Down stitching, jets

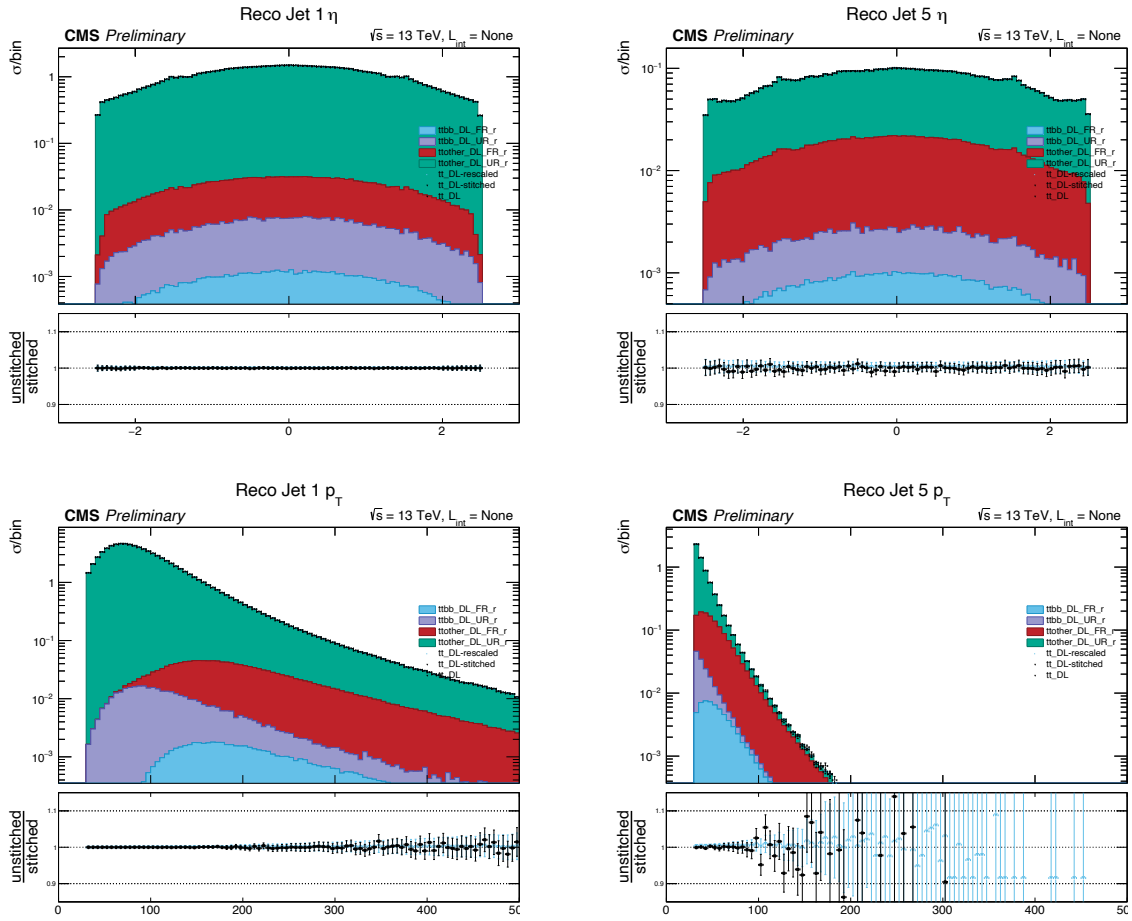

Figure D.41: Dilepton $t\bar{t}$ UE-Up stitching, jets

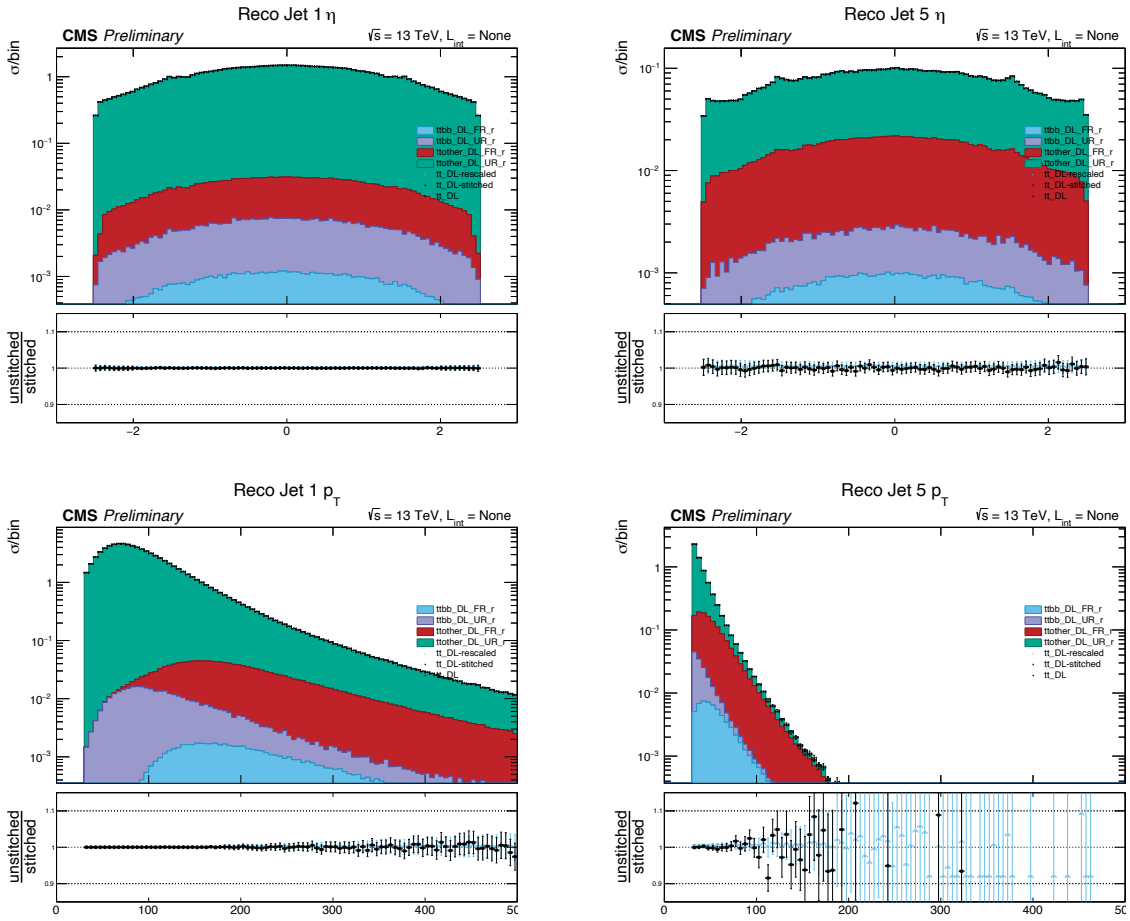

Figure D.42: Dilepton $t\bar{t}$ UE-Down stitching, jets

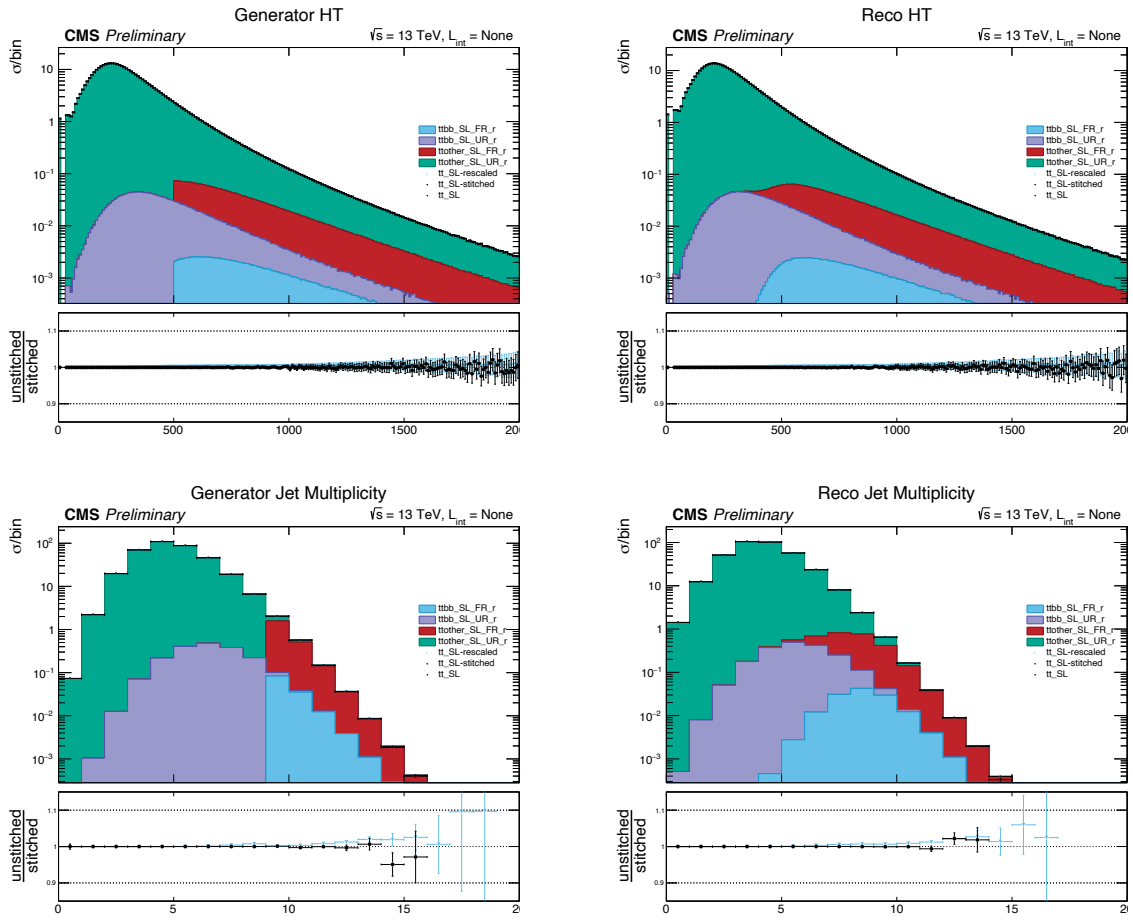

Figure D.43: Semileptonic $t\bar{t}$ stitching, filter variables: Black markers denote the unstitched sample, blue markers include the $t\bar{t}b\bar{b}$ correction, both are divided by the stitched sample. Filled histograms show the rescaled process from the $t\bar{t}b\bar{b}$ and non-$t\bar{t}b\bar{b}$ subprocesses, divided into the Filtered Region (FR) and Unfiltered Region (UR).

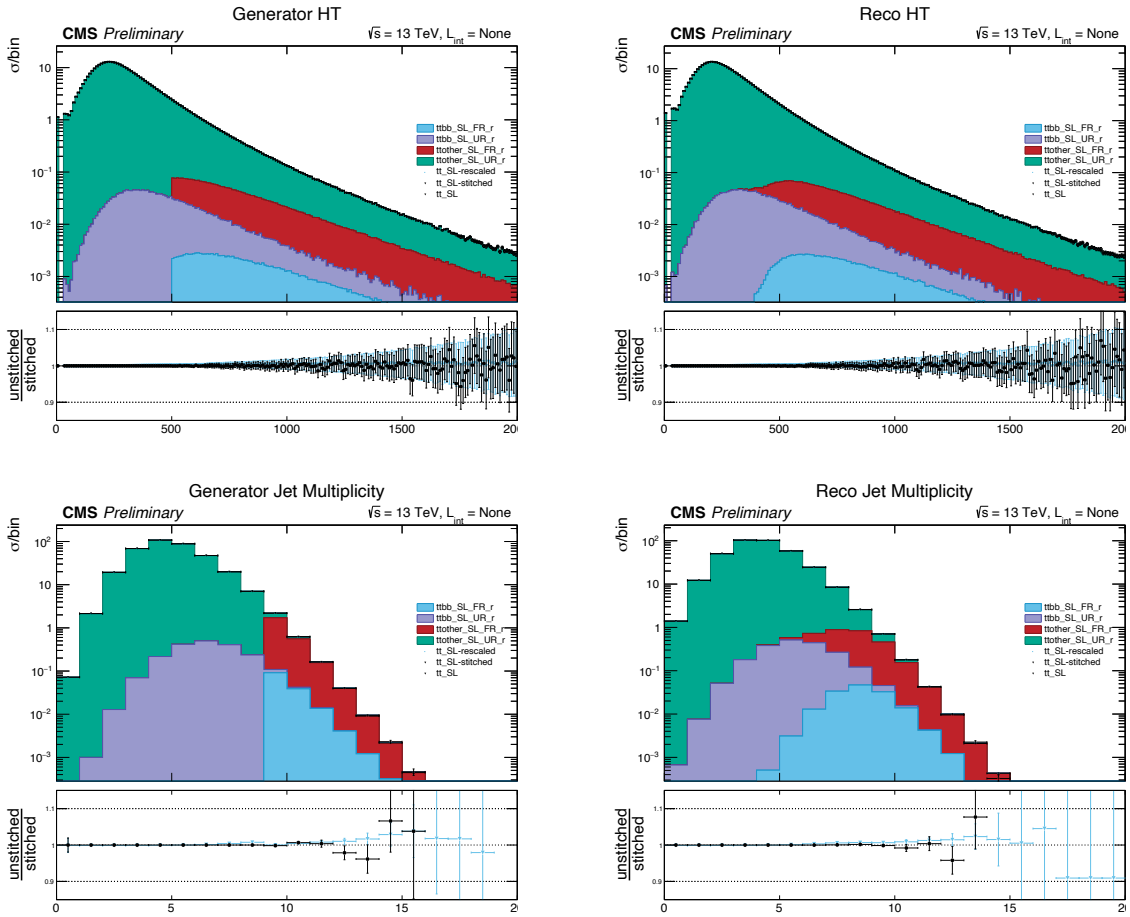

Figure D.44: Semileptonic $t\bar{t}$ h_{damp}-Up stitching, filter variables: Black markers denote the unstitched sample, blue markers include the $t\bar{t}b\bar{b}$ correction, both are divided by the stitched sample. Filled histograms show the rescaled process from the $t\bar{t}b\bar{b}$ and non-$t\bar{t}b\bar{b}$ subprocesses, divided into the Filtered Region (FR) and Unfiltered Region (UR).

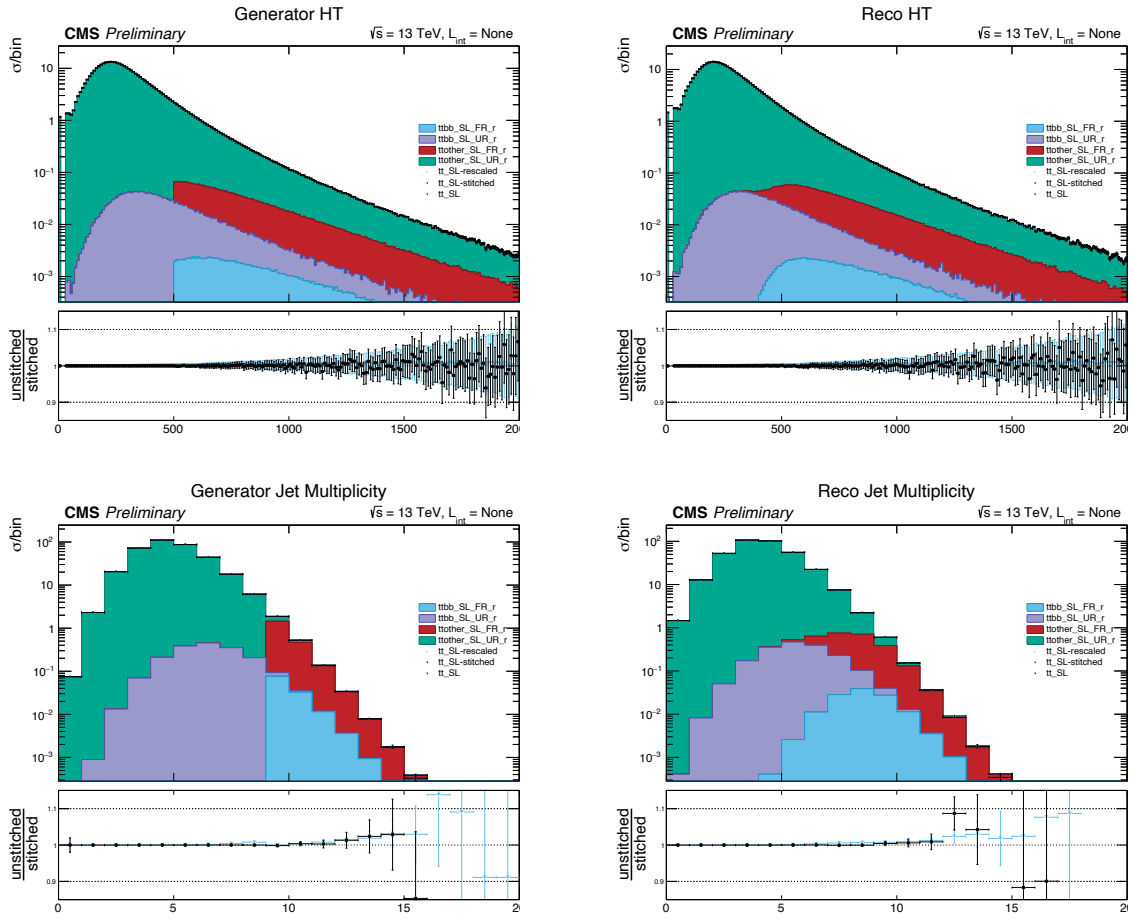

Figure D.45: Semileptonic $t\bar{t}$ h_{damp}-Down stitching, filter variables: Black markers denote the unstitched sample, blue markers include the $t\bar{t}b\bar{b}$ correction, both are divided by the stitched sample. Filled histograms show the rescaled process from the $t\bar{t}b\bar{b}$ and non-$t\bar{t}b\bar{b}$ subprocesses, divided into the Filtered Region (FR) and Unfiltered Region (UR).

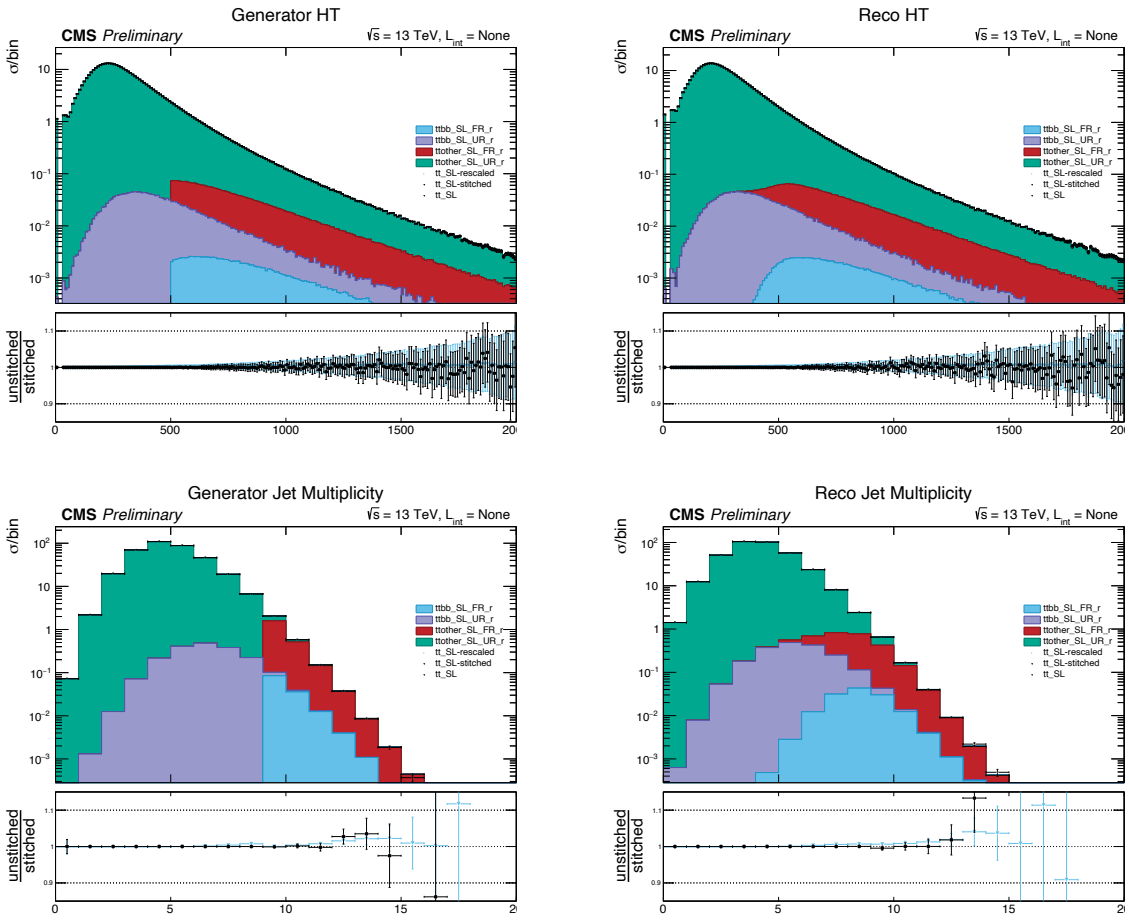

Figure D.46: Semileptonic $t\bar{t}$ UE-Up stitching, filter variables: Black markers denote the unstitched sample, blue markers include the $t\bar{t}b\bar{b}$ correction, both are divided by the stitched sample. Filled histograms show the rescaled process from the $t\bar{t}b\bar{b}$ and non-$t\bar{t}b\bar{b}$ subprocesses, divided into the Filtered Region (FR) and Unfiltered Region (UR).

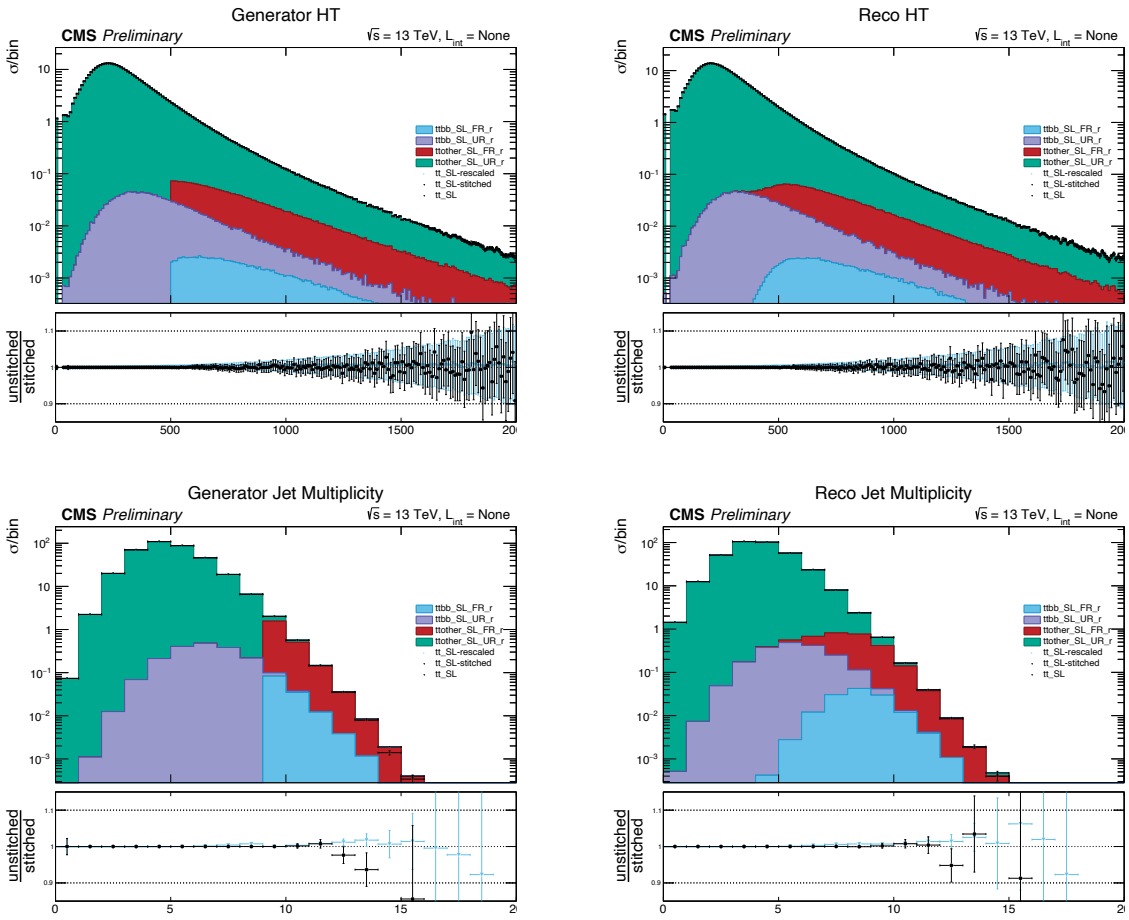

Figure D.47: Semileptonic $t\bar{t}$ UE-Down stitching, filter variables: Black markers denote the unstitched sample, blue markers include the $t\bar{t}b\bar{b}$ correction, both are divided by the stitched sample. Filled histograms show the rescaled process from the $t\bar{t}b\bar{b}$ and non-$t\bar{t}b\bar{b}$ subprocesses, divided into the Filtered Region (FR) and Unfiltered Region (UR).

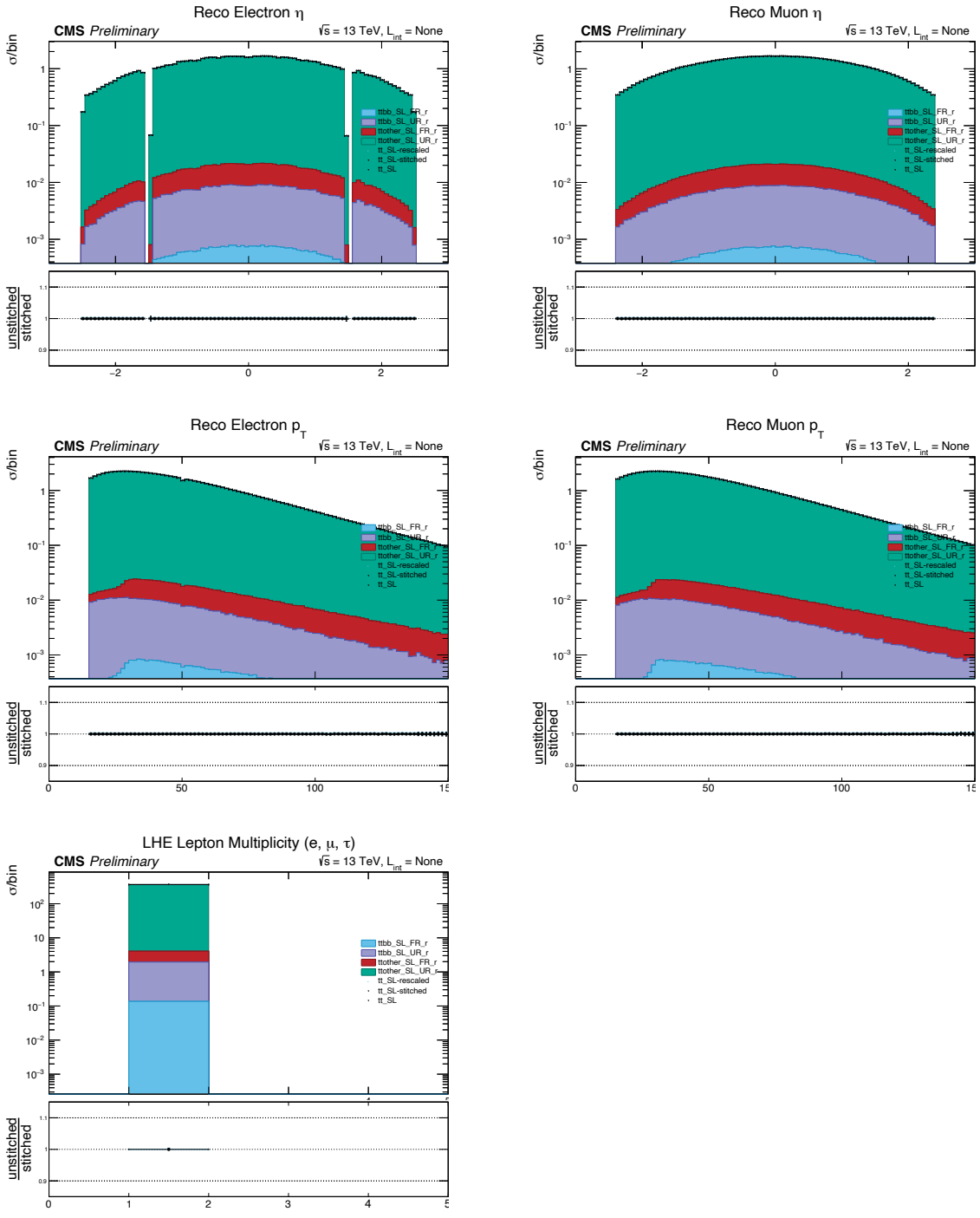

Figure D.48: Semileptonic $t\bar{t}$ stitching, leptons

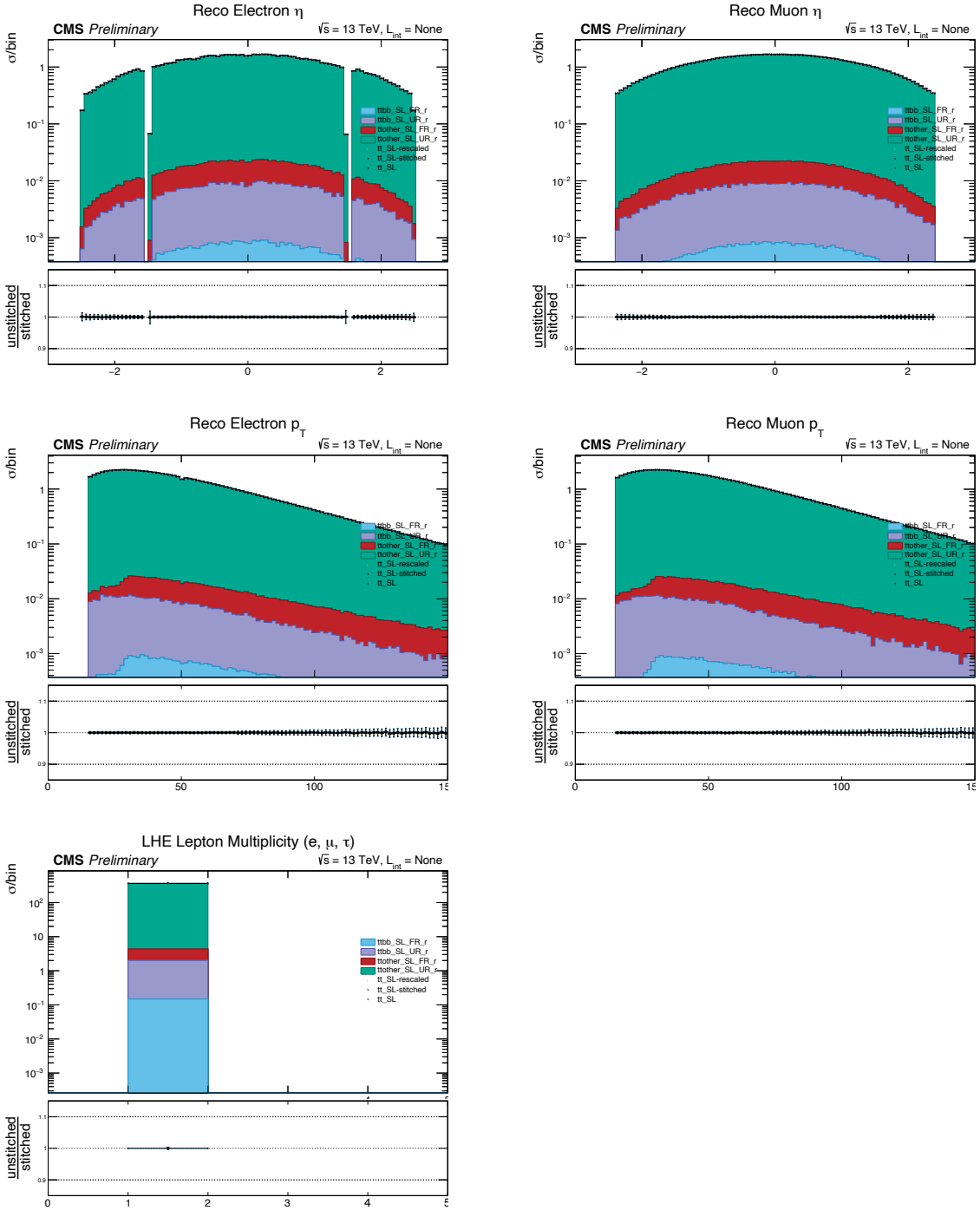

Figure D.49: Semileptonic $t\bar{t}$ h_{damp}-Up stitching, leptons

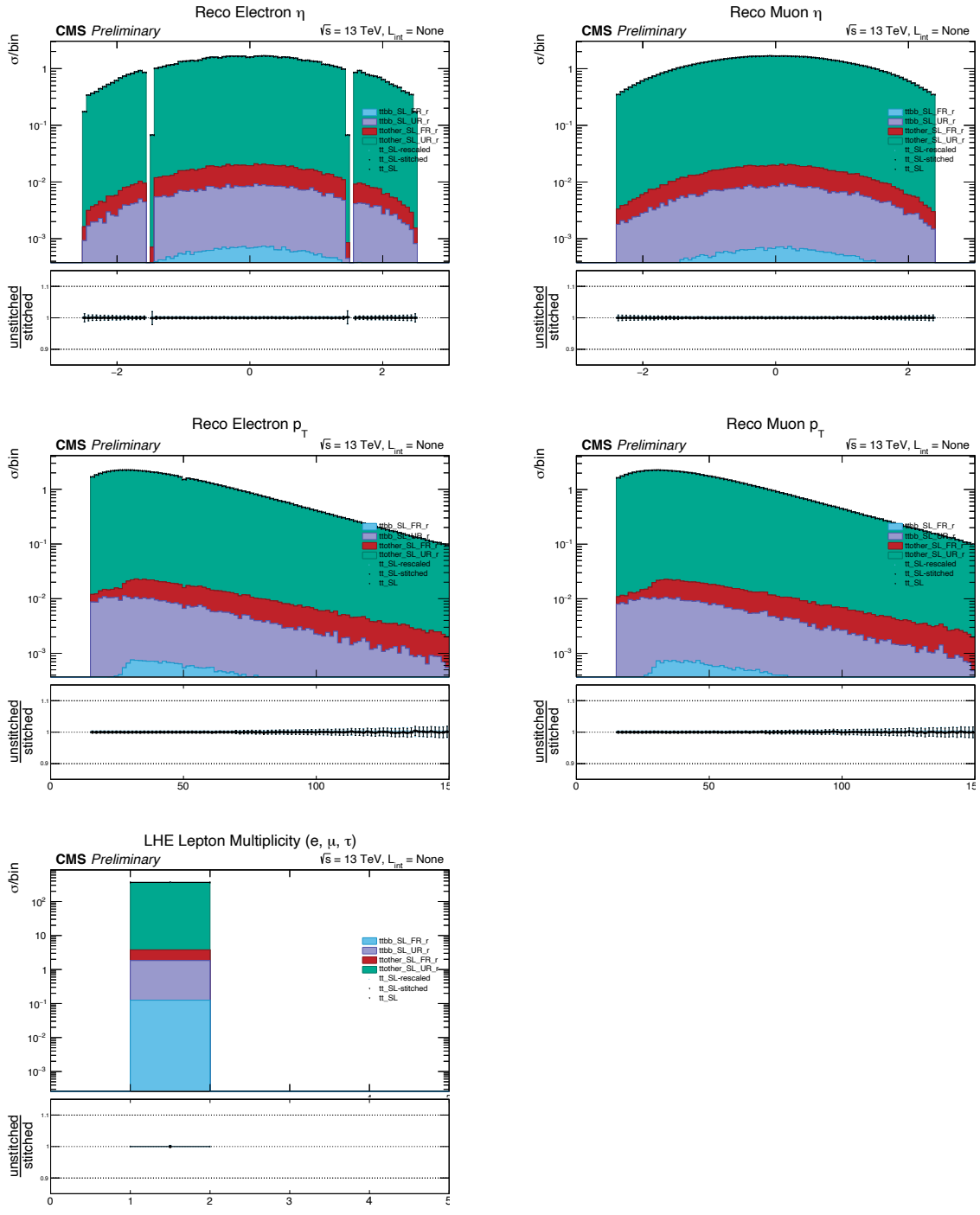

Figure D.50: Semileptonic $t\bar{t}$ h_{damp}-Down stitching, leptons

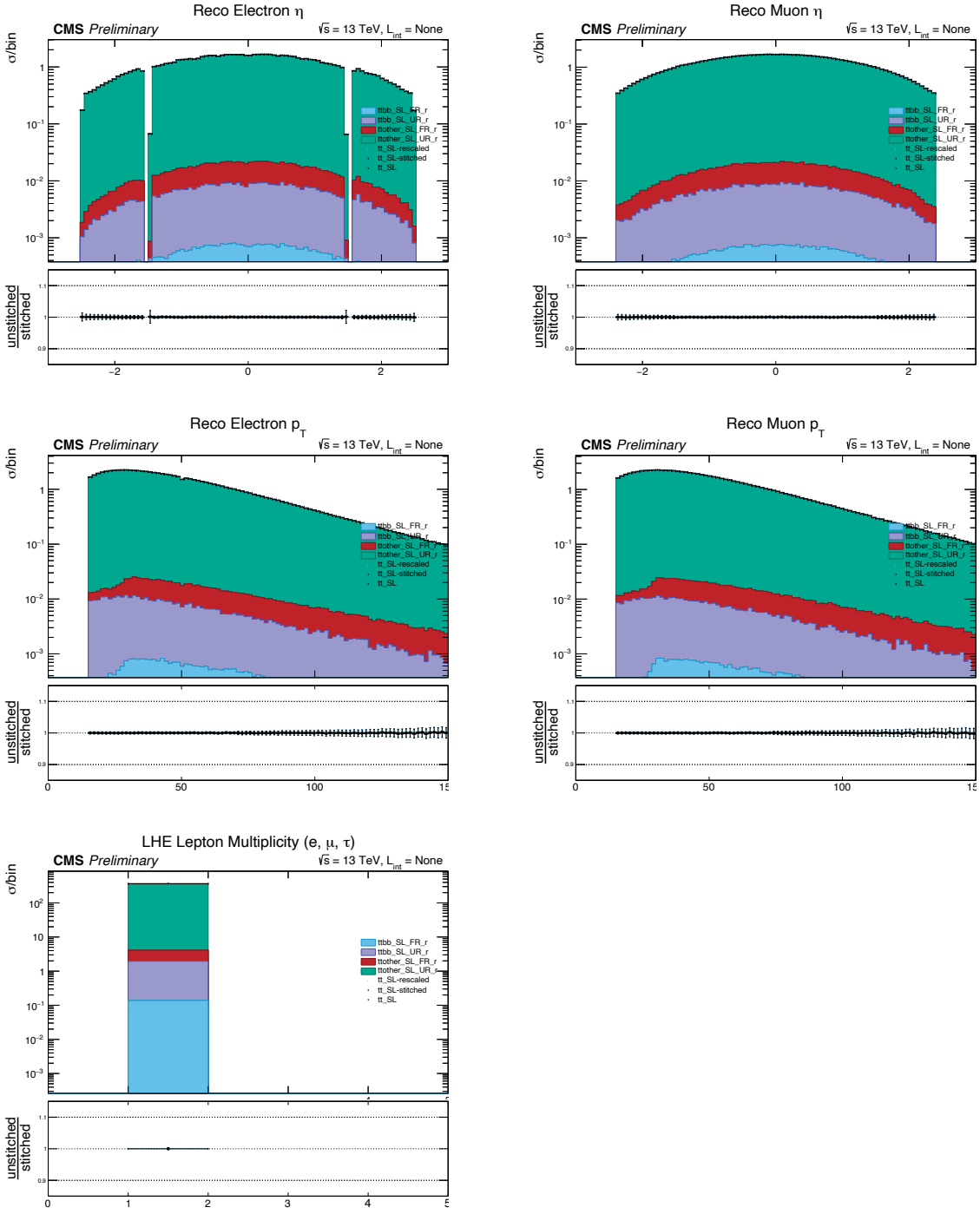

Figure D.51: Semileptonic $t\bar{t}$ UE-Up stitching, leptons

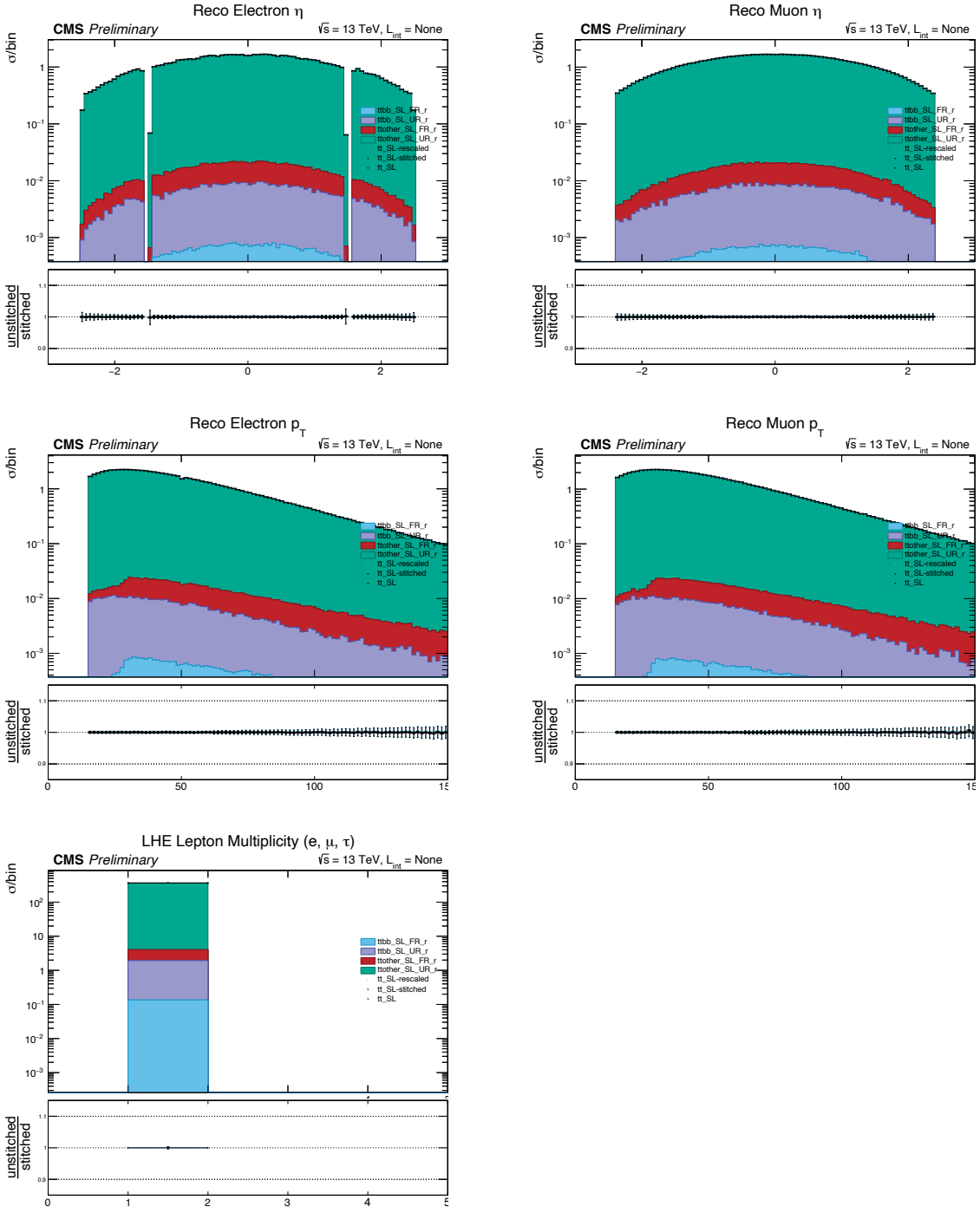

Figure D.52: Semileptonic $t\bar{t}$ UE-Down stitching, leptons

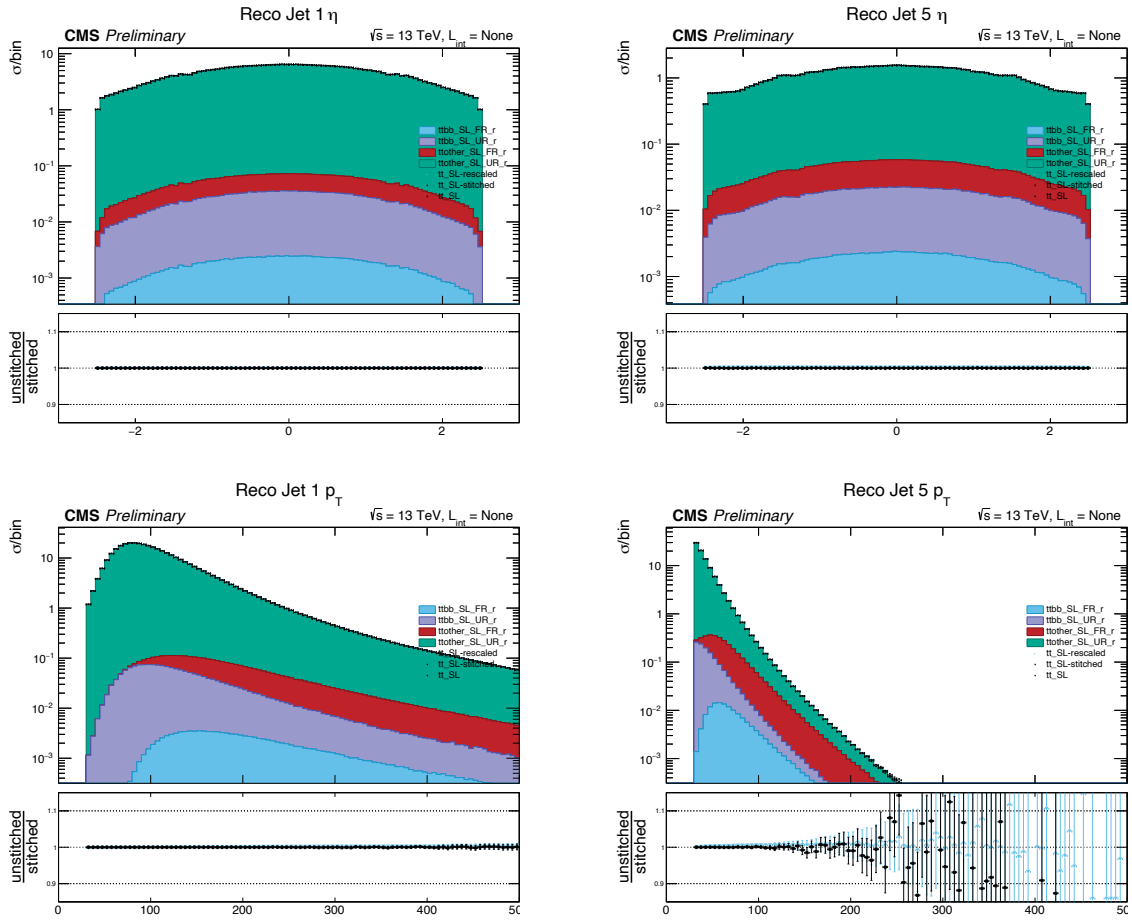

Figure D.53: Semileptonic $t\bar{t}$ stitching, jets

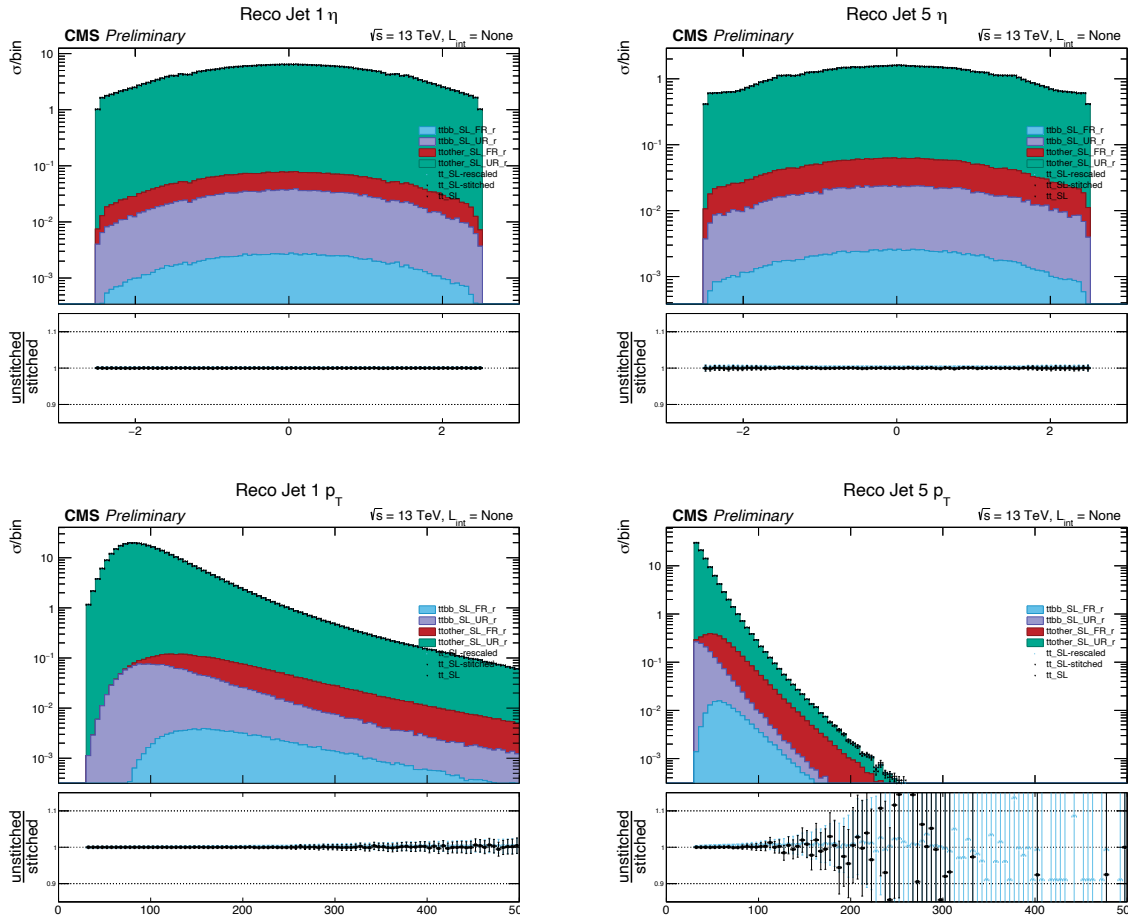

Figure D.54: Semileptonic $t\bar{t}$ h_{damp}-Up stitching, jets

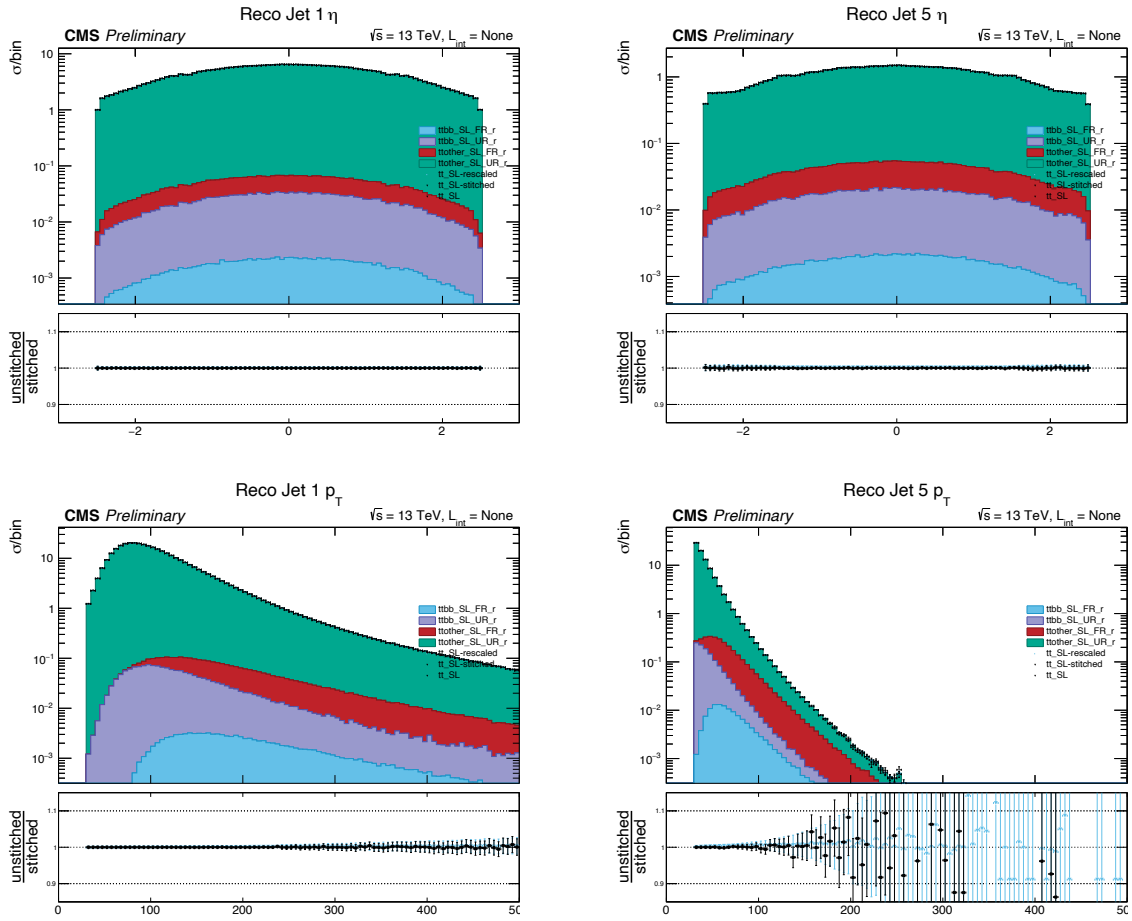

Figure D.55: Semileptonic $\mathrm{t\bar{t}}$ h_{damp}-Down stitching, jets

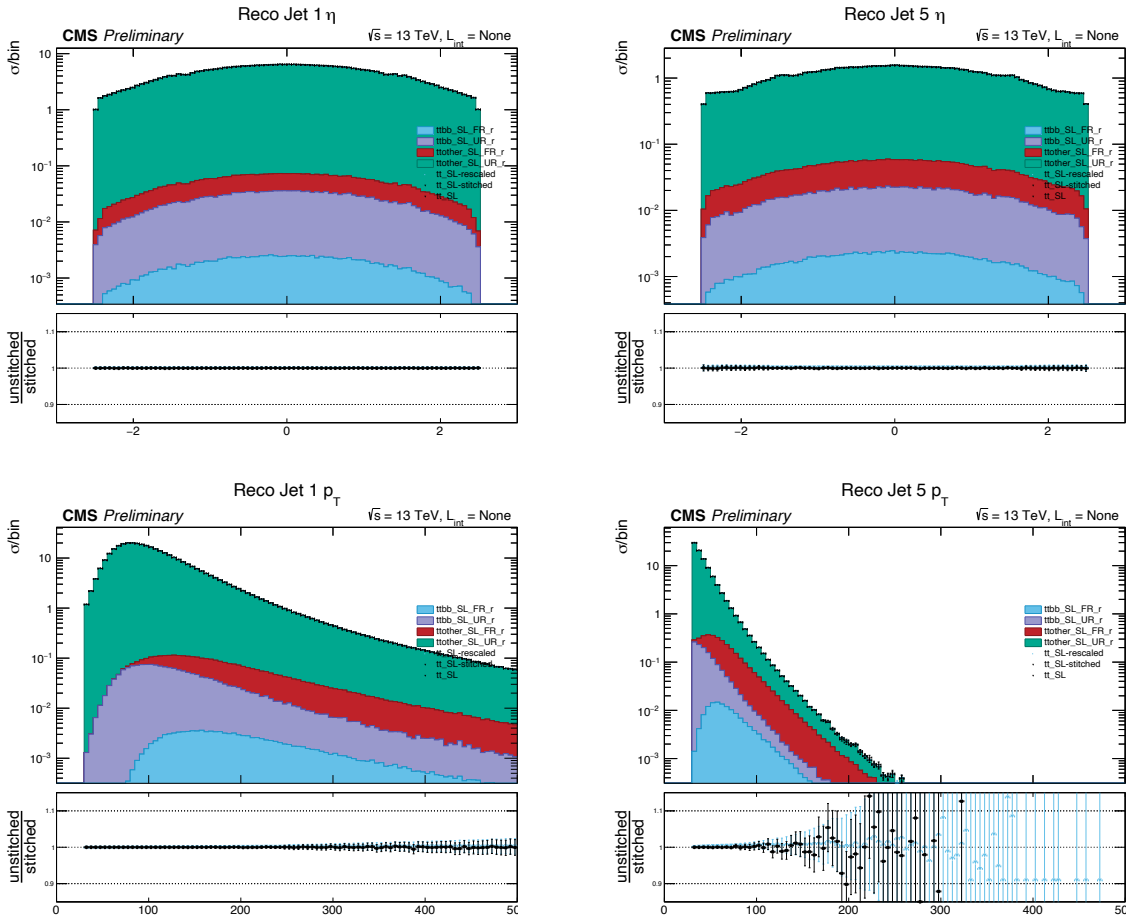

Figure D.56: Semileptonic $t\bar{t}$ UE-Up stitching, jets

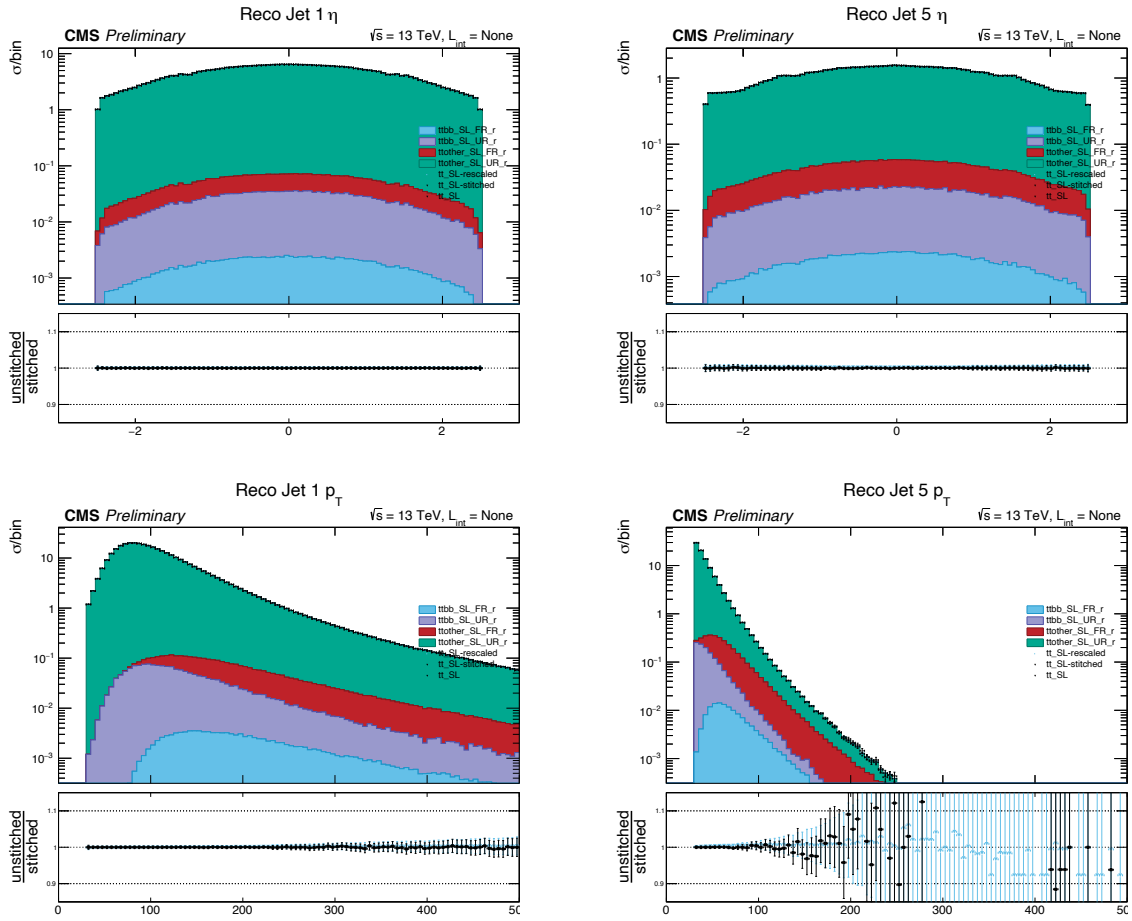

Figure D.57: Semileptonic $t\bar{t}$ UE-Down stitching, jets

D.5 The h_{damp} Studies

The h_{damp} variations have been studied using the systematic samples for the inclusive $\mathrm{t\bar{t}}$ as well as generator filtered variations akin to the ones used for nominal event simulation. These stitched systematic samples have, in total, about one fifth of the simulated events as the nominal in the analysis phase space. In this section, all studies have been performed on the Asimov dataset.

From the inclusive distributions in Section D.4, h_{damp} changes the p_{T} spectra and distribution of jets, shifting them simultaneously up or down (and therefore H_{T} as well). Within the analysis phase space, this manifests as principally a change in normalization (when considering the inclusive distributions, it is a pure shape variation). Due to insufficient numbers of simulation events, however, any residual shape is indistinguishable.

Multiple approaches to including this systematic variation were tried. The most rigorous was to include full H_{T} distributions as templates for the simultaneous fit. However, the result of this was a nuisance term constrained to just a few percent of the input. Upon study of the generated templates compared to the nominal distributions, it becomes apparent that there is not enough events to create sufficiently smooth distributions for this method.

In seeking another solution, h_{damp} was studied by summing the events of all lepton decay channels and either all 3 b tag multiplicity categories or all 5 jet multiplicity categories. Although this minimally doubles the number of unweighted events, the templates are still not trustworthy, even after application of smoothing algorithms. Shown in Figs. D.58-D.59 are the results, where the ratio plots have the full systematic uncertainty band in red and

the h_{damp} sub-components in green (Up) and blue (Down). Finally, aggregating both b tag and jet multiplicities as in Fig. D.60 in all decay channels produces demonstrates that we still cannot make a template accurately depicting the shape of these systematics. Although much of the fluctuations have been removed, the presence of multiple (near) crossover points, which are not well correlated between the two separate years, contraindicate the inclusion in the fit as a shape template. Moreover, this final aggregation suggests the h_{damp} variation have little real shape within the analysis phase space, at least for the H_{T} variable, principally being a normalization change. The effect seen in the simulation is a relative change of $+10\%/-7\%$ for the Up and Down variations, respectively.

The overall normalization of these were inserted as log-normal nuisances on a per-channel, per-b tag, per-jet multiplicity (that is, calculated individually in each of the analysis sub-categories). The result was again a large-impact, highly-constrained, one-sided nuisance, as seen in Fig. D.61. It was found that even when effectively reduced to a one-bin template, there were still large enough fluctuations between categories to introduce this artificial constraint. Reducing to a single asymmetric log-normal set to the measured $+10\%/-7\%$ produced an unconstrained nuisance. See Fig. D.62

Due to the fact that this implementation is indistinguishable, fit-wise, from the $\text{t}\bar{\text{t}}$ cross section uncertainty, they are merged together. No measure-able changes to the expected limit or significance were found. Including the high-H_{T} uncertainty, the total up and down variations are listed in the systematics section as the $\text{t}\bar{\text{t}}$ cross section uncertainty.

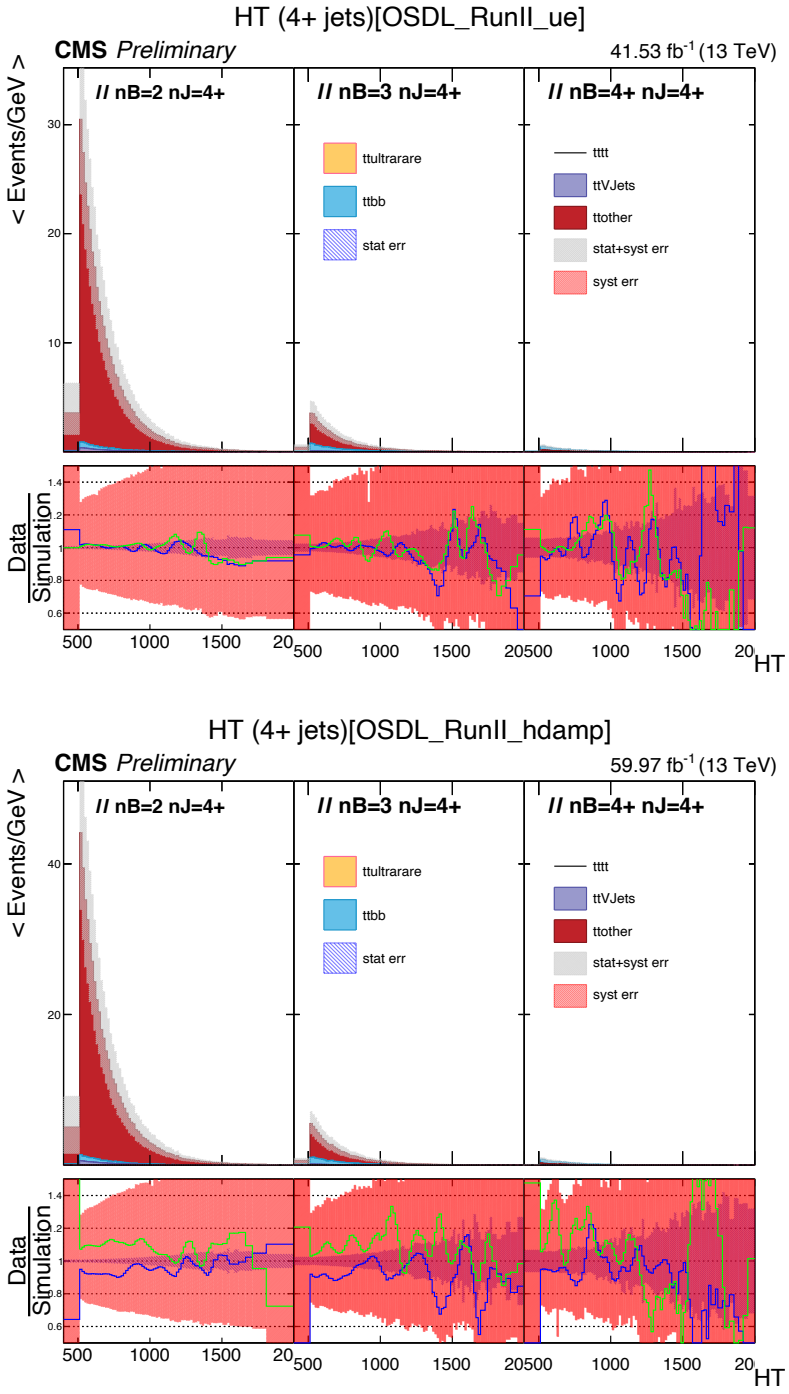

Figure D.58: In green and blue of the ratio plot are h_{damp} Up and Down, respectively. In this, lepton decay channels and Jet categories are merged, potentially showing the behavior of this variation by b tagged jet multiplicity. The templates are unreliably noisy relative to the nominal due to too small a Monte Carlo sample. 2017 is on top, 2018 on bottom.

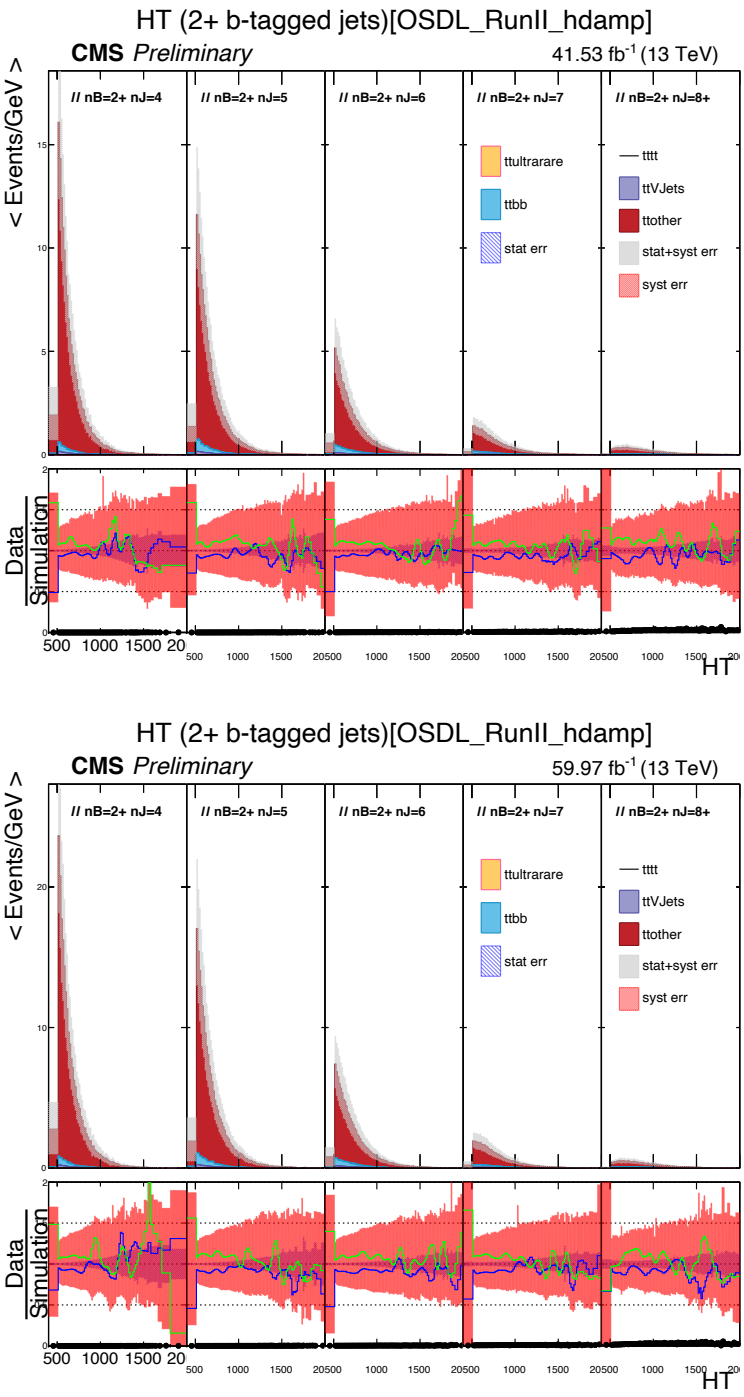

Figure D.59: In green and blue of the ratio plot are h_{damp} Up and Down, respectively. In this, lepton decay channels and b tagged jet categories are merged, potentially showing the behavior of this variation by jet multiplicity. The templates are unreliably noisy relative to the nominal due to too small a Monte Carlo sample. 2017 is on top, 2018 on bottom.

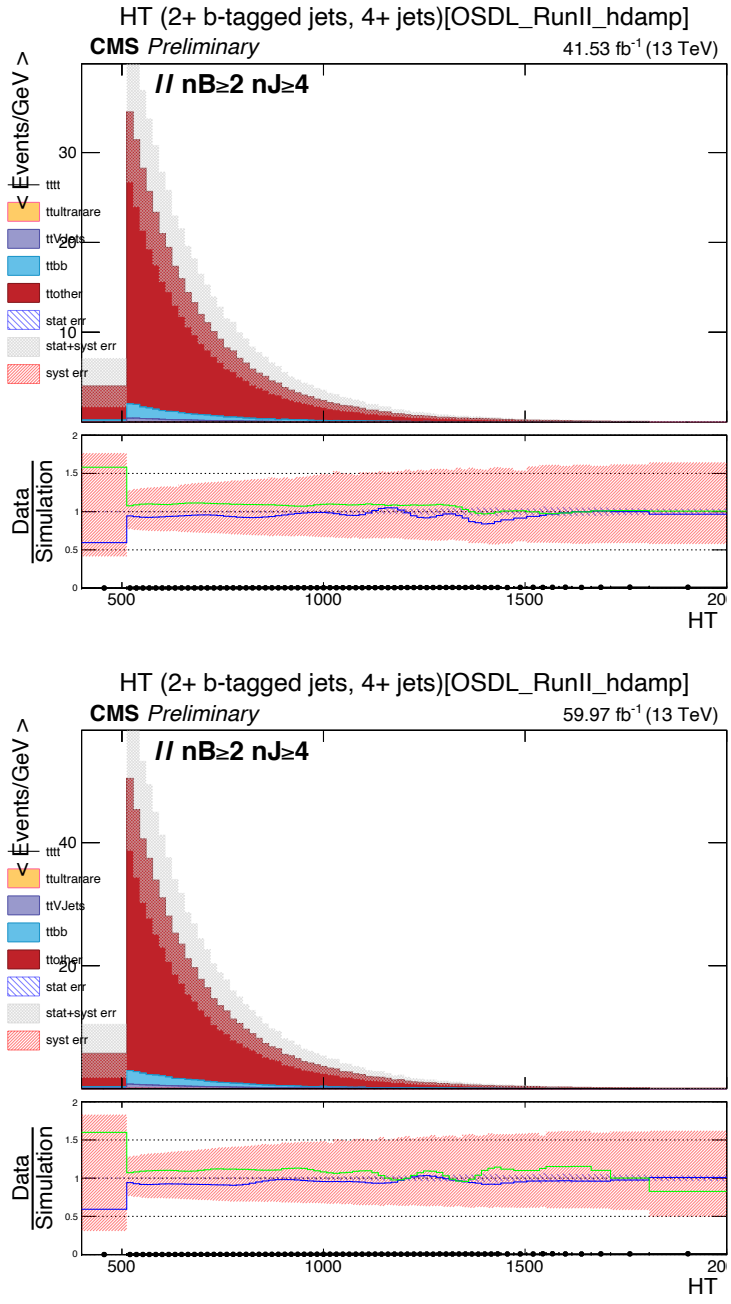

Figure D.60: In green and blue of the ratio plot are h_{damp} Up and Down, respectively. In this, lepton decay channels, jet multiplicity and b tagged jet multiplicity categories are merged. Although the templates are finally rid of the largest stochastic fluctuations relative to the nominal, they do not appear to be reliable descriptions of an underlying physics effect, given the multiple (near) crossover points for the Up and Down variations. The variations are not consistent between years, supporting this. 2017 is on top, 2018 on bottom.

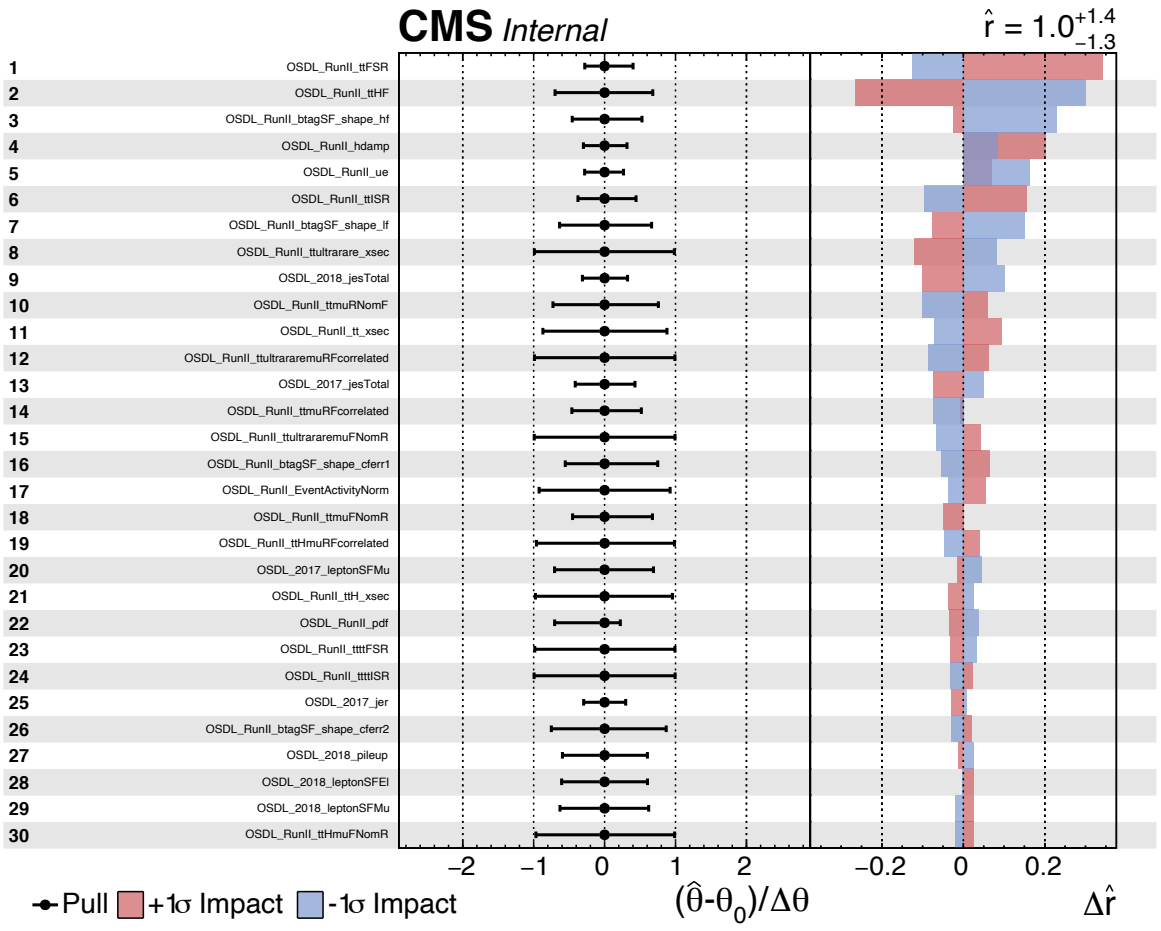

Figure D.61: Impacts with log-normal uncertainties calculated individually for each lepton decay x jet x b tagged jet category.

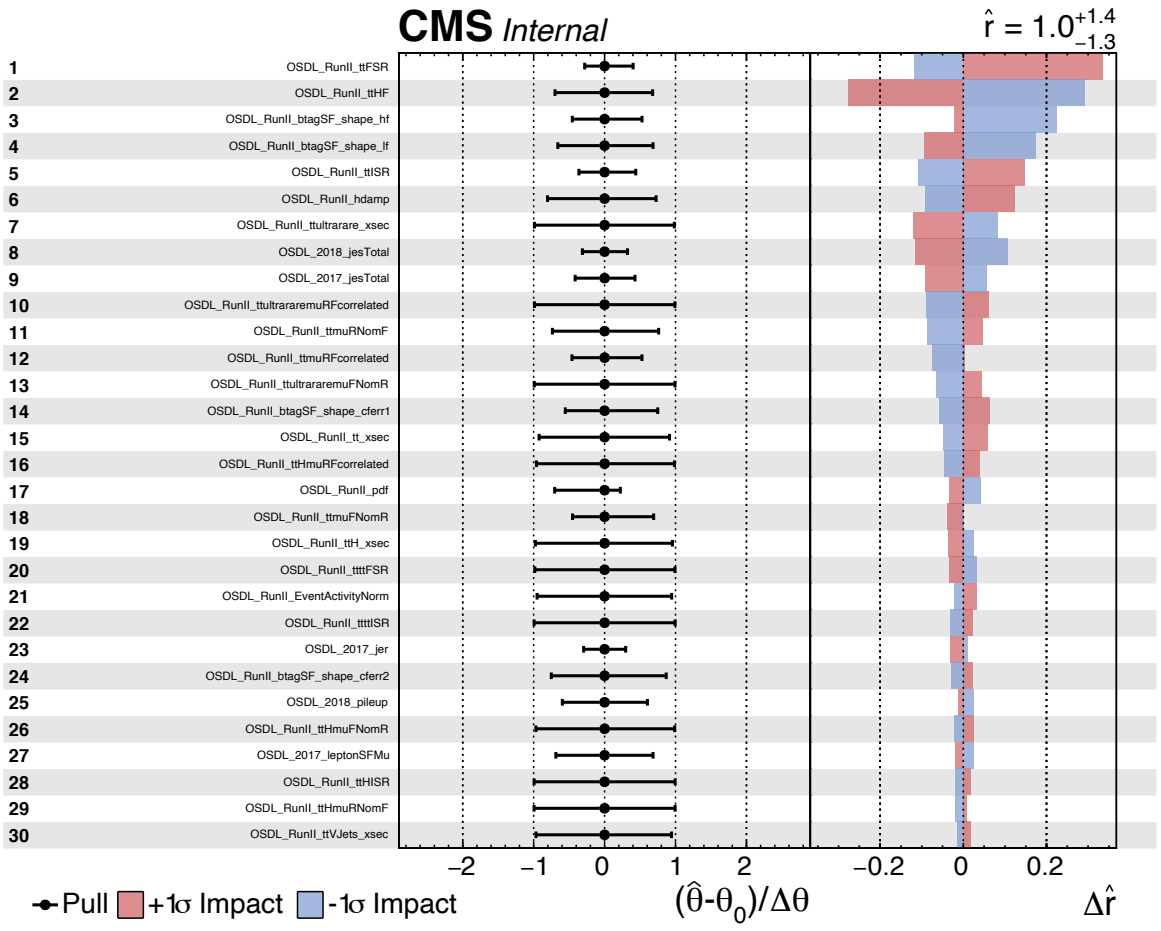

Figure D.62: Impacts with a single log-normal uncertainty, prior to merging hdamp in with the $t\bar{t}$ cross section uncertainty along with the Event Activity uncertainty.

D.6 Comparison of Signal and Background Shape in Signal Regions

The figures in this section show the shape comparison of total background to the $t\bar{t}t\bar{t}$ signal. Signal is normalized to the total background in the most signal-rich categories, shown as an aggregate of those with 3+ medium DeepJet btags and 7+ jets. The first set demonstrate the channel and era-dependent lepton variables (p_{T}, η, and number of tight ID'd leptons). Subsequent figures aggregate the 3 dilepton decay channels and simulation from 2017 and 2018. In general, these results show that the variables most closely associated with the leptonically decaying top quarks behave very similarly between signal and backgrounds. The electron and muon p_{T} spectra are skewed slightly higher in signal, and similarly so for the p_{T}^{miss}.

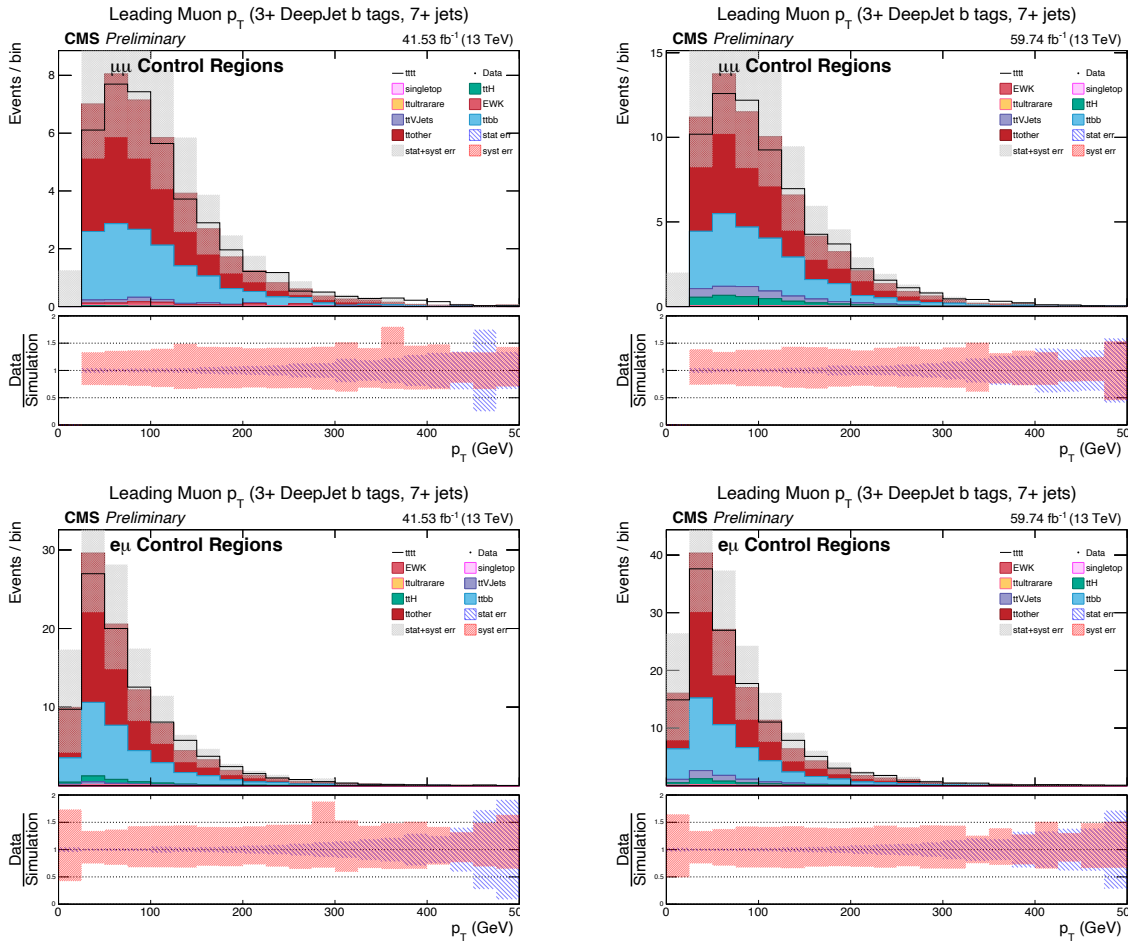

Figure D.63: Highest p_{T} muon for the 2017 and 2018 datasets. The left panels are for 2017 and the right panels are for 2018. The muons of the $\mu\mu$ c hannel are shown in the top row, the muons of the eμ channel are in the bottom row

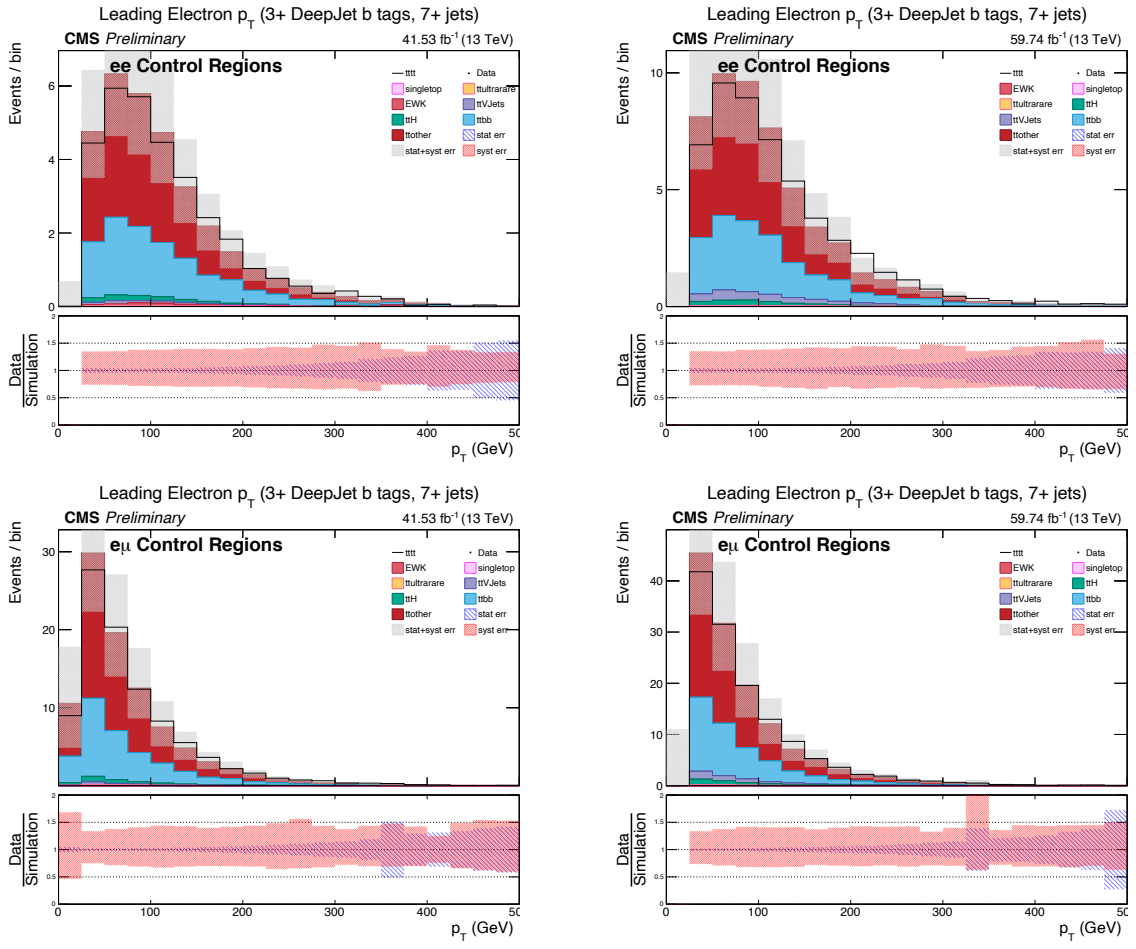

Figure D.64: Highest p_{T} electron for the 2017 and 2018 datasets. The left panels are for 2017 and the right panels are for 2018. The electrons of the ee c hannel are shown in the top row, the electrons of the eμ channel are in the bottom row

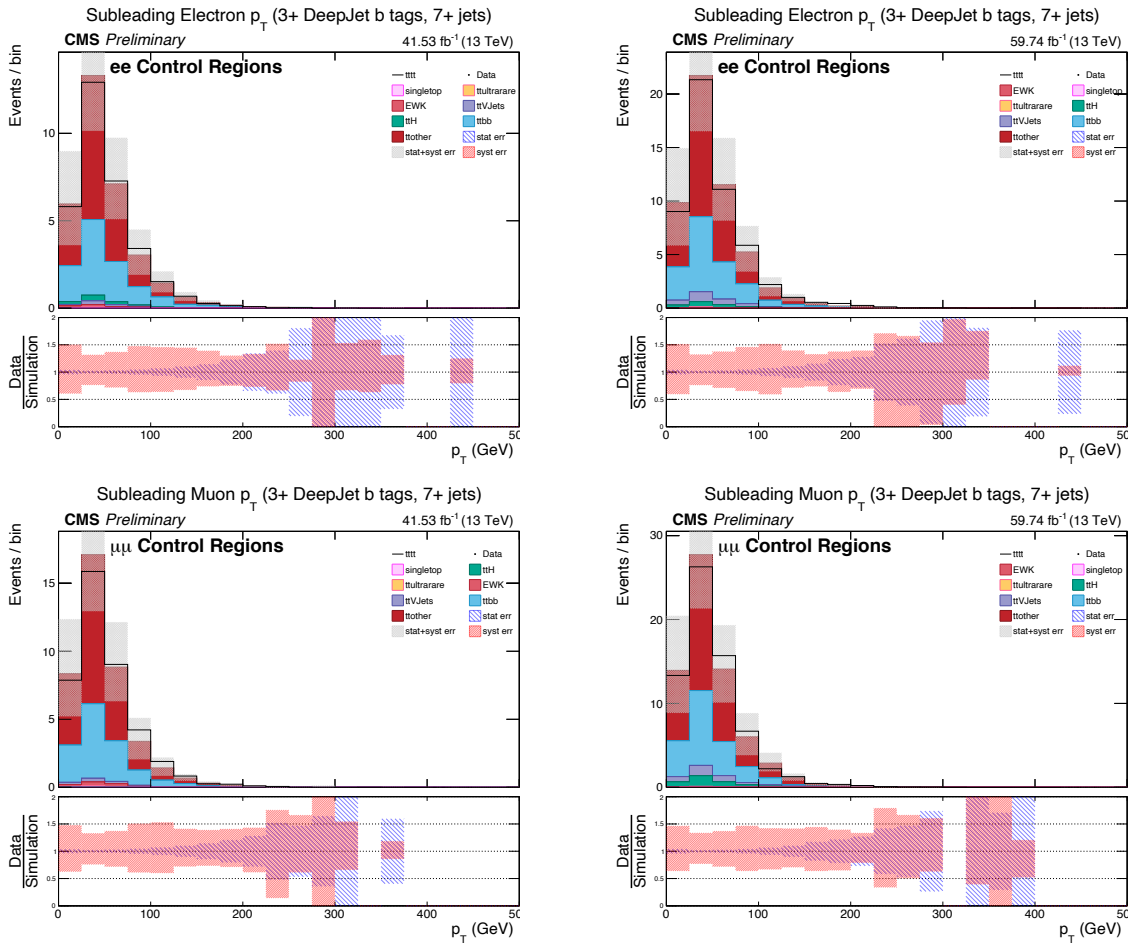

Figure D.65: Subleading lepton p_{T}

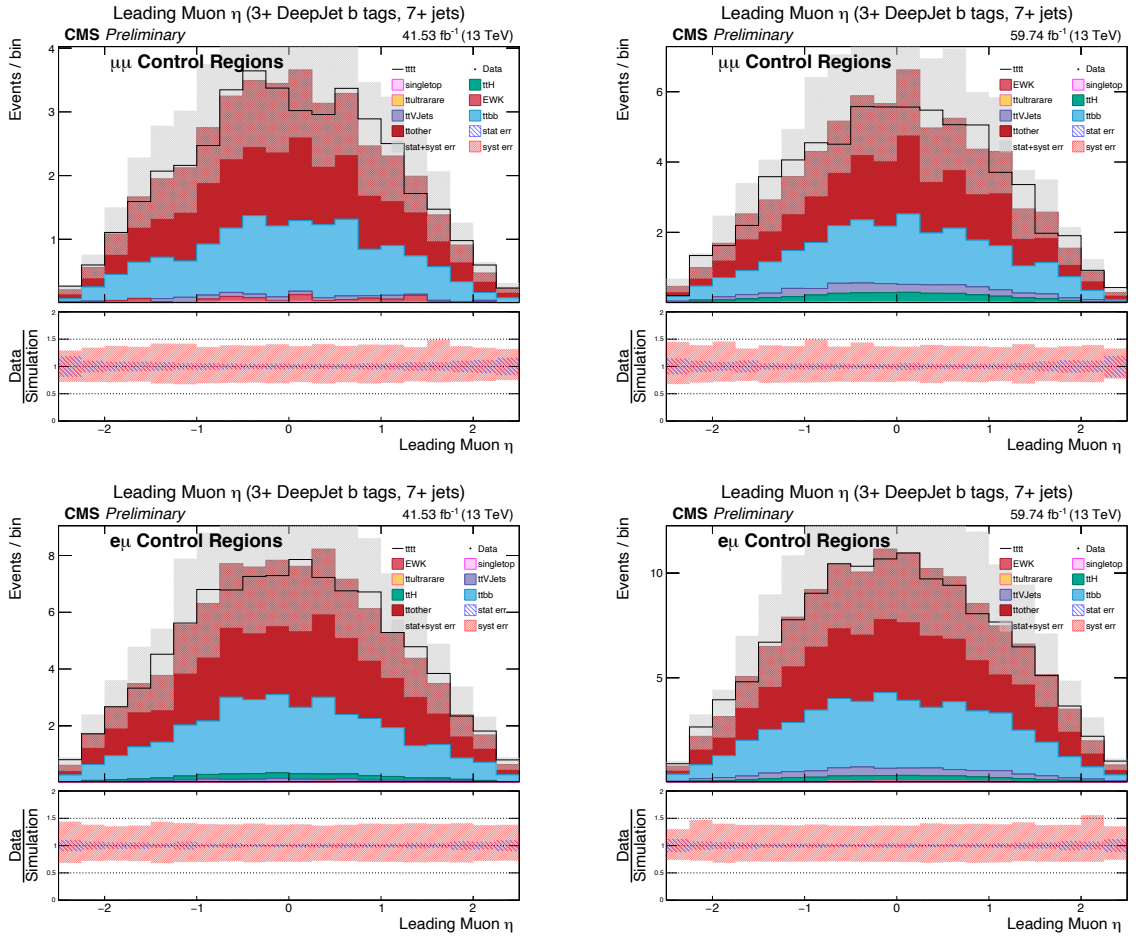

Figure D.66: Highest η muon for the 2017 and 2018 datasets. The left panels are for 2017 and the right panels are for 2018. The muons of the $\mu\mu$ c hannel are shown in the top row, the muons of the eμ channel are in the bottom row

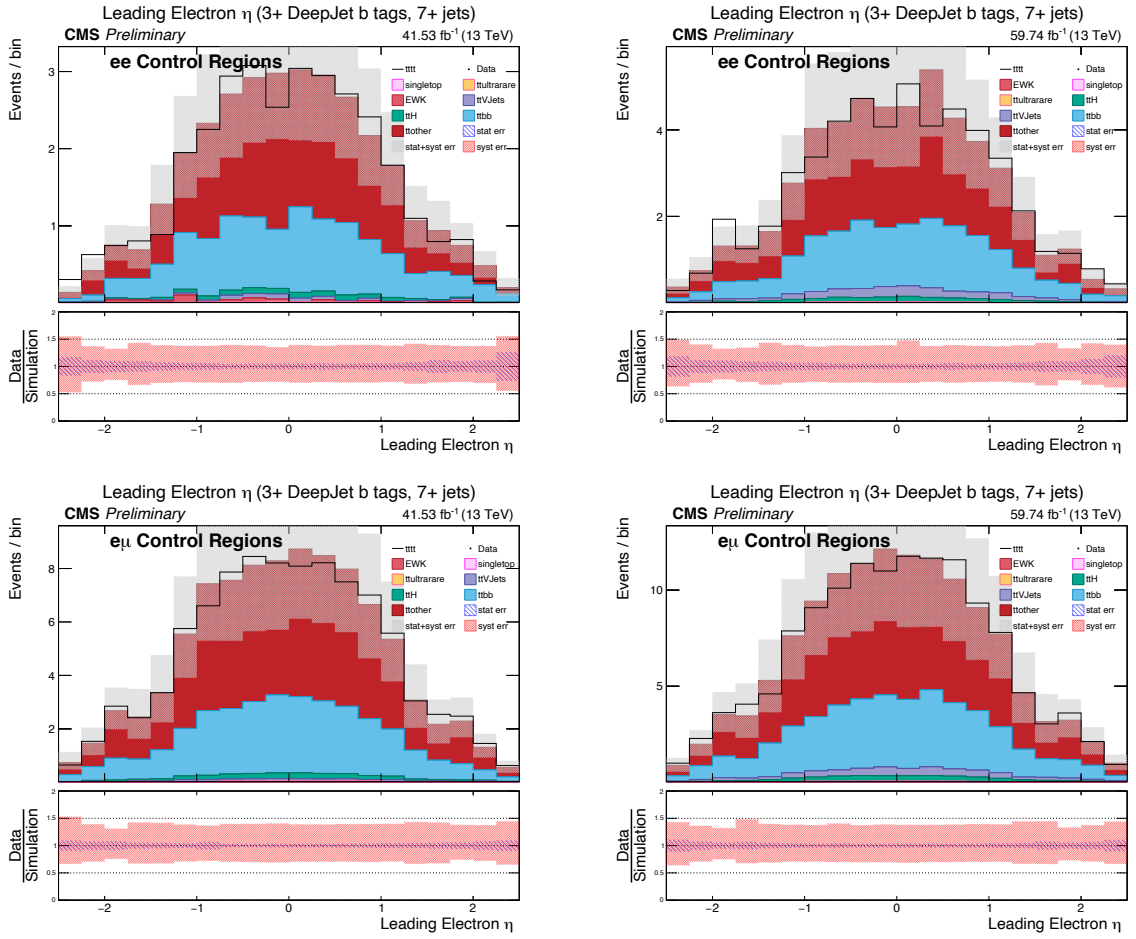

Figure D.67: Highest η electron for the 2017 and 2018 datasets. The left panels are for 2017 and the right panels are for 2018. The electrons of the ee c hannel are shown in the top row, the electrons of the eμ channel are in the bottom row

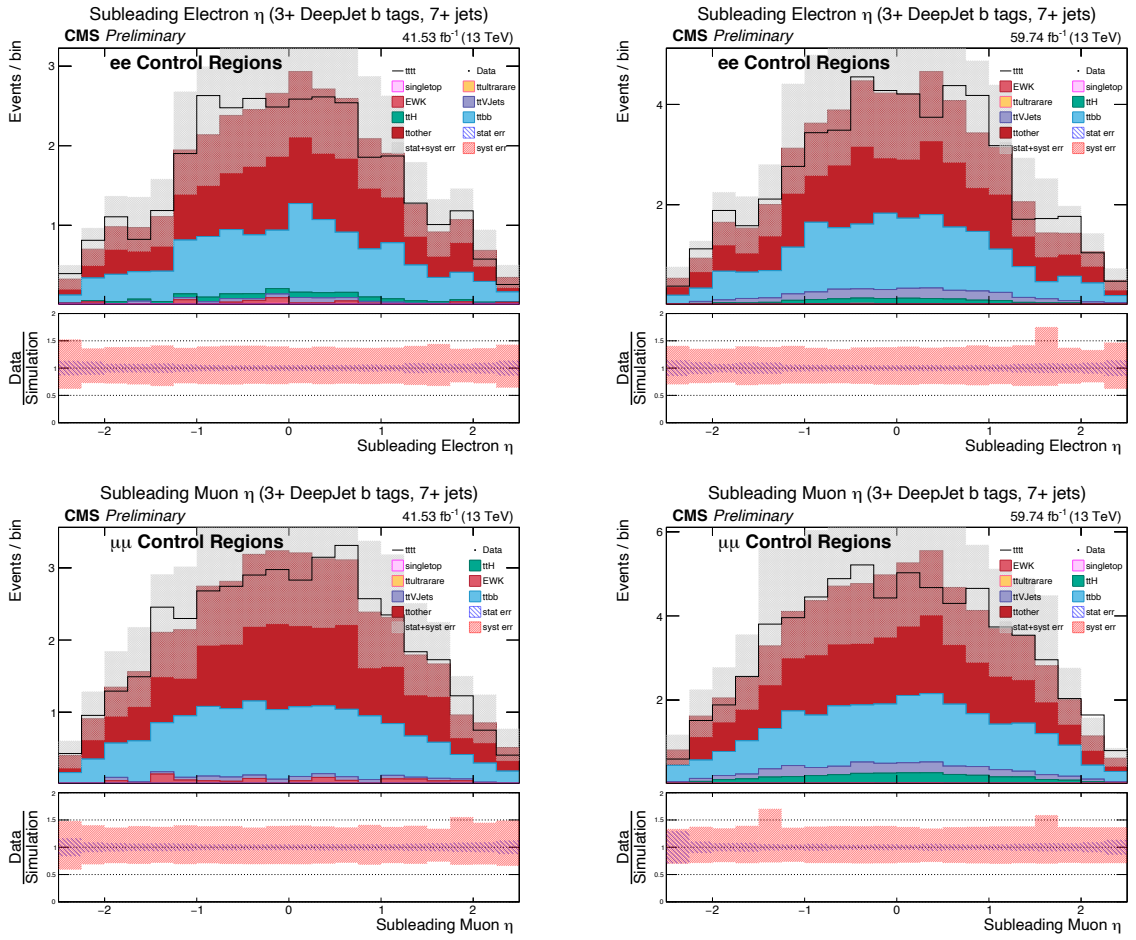

Figure D.68: Subleading lepton η

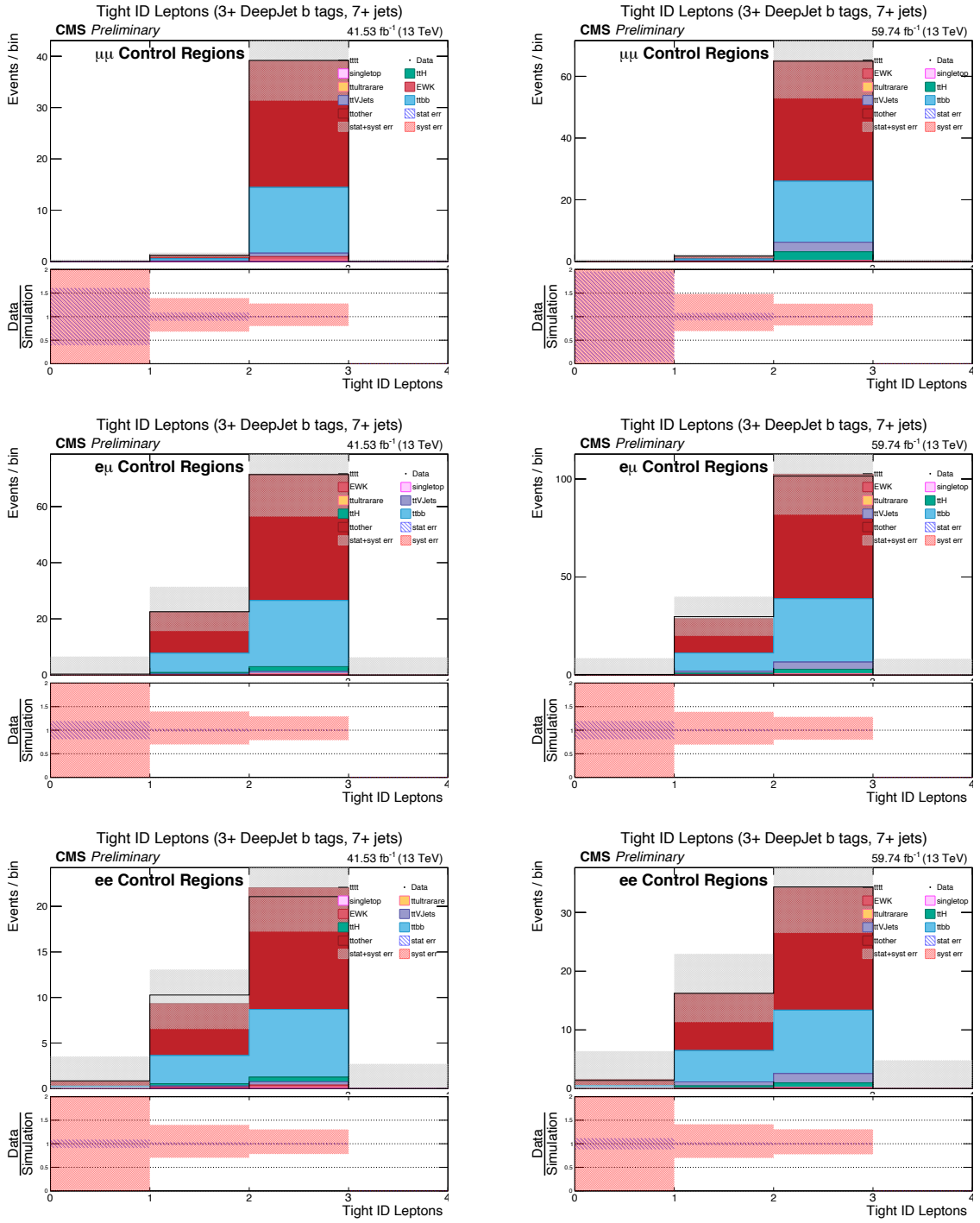

Figure D.69: Number of leptons passing tight lepton ID. The left panels are for 2017 and the right panels are for 2018. The $\mu\mu$ c hannels are shown in the top row, the eμ channels are in the middle, and the ee channels are in the bottom row.

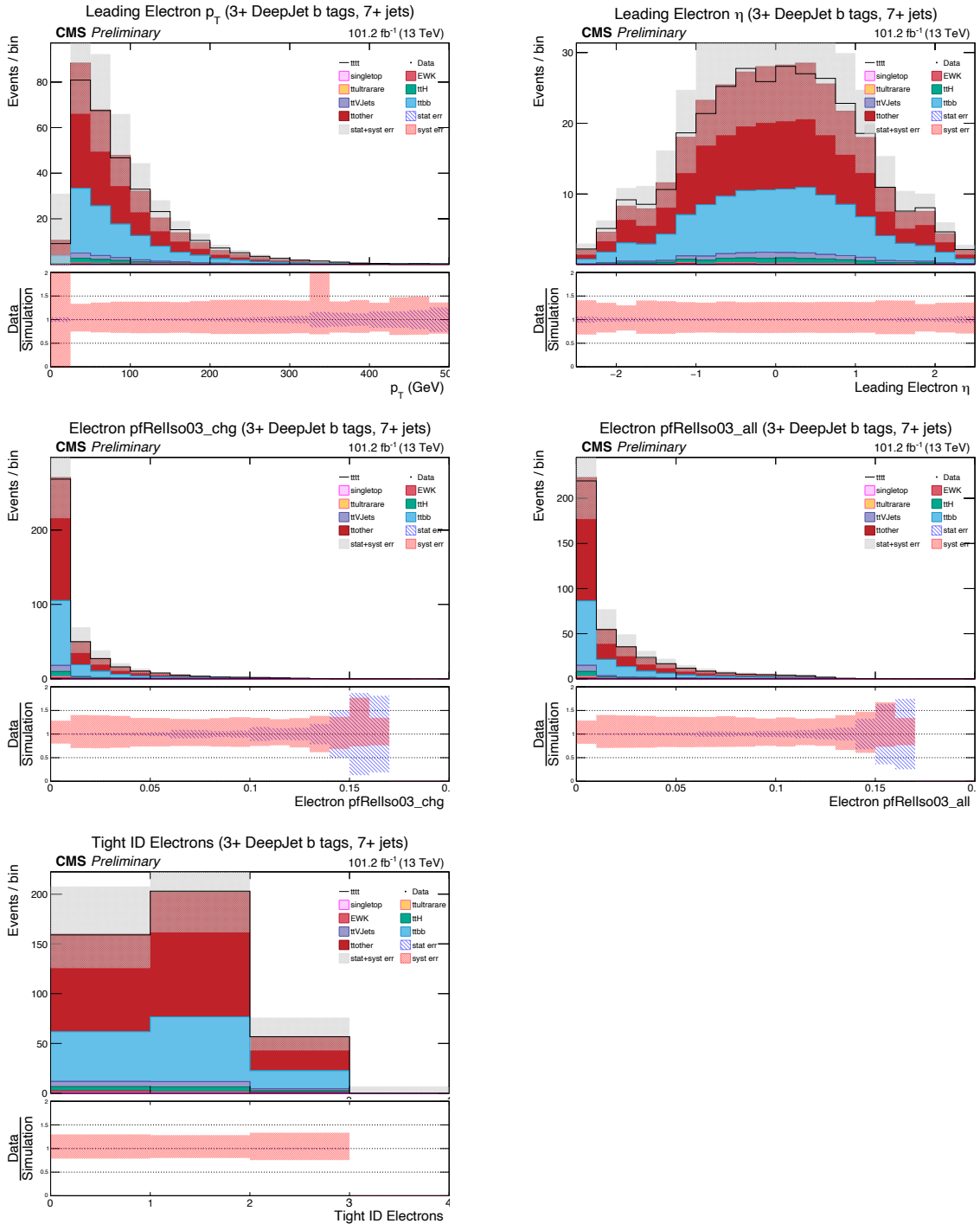

Figure D.70: Aggregate electron distributions.

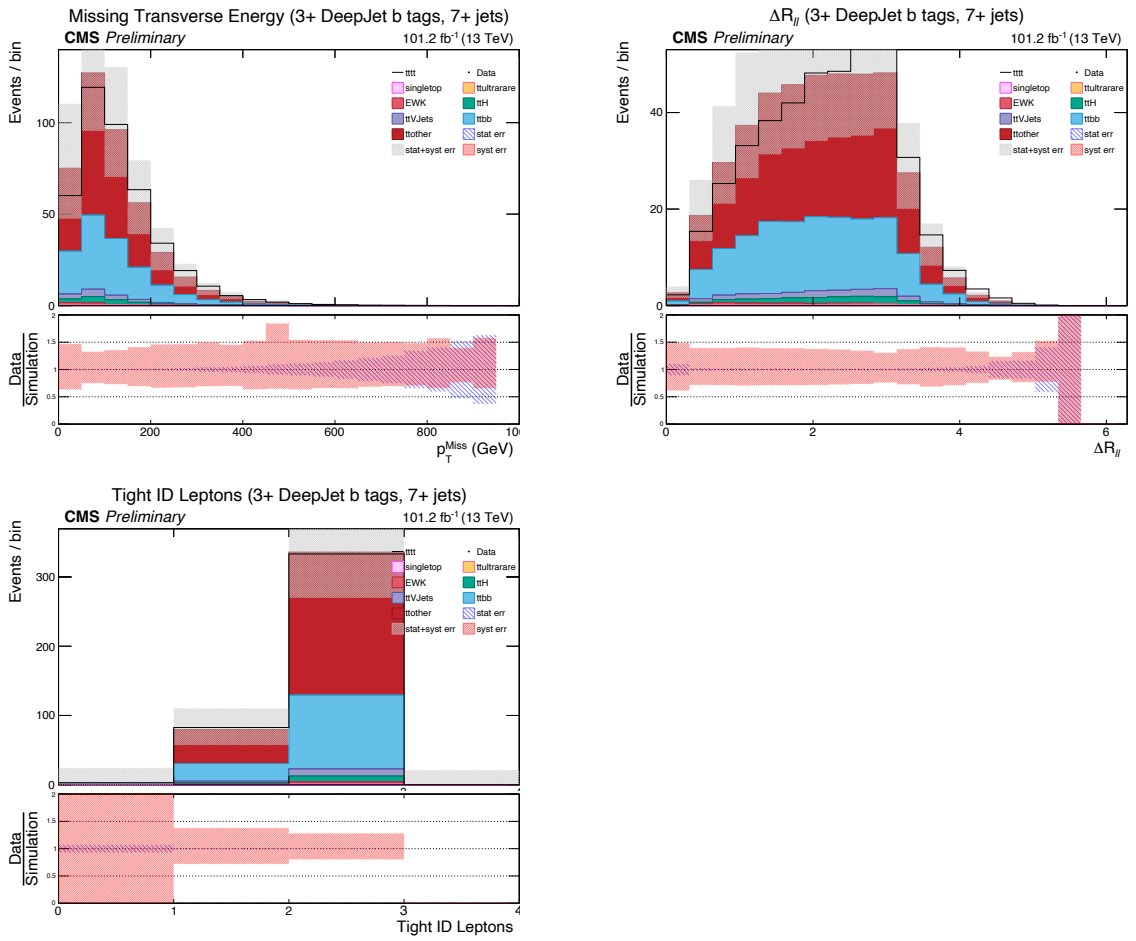

Figure D.71: p_{T}^{miss}, δ R between the two leptons, number of tight ID'd leptons across all channels and years.

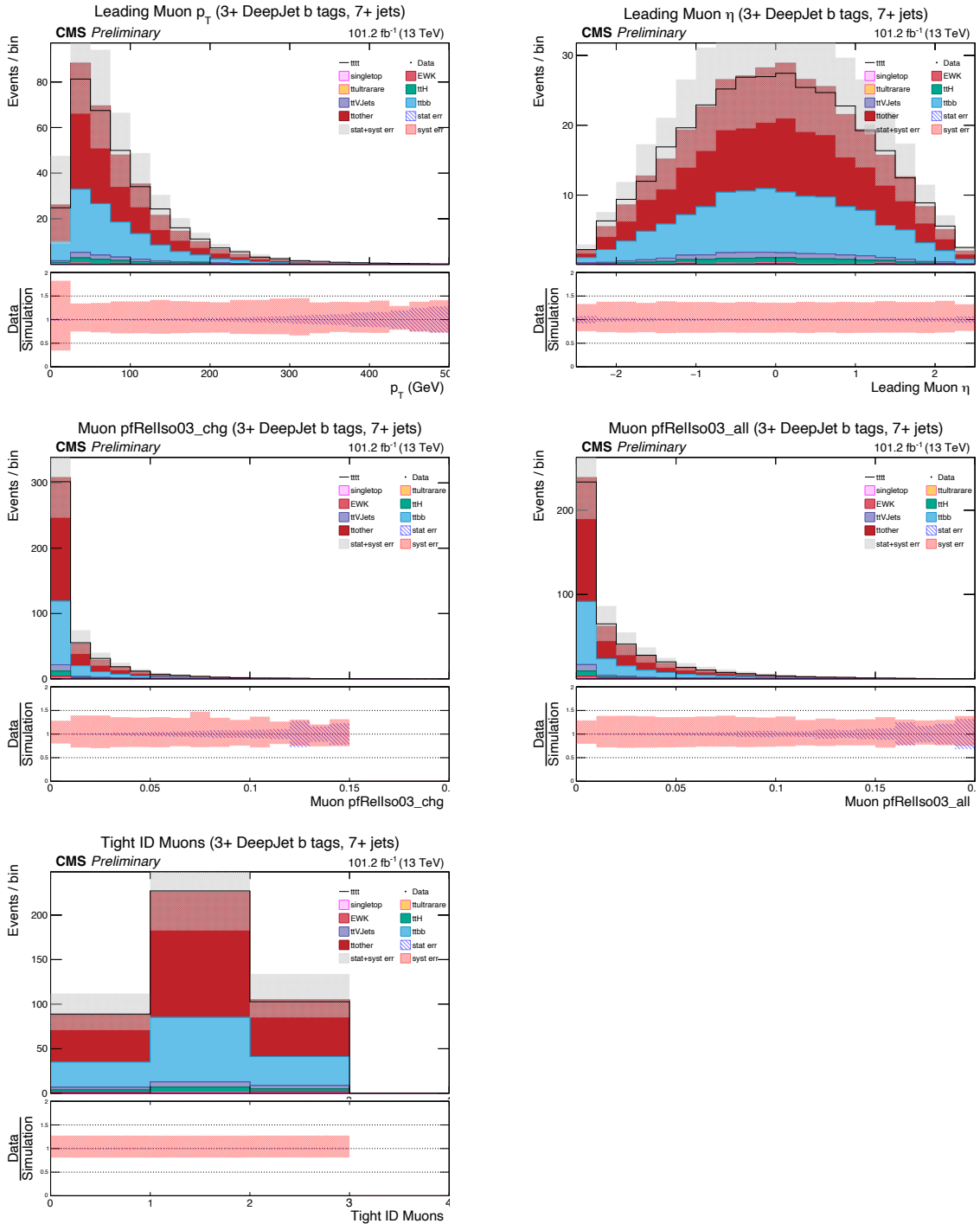

Figure D.72: Aggregate muon distributions.

D.7 Impacts from Control Region Fit

The higgs combine tool is used to perform a fit to data in the control regions, notably including those enriched in the $t\bar{t}b\bar{b}$ subprocess. The initial tests were performed with two sets of categories, one set restricted to the 3 b tag categories (2, 3, 4+) containing exactly 4 jets, and another set additionally using the categories with 2 b tags and 5 or 6 jets. Data and simulation from 2017 and 2018, including all lepton decay channels are employed. The impacts are shown below, first for the Asimov dataset, then the fit to data. Based on the impacts, we find that there is no statistical power to constrain the $t\bar{t}b\bar{b}$ uncertainty, and the correction to simultion is not obviously incompatible with the data. There are no concerning changes in rank nor one-sided uncertainties for those well-populated by templates in these control regions (Lacking signal-enriched regions, the impacts for the four top process will suffer from significant stochastic noise in these fits.)

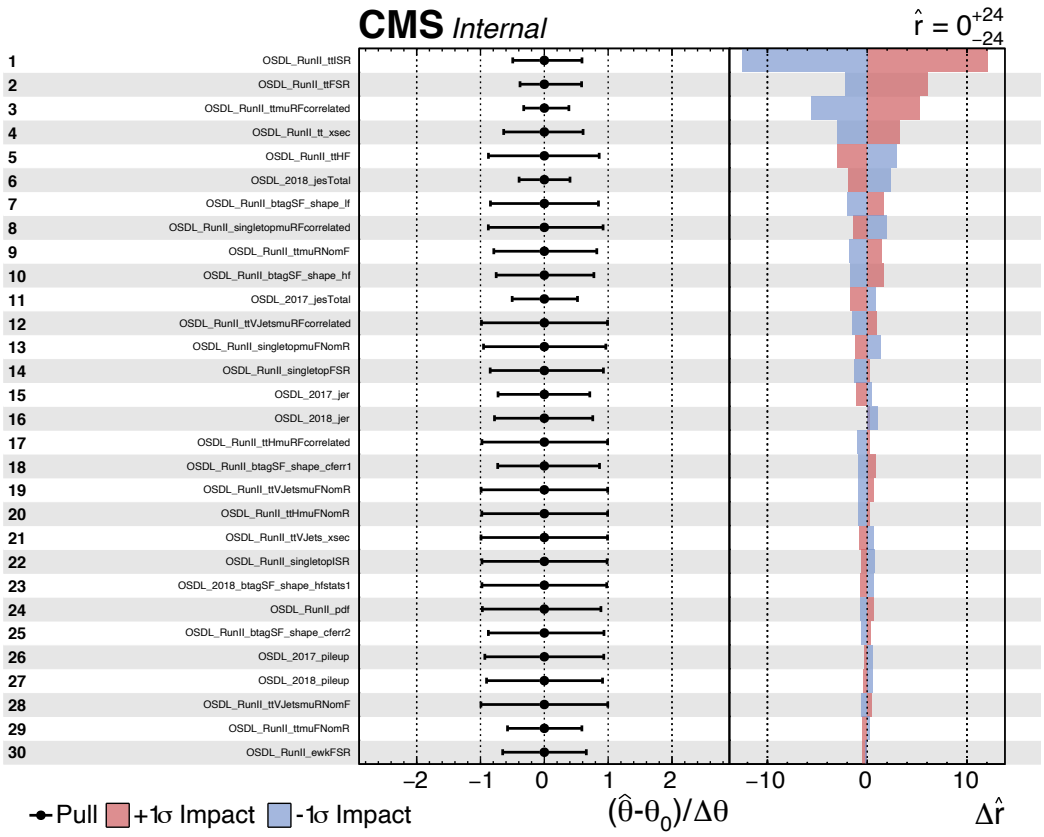

Figure D.73: Asimov impacts for the set of categories including exactly 4 jets and 2, 3, or 4+ b tags, or exactly 2 b tags and 5 or 6 jets

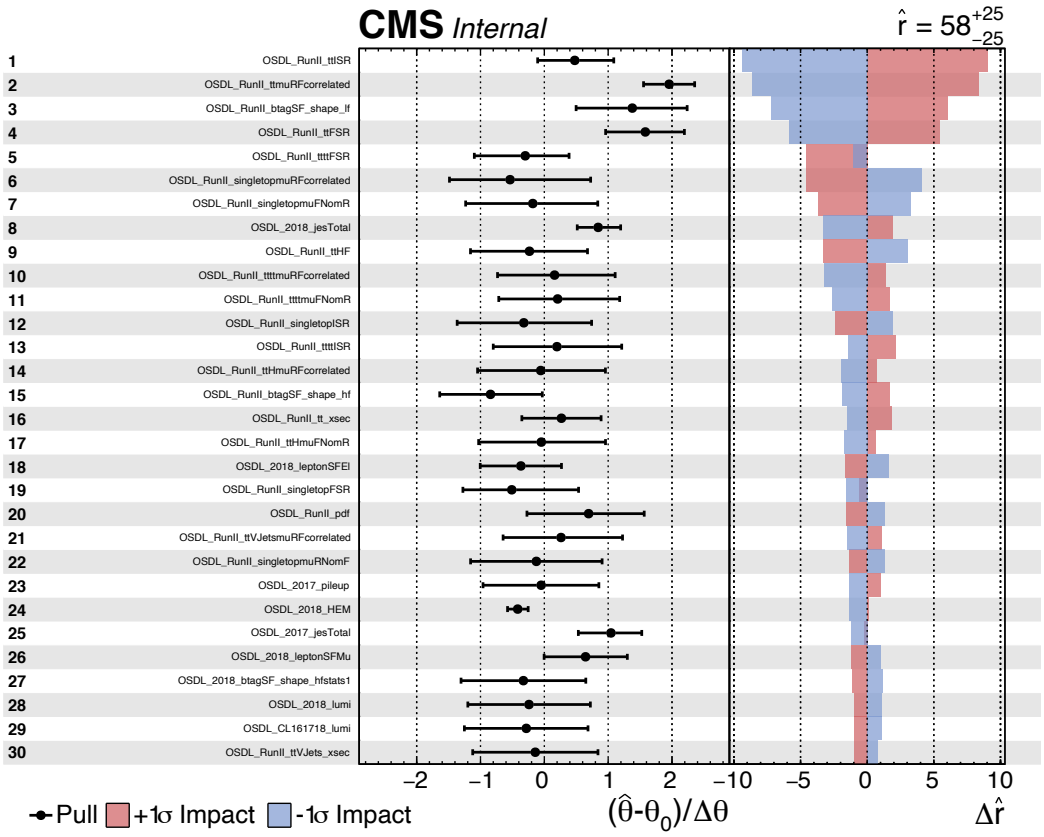

Figure D.74: Data impacts for the set of categories including exactly 4 jets and 2, 3, or 4+ b tags, or exactly 2 b tags and 5 or 6 jets

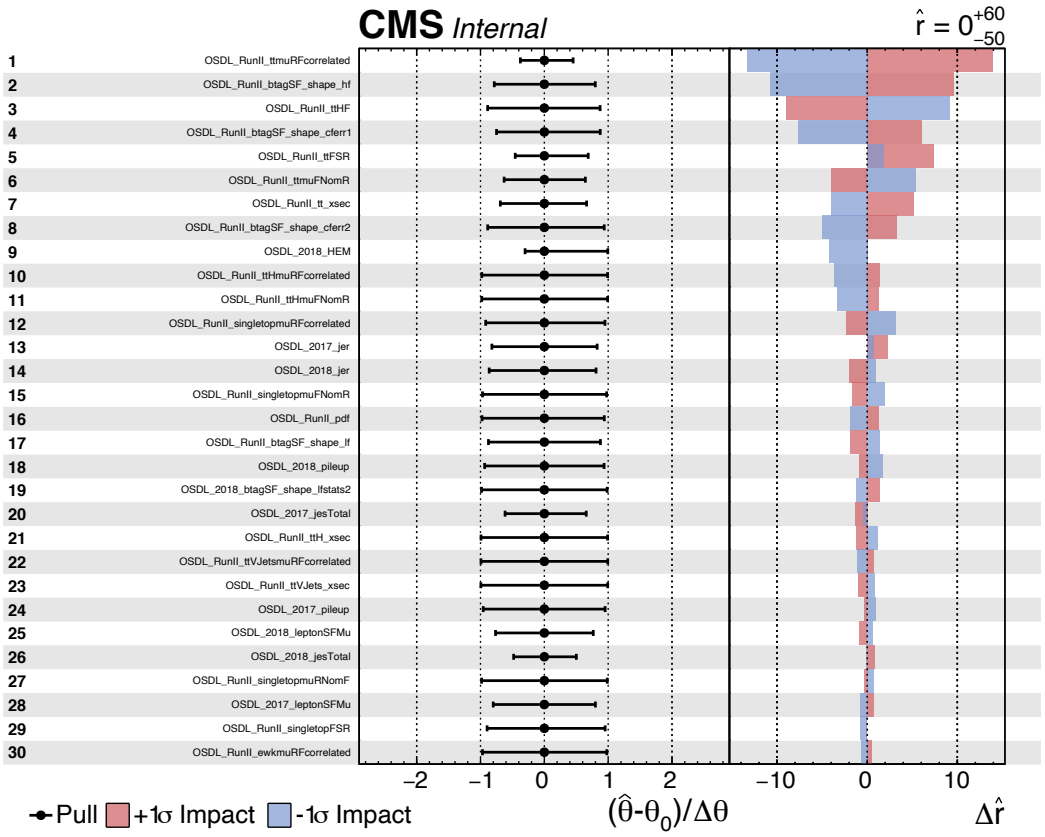

Figure D.75: Asimov impacts for the set of categories including exactly 4 jets and 2, 3, or 4+ b tags

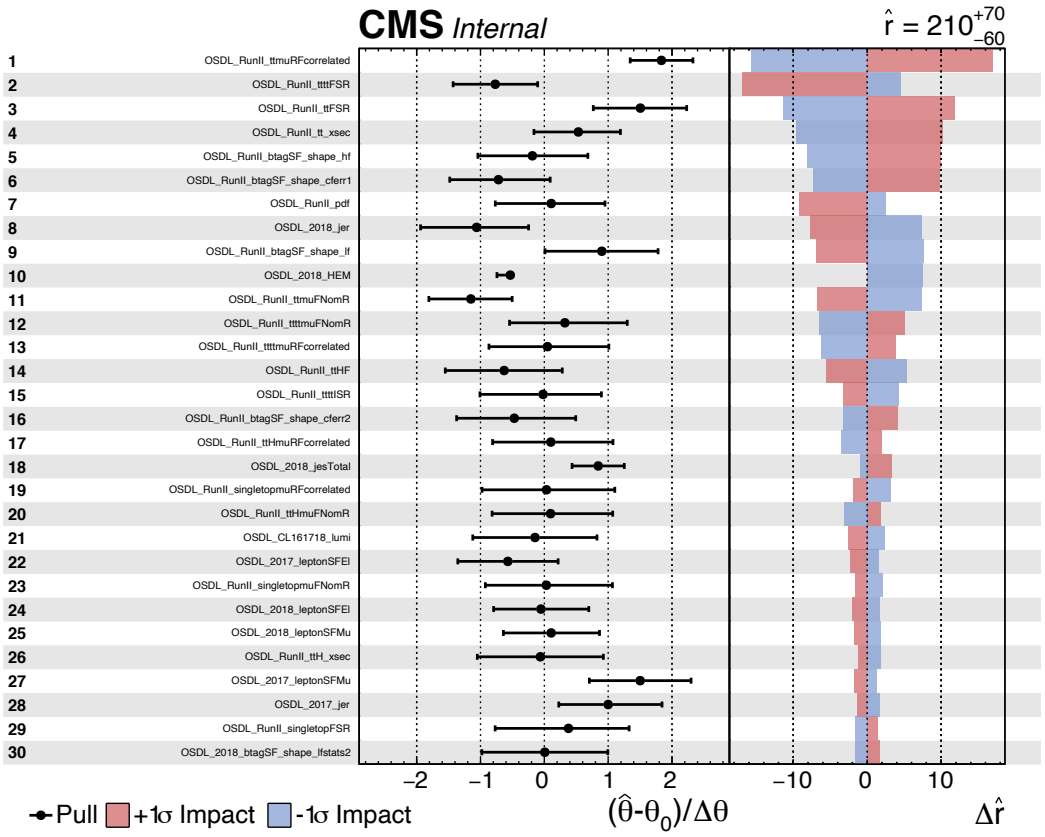

Figure D.76: Data impacts for the set of categories including exactly 4 jets and 2, 3, or 4+ b tags

D.8 Checks of Pileup Reweighting and Potential Mismodelling

The plots in Fig. D.77 demonstrate that the pileup reweighting applied to adjust the presumed pileup profile in 2018, prior to data-taking, succeeds in matching the pileup profile in real data as determined by the CMS BRIL group.

In further investigating potential sources of jet multiplicity mismodelling, spurred by the unblinding of the three channels being combined in CADI TOP-21-005 (single-lepton and all-hadronic, in addition to opposite-sign-dilepton) which demonstrated large jet multiplicity divergence in the other two channels (and a smaller one in OSDL), the ρ variables[144] were checked. The rho variables are per-event quantities designed to measure detector noise plus pileup contributions by taking the median of energies in the calorimeters, using "a grid of η-ϕ cells". These variables are effectively measures of different subtypes and subareas of unclustered energy, after the Particle Flow algorithm produces various objects like jets. Some of the ρ variables are then used in the Jet Energy Corrections, and due to their being correlated with pileup, are one potential way in which pileup bias may enter into the reconstruction of the event. As seen in Fig. D.80, the subvariable ρ_{All} demonstrates a large discrepancy between data and simulation in 2018 (seen with the data starting to trend upward while the simulation is trending down), above 30 good reconstructed primary vertices. Meanwhile, Fig. D.79 shows a weaker discrepancy for $\rho_{Central}$, and Fig. D.80 shows no substantial difference between data and simulation in the $\rho_{CentralChargedPileup}$ variable. Since in the OSDL channel the potential jet multiplicity discrepancies are correlated with dilepton decay channel ($e\mu$ showing an upward trend in data that is not present in the

background-only simulation), we conclude that this mismodelling is unrelated to the differences seen in the OSDL analysis for jet multiplicity, although it may potentially affect the other two analyses still.

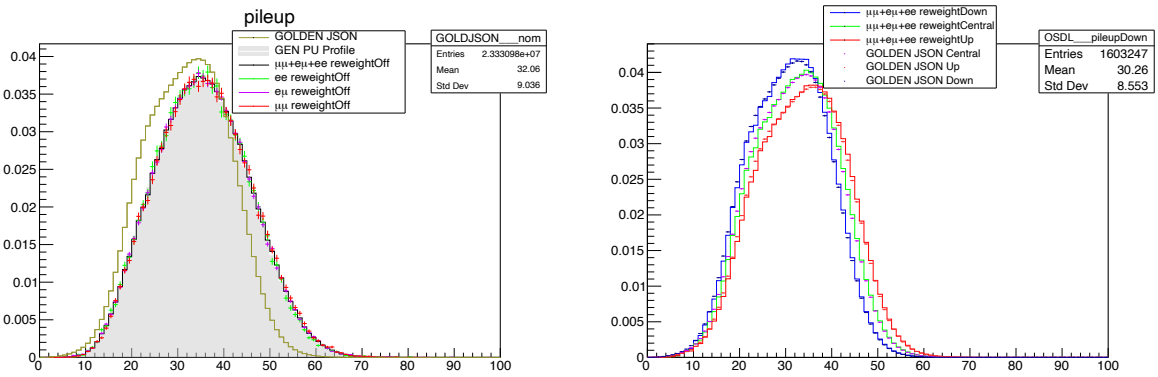

Figure D.77: Left is the simulation presumed pileup profile, GEN PU Profile; PU as derived by brilcalc, GOLDEN JSON; and our post-event selection simulations' injected pileup profiles in all channels, with PU reweighting turned off. Right is the simulation post-event selection in combined dilepton channel, with central, up, and down variations, plus corresponding GOLDEN JSON-derived targets. The right plot demonstrates the reweighting is functionioning as intended, and even the significant event cuts do not strongly bias the pileup profiles relative to the up and down shifts the reweighting algorithm targets.

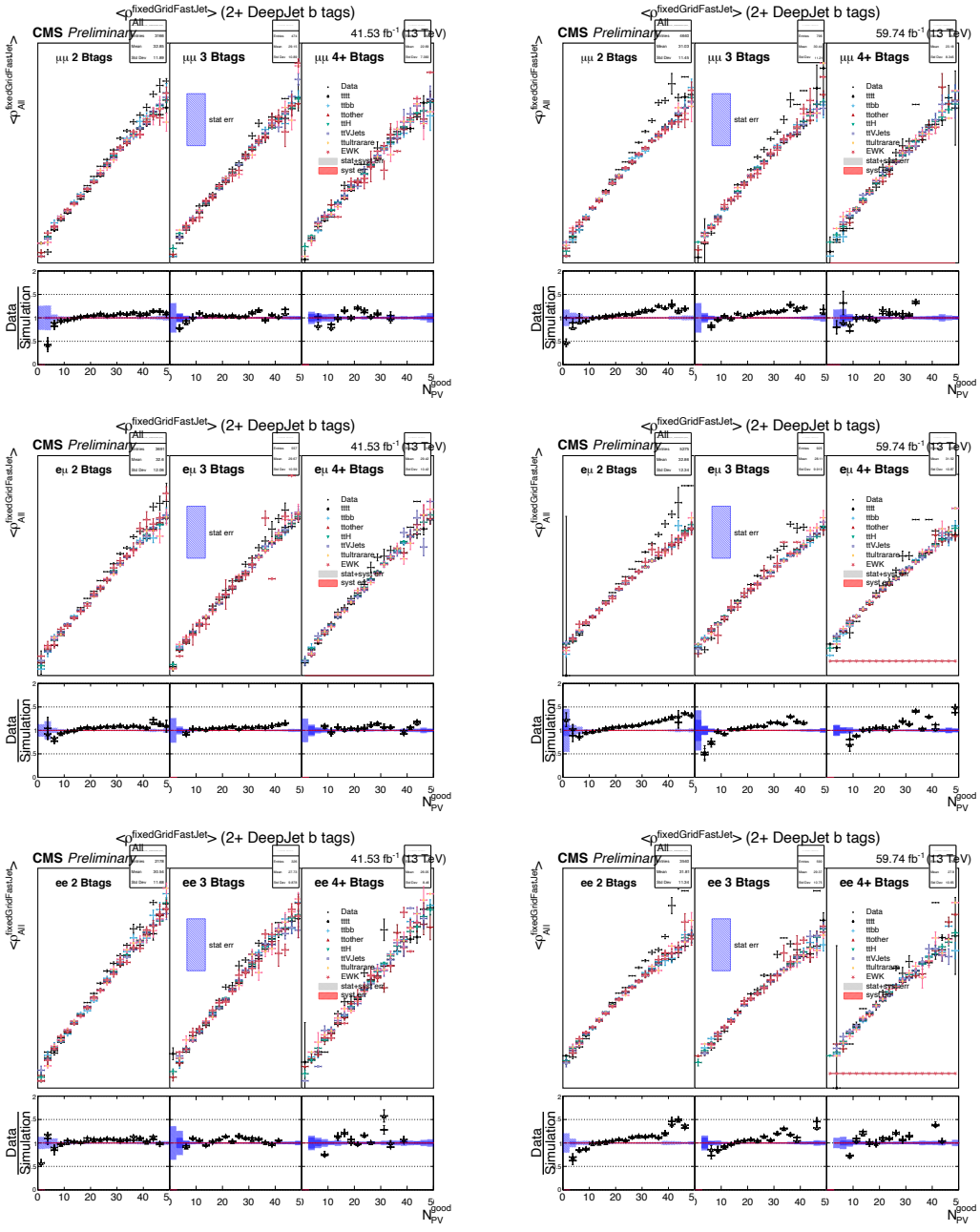

Figure D.78: Left 2017, right 2018. The $\mu\mu$ channel is shown in the top row, the eμ channel is in the middle row and the ee channel is in the bottom row. These figures show the average ρ_{All} value, with the various simulations and data plotted against the number of good pileup vertices reconstructed in the event. A divergence between data and simulation begins above 30 vertices in 2018, which does not appear to be correlated with jet multiplicity trends in the different lepton decay channels. The divergence is consistent across channels and btag multiplicities in 2018, and there is less of an effect in 2017. The ratio plot shows two ratios, one of data to $\mathrm{t\bar{t}b\bar{b}}$, and another for data to the non-$\mathrm{t\bar{t}b\bar{b}}$ subcomponent of $\mathrm{t\bar{t}}$

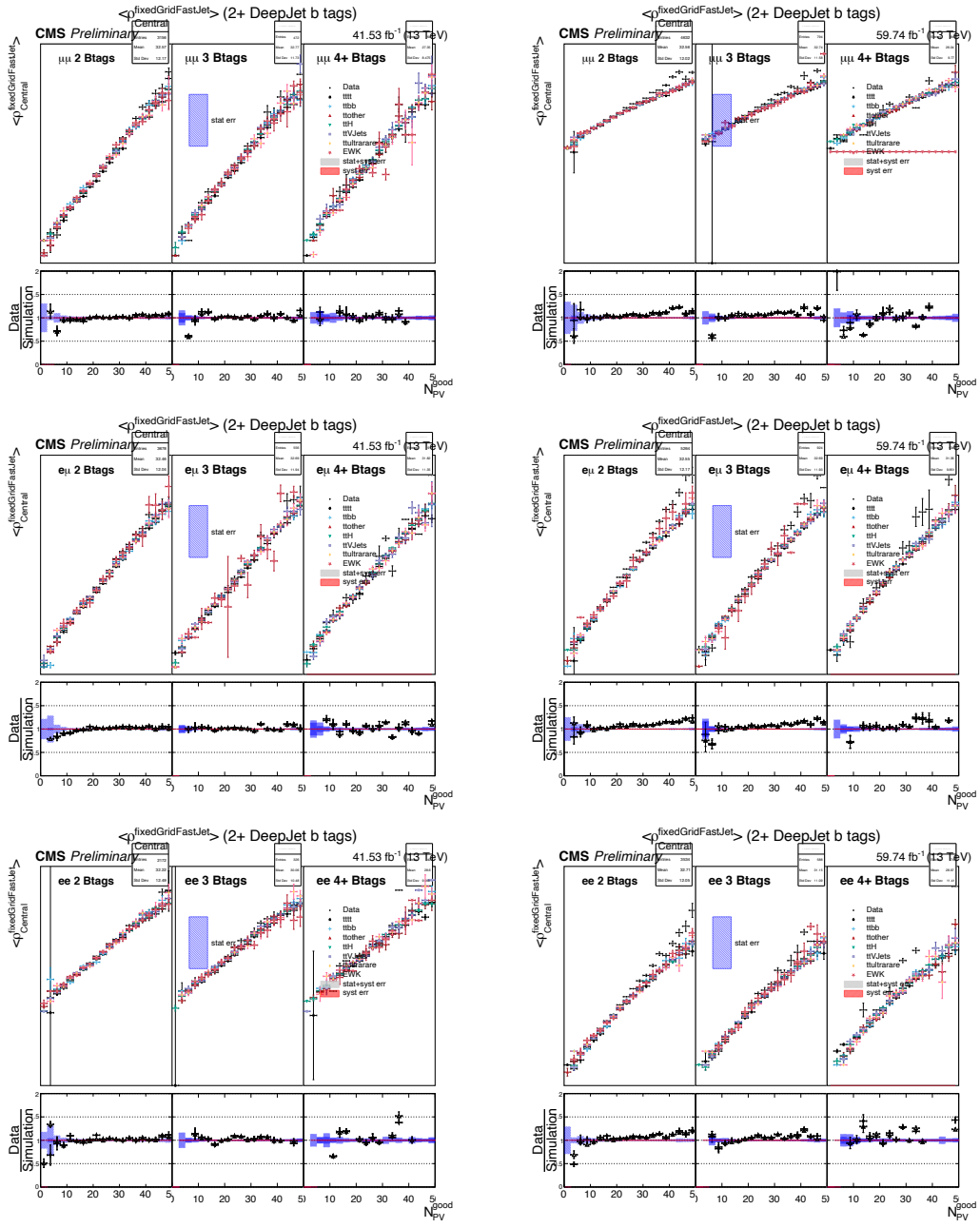

Figure D.79: Left 2017, right 2018. The $\mu\mu$ channel is shown in the top row, the eμ channel is in the middle row and the ee channel is in the bottom row. These figures show the average $\rho_{Central}$ value, with the various simulations and data plotted against the number of good pileup vertices reconstructed in the event. A divergence between data and simulation begins above 30 vertices in 2018, but it is less significant than the divergence seen in the ρ_{All} variable. The ratio plot shows two ratios, one of data to $\mathrm{t\bar{t}b\bar{b}}$, and another for data to the non-$\mathrm{t\bar{t}b\bar{b}}$ subcomponent of $\mathrm{t\bar{t}}$

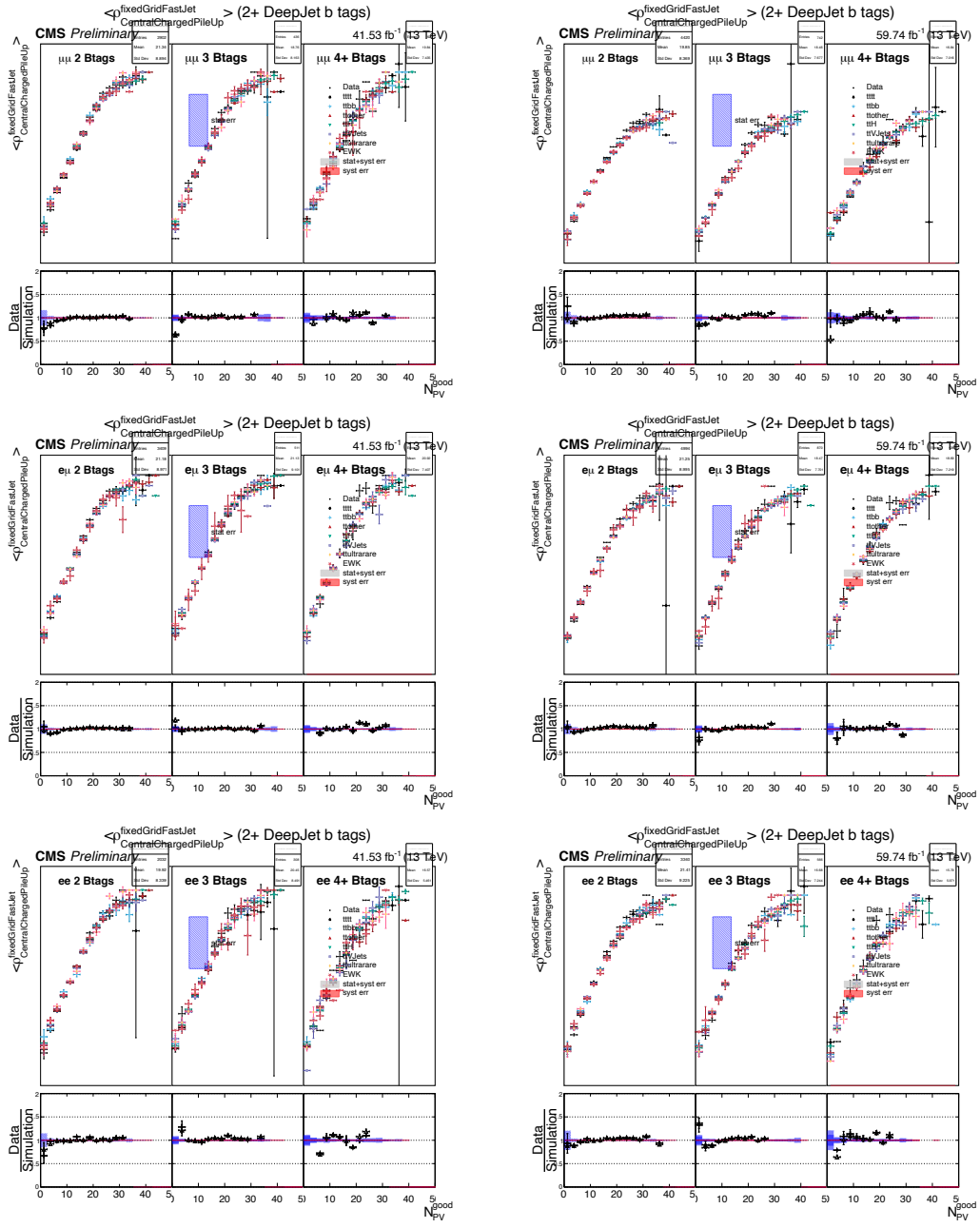

Figure D.80: Left 2017, right 2018. The $\mu\mu$ channel is shown in the top row, the eμ channel is in the middle row and the ee channel is in the bottom row. These figures show the average $\rho_{CentralChargedPileup}$ value, with the various simulations and data plotted against the number of good pileup vertices reconstructed in the event. Agreement between data and simulation appears good across all years, channels, and btag categories. The ratio plot shows two ratios, one of data to $t\bar{t}b\bar{b}$, and another for data to the non-$t\bar{t}b\bar{b}$ subcomponent of $t\bar{t}$

D.9 Resolution of H_{T} Variable in Signal

For the variable of interest to the analysis for fitting, H_{T}, we study the resolution in the main signal simulation. The procedure is as follows: first, we calculate the Generator H_{T} as the sum of Reco-matched GenJet p_{T} (When there is no match, $0p_{\mathrm{T}}$ is contributed to the Gen sum; Reco sum still includes all jets selected). For a bin in H_{T}, the quantity H_{T}^{Reco} - H_{T}^{Gen} is fit with a gaussian, and the standard deviation is set as the resolution for that H_{T} Bin. This procedure is repeated until all H_{T} bins are fit. In the next step, the resolution is fit as a function of H_{T} using the function $A - \exp(B * H_{\mathrm{T}} + C)$, where A, B, and C are the fit parameters. Some representative slices between 500 and 1100 GeV, and the final resolution fit are shown in Figs. D.81, D.82. In the final binning for fitting, a width close to the resolution (±20%) is used along with the requirements on the statistical uncertainty of the main $\mathrm{t\bar{t}}$ background. The change in limits and significance (apriori, asymptotic methods) are summarized for the 2017 + 2018 data in table D.4.

Table D.4: Summary of asymptotic apriori results for $\mathrm{t\bar{t}t\bar{t}}$ production using the RunII dataset, before and after switching to resolution-aware binning in H_{T}. A few % weakening of limits and significance are observed.

Era	Channel	Apriori limit $[\times\sigma^{SM}_{\mathrm{t\bar{t}t\bar{t}}}]$	Apriori limit [fb]	Apriori Significance [Std. Dev.]
Nonres. Binning	μμ	$5.38^{+5.62}_{-2.65}$	$64.5^{+67.4}_{-31.7}$	0.41σ
Res. Binning	μμ	$5.48^{+5.72}_{-2.68}$	$65.8^{+68.7}_{-32.1}$	0.4σ
Nonres. Binning	eμ	$3.66^{+3.77}_{-1.77}$	$43.9^{+45.2}_{-21.3}$	0.59σ
Res. Binning	eμ	$3.75^{+3.86}_{-1.82}$	$45.0^{+46.3}_{-21.8}$	0.57σ
Nonres. Binning	ee	$5.89^{+6.21}_{-2.88}$	$70.7^{+74.5}_{-34.5}$	0.37σ
Res. Binning	ee	$6.19^{+6.56}_{-3.05}$	$74.2^{+78.7}_{-36.5}$	0.36σ
Nonres. Binning	Combined	$2.66^{+2.66}_{-1.28}$	$32.0^{+31.9}_{-15.4}$	0.78σ
Res. Binning	Combined	$2.74^{+2.74}_{-1.31}$	$32.9^{+32.8}_{-15.7}$	0.76σ

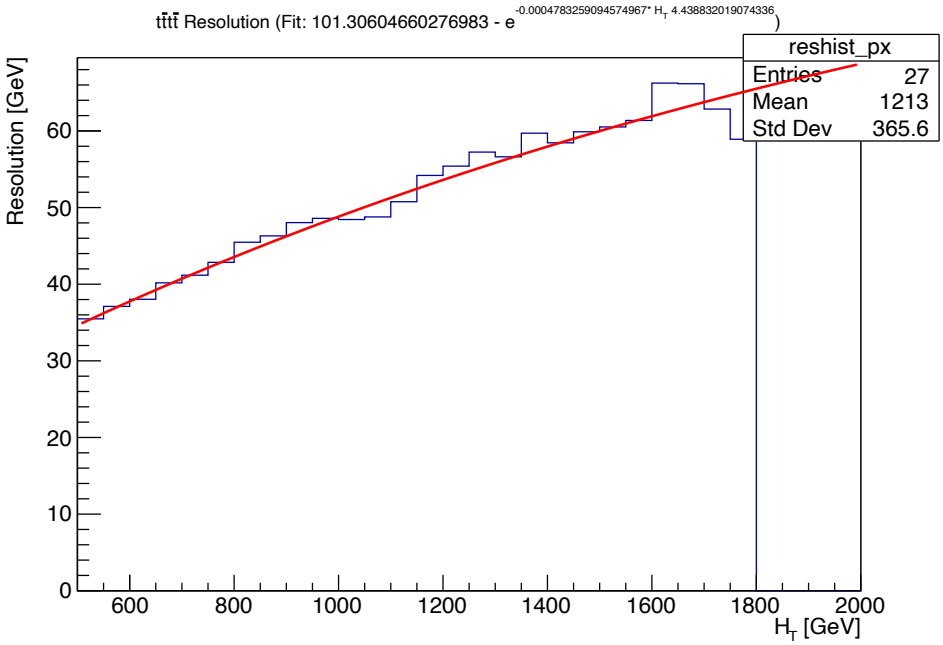

Figure D.81: The fitted resolution as a function of H_{T}, using 50 GeV slices in H_{T} for the $\mathrm{t\bar{t}t\bar{t}}$ signal simulation to determine the resolution for approximate rebinning.

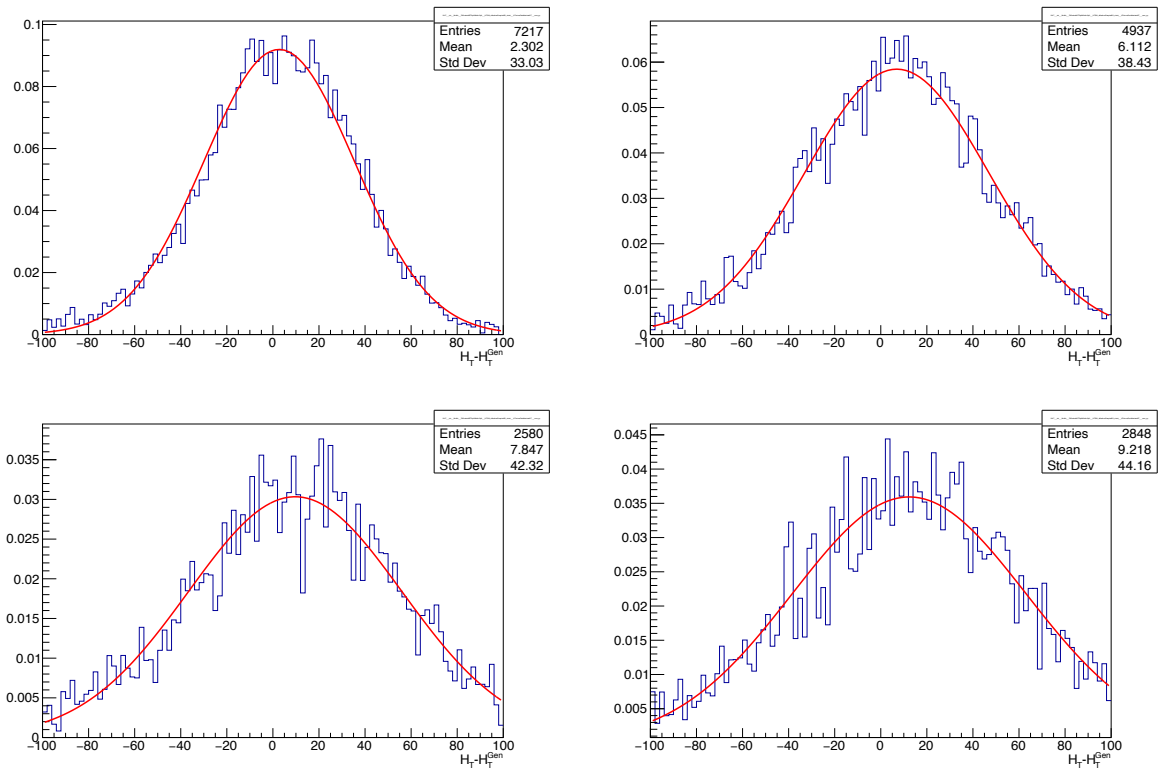

Figure D.82: Representative slices of the resolution, shown for the 50 GeV slices starting at 500 (top left), 700 (top right), 900 (bottom left), and 1150 (bottom right) GeV. For these slices, which spans the majority of the data analyzed, the resution grows from approximately 33 GeV to 45 GeV.

D.10 Checks for 2018 Data Reconstruction Differences

While investigating potential jet multiplicity modelling issues (as also noted in Appendix D.8), and following the recommendations of the CMS JMAR subgroup, the 2018 data were split into the individual data-taking periods. These are represented by scaling the data from each run period to the equivalent luminosity of the entire 2018 data taking period. The approximate luminosities for Run Periods A, B, C, and D are 14.0, 7.10, 6.94, and 31.95fb^{-1}. After scaling up the data and comparing to the simulation in the sum of jet multiplicities in the analysis phase space (4+) and divided into exactly 2 b-tags and 3+ b-tags, we find that there is no apparent distinction in jet multiplicities between the prompt reconstruction employed for Run D, versus the Re-Reconstruction calibrations, configuration and algorithms for Run A, B, and C. This is shown in Fig. D.83. Additionally, checking the H_{T} distribution, no statistically significant differences are found in 2018 (Fig. D.84). All distributions shown are pre-fit, with shape-only systematics bands included (shaded red)

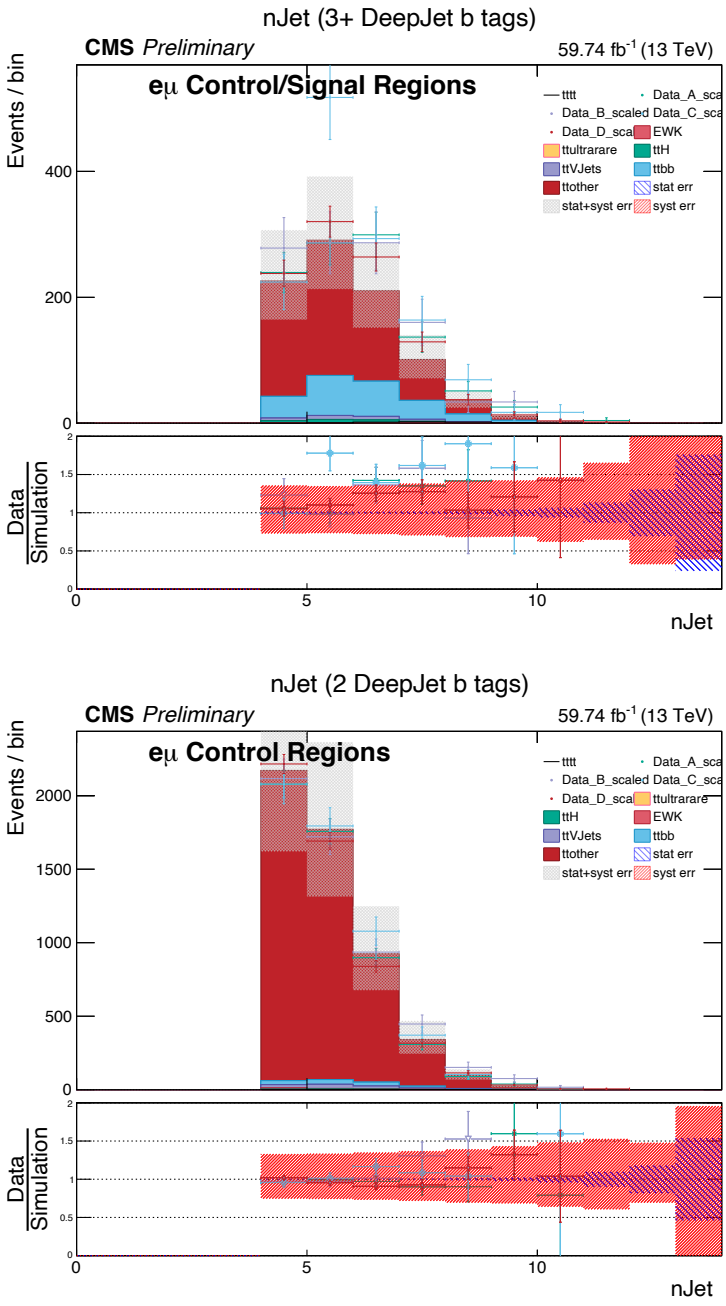

Figure D.83: Jet multiplicity comparison between run periods in 2018. Each run period is scaled up according to its integrated luminosity, such that each sub-period is treated like the full 2018 data taking period. No significant discrepancy is found between the prompt reconstruction of Run D versus earlier periods, within statistical uncertainties. 3+ b-tags are shown above, and exactly 2 b-tags below.

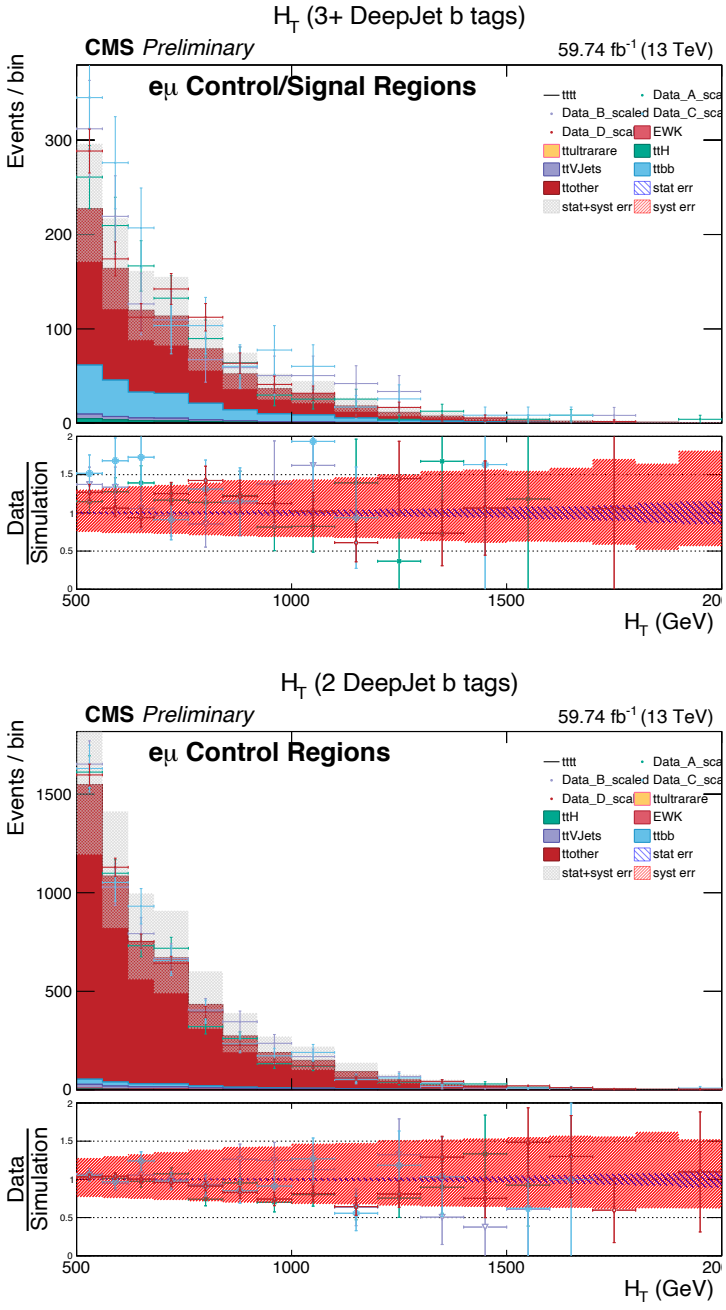

Figure D.84: H_{T} distribution for the control and the signal-enriched categories. No statistically significant differences are visually identifiable. 3+ b-tags are shown above, and exactly 2 b-tags below.

D.11 Parameterized b-tag Renormalization Maps

As described in Section 4.5, the b-tag yield renormalizations which preserve the process cross-section after the b-tag reshaping weights are applied are parameterized to capture additional dependencies not present in the original derivation of the scale-factors. Some examples from the most important processes in the analysis and some of their systematic variations are described in this section. As demonstrated in Figs. D.85, D.86, there are differences between years and processes for the average product of jet b-tag scalefactors (as these maps are the sum of event weights before and after the product of those b-tag SFs are applied, divided by eachother to provide the average ratio). Some systematic variations, such as FSR as on the middle and bottom-right plots of Fig D.85, show more dramatic variations in the ratio than other variations, like ISR on the left of the same figure. By comparing $t\bar{t}H$ and $t\bar{t}b\bar{b}$, while the jet-multiplicity and H_T dependence are similar, there are notable differenes between the processes. Because of these differences per-process, per-year, and per-systematic variation, a single renormalization-factor cannot be used. Other analyses such as those searching for $t\bar{t}H$ find effects on the shape of kinematic distributions, especially H_T, when no such parameterization is applied.

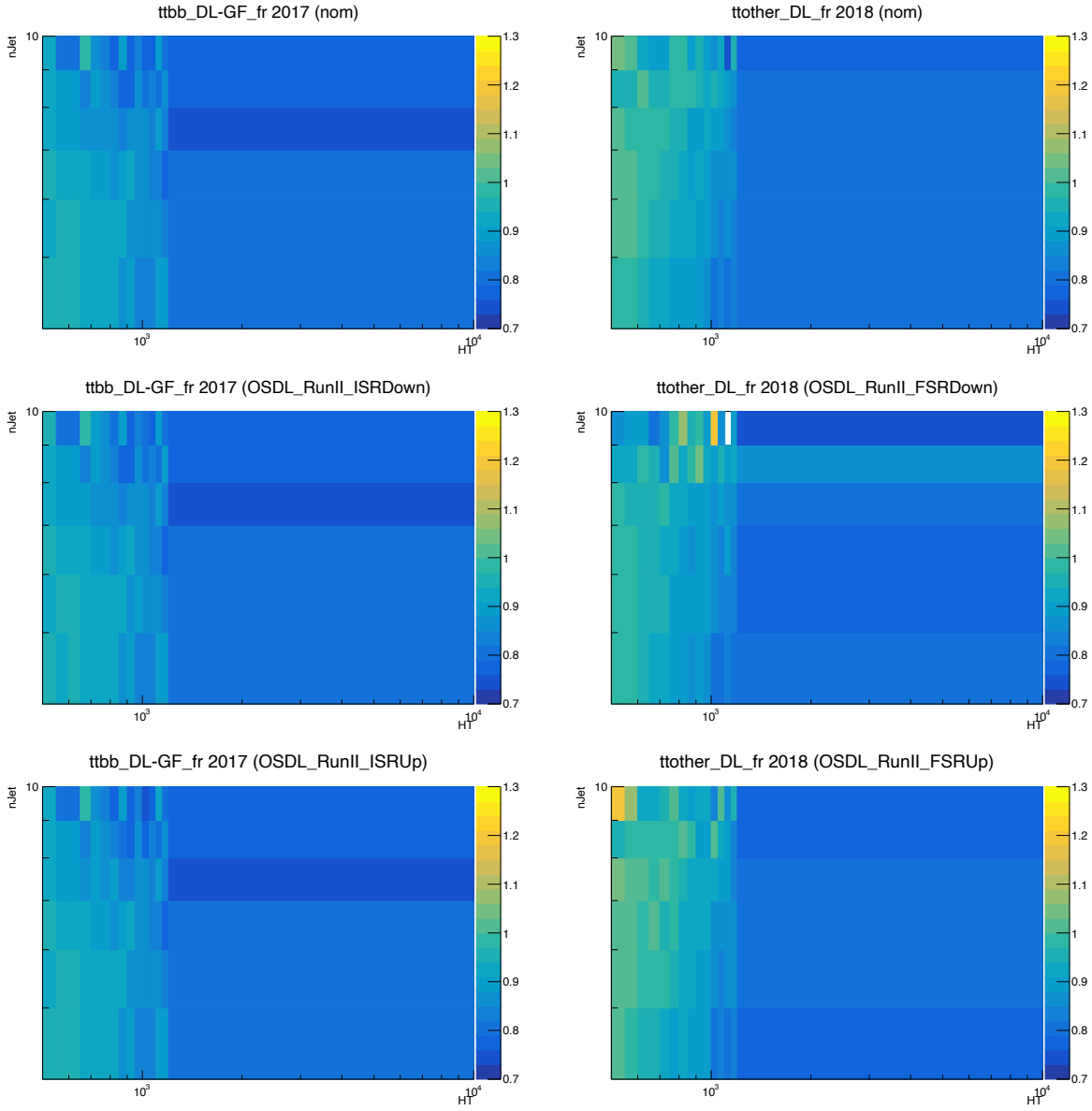

Figure D.85: On the left, renormalization factors parameterized in jet multiplicity and H_{T}, for $\mathrm{t\bar{t}b\bar{b}}$ in the gen-filtered phase space on the left and non-$\mathrm{t\bar{t}b\bar{b}}$ on the right. Left plots are from 2017, and right plots from 2018. For $\mathrm{t\bar{t}b\bar{b}}$, the top row is the nominal, and below that are the down and up variations for ISR. For non-$\mathrm{t\bar{t}b\bar{b}}$, nominal is on top and the down and up FSR variations. There is stronger scaling for FSR than ISR, and different normalization shapes for the nominal versions in different years and subprocesses.

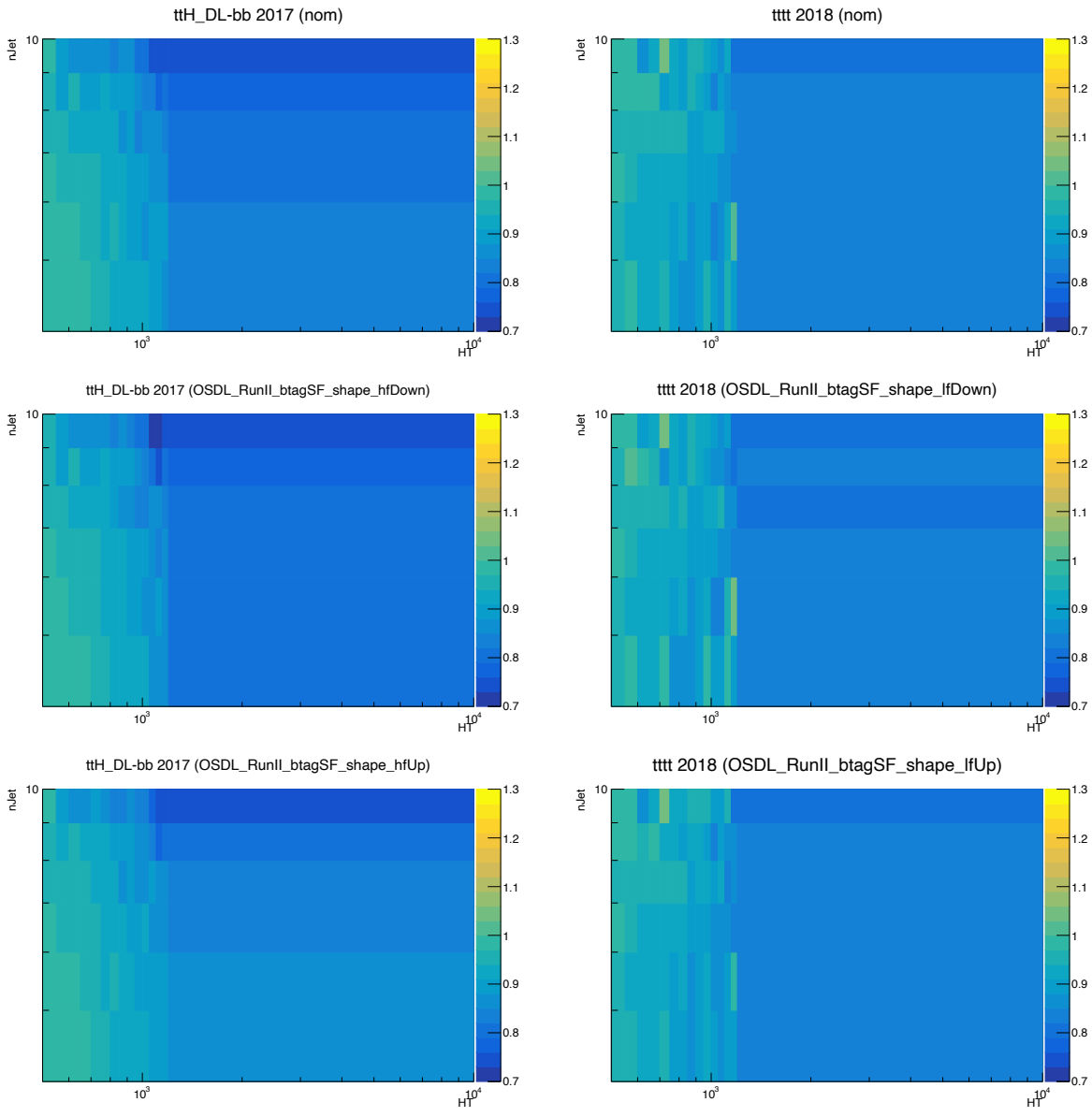

Figure D.86: On the left, renormalization factors parameterized in jet multiplicity and H_{T}, for $\mathrm{t\bar{t}H}$ on the left and $\mathrm{t\bar{t}t\bar{t}}$ on the right. Left plots are from 2017, and right plots from 2018. For $\mathrm{t\bar{t}H}$, the top row is the nominal, and below that are the down and up variations for b-tag heavy flavor contamination. For $\mathrm{t\bar{t}t\bar{t}}$, nominal is on top and the down and up b-tag light-flavor variations follow.

D.12 Updates to the $t\bar{t}b\bar{b}$ Treatment

Following a request from the Analysis Review Committee, the treatment of the $t\bar{t}b\bar{b}$ process has changed. Previously, $t\bar{t}b$ and $t\bar{t}\bar{b}$ events (identified as having one additional b-jet) were treated like $t\bar{t}$ events without any extra b-jets, for the purpose of grouping in templates and Heavy Flavor corrections. During review, it was requested to harmonize with the dedicated $t\bar{t}b\bar{b}$ measurement, in which these events are scaled like $t\bar{t}b\bar{b}$. While the dedicated analyses still keep these events separate as an additional background, they are modified in the fit the same way as the ideal signal events with two additional b-jets. In this analysis, as both of these are backgrounds, the events previously merged with "ttnobb" are now merged with $t\bar{t}b\bar{b}$, so that they are adjusted by the fit's nuisance parameters in the same way.

Multiple scenarios were studied to check step-by-step how this impacted the analysis. First, all results here used the new derivation of the Heavy Flavor reweighting scheme, in which $\sigma_{t\bar{t}b\bar{b}}$ is adjusted to match the dedicated analyses measurement (SF=1.36), with a reduced uncertainty derived only from the $\sigma_{t\bar{t}b\bar{b}}$ measurement's data uncertainty (to avoid double counting theory uncertainties), ($SF_{unc} = 8\%$). The results with the old event activity correction and old $t\bar{t}$ splitting scheme, but with this reduced HF uncertainty (which is approximately still correct even with the old derivation scheme) are labeled 'oldcorrection', while the others show results with the new correction and grouping schemes. The event activity correction can be derived and applied only to 'ttnobb' (1btagttnobb), or may be presumed to be equally applicable to all $t\bar{t}$ events (1btagttbar). Based on the majority of the 1 b-tag region being dominated by non-$t\bar{t}b(\bar{b})$ and to not perturb the $t\bar{t}b\bar{b}$ correc-

tion, we choose the 1btagttnobb scheme. The next two updates are shown, beyond this $t\bar{t}b\bar{b}$ change, in which the Parton Showering and Renormalization/Factorization scales are decorrelated between ttnobb and ttbb, and then with the JES uncertainties split into 11 correlated and uncorrelated (per-year) sources. The Asimov a-priori and a-posteriori results show stability and convergence in the Control Regions scheme (in which all eleven nominal jet x btag categories are used excluding the four categories per year and channel with (3 or 4+ b-tag) and (7 or 8+ jets)). Results are also shown for the full Nominal fit with this series of changes, for posterity, as a final demonstration of how these requested changes impact results on Asimov data a-priori and a-posteriori.

The top 40 impacts are shown for 3 cases, all corresponding to the scenario 'jes_reduced' in which the updated $t\bar{t}b\bar{b}$ grouping, correction, 1 b-tag event activity correction, decorrelation scheme, and 11-source reduced JES are all applied. First, in Fig. D.87, the impacts for Asimov and Observed are shown in the CR11 fit. For the full nominal fit, these are contained in Fig. D.88. Finally, as a test of the stability of the nuisance parameters in the fit, the observed impacts are shown between the CR11 and Nominal fit, Fig. D.89. These last show that all the nuisances of concern converge even when eliminating 4 of the 15 jet x b-tag categories used per channel and year.

Table D.5: Summary of asymptotic cross section limits for $t\bar{t}t\bar{t}$ production under different ttbb merging and correction scenarios, with Control Region (CR11) fits without the 4 most significant categories per channel/year and Nominal (NOM) fits with the usual categories.

Method	Regions	Apriori limit $[\times\sigma^{SM}_{t\bar{t}t\bar{t}}]$	Aposteriori limit $[\times\sigma^{SM}_{t\bar{t}t\bar{t}}]$
oldcorrection	CR11	$11.5^{+10.1}_{-5.3}$	$10.7^{+9.6}_{-4.95}$
1btagttbar	CR11	$10.3^{+9.07}_{-4.74}$	$11.7^{+9.92}_{-5.42}$
1btagttnobb	CR11	$10.9^{+9.59}_{-4.98}$	$11.7^{+9.51}_{-5.39}$
decorrelated	CR11	$10.9^{+9.40}_{-4.93}$	$11.8^{+10.30}_{-5.44}$
jes_reduced	CR11	$10.8^{+9.37}_{-4.96}$	$11.8^{+10.3}_{-5.43}$
oldcorrection	NOMINAL	$3.06^{+2.96}_{-1.45}$	$2.79^{+2.70}_{-1.33}$
1btagttbar	NOMINAL	$2.64^{+2.59}_{-1.25}$	$2.85^{+2.80}_{-1.36}$
1btagttnobb	NOMINAL	$2.88^{+2.77}_{-1.36}$	$2.96^{+2.87}_{-1.41}$
decorrelated	NOMINAL	$2.93^{+2.72}_{-1.36}$	$2.96^{+2.85}_{-1.40}$
jes_reduced	NOMINAL	$2.91^{+2.72}_{-1.35}$	$3.02^{+2.91}_{-1.43}$

Table D.6: Summary of asymptotic cross section limits for $t\bar{t}t\bar{t}$ production under different ttbb merging and correction scenarios, with Control Region (CR11) fits without the 4 most significant categories per channel/year and Nominal (NOM) fits with the usual categories.

Method	Regions	Apriori limit [fb]	Aposteriori limit [fb]
oldcorrection	CR11	$138.0^{+121.0}_{-63.6}$	$128.0^{+115.0}_{-59.4}$
1btagttbar	CR11	$123.0^{+109.0}_{-56.9}$	$140.0^{+119.0}_{-65.0}$
1btagttnobb	CR11	$131.0^{+115.0}_{-59.8}$	$140.0^{+114.0}_{-64.6}$
decorrelated	CR11	$130.0^{+113.0}_{-59.1}$	$142.0^{+123.0}_{-65.3}$
jes_reduced	CR11	$130^{+112}_{-59.5}$	$141^{+124}_{-65.2}$
oldcorrection	NOMINAL	$36.8^{+35.5}_{-17.4}$	$33.5^{+32.4}_{-16.0}$
1btagttbar	NOMINAL	$31.7^{+31.1}_{-15.0}$	$34.2^{+33.5}_{-16.3}$
1btagttnobb	NOMINAL	$34.5^{+33.2}_{-16.3}$	$35.5^{+34.4}_{-16.9}$
decorrelated	NOMINAL	$35.2^{+32.7}_{-16.3}$	$35.5^{+34.2}_{-16.8}$
jes_reduced	NOMINAL	$35.0^{+32.6}_{-16.3}$	$36.3^{+34.9}_{-17.1}$

Table D.7: Summary of asymptotic significances for $t\bar{t}t\bar{t}$ production under different ttbb merging and correction scenarios, with Control Region (CR11) fits without the 4 most significant categories per channel/year and Nominal (NOM) fits with the usual categories.

Method	Regions	Apriori signif. [Std. Dev.]	Aposteriori signif. [Std. Dev.]
oldcorrection	CR11	0.16σ	0.19σ
1btagttbar	CR11	0.18σ	0.17σ
1btagttnobb	CR11	0.17σ	0.17σ
decorrelated	CR11	0.17σ	0.17σ
jes_reduced	CR11	0.17σ	0.17σ
oldcorrection	NOMINAL	0.65σ	0.77σ
1btagttbar	NOMINAL	0.76σ	0.73σ
1btagttnobb	NOMINAL	0.70σ	0.70σ
decorrelated	NOMINAL	0.64σ	0.70σ
jes_reduced	NOMINAL	0.64σ	0.69σ

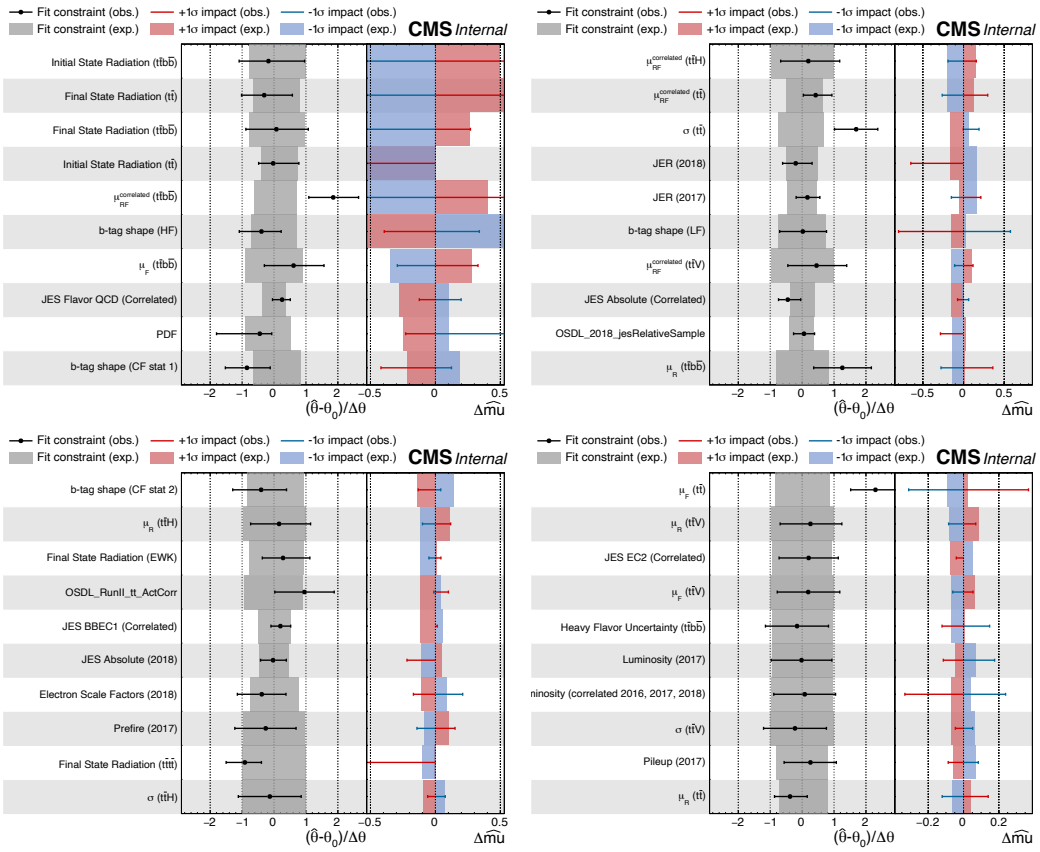

Figure D.87: Observed and Asimov impacts for the CR11 fit, for the scenario with jes split and the updated $\mathrm{t\bar{t}b\bar{b}}$ proscription used.

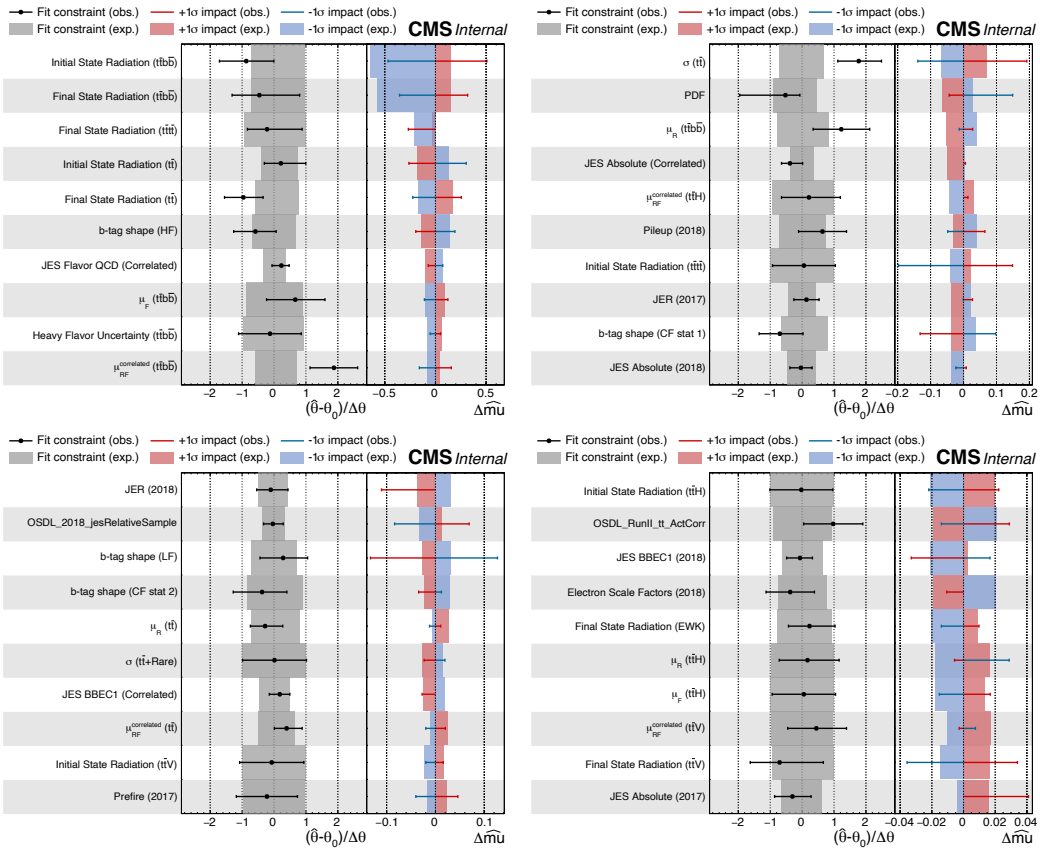

Figure D.88: Observed and Asimov impacts for the Nominal fit, for the scenario with jes split and the updated $\mathrm{t\bar{t}b\bar{b}}$ proscription used.

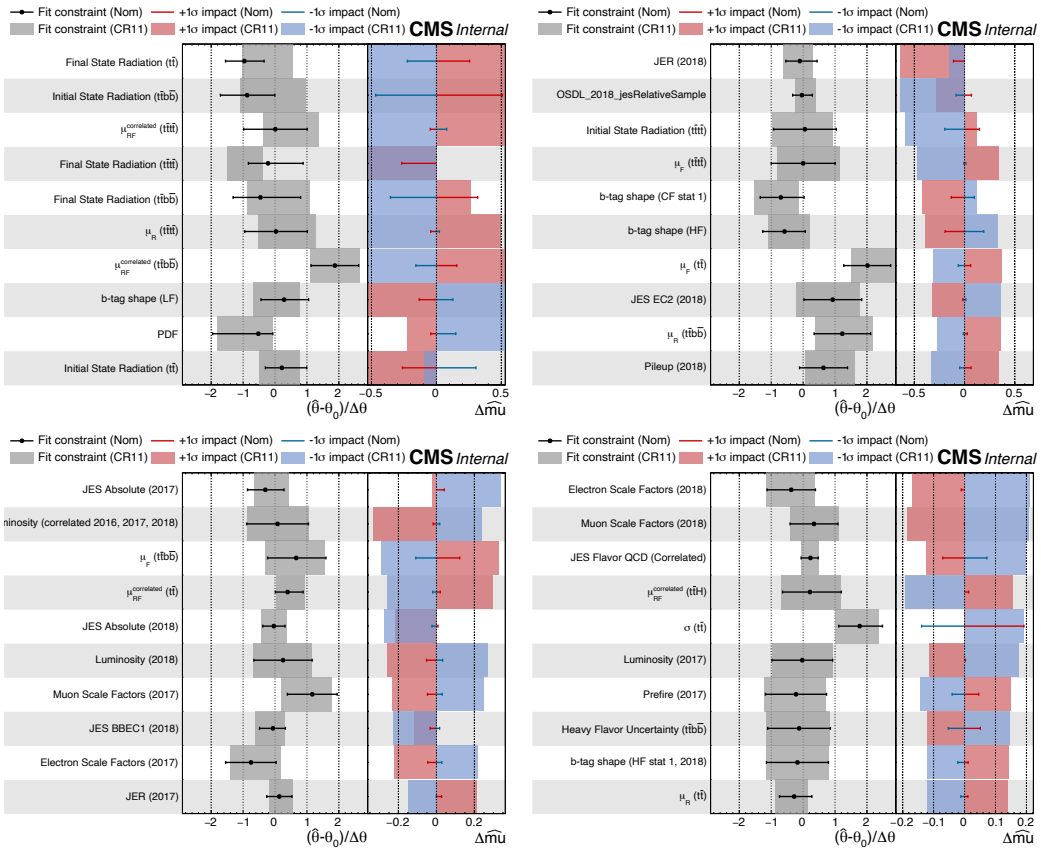

Figure D.89: Observed impacts for the CR11 and Nominal fit, for the scenario with jes split and the updated $\mathrm{t\bar{t}b\bar{b}}$ proscription used. This comparison shows that all the nuisances are consistend in the control-region fit and the full fit.

D.13 Goodness of Fit

The Higgs Combine tool contains a Goodness of Fit test, run in saturated mode [109]. The test is run on both the observed data as well as toys generated from the binned simulation, in each case calculating a test statistic. The observed test statistic ideally is compatible with the distribution of test statistics from the toys.

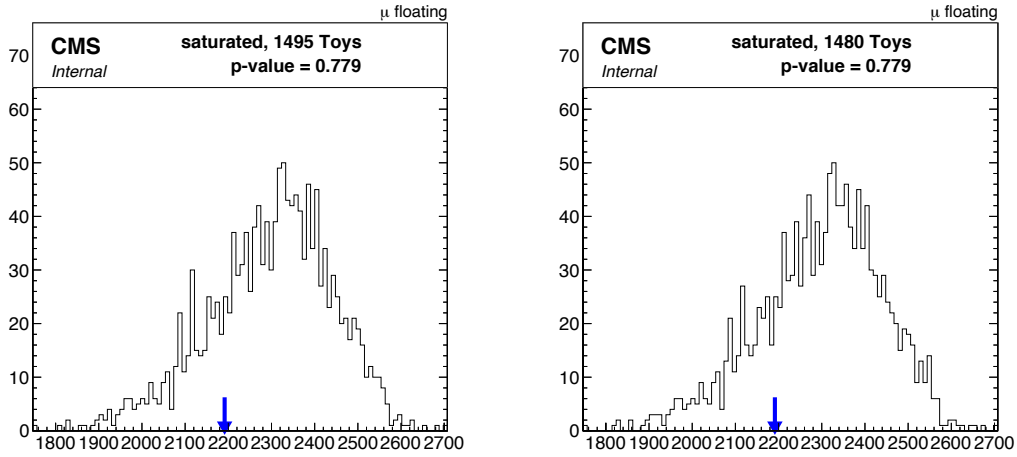

Figure D.90: Higgs Combine Goodness of Fit test run on the OSDL 2017 and 2018 data combined, where the fit is left unconstrained (floating). On left is the fit with rate effects removed from the signal (as for cross-section extraction and limits in fb) and on the right is the plot with rate effects remaining, the most significant of which come from the renormalization and factorization scales.

D.14 Channel Compatibility Test

The Channel Compatibility Check was run on the OSDL fit, partitioning by year and channel simultaneously, as well as by year alone. The commands used and their results are detailed below. This mode of Combine runs the fit with multiple signal-strength parameters, one for each channel. The output includes the fit for each channel, compared to all others being floated simultaneously and the unified "Nominal" fit.

```
combine -M ChannelCompatibilityCheck -m 690 --expectSignal 1
--datacard RunII_AllChan_AllSyst_HT.root
--group OSDL_2018_ElEl --group OSDL_2018_ElMu --group OSDL_2018_MuMu
--group OSDL_2017_ElEl --group OSDL_2017_MuMu --group OSDL_2017_ElMu


 --- ChannelCompatibilityCheck ---
Nominal fit  : r =  2.3732  -1.5392/+1.9331
Alternate fit: r =  0.0000  -0.0000/+2.7383   in channel OSDL_2017_ElEl
Alternate fit: r =  3.8725  -2.4177/+3.1729   in channel OSDL_2017_ElMu
Alternate fit: r =  1.6604  -1.6609/+3.4176   in channel OSDL_2017_MuMu
Alternate fit: r =  4.5315  -3.2620/+4.1536   in channel OSDL_2018_ElEl
Alternate fit: r =  5.5286  -2.4315/+3.3345   in channel OSDL_2018_ElMu
Alternate fit: r =  0.0000  -0.0000/+1.3888   in channel OSDL_2018_MuMu
Chi2-like compatibility variable: 6.04064
```

Splitting the fit into only two sub-channels, corresponding to unified dilepton decays for each of the two years gives the following results:

```
combine -M ChannelCompatibilityCheck -m 690 --expectSignal 1
--datacard RunII_AllChan_AllSyst_HT.root
--group OSDL_2018 --group OSDL_2017


 --- ChannelCompatibilityCheck ---
Nominal fit  : r =  2.3732  -1.5392/+1.9331
Alternate fit: r =  1.7949  +0.0000/+1.8889   in channel OSDL_2017
Alternate fit: r =  2.9730  +0.0000/+2.0679   in channel OSDL_2018
Chi2-like compatibility variable: 0.288225
```