

CMS Draft Analysis Note

The content of this note is intended for CMS internal use and distribution only

2019/10/29

Archive Hash: bf6fdbd6

Archive Date: 2019/10/29

Measurement of the differential $t\bar{t}$ production cross section for high- p_T top quarks in $e/\mu+jets$ final states at 13 TeV

Susan Dittmer, Louise Skinnari, and Julia Thom
Cornell University

Abstract

This analysis note describes a measurement of the differential top quark pair production cross section for top quarks with high transverse momentum (p_T) in pp collisions at a center-of-mass energy of 13 TeV. The measurement is performed using data collected at CMS during the 2016 run, corresponding to an integrated luminosity of 35.9 fb^{-1} . The measurement is performed for events where one top quark decays hadronically and is reconstructed as a single large-radius jet with $p_T > 400 \text{ GeV}$, while the other top quark decays leptonically to a b-jet, an electron or a muon, and a neutrino. Jet substructure techniques are used to identify the large-radius jet as a top jet candidate.

This box is only visible in draft mode. Please make sure the values below make sense.

PDFAuthor:

Louise Skinnari

PDFTitle:

Measurement of the differential $t\bar{t}$ production cross section for high- p_T top quarks in $e/\mu+jets$ final states at 13 TeV

PDFSubject:

CMS

PDFKeywords:

CMS, physics, top, boosted

Please also verify that the abstract does not use any user defined symbols

1 Contents

2	1	Introduction	3
3	2	Data and Monte Carlo samples	4
4	2.1	Data sample	4
5	2.2	Monte Carlo simulation	4
6	3	Object reconstruction	6
7	3.1	Primary vertex	6
8	3.2	Muons	6
9	3.3	Electrons	7
10	3.4	Jets	7
11	3.5	Missing transverse energy	9
12	4	Event selection	10
13	5	Background estimation	12
14	5.1	Data-driven QCD	12
15	6	Comparison of data and MC prediction	15
16	6.1	Event yields	15
17	6.2	Kinematic distributions	15
18	7	Systematic uncertainties	22
19	7.1	Experimental uncertainties	22
20	7.2	Theoretical uncertainties	23
21	8	Kinematic Fit	25
22	8.1	Likelihood template fit	25
23	8.2	Interpretation of results	26
24	9	Differential Cross Section	42
25	9.1	Phase Space	42
26	9.2	$t\bar{t}$ Simulation	43
27	9.3	Response Matrix	43
28	9.4	TUnfold	45
29	9.5	MC studies	51
30	9.6	Differential cross section for electron/muon channels	51
31	9.7	Differential cross section for combined channels	58
32	10	Conclusions	70
33	A	Lepton Optimization	73
34	A.1	Efficiency distributions	73
35	A.2	ROC curves	74
36	B	Trigger Efficiencies	78
37	C	Selection Optimization	81
38	C.1	2015 commissioning	81
39	C.2	Final optimization	84
40	D	Additional QCD studies	89
41	E	Additional Kinematic Distributions	95
42	F	Fit Cross Checks	120
43	G	Top tag SF p_T and $ \eta $ dependence	140

44	G.1	Top-tag SF binned in p_T	140
45	G.2	Top-tag SF binned in $ \eta $	140
46	G.3	Results	149
47	H	Unfolding Cross Checks	150
48	H.1	Unfolding Procedure	150
49	H.2	Unfolding Cross Checks for p_T unfolding	150
50	H.3	Unfolding Cross Checks for rapidity unfolding	165

DRAFT

51 1 Introduction

52 This analysis note presents a measurement of the differential top quark pair production ($t\bar{t}$)
53 cross section for top quarks with high transverse momentum $p_T > 400$ GeV at 13 TeV. Similar
54 measurements were performed at 8 TeV by CMS [1] and by ATLAS [2]. The measurement
55 presented in this analysis note is complimentary to other differential $t\bar{t}$ cross section measure-
56 ments from CMS at 13 TeV [3], which have reduced sensitivity to events where top quarks
57 are produced with large Lorentz boosts, resulting in a topology with merged decay products
58 ("boosted topology"). Other 13 TeV measurements identify events where each individual par-
59 ton from the sequential top quark decay is matched to one and only one jet. Since the angular
60 distance between the decay products of a particle of mass m and energy E is $\Delta\theta = 2m/E$,
61 requiring one parton to be matched to one jet effectively puts a maximum energy selection
62 of $2m/\Delta\theta$. By allowing partons to merge into a single jet, the boosted topology extends the
63 range of energies in the measurement. Accurately modeling this boosted top quark regime is
64 important as it is sensitive to new physics.

65 For the analysis, semileptonic $t\bar{t}$ events are selected where the hadronically decaying top quark
66 is boosted, such that its decay products are merged and reconstructed as a single large-radius
67 t -tagged jet, while the other top quark is resolved, decaying to a W boson and a b quark. The W
68 boson is in turn required to decay leptonically to a lepton (e/μ) and missing transverse energy
69 (E_T^{miss}) from the escaping neutrino.

DRAFT

70 **2 Data and Monte Carlo samples**

71 **2.1 Data sample**

72 The measurement is performed at 13 TeV using the full 2016 CMS pp data sample. Data are se-
 73 lected from the SingleMuon and SingleElectron data streams, using single muon/electron trig-
 74 gers without lepton isolation criteria applied. The HLT path HLT_ELE45_CALOIDVT_GSFTRKIDT-
 75 _PFJET200_PFJET50 is used for the e+jets channel, and the path HLT_MU40_ETA2P1_PFJET200-
 76 _PFJET50 is used for the μ +jets channel. The MiniAOD-v2 format is used, with the CMSSW_8_0_X
 77 release. The measurement utilizes the full 2016 dataset, Runs B-H. The 23Sep2016 re-reco is
 78 used for Runs B-G, and PromptReco for Run H. The Golden JSON is applied. This data sample
 79 corresponds to an integrated luminosity of 35867 pb^{-1} .

80 **2.2 Monte Carlo simulation**

81 Monte Carlo simulation is used to generate samples for the $t\bar{t}$ signal and to model most of the
 82 backgrounds. Signal $t\bar{t}$ events are simulated using POWHEG (v.2) [4–7] which is next-to-leading
 83 order (NLO) for the matrix element calculation, matched to PYTHIA8 (v.8.205) [8, 9] for the
 84 parton shower using tune CUETP8M2T4 [10]. A top quark of mass of 172.5 GeV is used. The
 85 $t\bar{t}$ simulation is normalized to the next-to-next-to-leading order (NNLO) calculation [11].

86 The dominant backgrounds contributing in this analysis include single top quark, W+jets,
 87 Z+jets, diboson, and QCD multijet projection. Single top quark production contains contrib-
 88 utions from t-channel, tW-channel, and s-channel processes, while diboson backgrounds in-
 89 clude contributions from WW, WZ, and ZZ processes. POWHEG is used to simulate both the
 90 tW-channel and t-channel processes, while aMC@NLO [12] is used to simulate the s-channel
 91 contribution. MadGraph MLM is used to simulate W+jets, Z+jets and QCD. PYTHIA8 is used
 92 to simulate the diboson processes. For all processes, the generator is interfaced to PYTHIA8 for
 93 the hadronization. Tune CUETP8M1 is used for all background processes. W+jets and QCD
 94 are generated binned in HT to increase statistics.

95 Table 1 provides a list of the MC samples used.

96 **2.2.1 Corrections and pileup reweighting**

97 Correction factors are applied to the simulated samples to match the conditions observed in
 98 data.

99 All simulated samples are reweighted to accurately model the pileup conditions in data. All
 100 MC samples include simulation of in-time and out-of-time pileup, though its exact distribution
 101 is modified to match the data. The average number of pileup events per bunch crossing (μ)
 102 is calculated from the measured instantaneous luminosity and the total inelastic cross section
 103 of 69.2 mb. To reweight the simulation, event weights are calculated from the ratio of the μ
 104 distribution used in the simulation and corresponding measured μ distribution.

Sample	N_{events}	σ [pb]	K-factor
TT_TuneCUETP8M2T4_13TeV-powheg-pythia8	77229341	831.76	-
ST_tW_top_5f_NoFullyHadronicDecays_13TeV-powheg_TuneCUETP8M1	8629641	19.3	-
ST_tW_antitop_5f_NoFullyHadronicDecays_13TeV-powheg_TuneCUETP8M1	8681541	19.3	-
ST_t-channel_top_4f_inclusiveDecays_13TeV-powhegV2-madspin-pythia8_TuneCUETP8M1	67240808	136.02	-
ST_t-channel_antitop_4f_inclusiveDecays_13TeV-powhegV2-madspin-pythia8_TuneCUETP8M1	38811017	80.95	-
ST_s-channel_4f_leptonDecays_13TeV-amcatnlo-pythia8	9651642	10.32 * 0.322	-
WJetsToLNu_HT-100To200_TuneCUETP8M1_13TeV-madgraphMLM-pythia8	39617787	1345	1.21
WJetsToLNu_HT-200To400_TuneCUETP8M1_13TeV-madgraphMLM-pythia8	19914590	359.7	1.21
WJetsToLNu_HT-400To600_TuneCUETP8M1_13TeV-madgraphMLM-pythia8	5796237	48.91	1.21
WJetsToLNu_HT-600To800_TuneCUETP8M1_13TeV-madgraphMLM-pythia8	14822888	12.05	1.21
WJetsToLNu_HT-800To1200_TuneCUETP8M1_13TeV-madgraphMLM-pythia8	6200954	5.501	1.21
WJetsToLNu_HT-1200To2500_TuneCUETP8M1_13TeV-madgraphMLM-pythia8	6324934	1.329	1.21
WJetsToLNu_HT-2500ToInf_TuneCUETP8M1_13TeV-madgraphMLM-pythia8	2384260	0.0322	1.21
DYJetsToLL_M-50_TuneCUETP8M1_13TeV-madgraphMLM-pythia8	42923575	5765	-
WW_TuneCUETP8M1_13TeV-pythia8	6987124	118.7	-
WZ_TuneCUETP8M1_13TeV-pythia8	2995828	44.9	-
ZZ_TuneCUETP8M1_13TeV-pythia8	990064	15.4	-
QCD_HT300to500_TuneCUETP8M1_13TeV-madgraphMLM-pythia8	17035891	347700	-
QCD_HT500to700_TuneCUETP8M1_13TeV-madgraphMLM-pythia8	18929951	32100	-
QCD_HT700to1000_TuneCUETP8M1_13TeV-madgraphMLM-pythia8	15629253	6831	-
QCD_HT1000to1500_TuneCUETP8M1_13TeV-madgraphMLM-pythia8	4767100	1207	-
QCD_HT1500to2000_TuneCUETP8M1_13TeV-madgraphMLM-pythia8	3970819	119.9	-
QCD_HT2000toInf_TuneCUETP8M1_13TeV-madgraphMLM-pythia8	1991645	25.24	-

Table 1: MC samples used for the analysis. All event counts are taken from [13] or the CMS DAS tool[14], as appropriate. Cross sections for single top are taken from [15], with the reduced tW-channel cross sections for the NoFullyHadronicDecays sample given in [16], with the branching ratio for lepton decays taken from [17]. Cross sections and K-factors for W+jets are taken from [18]. Drell-Yan cross section is taken from [19]. Cross sections for QCD are taken from [20].

105 3 Object reconstruction

106 All objects used in the analysis are reconstructed with the particle flow (PF) algorithm for event
 107 reconstruction. All PF candidates are identified as an electron, a muon, a photon, a charged
 108 hadron, or a neutral hadron. Charged hadrons that are associated with a pileup vertex are clas-
 109 sified as pileup candidates. These are ignored in the subsequent object reconstruction through
 110 the Charged Hadron Subtraction (CHS) scheme. Electron and muon objects are first identified
 111 from corresponding electron/muon PF candidates. Next, jet clustering is performed on all PF
 112 candidates that are not classified as pileup candidates. The jet clustering does not exclude the
 113 electron and muon PF candidates, even if these have already been assigned to electron/muon
 114 objects. A dedicated overlap removal is therefore applied at analysis level to avoid double-
 115 counting.

116 3.1 Primary vertex

117 Primary vertices are reconstructed by clustering tracks in the event. Each vertex candidate is
 118 required to be within 2 cm of the beam spot in the transverse direction, and within 24 cm along
 119 the beam axis. The weighted number of tracks used to reconstruct the vertex, referred to as the
 120 number of degrees of freedom, is required to be greater than 4. The vertex candidate with the
 121 highest $\sum_{\text{tracks}} p_T^2$ is defined as the event's primary vertex.

122 3.2 Muons

123 Muons are reconstructed from tracks in the inner silicon tracking detectors and from hits in the
 124 muon stations. Signal muons selected for the analysis are required to fulfill the medium iden-
 125 tification (ID) recommended by the Muon POG [21], corresponding to the following selection
 126 criteria:

- 127 • Loose muon ID
- 128 • Fraction of valid tracker hits > 0.8
- 129 • Satisfy either of the following:
 - 130 1. Good Global muon
 - 131 • Global muon
 - 132 • Normalized global-track chisquare < 3
 - 133 • Tracker-standalone position match < 12
 - 134 • Chisquare of kink-finder < 20
 - 135 • Segment compatibility > 0.303
 - 136 2. Tight segment compatibility
 - 137 • Segment compatibility > 0.451

138 Muons are further required to have $p_T > 50 \text{ GeV}$ and $|\eta| < 2.1$. Signal muons are also re-
 139 quired to be isolated according to the mini isolation algorithm [22], which is described fully in
 140 Appendix A. A cut of $\text{miniIso} < 0.1$ is used.

141 Veto muons are also required to fulfill the medium ID requirement and to have $p_T > 50 \text{ GeV}$
 142 and $|\eta| < 2.1$. No isolation requirement is applied to veto muons.

143 Studies were made of the most optimal ID and isolation requirements for the muon using 2015
 144 data. Comparisons were made between mini isolation and 2D separation, as well as tight
 145 vs. medium ID. 2D separation requires the muon to be separated from the closest jet, using a
 146 logical OR of a lower cut on $\Delta R(\ell, j)$ and a lower cut on the component of the muon momentum

orthogonal to the axis of the jet. Multiple working points were considered for each version of the isolation. These studies are shown in Appendix A.

Scale factors are applied to account for differences in modeling the muon reconstruction, ID, and trigger efficiencies. The muon reconstruction and ID SFs are provided by the Muon POG [23] as a function of muon η and p_T . The scale factors for the muon mini isolation are evaluated and correspond to $0.999 - 1.000 \pm 0.001$ in the **kinematic region** of this analysis, i.e. consistent with 1. The muon trigger SF is measured through a tag and probe method in a data control sample. The trigger SF measurement is described in detail in Appendix B.

3.3 Electrons

Electrons are reconstructed from **clusters in the ECAL** combined with a **track in the inner tracker**. **Signal electrons** selected for the analysis must fulfill the **tight ID** working point recommended by the EGamma POG [24], with the exception of the **relative isolation requirement**. Different cut values are used for barrel electrons as opposed to endcap. This corresponds to the following criteria:

- η width of the electron supercluster < 0.00998 (barrel), 0.0292 (endcap)
- η separation between the track and supercluster < 0.00308 (barrel), 0.00605 (endcap)
- ϕ separation between the track and supercluster < 0.0816 (barrel), 0.0394 (endcap)
- Hadronic energy fraction < 0.0414 (barrel), 0.0641 (endcap)
- $|1/E - 1/p| < 0.0129$
- Track impact parameter with respect to the primary vertex < 0.05 (barrel), 0.10 (endcap)
- Track separation from the primary vertex along the beam direction < 0.10 (barrel), 0.20 (endcap)
- ≤ 1 expected missing hits in the inner tracker
- Veto on electrons from photon conversions

Electrons are required to have $p_T > 50$ GeV and $|\eta| < 2.1$. Furthermore, **signal electrons** are required to have **miniIso** < 0.1 .

Veto electrons are required to **fulfill** the **medium ID requirement**, which consists of looser cuts on the same quantities listed above. Veto electrons are also required to have $p_T > 50$ GeV and $|\eta| < 2.1$. **No isolation requirement** is applied to veto electrons.

As with muons, studies were made to optimize the electron definition. These studies are also included in Appendix A.

Scale factors are applied to account for differences in modeling the electron ID and isolation efficiencies. Since the electron ID defintion does not include the relative isolation requirement used by the EGamma POG, the ID and isolation SFs computed by the SUS PAG are used [25]. The electron trigger SF is calculated a tag-and-probe method in a data control sample. The trigger efficiency measurement is described in detail in Appendix B.

3.4 Jets

Jets are reconstructed by clustering **particle-flow** objects using the **anti- k_t** algorithm. Two different jet collections are considered in this analysis. **Small- R** jets are clustered using a distance parameter of $R = 0.4$ and **large- R** jets using a distance parameter of $R = 0.8$. The CHS scheme,

188 mentioned above, is applied to both to exclude charged hadrons originating from pileup in
 189 interactions. Jets are required to fulfill the loose jet ID working point recommended by the JME
 190 POG [26]. This corresponds to the following criteria:

- 191 • Neutral hadron fraction < 0.99
- 192 • Neutral EM fraction < 0.99
- 193 • Charged EM fraction < 0.99
- 194 • Charged hadron fraction > 0
- 195 • Charged multiplicity > 0
- 196 • At least one constituent

197 Small- R jets are required to have $p_T > 50$ GeV and $|\eta| < 2.4$. Large- R jets are required to have
 198 $p_T > 400$ GeV and $|\eta| < 2.4$.

199 Since the PF candidates ultimately identified as muons or electrons are initially used in the jet
 200 clustering, the leptons must be cleaned from the jets at analysis level to avoid double-counting.
 201 Leptons which were included in the jet clustering are identified by matching the PF-references
 202 of the lepton to the PF-references of the jet constituents. Muons (electrons) which satisfy the
 203 respective medium ID requirement, $p_T > 50$ GeV and $|\eta| > 2.1$ (2.5) are cleaned from the jets.
 204 No reclustering is done after the jet cleaning.

????

205 Jet energy corrections are applied to jets in both data and simulation. L1FastJet, L2Relative,
 206 and L3Absolute corrections are applied to both data and simulation. An additional set of JECs
 207 (L2L3Residual) is applied only in data events to account for residual differences between data
 208 and simulation. The 'AK4PFchs' corrections are used for small- R jets and 'AK8PFchs' correc-
 209 tions for large- R jets. The latest JEC provided by the JME POG for 2016 data is employed
 210 (Summer16_23Sep2016V3). The JEC are applied to the cleaned jet.

211 The jet energy resolution (JER) in data is worse than predicted in simulation. For jets matched
 212 to a gen jet, the p_T is rescaled according to the prescription in [27] to match the resolution seen
 213 in data. For jets not matched to a gen jet, the p_T is smeared according to the same prescription.
 214 The JER smearing is applied after the jet energy corrections.

215 3.4.1 b tagging

216 AK4 jets are considered to be b-tagged if they pass the medium WP for the Combined Sec-
 217 ondary Vertex v2 tagger, $\text{CSVv2} > 0.8484$ [28]. Data/MC scale factors are used to correct the
 218 b-tagging efficiencies in simulation to match those in data. These scale factors are provided as a
 219 function of the p_T , η , and flavor of the b-jet candidate. The scale factors are used in combination
 220 with MC b-tagging efficiencies to reweight events in MC. Only one AK4 jet (the leading jet in
 221 the same hemisphere as the lepton) is considered for b-tagging. Because of this, the prescrip-
 222 tion for reweighting events as described in [29] is simplified. The event weight is scaled by SF_i
 223 if the jet is b-tagged and $(1 - \epsilon_i)/(1 - \epsilon_i * SF_i)$ if not, where SF_i and ϵ_i are the scale factors and
 224 MC efficiencies, respectively. The MC efficiencies are computed by determining the fraction
 225 of events in the signal sample where the b-jet candidate passes the b-tagging requirement, for
 226 events which pass the preselection.

227 3.4.2 t tagging

228 AK8 jets are considered to be t-tagged if they have a SoftDrop mass between 105 and 220 GeV
 229 and $\tau_{32} < 0.81$. This is roughly based on the loose top-tagging WP described in [30], with
 230 the subjet b-tagging requirement removed to increase signal acceptance. The top-tagging SF is

231 extracted simultaneously with the integrated cross section, as described in Section 8.1.

232 **3.5 Missing transverse energy**

233 The missing transverse energy in the event is measured based on PF quantities. The measured
234 missing E_T is corrected using type 1 corrections, which propagate the jet energy corrections to
235 the missing E_T . In addition, the missing E_T is corrected at analysis level to account for the JER
236 smearing and lepton cleaning.

DRAFT

237 4 Event selection

238 The measurement is performed using the lepton+jets final state, identified through an electron
 239 or a muon, a small- R jet assumed to be produced by the decay of the b quark from the leptonic
 240 top quark decay, and a large- R jet corresponding to the hadronically decaying top quark.

241 Events in both the muon and electron channels are required to satisfy the following quality
 242 cuts:

- 243 • At least one good primary vertex
- 244 • Event passes MET filter as defined for 80X in [31]:
 - 245 • HBHENoiseFilter
 - 246 • HBHENoiseIsoFilter
 - 247 • globalTightHalo2016Filter
 - 248 • EcalDeadCellTriggerPrimitiveFilter
 - 249 • goodVertices
 - 250 • eeBadScFilter

251 Events in the muon channel are required to pass a preselection defined as follows:

- 252 • Exactly one signal muon
- 253 • 0 additional veto leptons
- 254 • At least one leptonic jet
 - 255 • AK4 jet in the same hemisphere as the electron, $0.3 < \Delta R(\ell, \text{jet}) < \pi/2$
- 256 • At least one hadronic jet
 - 257 • AK8 jet in the opposite hemisphere from the electron, $\Delta R(\ell, \text{jet}) > \pi/2$
- 258 • Missing $E_T > 35$ GeV

259 Events in the electron channel are required to pass a preselection defined as follows:

- 260 • Exactly one signal electron
- 261 • 0 additional veto leptons
- 262 • At least one leptonic jet
 - 263 • AK4 jet in the same hemisphere as the electron, $0.3 < \Delta R(\ell, \text{jet}) < \pi/2$
- 264 • At least one hadronic jet
 - 265 • AK8 jet in the opposite hemisphere from the electron, $\Delta R(\ell, \text{jet}) > \pi/2$
- 266 • Missing $E_T > 50$ GeV
- 267 • 'Triangular' cut
 - 268 • $|\Delta\phi(e, E_T^{\text{miss}}) - 1.5| < 1.5 * \frac{E_T^{\text{miss}}}{110\text{GeV}}$
 - 269 • $|\Delta\phi(\text{lead jet}, E_T^{\text{miss}}) - 1.5| < 1.5 * \frac{E_T^{\text{miss}}}{110\text{GeV}}$

270 The higher E_T^{miss} cut and triangular cut in the electron channel are applied to further reduce
 271 QCD multijet background. The triangular cut is included to ensure that the missing transverse
 272 momentum vector does not point along the transverse direction of either the electron or the
 273 leading jet.

274 The preselection is the result of optimization studies shown in Appendix C.

275 Events which pass the preselection are further categorized by whether the b(t)-jet candidate

276 passes or fails the b(t)-tagging criteria. The b(t)-jet candidate is the highest- p_T leptonic(hadronic)
277 jet in the event. Events are divided into the following exclusive kinematic regions:

- 278 • 0 top-tag (0t): t-jet candidate fails the top-tagging requirement
- 279 • 1 top-tag 0 b-tag (1t0b): t-jet candidate passes the top-tagging requirement, but b-jet
280 candidate fails the b-tagging requirement
- 281 • 1 top-tag 1 b-tag (1t1b): both t-jet candidate and b-jet candidate pass their respective
282 tagging requirement

283 The kinematic regions are designed to have different admixtures of signal and background,
284 with the 0t region being the most background dominated and the 1t1b region being the most
285 signal dominated. These kinematic regions are used in the simultaneous likelihood fit to extract
286 the background normalizations and top-tagging SF.

DRAFT

287 **5 Background estimation**

288 Several sources of background may contribute to the final state and are considered in this
 289 measurement. The main backgrounds considered are non-signal $t\bar{t}$, single top quark, $W+jets$,
 290 $Z+jets$, diboson, and QCD multijet. Non-signal $t\bar{t}$ comprises dileptonic and hadronic final
 291 states, in addition to $\tau+jets$ where the tau does not decay to a muon or electron. Single top
 292 quark background is comprised of tW -, t -, and s -channel production of a single top or antitop.
 293 $W+jets$ refers to $W \rightarrow \ell\nu$ decays with associated jets. $Z+jets$ refers to $Z \rightarrow \ell\ell$ decays with as-
 294 sociated jets. Diboson includes WW , WZ , and ZZ production. QCD multijet contributes due to
 295 its high cross section, despite quality cuts on the hard lepton.

296 Non-signal $t\bar{t}$, single top, $W+jets$, $Z+jets$, and diboson backgrounds are modeled using MC.
 297 The same MC sample is used to model both non-signal and signal $t\bar{t}$, through applying or
 298 inverting a semileptonic filter at parton-level. The $W+jets$ sample is split into multiple bins in
 299 H_T to improve statistics, particularly in the high-energy tail.

300 QCD multijet background is modeled using a data-driven technique described in section 5.1
 301 below. A QCD MC sample is used to cross-check the data sideband used to extract the QCD
 302 shape. As with $W+jets$, the QCD MC sample is binned in H_T to improve statistics in the high-
 303 energy tail.

304 All MC samples used in the analysis are described in Table 1.

305 **5.1 Data-driven QCD**

306 QCD multijet background is modelled by looking at data in a sideband dominated by QCD.
 307 The sideband is defined by inverting the mini isolation requirement on the lepton, requiring
 308 instead that events contain exactly one lepton with $0.1 < \text{minIso} < 0.2$. The electron ID re-
 309 quirement is also relaxed from Tight to Medium.

310 The number of data events in the sideband, as well as predictions from each MC, are given in
 311 Table 2. QCD event counts are provided both as the prediction from QCD MC and as a data-
 312 driven number (data - MC prediction from other processes). The two measures of QCD agree
 313 within a reasonable level. The contribution from non-QCD processes in the sideband is 28% in
 314 the muon channel and 32% in the electron channel.

Sample	Event counts, $\mu+jets$	Event counts, $e+jets$
$t\bar{t}$ (signal)	426 ± 13	176 ± 8
$t\bar{t}$ (non-semilep)	211 ± 9	86 ± 6
Single top	51 ± 2	20 ± 1
$W+jets$	320 ± 4	144 ± 3
$Z+jets$	30 ± 12	15 ± 9
Diboson	5 ± 2	4 ± 1
QCD (MC)	3390 ± 196	1253 ± 118
QCD (Data - bkg)	2693	924
Data	3736	1369
% non-QCD in sideband	28%	33%

Table 2: Estimates of the signal and background yields, together with the observed number of events in data, in the QCD-dominated sideband defined by inverting the isolation requirement on the lepton. Uncertainties are statistical only.

315 The data-driven QCD is modelled by subtracting the predicted contributions from $t\bar{t}$, single

316 top, W+jets, Z+jets, and diboson from the observed data in the sideband. To verify that these
317 shapes appropriately describe the QCD kinematics in the signal region, the data-driven QCD
318 distributions are compared to the predictions from QCD MC in the signal region and in the
319 sideband. Comparing data-driven QCD to QCD MC in the sideband shows whether the two
320 methods of modelling QCD are consistent. Comparing the QCD distributions in the sideband
321 to those in the signal region shows whether there are any systematic shifts in the QCD kine-
322 matics between the two regions. If all three shapes are reasonably consistent, the data-driven
323 QCD is expected to accurately model the QCD kinematics in the signal region.

324 Selected comparisons of the QCD kinematics between the data-driven sideband, MC sideband,
325 and MC signal region are shown in Figure 1. For reference, additional comparisons are pro-
326 vided in Appendix D. In general, there is reasonable agreement between the MC prediction
327 for QCD in the sideband and the data-driven QCD estimate in the sideband. However, a sys-
328 tematic shift is observed in the electron η and p_T distributions where electrons in the signal
329 region are more likely to have high p_T and $|\eta|$ than electrons in the sideband region. The differ-
330 ences per-bin between either of the MC predictions and the data-driven estimate are generally
331 within 50%. The source of the disagreement may be traced to the mini isolation definition. As
332 described in Appendix A, the electron mini isolation cone is hollow in the endcap but not in
333 the barrel. In QCD events, objects reconstructed as electrons are most likely misidentified jets.
334 At high p_T , a light jet faking an electron could become narrow enough to fall within the inner
335 radius of the isolation cone in the endcap. The light jet would then appear to be fully isolated.
336 Thus, high p_T fake electrons in the endcap will contribute to the signal region rather than the
337 sideband.

general agreement and why we have discrepancies in the electron channel

338 A zeroth-order estimate of the QCD normalization is taken from the QCD MC prediction in the
339 signal region.

DRAFT

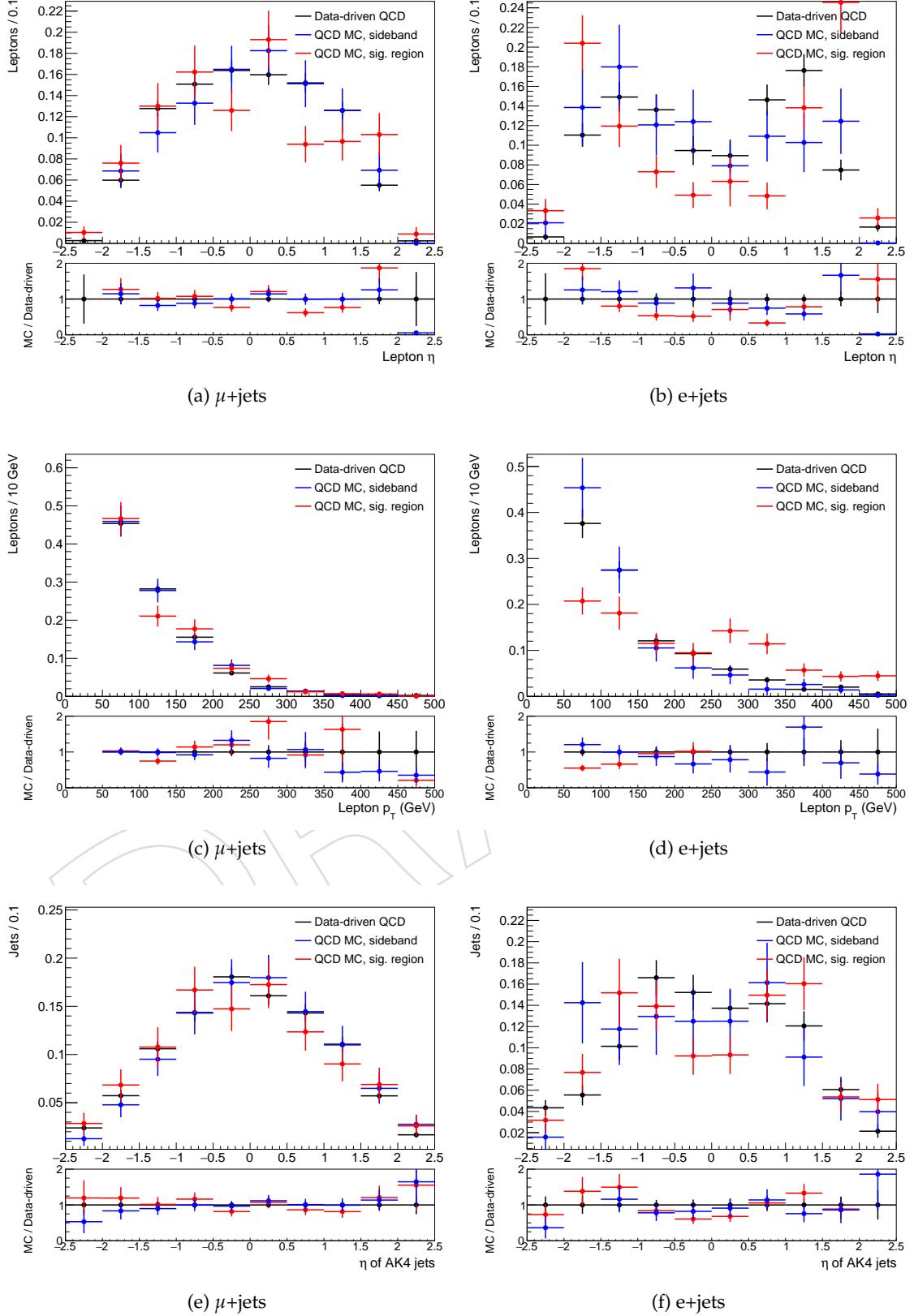


Figure 1: Comparisons of QCD kinematics as predicted by MC in the signal region, MC in the sideband region, and the data-driven technique. Comparisons for the muon channel are shown on the left and electron channel on the right. The largest difference is between the signal and sideband electron η and p_T distributions.

340 6 Comparison of data and MC prediction

341 In this section, the data observation is compared to the $t\bar{t}$ signal and the background predictions.
 342 Event yields and kinematic distributions are shown using the background predictions as
 343 described in Section 5. Both event yields and kinematic distributions are presented for the 0t,
 344 1t0b, and 1t1b kinematic regions. The pre-fit normalizations are not used for the final result,
 345 but these comparisons serve as a check of the modeling of the kinematic variables which are
 346 used for the fit.

347 6.1 Event yields

348 The pre-fit event yields in each kinematic region are listed in Table 3 for the muon channel and
 349 in Table 4 for the electron channel. The quoted uncertainties are statistical only.

Sample	Preselection	Number of events (μ +jets channel)		
		0t	1t0b	1t1b
$t\bar{t}$ (signal)	26491 ± 102	17290 ± 118	4819 ± 72	4382 ± 41
$t\bar{t}$ (non-semilep)	4748 ± 43	4145 ± 46	318 ± 19	285 ± 10
Single top	3257 ± 16	2871 ± 17	249 ± 6	137 ± 3
W+jets	30861 ± 43	27155 ± 46	3566 ± 15	139 ± 3
Z+jets	3070 ± 122	2726 ± 128	318 ± 42	26 ± 11
Diboson	595 ± 19	552 ± 19	40 ± 5	3 ± 1
QCD	2992 ± 72	2741 ± 69	197 ± 20	53 ± 8
Total	72014 ± 186	57481 ± 200	9508 ± 90	5025 ± 45
Data	60672	49137	7348	4187

Table 3: Estimates of the signal and background yields, together with the observed number of events in data, in each kinematic region for the muon channel. The quoted uncertainties are statistical only.

Sample	Preselection	Number of events (e+jets channel)		
		0t	1t0b	1t1b
$t\bar{t}$ (signal)	16492 ± 80	10417 ± 93	3168 ± 59	2907 ± 33
$t\bar{t}$ (non-semilep)	3475 ± 37	3036 ± 39	249 ± 16	190 ± 8
Single top	2203 ± 13	1944 ± 14	165 ± 5	93 ± 3
W+jets	18306 ± 36	16083 ± 37	2143 ± 12	80 ± 2
Z+jets	1259 ± 77	1075 ± 83	167 ± 31	16 ± 9
Diboson	391 ± 15	361 ± 16	27 ± 4	2 ± 1
QCD	2630 ± 112	2345 ± 103	263 ± 45	22 ± 7
Total	44755 ± 167	35261 ± 172	6183 ± 83	3312 ± 36
Data	39313	31559	4801	2953

Table 4: Estimates of the signal and background yields, together with the observed number of events in data, in each kinematic region for the electron channel. The quoted uncertainties are statistical only.

350 6.2 Kinematic distributions

351 A variety of kinematic distributions are studied to compare data to the signal and background
 352 predictions. Selected distributions, namely those used in the combined likelihood fit and those

used to extract the differential cross section, are shown in this section. Additional control plots are shown in Appendix E. Separate kinematic distributions are shown for events in the 0t, 1t0b, and 1t1b regions. All figures shown in this section use the pre-fit normalizations as given in Tables 3 and 4. The 0t region is dominated by W+jets, with significant contributions from QCD and t̄ signal. The 1t0b region is dominated by t̄ signal with large contributions from W+jets and some from QCD. The 1t1b region is fully dominated by the t̄ signal, with only small background contributions from QCD, non-semileptonic t̄, W+jets, and single top.

The kinematic distributions used in the likelihood fit are the AK4 jet η in the 0t and 1t0b regions and the AK8 jet SoftDrop mass in the 1t1b region. The a priori versions of these distributions are shown in Figures 2 and 3 for the muon and electron channels, respectively. The Monte Carlo predictions are seen to overestimate the data in terms of the yield. Figures 4 and 6 show the p_T spectrum for the hadronic top candidate for each of the three regions for the muon and electron channels. Finally, Figures 5 and 7 show the rapidity distribution for the hadronic top candidate in each of the three regions for the muon and electron channels.

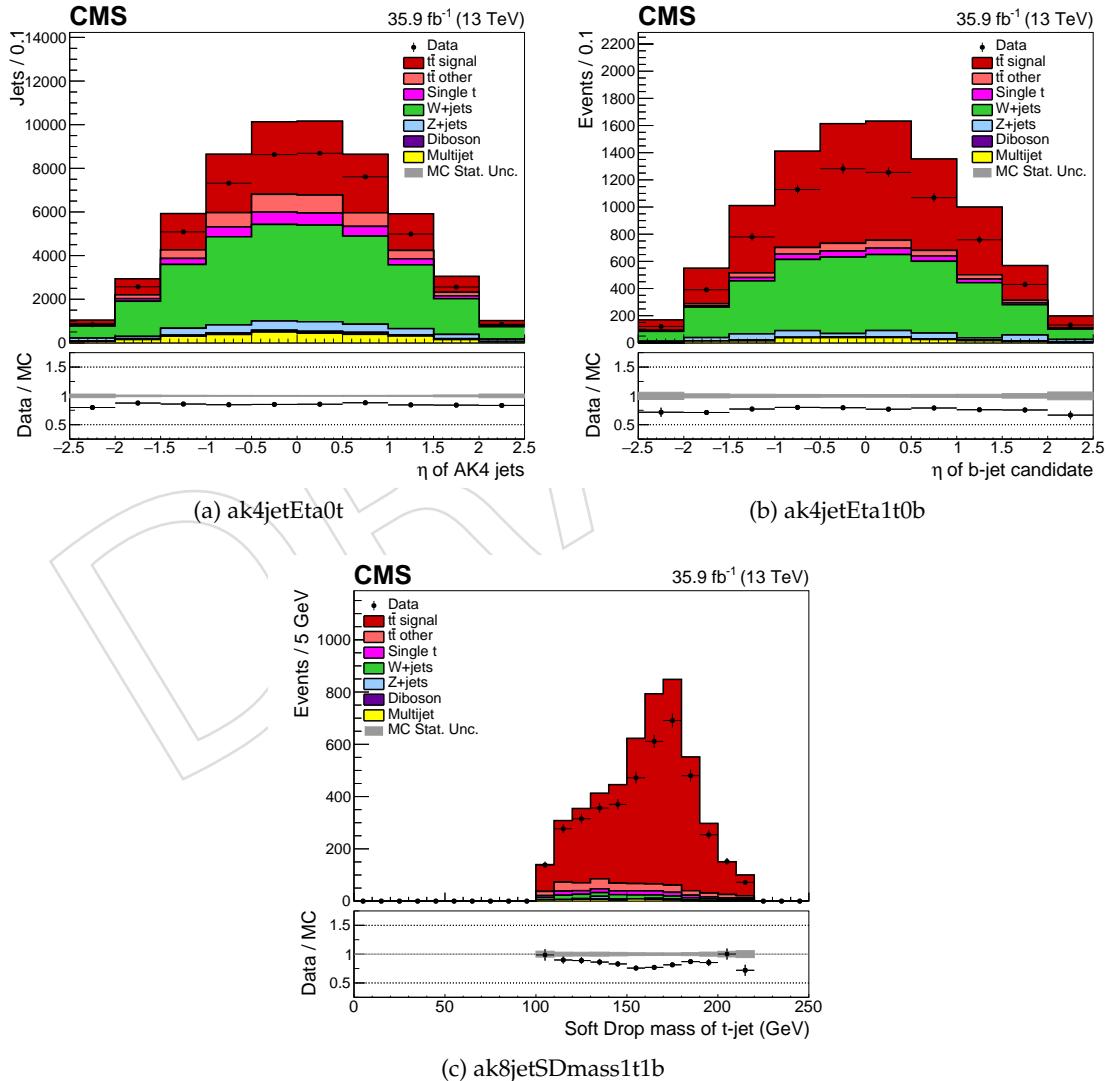


Figure 2: A priori kinematic distributions used in the likelihood fit, for events in the muon channel.

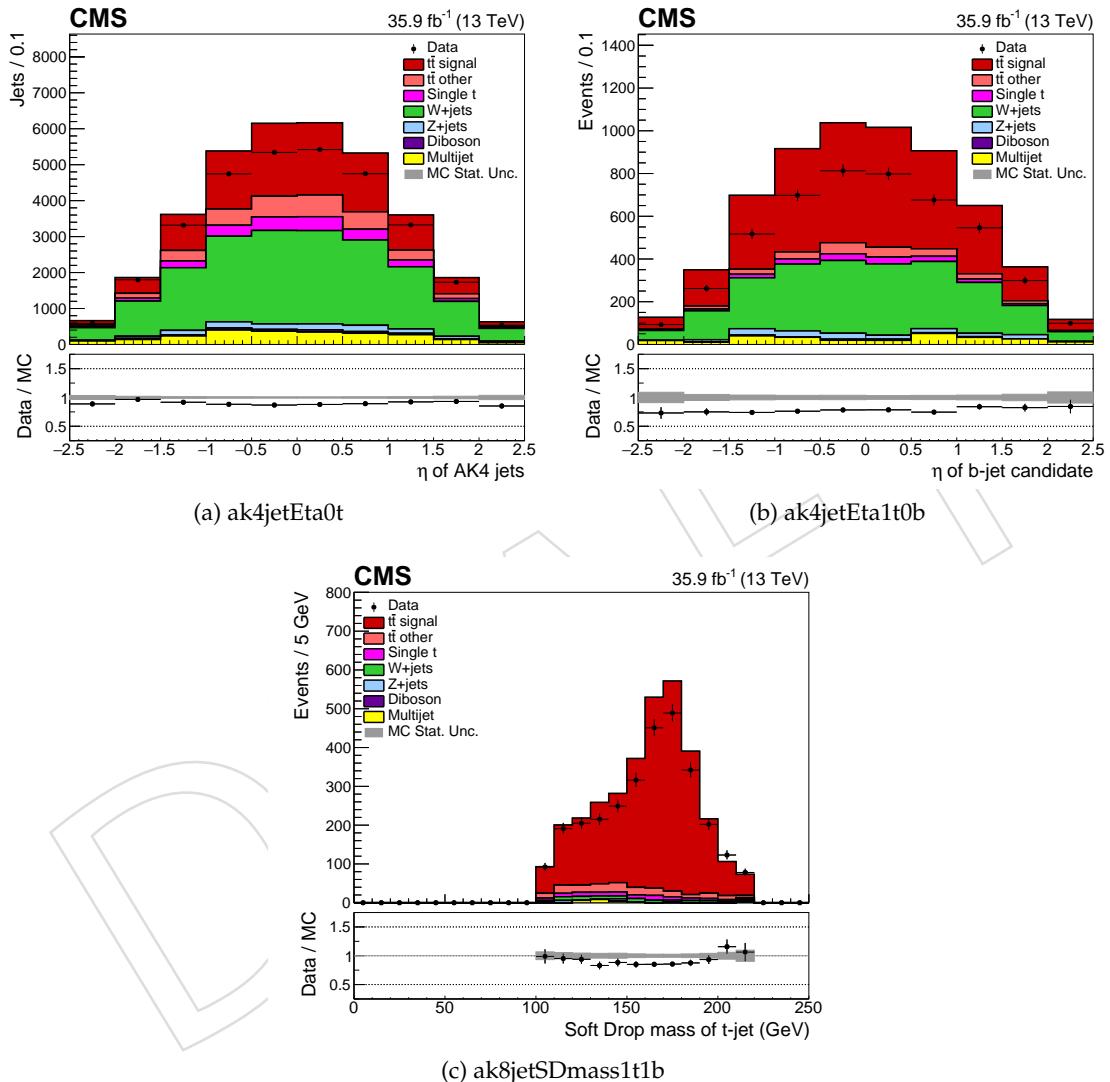


Figure 3: A priori kinematic distributions used in the likelihood fit, for events in the electron channel.

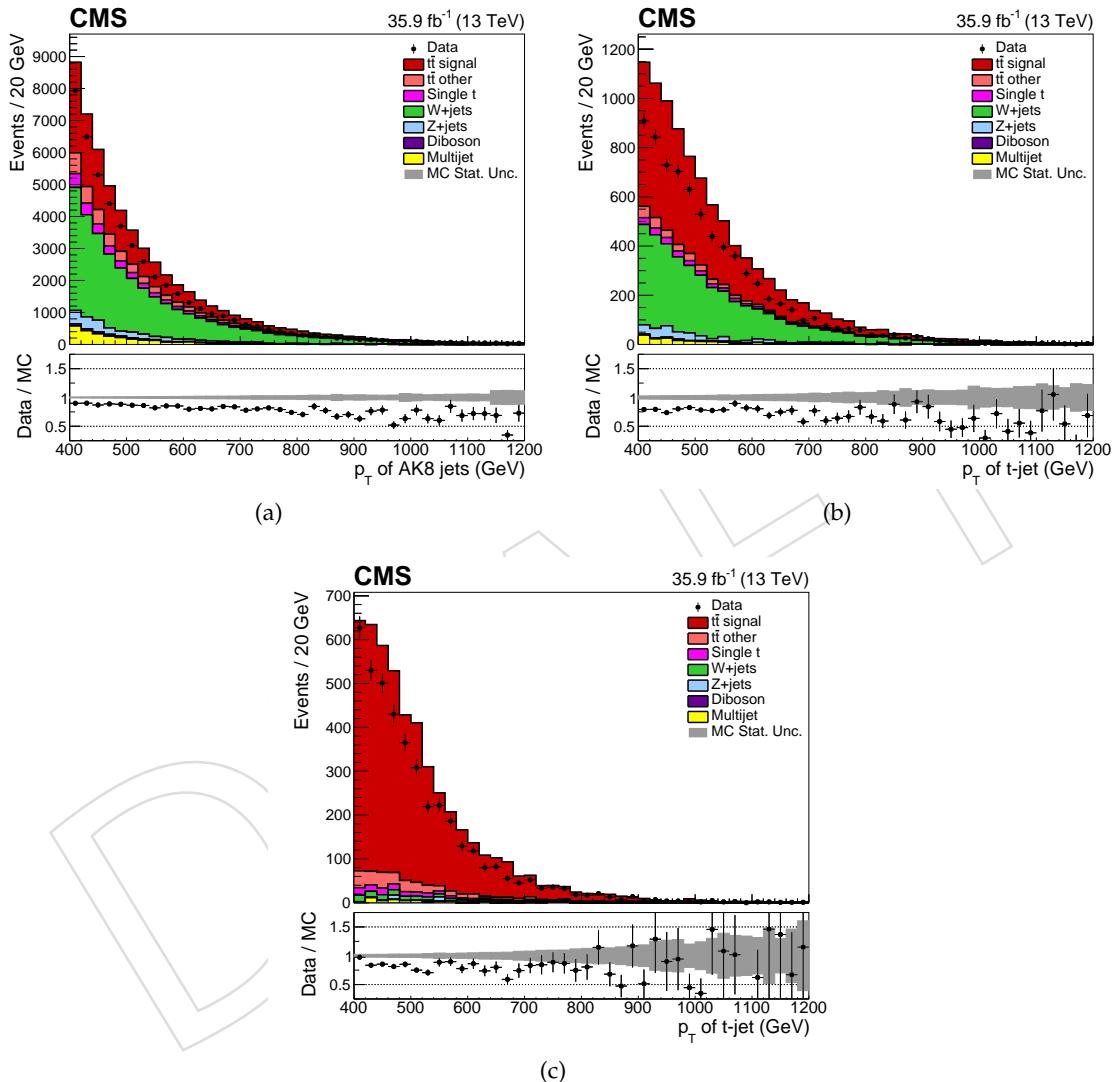


Figure 4: p_T of the hadronic top candidate in the $0t$, $1t0b$, and $1t1b$ regions, for events in the muon channel

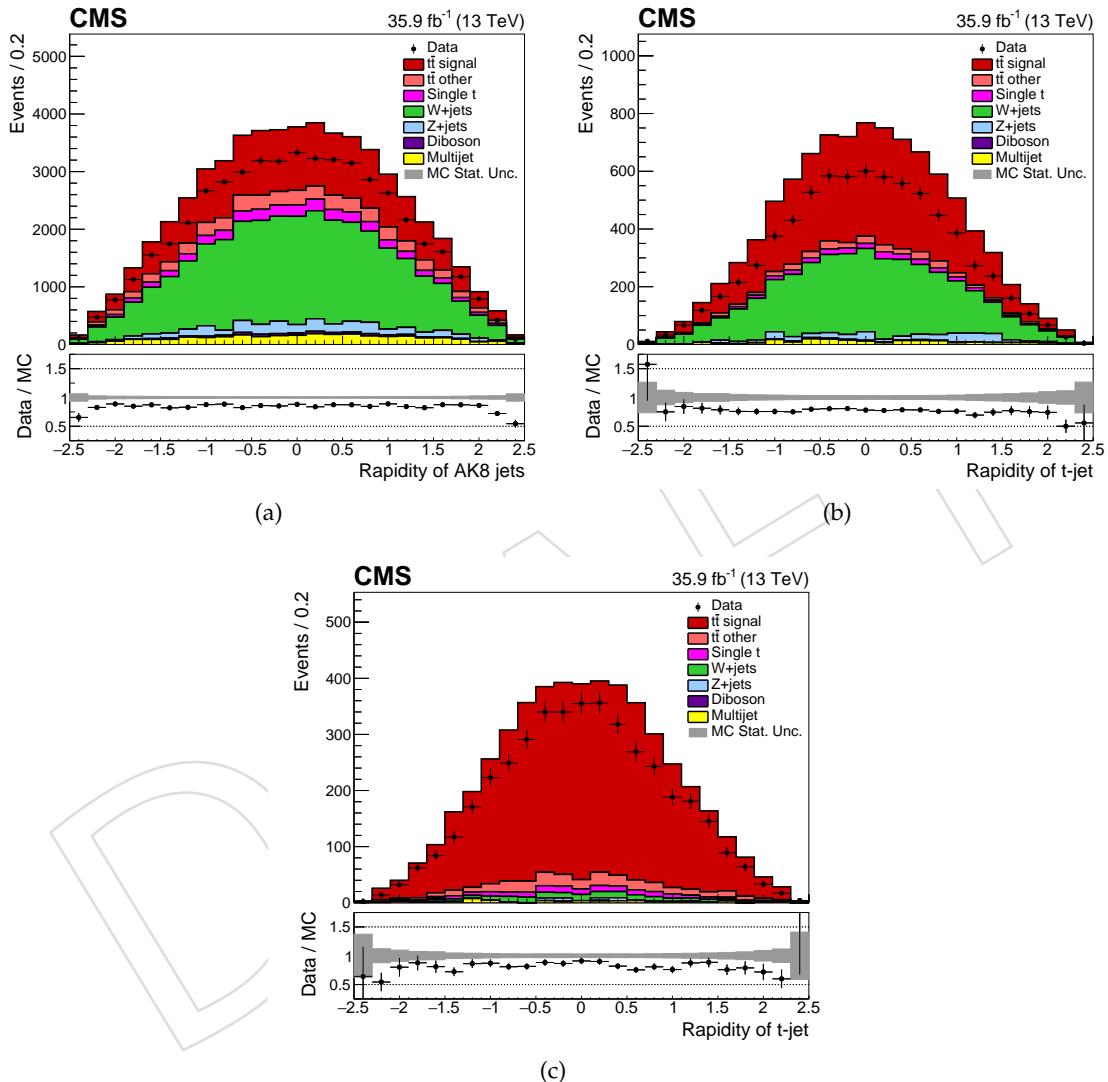


Figure 5: Rapidity of the hadronic top candidate in the 0t, 1t0b, and 1t1b regions, for events in the muon channel

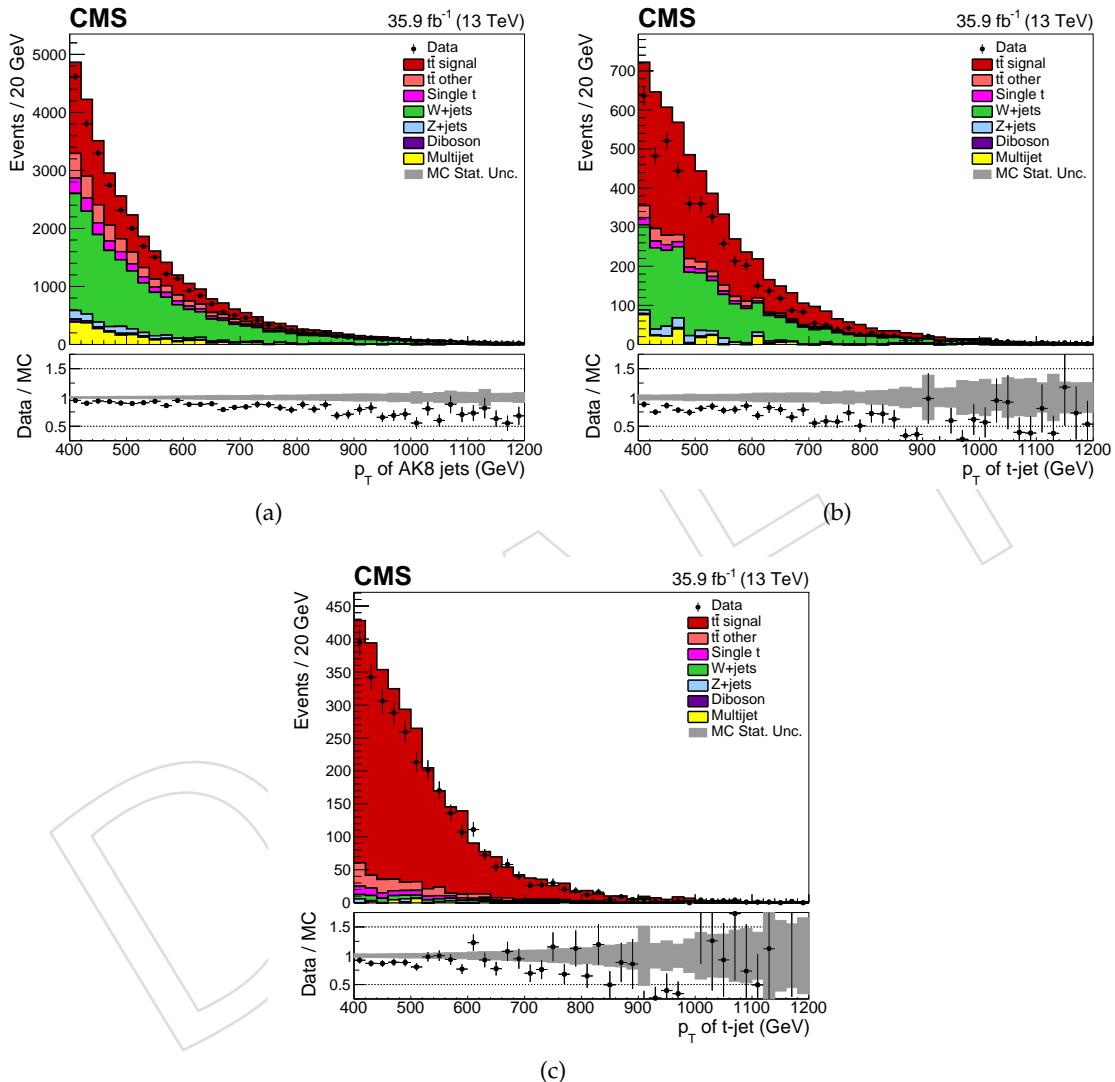


Figure 6: p_T of the hadronic top candidate in the 0t, 1t0b, and 1t1b regions, for events in the electron channel

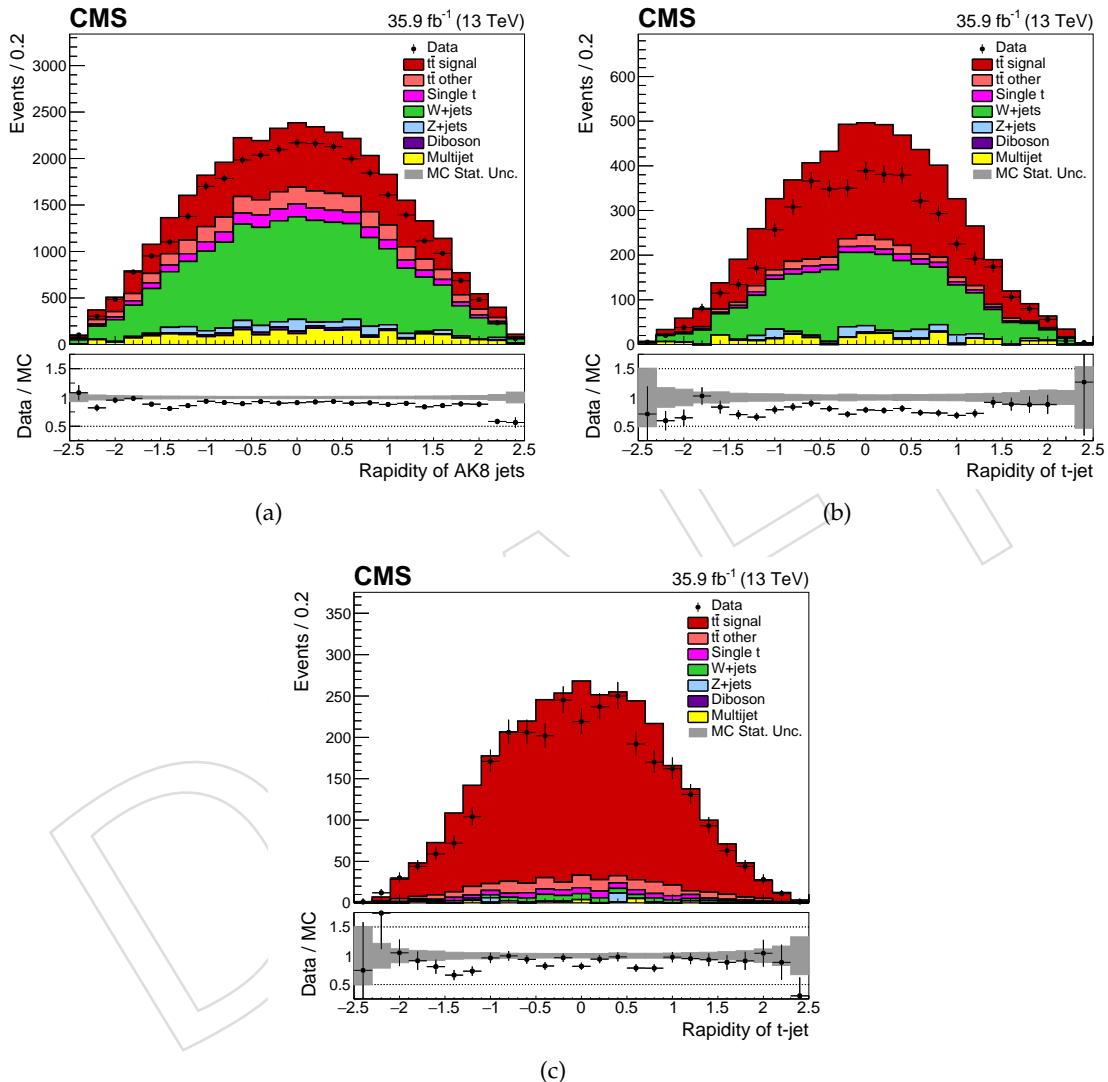


Figure 7: Rapidity of the hadronic top candidate in the 0t, 1t0b, and 1t1b regions, for events in the electron channel

367 7 Systematic uncertainties

368 Many sources of systematic uncertainty can impact the cross section measurement. These may
369 affect the background normalizations, the shape of kinematic variables, or the signal accep-
370 tance. The uncertainties may be experimental or theoretical. Where possible, uncertainties are
371 used which are provided by the appropriate object performance groups. Many experimental
372 uncertainties are additionally treated as nuisance parameters in the combined likelihood fit
373 described in Section 8.1.

374 7.1 Experimental uncertainties

375 A number of experimental uncertainties affect the cross section measurement. Among these,
376 the dominant uncertainties come from the b-tagging and top-tagging efficiencies. The various
377 sources of experimental uncertainty are listed below, in roughly decreasing order of impact.

378 The top-tagging efficiency has been evaluated by the JME group, as a flat value and as a func-
379 tion of top jet p_T [30]. As this evaluation is carried out in a region very similar to our signal
380 region, we instead simultaneously extract the top-tagging scale factor with the signal strength
381 and background normalizations in the likelihood fit. This is described in Section 8.1. The input
382 to the fit is an ad-hoc flat scale factor of 1.0, with an uncertainty of 25%. Separate SFs are used
383 for the tagging of true top jets, applied to the $t\bar{t}$ and single top tW-channel samples, and the
384 mistag of fake top jets, applied to the t- and s-channel single top, W+jets, Z+jets, diboson, and
385 QCD multijet samples.

386 The b-tagging efficiency is evaluated by the BTV POG as a function of the b-jet p_T , η , and
387 flavour [28]. Events are weighted according to whether the b-jet candidate passes or fails the
388 b-tagging requirement, as described in 3.4.1. The b-tagging uncertainty is assessed by varying
389 the b-tagging SF within the provided uncertainties. The b-tagging SF for all flavours are varied
390 up or down simultaneously, effectively treating the tag and mistag SF as fully correlated.

391 The uncertainty on the jet energy scale is assessed as an uncertainty on the JEC provided by the
392 JME POG. The uncertainty σ is used to scale the jet four-vector by $(1 \pm \sigma)$. The same approach
393 is used for AK4 and AK8 jets.

394 The uncertainty on the jet energy resolution is assessed by varying the JER SF provided by
395 the JME POG within the provided uncertainty. The hybrid smearing described in [27] is then
396 performed with the new SF.

397 The pileup reweighting that is applied to the simulated samples assume a total inelastic cross
398 section of 69.2 mb. An associated systematic uncertainty is derived by applying different pileup
399 weights obtained using a total cross section scaled by $\pm 4.6\%$.

400 The performance of the lepton identification, reconstruction, trigger and isolation provides a
401 small source of uncertainty. The data/MC SF for the muon identification and reconstruction are
402 provided by the muon POG, while the trigger SF is calculated using a tag and probe technique.
403 The muon uncertainty is assessed by simultaneously varying these SF up or down within their
404 uncertainty. The data/MC SF for the electron track reconstruction, identification, and isolat-
405 tion are provided by the SUSY group using their tag-and-probe tool [25]. The electron trigger
406 efficiency is calculated using a separate tag and probe technique. The electron uncertainty is
407 assessed by simultaneously varying these values up or down within their uncertainty.

408 An ad-hoc a priori uncertainty of 30% is applied to the single top, W+jets, Z+jets, and diboson
409 rates. An additional uncertainty on the flavor composition of the W+jets sample is assessed
410 by dividing the sample into light and heavy flavor components using MC truth information.

411 These components are allowed to vary separately with 30% normalization uncertainty. A 50%
 412 uncertainty is assumed on the QCD rate, to reflect the combined uncertainty in the normaliza-
 413 tion and the extraction of the kinematic distributions from the sideband.

414 The uncertainty in the measurement of the integrated luminosity is 2.5% [32].

415 7.2 Theoretical uncertainties

416 A number of theoretical uncertainties affect the expected top quark kinematics, and the relation
 417 between the true top quark and reconstructed top jet. Most of these uncertainties correspond to
 418 parameters of the POWHEG+PYTHIA8 simulation used to generate $t\bar{t}$ events. Some uncertain-
 419 ties may be described by reweighting events in the nominal sample, while others are modeled
 420 by generating separate samples corresponding to different parameters in the simulation. The
 421 dedicated samples corresponding to certain theory systematic shifts are given in Table 5.

422 The uncertainty in the choice of renormalization and factorization scale is assessed by sep-
 423 arately varying the renormalization and factorization scales by a factor of [2.0,0.5]. The un-
 424 physical anticorrelated variations are discarded, yielding a total of 7 combinations of the renor-
 425 malization and factorization scales. The event weight is determined for each variation and the
 426 envelope is taken as the scale uncertainty, according to the prescription in [33]. These variations
 427 are applied using the event weights provided in the LHEEventInfo [34].

428 The signal is generated using the NNPDF30 PDF set. To model the PDF uncertainty, 100 MC
 429 replicas are drawn from the PDF probability distribution. The event weights corresponding to
 430 these MC replicas are stored in the LHEEventInfo. The standard deviation of these weights is
 431 taken as the PDF uncertainty.

432 The uncertainty in the ISR and FSR shower scales is assessed by varying α_S^{ISR} and α_S^{FSR} up
 433 and down by factors of 2 [35, 36]. The nominal values assumed are $\alpha_S^{ISR} = 0.1108^{+0.0145}_{-0.0142}$, as
 434 tuned in [36], and $\alpha_S^{FSR} = 0.1365$, as extracted from LEP, see again [36]. Separate samples are
 435 generated for the up and down ISR and FSR variations (i.e. for up, factor 2, and for down, factor
 436 1/2). Furthermore, the resulting FSR uncertainty from varying the FSR by factors of 2 and 1/2
 437 are re-scaled to $\sqrt{2}$ and $1/\sqrt{2}$, as recommended by the TOP group based on measurements of
 438 event shapes at LEP (see again [35]).

439 In the Powheg ME-PS matching scheme, the resummation damping factor h_{damp} is used to
 440 regulate high- p_T radiation. Uncertainties in h_{damp} are parameterized by generating separate
 441 samples with h_{damp} varied by $\pm 1\sigma$ (specifically, $h_{damp} = 1.58^{+0.66}_{-0.59} \times m_{top}$).

442 Uncertainties in the underlying event are parameterized by generating separate samples with
 443 the tune CUETP8M2T4 parameters varied by $\pm 1\sigma$ [35, 36].

444 Uncertainties in the color reconnection are assessed by generating a separate sample with a
 445 color reconnection model (MPI-based + QCD-inspired + gluon move) applied [37].

446 The uncertainty in the top mass is small and not added as a systematic uncertainty in the
 447 analysis. It is verified that the impact is small by comparing the nominal simulation with
 448 samples with top mass varied by ± 1 GeV, in terms of top p_T and rapidity shapes and the signal
 449 acceptance. The impact on the shape is basically zero, there is a 1-2% impact on the acceptance,
 450 which is small compared to the other theory uncertainties.

Systematic	Parameter	Variation	Sample	N_{events}
Shower Scales	α_S^{ISR}	Up	TT_TuneCUETP8M2T4_13TeV-powheg-isrup-pythia8	59033604
		Down	TT_TuneCUETP8M2T4_13TeV-powheg-isrdown-pythia8	58999580
	α_S^{FSR}	Up	TT_TuneCUETP8M2T4_13TeV-powheg-fsrup-pythia8	59230899
		Down	TT_TuneCUETP8M2T4_13TeV-powheg-fsrdown-pythia8	59306906
ME-PS matching	hdamp	Up	TT_hdampUP.TuneCUETP8M2T4_13TeV-powheg-pythia8	58858606
		Down	TT_hdampDOWN.TuneCUETP8M2T4_13TeV-powheg-pythia8	58163976
Underlying event	Tune CUETP8M2T4	Up	TT_TuneCUETP8M2T4up_13TeV-powheg-pythia8	58953660
		Down	TT_TuneCUETP8M2T4down_13TeV-powheg-pythia8	58338240
Color reconnection	Different model	On	TT_TuneCUETP8M2T4_erdON_13TeV-powheg-pythia8	59882210
Hadronization	Herwig++ vs. Pythia8		TT_TuneEE5C_13TeV-powheg-herwigpp	59174465

Table 5: MC samples used for theory systematic variations, following prescription in [35]

451 8 Kinematic Fit

452 The integrated signal strength is measured using a likelihood template fit to kinematic variables using the HiggsCombine framework [38]. Also the top-tagging SFs, lepton SFs, and background normalizations are treated as nuisance parameters, and simultaneously extracted with the cross section. These posterior background yields, top-tag SFs, and lepton SFs are later used in the unfolded measurement in extracting the differential cross section, described in Section 456 9.4.

458 8.1 Likelihood template fit

459 The 0t, 1t0b, and 1t1b exclusive kinematic regions are fit simultaneously in the likelihood fit, 460 normalizing each background component to the same cross section in all regions. The fit result 461 is expressed in terms of a multiplicative factor, the signal strength r , applied to the input $t\bar{t}$ cross 462 section. Different variables are used in each region to discriminate $t\bar{t}$ from background. The 463 AK4 jet η is used in the 0t and 1t0b regions while the AK8 jet SoftDrop mass is used in the 1t1b 464 region. These distributions were chosen to have good discrimination between $t\bar{t}$, W+Jets, and 465 QCD. Figures 2 and 3 show the kinematic distributions prior to the fit for the muon and electron 466 channels. The $t\bar{t}$ signal and background distributions are merged into a single distribution for 467 the purpose of the fit, essentially constraining the semileptonic branching ratio to be exactly 468 equal to that provided by MC.

469 Background normalizations and systematic uncertainties are treated as nuisance parameters in 470 the fit. The uncertainties due to lepton SFs, jet energy scale, jet energy resolution, b-tagging 471 efficiency, and top-tagging efficiency are treated as uncertainties on the shape of the input tem- 472 plates. Two nuisance parameters are used to describe the top-tagging uncertainty: a top-tag 473 SF nuisance parameter is applied to the $t\bar{t}$ and tW-channel single top, where we expect a true 474 top jet, while a top-mis-tag SF is applied to t- and s-channel single top, W+jets, Z+jets, diboson, 475 and QCD. The uncertainties on the single top, W+jets, Z+jets, diboson, and QCD rates are 476 treated as rate uncertainties on their respective samples. Separate rate uncertainties are applied 477 to the W+light and W+heavy flavor components of the W+jets sample. The event categories 478 which are fit are designed such that the top-tagging efficiency is constrained by the relative 479 population of events in the three categories. The varying admixtures of signal and background 480 between the different regions also offer more power to determine the background rates. The 481 measurement of the cross section is correlated with the various nuisance parameters, with the 482 strongest correlation being with the top-tagging efficiency, as expected.

483 The HiggsCombine framework is used to assign log-normal prior constraints for each nuisance 484 parameter. For the shape uncertainties, the nuisance parameter is used to interpolate 485 between the nominal kinematic distribution and distributions corresponding to 1σ variations 486 in the given uncertainty.

487 Additional uncertainties originate from difficulty in the theoretical modeling, as described in 488 Section 7.2. These theoretical uncertainties are evaluated independently from the likelihood fit.

489 The likelihood template fit is first performed separately for the individual e/ μ channels. To 490 combine the channels, the likelihood fit is also evaluated by fitting e/ μ +jets events together, 491 with all nuisance parameters generally constrained to be the same in both channels. The muon 492 and electron SF nuisance parameters are treated separately, as are the muon and electron QCD 493 rates.

494 The posterior kinematic distributions in the three fit categories are shown for the muon-only 495 channel in Figure 8 and for the electron-only fit in Figure 9. The posterior kinematic distribu-

496 tions for the combined fit in both the muon and electron channels are shown in Figures 10 and
 497 11, respectively.

498 The fitted values of the nuisance parameters and their correlations are shown for the muon
 499 channel in Figure 12, for the electron channel in Figure 13, and for the combined fit in Figure
 500 14. All nuisance parameters are a priori distributed according to a gaussian with mean 0 and
 501 standard deviation 1. Therefore, a nuisance parameter of 0 indicates no change to the input
 502 systematic central value; a nuisance parameter uncertainty of 1 indicates that the fit does not
 503 constrain the uncertainty.

504 In general, most of the posterior nuisance parameters show agreement with their a priori val-
 505 ues. The nuisance parameters most constrained by the fit are the top-tag and top-mis-tag SFs,
 506 as expected. A slight tension between the muon and electron channels causes more disagree-
 507 ment between the prior and posterior nuisance parameters in the combined fit. The largest
 508 (anti)correlations in the single-channel fits are between the signal strength and top-tag SF, top-
 509 tag SF and top-mis-tag SF, and W+light and W+HF normalizations. This is as expected. In the
 510 combined fit, there is also an anticorrelation between the signal strength and the single top nor-
 511 malization, and a correlation between the muon and electron SFs and the muon and elec-
 512 tron QCD rates. The latter correlations are expected, as the nuisance parameters have similar effects
 513 on the fit.

514 The posterior event counts which account for all posterior parameters are given in Table 6 for
 515 the fit in the muon channel. Table 7 gives the posterior event counts for the electron channel,
 516 and Table 8 the posterior event counts in both channels for the combined fit.

517 Figures 15 and 16 show the p_T spectrum for the hadronic top candidate for each of the three
 518 regions for the muon and electron channels. Finally, Figures 17 and 18 show the rapidity dis-
 519 tribution for the hadronic top candidate in each of the three regions for the muon and electron
 520 channels. These distributions use the posterior top-tag SF, lepton SF, and background normal-
 521 izations, but not the posterior JER, JEC, or b-tag SF.

522 Multiple cross-checks were performed to assess the fit stability. Details on these cross-checks
 523 may be found in Appendix F. As a result of these checks, an additional systematic uncertainty
 524 was assigned on the posterior lepton SF and background normalizations.

525 To account for a possible dependence of the top-tag SF on top jet p_T or $|\eta|$, fit cross-checks were
 526 performed dividing the kinematic regions into two bins of top jet p_T or $|\eta|$. The top-tag and
 527 mistag SFs were extracted separately in both regions, while the other nuisance parameters were
 528 treated as fully correlated. The difference between the binned and flat top-tag SFs was treated
 529 as a shape uncertainty on the top-tag SF. More information on these cross-checks can be found
 530 in Appendix G.

531 8.2 Interpretation of results

532 8.2.1 Posterior top-tagging scale factor

533 An important goal of the likelihood template fit is to constrain the top-tagging scale factor and
 534 its uncertainty by simultaneously fitting with the signal cross section. The posterior nuisance
 535 parameters for the top-tag and top-mis-tag scale factors are given in Table 9 for the muon
 536 channel, electron channel, and combined fits. There is some tension between the results in the
 537 electron and muon channels, but the results bracket and agree with the combined result. The
 538 posterior nuisance parameter from the combined fit is used to determine the posterior top-tag
 539 SF and top-mis-tag SF.

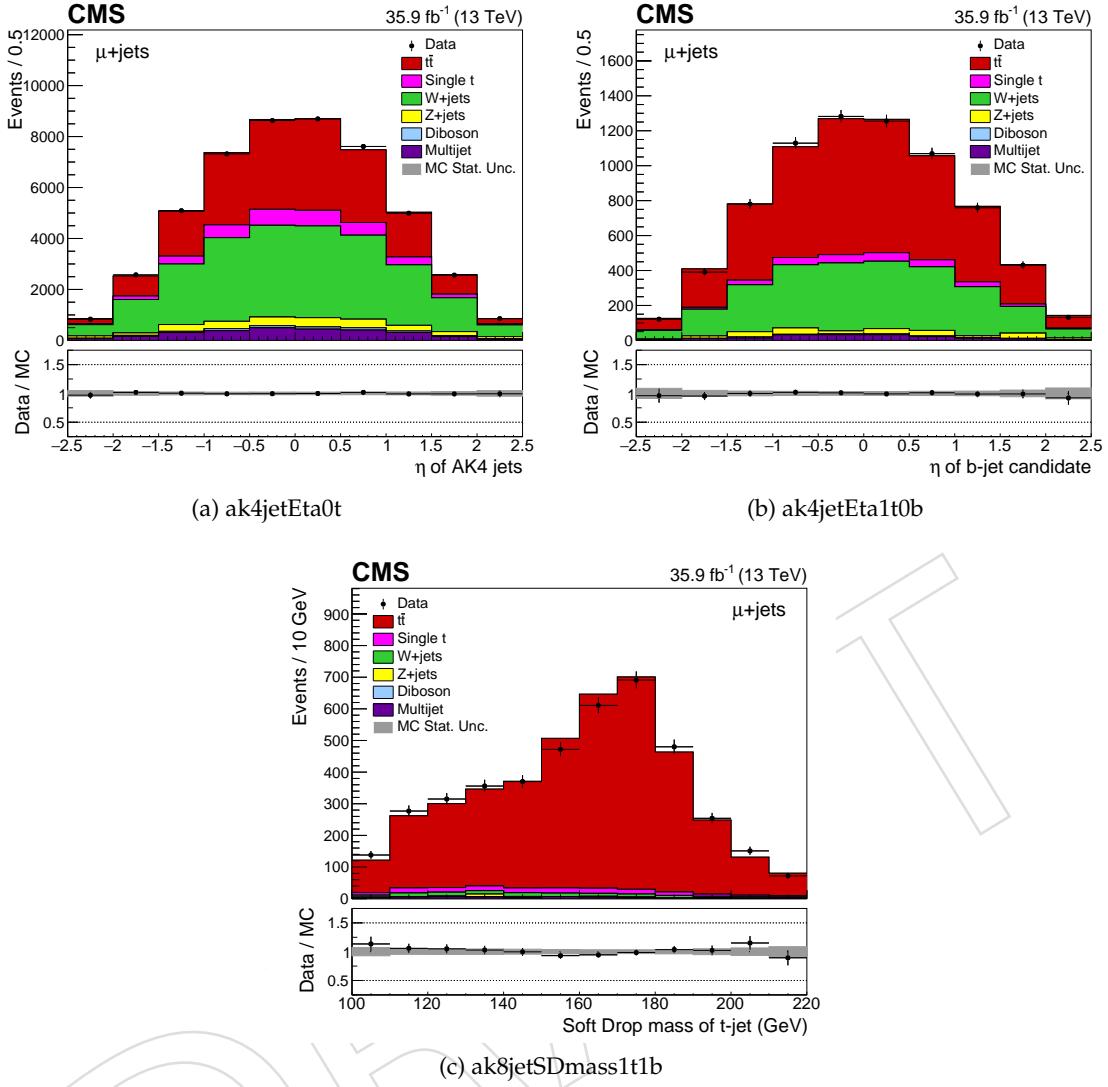


Figure 8: Post-fit kinematic distributions for the fit in the muon channel. The posterior agreement between data and MC is within uncertainties.

Sample	Number of events ($\mu+$ jets channel)		
	0t	1t0b	1t1b
tt	18300 ± 1744	4226 ± 189	3865 ± 94
SingleTop	3209 ± 719	263 ± 81	144 ± 39
WJets	22070 ± 2919	2456 ± 336	104 ± 21
ZJets	2249 ± 597	216 ± 68	18 ± 10
Diboson	550 ± 176	34 ± 12	2 ± 1
QCD	2761 ± 1219	166 ± 79	46 ± 25
total	49140 ± 3735	7360 ± 408	4179 ± 107
Data	49137	7348	4187

Table 6: Post-fit signal and background normalizations in the 0t, 1t0b, and 1t1b regions, together with observation in data, for the fit in the muon channel. The uncertainties include all post-fit experimental uncertainties.

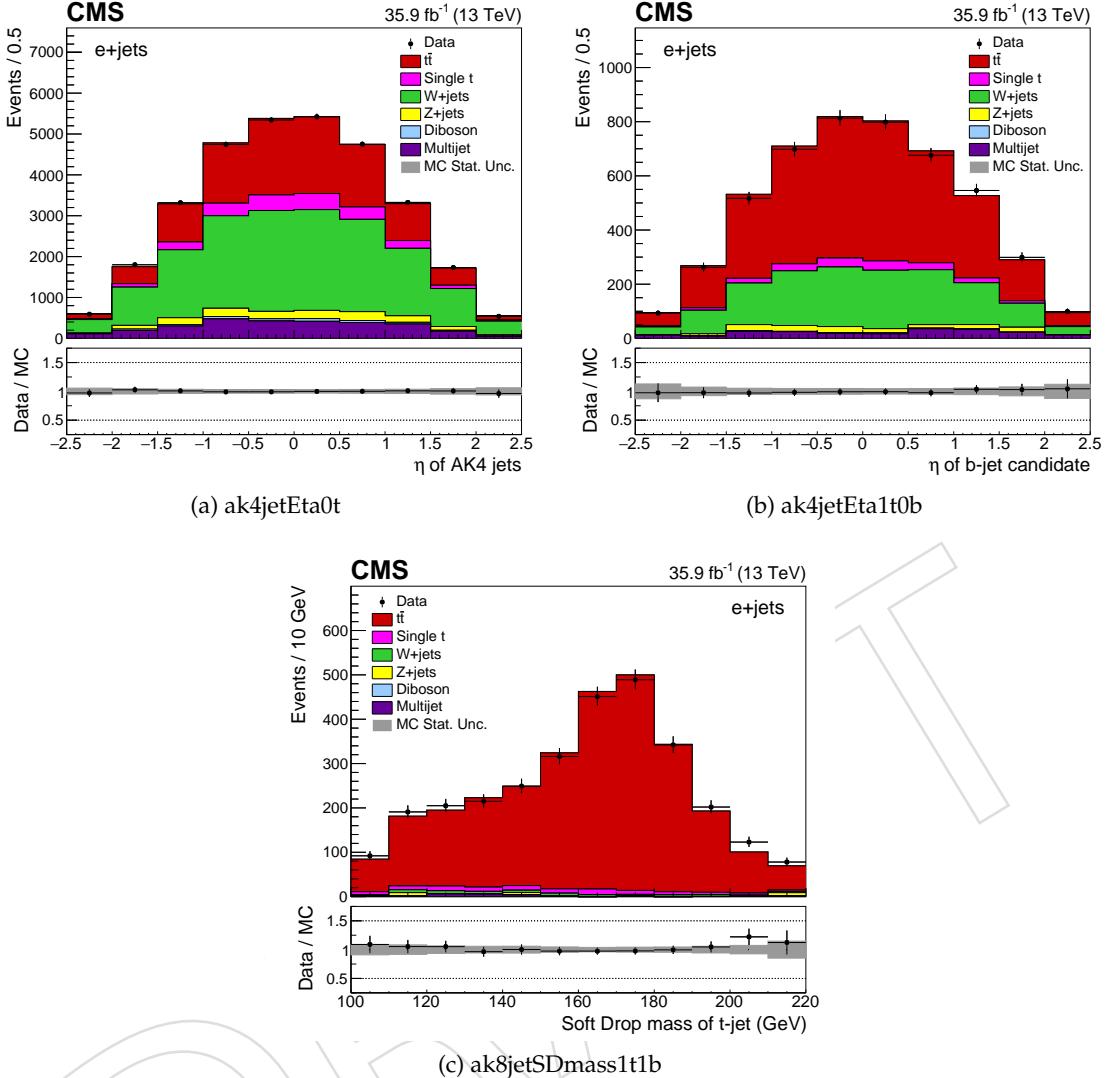


Figure 9: Post-fit kinematic distributions for the fit in the electron channel.

Sample	Number of events (e+jets channel)		
	0t	1t0b	1t1b
tt	9668 ± 1314	2907 ± 126	2729 ± 68
SingleTop	1978 ± 400	178 ± 52	98 ± 27
WJets	15300 ± 2320	1384 ± 205	56 ± 12
ZJets	1351 ± 378	134 ± 46	19 ± 17
Diboson	378 ± 104	20 ± 7	2 ± 1
QCD	2879 ± 1328	208 ± 111	22 ± 23
total	31555 ± 3030	4831 ± 274	2925 ± 80
Data	31559	4801	2953

Table 7: Post-fit signal and background normalizations in the 0t, 1t0b, and 1t1b regions, together with observation in data, for the fit in the electron channel. The uncertainties include all post-fit experimental uncertainties.

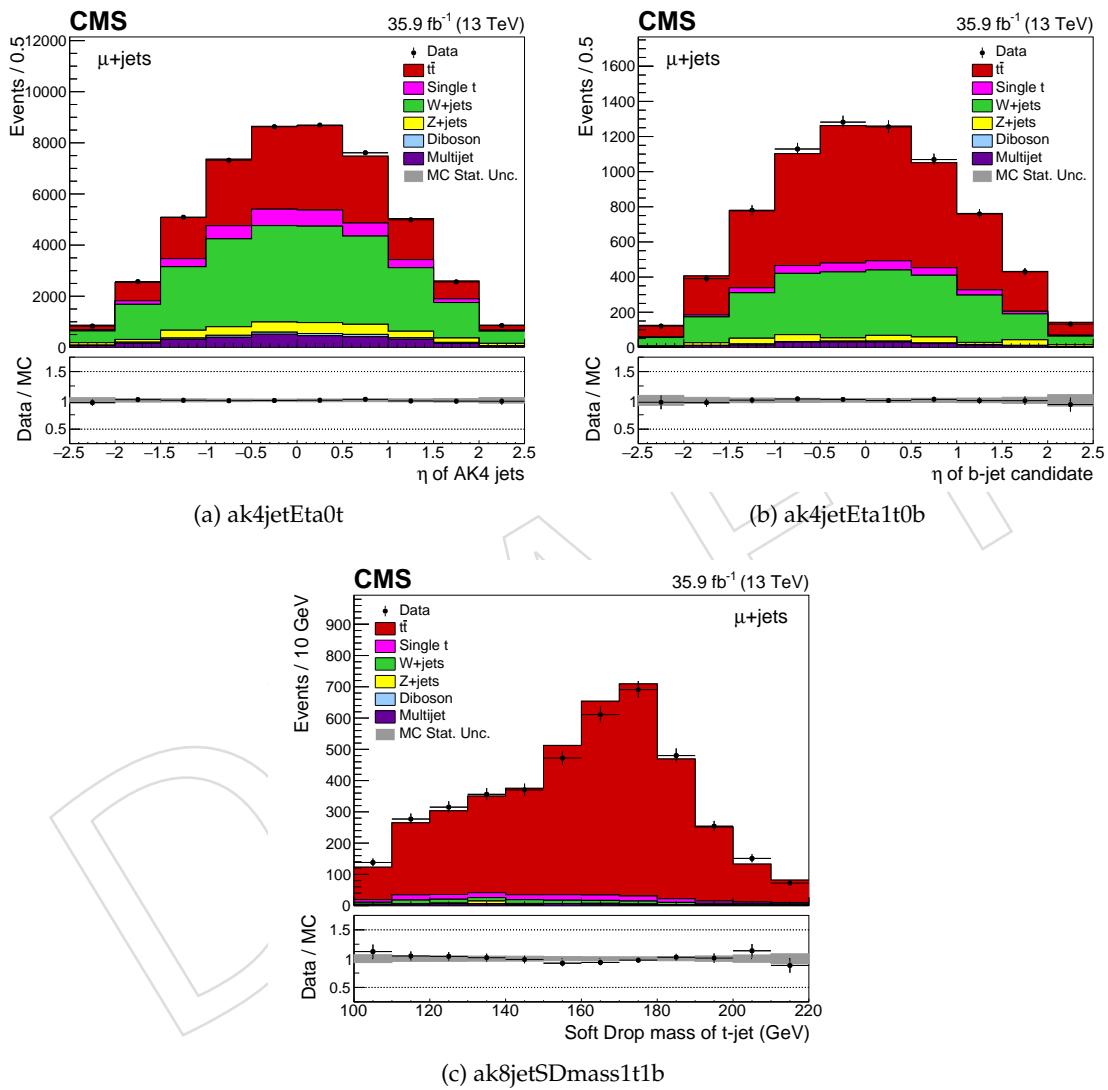


Figure 10: Post-fit kinematic distributions for the combined fit, for events in the muon channel.

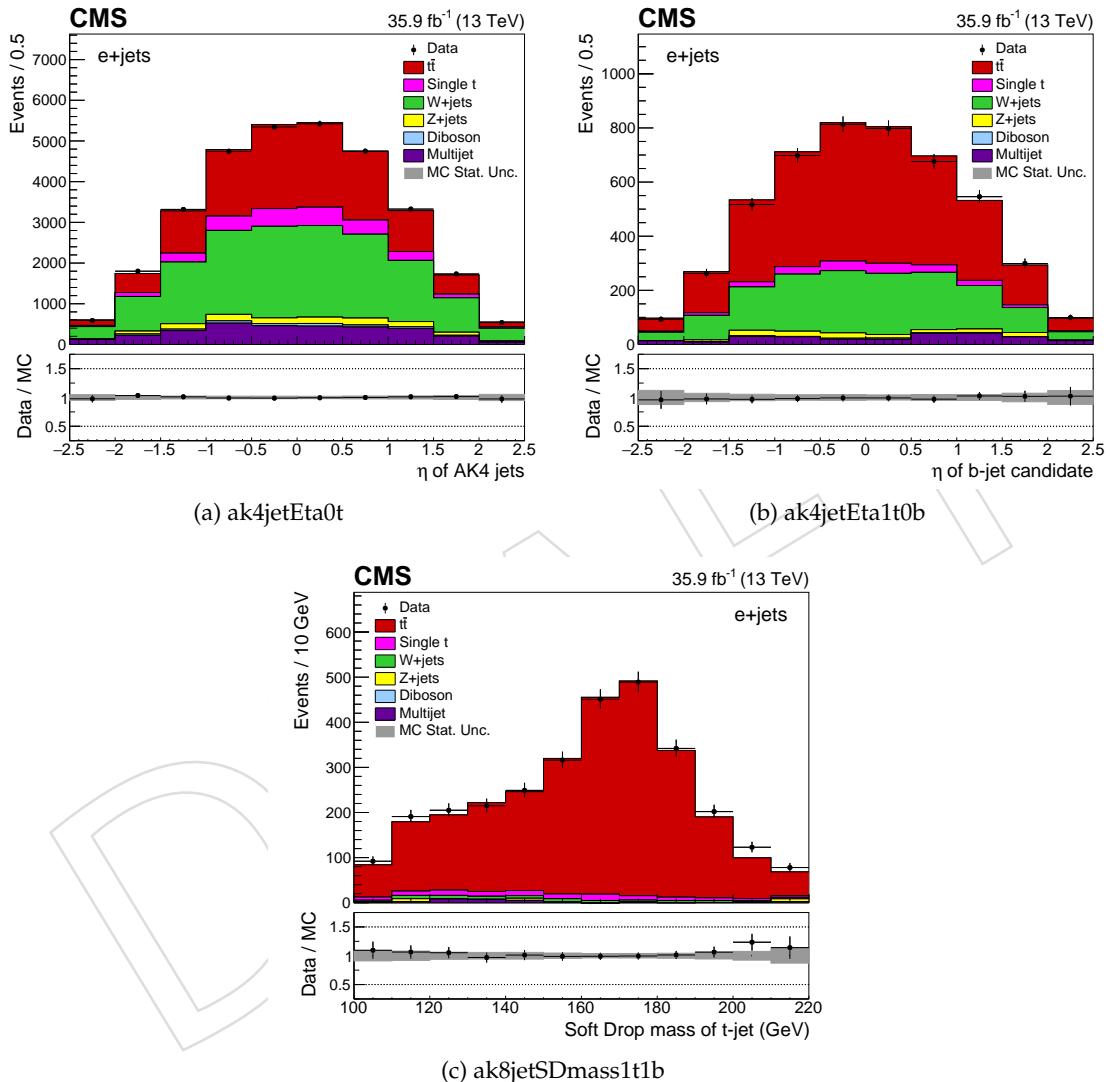


Figure 11: Post-fit kinematic distributions for the combined fit, for events in the electron channel.

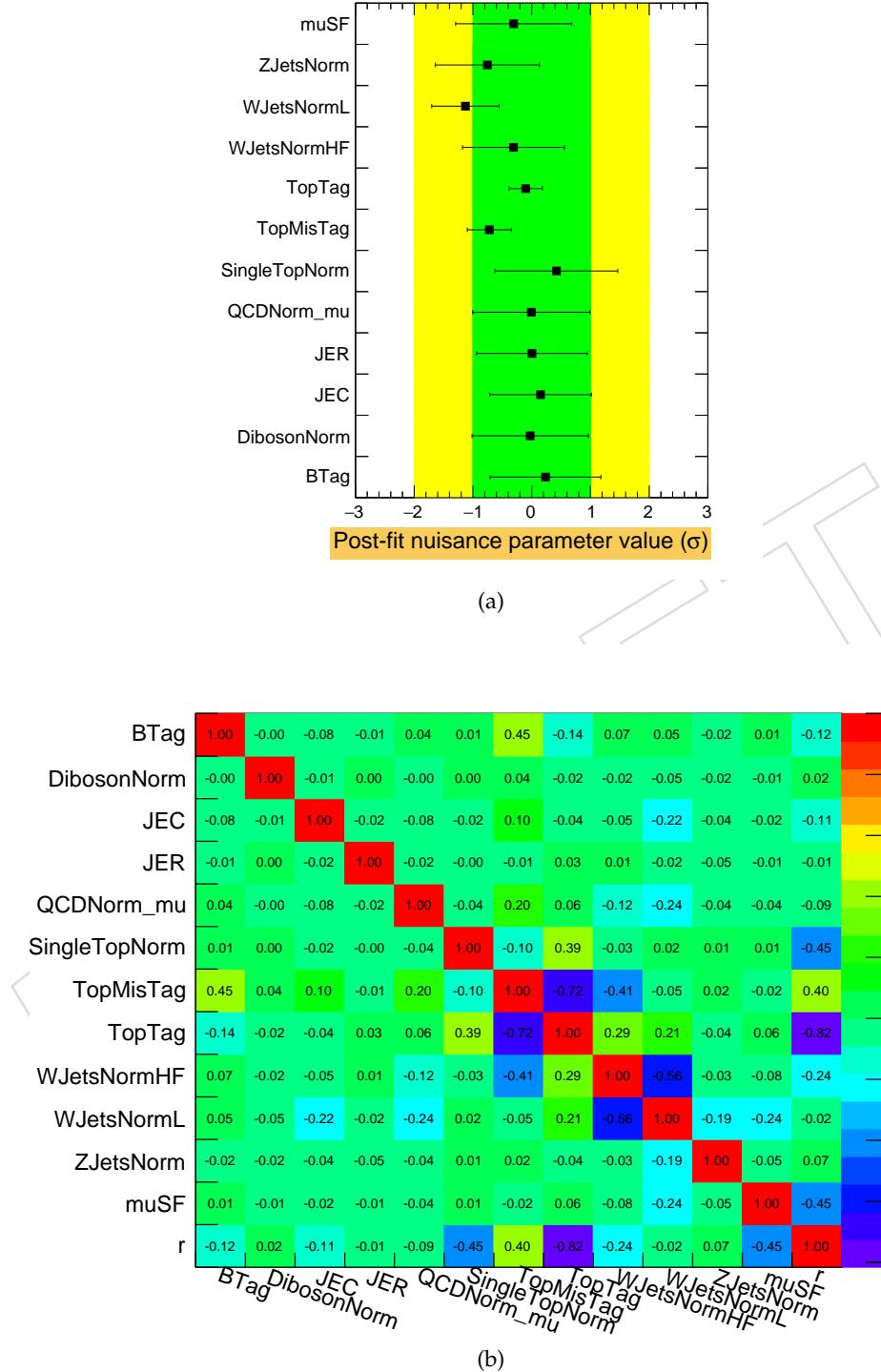


Figure 12: Post-fit nuisance parameters and their correlations for the fit in the muon channel. Nuisance parameters compatible with zero indicate no modification to the systematic by the fit; nuisance parameter uncertainties of 1σ indicate no constraining on the systematic uncertainty due to the fit.

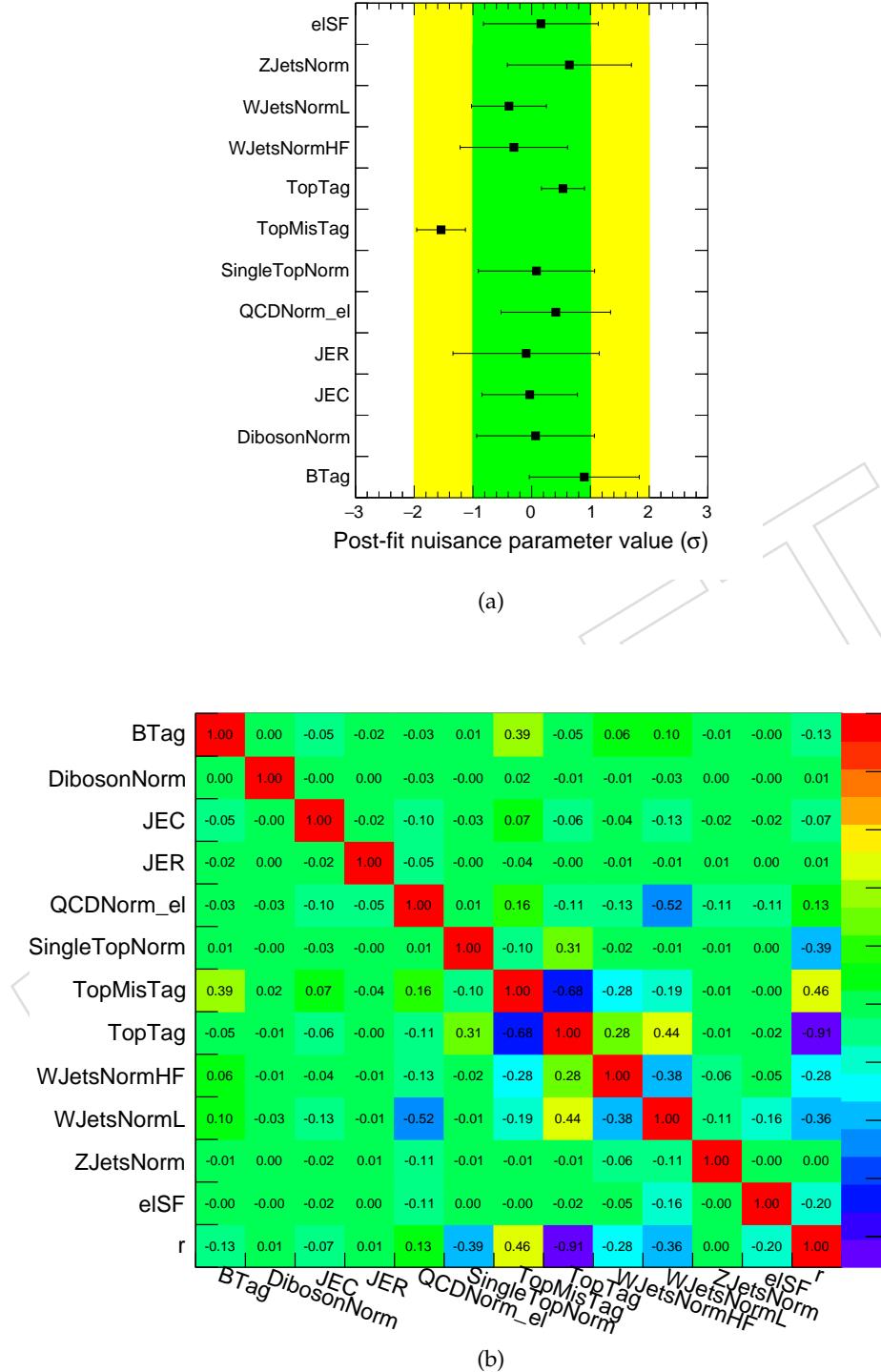


Figure 13: Post-fit nuisance parameters and their correlations for the fit in the electron channel. Nuisance parameters compatible with zero indicate no modification to the systematic by the fit; nuisance parameter uncertainties of 1σ indicate no constraining on the systematic uncertainty due to the fit.

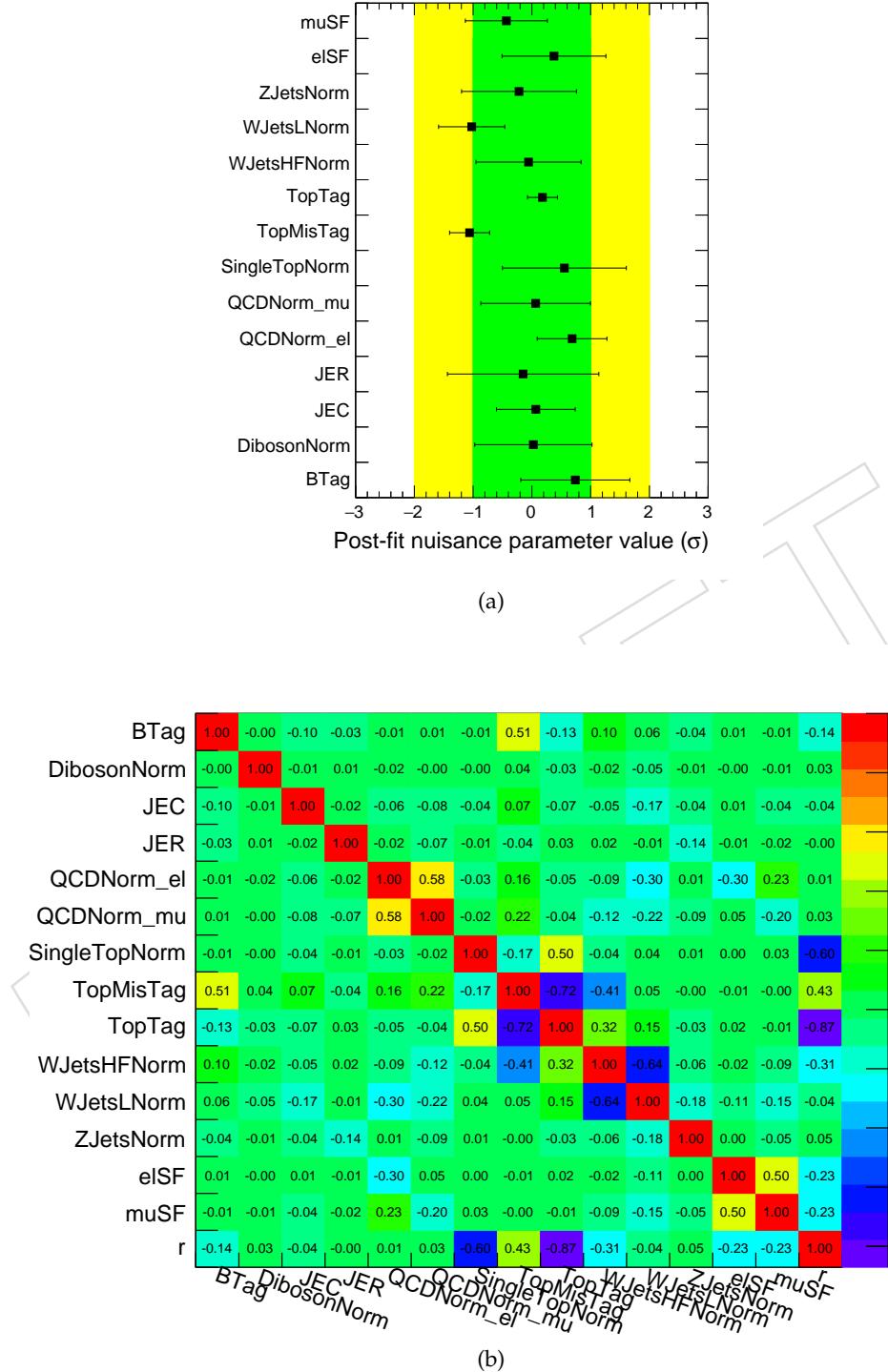


Figure 14: Post-fit nuisance parameters and their correlations for the fit in the combined channel. Nuisance parameters compatible with zero indicate no modification to the systematic by the fit; nuisance parameter uncertainties of 1σ indicate no constraining on the systematic uncertainty due to the fit.

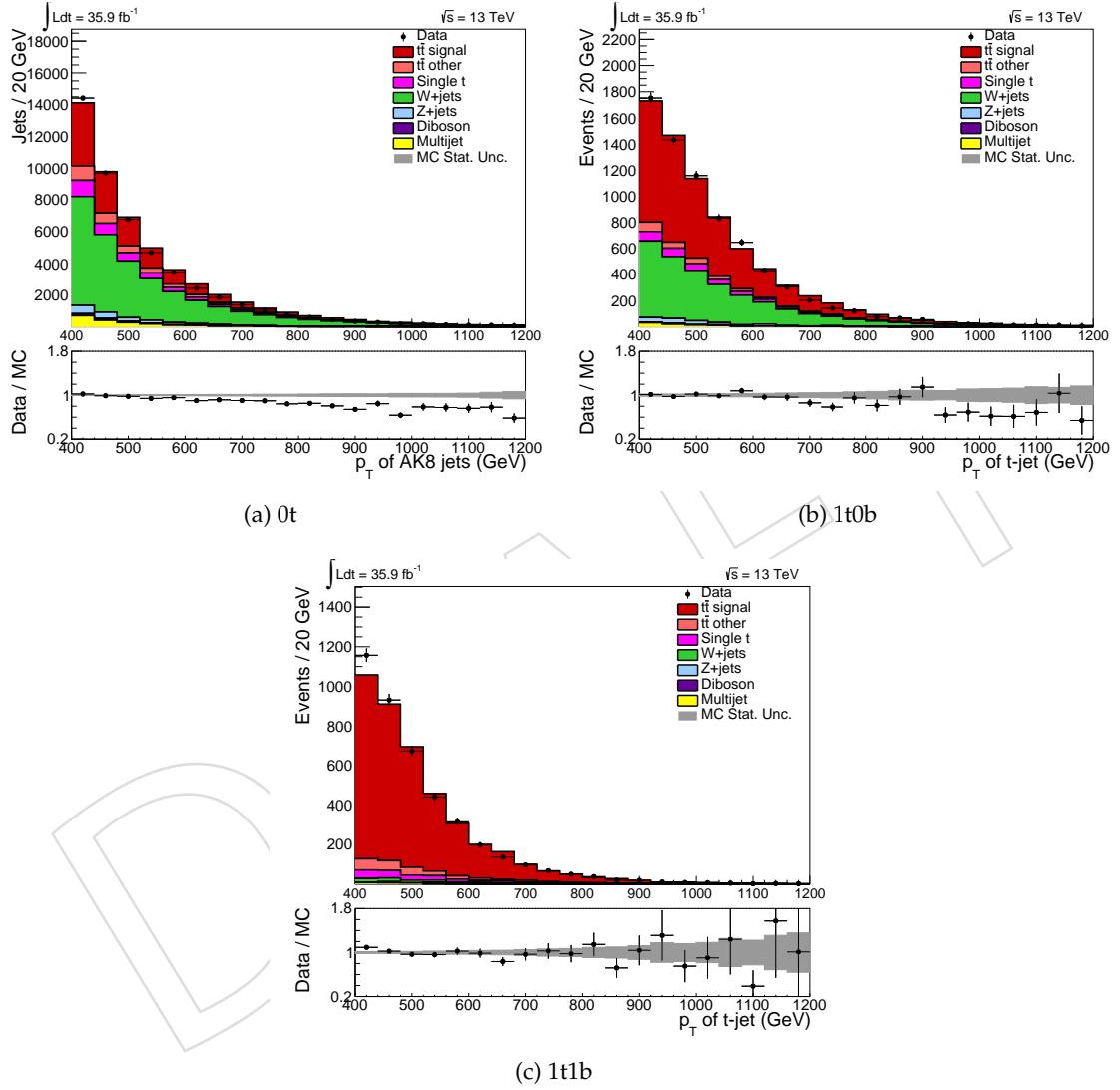


Figure 15: p_T of the hadronic top candidate in the 0t, 1t0b, and 1t1b regions, for events in the muon channel. Posterior top-tag SF, lepton SF, and background normalizations are used.

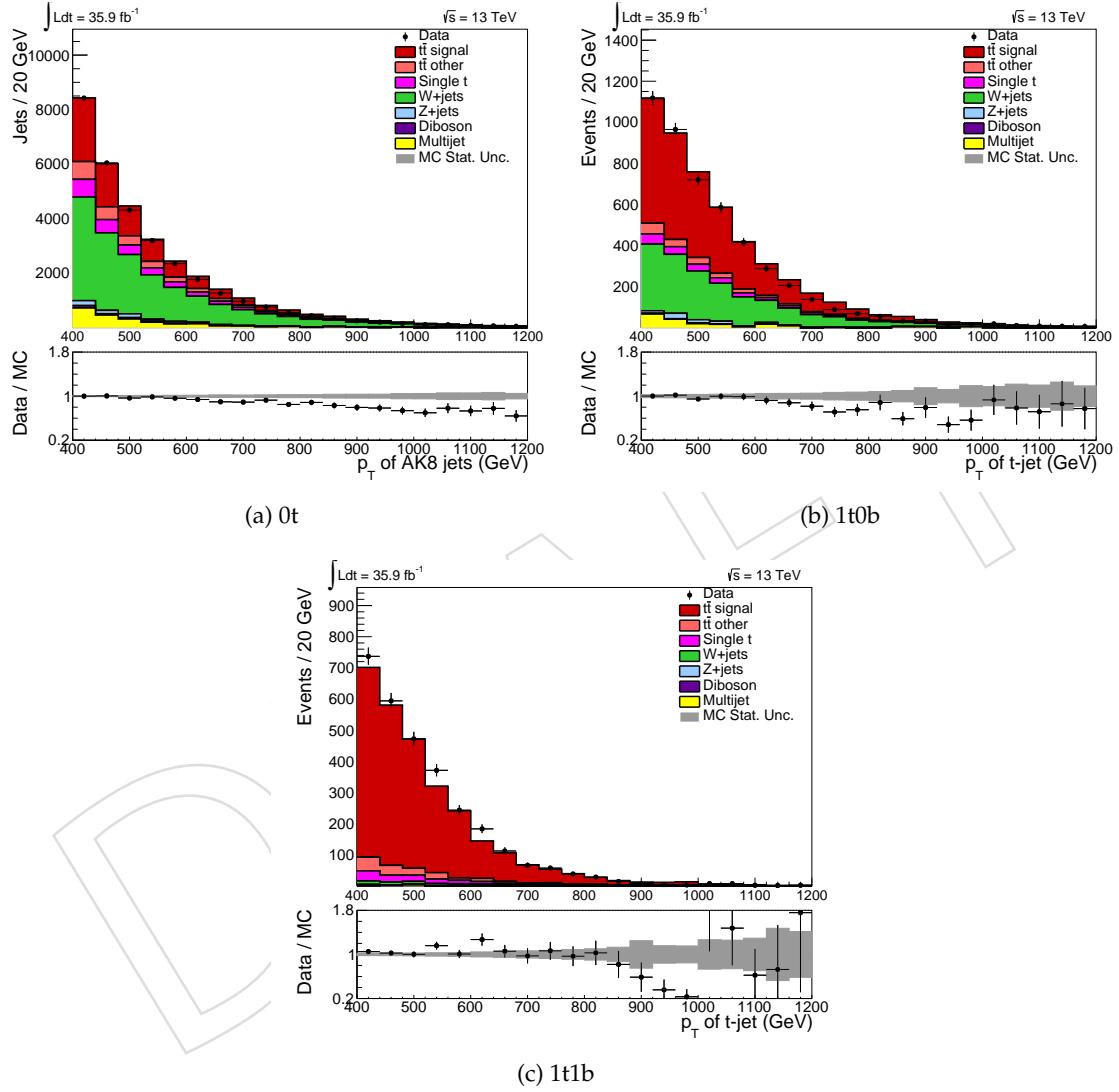


Figure 16: p_T of the hadronic top candidate in the 0t, 1t0b, and 1t1b regions, for events in the electron channel. Posterior top-tag SF, lepton SF, and background normalizations are used.

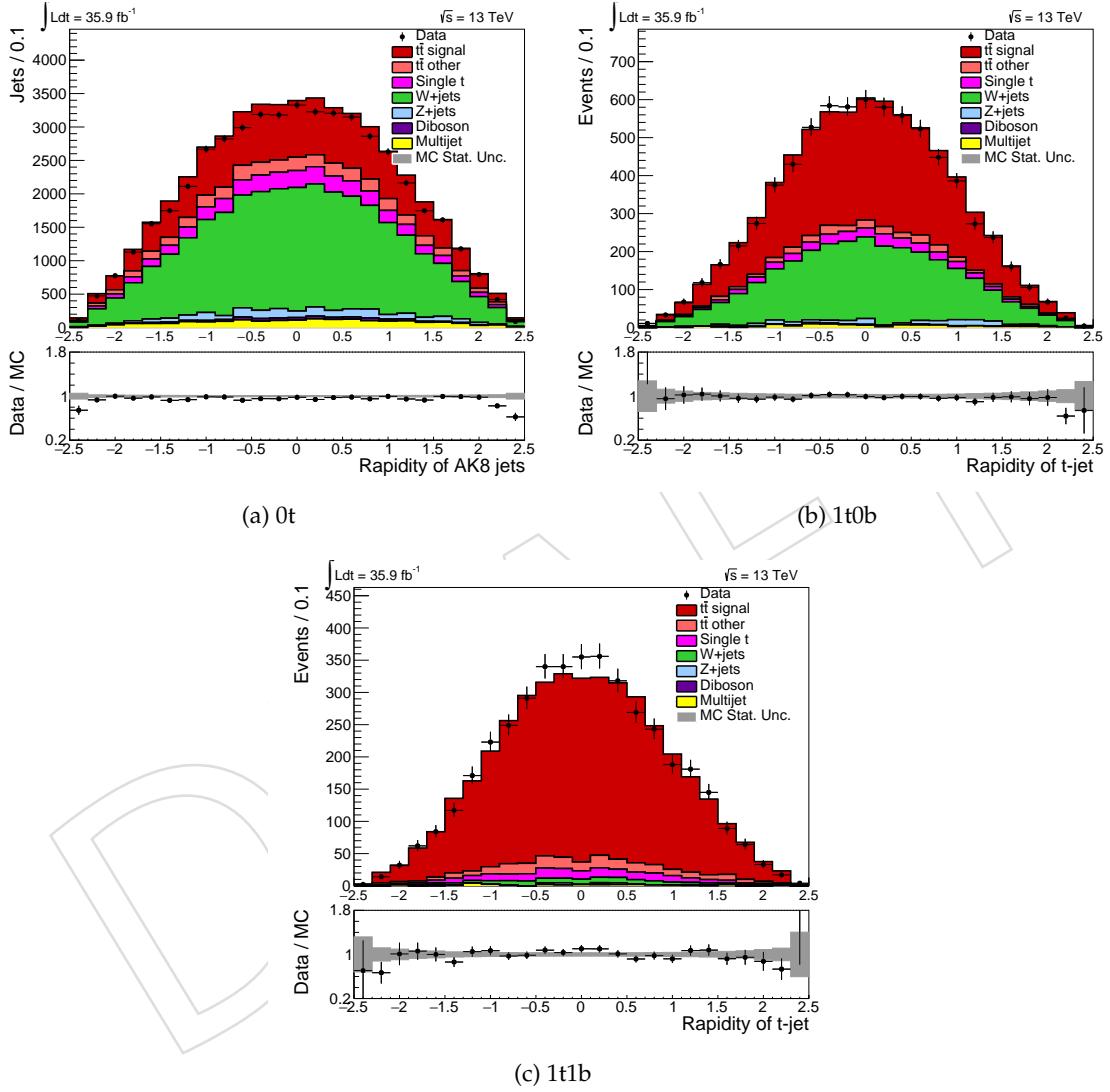


Figure 17: Rapidity of the hadronic top candidate in the 0t, 1t0b, and 1t1b regions, for events in the muon channel. Posterior top-tag SF, lepton SF, and background normalizations are used.

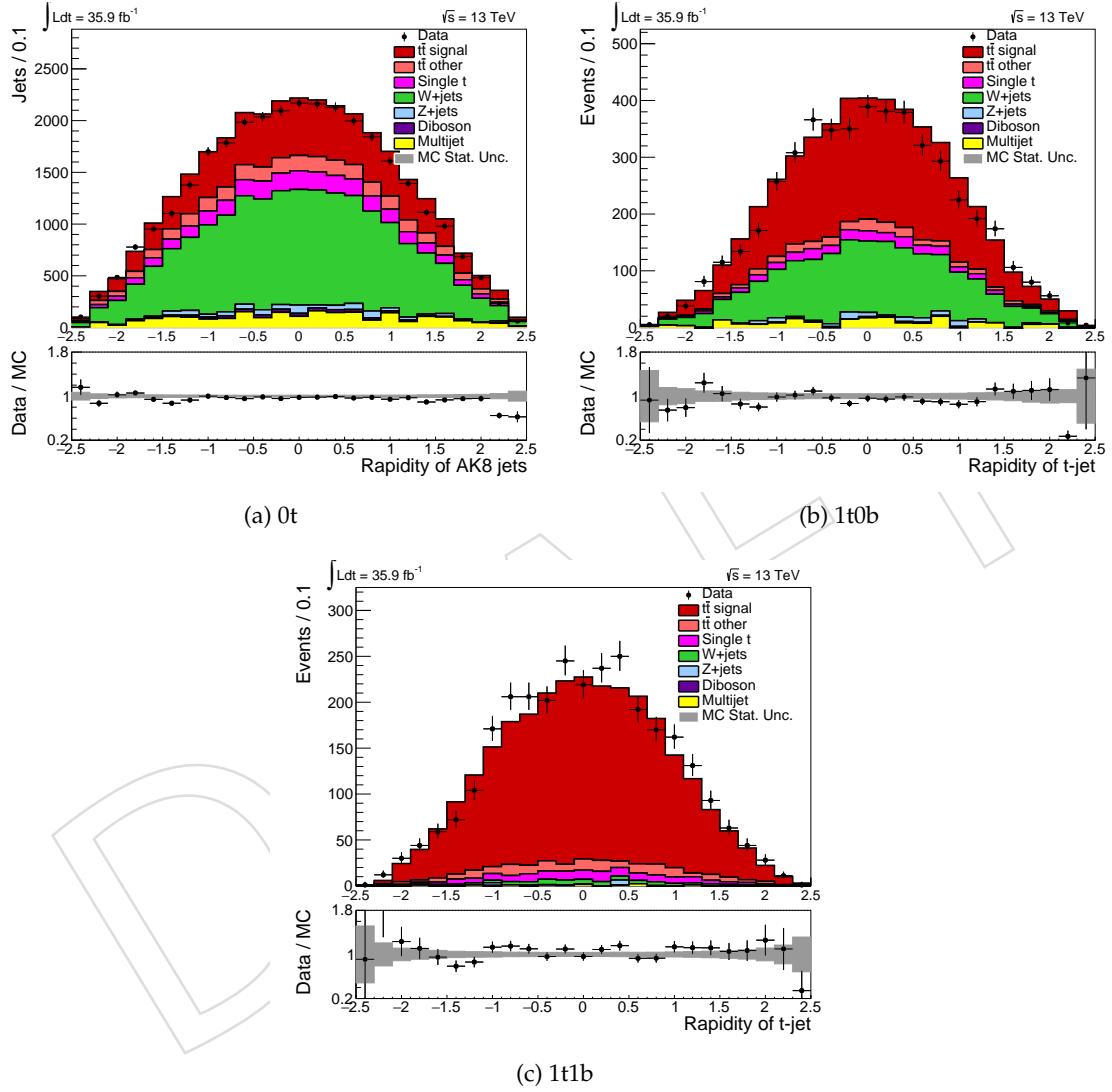


Figure 18: Rapidity of the hadronic top candidate in the 0t, 1t0b, and 1t1b regions, for events in the electron channel. Posterior top-tag SF, lepton SF, and background normalizations are used.

Sample	Number of events (μ +jets channel)		
	0t	1t0b	1t1b
t <bar>t</bar>	16772 \pm 1438	4245 \pm 174	3905 \pm 80
SingleTop	3286 \pm 587	282 \pm 68	153 \pm 34
WJets	23104 \pm 2871	2368 \pm 318	105 \pm 20
ZJets	2582 \pm 680	234 \pm 69	19 \pm 10
Diboson	557 \pm 155	31 \pm 10	2 \pm 1
QCD	2833 \pm 1207	159 \pm 76	43 \pm 22
total	49135 \pm 3549	7320 \pm 383	4228 \pm 93
Data	49137	7348	4187

Sample	Number of events (e+jets channel)		
	0t	1t0b	1t1b
t <bar>t</bar>	10707 \pm 938	2835 \pm 116	2670 \pm 66
SingleTop	2267 \pm 403	191 \pm 47	107 \pm 24
WJets	13945 \pm 1742	1445 \pm 194	62 \pm 12
ZJets	1068 \pm 295	118 \pm 37	17 \pm 15
Diboson	373 \pm 105	22 \pm 7	2 \pm 1
QCD	3200 \pm 735	242 \pm 80	31 \pm 30
total	31560 \pm 2171	4854 \pm 247	2889 \pm 79
Data	31559	4801	2953

Table 8: Post-fit signal and background normalizations in the 0t, 1t0b, and 1t1b regions, together with observation in data, for the fit in the combined channel. The uncertainties include all post-fit experimental uncertainties.

Fit	Posterior nuisance parameters & scale factors				
	nuisance parameter	Top-tag		Top-mis-tag	
		SF		nuisance parameter	SF
Muon-only	-0.10 \pm 0.28	0.98 \pm 0.06		-0.72 \pm 0.38	0.85 \pm 0.08
Electron-only	0.53 \pm 0.37	1.12 \pm 0.09		-1.54 \pm 0.41	0.71 \pm 0.07
Combined	0.18 \pm 0.25	1.04 \pm 0.06		-1.06 \pm 0.34	0.79 \pm 0.06

Table 9: Posterior top-tag and top-mis-tag SF nuisance parameters given by the fits in the muon, electron, and combined channels. There is reasonable agreement between the results.

In addition to the uncertainty from the fit, shape uncertainties on the top-tag and top-mis-tag SFs are introduced to cover the dependence of the top-tag SF on the top jet p_T or $|\eta|$. The derivation of these uncertainties is described in Appendix G. The resulting shape uncertainty due to the $p_T(|\eta|)$ dependence is 4-8% (1-5%) for the top-tag SF and 1-13% (2-3%) for the top-mis-tag SF. As a final check, the fitted top-tag SF is compared to the value provided in [30]. The top-tag SF we calculate shows very good agreement, and comparable uncertainty, with the provided top-tag SF of $1.06^{+0.09}_{-0.04}$.

8.2.1.1 Posterior lepton scale factors The posterior nuisance parameters for the muon and electron combined ID / reconstruction / isolation SFs are given in Table 10 for the muon channel, electron channel, and combined fits. The posterior nuisance parameters from the combined fit are used to determine the posterior muon and electron SFs.

Fit	Posterior nuisance parameters	
	Muon SF	Electron SF
Muon-only	-0.31 ± 0.99	N/A
Electron-only	N/A	0.16 ± 0.98
Combined	-0.43 ± 0.70	0.38 ± 0.88

Table 10: Posterior muon and electron SF nuisance parameters given by the fits in the muon, electron, and combined channels. There is reasonable agreement between the results.

In addition to the uncertainty from the fit, an additional systematic uncertainty on the muon and electron SF nuisance parameters is assessed as a result of a set of fit cross checks, described in detail in F. The final posterior nuisance parameters are given in Table 11. The final nuisance parameters are used to determine the posterior lepton SF, using the interpolation of the SFs corresponding to nuisance parameters of 0, 1, and -1.

Fit	Nuisance parameter
Muon SF	-0.43 ± 0.70 (fit) ± 0.29 (sys)
Electron SF	0.38 ± 0.88 (fit) ± 0.09 (sys)

Table 11: Posterior muon and electron SF nuisance parameters with total uncertainty.

8.2.2 Posterior background normalizations

In addition to the fitted top-tag SF, fitted values of the background rates are used to construct the background-subtracted data distribution which is unfolded to produce the differential cross section. The posterior nuisance parameters for the background rates are given in Table 12 for the muon channel, electron channel, and combined fits. While the nuisance parameters determined from the separate fits do not necessarily bracket the nuisance parameters determined in the combined fit, the results generally agree within uncertainties. The posterior nuisance parameters from the combined fit are used to determine the posterior background rates.

In addition to the uncertainty from the fit, an additional systematic uncertainty on the background normalization nuisance parameters is assessed as a result of the fit checks performed in Appendix F. The final posterior nuisance parameters, along with the resulting ratio of posterior to prior background rate, are given in Table 13.

8.2.3 Signal strength

The posterior signal strength describes the relation between the measured inclusive $t\bar{t}$ cross section and the MC prediction. While the signal strength is not directly used to determine

Fit	Posterior nuisance parameters					
	Single top	W+light	W+heavy	Z+Jets	Diboson	QCD
Muon	0.42 ± 1.04	-1.13 ± 0.57	-0.31 ± 0.87	-0.75 ± 0.89	-0.03 ± 0.99	-0.01 ± 1.00
Electron	0.08 ± 0.99	-0.39 ± 0.64	-0.30 ± 0.92	0.64 ± 1.06	0.07 ± 1.00	0.41 ± 0.93
Combined	0.55 ± 1.05	-1.02 ± 0.56	-0.06 ± 0.89	-0.22 ± 0.98	0.03 ± 1.00	0.06 ± 0.93 (mu) 0.68 ± 0.59 (el)

Table 12: Posterior background normalization nuisance parameters given by the fits in the muon, electron, and combined channels. There is reasonable agreement between the results in the different channels.

Fit	Nuisance parameter	Posterior factor
Single top	0.55 ± 1.05 (fit) ± 0.22 (sys)	$1.16^{+0.37}_{-0.29}$
W+light	-1.02 ± 0.56 (fit) ± 0.77 (sys)	$0.77^{+0.21}_{-0.17}$
W+heavy flavor	-0.06 ± 0.89 (fit) ± 0.58 (sys)	$0.98^{+0.32}_{-0.23}$
Z+Jets	-0.22 ± 0.98 (fit) ± 0.36 (sys)	$0.94^{+0.30}_{-0.22}$
Diboson	0.03 ± 1.00 (fit) ± 0.05 (sys)	$1.01^{+0.30}_{-0.23}$
QCD (mu)	0.06 ± 0.93 (fit) ± 1.70 (sys)	$1.02^{+1.23}_{-0.55}$
QCD (el)	0.68 ± 0.59 (fit) ± 1.08 (sys)	$1.32^{+0.85}_{-0.52}$

Table 13: Posterior background normalization nuisance parameters with total uncertainty, along with corresponding posterior normalization factors. Normalization factors applied to a priori background normalizations yields posterior values. The lognormal a priori background normalization uncertainties are 30% for the single top, W+light, W+heavy flavor, Z+jets, and diboson components and 50% for the QCD components.

571 the background-subtracted data distribution, it is implicitly related to the normalization of
572 the final background-subtracted data distribution. Therefore, the signal strength remains a
573 useful indicator of the overall fit result. The signal strengths are given in Table 14 for the
574 muon channel, electron channel, and combined fits. In addition to the uncertainty from the fit,
575 an additional systematic uncertainty on the background normalization nuisance parameters is
576 assessed as a result of the fit checks performed in Appendix F. The measured signal strengths
577 indicate that MC overpredicts the rate of $t\bar{t}$ production in the boosted regime.

Fit	Signal strength
Muon-only	0.85 ± 0.07 (fit) ± 0.06 (sys)
Electron-only	0.77 ± 0.07 (fit) ± 0.09 (sys)
Combined	0.81 ± 0.05 (fit) ± 0.07 (sys)

Table 14: Signal strength given by the fits in the muon, electron, and combined channels. There is reasonable agreement between the results.

DRAFT

578 9 Differential Cross Section

579 This section presents the cross section measurement. The differential $t\bar{t}$ cross section is mea-
 580 sured as a function of the transverse momentum (p_T) and rapidity (y) of the hadronically de-
 581 caying top quark. The measurements are performed at particle level within a fiducial region
 582 similar to the reconstructed phase space and at parton level in the full phase space. The cross
 583 sections are extrapolated from the measured $p_T/y(\text{top})$ distribution through an unfolding pro-
 584 cedure, accounting for signal reconstruction efficiencies and bin migrations. The measured
 585 $p_T(\text{top})$ distribution is extracted from the signal-dominated 1t1b region. Background contribu-
 586 tions are subtracted from the data distribution, using the posterior normalizations from the
 587 combined maximum likelihood fit given in Table 8. The posterior top-tagging efficiency and
 588 lepton ID / trigger / reconstruction efficiencies for the combined fit are applied to the MC in
 589 place of their a priori values. The background-subtracted data distribution is used as the mea-
 590 sured $t\bar{t}$ distribution. The unfolding uses the TUnfold method [39, 40], and relies on a response
 591 matrix mapping the p_T/y of the top-tagged jet to the top quark (particle-level top jet) p_T/y .
 592 This response matrix is derived from the semileptonic $t\bar{t}$ POWHEG+PYTHIA8 MC.

593 The systematic uncertainty on the unfolded measurement receives contributions from both
 594 experimental and theoretical sources. The experimental uncertainties, described in Section 7.1,
 595 include the lepton ID / trigger / reconstruction efficiency, luminosity, pileup corrections, jet
 596 energy scale, jet energy resolution, b tagging efficiency, top tagging efficiency, and background
 597 normalizations. The a posteriori values from the likelihood fit are used for the top tagging
 598 efficiency, background rates, and lepton ID / trigger / reconstruction efficiency, while the a
 599 priori values are used for the remaining uncertainties. The theoretical uncertainties due to
 600 PDF, renormalization and factorization scales, shower scales, underlying event, matching scale,
 601 color reconnection, and parton shower are also considered. unfolding uncertainties

602 The unfolding is performed separately in the muon and electron channels as a cross-check, and
 603 for the combined lepton+jets phase space (primary measurement).

604 9.1 Phase Space

605 9.1.1 Parton level

606 The parton level phase space includes semileptonic events where the p_T of the hadronically de-
 607 caying top quark is greater than 400 GeV. Semileptonic ($e/\mu+\text{jets}$) events are selected at parton
 608 level in the following way. First, the event is required to contain exactly one prompt lepton.
 609 Prompt leptons may be electrons, muons, or taus, defined as follows:

- 610 • Prompt electron: $|\text{PDGID}| = 11$, `ISPROMPTFINALSTATE`
- 611 • Prompt muon: $|\text{PDGID}| = 13$, `ISPROMPTFINALSTATE`
- 612 • Prompt tau: $|\text{PDGID}| = 15$, `ISPROMPTDECAYED`

613 In addition, the event must contain a final state muon or electron, where the final state particle
 614 must satisfy `ISPROMPTFINALSTATE` or `ISDIRECTPROMPTTAUDECAYPRODUCTFINALSTATE`. The
 615 hadronically decaying top quark is identified as the top or antitop quark through the sign of the
 616 final state lepton. Specifically, the on-shell top quark after FSR is used, identified through the
 617 `ISLASTCOPY` flag (status code 62 in Pythia8). The `ISPROMPTFINALSTATE`, `ISPROMPTDECAYED`,
 618 `ISDIRECTPROMPTTAUDECAYPRODUCTFINALSTATE`, and `ISLASTCOPY` flags are all `GENSTA-`
 619 `TUSFLAGS` stored for all genParticles in MINIAOD. Further information may be found in [33].
 620 The differential cross section at parton level is given as a function of the hadronically decaying
 621 top quark p_T and y . This parton-level selection follows the guidelines given in [41].

9.1.2 Particle level

The particle level phase space is set up to mimic the kinematic selections at reconstructed level. Particle-level AK8 jets, defined through the CMS ParticleLevelProducer [42], are selected if they fulfill $p_T > 400 \text{ GeV}$, $|\eta| < 2.4$, and $105 < m_{jet} < 220 \text{ GeV}$, and are then referred to as particle-level top jets. Particle-level AK4 jets are selected if they have $p_T > 50 \text{ GeV}$, $|\eta| < 2.4$, and are flagged as a b -jet (through its associated PDG ID), referred to as particle-level b jets. Particle-level electrons (muons) are selected if they have $p_T > 50 \text{ GeV}$, $|\eta| < 2.5(2.1)$, and have PDG ID 11 (13). To pass the particle-level selection, an event must contain at least one particle-level top jet, at least one particle-level b jet, and at least one particle-level electron or muon.

9.2 $t\bar{t}$ Simulation

Since the unfolding procedure is very sensitive to statistical fluctuations, dedicated samples with high $t\bar{t}$ invariant mass are used to improve statistics. Samples are generated for $700 \text{ GeV} < M_{t\bar{t}} < 1000 \text{ GeV}$ and $M_{t\bar{t}} > 1000 \text{ GeV}$. In addition, an extension to the inclusive $t\bar{t}$ sample is used to improve statistics at lower energies. These samples are described in Table 15. To prevent overlap between the samples, a cut of $M_{t\bar{t}} < 700 \text{ GeV}$ is applied to the inclusive $t\bar{t}$ samples. $M_{t\bar{t}}$ is defined as the invariant mass of the hard process top quarks prior to FSR, identified via the ISHARDPROCESS flag. The acceptance for the high- $M_{t\bar{t}}$ samples is calculated from simulation, using the fraction of events in the inclusive $t\bar{t}$ sample with $700 \text{ GeV} < M_{t\bar{t}} < 1000 \text{ GeV}$ and $M_{t\bar{t}} > 1000 \text{ GeV}$. The acceptance for $700 \text{ GeV} < M_{t\bar{t}} < 1000 \text{ GeV}$ is 0.0967 and the acceptance for $M_{t\bar{t}} > 1000 \text{ GeV}$ is 0.0256. To cross-check the stitching of the combined high-statistics sample, certain kinematic distributions in the stitched sample are compared against the kinematic distributions from the original inclusive sample. These comparisons are shown in Figure 19. $\text{acceptance} = \# \text{events in } [700, 1000] \text{ from Nominal ttbar} / \# \text{events in the Mtt 700-1000 sample}$

Sample	N_{events}	Acceptance
TT_TuneCUETP8M2T4_13TeV-powheg-pythia8	77229341	1.0
TT_Mtt-700to1000_TuneCUETP8M2T4_13TeV-powheg-pythia8	78006311	1.0
TT_Mtt-1000toInf_TuneCUETP8M2T4_13TeV-powheg-pythia8	38578334	0.0967
TT_Mtt-1000toInf_TuneCUETP8M2T4_13TeV-powheg-pythia8	24495211	0.0256

Table 15: MC samples used to improve statistics for the unfolding. All other kinematic plots and fit use only the first part of the inclusive sample. All samples have cross section 831.76 pb^{-1} . The $M_{t\bar{t}}$ acceptances are determined by calculating the fraction of events in the inclusive sample with $M_{t\bar{t}}$ in the given range.

9.3 Response Matrix

A response matrix is used to map the p_T/y of the top-tagged jet to the top quark (particle-level top jet) p_T/y . The response matrix is generated from $t\bar{t}$ MC. The binning of the response matrix is chosen such that the purity and stability in each bin is generally greater than 50%. The purity is defined as the fraction of events reconstructed in a given bin which have a true p_T/y in the same bin (purity = $N(\text{rec}_i \& \text{gen}_i) / N(\text{rec}_i)$ for bin i), whereas the stability is the fraction of events with a true p_T/y in a given bin which are reconstructed in the same bin (stability = $N(\text{rec}_i \& \text{gen}_i) / N(\text{gen}_i)$ for bin i). The binning is also chosen such that the resolution in each bin is generally smaller than the bin width. The resolution is defined as the RMS of the reconstructed top jet p_T/y for events in a given bin of top quark (particle-level top jet) p_T/y . Alternatively, the resolution may be parameterized as the width of a Gaussian fit to the reconstructed top jet p_T/y . In addition to the purity, stability, and resolution, the efficiency (ratio of

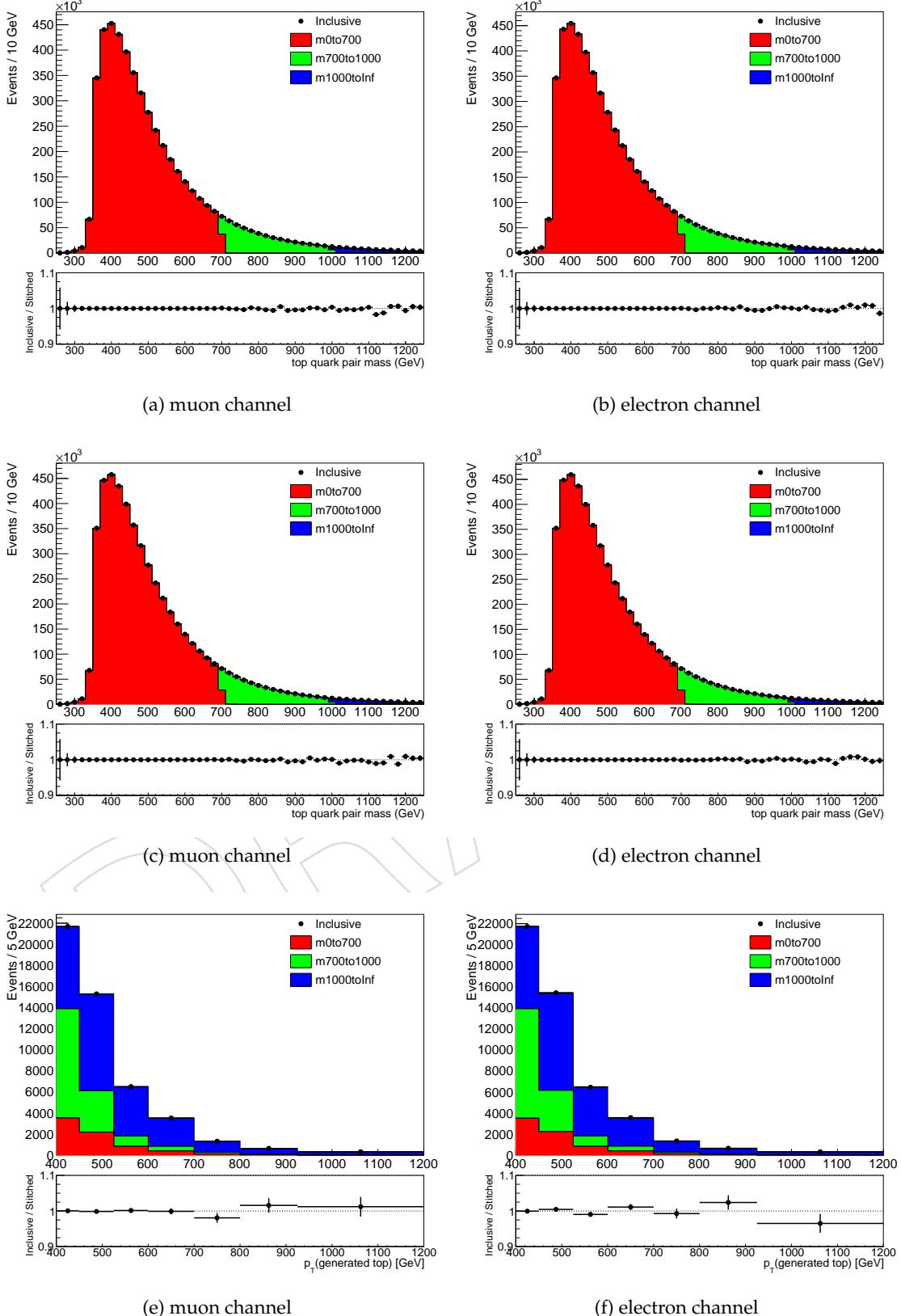


Figure 19: Comparison of kinematics between the original inclusive and stitched high-statistics $t\bar{t}$ samples. The kinematics shown are $M_{t\bar{t}}$ for the hard process tops prior to FSR (a,b), parton-level $M_{t\bar{t}}$ post-FSR (c,d), and parton-level top quark p_T (e,f). The stitched sample kinematics match the original inclusive sample, with improved statistics.

657 number of reconstructed events to true in a given bin) and relative statistical uncertainty are
 658 also considered when optimizing the binning. Figures 20, 21, 22, and 23 show the purity,
 659 stability, efficiency, resolution, and statistical uncertainty for the chosen binning for parton-level
 660 unfolding vs p_T and y , and for particle-level unfolding vs p_T and y . The response matrices
 661 using this binning are shown in Figures 24, 25, 26, and 27. Each bin of generated top quark
 662 (particle-level top jet) p_T/y is split into two bins of reconstructed top p_T/y , in order to satisfy a
 663 requirement of the TUnfold algorithm.

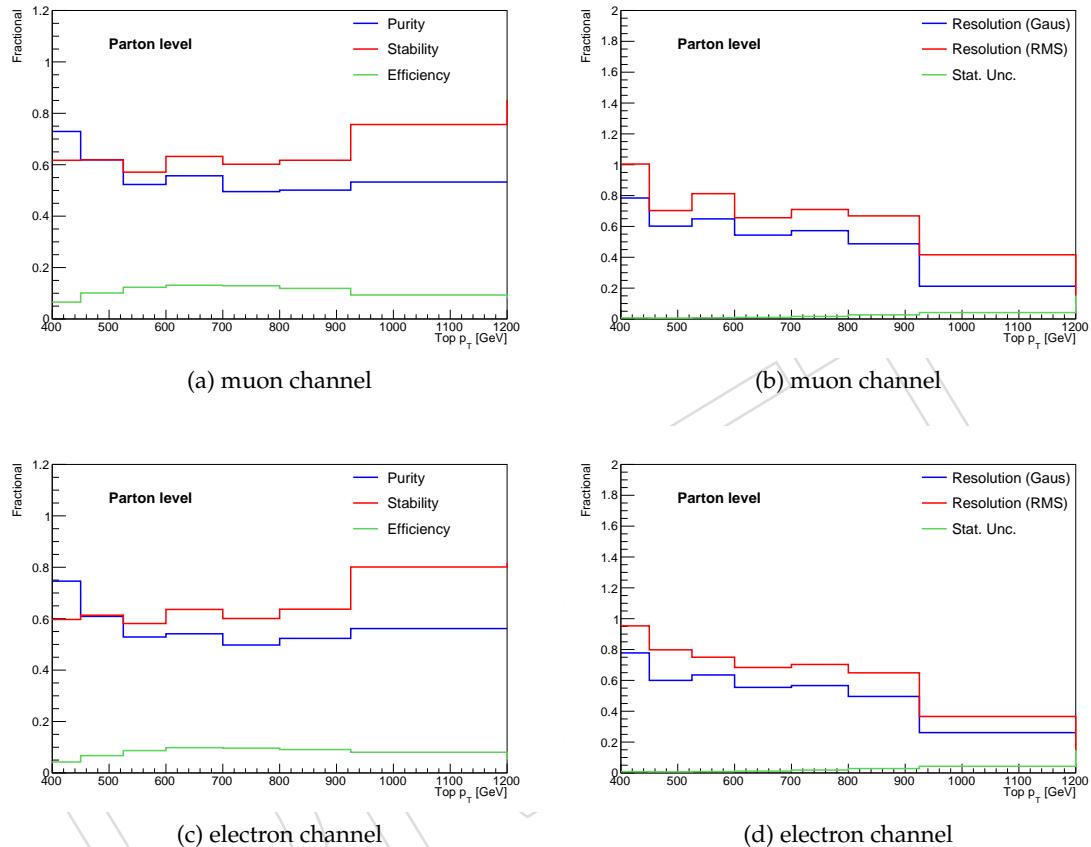


Figure 20: **Parton level** unfolding vs top quark p_T . Plots of purity, stability, and efficiency (left), as well as relative resolution and statistical uncertainty (right), for the muon and electron channels. The relative resolution is defined as the resolution divided by the bin width, while the relative statistical uncertainty is the statistical uncertainty divided by the bin contents. The binning has been chosen such that the purity and stability are generally greater than 50% in all bins, and the RMS resolution is less than the bin width.

664 9.4 TUnfold

665 The TUnfold algorithm is based on a least squares minimization with Tikhonov regularization.
 666 Specifically, the algorithm finds the stationary point of

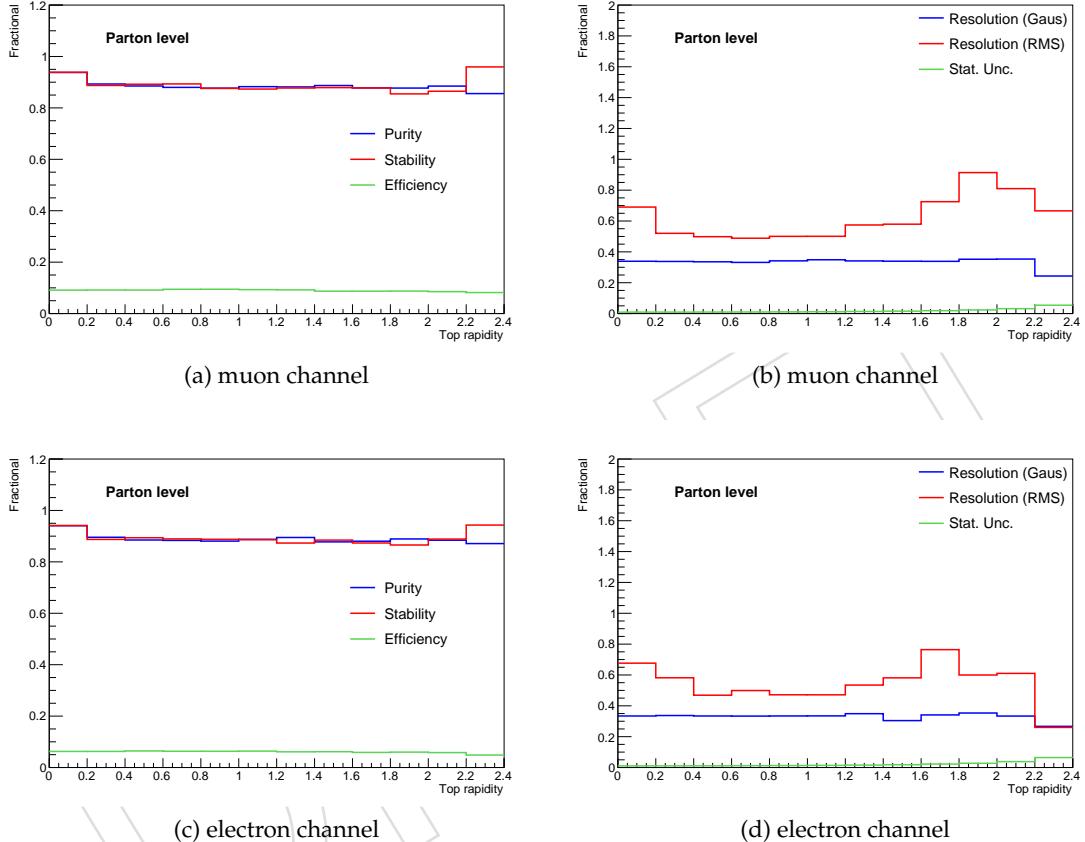


Figure 21: **Parton level** unfolding vs top quark y . Plots of purity, stability, and efficiency (left), as well as relative resolution and statistical uncertainty (right), for the muon and electron channels. The relative resolution is defined as the resolution divided by the bin width, while the relative statistical uncertainty is the statistical uncertainty divided by the bin contents. The binning has been chosen such that the purity and stability are generally greater than 50% in all bins, and the RMS resolution is less than the bin width.

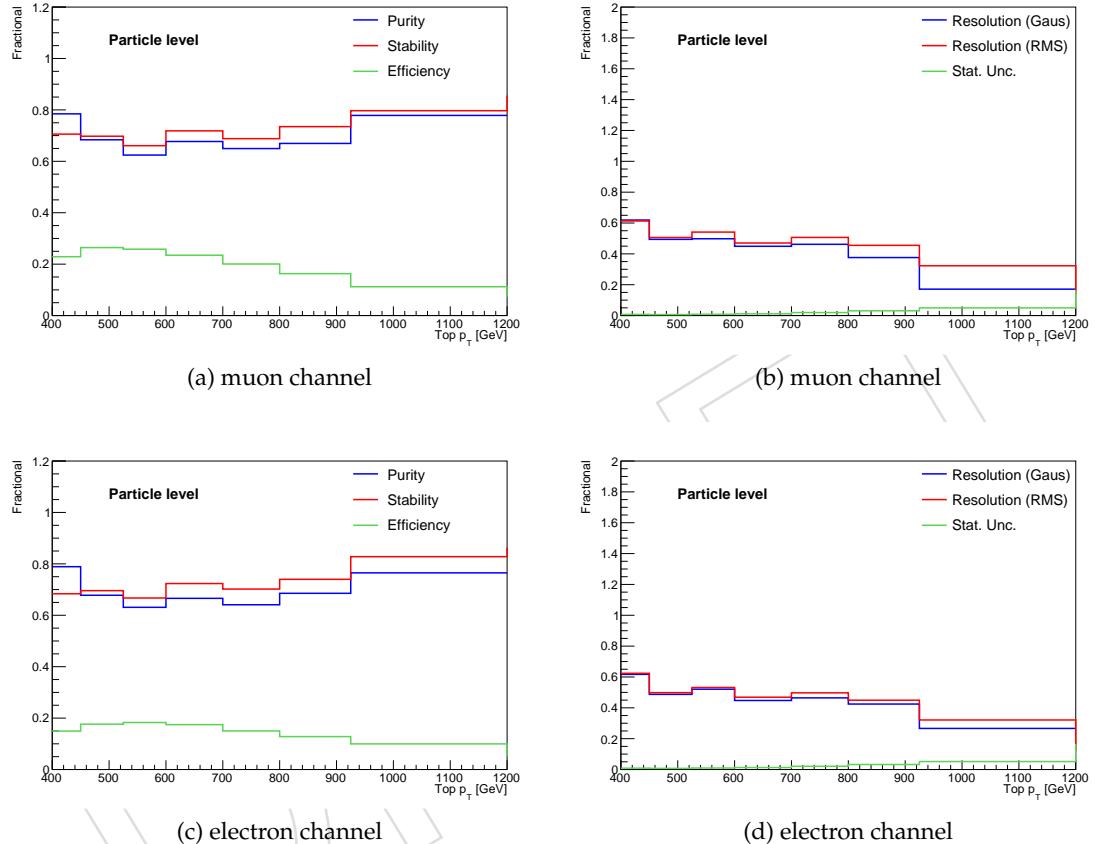


Figure 22: **Particle level** unfolding vs top jet p_T . Plots of purity, stability, and efficiency (left), as well as relative resolution and statistical uncertainty (right), for the muon and electron channels. The relative resolution is defined as the resolution divided by the bin width, while the relative statistical uncertainty is the statistical uncertainty divided by the bin contents. The binning has been chosen such that the purity and stability are generally greater than 50% in all bins, and the RMS resolution is less than the bin width.

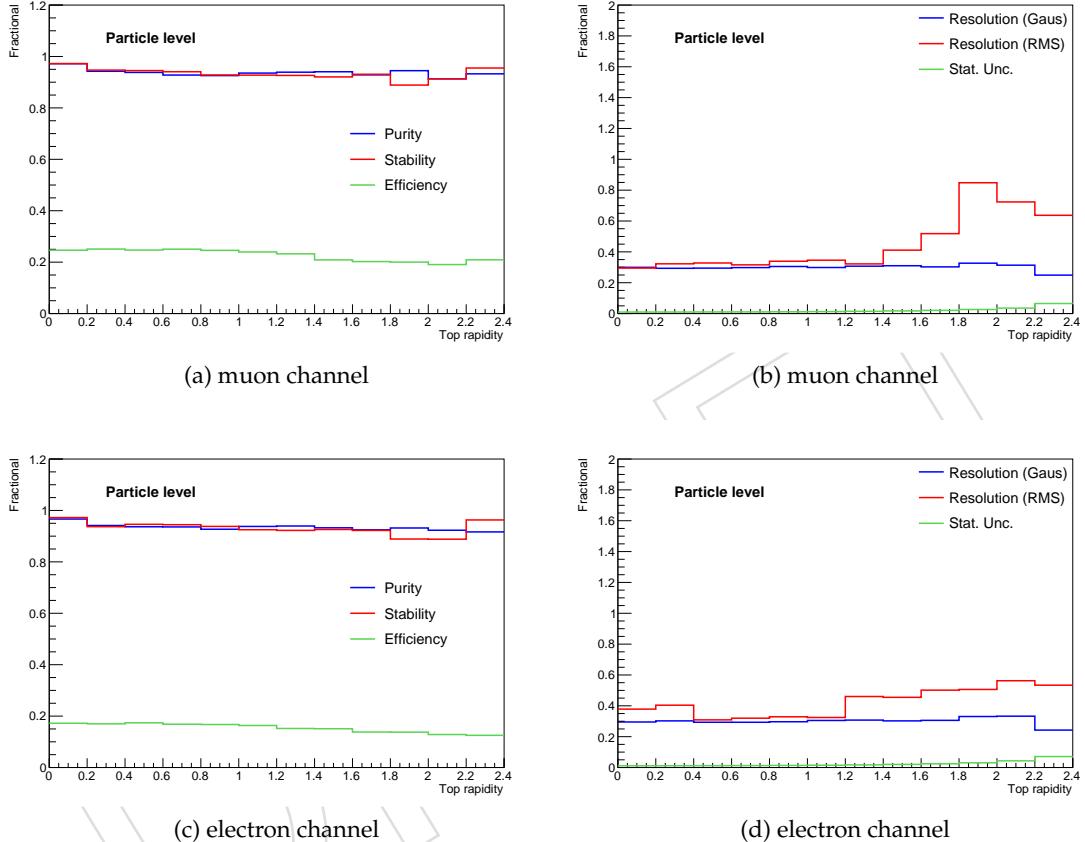


Figure 23: **Particle level** unfolding vs top jet y . Plots of purity, stability, and efficiency (left), as well as relative resolution and statistical uncertainty (right), for the muon and electron channels. The relative resolution is defined as the resolution divided by the bin width, while the relative statistical uncertainty is the statistical uncertainty divided by the bin contents. The binning has been chosen such that the purity and stability are generally greater than 50% in all bins, and the RMS resolution is less than the bin width.

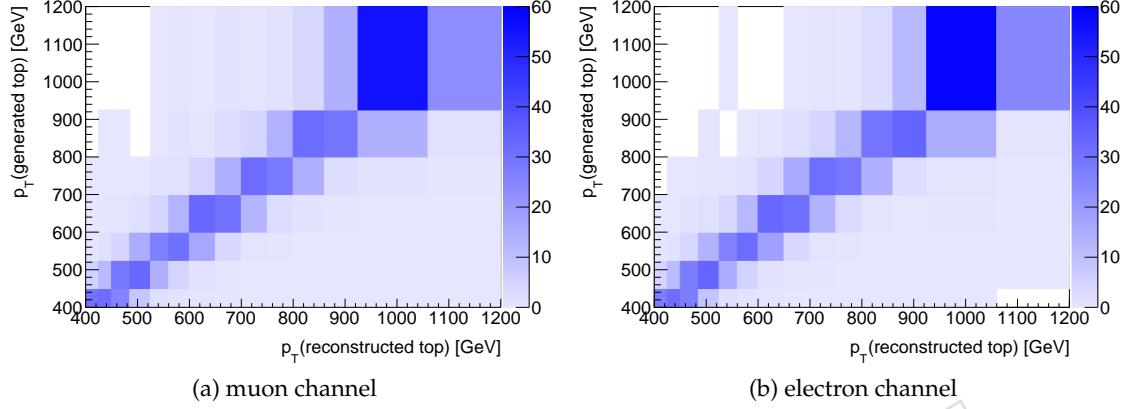


Figure 24: **Parton level:** Response matrices mapping the p_T of the top-tagged jet to the top quark p_T in the muon and electron channels. Response matrices are normalized to 100% in each bin of top quark p_T .

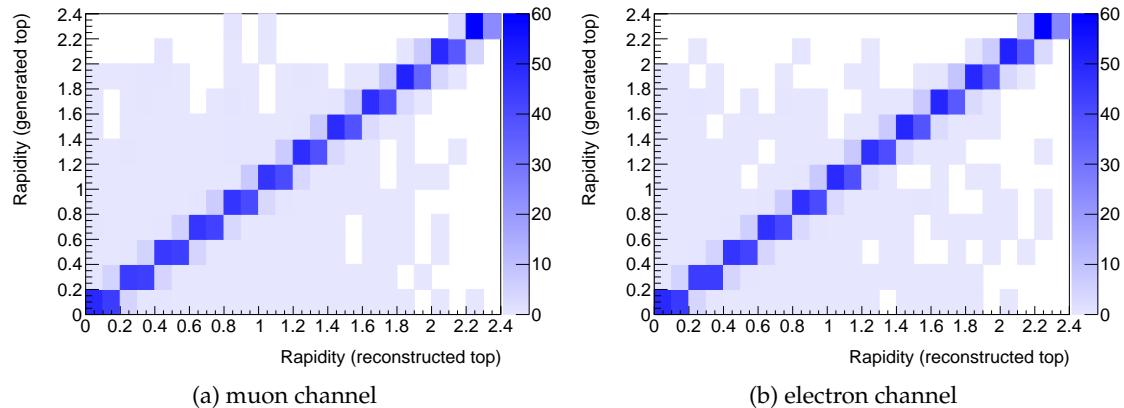


Figure 25: **Parton level:** Response matrices mapping the y of the top-tagged jet to the top quark y in the muon and electron channels. Response matrices are normalized to 100% in each bin of top quark y .

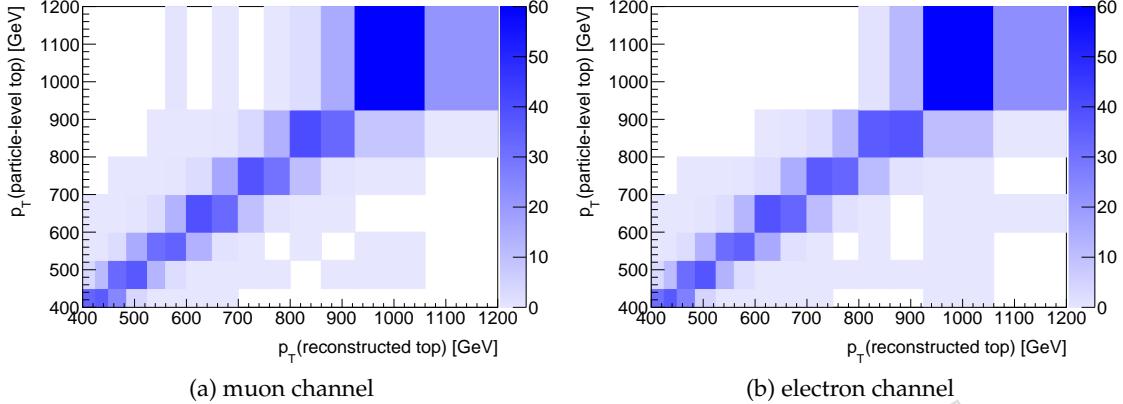


Figure 26: **Particle level:** Response matrices mapping the p_T of the top-tagged jet to the particle-level top jet p_T in the muon and electron channels. Response matrices are normalized to 100% in each bin of particle-level top jet p_T .

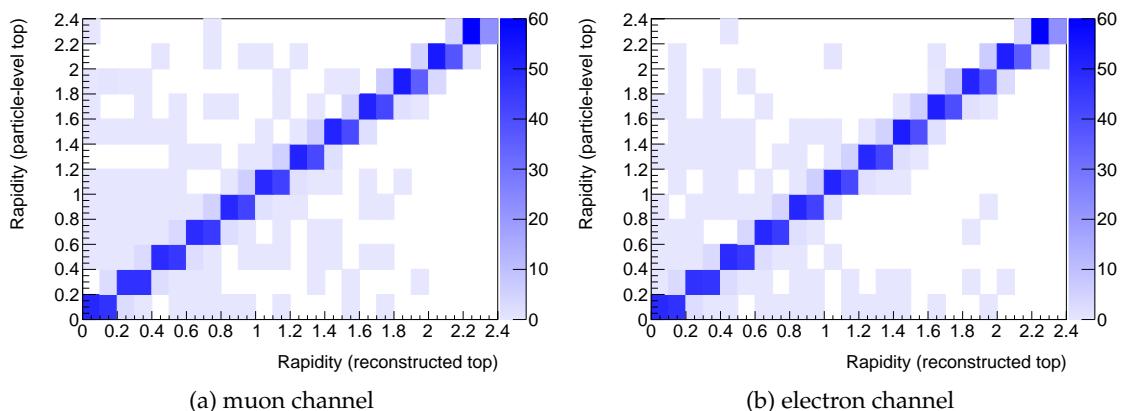


Figure 27: **Particle level:** Response matrices mapping the y of the top-tagged jet to the particle-level top jet y in the muon and electron channels. Response matrices are normalized to 100% in each bin of particle-level top jet y .

$$\mathcal{L}(x, \lambda) = \mathcal{L}_1 + \mathcal{L}_2 + \mathcal{L}_3 \quad (1)$$

$$\mathcal{L}_1 = (y - Ax)^T V_{yy}^{-1} (y - Ax) \quad (2)$$

$$\mathcal{L}_2 = \tau^2 (x - f_b x_0)^T (L^T L) (x - f_b x_0) \quad (3)$$

$$\mathcal{L}_3 = \lambda (Y - e^T x) \quad (4)$$

$$Y = \sum_i y_i \quad (5)$$

$$e_j = \sum_i A_{ij} \quad (6)$$

667 Here, y is the measured result (with covariance matrix V_{yy}) and x is the unfolded measurement,
 668 with A the response matrix linking them. The term \mathcal{L}_1 describes the least-squares minimiza-
 669 tion, while \mathcal{L}_2 parameterizes the regularization. \mathcal{L}_3 is an optional area constraint. The regular-
 670 ization is described by the regularization strength τ , the regularization mode L , and the bias x_0
 671 and its strength f_b .

672 A number of studies were performed to determine the optimal unfolding procedure. Details
 673 on these studies may be found in Appendix H. From these studies a non-regularized unfolding
 674 is chosen to be optimal, setting $\tau = 0$.

675 9.5 MC studies

676 Closure tests are performed to ensure that the procedure is bias-free. The $t\bar{t}$ MC sample is
 677 divided into two halves based on event ID, with the response matrix taken from one half of
 678 the sample and the top p_T/y distributions taken from the other half. To better measure the
 679 bias, 1000 toy distributions are generated by randomly sampling the measured distribution.
 680 These toy distributions are then unfolded and compared to the truth distribution. The mean
 681 difference with respect to truth represents the bias in the unfolding. The average statistical and
 682 total uncertainties on the unfolded result are also computed.

683 As a further test, the p_T unfolding is performed with top p_T reweighting. Events are reweighted
 684 by $(1.0 + 0.0004 \times p_T)^{\pm 1}$, where p_T refers to the top quark p_T . The reweighted measured distri-
 685 bution is then unfolded with the unweighted response matrix, and compared to the reweighted
 686 truth distribution. 1000 toy distributions are also generated for the reweighted samples, and
 687 the above procedure followed to measure the bias, statistical uncertainty, and total uncertainty
 688 in the unfolding.

689 The unfolded distributions and corresponding biases for the p_T unfolding are shown in Figures
 690 28 and 29 for the muon and electron channels, respectively, for the parton level unfolding, and
 691 in Figures 32 and 33 for the particle level unfolding. Corresponding plots for the rapidity
 692 unfolding are shown in Figures 30, 31, 34, and 35. The unfolding shows a reasonable level
 693 of bias, even when the truth distribution is deliberately perturbed from the response matrix
 694 expectation.

695 9.6 Differential cross section for electron/muon channels

696 The response matrices derived from the full $t\bar{t}$ MC sample are used next, to unfold the back-
 697 ground subtracted data to parton-level. The distributions of number of events in bins of top
 698 quark (particle level top jet) p_T/y are corrected for the bin width and for the theoretical branch-
 699 ing ratio for $t\bar{t} \rightarrow e/\mu + \text{jets}$.

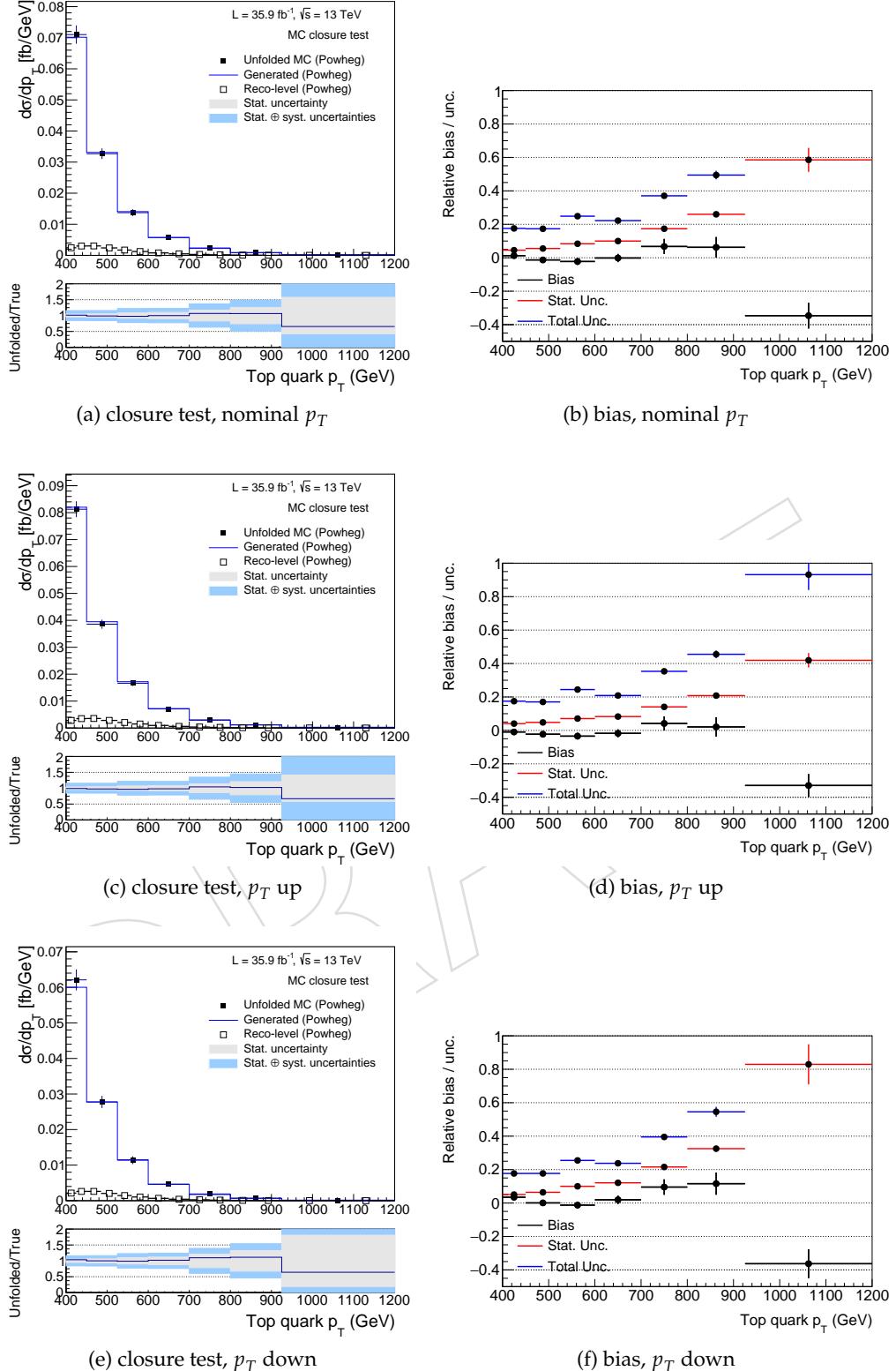


Figure 28: **Parton level / muon channel:** Closure test (left) and bias (right) for the non-regularized p_T unfolding in the muon channel, for the nominal (top), p_T up (middle), and p_T down (bottom) distributions. Closure test is performed by unfolding one half of the MC sample with the other half. Bias, and corresponding average statistical and total uncertainty, are drawn from 1000 unfolded toys.

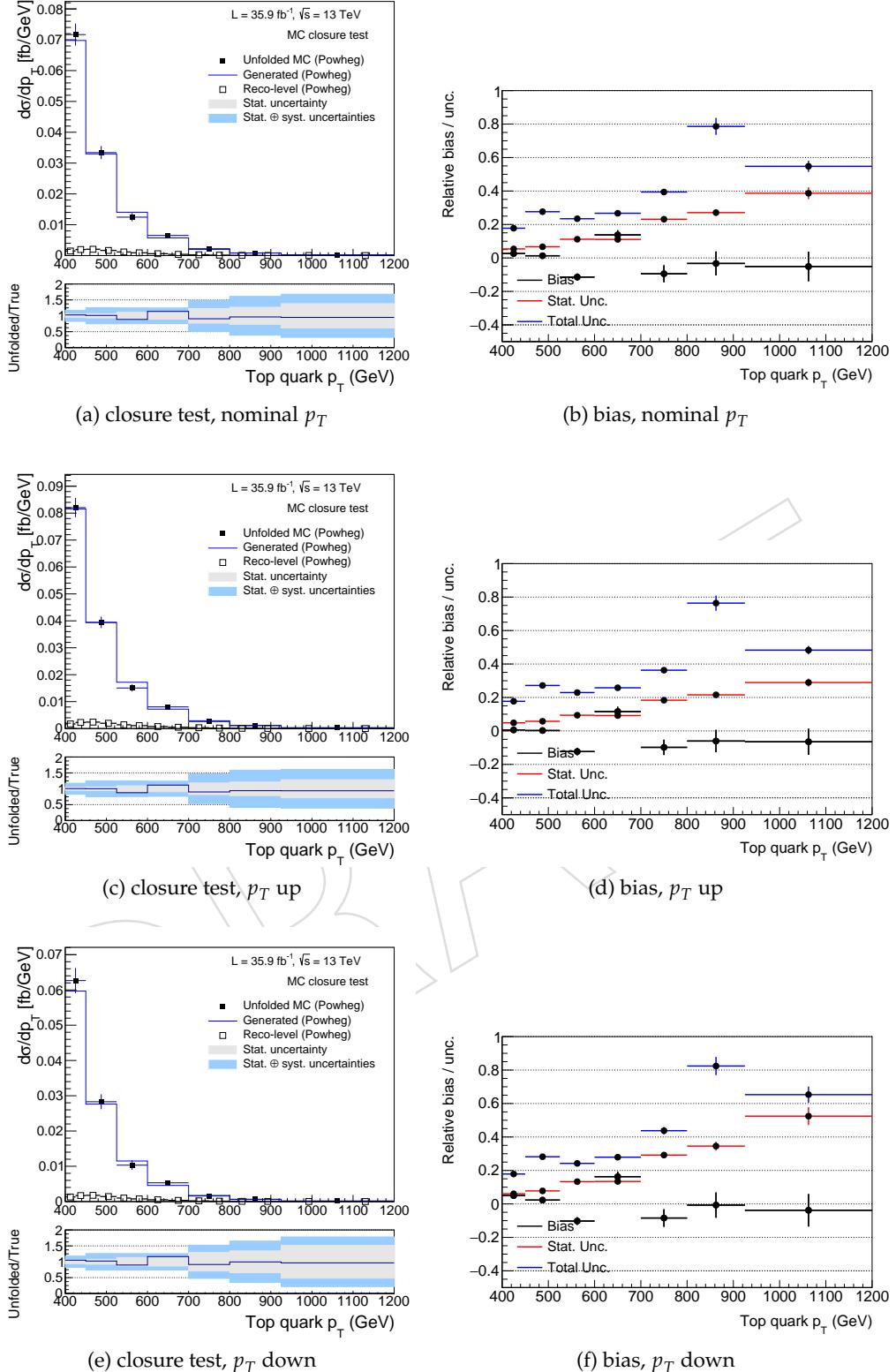


Figure 29: **Parton level / electron channel:** Closure test (left) and bias (right) for the non-regularized p_T unfolding in the electron channel, for the nominal (top), p_T up (middle), and p_T down (bottom) distributions. Closure test is performed by unfolding one half of the MC sample with the other half. Bias, and corresponding average statistical and total uncertainty, are drawn from 1000 unfolded toys.

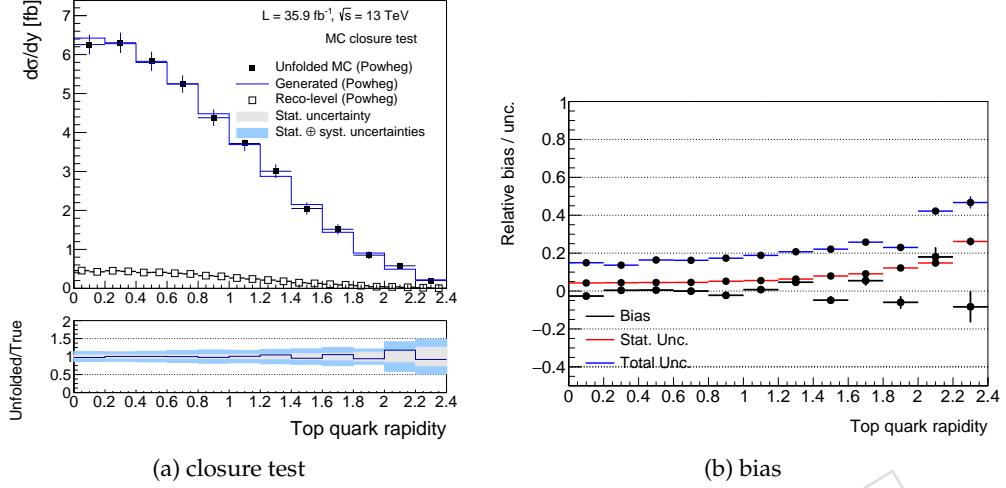


Figure 30: **Parton level / muon channel:** Closure test (left) and bias (right) for the non-regularized rapidity unfolding in the muon channel. Closure test is performed by unfolding one half of the MC sample with the other half. Bias, and corresponding average statistical and total uncertainty, are drawn from 1000 unfolded toys.

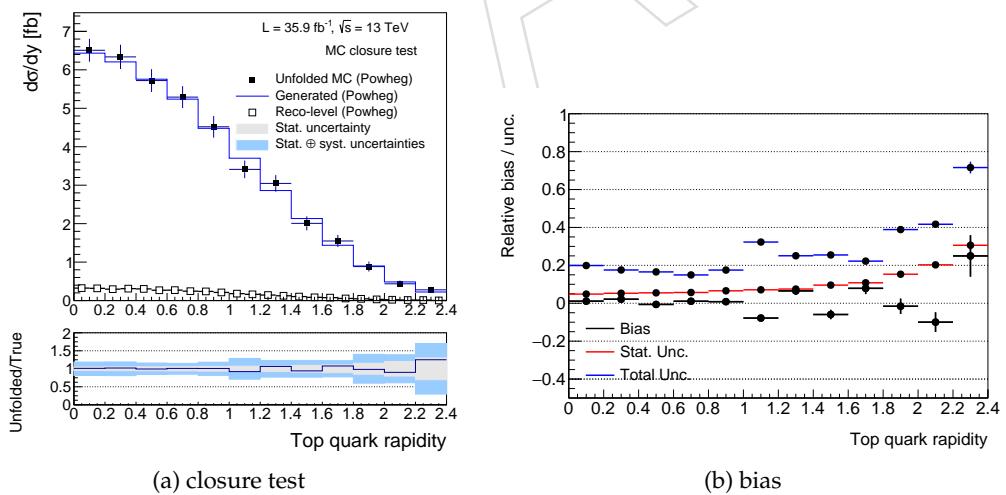


Figure 31: **Parton level / electron channel:** Closure test (left) and bias (right) for the non-regularized rapidity unfolding in the electron channel. Closure test is performed by unfolding one half of the MC sample with the other half. Bias, and corresponding average statistical and total uncertainty, are drawn from 1000 unfolded toys.

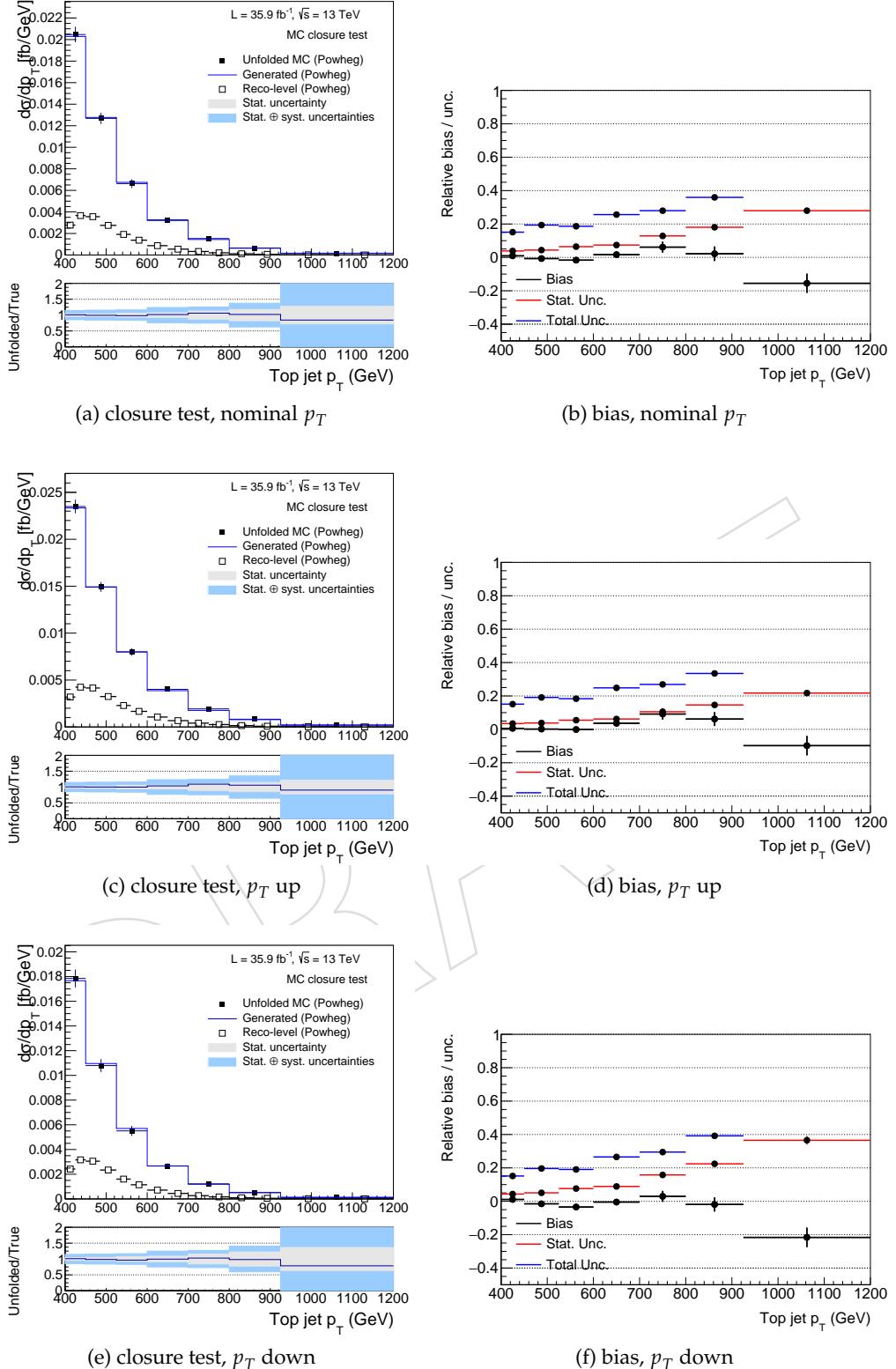


Figure 32: Particle level / muon channel: Closure test (left) and bias (right) for the non-regularized p_T unfolding in the muon channel, for the nominal (top), p_T up (middle), and p_T down (bottom) distributions. Closure test is performed by unfolding one half of the MC sample with the other half. Bias, and corresponding average statistical and total uncertainty, are drawn from 1000 unfolded toys.

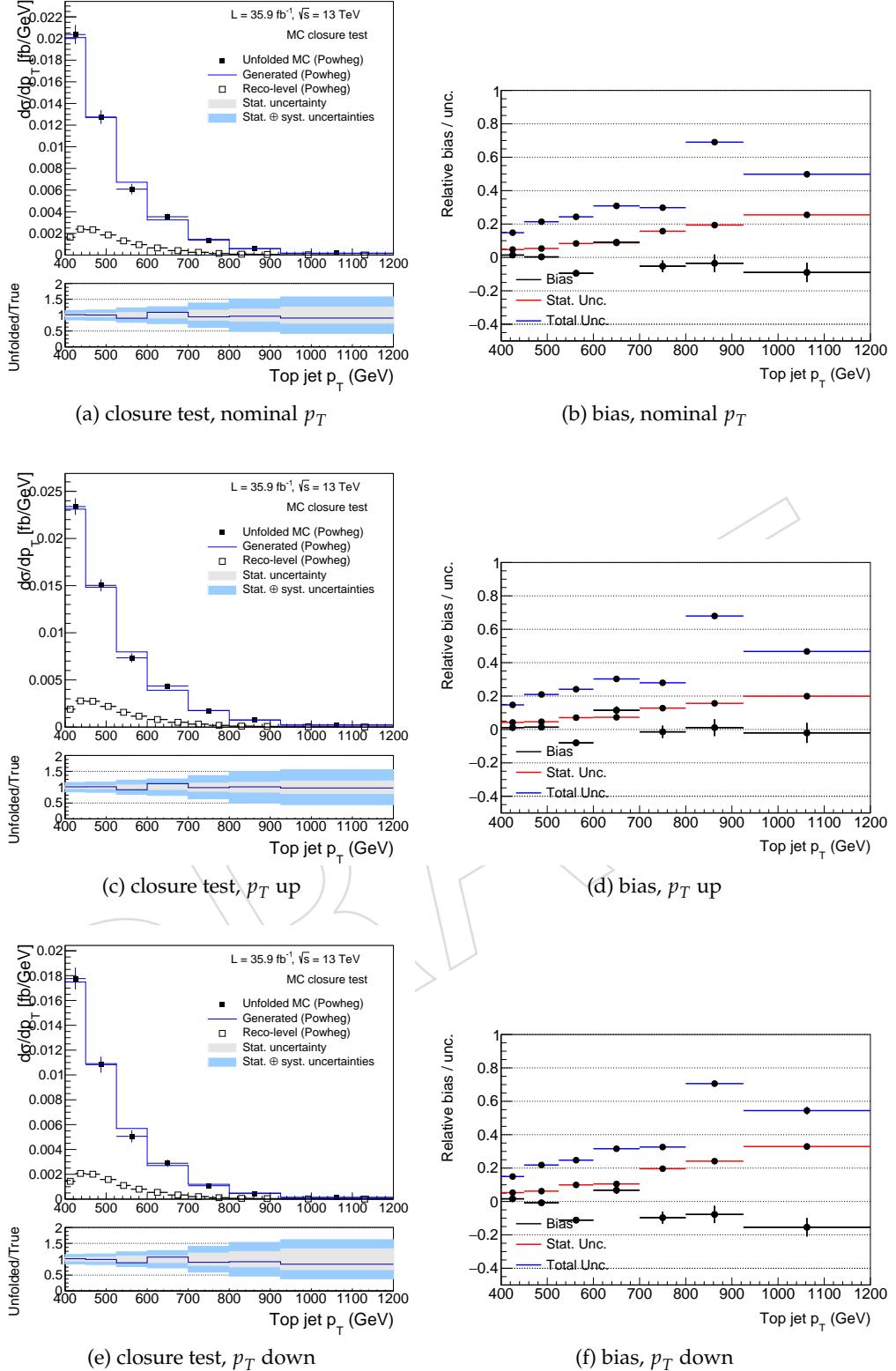


Figure 33: **Particle level / electron channel:** Closure test (left) and bias (right) for the non-regularized p_T unfolding in the electron channel, for the nominal (top), p_T up (middle), and p_T down (bottom) distributions. Closure test is performed by unfolding one half of the MC sample with the other half. Bias, and corresponding average statistical and total uncertainty, are drawn from 1000 unfolded toys.

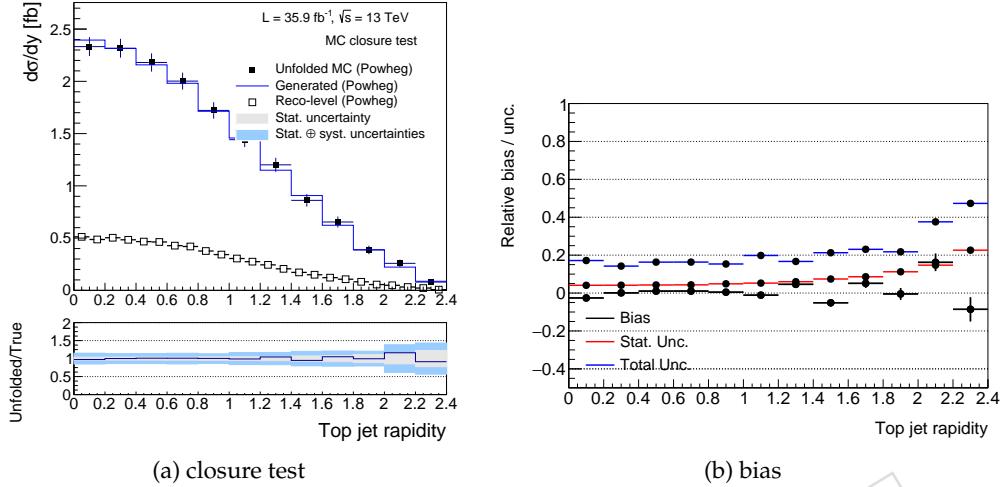


Figure 34: **Particle level / muon channel:** Closure test (left) and bias (right) for the non-regularized rapidity unfolding in the muon channel. Closure test is performed by unfolding one half of the MC sample with the other half. Bias, and corresponding average statistical and total uncertainty, are drawn from 1000 unfolded toys.

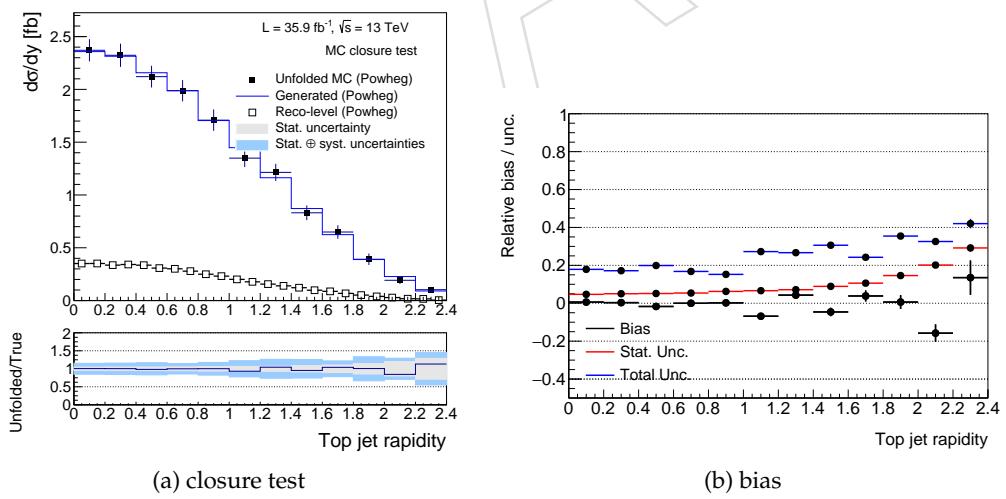


Figure 35: **Particle level / electron channel:** Closure test (left) and bias (right) for the non-regularized rapidity unfolding in the electron channel. Closure test is performed by unfolding one half of the MC sample with the other half. Bias, and corresponding average statistical and total uncertainty, are drawn from 1000 unfolded toys.

700 The propagation of the systematic uncertainties is handled by TUnfold. For a non-regularized
 701 unfolding with no area constraint, the covariance between bins of the unfolded result is given
 702 as

$$V_{xx} = \left(A^T V_{yy}^{-1} A \right)^{-1} \quad (7)$$

703 where V_{yy} is the covariance between bins of the measured result and A is the normalized re-
 704 sponse matrix. Normalization and statistical uncertainties on the backgrounds contribute to
 705 V_{yy} , as do statistical uncertainties on the measured result in data. Systematic and statistical
 706 uncertainties on A are propagated to V_{xx} through partial derivatives. Details on the treatment
 707 of systematic errors may be found in [39].

708 The unfolded distributions at parton level for the muon and electron channels are shown in
 709 Figure 36. Breakdowns of the systematic uncertainties are shown in Figure 37.

710 The corresponding particle-level results are shown in Figure 38 and 39.

711 Figure 40 and 41 shows the total correlation matrices between bins of the unfolded result for
 712 the muon and electron channels at parton and particle levels. The correlation matrix C_{xx} is
 713 related to the covariance matrix V_{xx} via:

$$C_{xx}^{ij} = \frac{V_{xx}^{ij}}{\sqrt{V_{xx}^{ii} * V_{xx}^{jj}}} \quad (8)$$

714 9.7 Differential cross section for combined channels

715 The combination of the electron and muon channels is performed by merging the muon and
 716 electron measured distributions and response matrices. The unfolding is then performed using
 717 the merged measured distribution and response matrices. Merged response matrices are used
 718 for the systematic uncertainties, and merged backgrounds are subtracted from the measured
 719 data. Muon channel and electron channel QCD are treated as separate sources of background,
 720 and maintain their original rates and uncertainties.

721 The unfolded cross section results are shown in Figs. 42 and 43 as a function of p_T and $|y|$,
 722 respectively. The measurements are compared to the prediction from POWHEG +PYTHIA 8. The
 723 breakdown of the sources of systematic uncertainty are shown in Figs. 44 and 45.

724 The total covariance matrix is formed by adding the non-normalized covariance matrices for
 725 the muon and electron channels. The total covariance matrix is then normalized as described
 726 above. The resultant correlation matrices are shown in Figure 46 (parton level) and 47 (particle
 727 level).

728 The unfolded results, with uncertainties, are given in Tables 16 and 17 for parton and particle
 729 level for unfolded results vs p_T and in Tables 18 and 19 for unfolded results vs y .

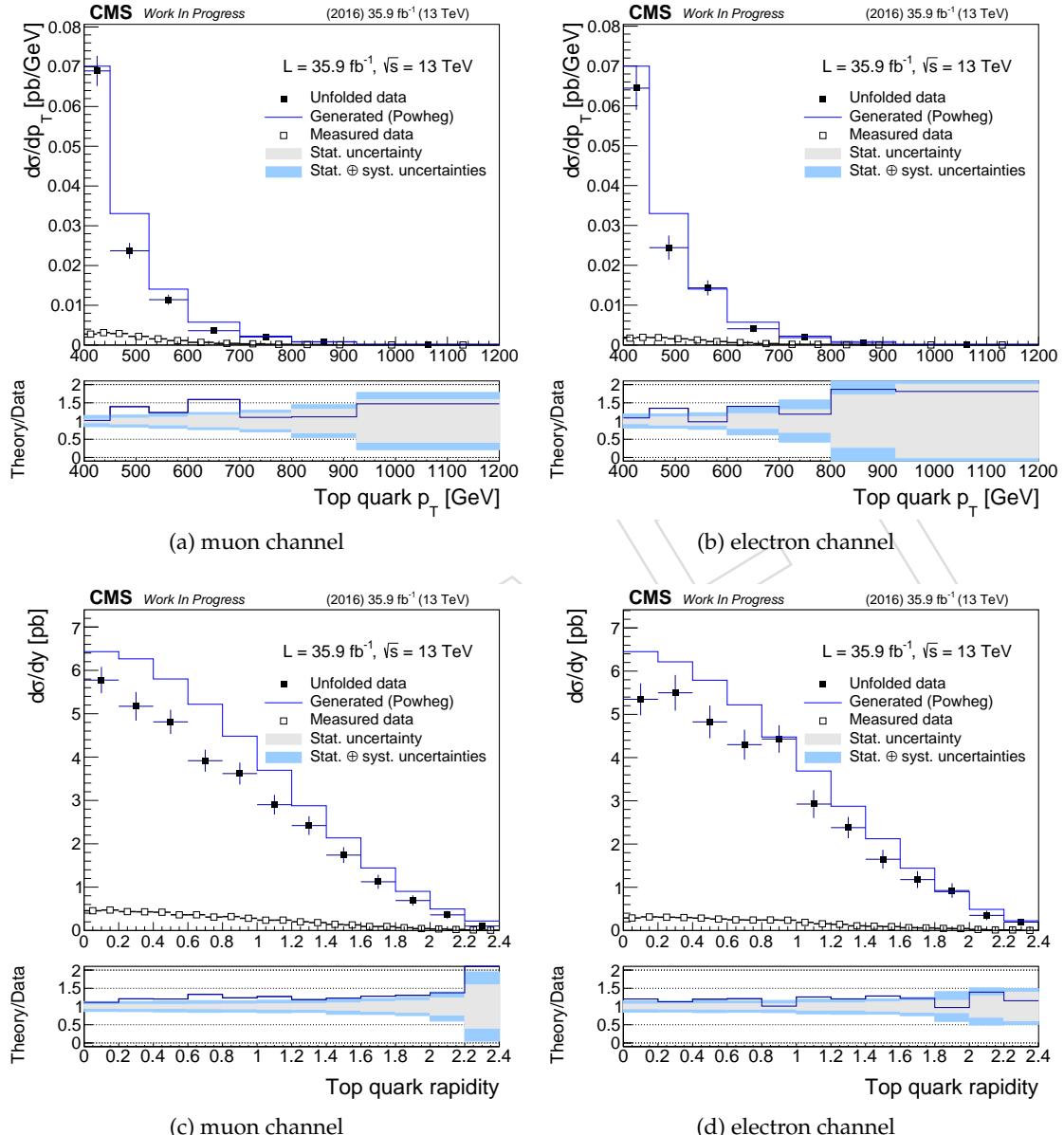


Figure 36: **Parton level:** Cross section in bins of the hadronic top quark p_T and y , unfolded to full partonic phase space, including all systematic uncertainties, for (a/c) μ +jets and (b/d) e+jets final states.

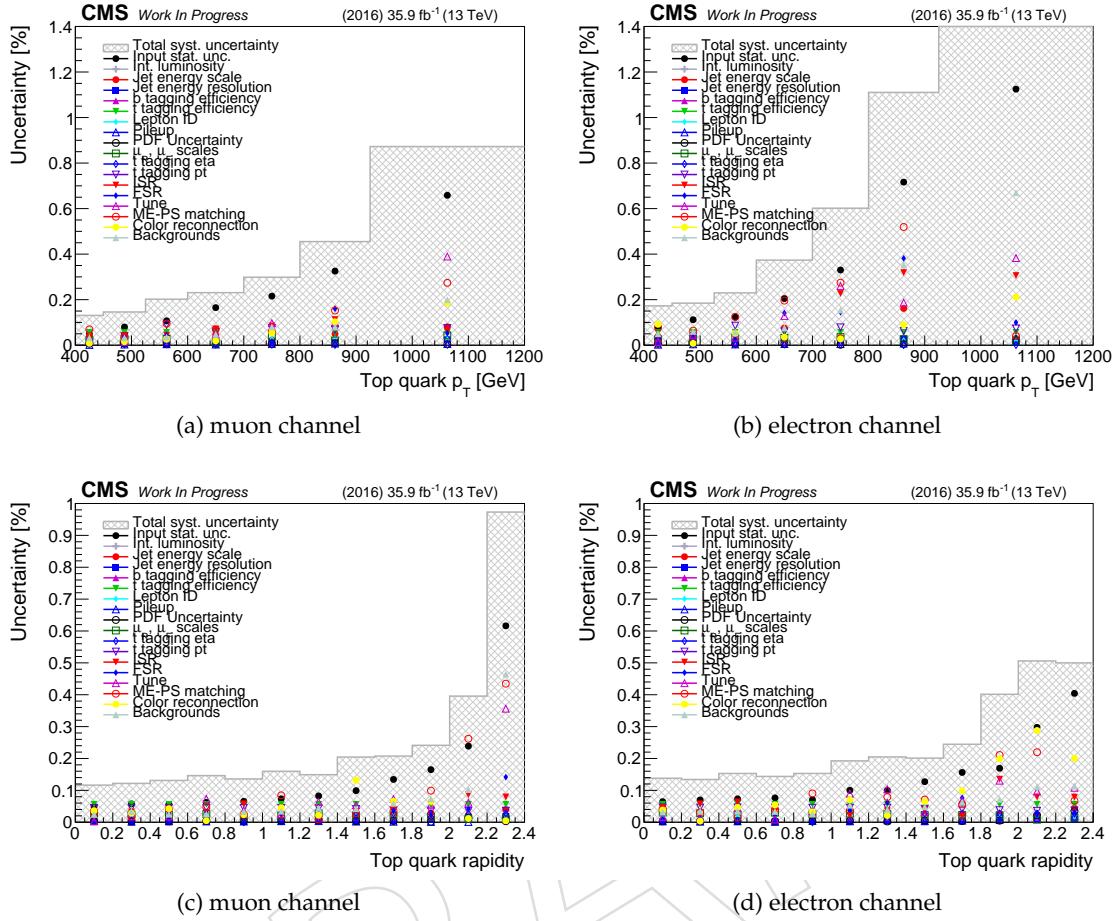


Figure 37: **Parton level:** Different sources of systematic uncertainty on the cross section as a function of parton-level top quark p_T and y for (a/c) the muon and (b/d) the electron channel.

p_T [GeV]	d σ /d p_T [fb/GeV] at parton-level (combined result)						
	data	stat. [%]	exp. [%]	th. [%]	lumi [%]	total [%]	PowhegPythia8
[400.0, 450.0]	66.48	4.09	8.27	8.90	2.50	13.06	70.04
[450.0, 525.0]	23.60	6.61	9.22	7.02	2.50	13.57	33.03
[525.0, 600.0]	12.50	8.10	11.41	10.34	2.50	17.58	14.05
[600.0, 700.0]	3.70	13.14	11.36	10.22	2.50	20.31	5.73
[700.0, 800.0]	2.02	17.63	13.11	21.89	2.50	31.11	2.23
[800.0, 925.0]	0.57	35.63	20.96	28.89	2.50	50.49	0.86
[925.0, 1200.0]	0.12	59.45	33.43	33.72	2.50	76.12	0.21

Table 16: Cross section in bins of parton-level top quark p_T . The data is compared to the prediction from POWHEG+PYTHIA8 $t\bar{t}$ Monte Carlo. The uncertainties on the unfolded data are separated into statistical, experimental systematics, theoretical, and luminosity uncertainties.

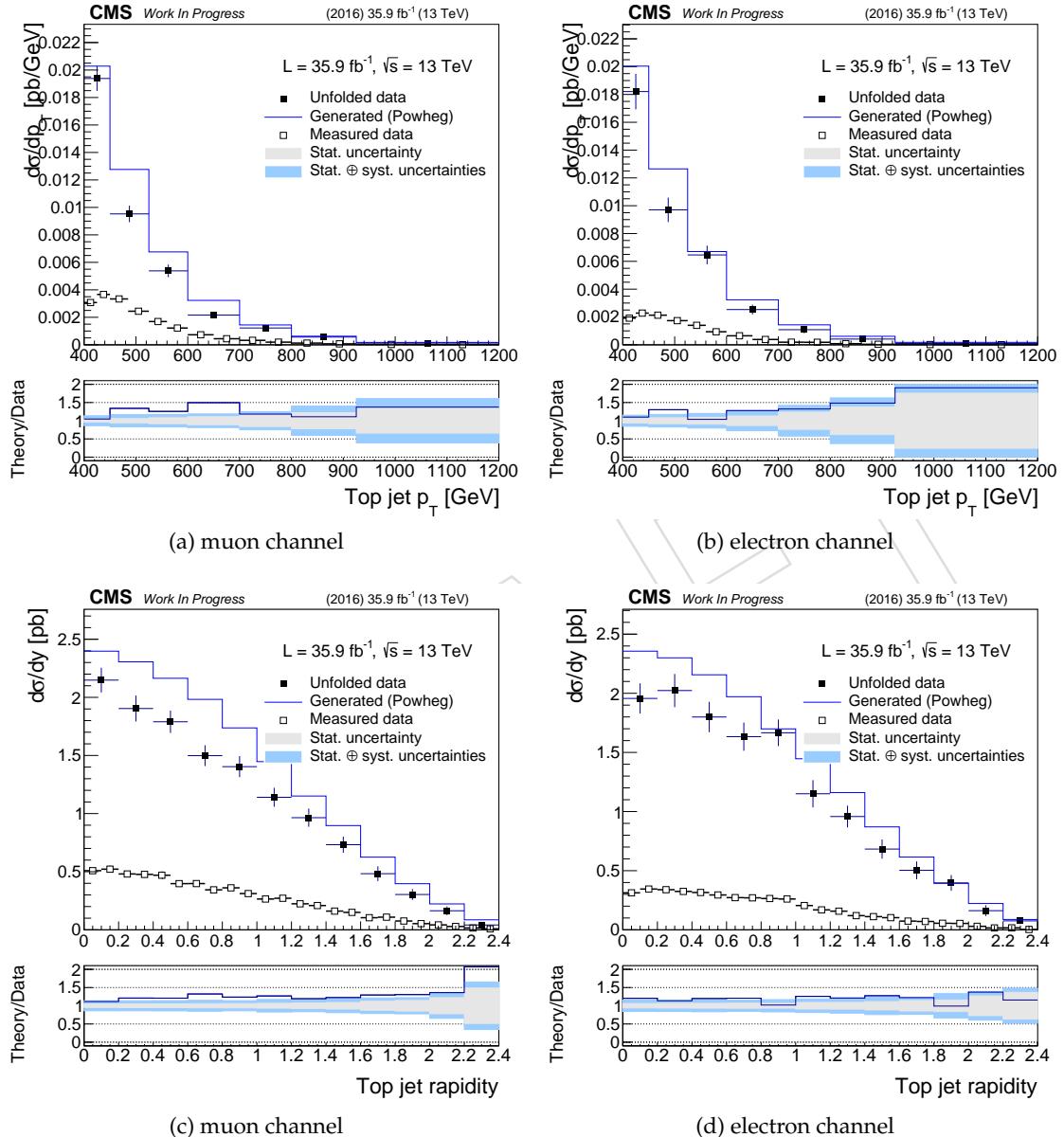


Figure 38: **Particle level:** Cross section in bins of the particle-level top jet p_T and y , unfolded to the fiducial particle level phase space, including all systematic uncertainties, for (a/c) μ +jets and (b/d) e +jets final states.

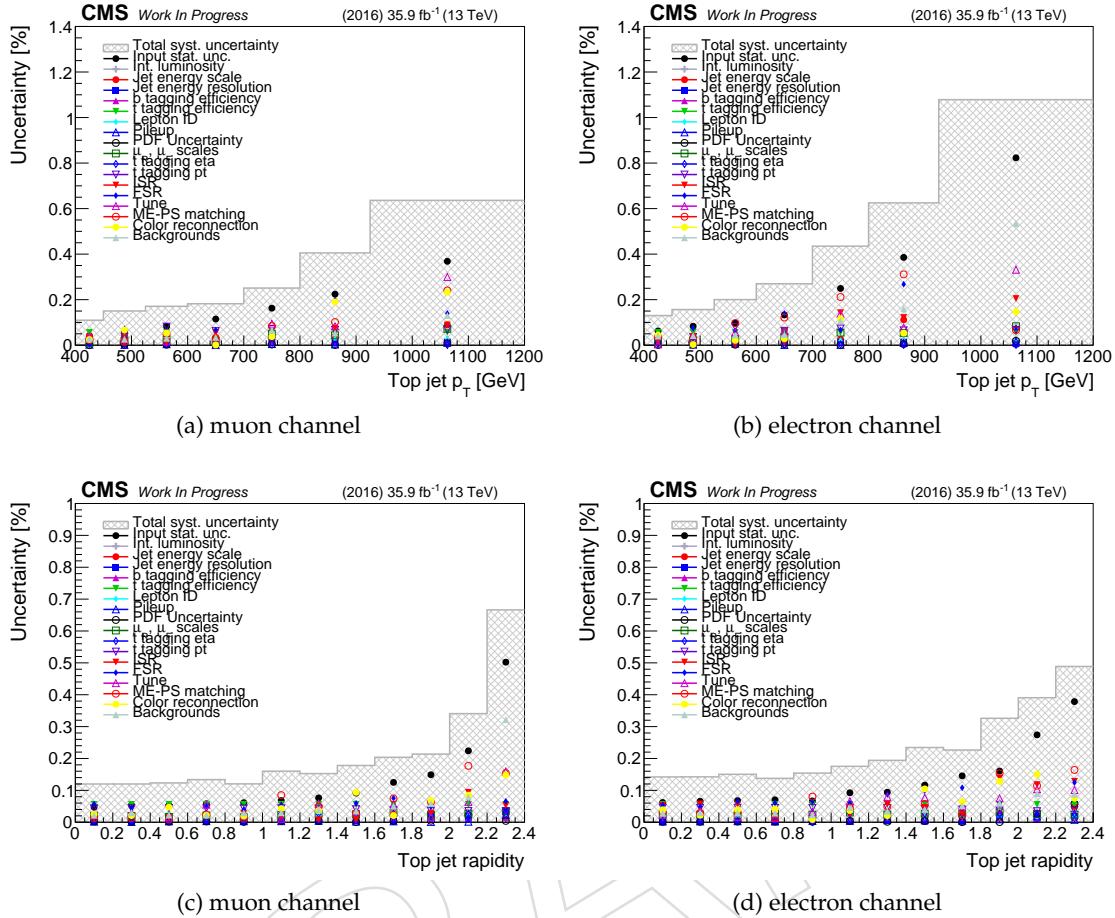


Figure 39: **Particle level:** Different sources of systematic uncertainty on the cross section as a function of particle-level top jet p_T and y for (a/c) the muon and (b/d) the electron channel.

p_T [GeV]	d σ /d p_T [fb/GeV] at particle-level (combined result)						
	data	stat. [%]	exp. [%]	th. [%]	lumi [%]	total [%]	PowhegPythia8
[400.0, 450.0]	18.68	3.49	8.17	4.08	2.50	10.09	20.16
[450.0, 525.0]	9.46	4.91	8.54	8.93	2.50	13.52	12.71
[525.0, 600.0]	5.75	6.35	10.73	9.14	2.50	15.66	6.73
[600.0, 700.0]	2.27	8.84	10.44	11.74	2.50	18.20	3.23
[700.0, 800.0]	1.18	13.46	11.67	17.56	2.50	25.14	1.44
[800.0, 925.0]	0.47	21.38	14.28	26.74	2.50	37.18	0.62
[925.0, 1200.0]	0.10	39.32	25.93	27.86	2.50	54.78	0.17

Table 17: Cross section in bins of particle-level top jet p_T . The data is compared to the prediction from POWHEG+PYTHIA8 $t\bar{t}$ Monte Carlo. The uncertainties on the unfolded data are separated into statistical, experimental systematics, theoretical, and luminosity uncertainties.

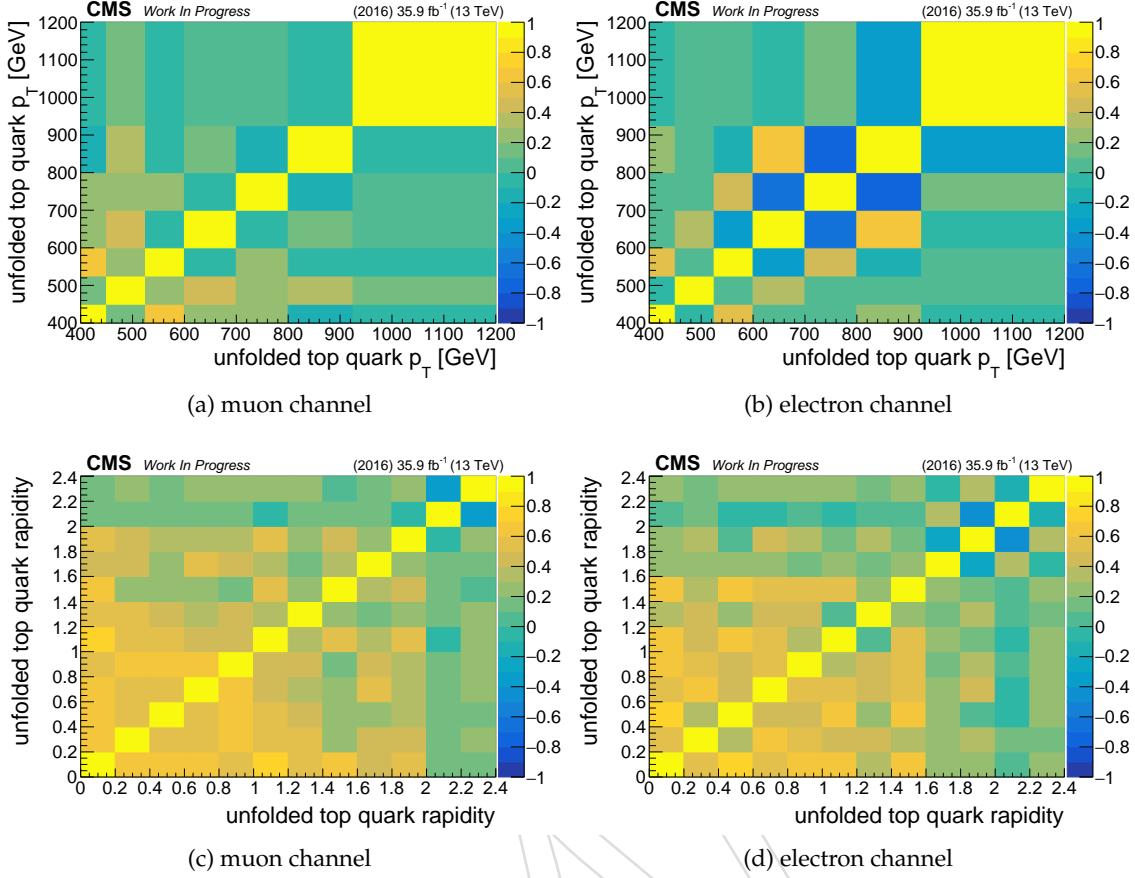


Figure 40: **Parton level:** Total correlation matrix between bins for the unfolded result in (a/c) the muon and (b/d) the electron channel for top quark p_T and y . Correlation matrix is normalized to have a maximum correlation of 100%.

y	d σ /dy [fb] at parton-level (combined result)						
	data	stat. [%]	exp. [%]	th. [%]	lumi [%]	total [%]	PowhegPythia8
[0.0, 0.2]	5519.22	3.94	8.06	6.46	2.50	11.33	6441.16
[0.2, 0.4]	5230.29	4.45	8.21	4.86	2.50	10.81	6240.35
[0.4, 0.6]	4769.04	4.46	8.12	6.09	2.50	11.36	5795.48
[0.6, 0.8]	4040.88	4.83	8.08	6.94	2.50	11.96	5219.17
[0.8, 1.0]	3897.07	4.79	8.09	6.94	2.50	11.95	4479.15
[1.0, 1.2]	2870.20	6.00	8.88	10.28	2.50	15.05	3692.71
[1.2, 1.4]	2352.57	6.38	8.87	6.49	2.50	12.95	2875.90
[1.4, 1.6]	1684.24	7.84	8.94	13.13	2.50	17.89	2129.15
[1.6, 1.8]	1136.01	10.19	9.42	7.46	2.50	15.95	1441.36
[1.8, 2.0]	772.55	11.88	10.09	15.54	2.50	22.15	900.36
[2.0, 2.2]	354.44	18.60	11.60	19.86	2.50	29.68	492.38
[2.2, 2.4]	136.12	34.76	24.72	25.73	2.50	49.87	221.84

Table 18: Cross section in bins of parton-level top quark p_T . The data is compared to the prediction from POWHEG+PYTHIA8 $t\bar{t}$ Monte Carlo. The uncertainties on the unfolded data are separated into statistical, experimental systematics, theoretical, and luminosity uncertainties.

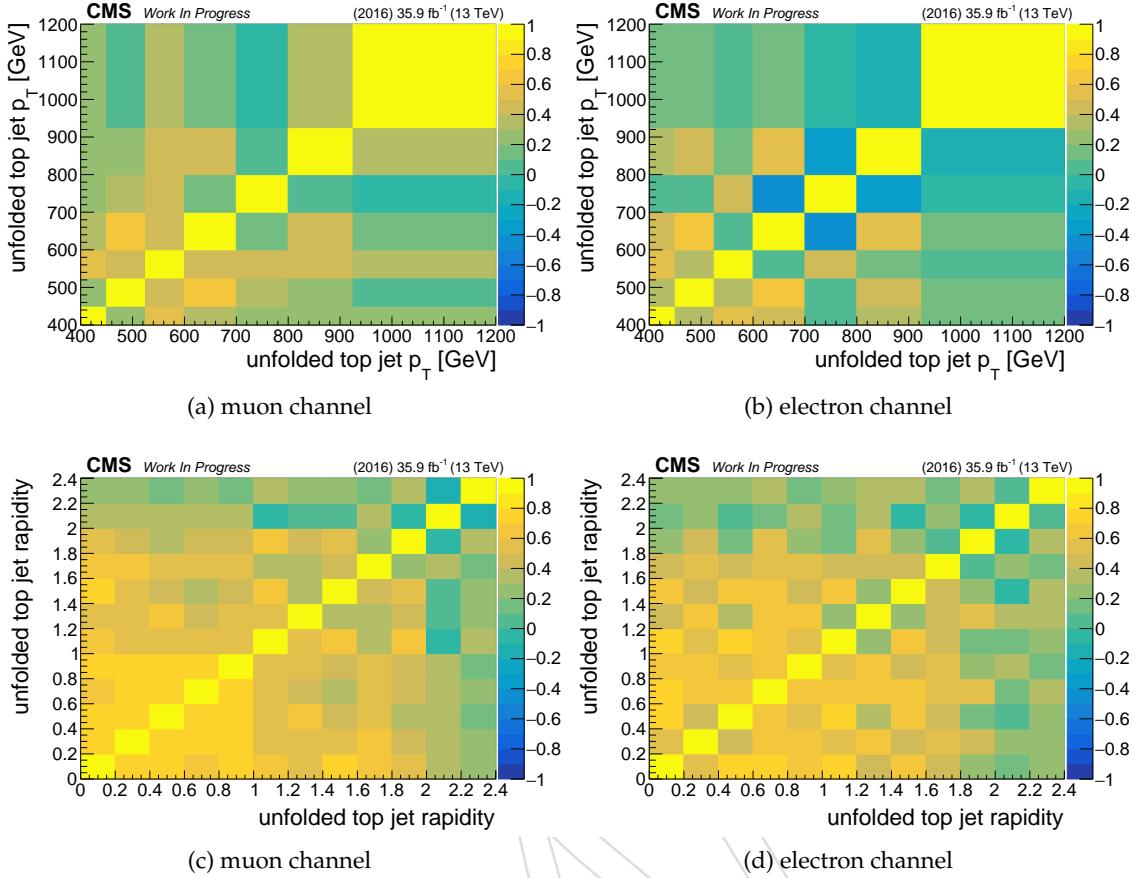


Figure 41: **Particle level:** Total correlation matrix between bins for the unfolded result in (a/c) the muon and (b/d) the electron channel for top quark p_T and y . Correlation matrix is normalized to have a maximum correlation of 100%.

y	d σ /dy [fb] at particle-level (combined result)						
	data	stat. [%]	exp. [%]	th. [%]	lumi [%]	total [%]	PowhegPythia8
[0.0, 0.2]	2036.81	3.77	7.90	7.06	2.50	11.52	2376.74
[0.2, 0.4]	1925.25	4.15	8.04	6.72	2.50	11.54	2302.77
[0.4, 0.6]	1774.95	4.16	7.96	6.97	2.50	11.64	2160.22
[0.6, 0.8]	1539.66	4.48	7.94	6.93	2.50	11.72	1976.82
[0.8, 1.0]	1486.72	4.50	7.94	6.98	2.50	11.76	1718.00
[1.0, 1.2]	1129.21	5.53	8.60	10.84	2.50	15.11	1446.80
[1.2, 1.4]	944.80	5.93	8.61	7.63	2.50	13.18	1155.81
[1.4, 1.6]	702.74	7.21	8.58	13.40	2.50	17.65	883.50
[1.6, 1.8]	486.52	9.48	8.96	10.77	2.50	17.10	620.24
[1.8, 2.0]	335.60	10.96	9.49	12.91	2.50	19.57	394.85
[2.0, 2.2]	161.22	17.42	10.31	21.55	2.50	29.67	222.58
[2.2, 2.4]	54.75	30.11	18.59	21.77	2.50	41.62	86.75

Table 19: Cross section in bins of particle-level top jet p_T . The data is compared to the prediction from POWHEG+PYTHIA8 $t\bar{t}$ Monte Carlo. The uncertainties on the unfolded data are separated into statistical, experimental systematics, theoretical, and luminosity uncertainties.

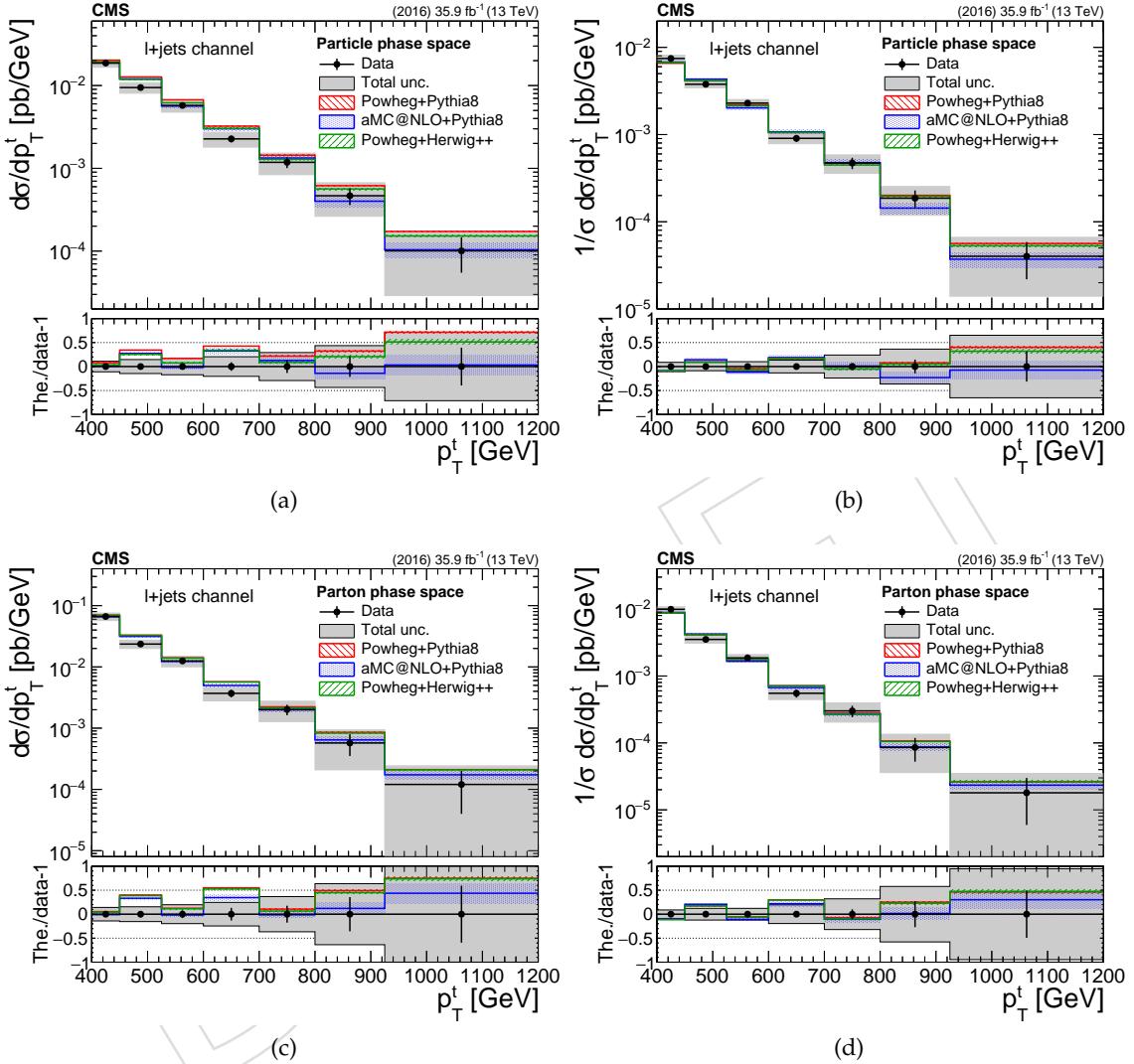


Figure 42: Differential cross section measurements at particle level (top row) and parton level (bottom row) as a function of the particle-level t jet p_T or the parton-level top quark p_T for the ℓ +jets analysis. Both absolute (left column) and normalized (right column) cross sections are shown. The uncertainties on the data markers are the total statistical uncertainty, while the grey band shows the total statistical and systematic uncertainty.

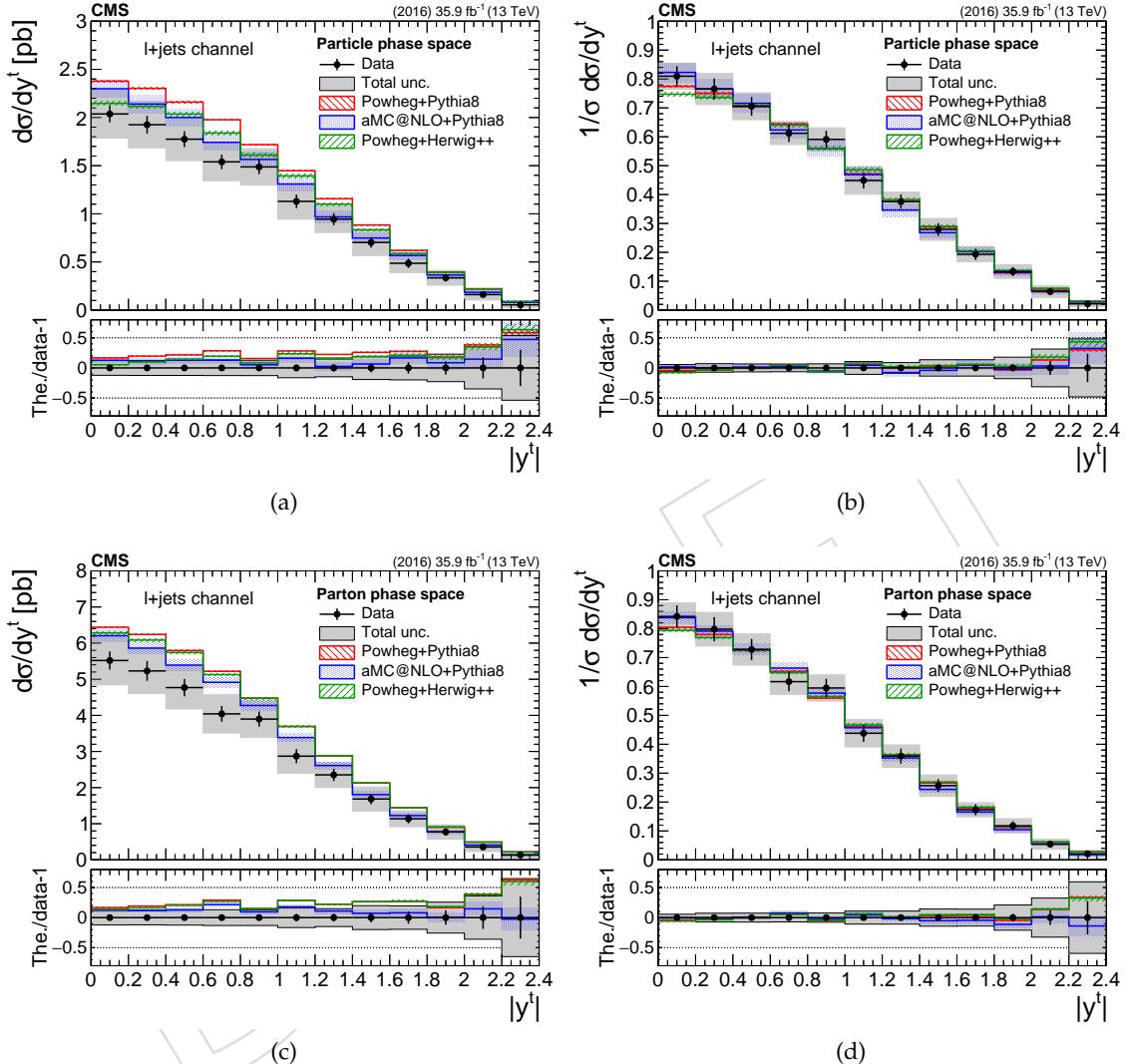


Figure 43: Differential cross section measurements at particle level (top row) and parton level (bottom row) as a function of the particle-level t jet $|y|$ or the parton-level top quark $|y|$ for the ℓ +jets analysis. Both absolute (left column) and normalized (right column) cross sections are shown. The uncertainties on the data markers are the total statistical uncertainty, while the grey band shows the total statistical and systematic uncertainty.

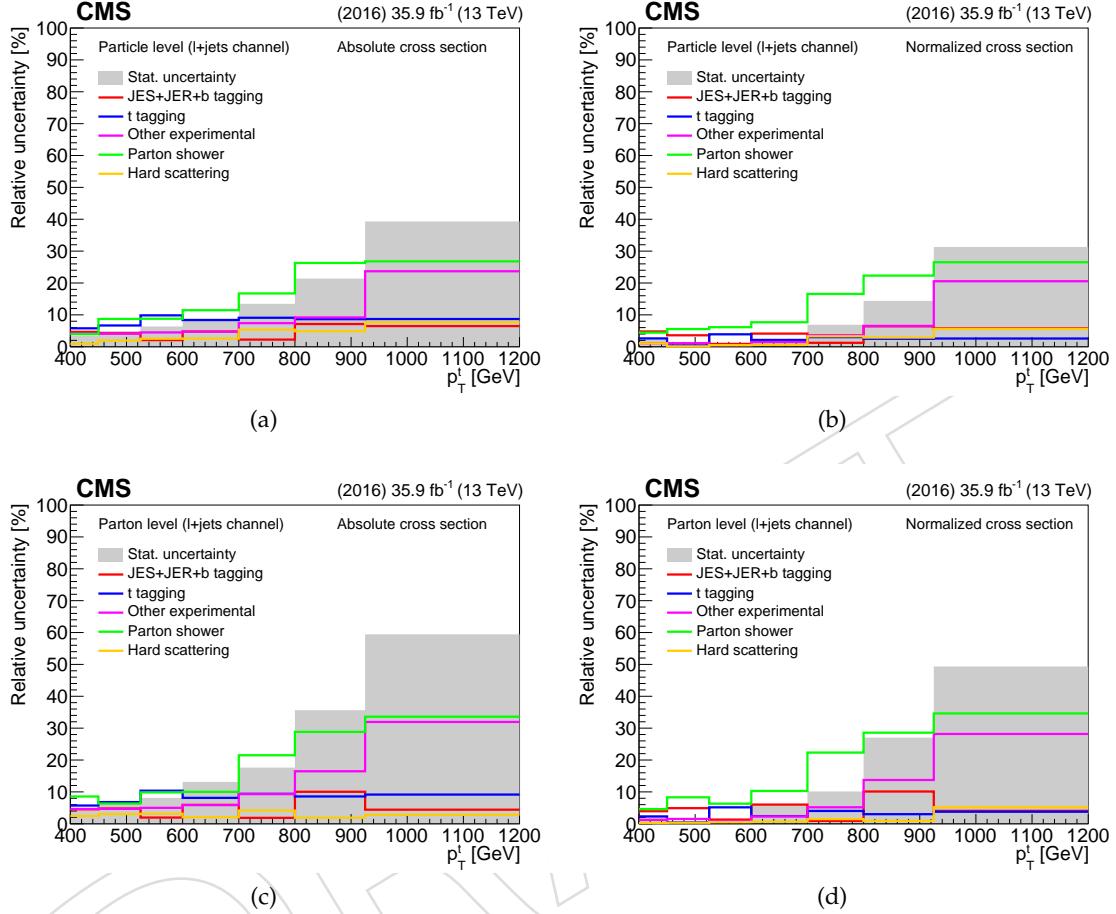


Figure 44: Breakdown of the sources of systematic uncertainty affecting the differential cross section measurements in the ℓ +jets channel at the particle level (top row) and at the parton level (bottom row) as a function of the particle-level t jet p_T or the parton-level top quark p_T . Both the systematic uncertainties in the absolute (left column) and the normalized (right column) cross sections are shown. "JES+JER+b tagging" includes uncertainties due to the jet energy scale, jet energy resolution, and small- R jet b tagging efficiency; " t tagging" is the uncertainty associated with the large- R jet t tagging efficiency; "Other experimental" includes the uncertainties originating from the background estimate, pileup modeling, lepton identification and trigger efficiency, and measurement of the integrated luminosity; "Parton shower" includes contributions from initial and final state radiation, underlying event tune, ME-PS matching, and color reconnection; "Hard scattering" includes the uncertainty due to PDFs as well as renormalization and factorization scales.

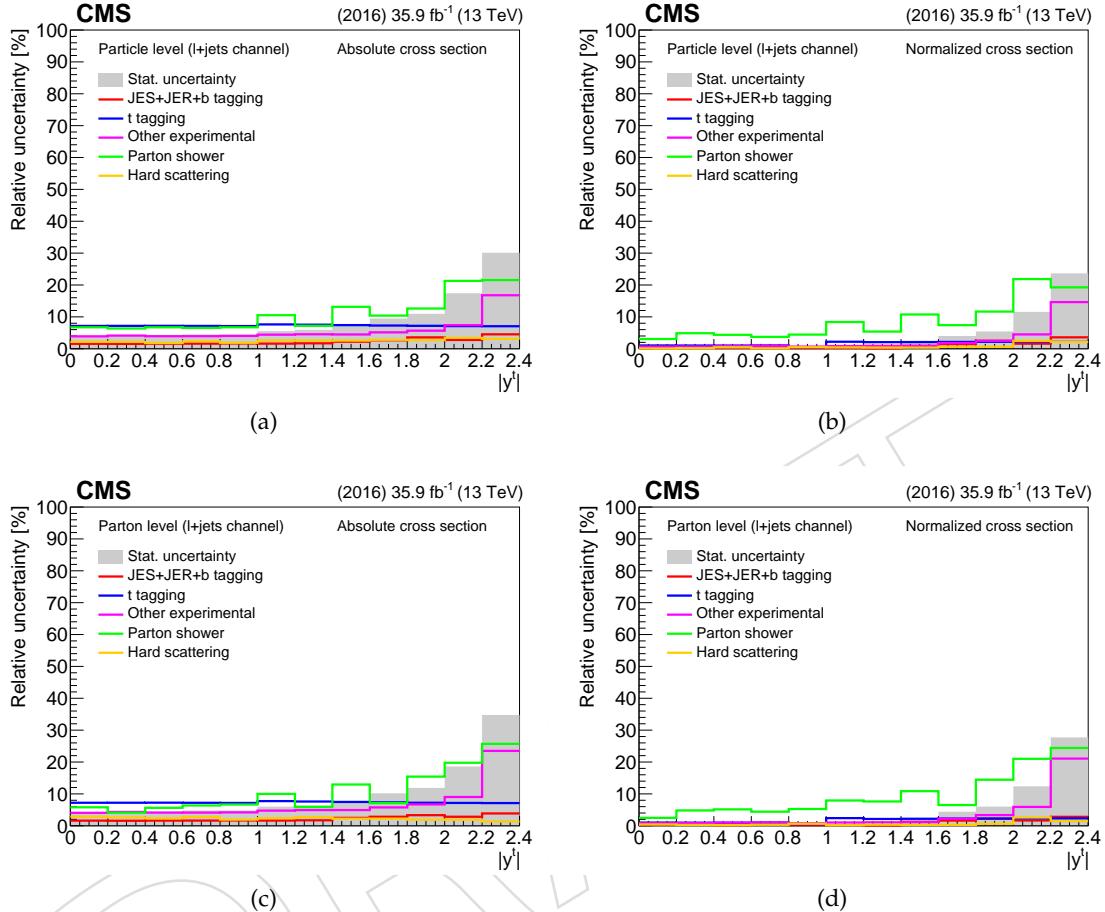


Figure 45: Breakdown of the sources of systematic uncertainty affecting the differential cross section measurements in the ℓ +jets channel at the particle level (top row) and at the parton level (bottom row) as a function of the particle-level t jet $|y|$ or the parton-level top quark $|y'|$. Both the systematic uncertainties in the absolute (left column) and the normalized (right column) cross sections are shown. "JES+JER+b tagging" includes uncertainties due to the jet energy scale, jet energy resolution, and small- R jet b tagging efficiency; " t tagging" is the uncertainty associated with the large- R jet t tagging efficiency; "Other experimental" includes the uncertainties originating from the background estimate, pileup modeling, lepton identification and trigger efficiency, and measurement of the integrated luminosity; "Parton shower" includes contributions from initial and final state radiation, underlying event tune, ME-PS matching, and color reconnection; "Hard scattering" includes the uncertainty due to PDFs as well as renormalization and factorization scales.

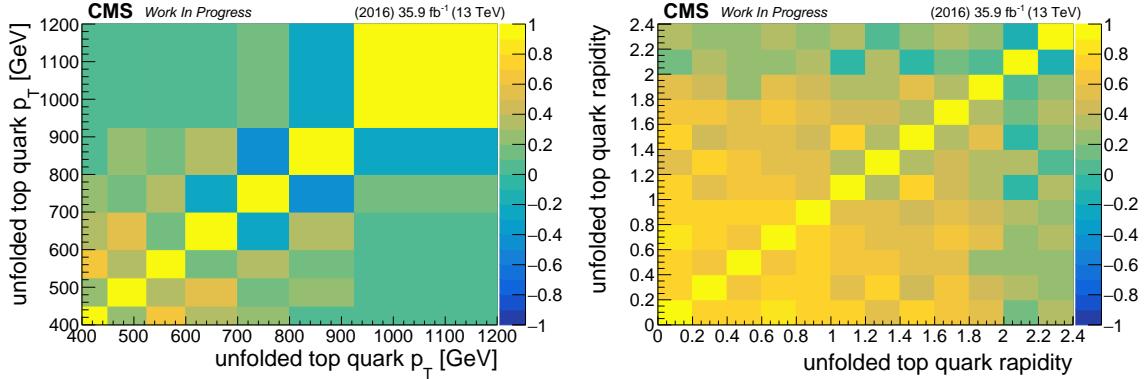


Figure 46: **Parton level:** Total correlation matrix between bins for the unfolded result in the combined channel for the parton level p_T (left) and y (right) results. Correlation matrix is normalized to have a maximum correlation of 100%.

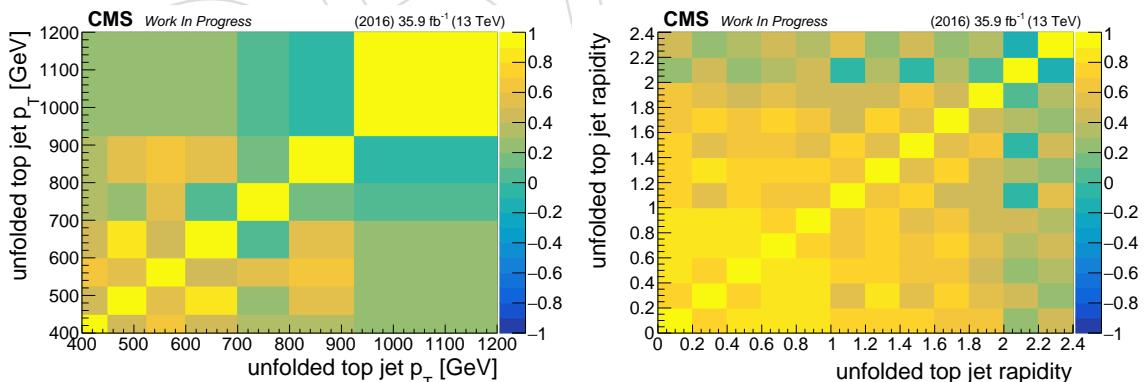


Figure 47: **Particle level:** Total correlation matrix between bins for the unfolded result in the combined channel for the particle level p_T (left) and y (right) results. Correlation matrix is normalized to have a maximum correlation of 100%.

730 10 Conclusions

731 This analysis note presents a measurement of the $t\bar{t}$ production cross section in the boosted
732 regime at 13 TeV using CMS data. The differential cross section is measured as a function of
733 the p_T/y of the hadronically decaying top quark for the region with $p_T(\text{top}) > 400$ GeV at
734 particle and parton levels. The analysis uses the lepton (e/μ)+jets final states, as identified
735 through an electron/muon and b-tagged jet from the leptonically decaying top quark, and a
736 single top-tagged jet corresponding to the hadronically decaying top quark. By taking advan-
737 tage of the rich event topology, backgrounds are greatly reduced. The remaining backgrounds
738 are modeled using MC for the shapes (data sideband for QCD production) while the back-
739 ground normalizations are extracted jointly with the top tagging efficiency through likelihood
740 template fits to three kinematic regions. Measured $t\bar{t}$ distributions are constructed by subtract-
741 ing backgrounds from data, using the appropriate fitted background normalizations and top
742 tagging efficiency.

743 The differential cross section as a function of top quark p_T/y is extrapolated from the measured
744 top jet p_T/y distribution through an unfolding procedure. The unfolding corrects for signal
745 acceptance and bin migrations, extrapolating to either a fiducial region at particle level, or to
746 the full phase space at parton level. The measurements are compared to the theory prediction
747 from POWHEG MC which is observed to overestimate the data.

DRAFT

748 References

- 749 [1] CMS Collaboration, "Measurement of the integrated and differential t-tbar production
750 cross sections for high-pt top quarks in pp collisions at $\sqrt{s} = 8$ TeV",
751 arXiv:1605.00116.
- 752 [2] ATLAS Collaboration, "Measurement of the differential cross-section of highly boosted
753 top quarks as a function of their transverse momentum in $\sqrt{s} = 8$ TeV proton-proton
754 collisions using the ATLAS detector", *Phys. Rev. D* **93** (2016) 032009,
755 doi:10.1103/PhysRevD.93.032009, arXiv:1510.03818.
- 756 [3] CMS Collaboration, "Measurement of the inclusive and differential tt production cross
757 sections in lepton + jets final states at 13 tev", CMS Physics Analysis Summary
758 CMS-PAS-TOP-16-008, 2016.
- 759 [4] P. Nason, "A new method for combining NLO QCD with shower Monte Carlo
760 algorithms", *JHEP* **11** (2004) 040, doi:10.1088/1126-6708/2004/11/040,
761 arXiv:hep-ph/0409146.
- 762 [5] S. Frixione, P. Nason, and C. Oleari, "Matching NLO QCD computations with Parton
763 Shower simulations: the POWHEG method", *JHEP* **11** (2007) 070,
764 doi:10.1088/1126-6708/2007/11/070, arXiv:0709.2092.
- 765 [6] S. Alioli, P. Nason, C. Oleari, and E. Re, "A general framework for implementing NLO
766 calculations in shower Monte Carlo programs: the POWHEG BOX", *JHEP* **06** (2010) 043,
767 doi:10.1007/JHEP06(2010)043, arXiv:1002.2581.
- 768 [7] J. M. Campbell, R. K. Ellis, P. Nason, and E. Re, "Top-pair production and decay at NLO
769 matched with parton showers", *JHEP* **04** (2015) 114,
770 doi:10.1007/JHEP04(2015)114, arXiv:1412.1828.
- 771 [8] T. Sjöstrand, S. Mrenna, and P. Z. Skands, "PYTHIA 6.4 Physics and Manual", *JHEP* **05**
772 (2006) 026, doi:10.1088/1126-6708/2006/05/026, arXiv:hep-ph/0603175.
- 773 [9] T. Sjöstrand, S. Mrenna, and P. Z. Skands, "A Brief Introduction to PYTHIA 8.1", *Comput.
774 Phys. Commun.* **178** (2008) 852–867, doi:10.1016/j.cpc.2008.01.036,
775 arXiv:0710.3820.
- 776 [10] P. Skands, S. Carrazza, and J. Rojo, "Tuning PYTHIA 8.1: the Monash 2013 Tune", *Eur.
777 Phys. J.* **C74** (2014), no. 8, 3024, doi:10.1140/epjc/s10052-014-3024-y,
778 arXiv:1404.5630.
- 779 [11] M. Czakon and A. Mitov, "Top++: a program for the calculation of the top-pair
780 cross-section at hadron colliders", *Comput. Phys. Commun.* **185** (2014) 2930,
781 doi:10.1016/j.cpc.2014.06.021, arXiv:1112.5675.
- 782 [12] J. Alwall et al., "The automated computation of tree-level and next-to-leading order
783 differential cross sections, and their matching to parton shower simulations", *JHEP* **07**
784 (2014) 079, doi:10.1007/JHEP07(2014)079, arXiv:1405.0301.
- 785 [13] <https://twiki.cern.ch/twiki/bin/view/CMS/B2GANEDMNTuples80X>.
- 786 [14] <https://cmsweb.cern.ch/das>.
- 787 [15] <https://twiki.cern.ch/twiki/bin/view/LHCPhysics/SingleTopRefXsec>.

- 788 [16] "Cms an-16-020",.
- 789 [17] <https://twiki.cern.ch/twiki/bin/view/Main/EdbrBackup>.
- 790 [18] "Cms an-15-107",.
- 791 [19] "Cms an-17-012",.
- 792 [20] "Cms an-15-136",.
- 793 [21] <https://twiki.cern.ch/twiki/bin/viewauth/CMS/SWGuideMuonIdRun2>.
- 794 [22] <https://twiki.cern.ch/twiki/bin/view/CMS/MiniIsolationSUSY>.
- 795 [23] <https://twiki.cern.ch/twiki/bin/view/CMS/MuonWorkInProgressAndPagResults>.
- 796 [24] <https://twiki.cern.ch/twiki/bin/view/CMS/CutBasedElectronIdentificationRun2>.
- 797 [25] <https://twiki.cern.ch/twiki/bin/viewauth/CMS/SUSLeptonSF>.
- 798 [26] <https://twiki.cern.ch/twiki/bin/viewauth/CMS/JetID>.
- 799 [27] <https://twiki.cern.ch/twiki/bin/viewauth/CMS/JetResolution>.
- 800 [28] <https://twiki.cern.ch/twiki/bin/viewauth/CMS/BtagRecommendation80XReReco>.
- 801 [29] <https://twiki.cern.ch/twiki/bin/viewauth/CMS/BTagSFMethods>.
- 802 [30] "Top-tagging WPs 80X".
803 <https://twiki.cern.ch/twiki/bin/viewauth/CMS/JetTopTagging>.
- 804 [31] <https://twiki.cern.ch/twiki/bin/view/CMS/MissingETOptionalFiltersRun2>.
- 805 [32] <https://twiki.cern.ch/twiki/bin/view/CMS/TWikiLUM>.
- 806 [33]
807 <https://indico.cern.ch/event/459797/contributions/1961581/attachments/1181555/1800214/mcaod->
808 Feb15-2016.pdf.
- 809 [34] <https://twiki.cern.ch/twiki/bin/viewauth/CMS/LHEReaderCMSSW>.
- 810 [35] <https://twiki.cern.ch/twiki/bin/viewauth/CMS/TopSystematics>.
- 811 [36] CMS Collaboration Collaboration, "Investigations of the impact of the parton shower
812 tuning in Pythia 8 in the modelling of $t\bar{t}$ at $\sqrt{s} = 8$ and 13 TeV", Technical Report
813 CMS-PAS-TOP-16-021, CERN, Geneva, 2016.
- [37]
https://indico.cern.ch/event/613637/contributions/2489492/attachments/1419601/2174856/color_eco
- [38] <https://twiki.cern.ch/twiki/bin/viewauth/CMS/SWGuideHiggsAnalysisCombinedLimit>.
- [39] S. Schmitt, "TUnfold, an algorithm for correcting migration effects in high energy physics",
816 (2012) arXiv:1205.6201v4.
- [40] <http://www.desy.de/~sschmitt/tunfold.html>.
- [41] <https://twiki.cern.ch/twiki/bin/viewauth/CMS/TopUnfolding>.
- [42] <https://twiki.cern.ch/twiki/bin/view/CMS/ParticleLevelProducer>.

820 A Lepton Optimization

821 To determine the optimal lepton definition, several studies were made of different options for
 822 the lepton ID and isolation.

823 For ID, the medium and tight working points as defined by the EGamma and Muon POGs
 824 were considered. The relative isolation requirement of the electron IDs was dropped.

825 For isolation, two methods were considered: mini isolation and 2D separation. Mini isolation
 826 is similar to relative isolation, in that it requires the sum p_T of the tracks in a cone around the
 827 lepton to be less than a given fraction of the lepton p_T . In contrast to relative isolation, the cone
 828 size is variable with respect to lepton p_T . Specifically, mini isolation is defined as

$$\text{miniIso} = \frac{1}{p_T^\ell} * [\text{CH} + \max(0, \text{NH} + \text{PH} - 0.5 * \text{PU} * \rho * A_{\text{eff}} * (\frac{R}{0.3})^2)] \quad (9)$$

829 where CH, NH, PH, and PU are the $\sum p_T$ of the charged hadrons, neutral hadrons, photons,
 830 and pileup, respectively, in a cone of inner width r and outer width R around the lepton. A_{eff}
 831 is the effective area, and ρ is the residual energy from pileup. R varies with the lepton p_T as

$$R = \begin{cases} 0.2 & p_T^\ell < 50 \text{ GeV} \\ \frac{10}{p_T^\ell} & 50 \text{ GeV} < p_T < 200 \text{ GeV} \\ 0.05 & p_T^\ell > 200 \text{ GeV} \end{cases} \quad (10)$$

832 The inner cone width r is dependent on the lepton and the isolation constituent, and is defined
 833 in Table 20.

Constituent	Muon	Electron (barrel)	Electron (endcap)
NH	0.0001	0.0	0.015
CH	0.01	0.0	0.0
PH	0.01	0.0	0.08
PU	0.01	0.0	0.015

Table 20: Inner radii of the cone around the lepton in mini isolation. Constituents are not added to the mini isolation sum if they are separated from the lepton by less than this amount.

834 2D separation requires the lepton to be well-separated from nearby jets. This is defined as
 835 requiring:

$$\Delta R(\ell, \text{closest jet}) > X \text{ OR } p_T^{rel}(\ell, \text{closest jet}) > Y \text{ GeV} \quad (11)$$

836 where $p_T^{rel}(\ell, \text{closest jet})$ is defined as the component of the lepton p_T orthogonal to the axis of
 837 the closest jet. All jets with $p_T > Z$ and $|\eta| < 3.0$ are considered.

838 A.1 Efficiency distributions

839 As a first look at the performance of the different isolation definitions, the signal and back-
 840 ground efficiencies were determined for several isolation working points as a function of the
 841 top candidate p_T and leading lepton p_T . The working points chosen were as follows:

- 842 • Mini isolation, $A = 0.1$

- 843 • Mini isolation, A = 0.2
- 844 • 2D separation, X = 0.4, Y = 45(mu)/25(el), Z = 25
- 845 • 2D separation, X = 0.4, Y = 35(mu)/25(el), Z = 45
- 846 • 2D separation, X = 0.4, Y = 35, Z = 15
- 847 • 2D separation, X = 0.5, Y = 25, Z = 25

848 The efficiency is defined as the number of events with exactly one lepton which passes the
 849 isolation requirements, divided by the number of events with at least one lepton. In both cases
 850 the lepton was required to pass the medium ID. In addition, events were required to have 0
 851 veto leptons from the opposite channel. The veto lepton definition (medium ID, no isolation
 852 requirement) was kept the same for this and all other studies. An additional preselection was
 853 also applied, requiring at least one t-jet candidate (AK8 jet with $p_T > 400$ GeV) and at least
 854 one b-jet candidate (AK4 jet with $p_T > 50$ GeV). The signal efficiency was computed using
 855 the semileptonic $t\bar{t}$ sample, and the background efficiency was computed using the QCD MC
 856 sample.

857 The signal and background efficiencies as a function of the top jet candidate p_T and the leading
 858 lepton p_T are shown in Figure 48 for the muon channel and Figure 49 for the electron channel.
 859 In both channels, mini isolation gives a better signal efficiency – both higher and more uniform
 860 than 2D separation. Overall, mini isolation also has a lower and flatter background efficiency,
 861 particularly at high p_T . A mini isolation cut of 0.1 gives slightly worse signal efficiency, but
 862 relatively larger background rejection. Therefore a mini isolation cut of 0.1 appears to be the
 863 most optimal choice of isolation in both the muon and electron channels.

864 A.2 ROC curves

865 To verify our choice of isolation working point and to determine the most optimal ID require-
 866 ment, we want to determine which combination of ID and type of isolation offers the best
 867 overall performance. To determine this, we create ROC curves for each combination of ID and
 868 isolation. The ROC curves are generated by varying the isolation parameters. For mini isol-
 869 ation the relevant parameter is A; for 2D separation X, Y, and Z are all varied. A is scanned from
 870 0.0 to 0.3 in steps of 0.01. X is scanned from 0.4 to 0.55 in steps of 0.05, Y is scanned from 20 to
 871 45 in steps of 5, and Z is scanned from 15 to 45 in steps of 5.

872 For each choice of parameters, the signal and background efficiencies are computed. In this
 873 case, the efficiency is the number of leptons passing the ID and trigger requirements, divided by
 874 the number of total leptons. This is done for events with at least one t-jet candidate, at least one
 875 b-jet candidate, and at least one lepton. The signal efficiency is computed using semileptonic
 876 $t\bar{t}$, while the background efficiency is computed using QCD MC.

877 The ROC curves for the muon and electron channels are shown in Figure 50. In both the muon
 878 and electron channels mini isolation gives better performance than 2D separation, which sup-
 879 ports the conclusions from the efficiency comparison above. In the muon channel, the medium
 880 ID gives the best performance, while in the electron channel the tight ID is more optimal.

881 The final choice of lepton selection is thus medium ID and $\text{miniIso} > 0.1$ for muons, and tight
 882 ID and $\text{miniIso} > 0.1$ for electrons. Signal and background efficiencies for these working points
 883 are shown in Figure 51.

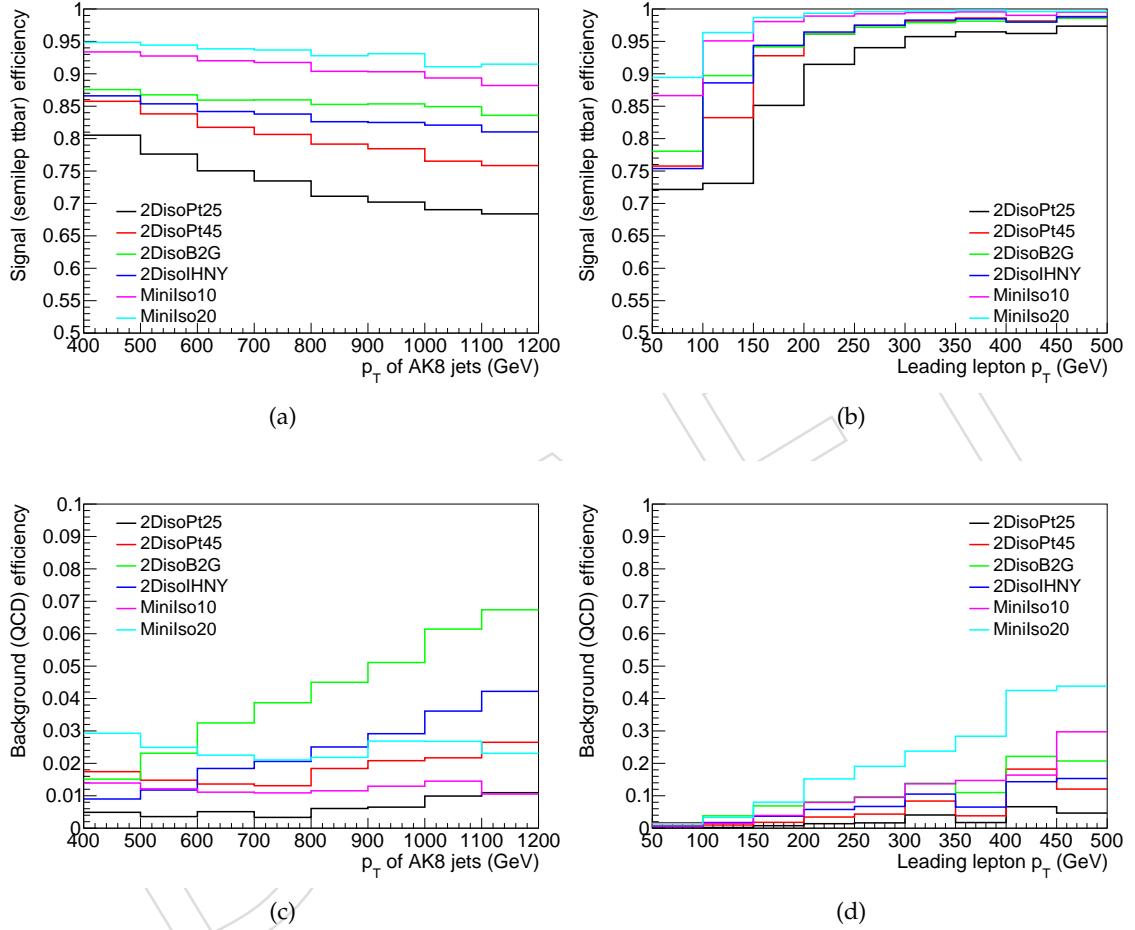


Figure 48: Signal and background efficiencies for different isolation working points in the $\mu + \text{jets}$ channel. All events are required to have 0 veto electrons, at least one t-jet candidate, and at least one b-jet candidate.

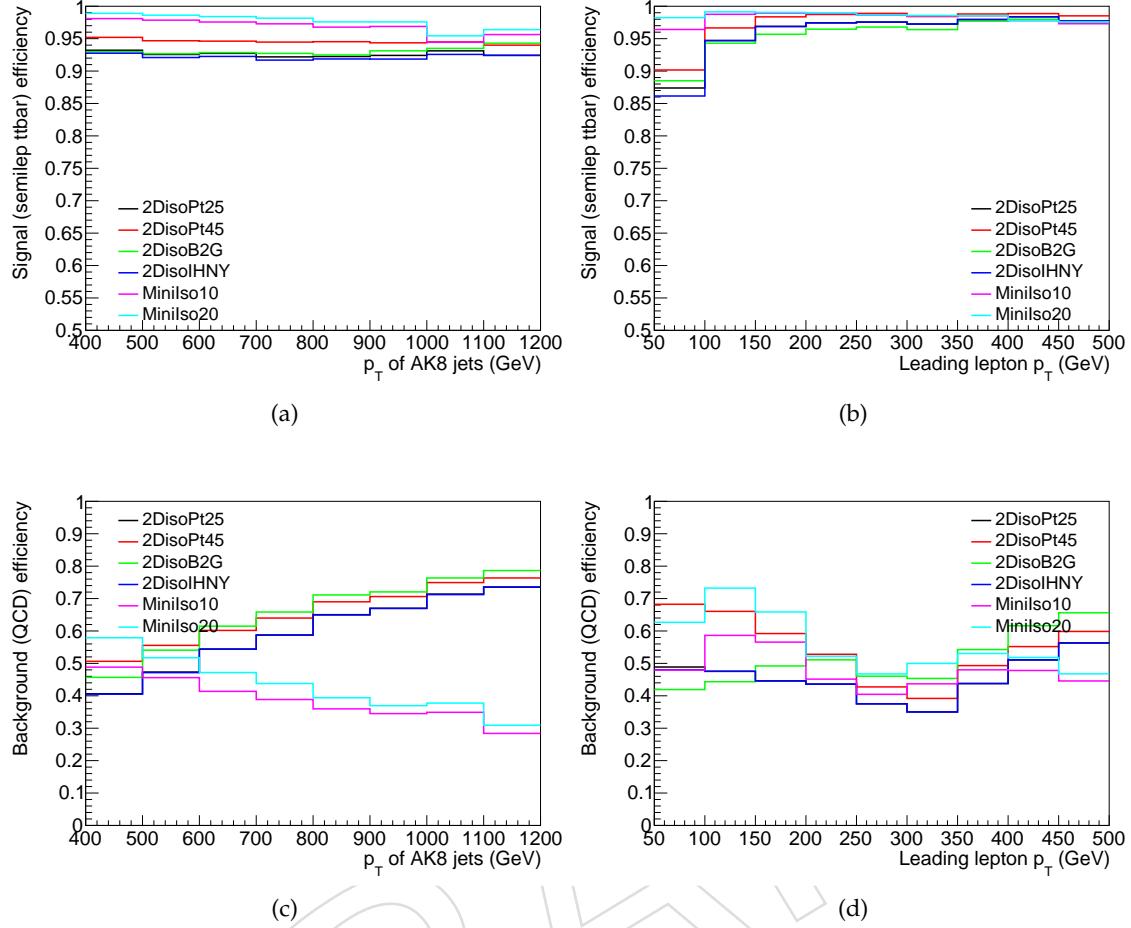


Figure 49: Signal and background efficiencies for different isolation working points in the e+jets channel. All events are required to have 0 veto muons, at least one t-jet candidate, and at least one b-jet candidate.

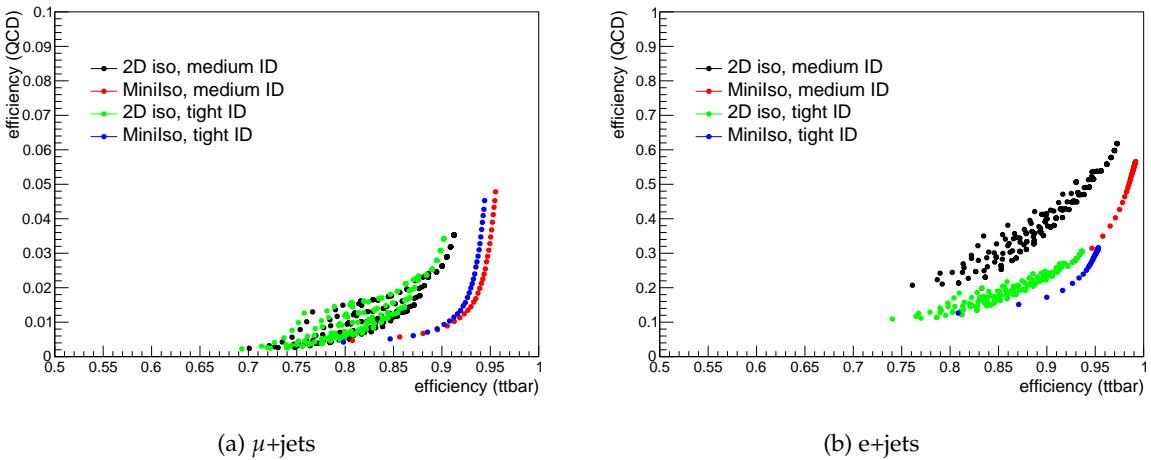


Figure 50: ROC curves for each choice of lepton selection.

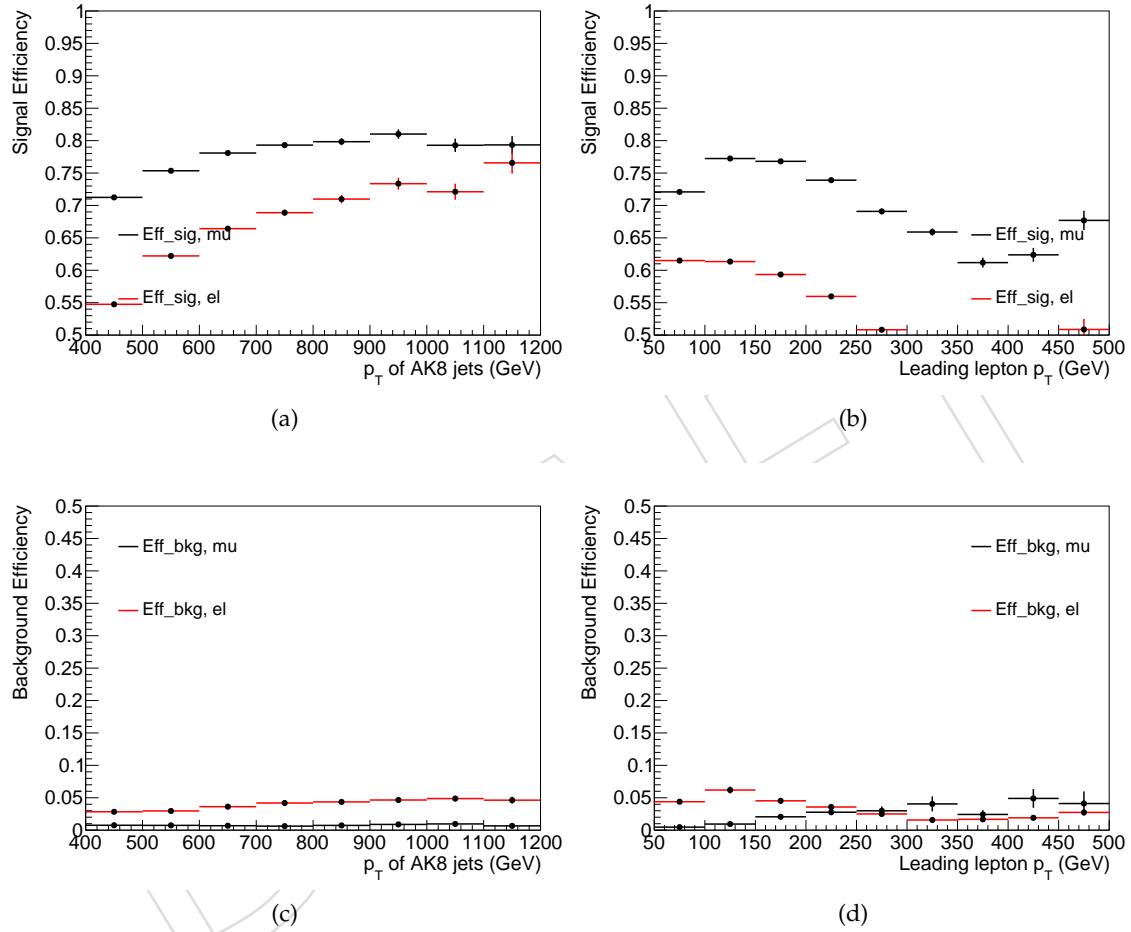


Figure 51: Signal and background efficiencies for the final choice of ID and isolation requirements in the muon and electron channels. Efficiencies are shown with respect to t-jet candidate p_T (left) and leading lepton p_T (right).

884 B Trigger Efficiencies

885 The trigger efficiencies for the muon trigger HLT_MU40_ETA2P1_PFJET200_PFJET50 and elec-
 886 tron trigger HLT_ELE45_CALOIDVT_GSFTRKIDT_PFJET200_PFJET50 were measured in data
 887 and MC using a tag-and-probe technique in an $e\mu$ control sample. To measure the muon trigger
 888 efficiency, data from the SingleElectron stream were collected using the HLT_ELE32_ETA2P1-
 889 _WPTIGHT_GSF trigger. To measure the electron trigger efficiency, data from the SingleMuon
 890 stream were collected using the HLT_MU50 trigger. Data were further required to pass the
 891 following offline selection:

- 892 • ==1 muon with $p_T > 50$ GeV, $|\eta| < 2.1$, Medium ID, and $\text{miniIso} < 0.1$
- 893 • ==1 electron with $p_T > 50$ GeV, $|\eta| < 2.5$, Tight ID, and $\text{miniIso} < 0.1$
- 894 • Leading AK4 jet has $p_T > 250$ GeV, $|\eta| < 2.4$
- 895 • Sub-leading AK4 jet has $p_T > 50$ GeV, $|\eta| < 2.4$
- 896 • Missing $E_T < 75$ GeV

897 The selection was designed to select real $e\mu$ events from $t\bar{t}$ decay, while remaining on the
 898 plateau for the tag triggers. Events passing the preselection and tag trigger were tested as
 899 to whether they passed the corresponding probe trigger. The number of tagged events passing
 900 the preselection in data and MC is given in Table 21, along with the number of events passing
 901 the trigger and the flat trigger efficiency. The trigger efficiency was also computed as a function
 902 of the lepton p_T and η , as well as the p_T of the leading and sub-leading jets. These efficiencies
 903 are given in Figures 52 and 53 for the muon and electron triggers, respectively.

Trigger	Muon trigger		Electron trigger	
	MC	Data	MC	Data
Events passing preselection	7300	2960	8928	3314
Events passing trigger	6704	2594	8717	3160
Trigger efficiency	0.92	0.88	0.98	0.96

Table 21: Number of events passing the tag trigger and preselection, along with number and fraction of events passing the probe trigger, in data and MC for the muon and electron triggers. Trigger efficiencies in MC take into account non-unity event weights.

904 The trigger efficiencies and data / MC SF overall show little dependence on lepton p_T or η ,
 905 and only a slight dependence on the p_T of the leading and subleading jets. The trigger SF as a
 906 function of lepton p_T was chosen to correct MC events.

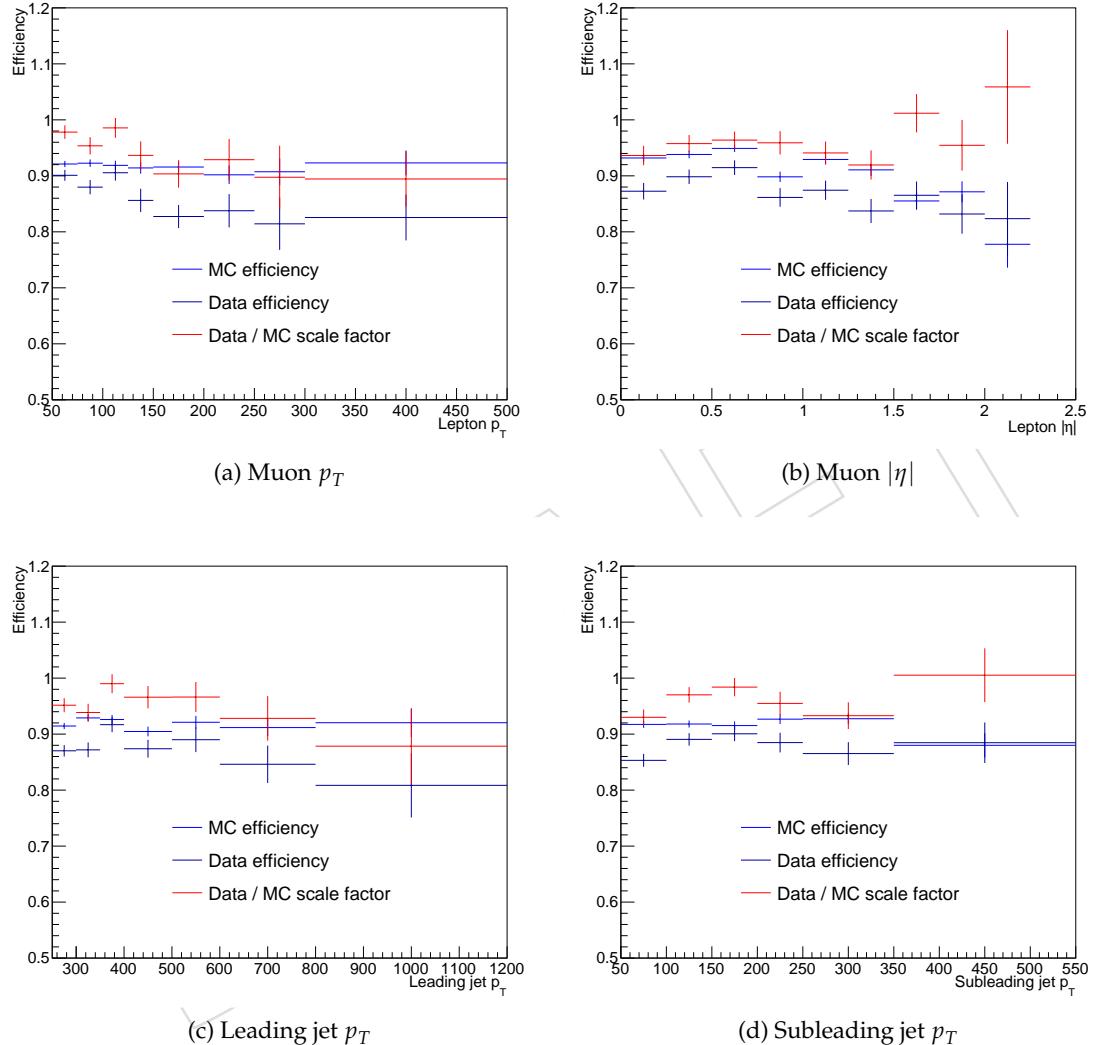


Figure 52: Muon trigger efficiencies in data and MC, as well as data / MC SF, as a function of muon p_T and $|\eta|$, leading jet p_T , and subleading jet p_T . Efficiencies are generally flat with respect to muon kinematics, and only depend slightly on jet p_T .

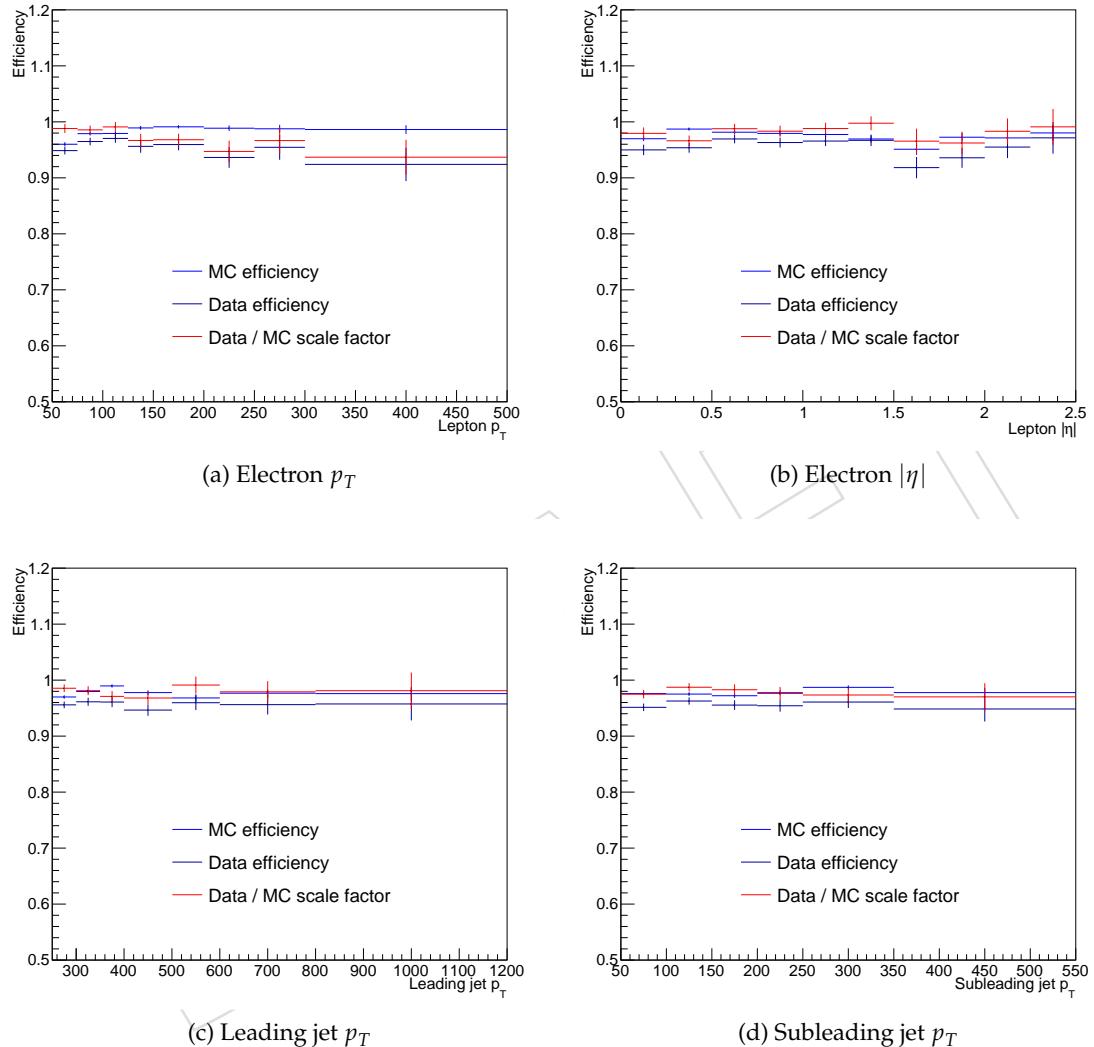


Figure 53: Electron trigger efficiencies in data and MC, as well as data / MC SF, as a function of electron p_T and $|\eta|$, leading jet p_T , and subleading jet p_T . Efficiencies are generally flat with respect to electron kinematics, and only depend slightly on jet p_T .

907 C Selection Optimization

908 This appendix contains the tests performed to optimize the preselection used in the analysis.
 909 Two sets of optimization studies were performed. The first, using 2015 data, was performed
 910 when first commissioning the analysis. A second set of optimization studies was then per-
 911 formed using the full 2016 data, in order to optimize QCD rejection. The optimization studies
 912 are summarized below.

913 C.1 2015 commissioning

914 When initially commissioning the analysis on 2015 data, studies were done to determine the
 915 optimal event selection for the analysis. In these studies, events were considered which passed
 916 the following looser selection:

- 917 • Exactly one signal muon or electron
- 918 • Zero additional veto leptons
- 919 • At least one AK8 jet
- 920 • At least one AK4 jet

921 The additional cuts which were considered are as follows:

- 922 • Cut on ΔR between AK4 jet and lepton
- 923 • Cut on ΔR between AK8 jet and lepton
- 924 • Cut on missing E_T
- 925 • Cut on $H_T (\sum_{jets} p_T)$
- 926 • Cut on H_T^{lep} (missing $E_T + p_T^\ell$)
- 927 • Triangular cut in electron channel

To determine if adding a given cut would optimize the selection, the quantity in question was plotted to determine whether the signal and background shapes looked different by eye. In addition, the discriminant

$$\sigma = \frac{S}{\sqrt{S + B}} \quad (12)$$

928 was computed while scanning over possible cut values. A cut on the quantity was considered
 929 if the cut maximized σ while remaining physically well-motivated.

930 The distributions of ΔR between the lepton and the AK4 / AK8 jets, along with the correspond-
 931 ing σ distributions, are shown in Figure 54 for the muon channel and Figure 55 for the electron
 932 channel. Physical motivations suggest the lepton should be separated from the AK8 jet and
 933 close to the AK4 jet. However, the σ scan does not indicate a particular preference for a cut on
 934 the ΔR separation in either case. Because of this, the cuts were chosen based on what was used
 935 in the 8 TeV analysis. The AK4 jet was required to satisfy $0.1 < \Delta R(\ell, j) < \pi/2$, and the AK8
 936 jet was required to satisfy $\Delta R(\ell, j) > \pi/2$.

937 The distributions of the missing E_T , H_T , and H_T^{lep} , along with the corresponding σ distributions,
 938 are shown in Figure 56 for the muon channel and Figure 57 for the electron channel. These
 939 distributions are produced with the hemisphere cuts determined above applied. The only cut
 940 which maximizes σ is a cut at 35 GeV on missing E_T in the electron channel. To be consistent,
 941 a missing E_T cut of 35 GeV was chosen for both channels. While cuts at 500 GeV for H_T or
 942 150 GeV for H_T^{lep} would remain optimal, they would not cut out a large number of signal or

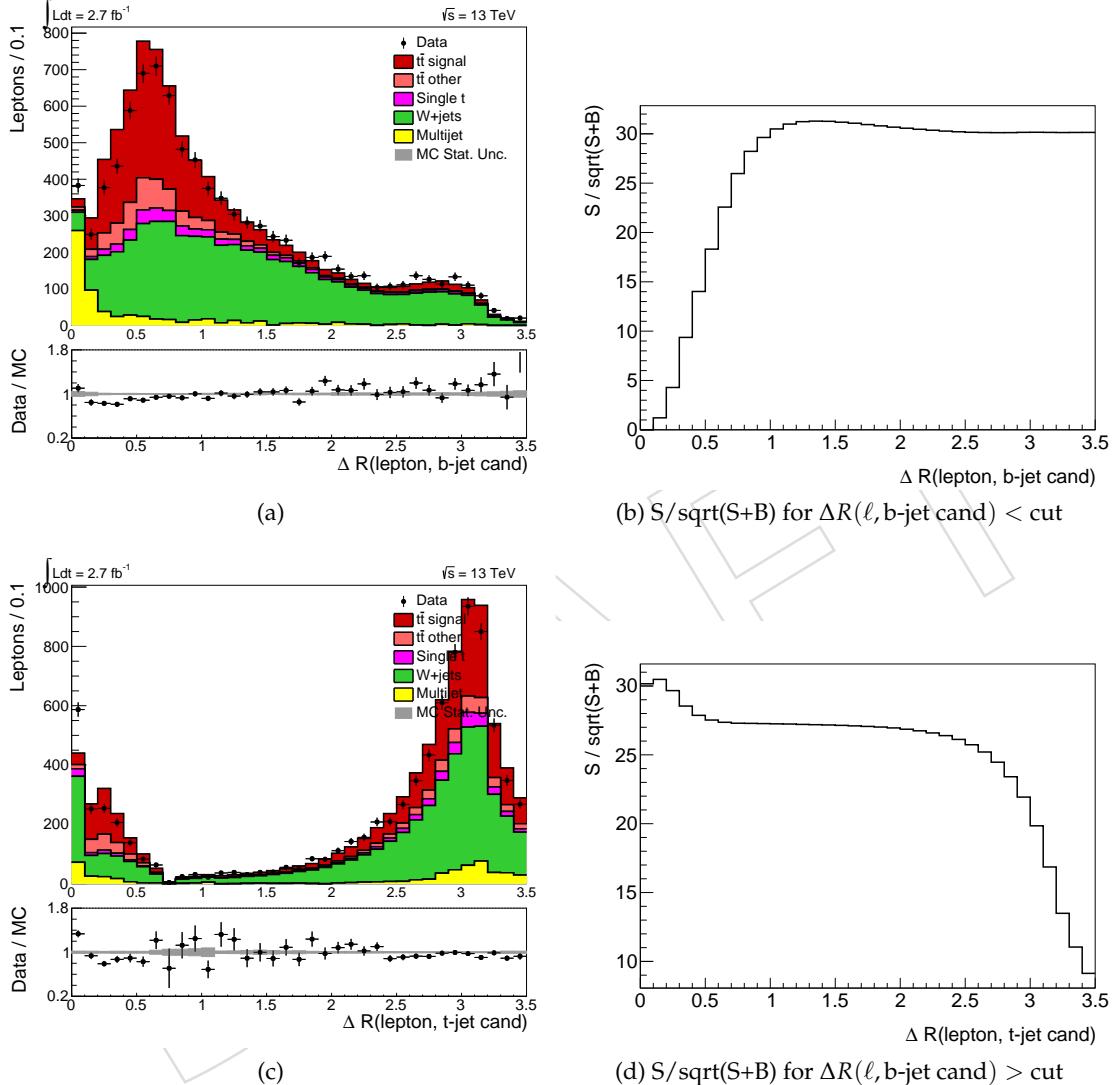


Figure 54: Studies on the effect of a ΔR cut between the muon and the AK4 or AK8 jet. The kinematic distributions of each quantity are shown on the left, and the distribution of σ given a cut at the given value are shown on the right.

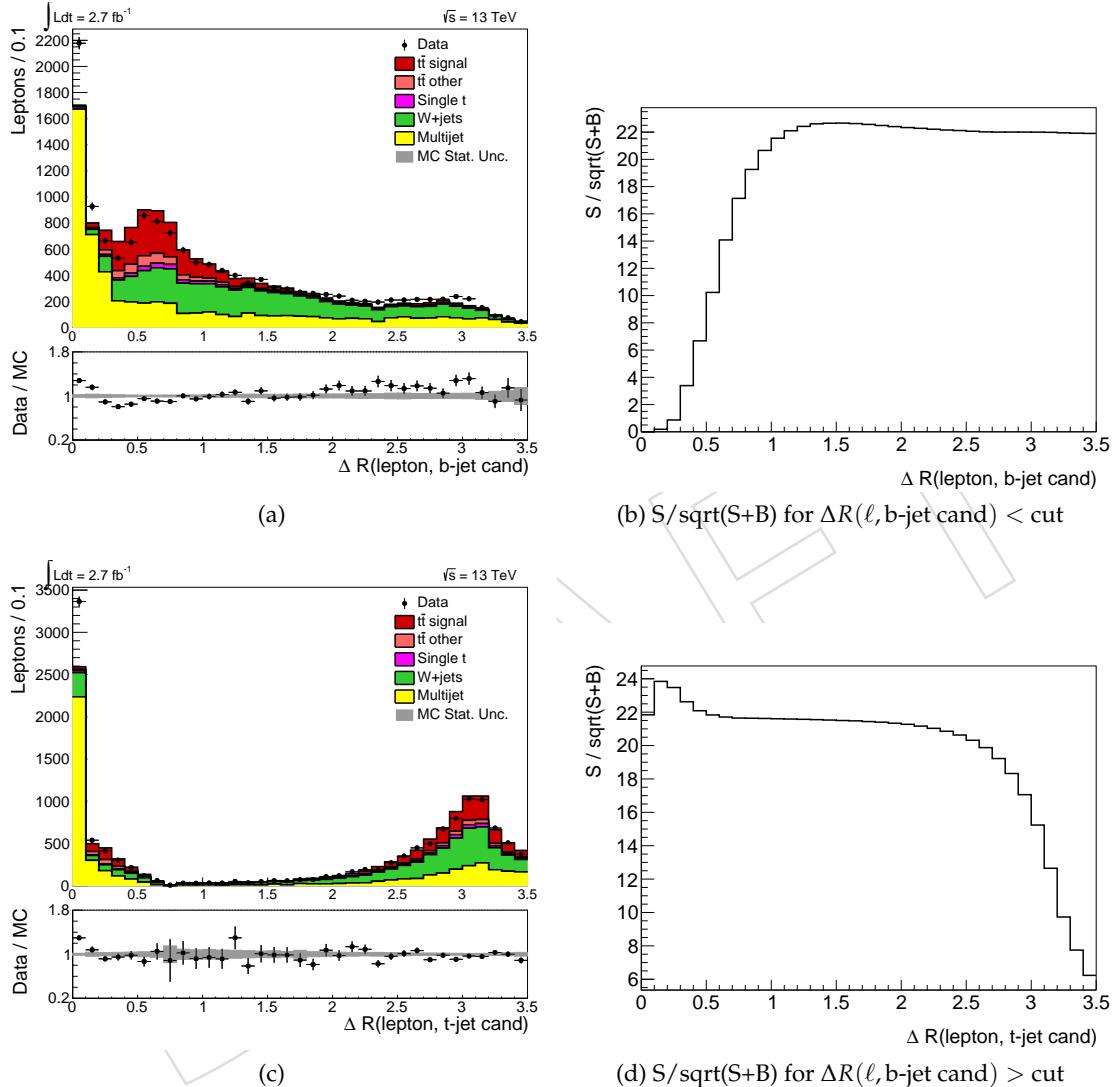


Figure 55: Studies on the effect of a ΔR cut between the electron and the AK4 or AK8 jet. The kinematic distributions of each quantity are shown on the left, and the distribution of σ given a cut at the given value are shown on the right.

background events. Because of this, no cut was chosen for H_T or H_T^{lep} .

Finally, a triangular cut was considered on the electron channel. The number of signal and background events, along with σ , for the electron channel is shown in Table 22. While the value of σ is the same with and without the cut, the benefit of eliminating QCD background was judged to be higher than retaining signal. Thus the triangular cut was applied in the electron channel.

Sample	No triangular cut	Triangular cut
t̄t (signal)	2056 ± 5	1629 ± 4
t̄t (non-semilep)	401 ± 2	362 ± 2
Single top	257 ± 4	215 ± 3
W+jets	2659 ± 6	1937 ± 5
QCD (MC)	2435 ± 49	680 ± 24
Data	7278	4342
σ	23	23

Table 22: Event counts in the electron channel, along with the calculated value of σ , for events passing the selection with and without a triangular cut.

C.2 Final optimization

To further improve QCD rejection, especially in the electron channel, a second set of optimization studies was performed with the full 2016 dataset. In this case, the optimization studies were performed using events which passed the preselection as determined in the previous section. Specifically, events were required to pass

- Exactly one signal muon(electron)
- 0 additional veto leptons
- At least one leptonic jet
 - AK4 jet in the same hemisphere as the electron, $0.1 < \Delta R(\ell, \text{jet}) < \pi/2$
- At least one hadronic jet
 - AK8 jet in the opposite hemisphere from the electron, $\Delta R(\ell, \text{jet}) > \pi/2$
- Missing $E_T > 35 \text{ GeV}$
- ‘Triangular’ cut in the electron channel
 - $|\Delta\phi(e, E_T^{\text{miss}}) - 1.5| < 1.5 * \frac{E_T^{\text{miss}}}{75 \text{ GeV}}$
 - $|\Delta\phi(\text{leadjet}, E_T^{\text{miss}}) - 1.5| < 1.5 * \frac{E_T^{\text{miss}}}{75 \text{ GeV}}$

The following further optimizations were considered:

- Increase of minimum ΔR between AK4 jet and lepton
- Increase of cut on missing E_T
- Increase of missing E_T scale in triangular cut

As previously, the optimal selection was determined by maximizing $\sigma = \frac{S}{\sqrt{S+B}}$. In this case, B is the number of QCD events passing the selection, while S is the number of t̄t events.

The missing E_T distributions, along with the corresponding σ distributions, are shown for the muon and electron channels in Figure 58. Increasing the cut on missing E_T in the electron

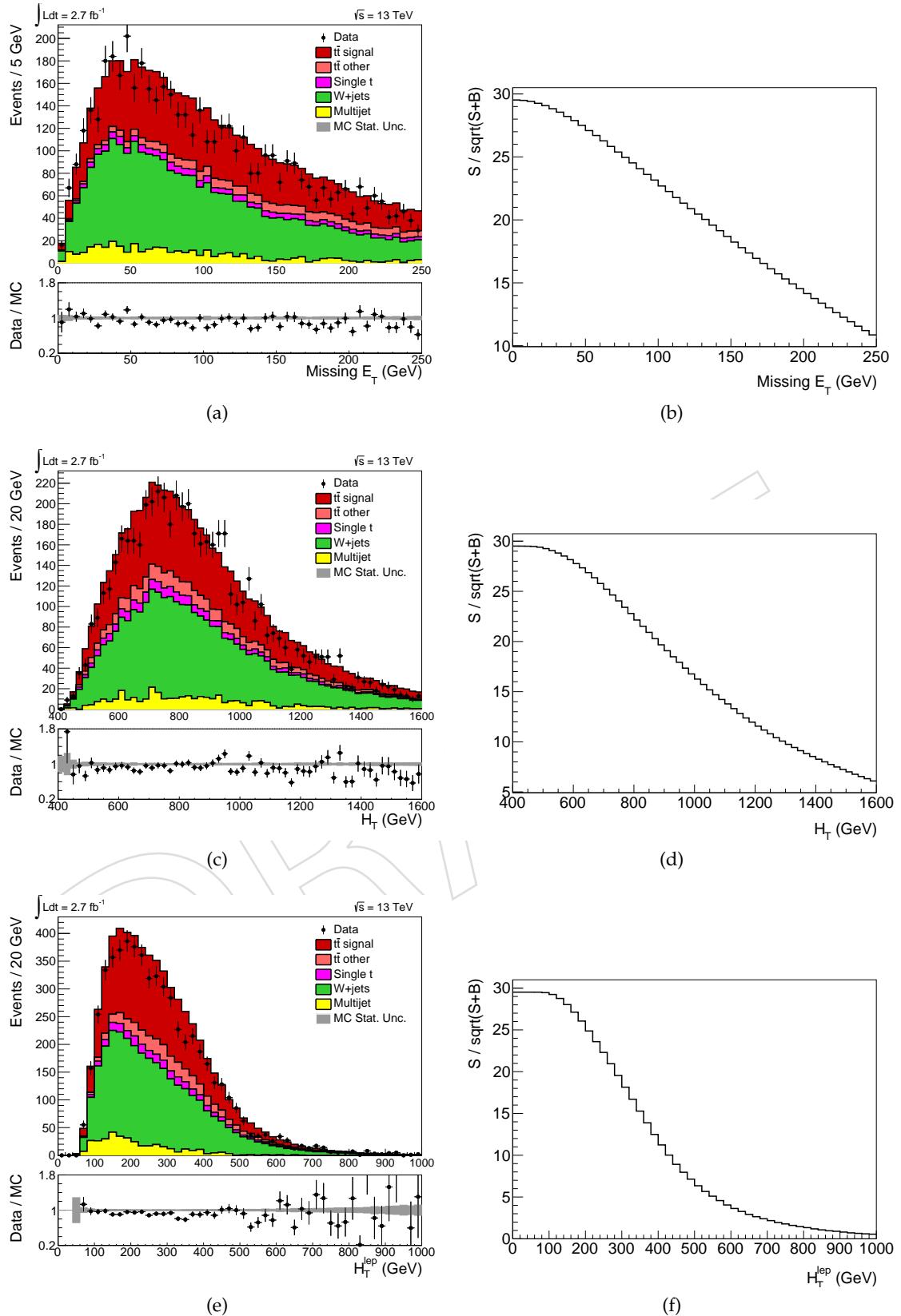


Figure 56: Studies on the effect of a cut on missing E_T , H_T , or H_T^{lep} for events in the muon channel. The kinematic distributions of each quantity are shown on the left, and the distribution of σ given a cut at the given value are shown on the right.

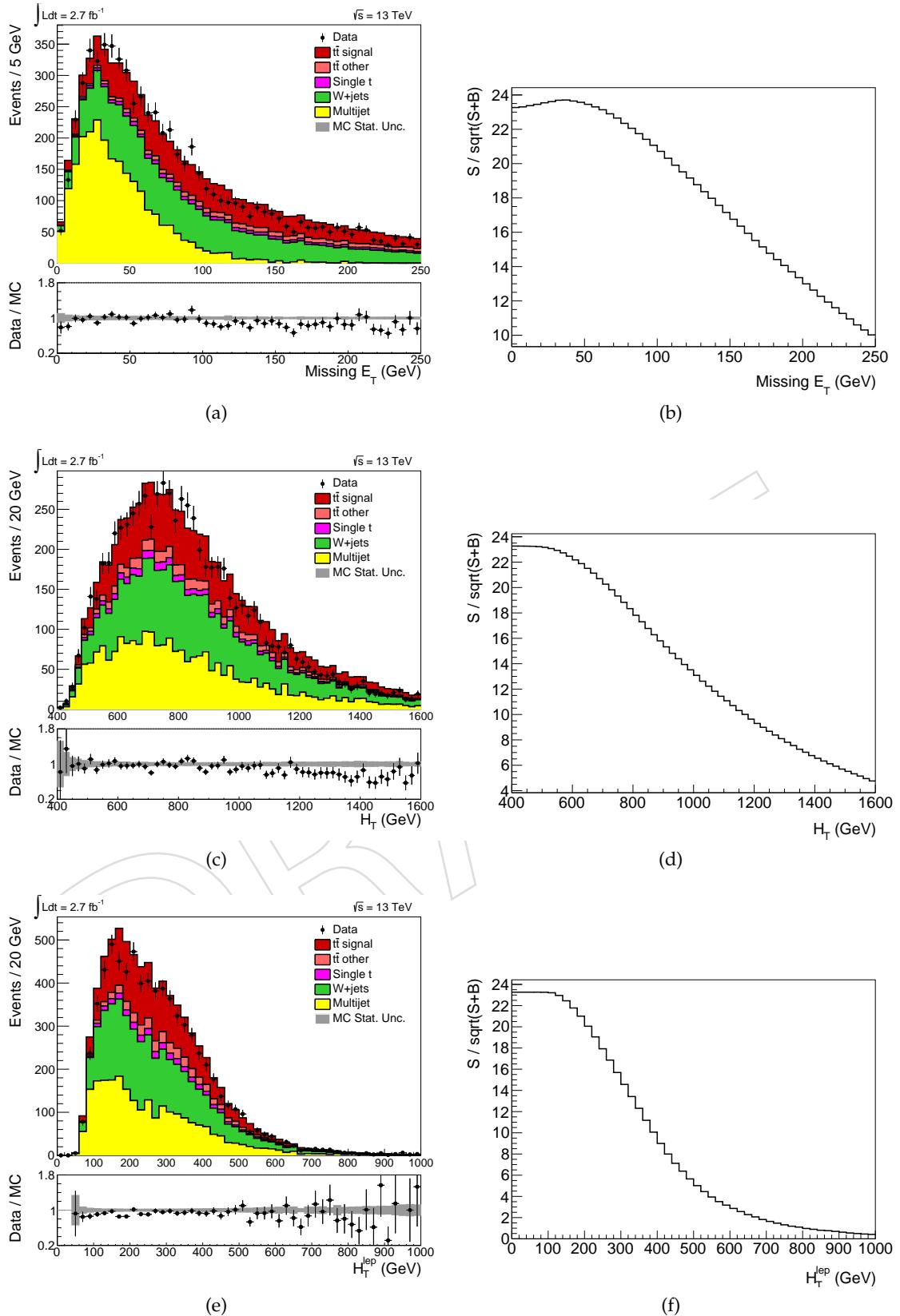


Figure 57: Studies on the effect of a cut on missing E_T , H_T , or H_T^{lep} for events in the electron channel. The kinematic distributions of each quantity are shown on the left, and the distribution of σ given a cut at the given value are shown on the right.

channel to 50 GeV is shown to be comparably optimal to the current selection. The cut is thus increased to 50 GeV to benefit from the greater reduction in QCD.

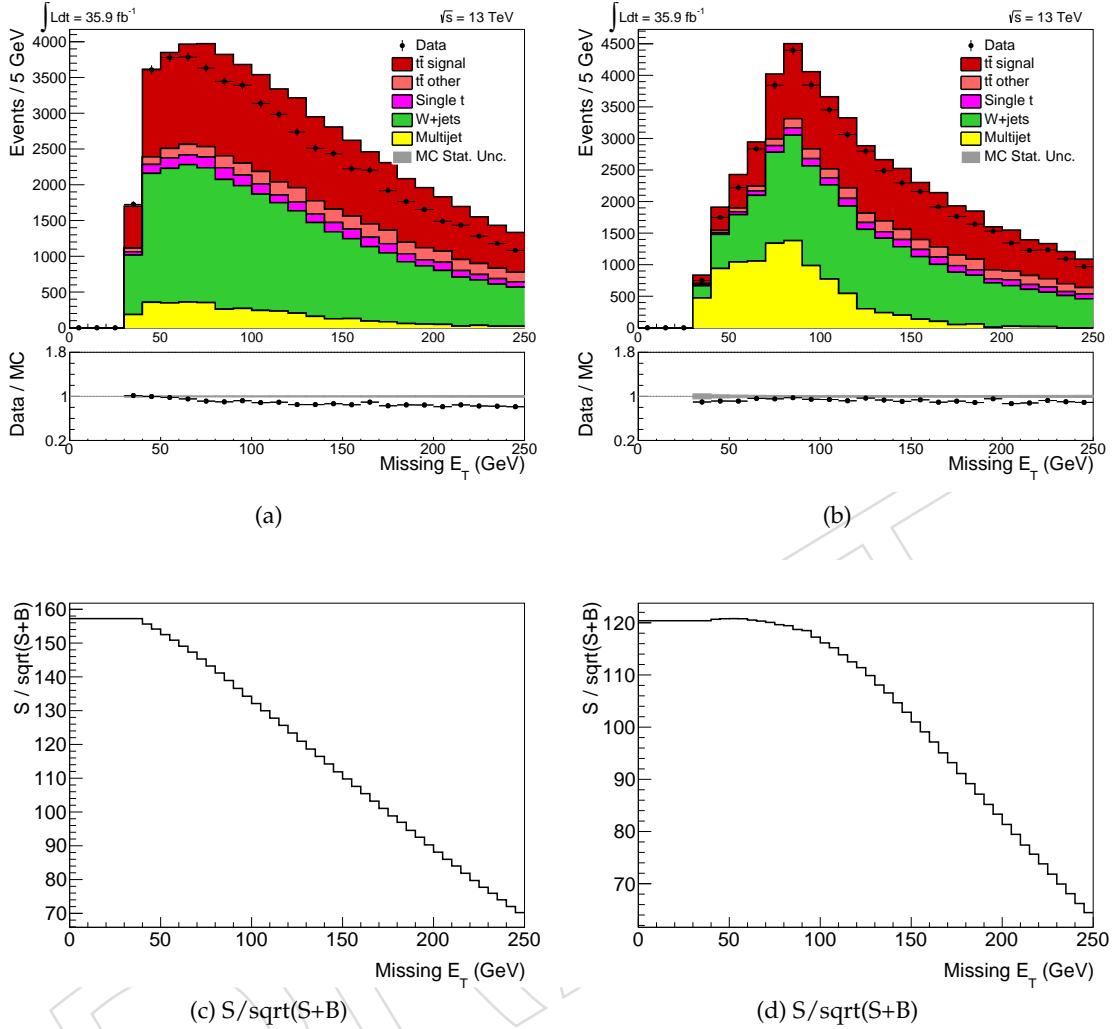


Figure 58: Studies on the effect of increasing the missing E_T cut in the muon (left) and electron (right) channels. The top plots show the missing E_T distribution for events passing the preselection, while the bottom plots show the distribution of σ given a cut at the given value.

The distributions of σ corresponding to different values of the missing E_T scales in the triangular cut are shown in Figure 59 for the missing E_T separation from the jet and lepton. Both σ distributions peak around 110 GeV. The missing E_T scale in the triangular cut is thus increased to 110 GeV for both components.

Instead of plotting the σ distribution for the minimum separation between the lepton and the b-jet, the assessment of the optimal cut is done by eye. The $\Delta R(e/\mu, b\text{-jet})$ distributions are shown in Figure 60. Increasing the minimum value of $\Delta R(e/\mu, b\text{-jet})$ to 0.3 removes a significant fraction of QCD, while retaining most signal events.

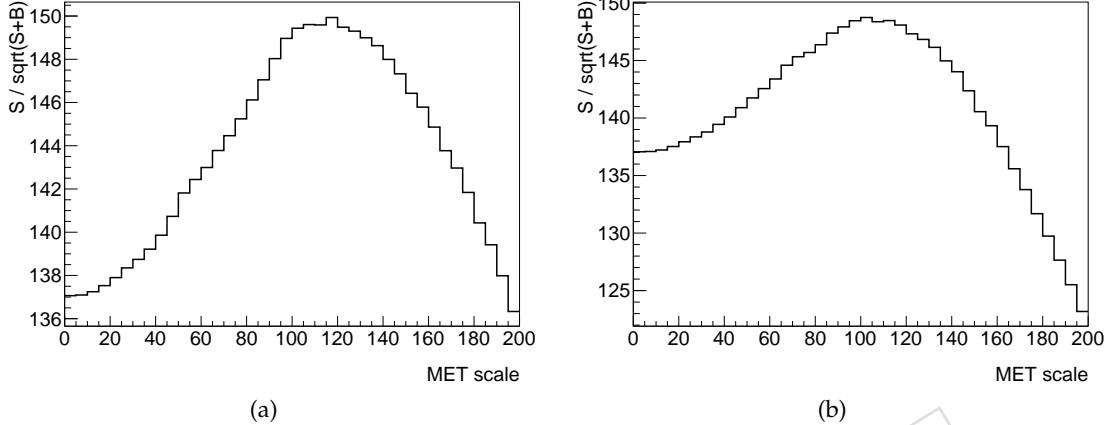


Figure 59: Studies on the effect of increasing the missing E_T scale in the electron channel triangular cut, for the electron (left) and jet (right) components. Both plots show σ for the events passing the triangular cut with the corresponding missing E_T scale.

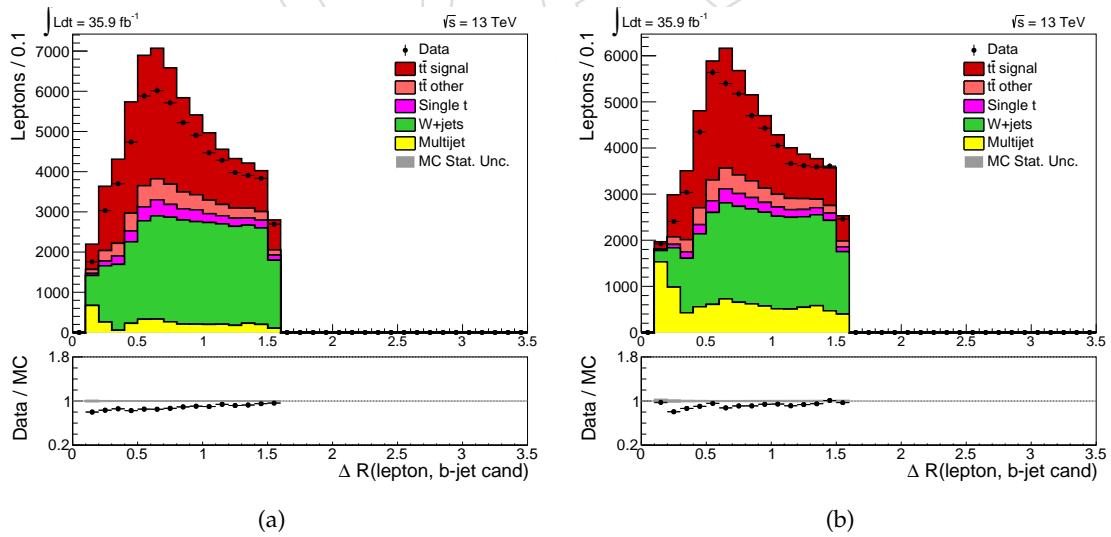


Figure 60: Separation between the b-jet candidate and the muon (right) or electron (left), for events passing the preselection.

982 D Additional QCD studies

983 This appendix contains comparisons of the QCD predictions given by the data-driven side-
 984 band, MC sideband, and MC signal region for many kinematic distributions. In general, there
 985 is reasonable agreement between the MC prediction for QCD in the sideband and the data-
 986 driven QCD estimate in the sideband. However, some systematic shifts may be observed. The
 987 largest difference between the signal region and sideband region distributions is in the lepton
 988 η in the electron channel, where the signal region distribution peaks at higher $|\eta|$ than the side-
 989 band. The same is true (to a lesser degree) for the AK4 jet η in the electron channel. In the muon
 990 channel, the number of jets is lower in the signal region than the sideband, and the muon p_T is
 991 higher. The comparison between the QCD MC prediction in the signal region and the shapes
 992 in the sideband is dominated by the low QCD statistics in the signal region. The QCD shape in
 993 the signal region generally agrees with that in the sideband within 100%.

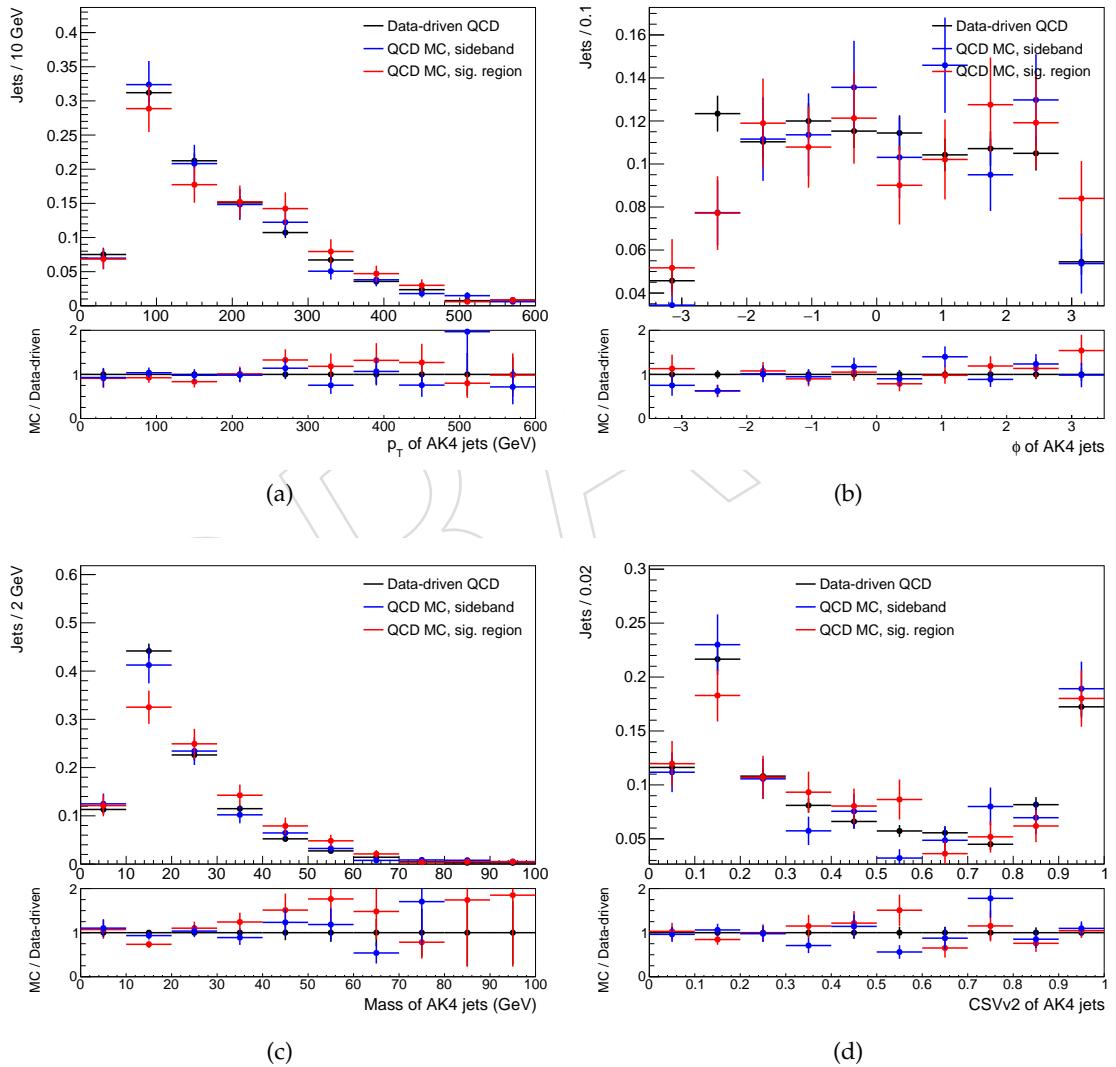


Figure 61: Comparisons of AK4 jet properties for QCD events in the muon channel as predicted by MC in the signal region, MC in the sideband region, and the data-driven technique.

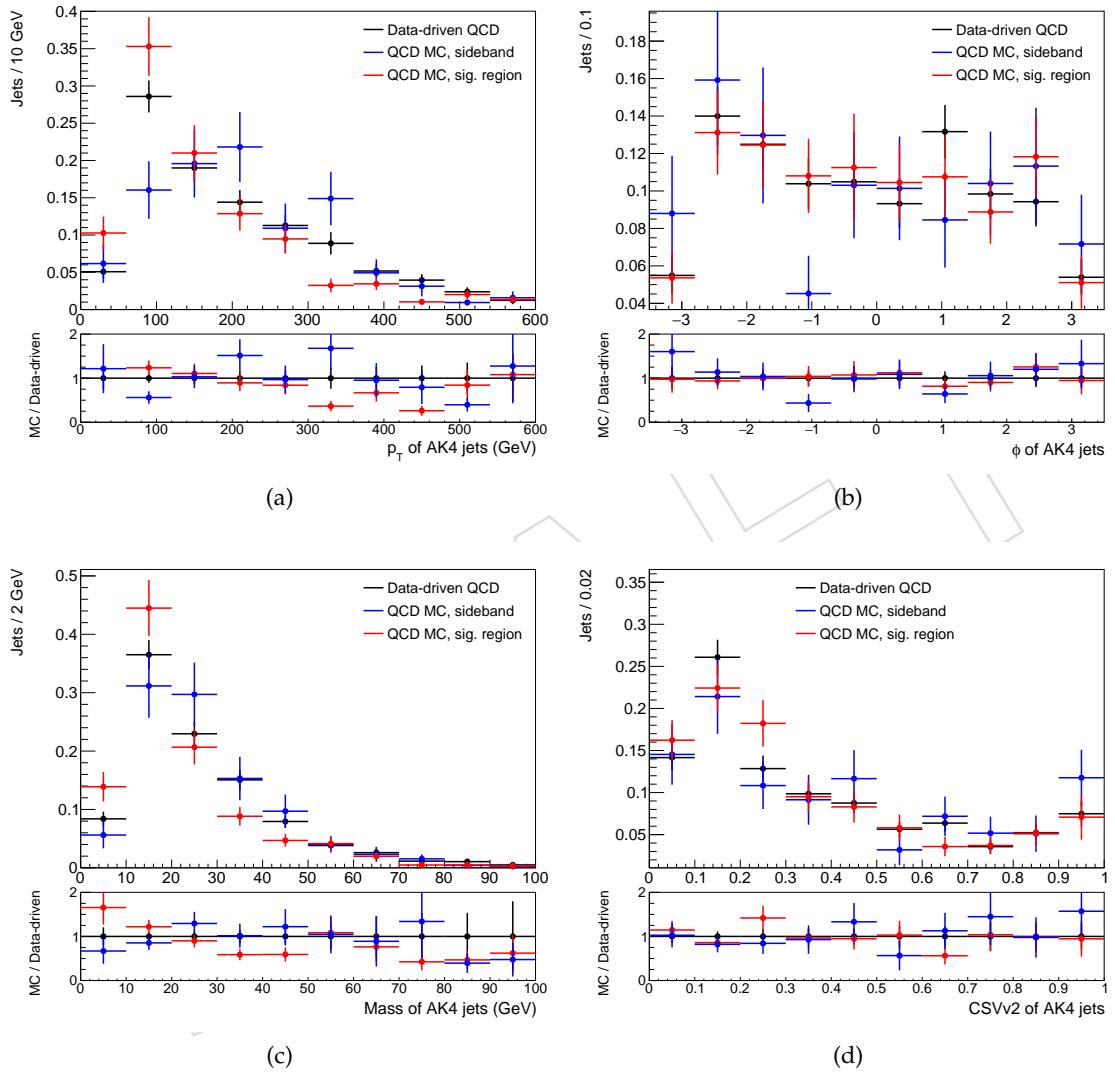


Figure 62: Comparison of AK4 jet properties for QCD events in the electron channel as predicted by MC in the signal region, MC in the sideband region, and the data-driven technique

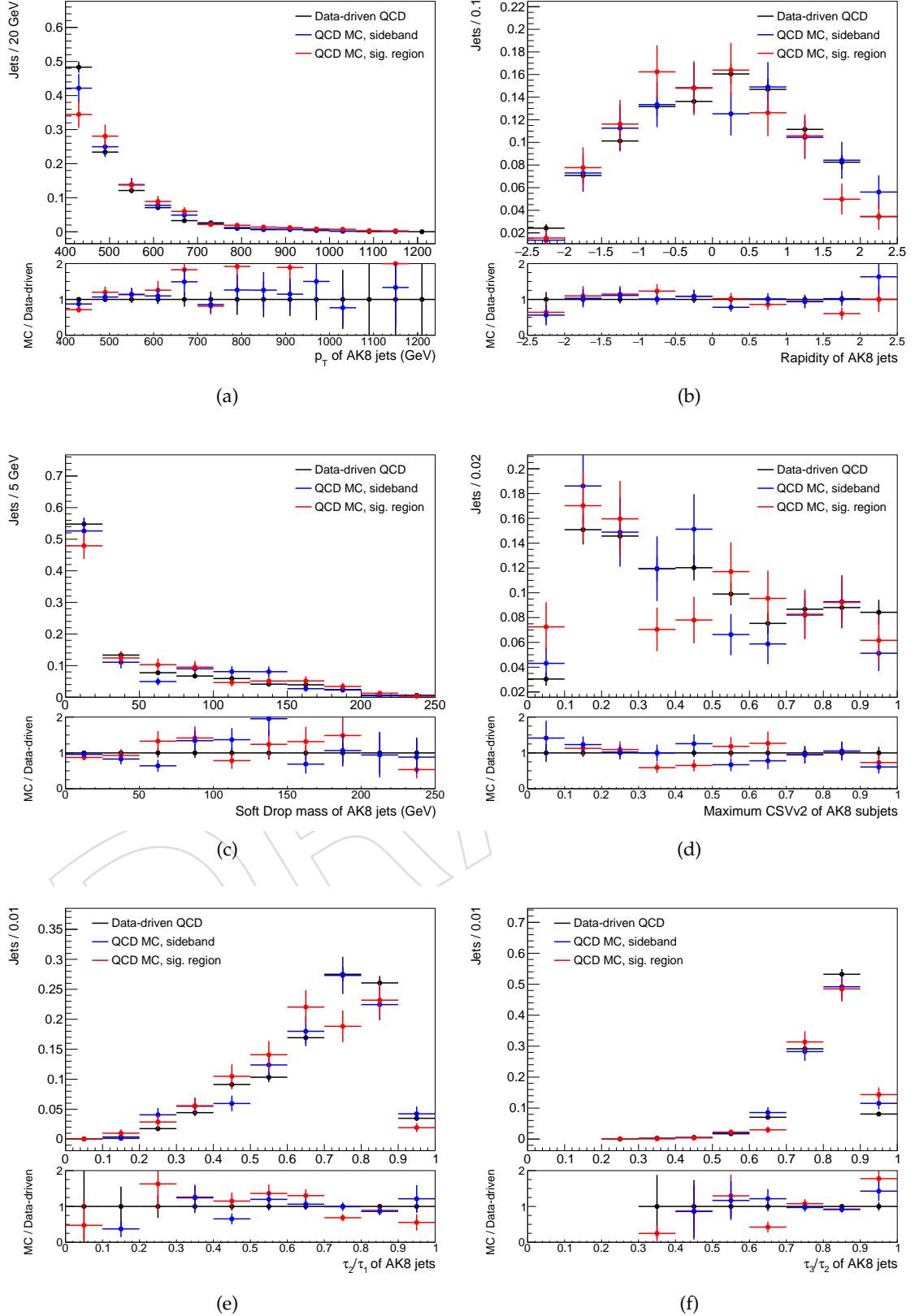


Figure 63: Comparison of AK8 jet properties for QCD events in the muon channel as predicted by MC in the signal region, MC in the sideband region, and the data-driven technique

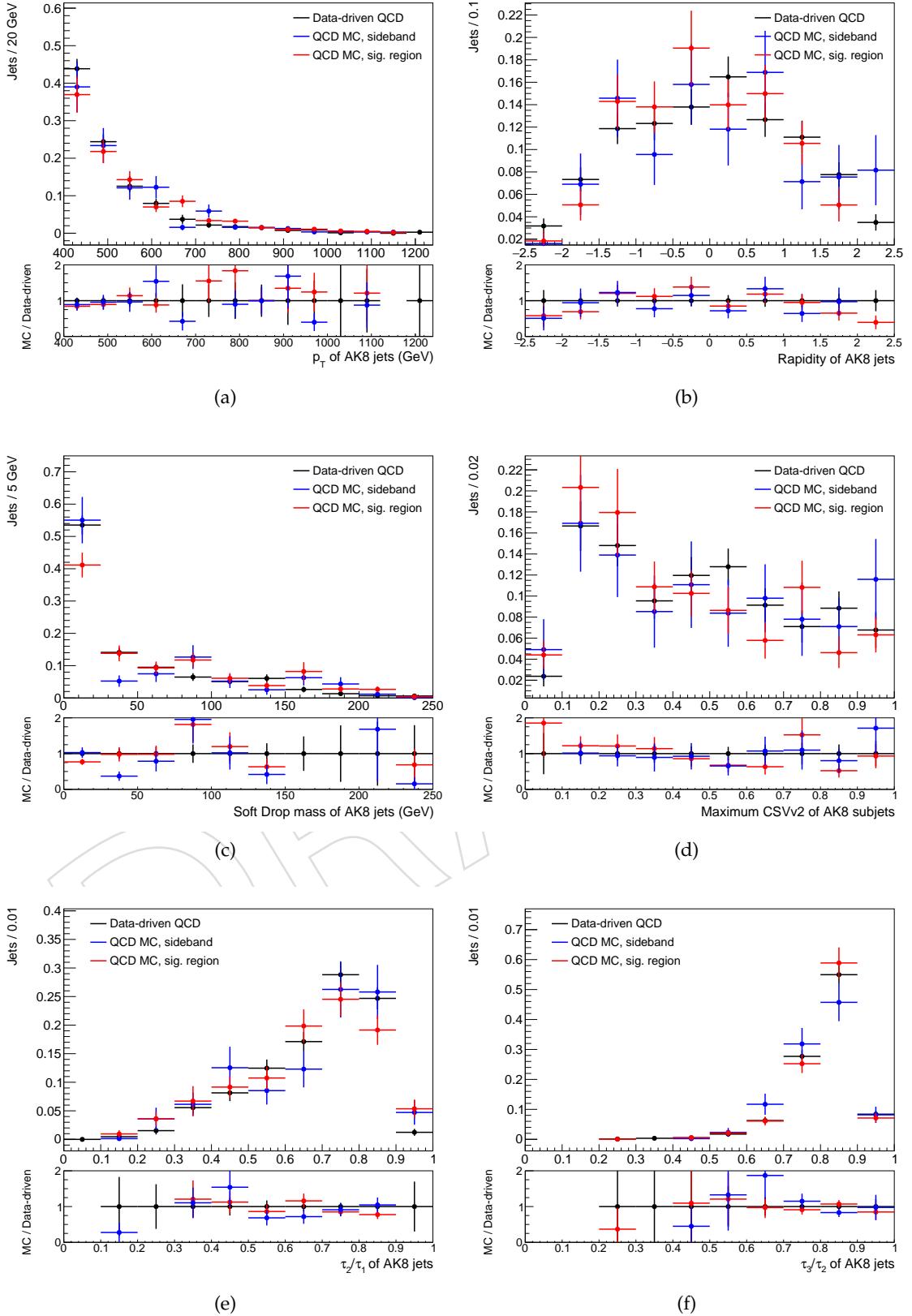


Figure 64: Comparison of AK8 jet properties for QCD events in the electron channel as predicted by MC in the signal region, MC in the sideband region, and the data-driven technique

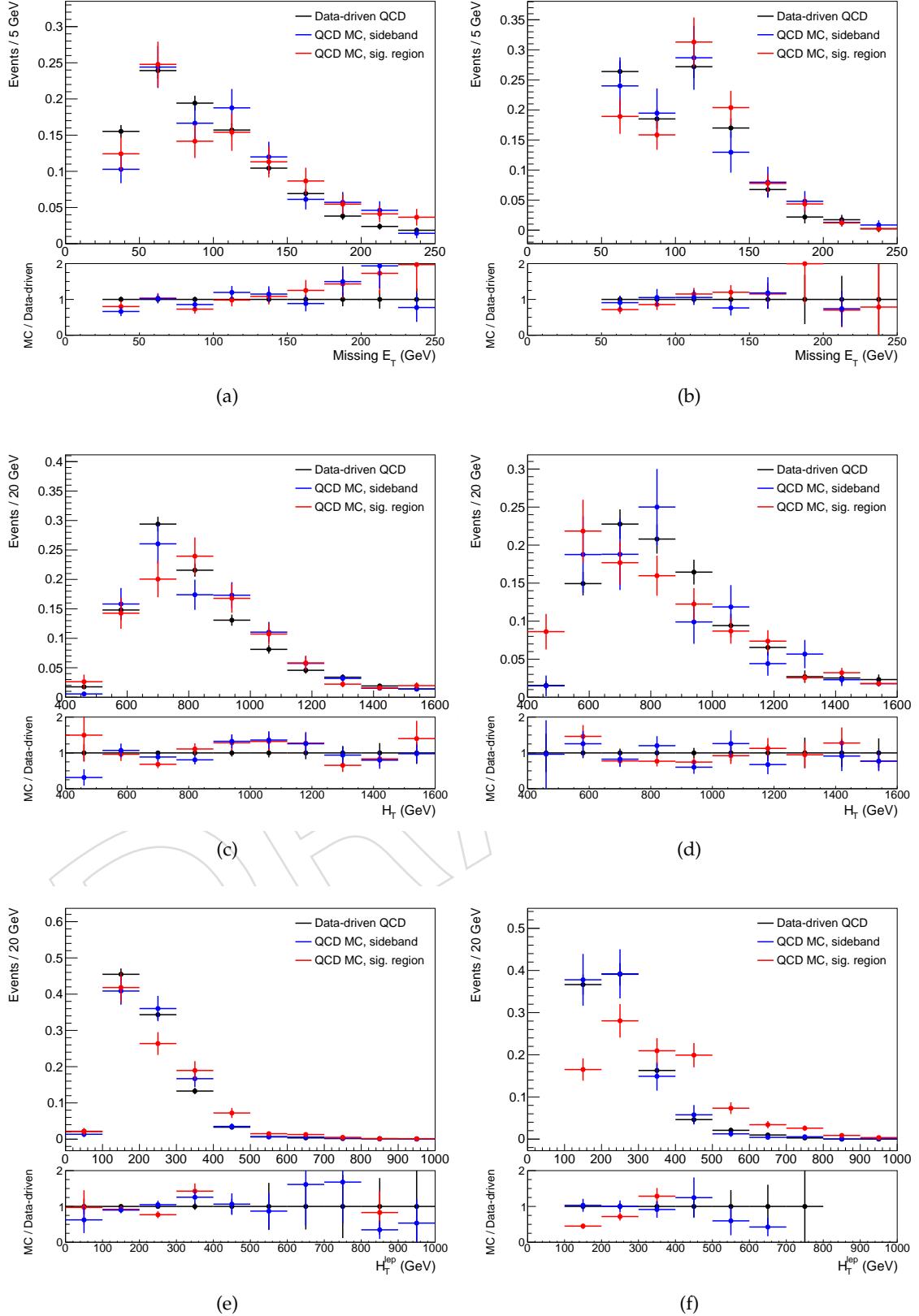


Figure 65: Comparison of event properties for QCD events in the muon channel (left column) and electron channel (right column) as predicted by MC in the signal region, MC in the sideband region, and the data-driven technique

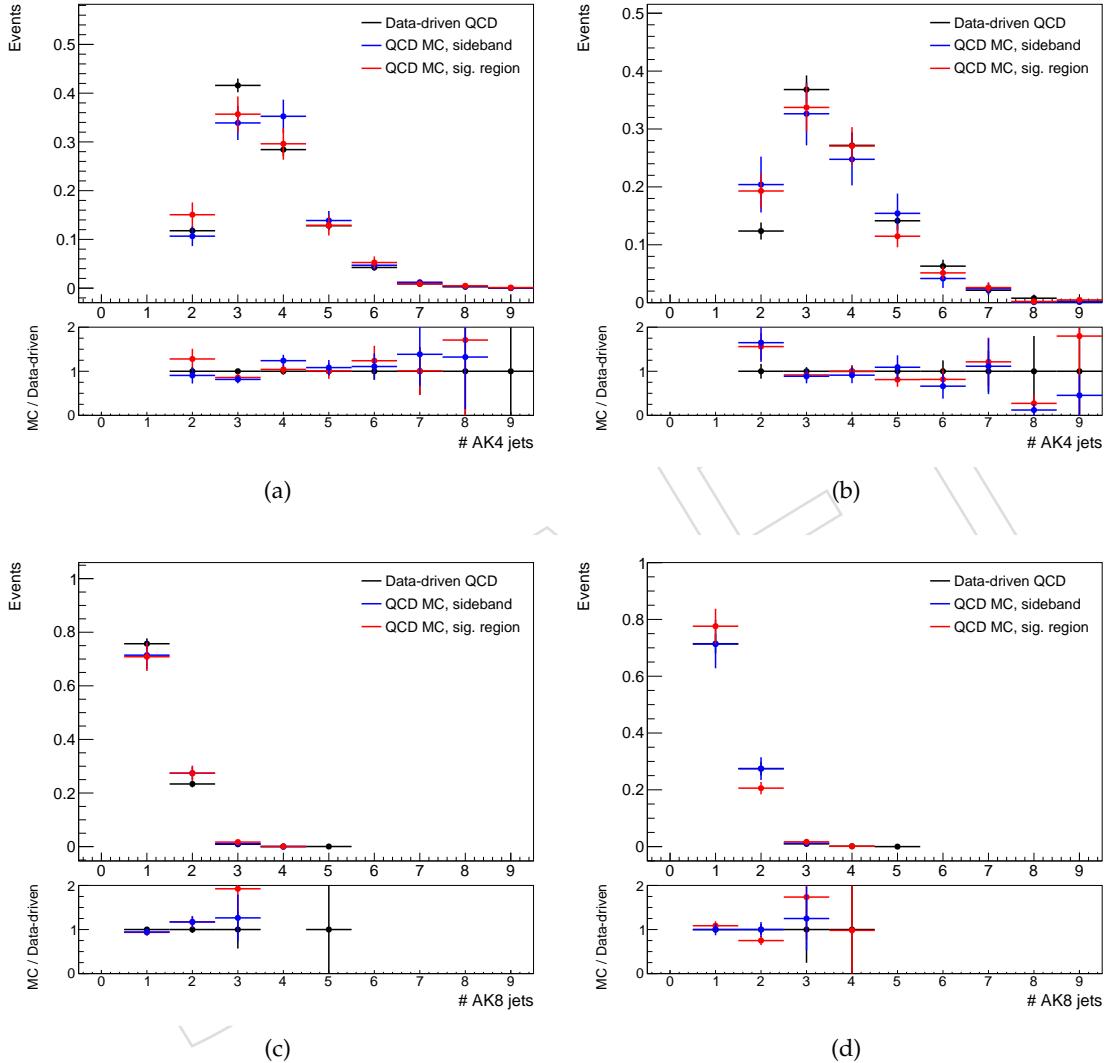


Figure 66: Comparison of event properties for QCD events in the muon channel (left column) and electron channel (right column) as predicted by MC in the signal region, MC in the sideband region, and the data-driven technique

994 E Additional Kinematic Distributions

995 This appendix contains comparisons between data and signal and background predictions for
996 a variety of kinematic distributions. Kinematic distributions are shown for events passing the
997 preselection, as well as separate distributions for events in the 0t, 1t0b, and 1t1b regions. All
998 figures shown use the pre-fit normalizations as given in Tables 3 and 4. The 0t region is dom-
999 inated by W+jets, with significant contributions from QCD and tt> signal. The 1t0b region is
1000 dominated by tt> signal with large contributions from W+jets and some from QCD. The 1t1b re-
1001 gion is fully dominated by the tt> signal, with only small background contributions from QCD,
1002 non-semileptonic tt>, W+jets, and single top.

DRAFT

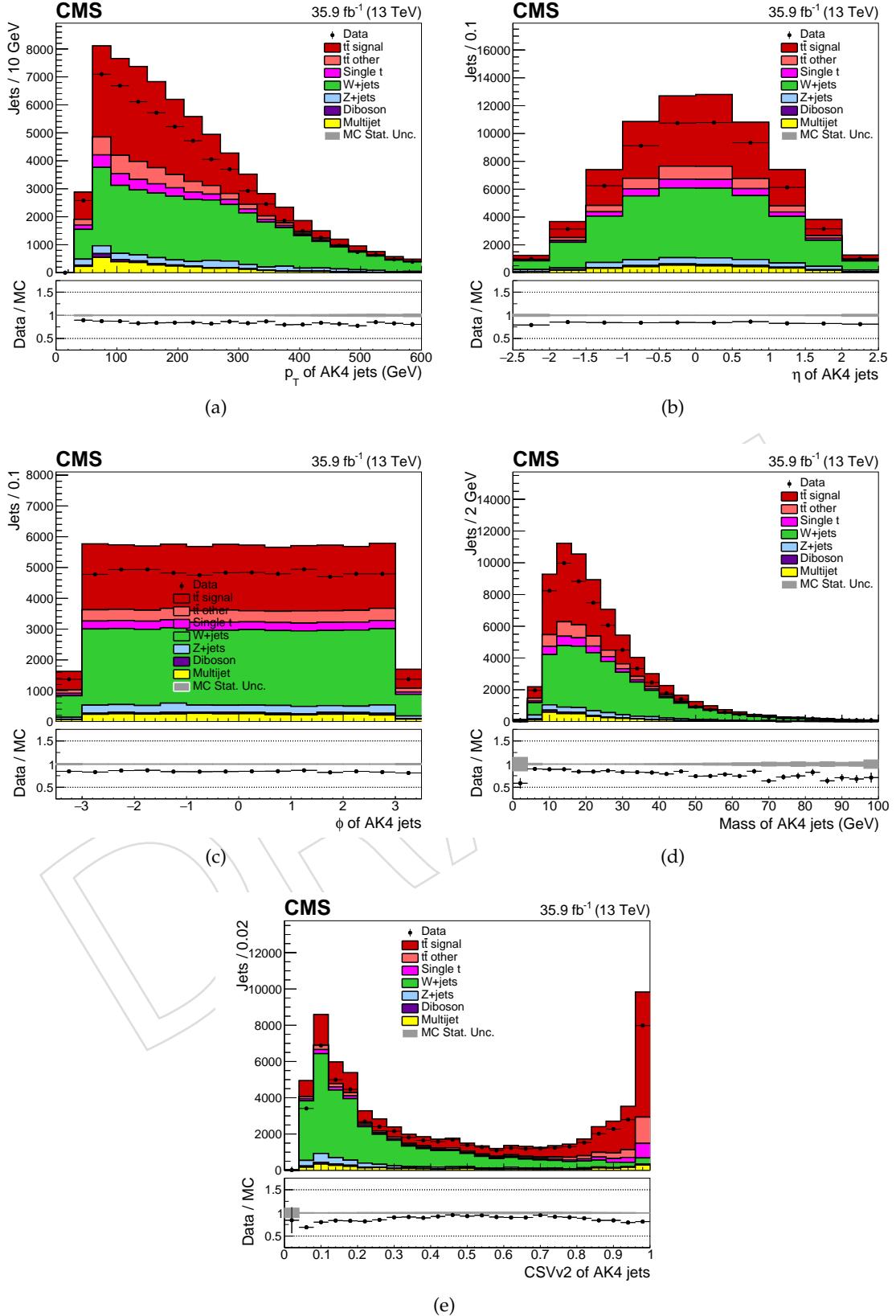


Figure 67: AK4 jet properties in the muon channel for events passing the preselection

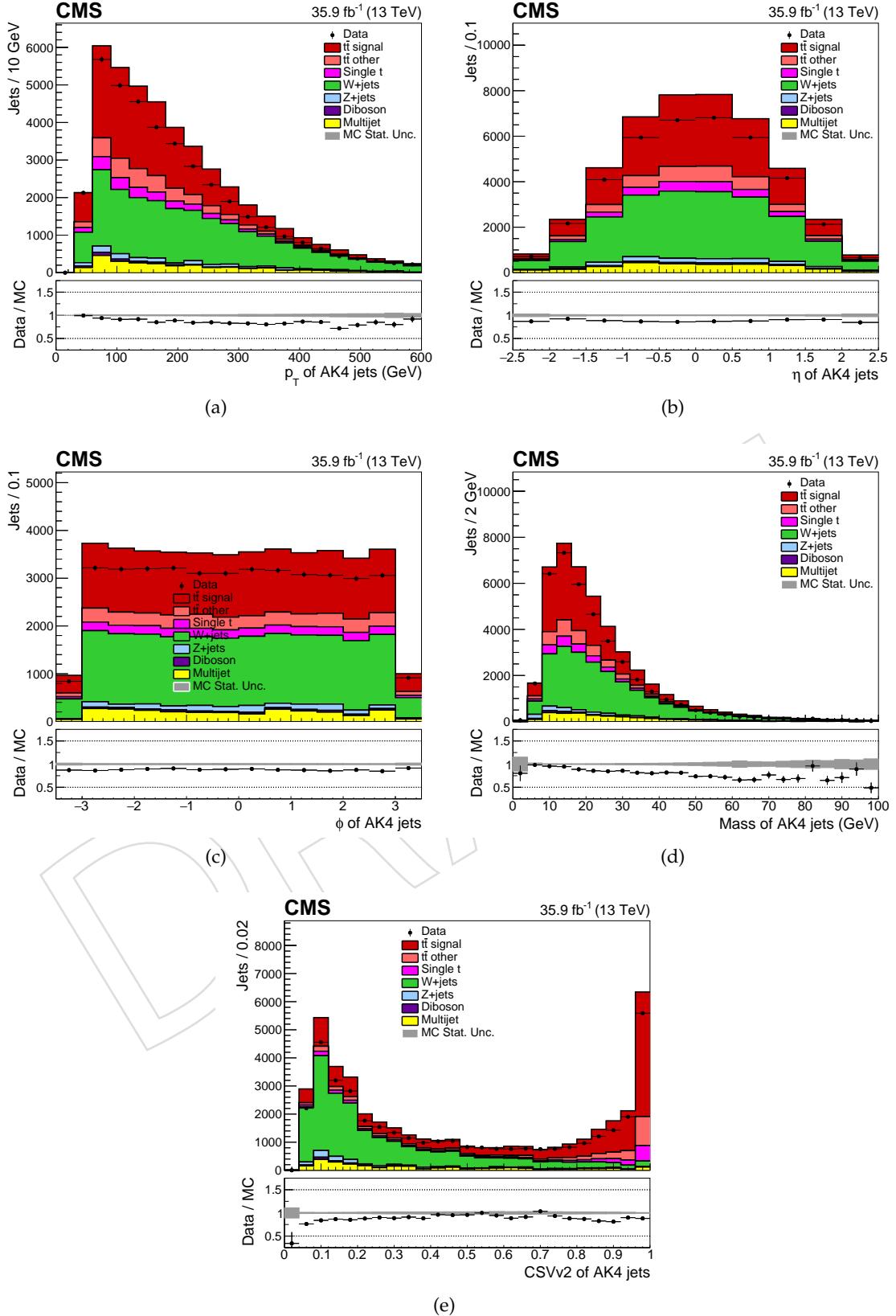


Figure 68: AK4 jet properties in the electron channel for events passing the preselection

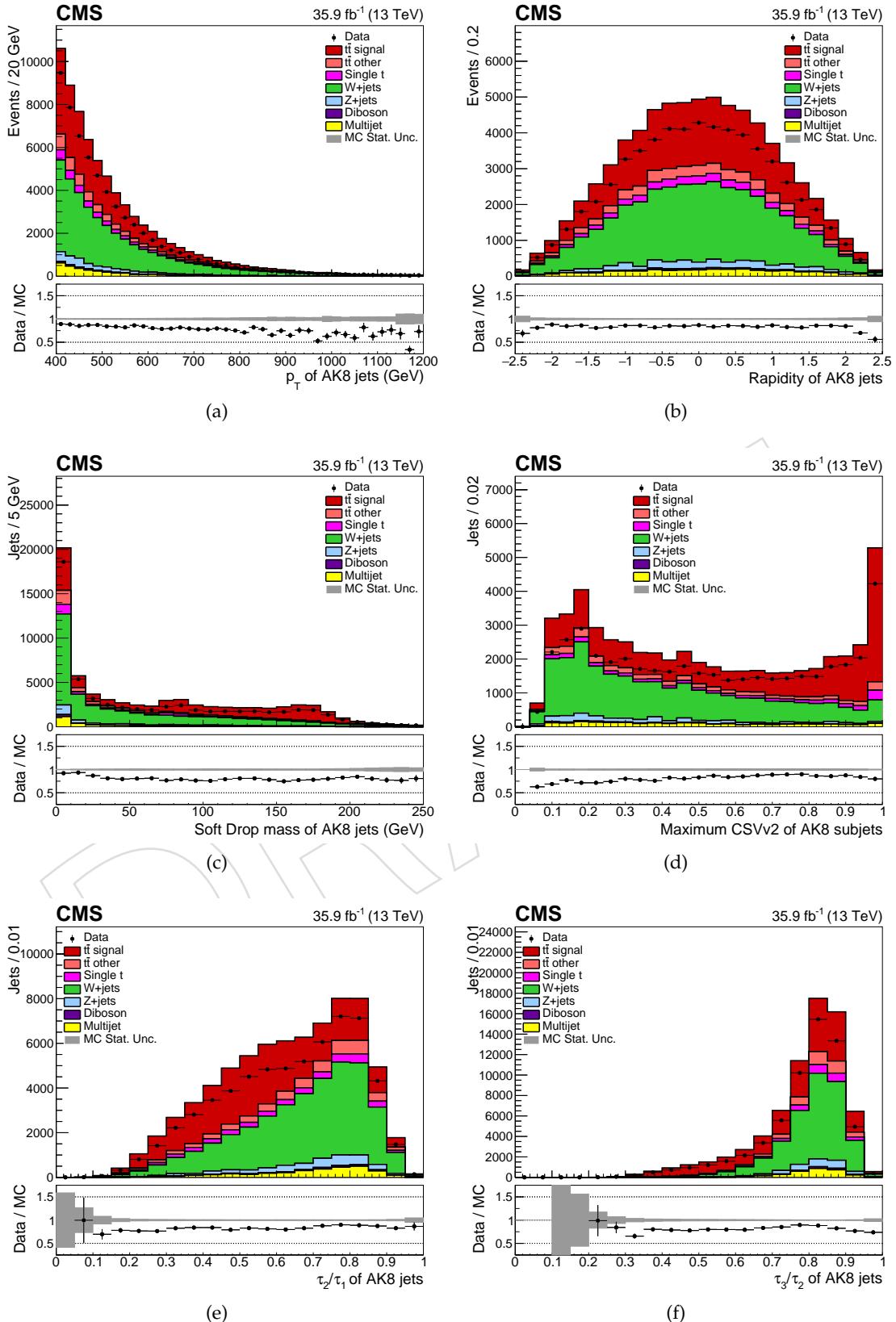


Figure 69: AK8 jet properties in the muon channel for events passing the preselection

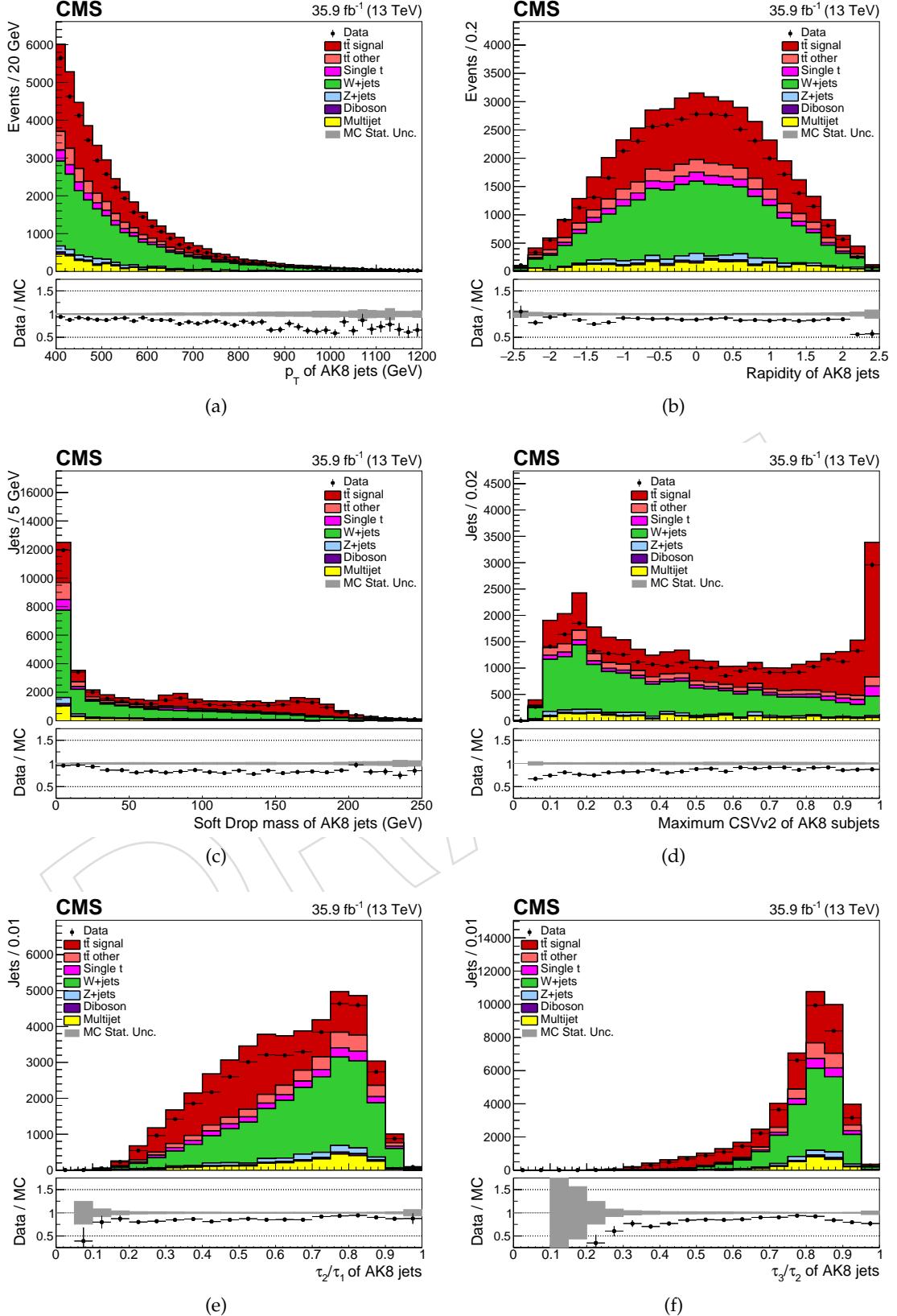


Figure 70: AK8 jet properties in the electron channel for events passing the preselection

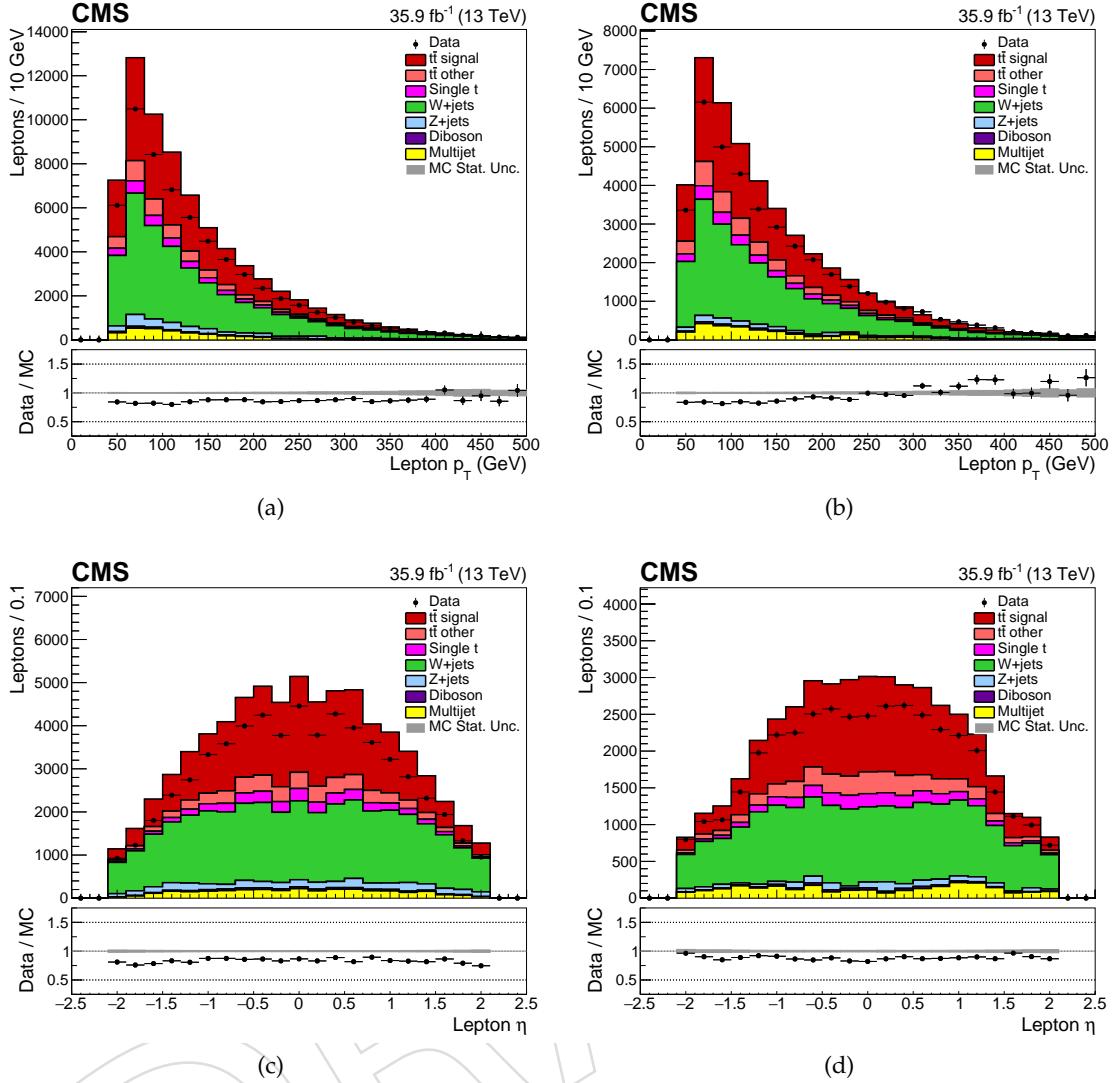


Figure 71: Lepton properties in the muon channel (left column) and electron channel (right column) for events passing the preselection

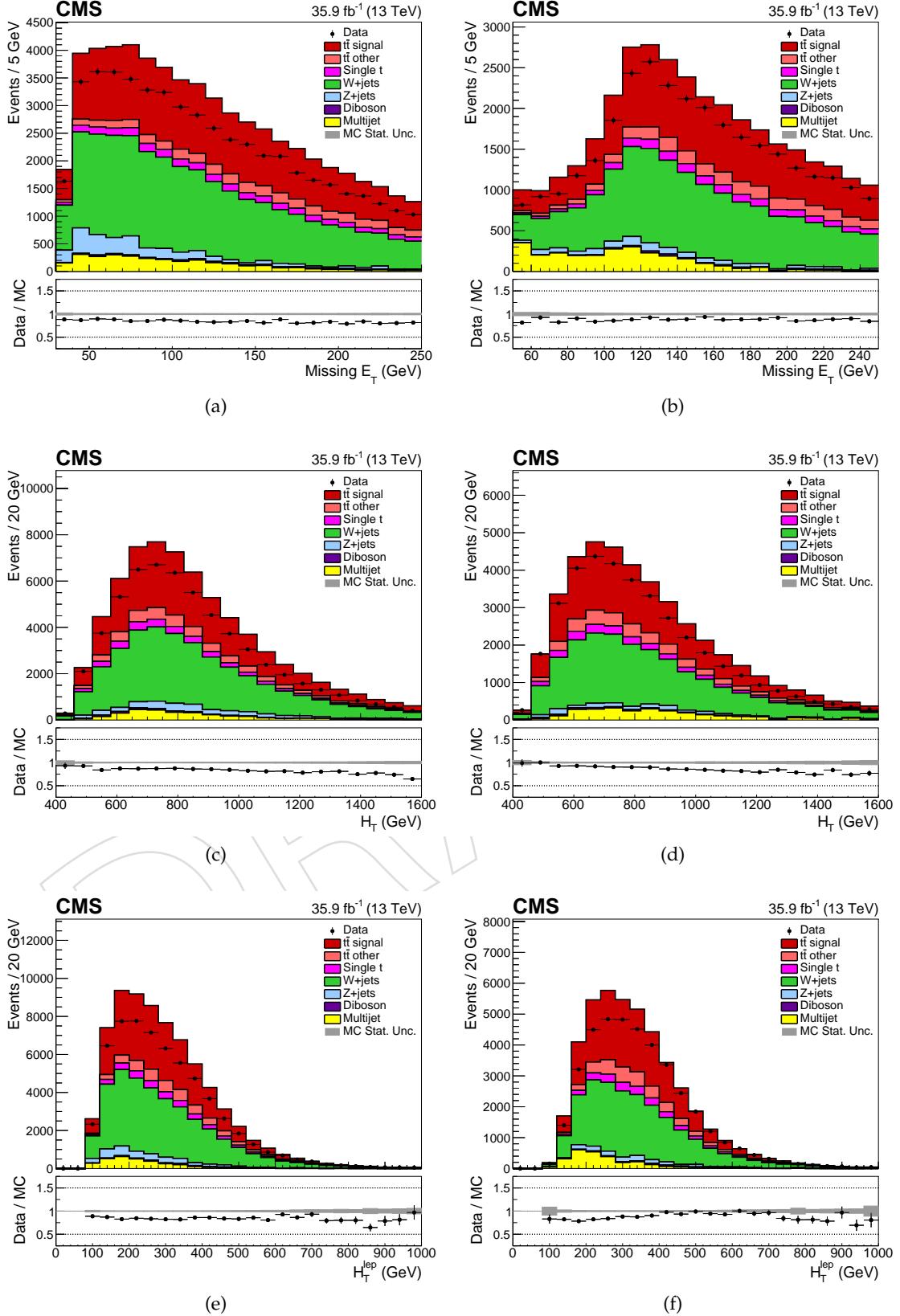


Figure 72: Event properties in the muon channel (left column) and electron channel (right column) for events passing the preselection

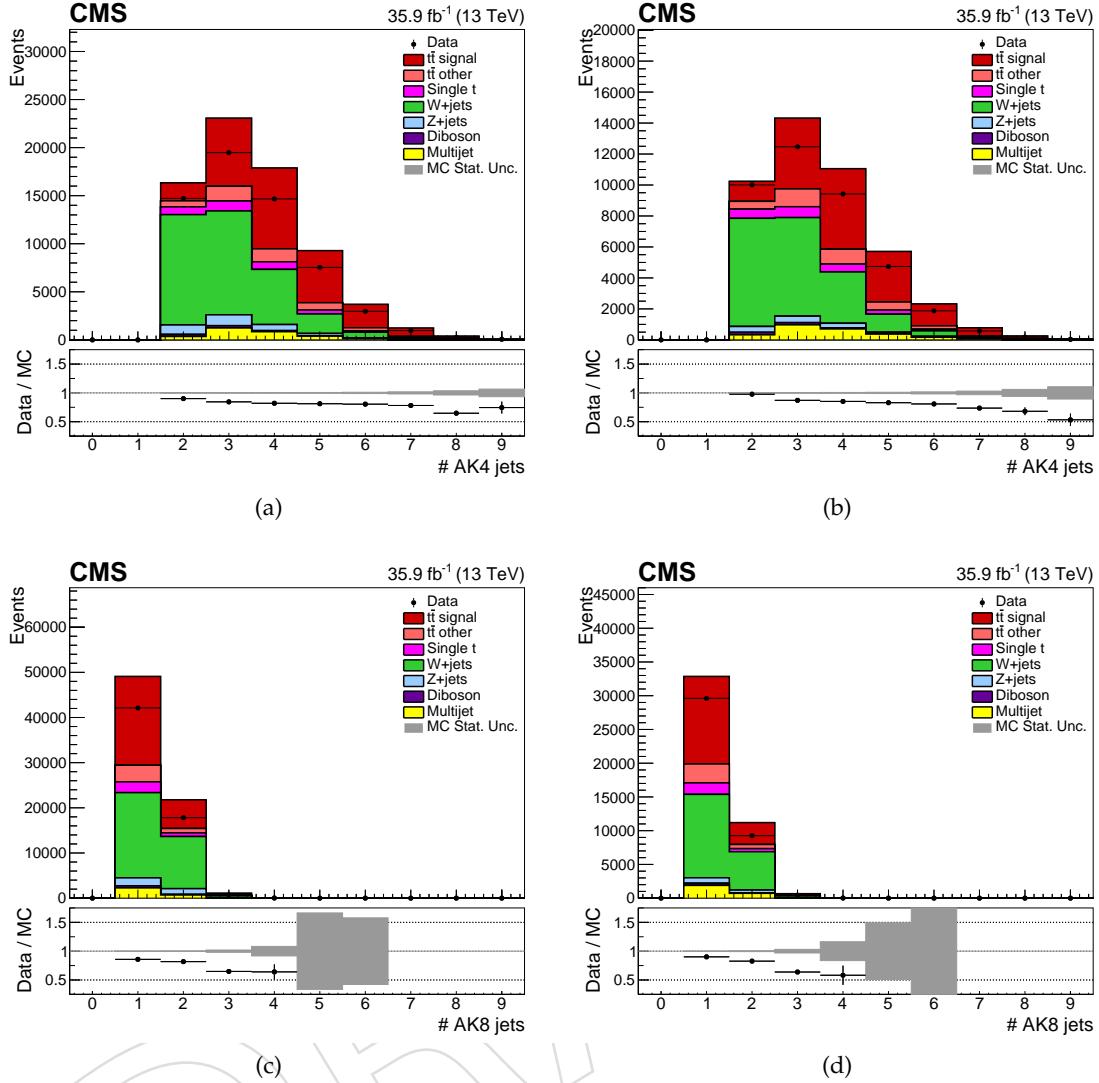


Figure 73: Event properties in the muon channel (left column) and electron channel (right column) for events passing the preselection

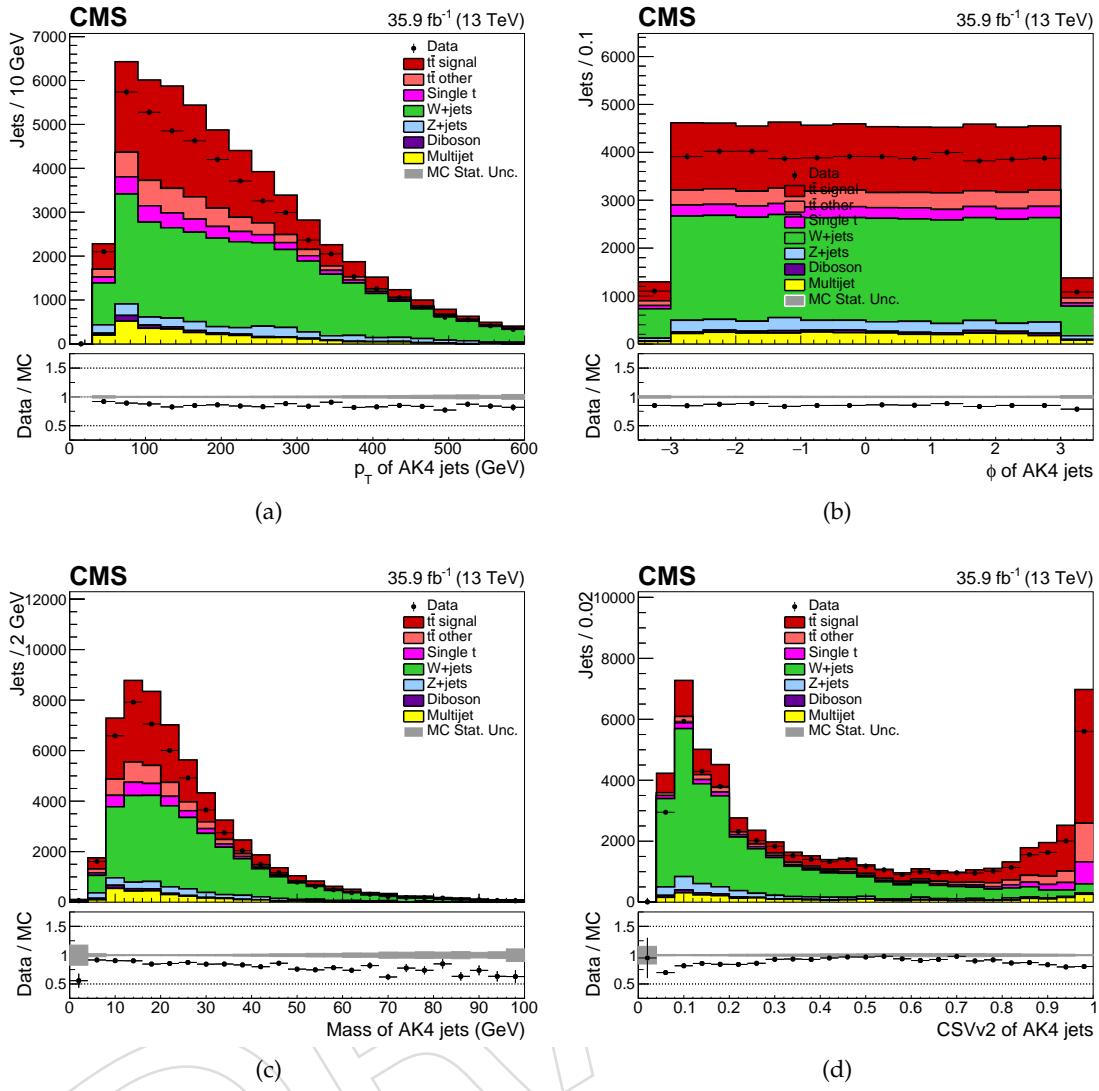


Figure 74: AK4 jet properties in the muon channel for events with 0 t-tags

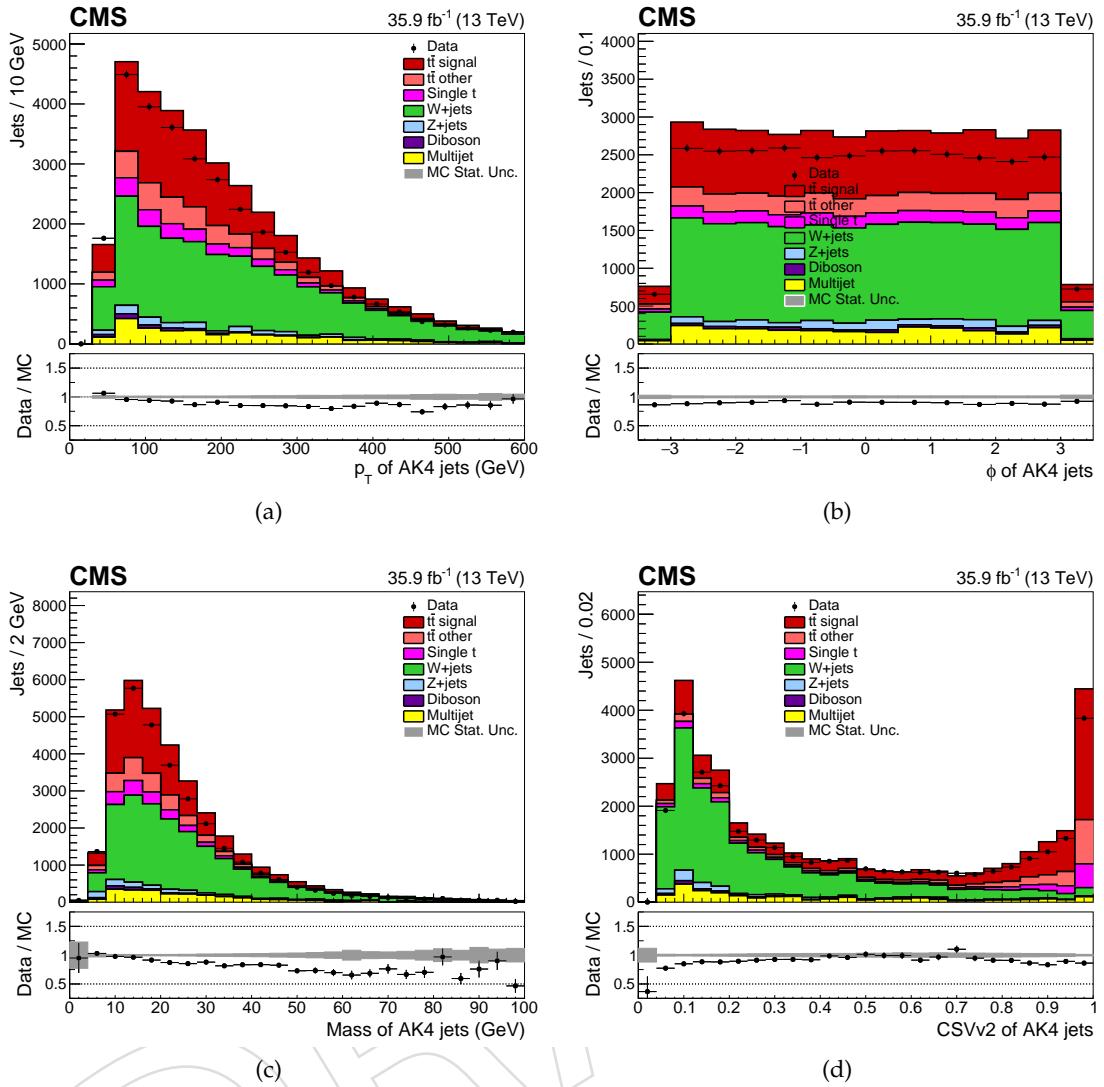


Figure 75: AK4 jet properties in the electron channel for events with 0 t-tags

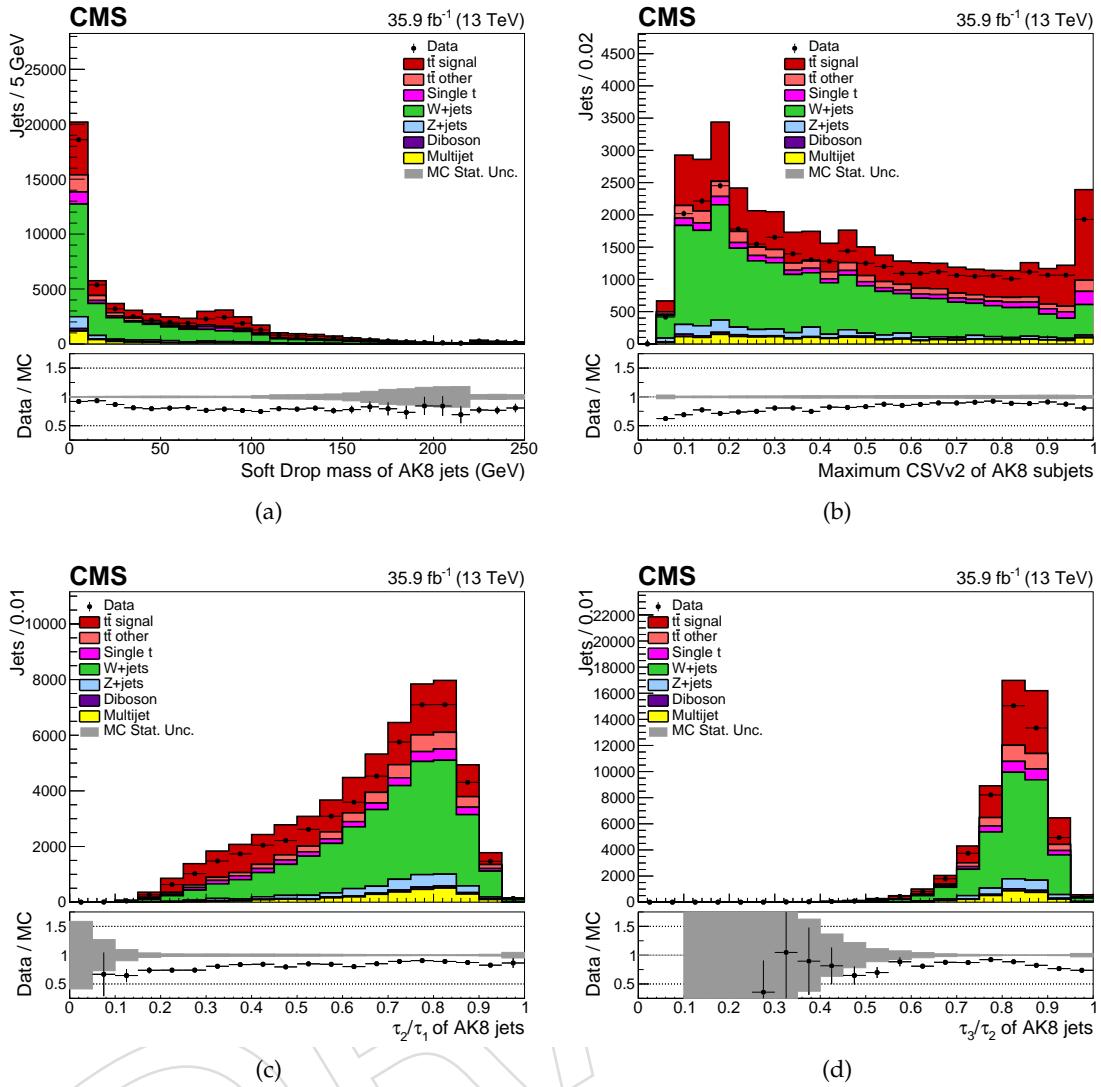


Figure 76: AK8 jet properties in the muon channel for events with 0 t-tags

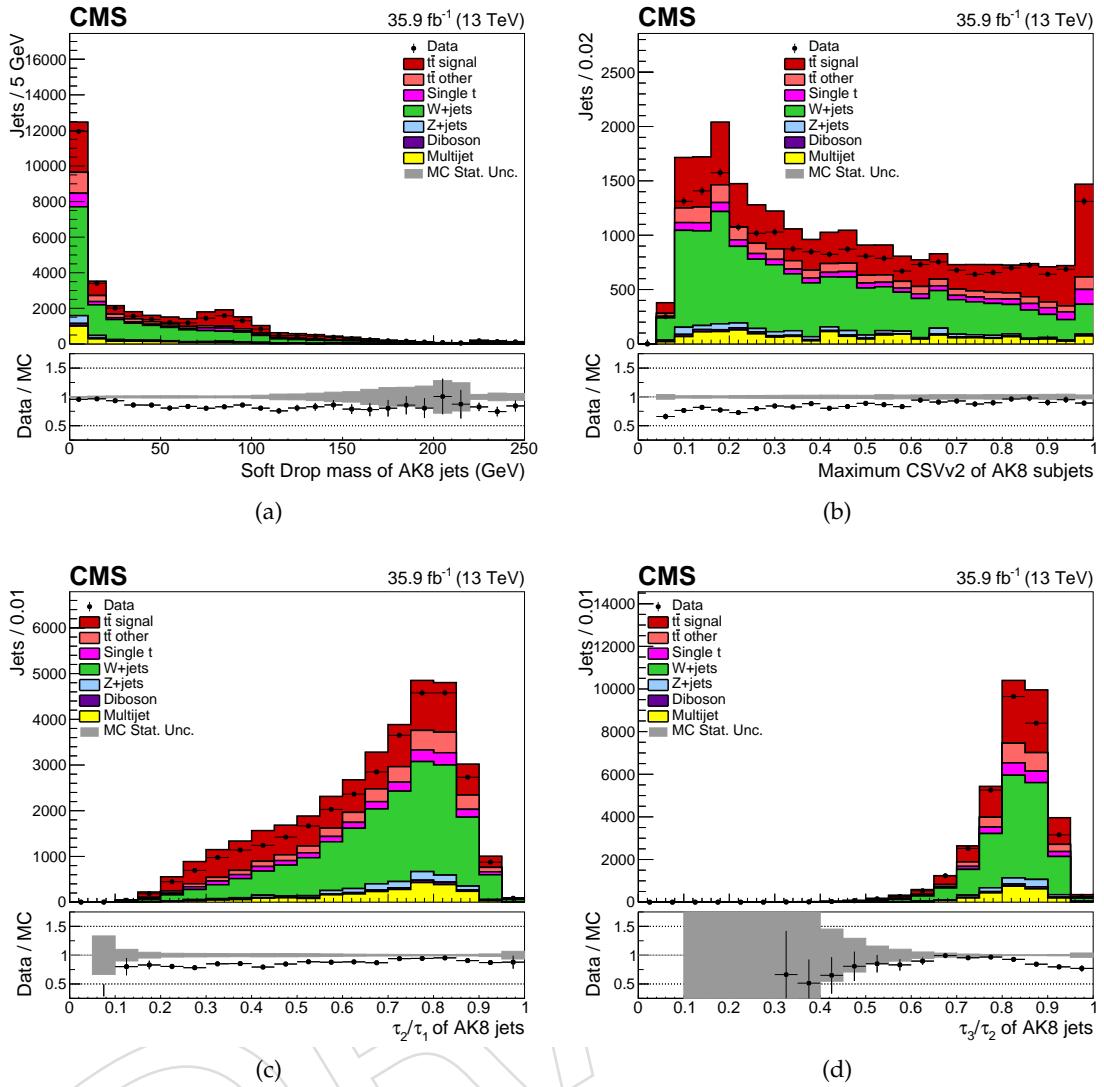


Figure 77: AK8 jet properties in the electron channel for events with 0 t-tags

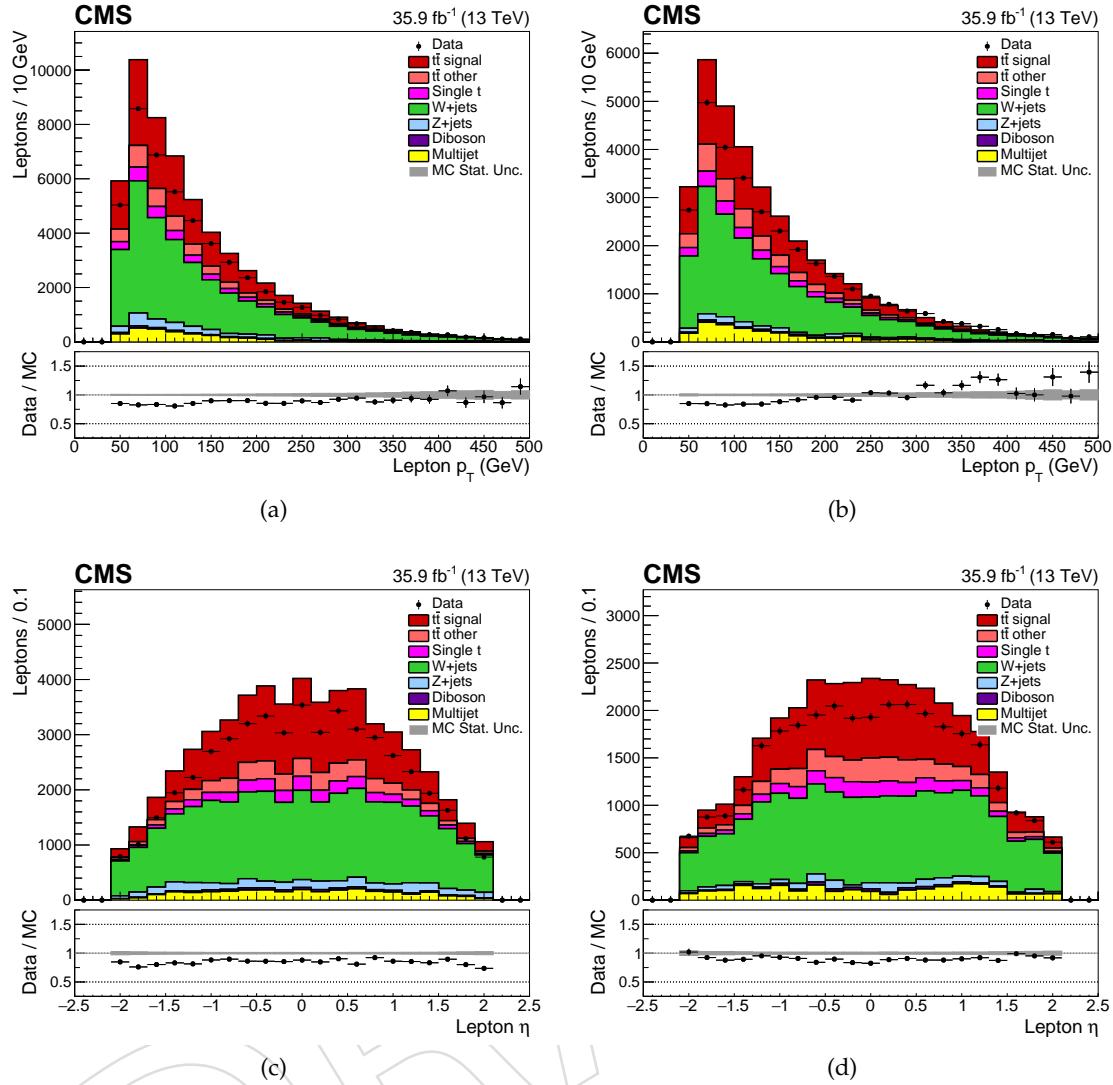


Figure 78: Lepton properties for events in the muon channel (left column) and electron channel (right column) with 0 t-tags

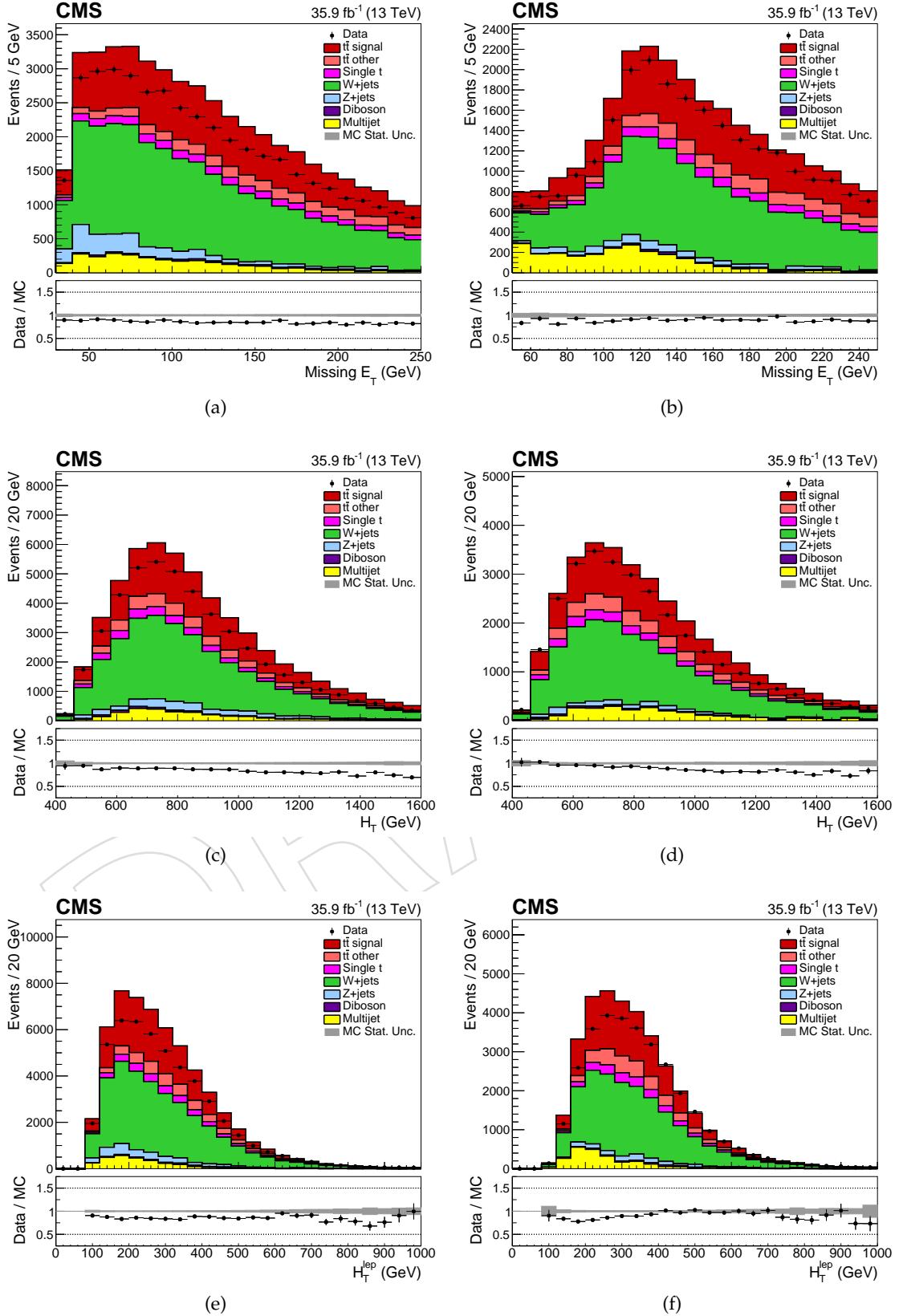


Figure 79: Event properties in the muon channel (left column) and electron channel (right column) for events with 0 t-tags

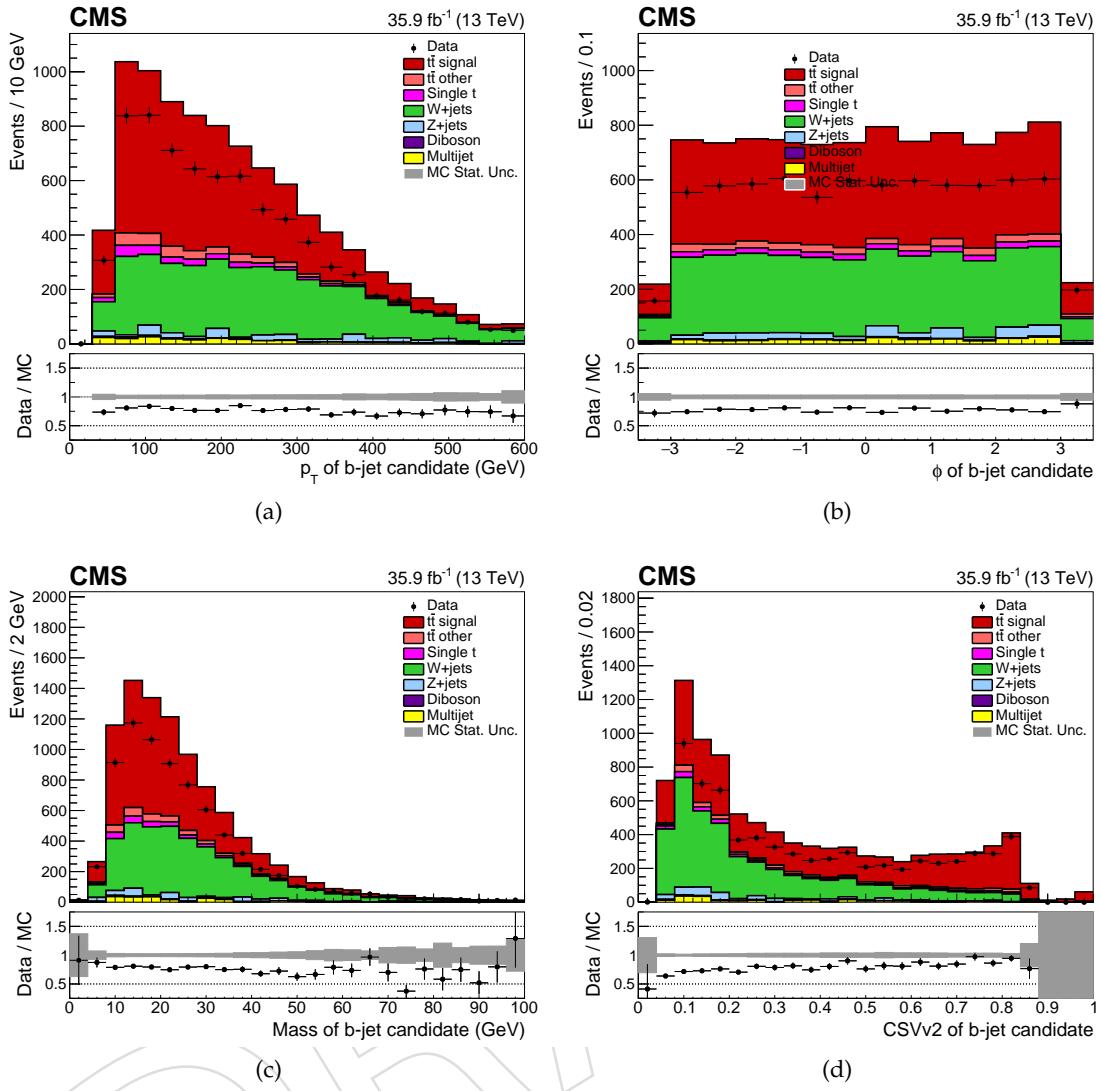


Figure 80: AK4 jet properties in the muon channel for events in the ≥ 1 t-tag, 0 b-tag region

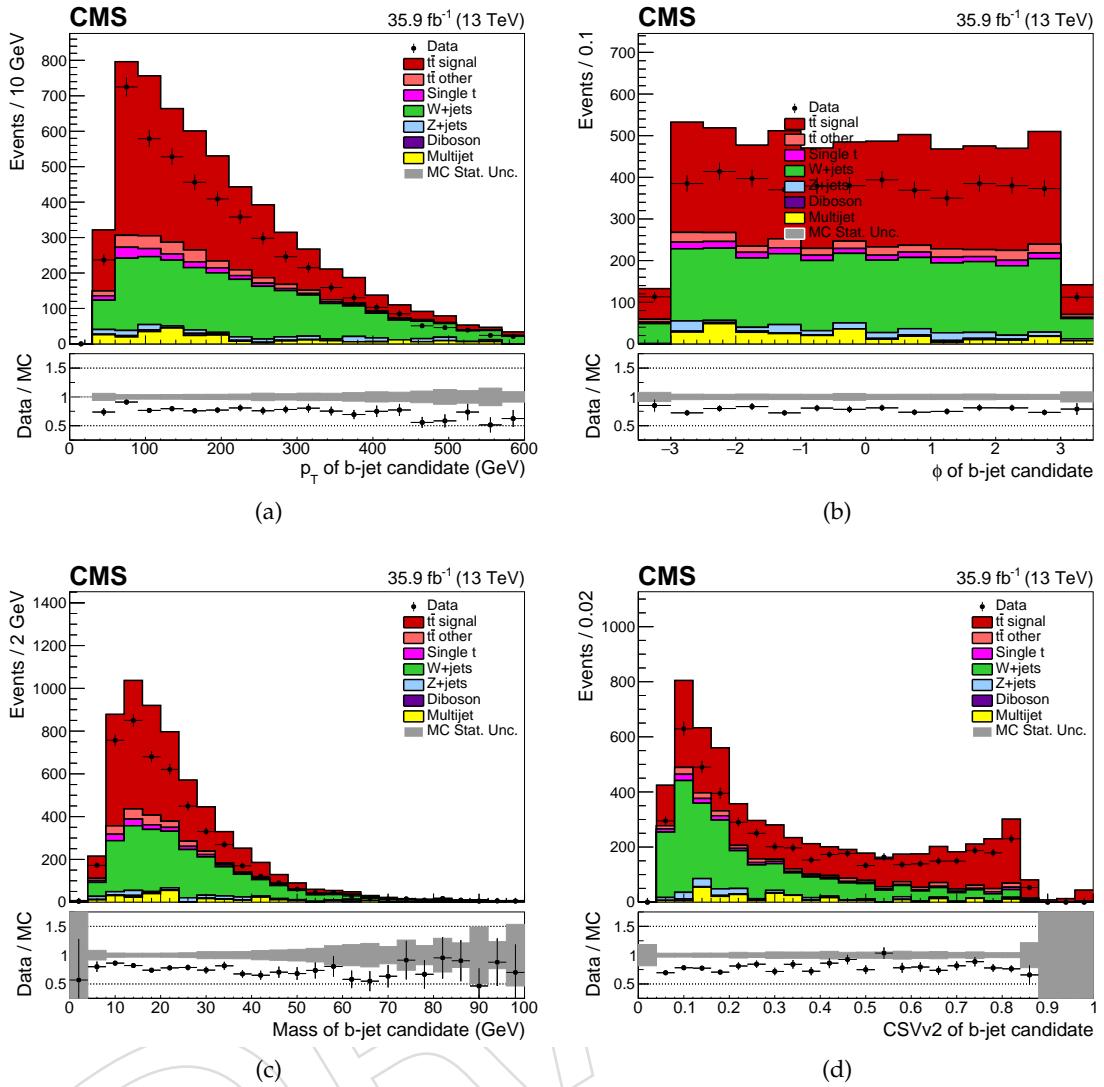


Figure 81: AK4 jet properties in the electron channel for events in the ≥ 1 t-tag, 0 b-tag region

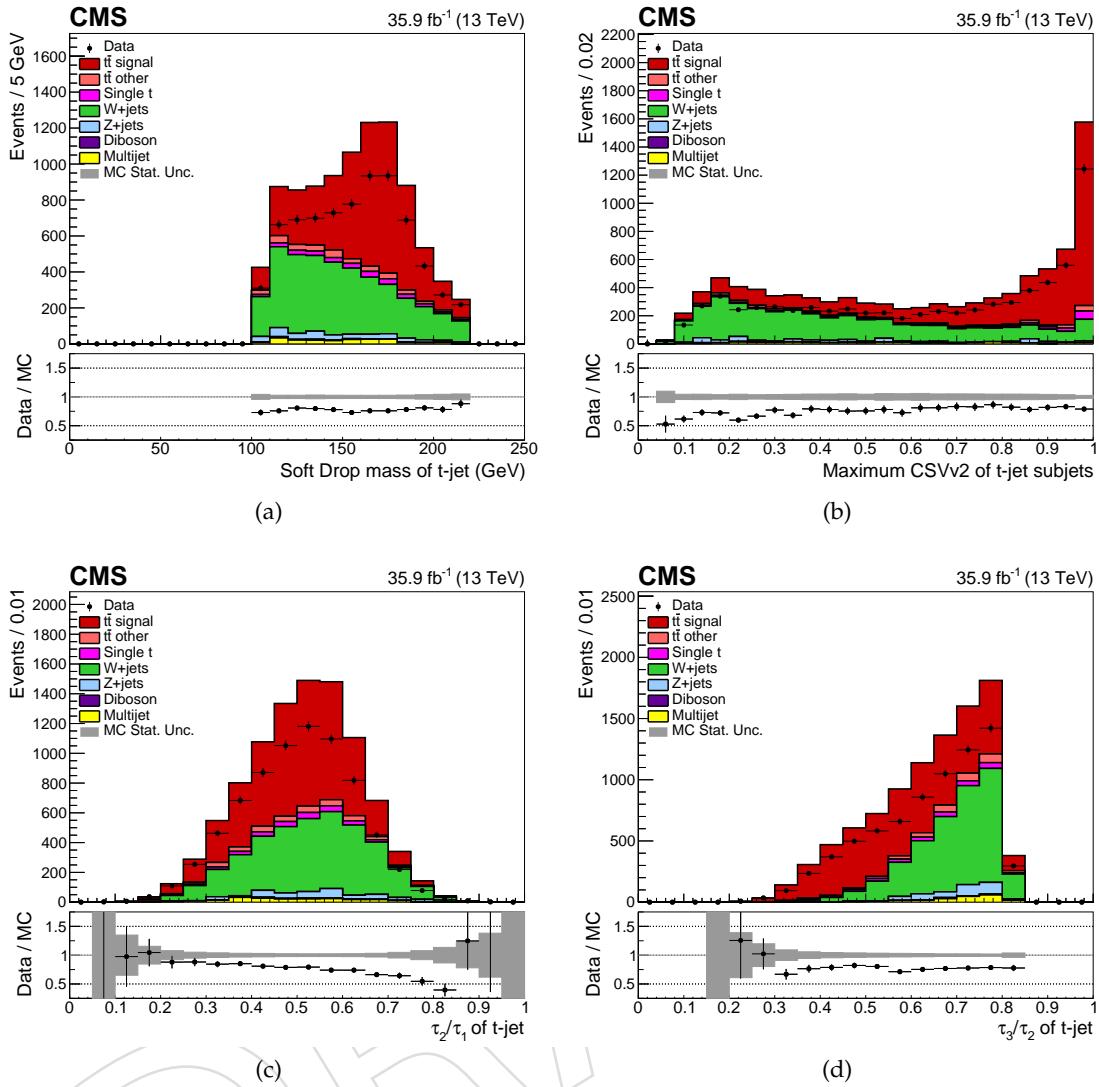


Figure 82: AK8 jet properties in the muon channel for events in the ≥ 1 t-tag, 0 b-tag region

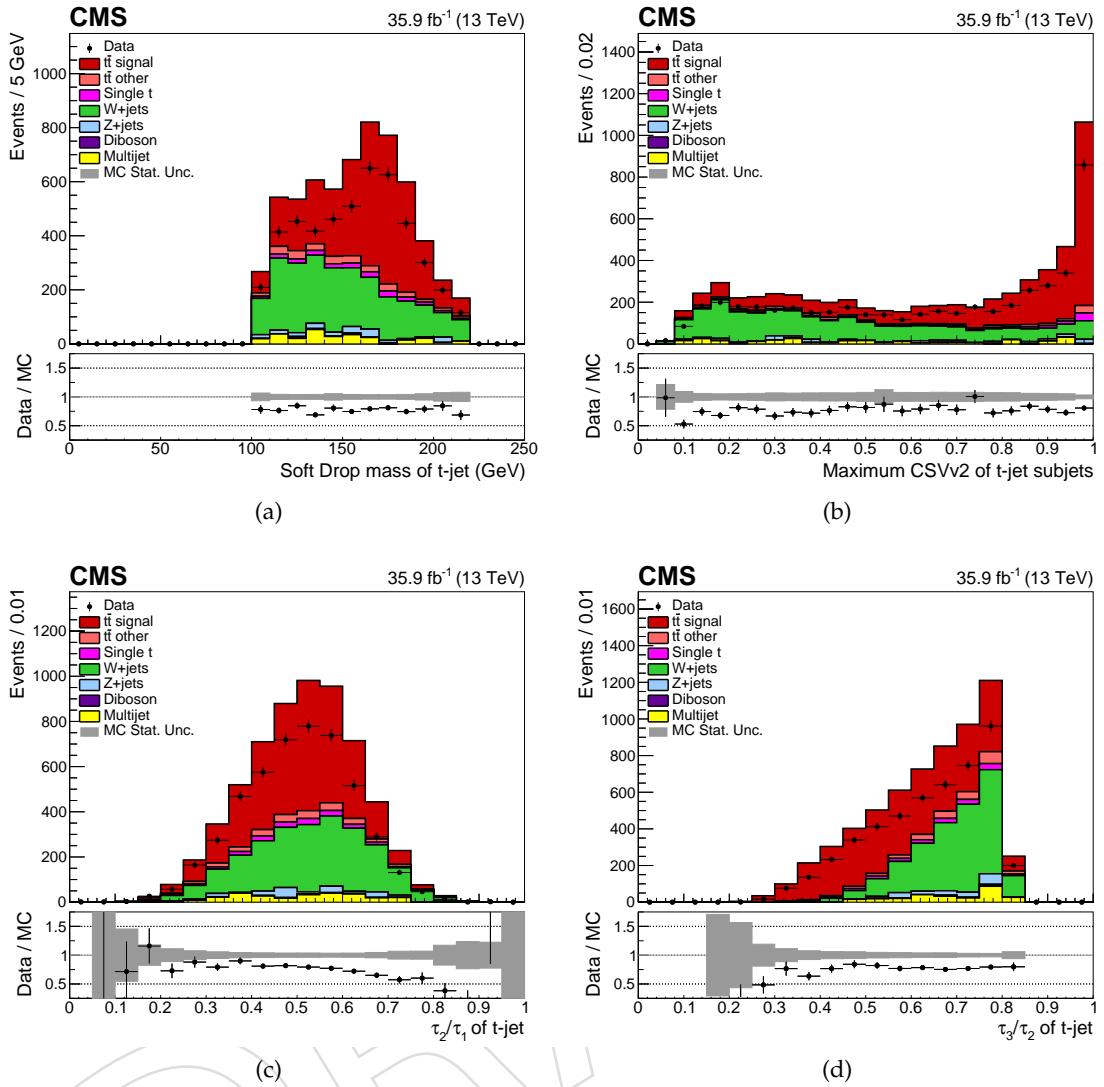


Figure 83: AK8 jet properties in the electron channel for events in the ≥ 1 t-tag, 0 b-tag region

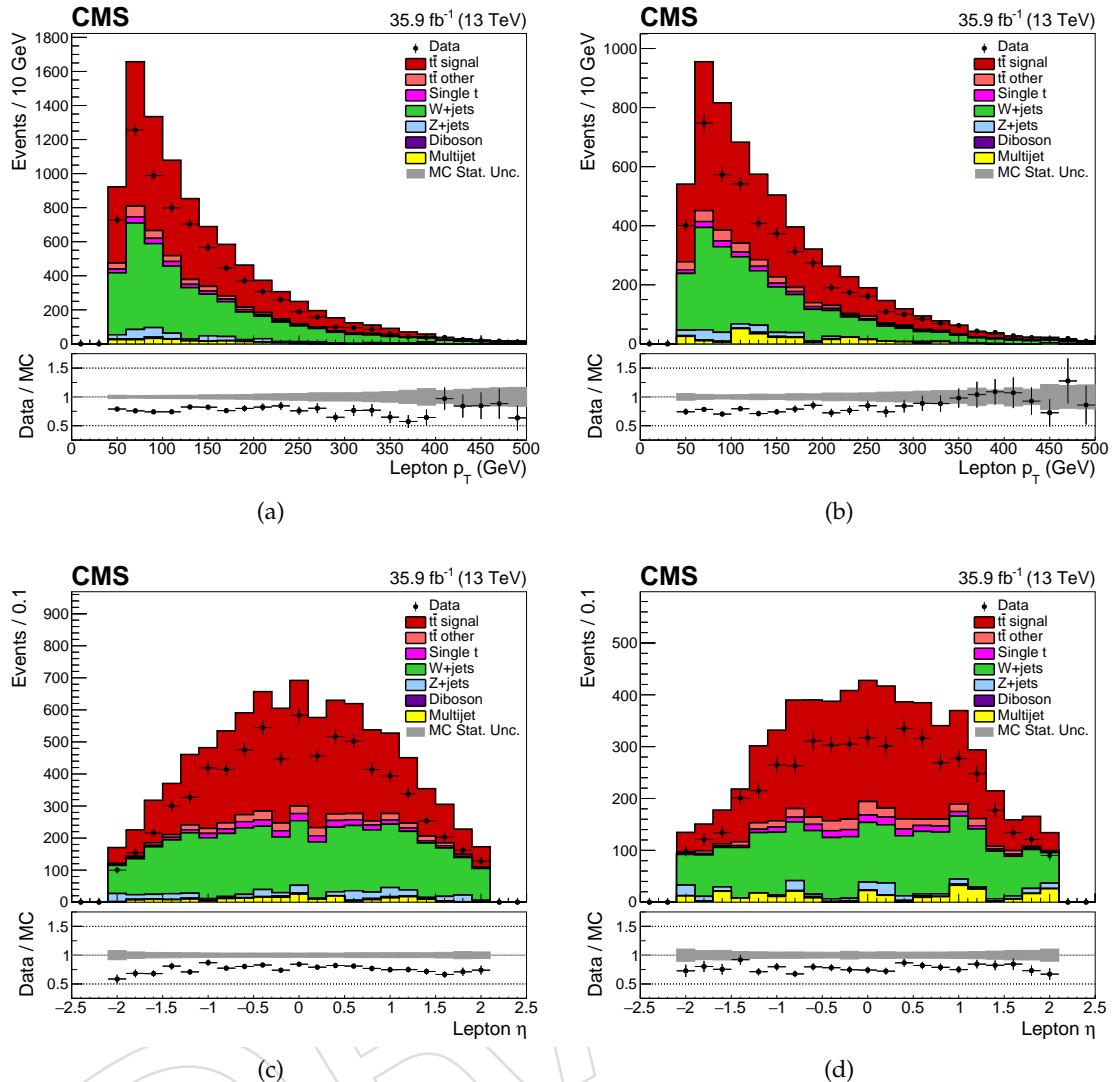


Figure 84: Lepton properties for events in the muon channel (left column) and electron channel (right column) in the ≥ 1 t-tag, 0 b-tag region

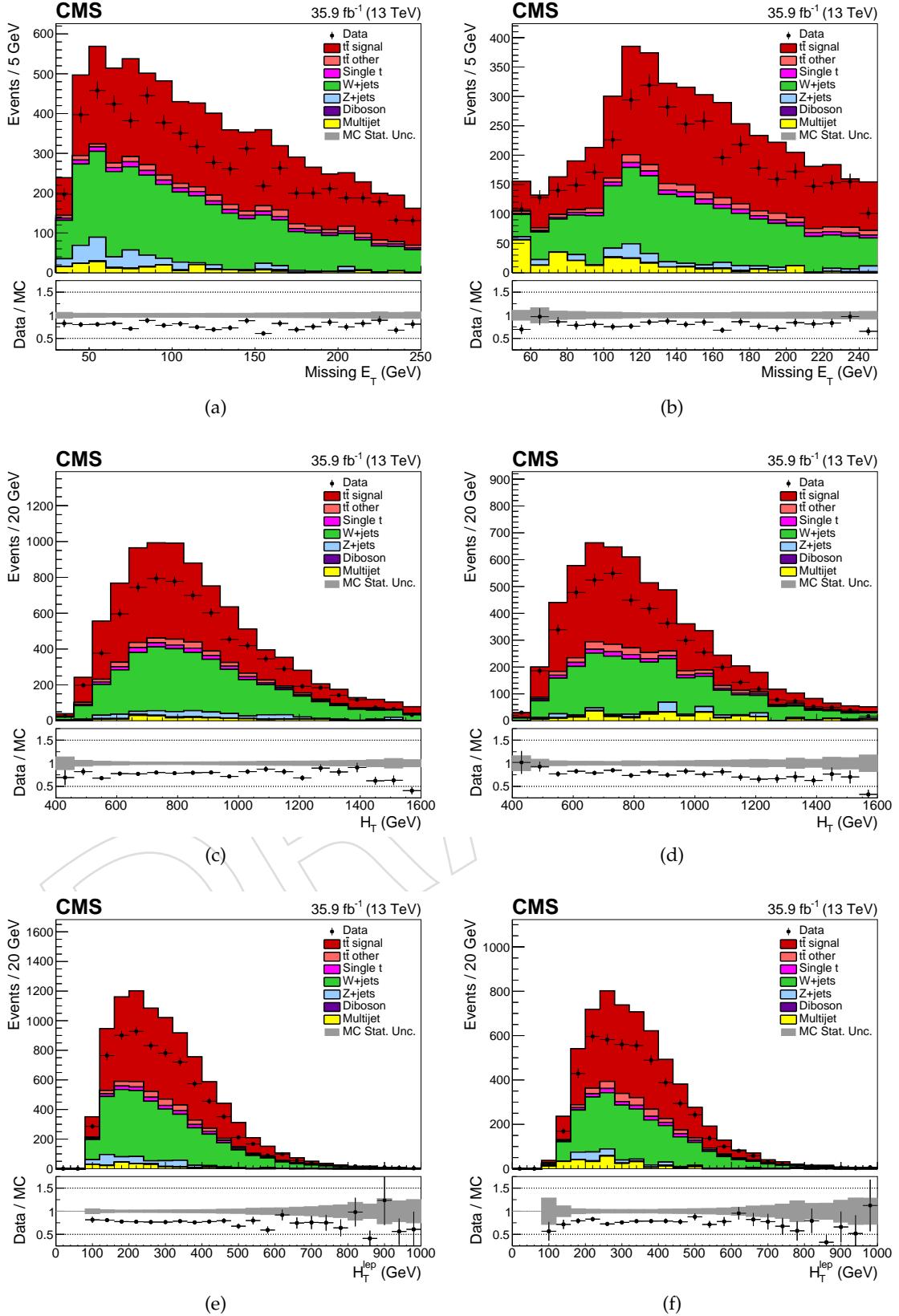


Figure 85: Event properties in the muon channel (left column) and electron channel (right column) for events in the ≥ 1 t-tag, 0 b-tag region

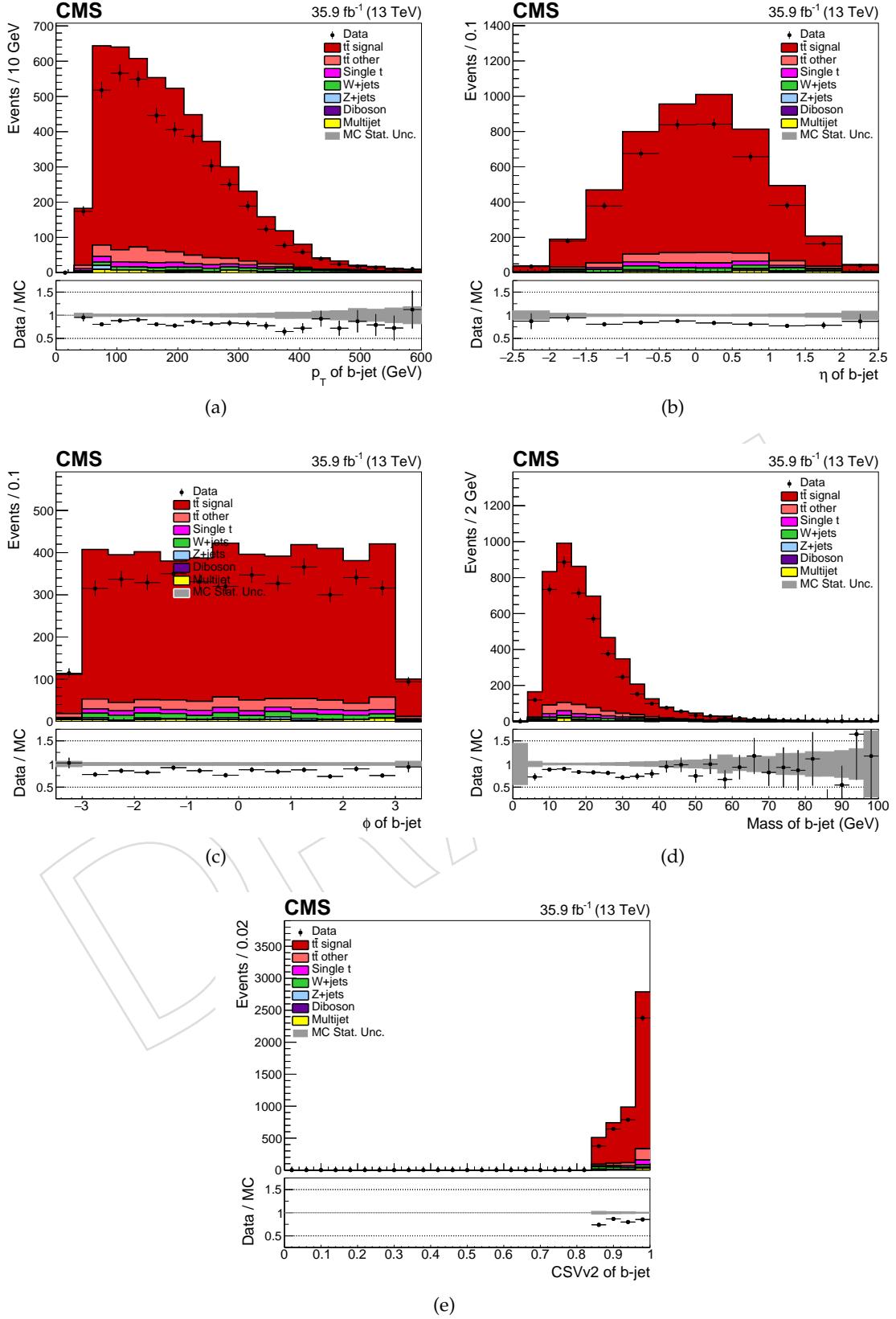


Figure 86: AK4 jet properties in the muon channel for events in the ≥ 1 t-tag, ≥ 1 b-tag region

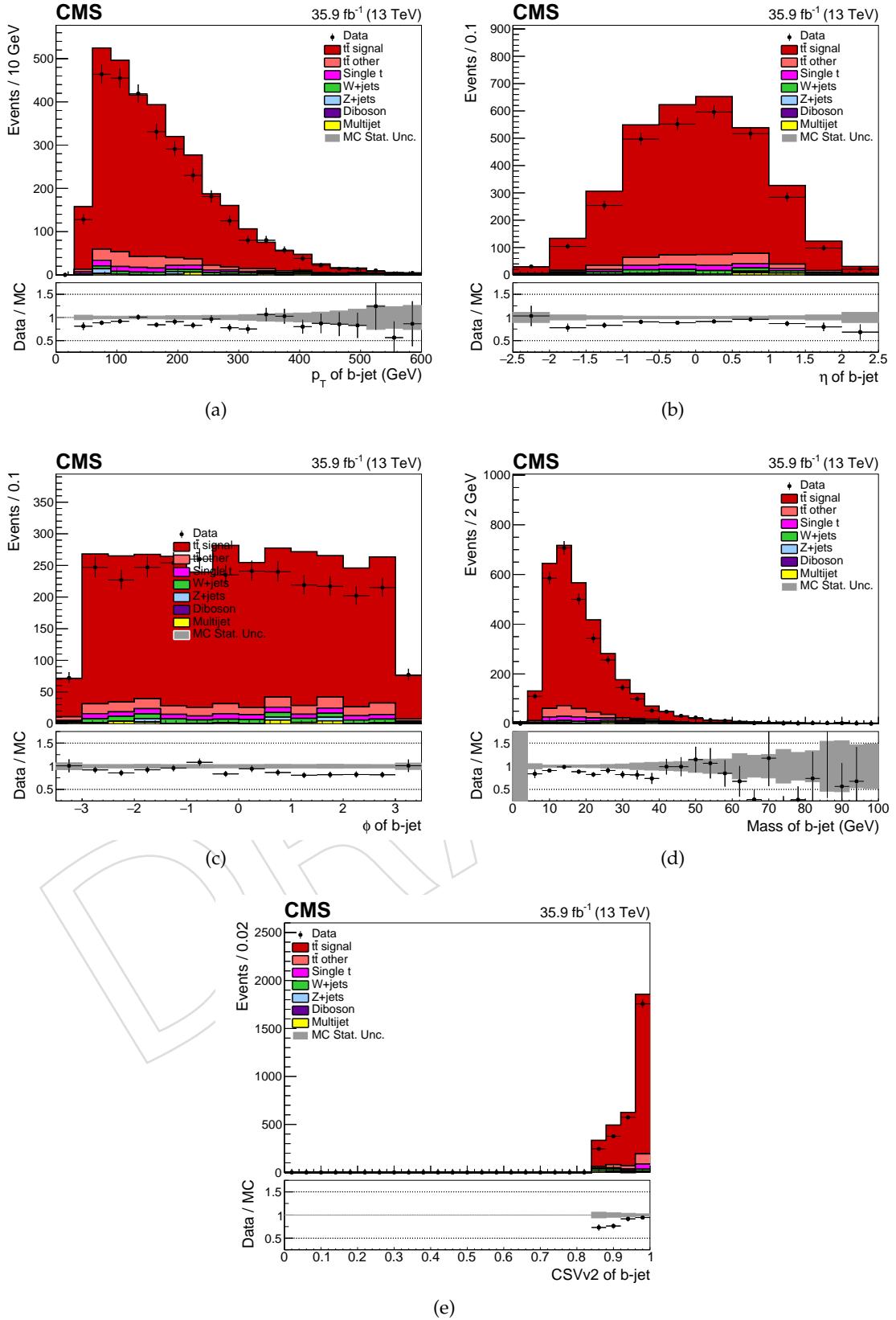


Figure 87: AK4 jet properties in the electron channel for events in the ≥ 1 t-tag, ≥ 1 b-tag region

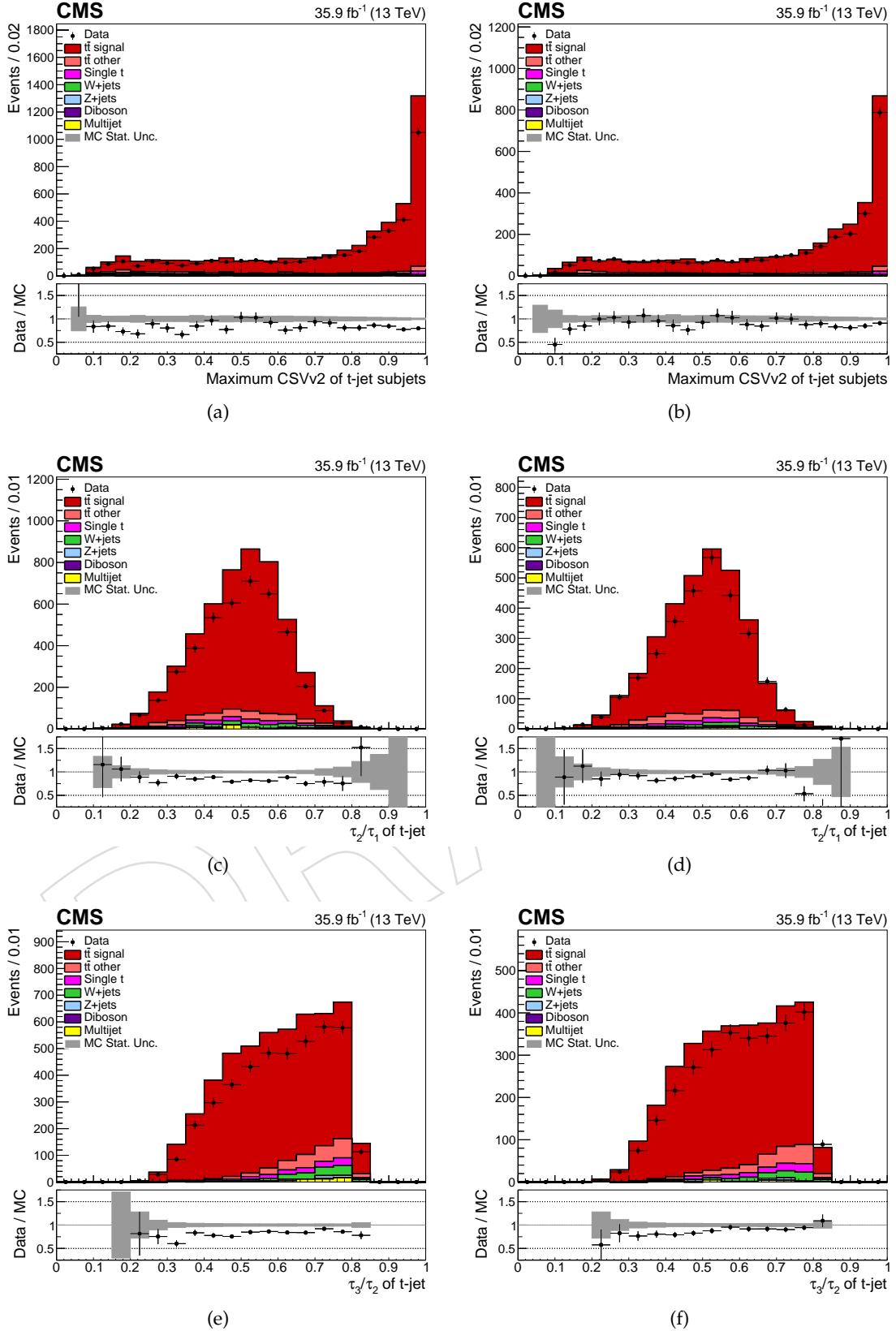


Figure 88: AK8 jet properties in the muon channel (left column) and electron channel (right column) for events in the ≥ 1 t-tag, ≥ 1 b-tag region

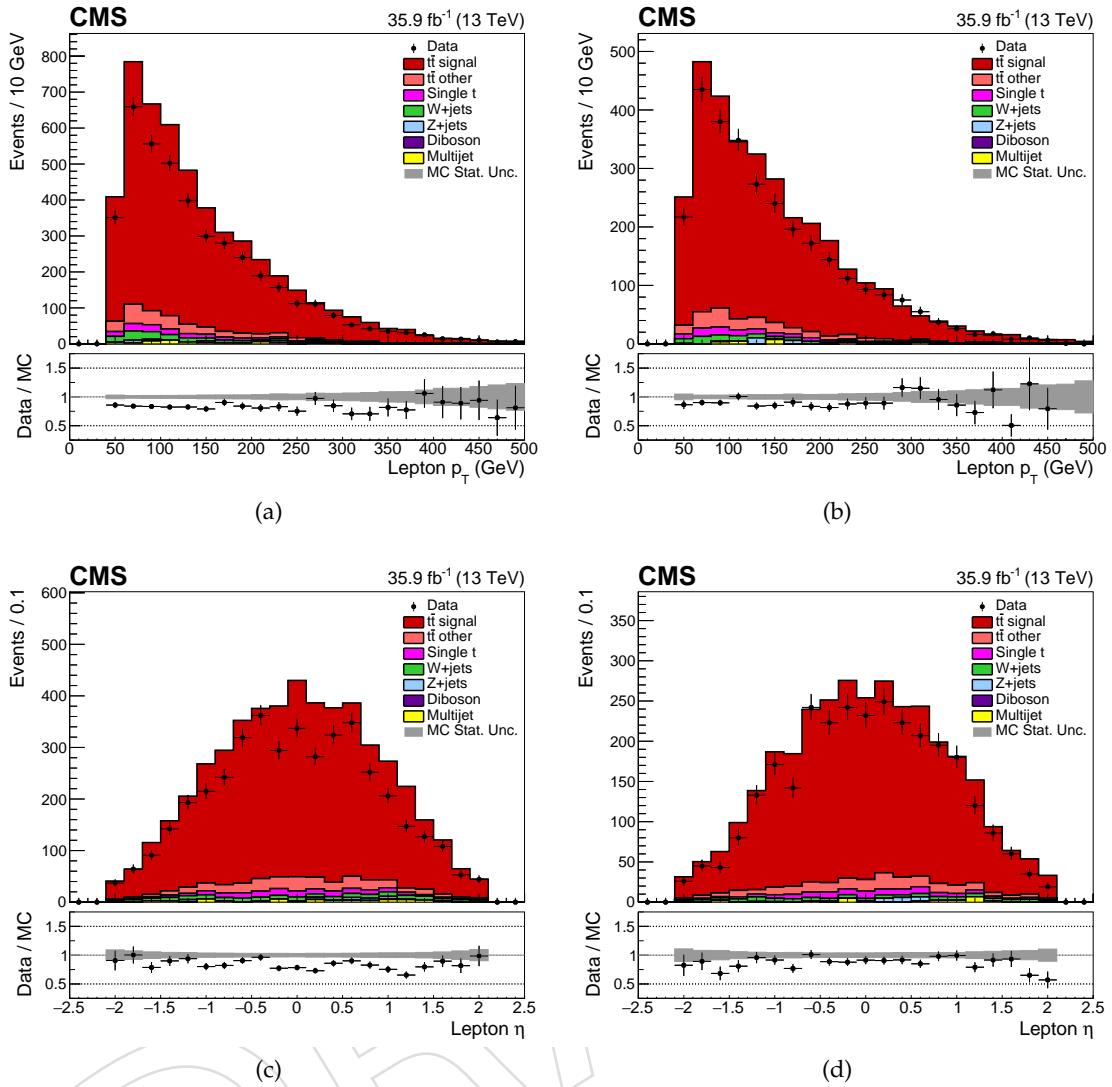


Figure 89: Lepton properties in the muon channel (left column) and electron channel (right column) for events in the ≥ 1 t-tag, ≥ 1 b-tag region

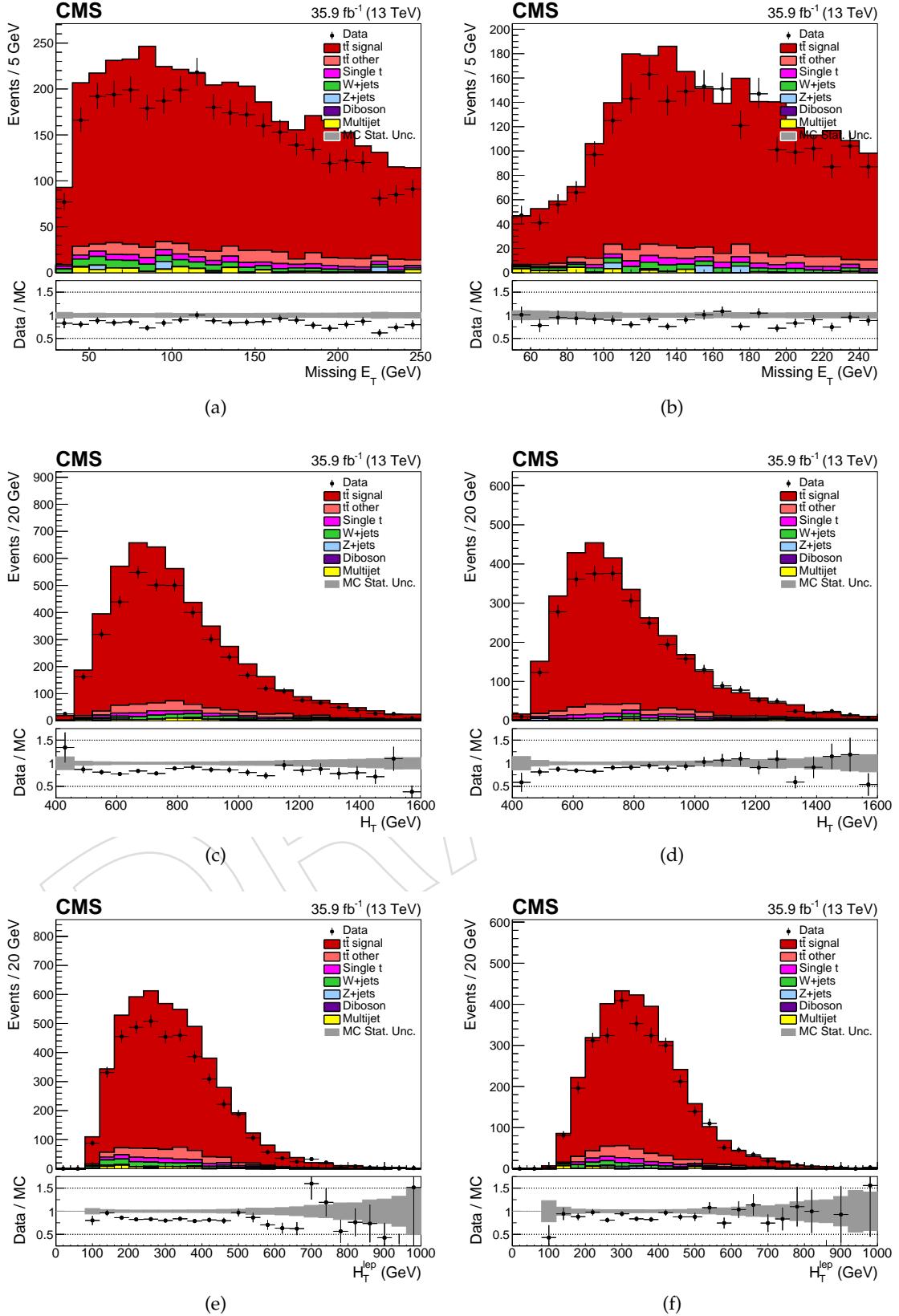


Figure 90: Event properties in the muon channel (left column) and electron channel (right column) for events in the ≥ 1 t-tag, ≥ 1 b-tag region

1003 F Fit Cross Checks

1004 In order to assess the stability of the kinematic likelihood fit, several cross checks were per-
 1005 formed. Particular emphasis was placed on understanding the fit dependence on the choice of
 1006 fit templates. Fits were performed using the AK4 jet $|\eta|$, lepton η , and lepton $|\eta|$ in place of
 1007 the AK4 jet η in the 0 top-tag and 1 top-tag 0 b-tag regions. Since the 1 top-tag 1 b-tag region
 1008 was dominated by signal, no template variation was performed in this region. In addition, fits
 1009 were performed using a QCD template modeled using MC in place of data, to assess the effect
 1010 of the QCD mismodelling on the fit result. A summary of the various template choices is given
 1011 in Table 23. In addition, a counting fit was performed fitting the event counts in the 0 top-tag,
 1012 1 top-tag 0 b-tag, and 1 top-tag 1 b-tag regions. As in the nominal case, the fit variations were
 1013 first performed separately in the muon and electron channels, then simultaneously combining
 1014 both channels.

Fit	Templates			QCD
	0 top-tag	1 top-tag 0 b-tag	1 top-tag 1 b-tag	
AK4 $ \eta $	AK4 jet $ \eta $	AK4 jet $ \eta $	AK8 jet SoftDrop mass	Data
Lepton η	Lepton η	Lepton η	AK8 jet SoftDrop mass	Data
Lepton $ \eta $	Lepton $ \eta $	Lepton $ \eta $	AK8 jet SoftDrop mass	Data
AK4 η , MC QCD	AK4 jet η	AK4 jet η	AK8 jet SoftDrop mass	MC
AK4 $ \eta $, MC QCD	AK4 jet $ \eta $	AK4 jet $ \eta $	AK8 jet SoftDrop mass	MC
Counts	Fit 0t, 1t0b, 1t1b event counts			N/A

Table 23: Fit templates used for the different fit variants, along with method of modelling the QCD shape.

- 1015 The nuisance parameters for the fit variants are shown in Figures 91, 94, 97, 100, 103, and 106
 1016 for the AK4 $|\eta|$, lepton η , lepton $|\eta|$, MC QCD AK4 η , MC QCD AK4 $|\eta|$, and counts fits respec-
 1017 tively.
- 1018 The correlations between nuisance parameters are shown in Figures 92, 95, 98, 101, 104, and 107
 1019 for the AK4 $|\eta|$, lepton η , lepton $|\eta|$, MC QCD AK4 η , MC QCD AK4 $|\eta|$, and counts fits respec-
 1020 tively.
- 1021 The posterior kinematic distributions in the three fit categories are shown in Figures 93, 96, 99, 102, 105,
 1022 and 108 for the AK4 $|\eta|$, lepton η , lepton $|\eta|$, MC QCD AK4 η , MC QCD AK4 $|\eta|$, and counts
 1023 fits respectively. The posterior kinematic distributions are shown for the combined fit result.

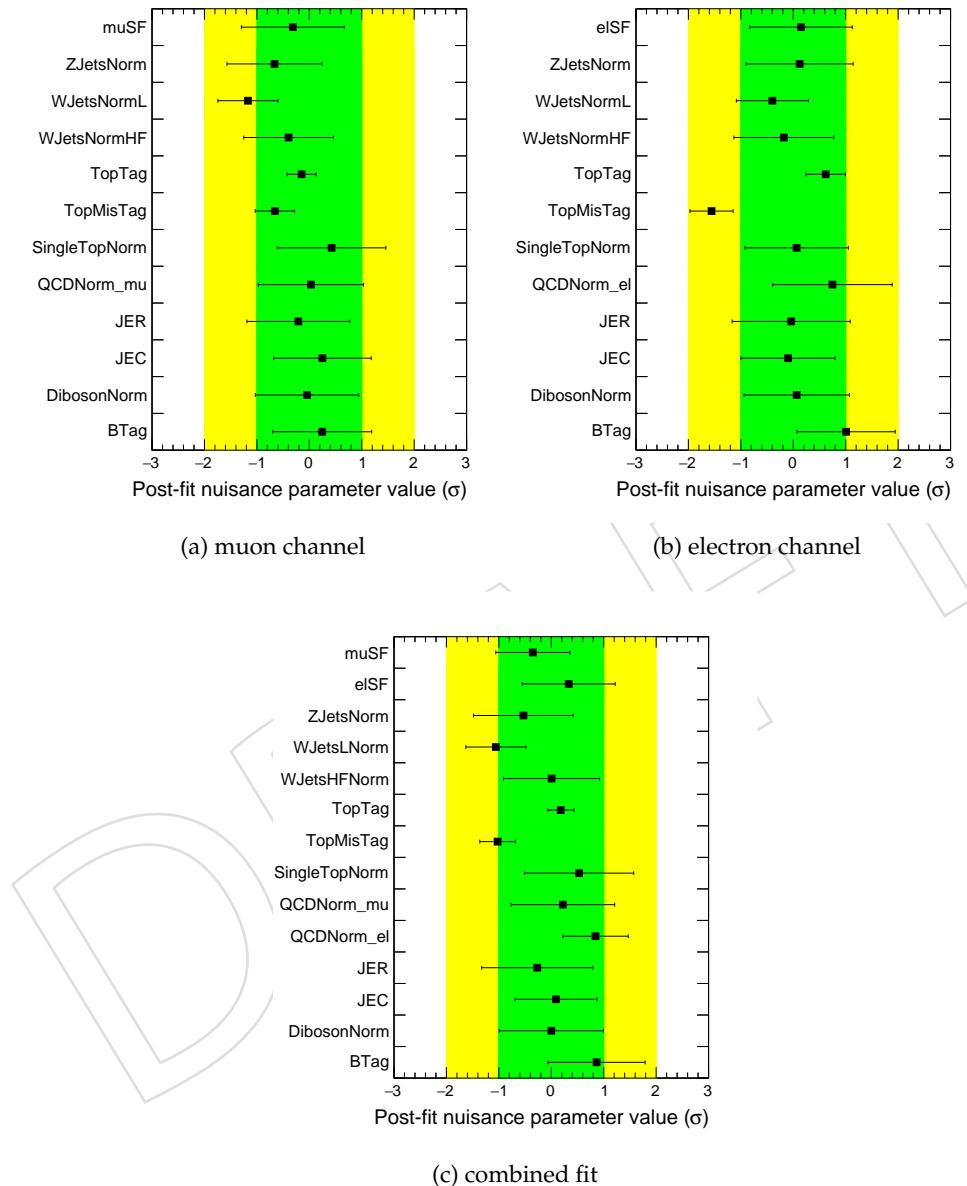


Figure 91: Post-fit nuisance parameters for the AK4 $|\eta|$ fit. Nuisance parameters compatible with zero indicate no modification to the systematic by the fit; nuisance parameter uncertainties of 1σ indicate no constraint on the systematic uncertainty due to the fit.

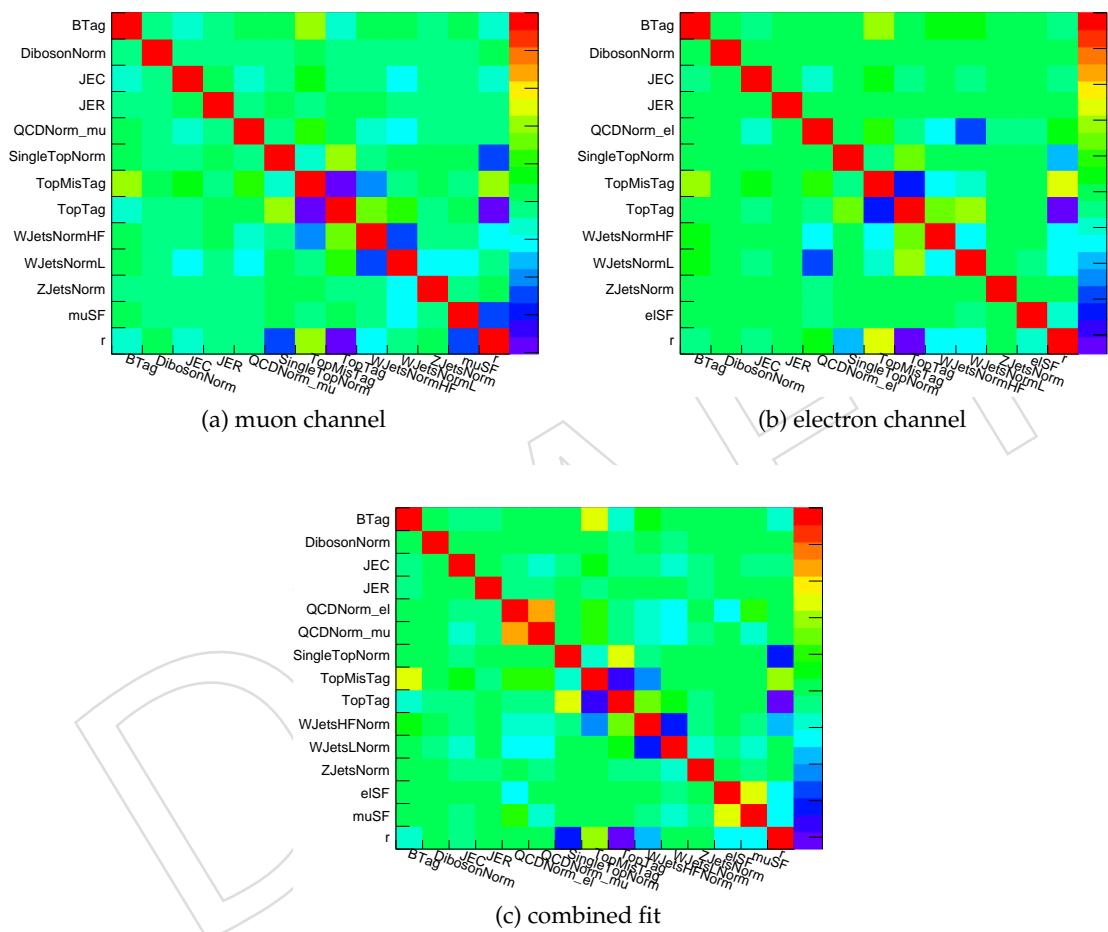


Figure 92: Correlations between nuisance parameters for the AK4 $|\eta|$ fit.

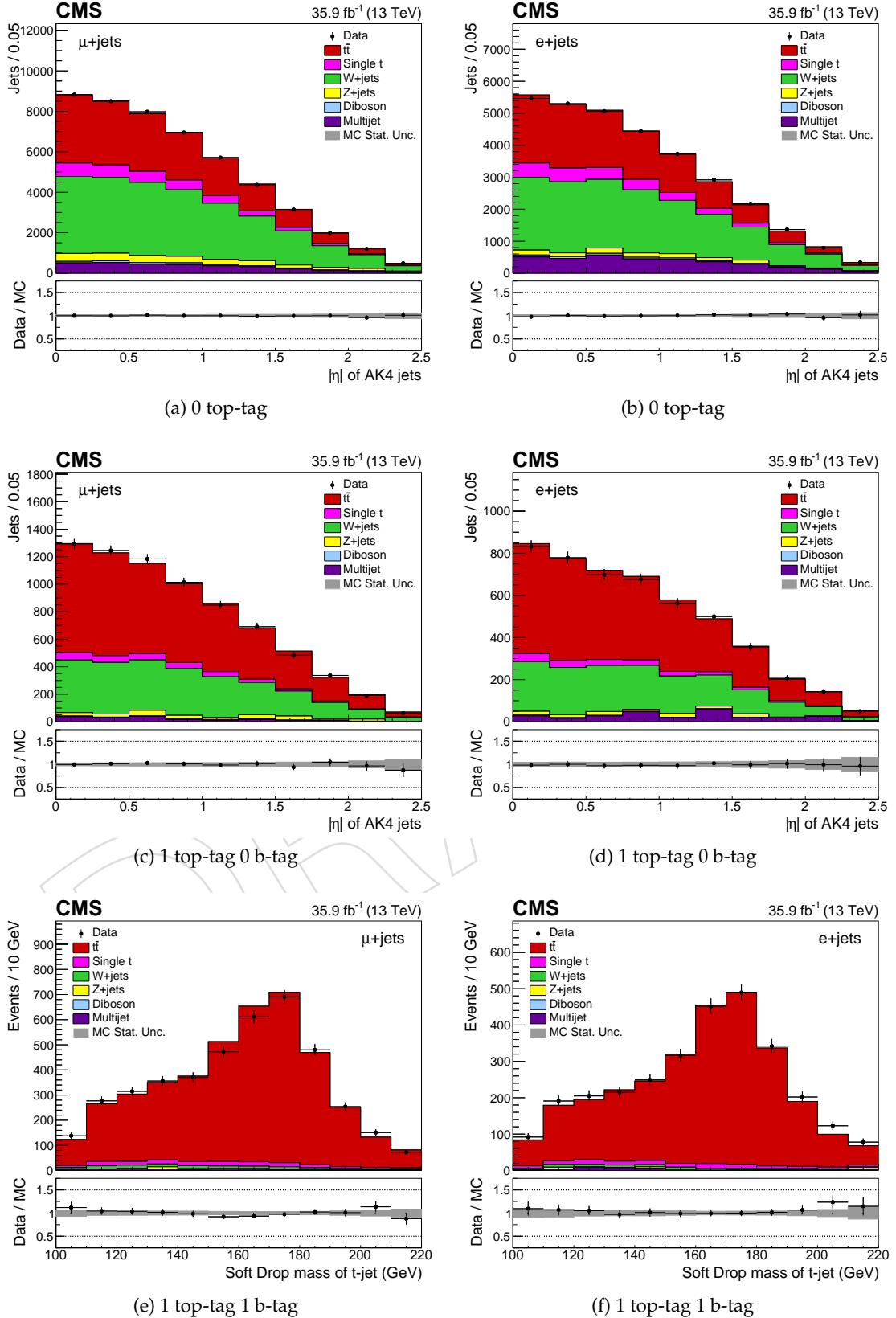


Figure 93: Post-fit kinematic distributions in the muon (left) and electron (right) channels for the AK4 $|\eta|$ fit combining both channels. For this fit, the QCD shape is taken from a data sideband.

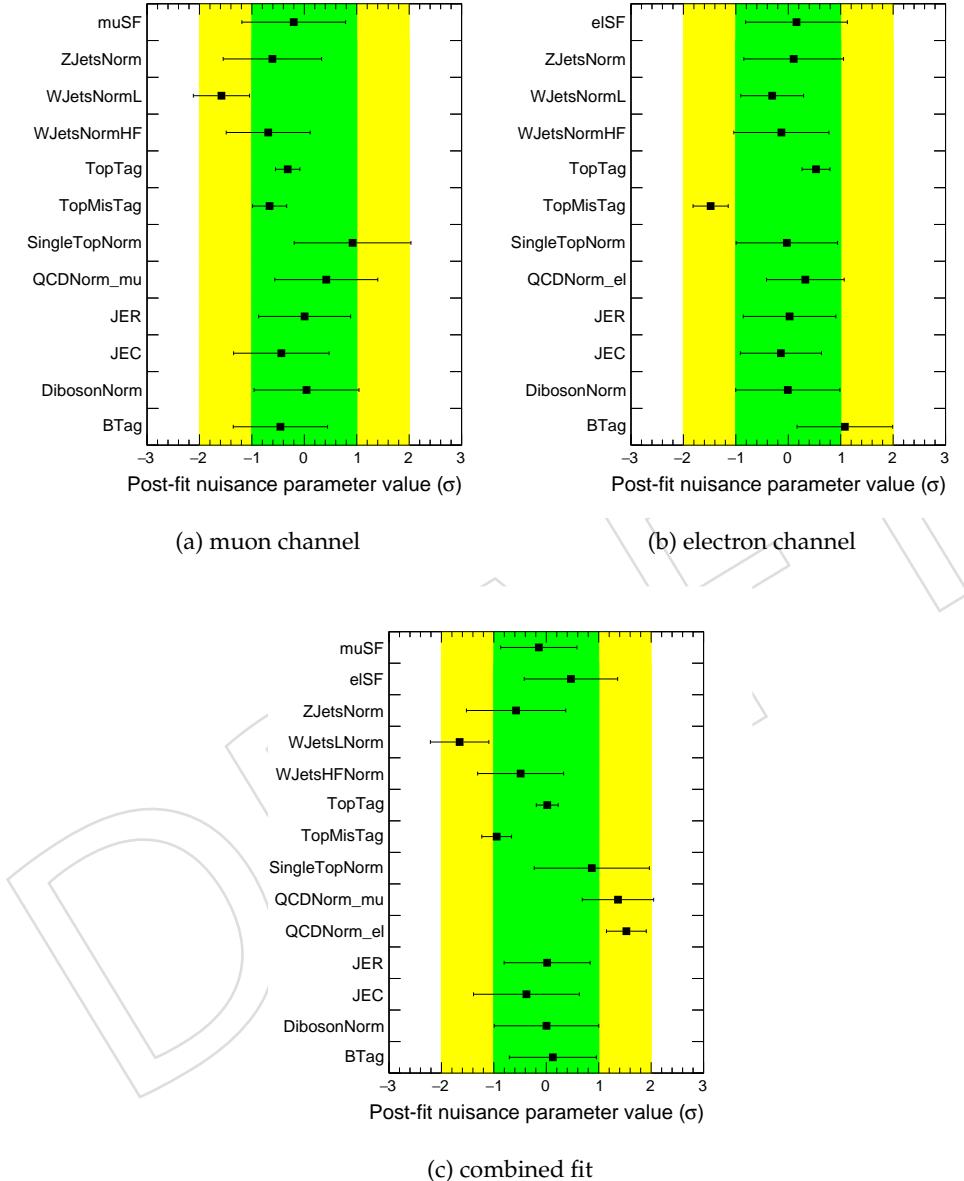


Figure 94: Post-fit nuisance parameters for the lepton η fit. Nuisance parameters compatible with zero indicate no modification to the systematic by the fit; nuisance parameter uncertainties of 1σ indicate no constraint on the systematic uncertainty due to the fit.

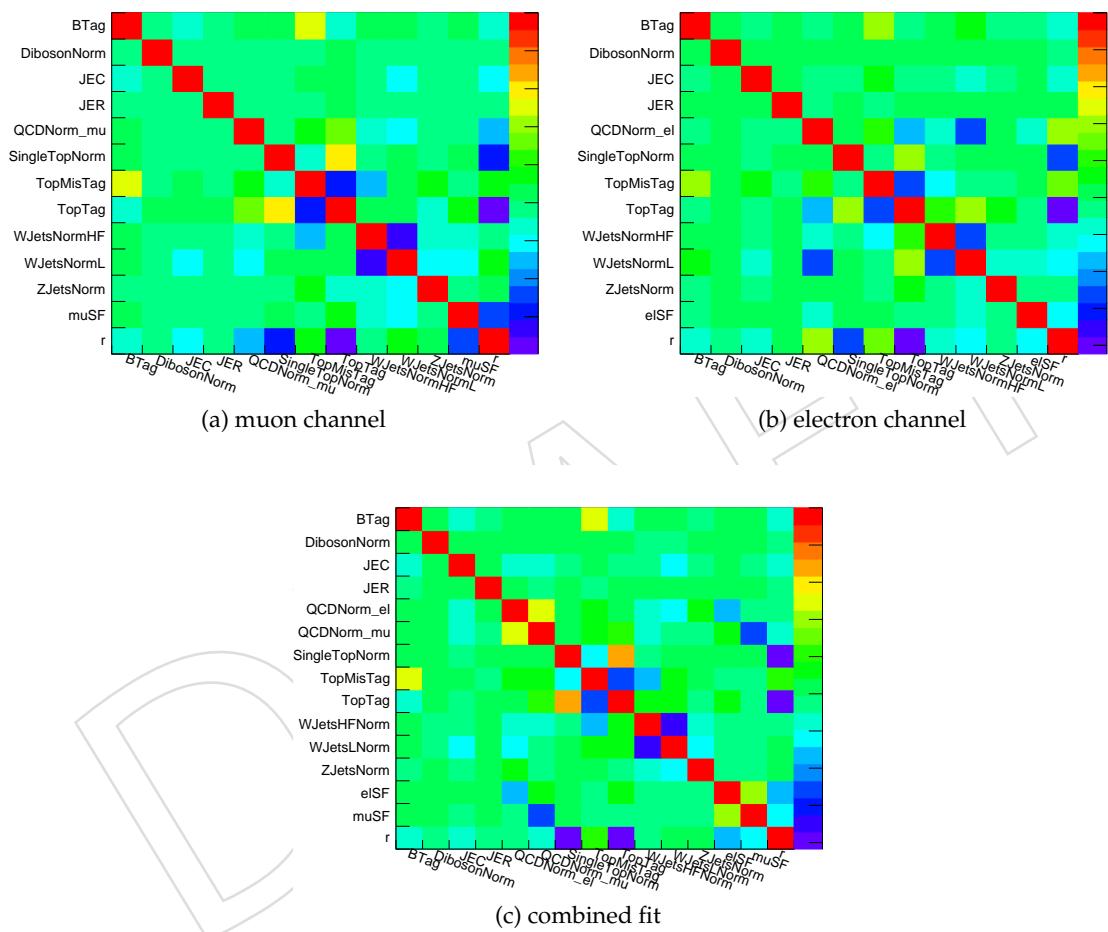


Figure 95: Correlations between nuisance parameters for the lepton η fit.

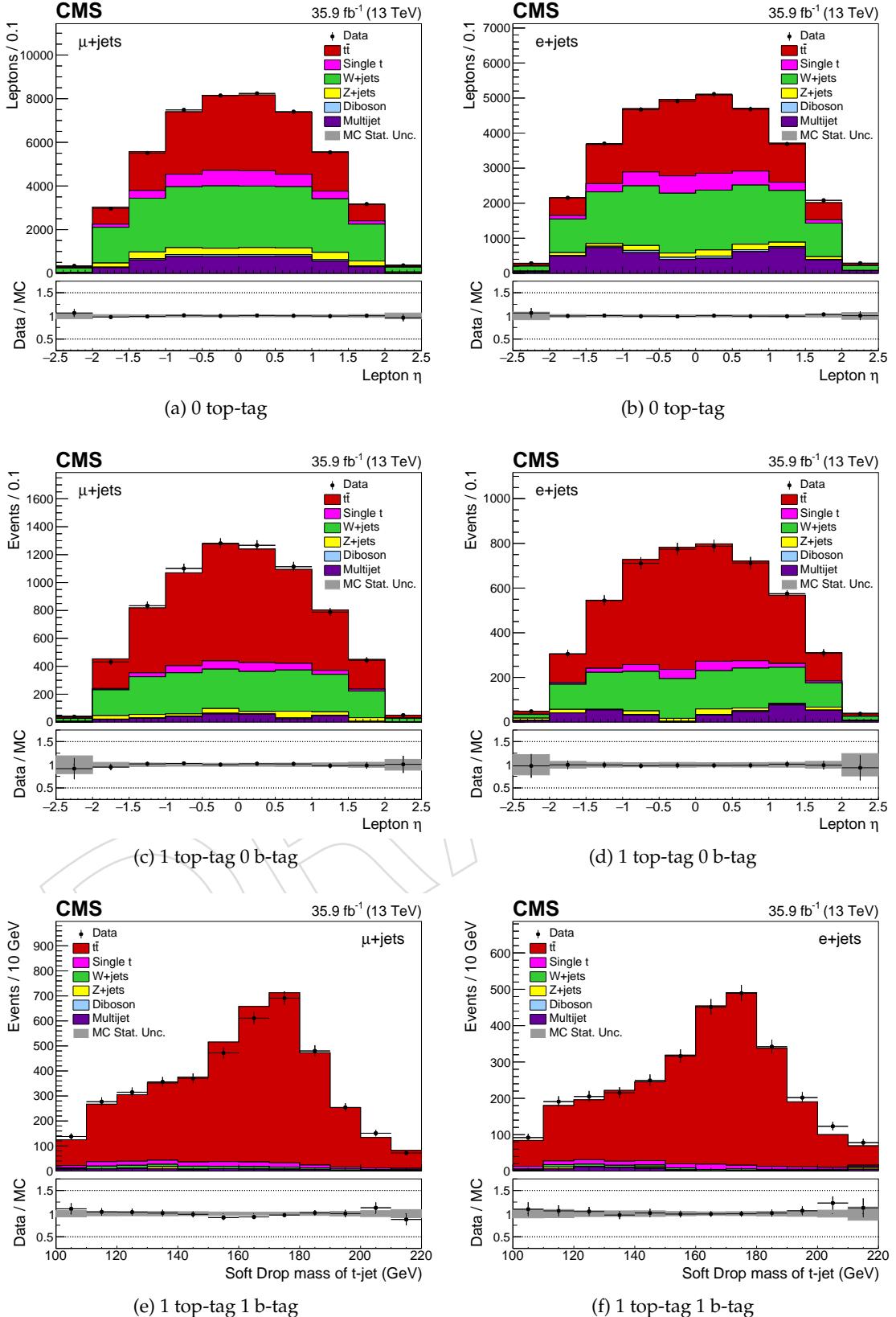


Figure 96: Post-fit kinematic distributions in the muon (left) and electron (right) channels for the lepton η fit combining both channels. For this fit, the QCD shape is taken from a data sideband.

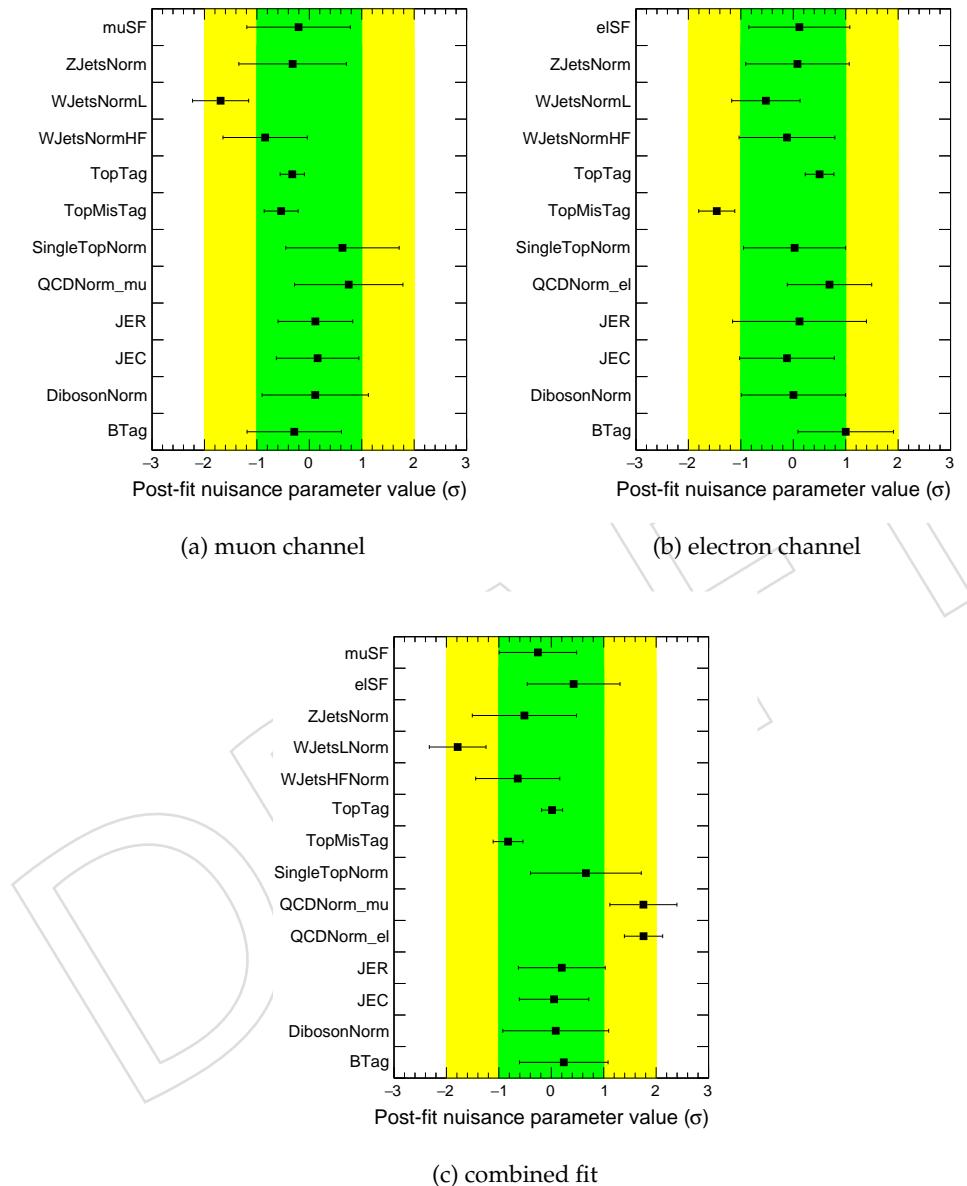


Figure 97: Post-fit nuisance parameters for the lepton $|\eta|$ fit. Nuisance parameters compatible with zero indicate no modification to the systematic by the fit; nuisance parameter uncertainties of 1σ indicate no constraint on the systematic uncertainty due to the fit.

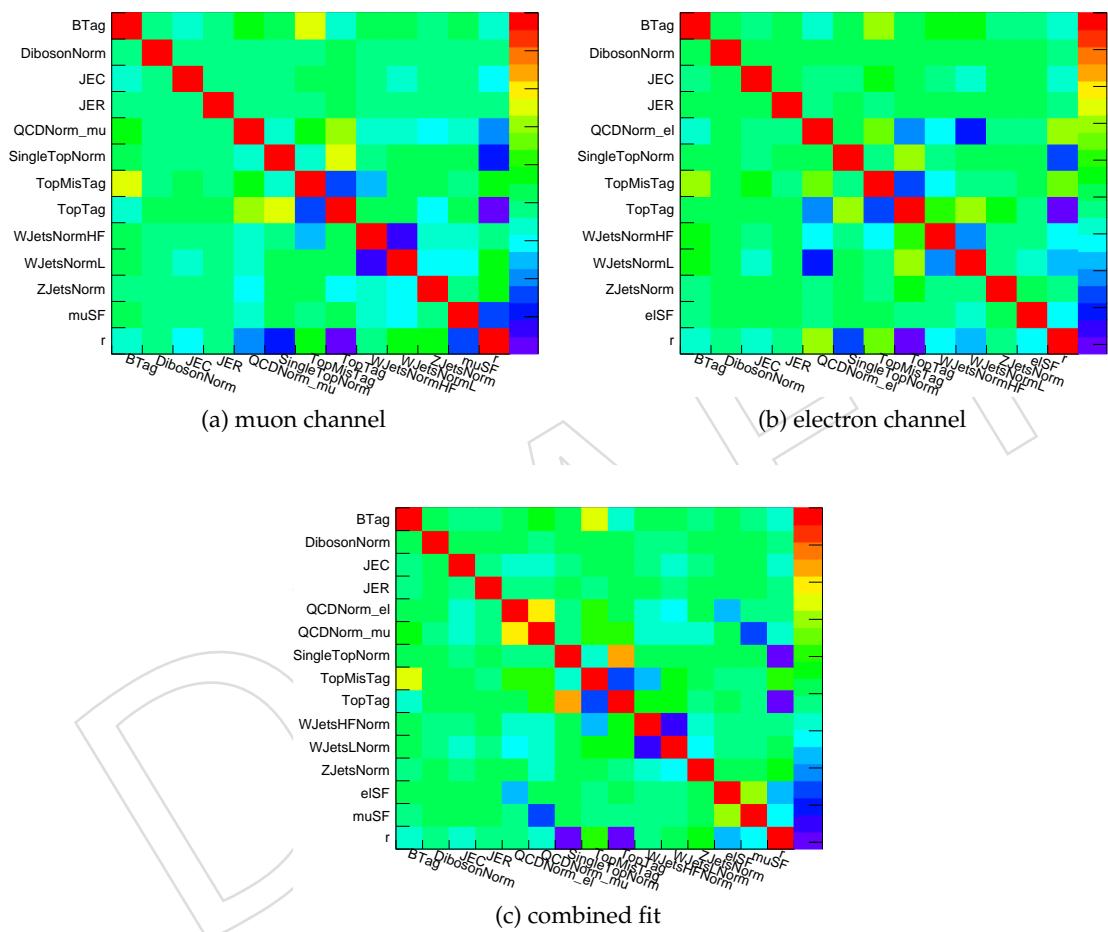


Figure 98: Correlations between nuisance parameters for the lepton $|\eta|$ fit.

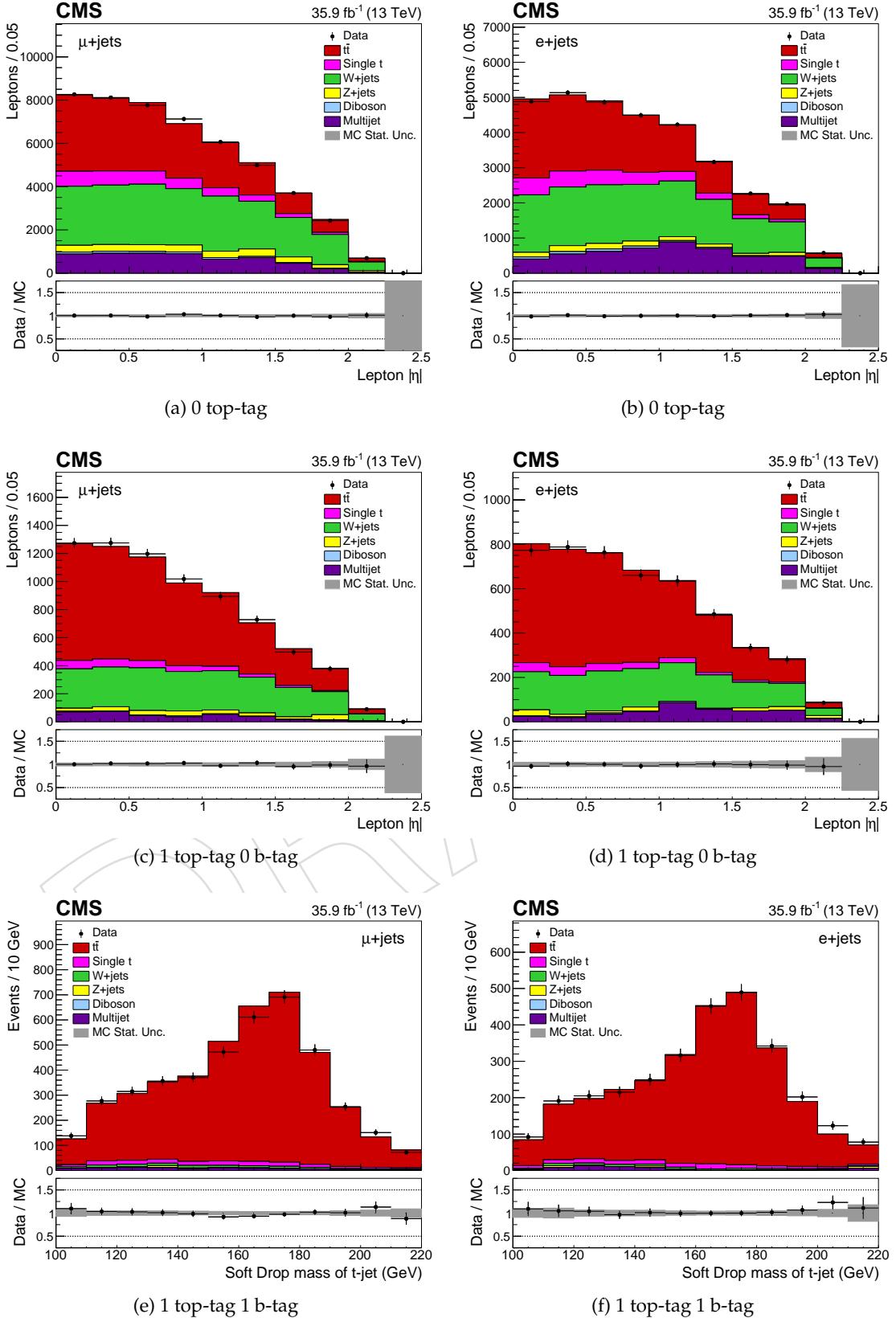


Figure 99: Post-fit kinematic distributions in the muon (left) and electron (right) channels for the lepton $|\eta|$ fit combining both channels. For this fit, the QCD shape is taken from a data sideband.

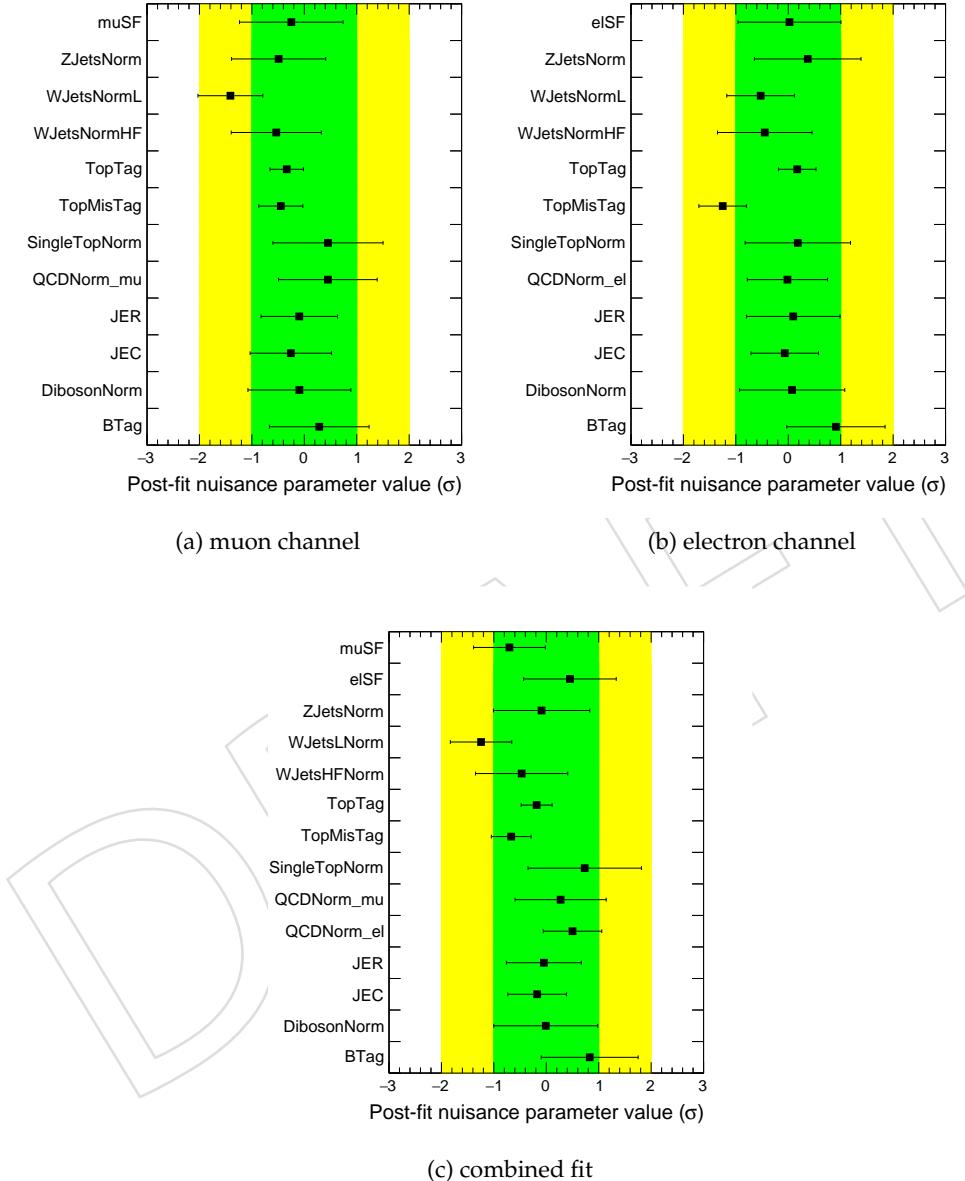


Figure 100: Post-fit nuisance parameters for the MC QCD AK4 η fit. Nuisance parameters compatible with zero indicate no modification to the systematic by the fit; nuisance parameter uncertainties of 1σ indicate no constraint on the systematic uncertainty due to the fit.

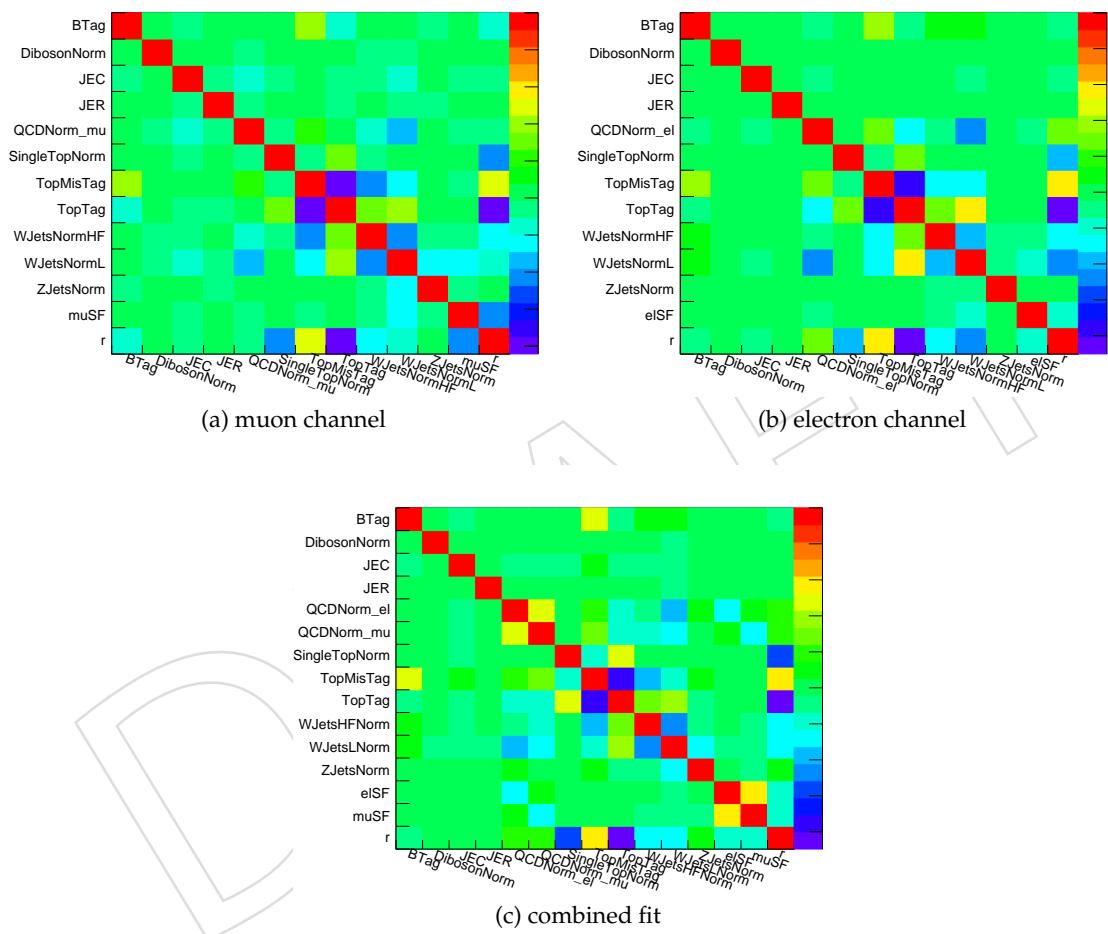


Figure 101: Correlations between nuisance parameters for the MC QCD AK4 η fit.

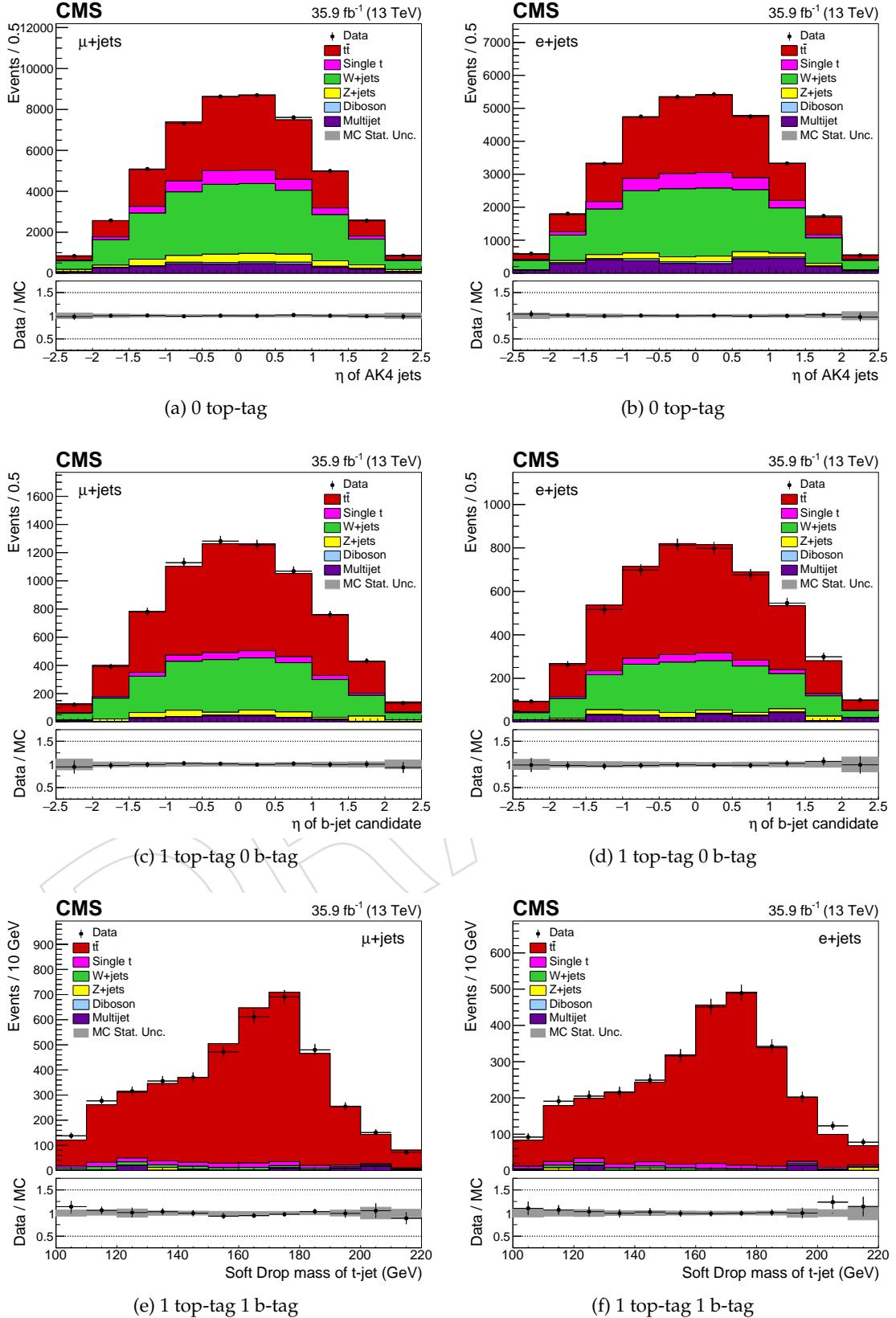


Figure 102: Post-fit kinematic distributions in the muon (left) and electron (right) channels for the MC QCD AK4 η fit combining both channels. For this fit, the QCD shape is taken from MC.

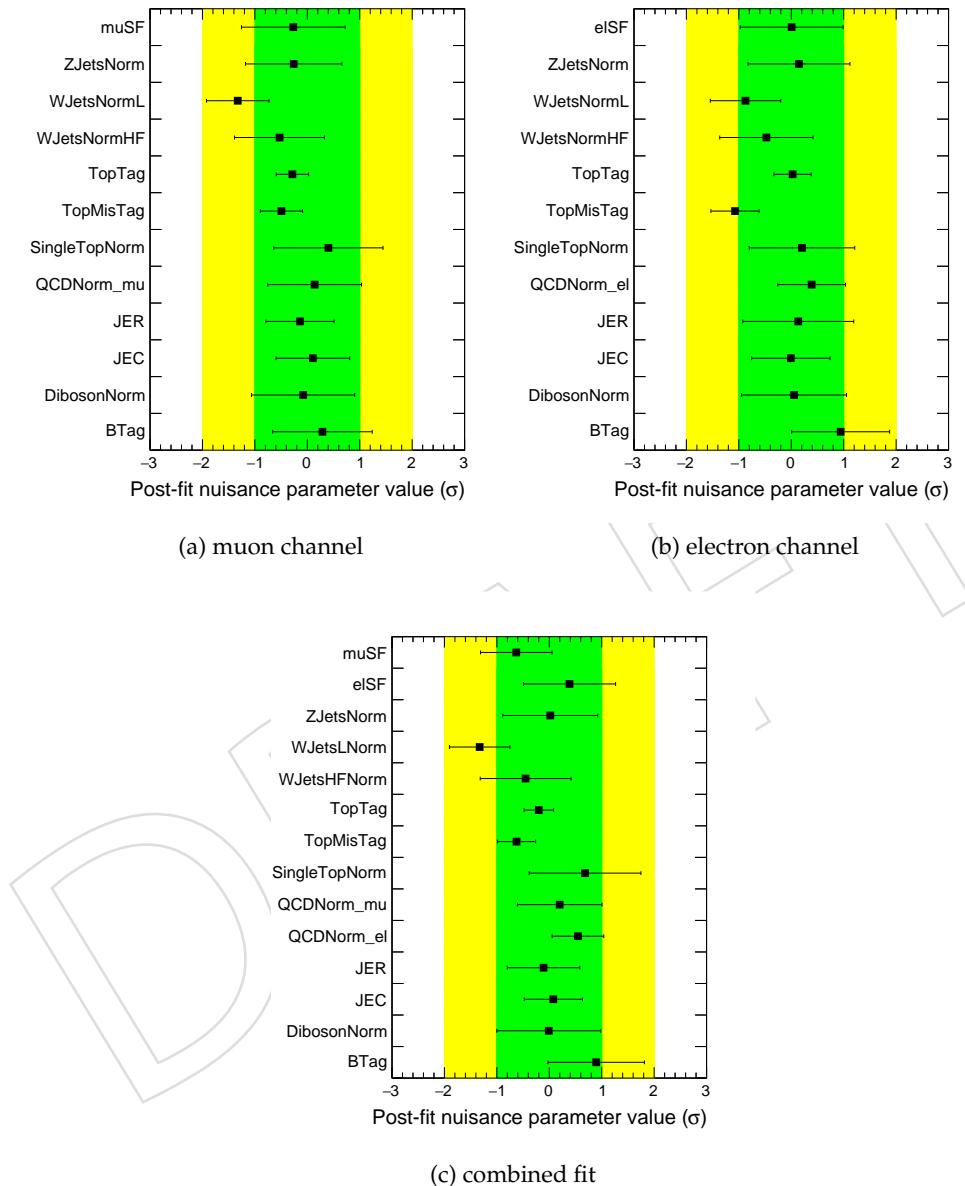


Figure 103: Post-fit nuisance parameters for the MC QCD AK4 $|\eta|$ fit. Nuisance parameters compatible with zero indicate no modification to the systematic by the fit; nuisance parameter uncertainties of 1σ indicate no constraint on the systematic uncertainty due to the fit.

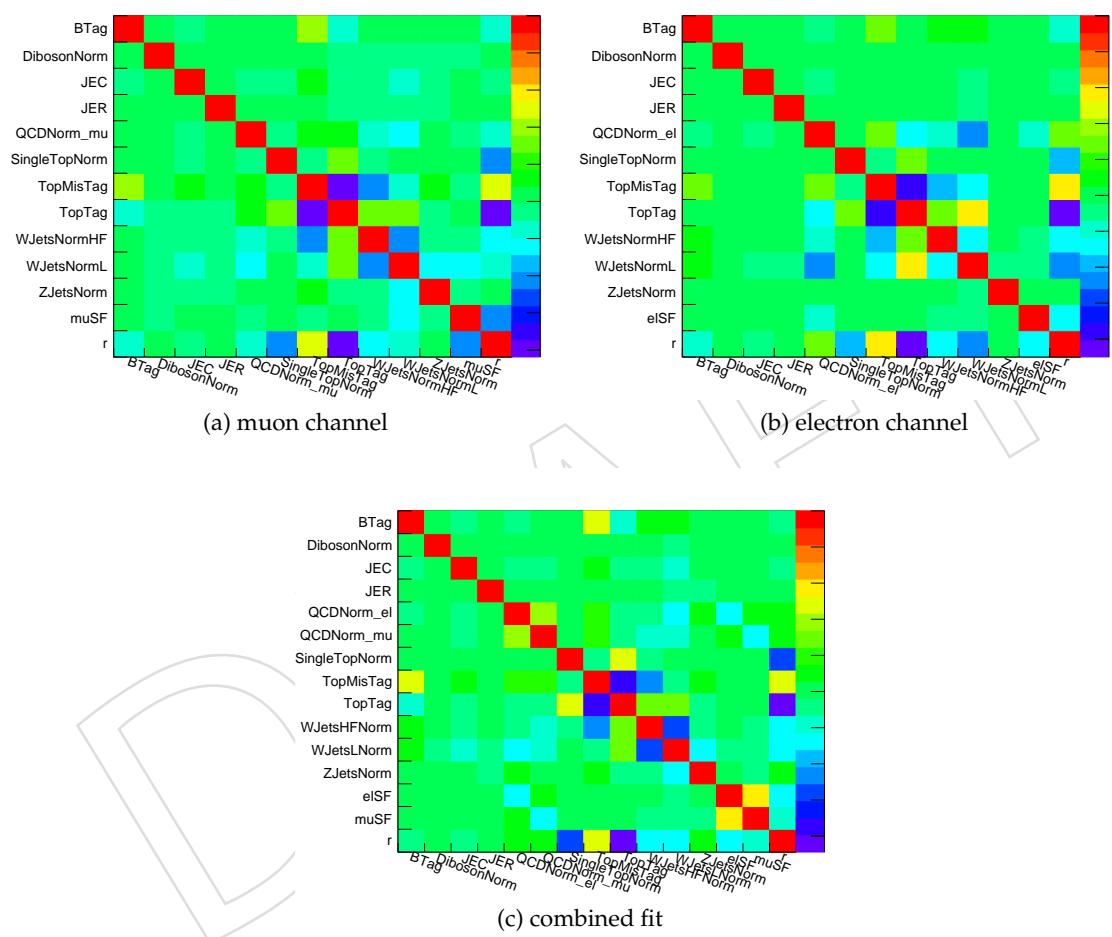


Figure 104: Correlations between nuisance parameters for the MC QCD AK4 $|\eta|$ fit.

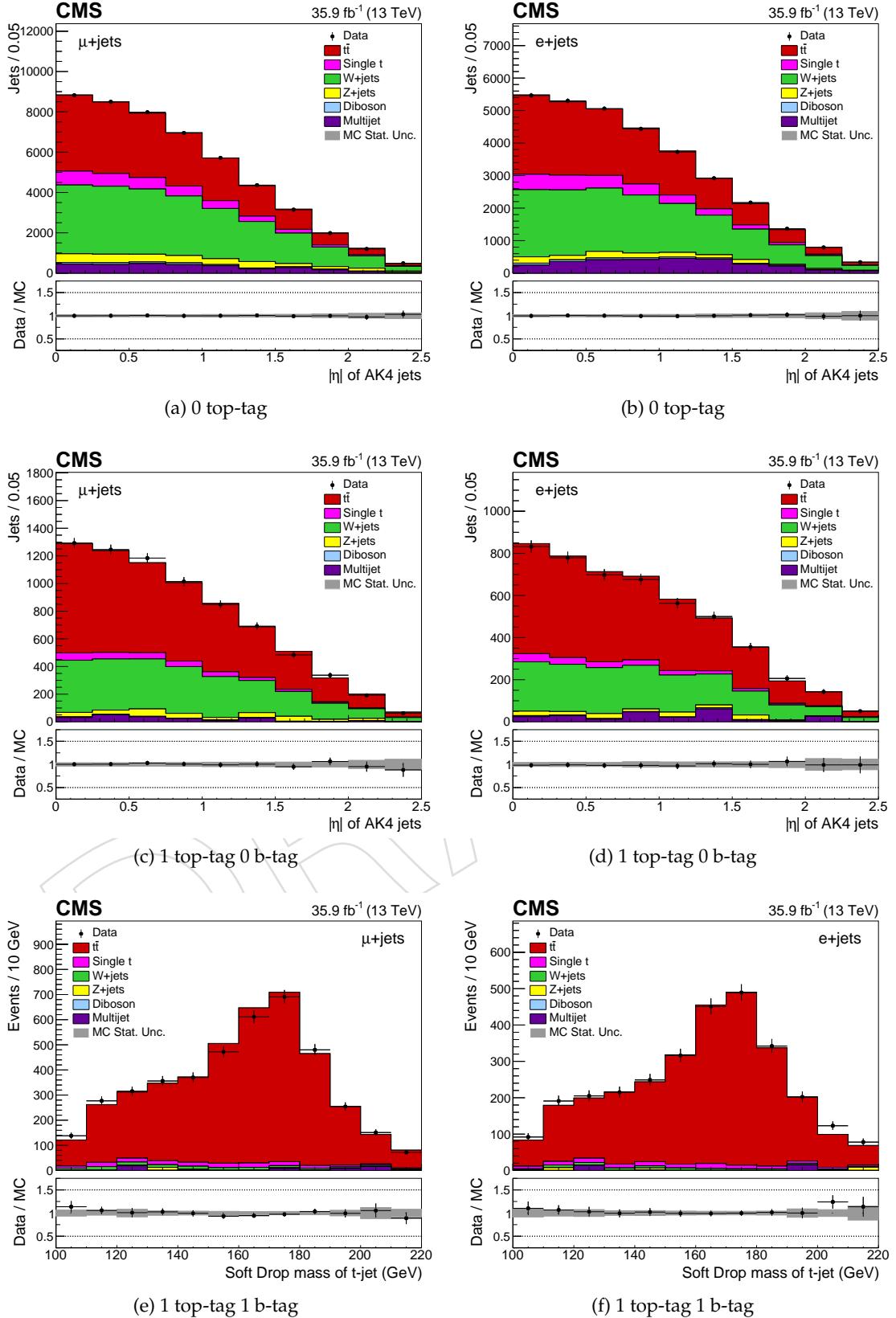


Figure 105: Post-fit kinematic distributions in the muon (left) and electron (right) channels for the MC QCD AK4 $|\eta|$ fit combining both channels. For this fit, the QCD shape is taken from MC.

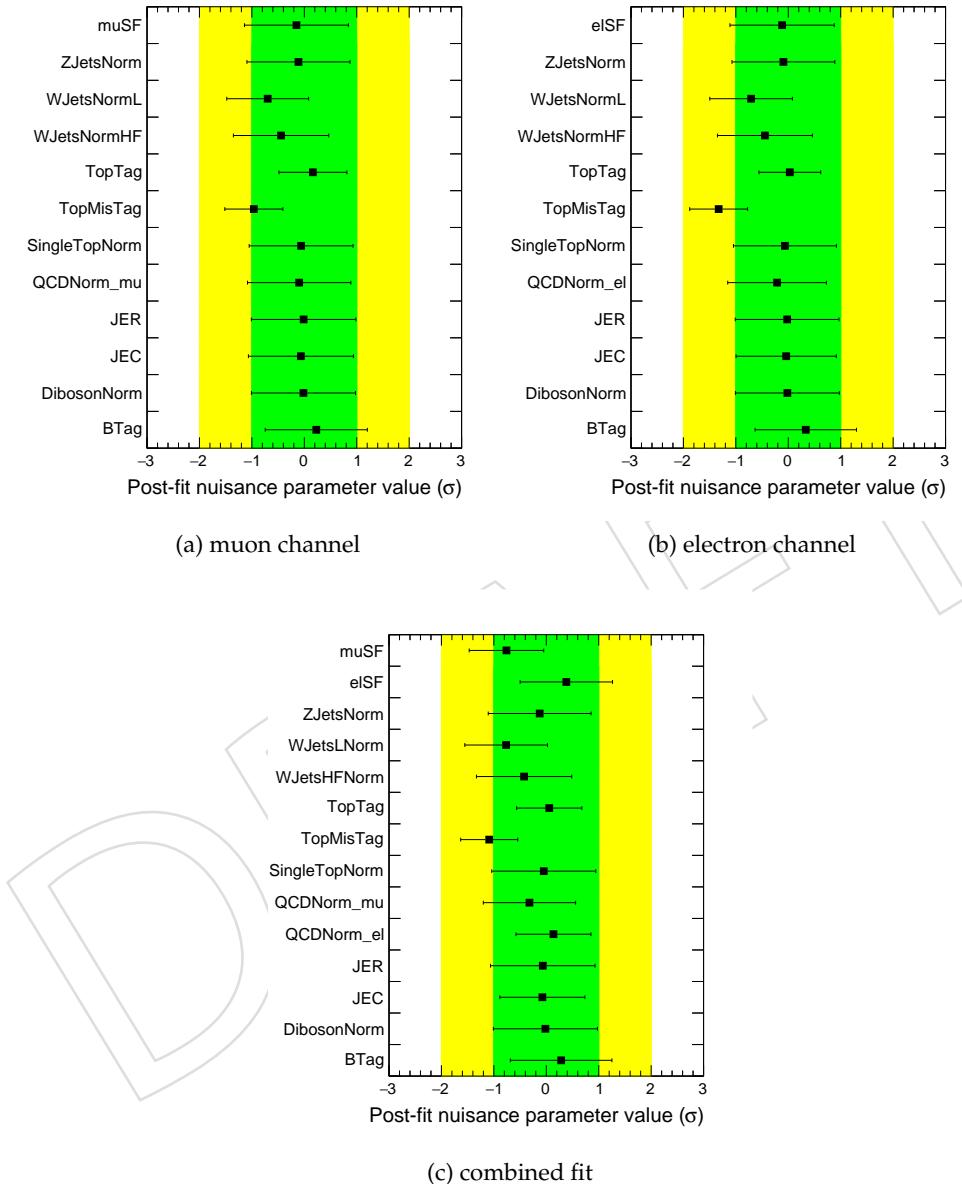


Figure 106: Post-fit nuisance parameters for the event count only fit. Nuisance parameters compatible with zero indicate no modification to the systematic by the fit; nuisance parameter uncertainties of 1σ indicate no constraint on the systematic uncertainty due to the fit.

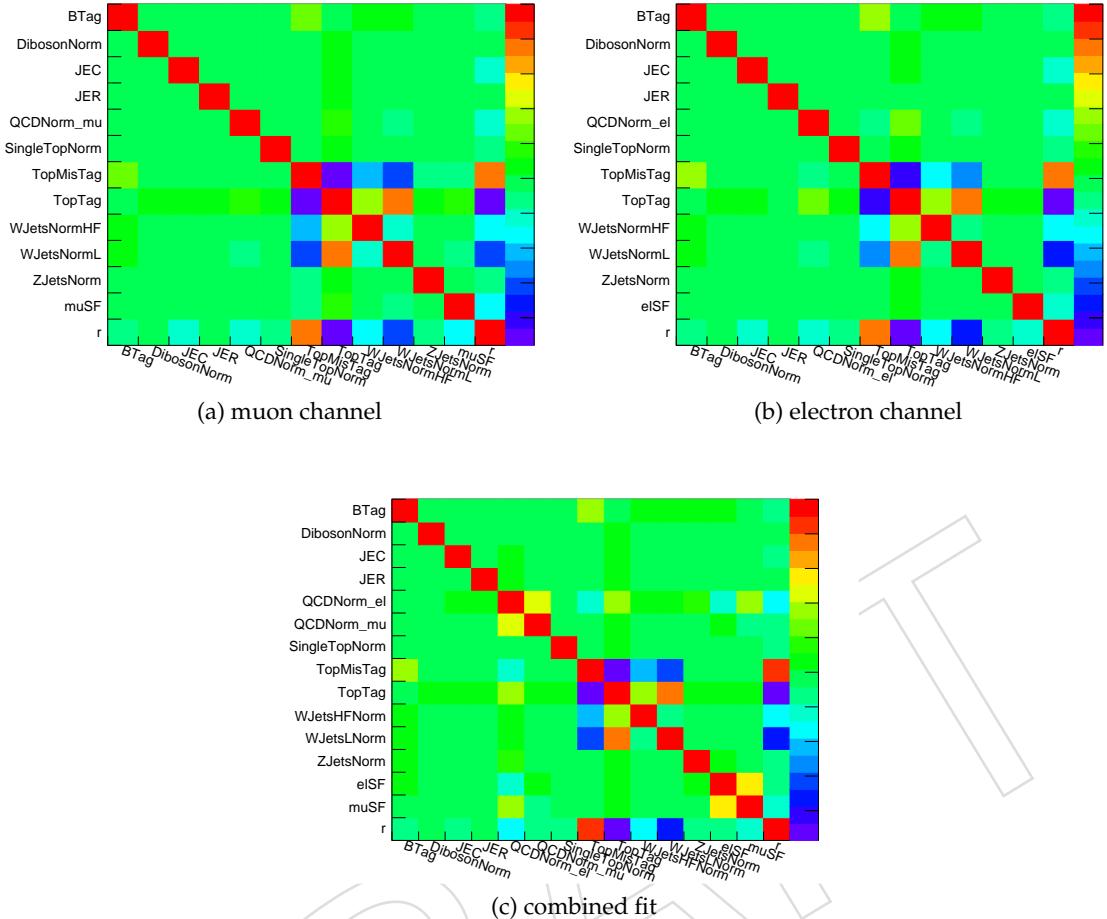


Figure 107: Correlations between nuisance parameters for the event count only fit.

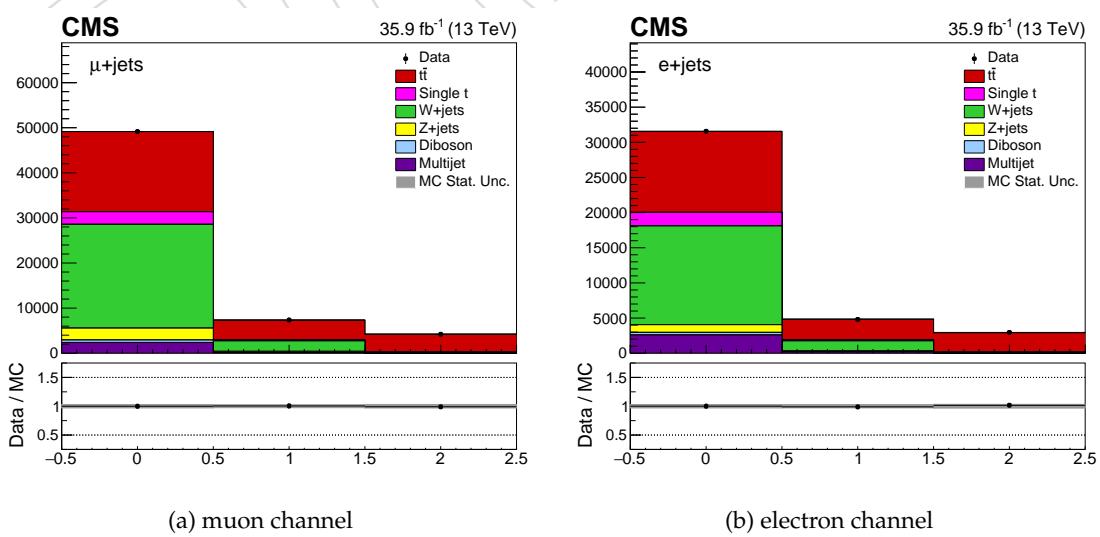


Figure 108: Fitted event counts in the 0 top-tag, 1 top-tag 0 b-tag, and 1 top-tag 1 b-tag regions for the event count only fit.

Table 24 gives a summary of the signal strengths for the nominal fit and fit variants in the muon, electron, and combined channels. While the signal strength is not used directly to construct the background-subtracted data distribution, it is implicitly related to the normalization. The signal strength thus serves as a good overall indicator of the fit result, which can be used to compare between fits. The AK4 η and $|\eta|$ fits have very similar results, as do the lepton η and $|\eta|$ fits and the MC QCD AK4 η and $|\eta|$ fits. The lepton and AK4 fits agree better than the AK4 and MC QCD AK4 fits, though all fits agree within uncertainties. All the template fits are well within the uncertainty bounds of the event count only fit. The largest divergences from the nominal among the template fits are the lepton η fit in the muon channel (0.06) and the MC QCD AK4 $|\eta|$ fit in the electron and combined channels (0.09 and 0.07, respectively).

Fit	muon channel	electron channel	combined channel
AK4 η (nominal)	0.85 ± 0.07	0.77 ± 0.07	0.81 ± 0.05
AK4 $ \eta $	0.86 ± 0.06	0.75 ± 0.07	0.80 ± 0.05
Lepton η	0.91 ± 0.07	0.77 ± 0.05	0.85 ± 0.05
Lepton $ \eta $	0.90 ± 0.07	0.77 ± 0.06	0.84 ± 0.05
MC QCD AK4 η	0.89 ± 0.07	0.84 ± 0.07	0.88 ± 0.06
MC QCD AK4 $ \eta $	0.88 ± 0.07	0.86 ± 0.08	0.88 ± 0.06
Event counts	0.80 ± 0.13	0.89 ± 0.13	0.85 ± 0.13

Table 24: Posterior signal strengths in the muon, electron, and combined channels for each fit variant. The nominal fit is shown as well for reference.

Table 25 gives a summary of the posterior top(mis)tag nuisance parameters for the nominal fit and fit variants. The values shown are taken from the combined fit. The AK4 η and $|\eta|$ fits have very similar results, as do the lepton η and $|\eta|$ fits and the MC QCD AK4 η and $|\eta|$ fits. There is some variation between the AK4, lepton, and MC QCD AK4 fits, though the fits all agree within uncertainties and all fall well within the uncertainty bounds of the event count only fit.

For the top-tag SF and top-mis-tag SF, the largest variation with respect to the nominal is the MC QCD AK4 $|\eta|$ fit, yielding systematic variations of 0.38 and 0.43, respectively. The top tag SF fit variation is not assigned as a systematic uncertainty; rather the fit variants are simply used as cross checks. Instead, a systematic uncertainty on the top tag SF arising from the top jet p_T and $|\eta|$ dependence is assessed by performing the fit divided in bins of top jet p_T and $|\eta|$. This is described in more detail in Appendix G.

Fit	Top-tag SF	Top-mis-tag SF
AK4 η (nominal)	0.18 ± 0.25	-1.06 ± 0.34
AK4 $ \eta $	0.18 ± 0.25	-1.02 ± 0.34
Lepton η	0.02 ± 0.21	-0.95 ± 0.28
Lepton $ \eta $	0.02 ± 0.20	-0.82 ± 0.28
MC QCD AK4 η	-0.18 ± 0.30	-0.67 ± 0.38
MC QCD AK4 $ \eta $	-0.20 ± 0.28	-0.63 ± 0.36
Event counts	0.06 ± 0.62	-1.09 ± 0.54

Table 25: Posterior top(mis)tag SF nuisance parameters for each fit variant in the combined channel. The nominal fit is shown as well for reference. While the results in the muon and electron channels are not shown, they agree with the combined result for all fits.

Table 26 gives a summary of the posterior background normalization nuisance parameters for the nominal fit and fit variants. The values shown are taken from the combined fit. The posterior background normalization nuisance parameters show noticeable variation, though for

1048 the most part the results agree within uncertainties. The single top normalization is generally
 1049 scaled up slightly but not constrained by the fit. The variation between fits is smaller than the
 1050 fit uncertainty. The W+light normalization is scaled down and somewhat constrained by the
 1051 fits. The lepton and AK4 fits show the most variation but still agree within uncertainties. The
 1052 W+heavy flavor normalization is generally scaled down slightly with little constraint, though
 1053 it is scaled up slightly in the MC QCD AK4 $|\eta|$ fit. The latter fit shows the most tension with
 1054 respect to the others. The Z+jets normalization is scaled down slightly and not constrained
 1055 by the fits, with slight variations depending on the fit template. The diboson normalization is
 1056 generally not affected by the fits. The muon channel QCD normalization is only slightly scaled
 1057 up for the AK4 and MC QCD AK4 fits, but is scaled up significantly for the lepton fits. The nor-
 1058 malization is more constrained for the latter fits as well. A similar trend occurs for the electron
 1059 channel QCD normalization, though the normalization is generally scaled up and constrained
 1060 more strongly. Not all fit results agree with the result of the event count only fit.

1061 The largest variation between the the nominal fit and the other template fits is assigned as a
 1062 systematic uncertainty on each posterior background normalization nuisance parameter. The
 1063 largest varition from nominal in the single top and Z+jets normalizations is the lepton η , with
 1064 systematics of 0.32 and 0.36, respectively. The largest variation from nominal in the W+light,
 1065 W+heavy flavor, diboson, muon QCD, and electron QCD normalizations is the lepton $|\eta|$, with
 1066 systematics of 0.77, 0.58, 0.05, 1.70, and 1.08, respectively.

Fit	Single Top	W+light	W+heavy	Z+jets	Diboson	QCD (mu)	QCD (el)
AK4 η (nominal)	0.55 ± 1.05	-1.02 ± 0.56	-0.06 ± 0.89	-0.22 ± 0.98	0.03 ± 1.00	0.06 ± 0.93	0.68 ± 0.59
AK4 $ \eta $	0.53 ± 1.04	-1.06 ± 0.57	0.01 ± 0.92	-0.53 ± 0.95	0.00 ± 0.99	0.22 ± 0.99	0.84 ± 0.62
Lepton η	0.87 ± 1.10	-1.65 ± 0.56	-0.49 ± 0.82	-0.58 ± 0.95	0.00 ± 1.00	1.37 ± 0.68	1.53 ± 0.38
Lepton $ \eta $	0.66 ± 1.06	-1.79 ± 0.54	-0.64 ± 0.80	-0.51 ± 0.99	0.08 ± 1.01	1.76 ± 0.64	1.76 ± 0.37
MC QCD AK4 η	0.73 ± 1.08	-1.24 ± 0.59	-0.47 ± 0.88	-0.09 ± 0.92	-0.01 ± 0.99	0.27 ± 0.87	0.50 ± 0.56
MC QCD AK4 $ \eta $	0.68 ± 1.07	-1.33 ± 0.58	0.45 ± 0.87	0.02 ± 0.91	-0.01 ± 0.99	0.20 ± 0.81	0.55 ± 0.49
Event counts	-0.05 ± 0.99	-0.77 ± 0.79	-0.42 ± 0.91	-0.13 ± 0.98	-0.02 ± 0.99	-0.32 ± 0.88	0.14 ± 0.72

Table 26: Posterior background normalization nuisance parameters for each fit variant in the combined channel. The nominal fit is shown as well for reference.

1067 G Top tag SF p_T and $|\eta|$ dependence

1068 To assess the dependence of the top-tag SF on the top jet candidate kinematics, the top-tag SF
 1069 was extracted in two bins of top jet candidate p_T and $|\eta|$. The difference between these binned
 1070 top-tag SFs and the flat top-tag SF was assessed as a shape uncertainty on the top-tag and top-
 1071 mis-tag SFs used in the unfolding. To extract the top-tag SF binned in top jet p_T ($|\eta|$), the fit was
 1072 performed with the 0t, 1t0b, and 1t1b kinematic regions further subdivided into regions of low
 1073 and high top jet p_T ($|\eta|$). All nuisances aside from the top-tag and top-mis-tag SFs were treated
 1074 as fully correlated between the low and high top jet p_T ($|\eta|$) regions. The fit templates used
 1075 match those of the nominal fit. As in the nominal case, the fit variations were first performed
 1076 separately in the muon and electron channels, then simultaneously combining both channels.

1077 G.1 Top-tag SF binned in p_T

1078 To assess the top-tag SF dependence on top jet p_T , the top-tag SF was extracted separately for
 1079 top jet candidates with p_T greater than or less than 500 GeV. The fit was performed simultane-
 1080ously in six kinematic regions – the 0t, 1t0b, and 1t1b regions of the original fit, further divided
 1081 into regions of low and high top jet p_T . In the a priori kinematic distributions, a MC excess
 1082 was observed in the 0t, top jet $p_T > 500$ GeV region. As this region is dominated by W+jets
 1083 background, the excess was believed to be due to mismodelling of the W+jets contribution at
 1084 high jet p_T . To compensate for this, an additional nuisance parameter was introduced to the fit
 1085 to renormalize the W+jets contribution in the top jet $p_T > 500$ region.

1086 The posterior nuisance parameters and their correlations are shown in Figures 109 and 110, re-
 1087 spectively. The posterior nuisances for the top jet p_T binned fit generally agree with those from
 1088 the nominal fit, shown in Section 8.1. The correlations between nuisance parameters match
 1089 those of the nominal fit, with the top-tag SFs for low and high top jet p_T being strongly corre-
 1090 lated. The low posterior WJetsPt nuisance parameter indicates that the fit is suppressing the
 1091 W+jets normalization in the top jet $p_T > 500$ GeV region. This parameter is anticorrelated with
 1092 the JEC, as expected since both affect the top jet p_T distribution for W+jets. The posterior signal
 1093 strength is 0.84 ± 0.05 , agreeing well with the nominal value of 0.81 ± 0.05 . The posterior kine-
 1094 matic distributions are shown in Figures 111 and 112 for the low and high top jet p_T regions,
 1095 respectively.

1096 G.2 Top-tag SF binned in $|\eta|$

1097 To assess the top-tag SF dependence on top jet $|\eta|$, the top-tag SF was extracted separately for
 1098 top jet candidates with $|\eta|$ greater than or less than 1.0. The fit was performed simultaneoulsy
 1099 in six kinematic regions – the 0t, 1t0b, and 1t1b regions of the original fit, further divided into
 1100 regions with low and high top jet $|\eta|$.

1101 The posterior nuisance parameters and their correlations are shown in Figures 113 and 114, re-
 1102 spectively. The posterior nuisances for the top jet $|\eta|$ binned fit agree well with those from the
 1103 nominal fit, shown in Section 8.1. The correlations between nuisance parameters also match
 1104 those of the nominal fit, with the top-tag SFs for low and high top jet $|\eta|$ being strongly corre-
 1105 lated. The posterior signal strength is 0.82 ± 0.05 , agreeing well with the nominal value of
 1106 0.81 ± 0.05 . The posterior kinematic distributions are shown in Figures 115 and 116 for the low
 1107 and high top jet $|\eta|$ regions, respectively.

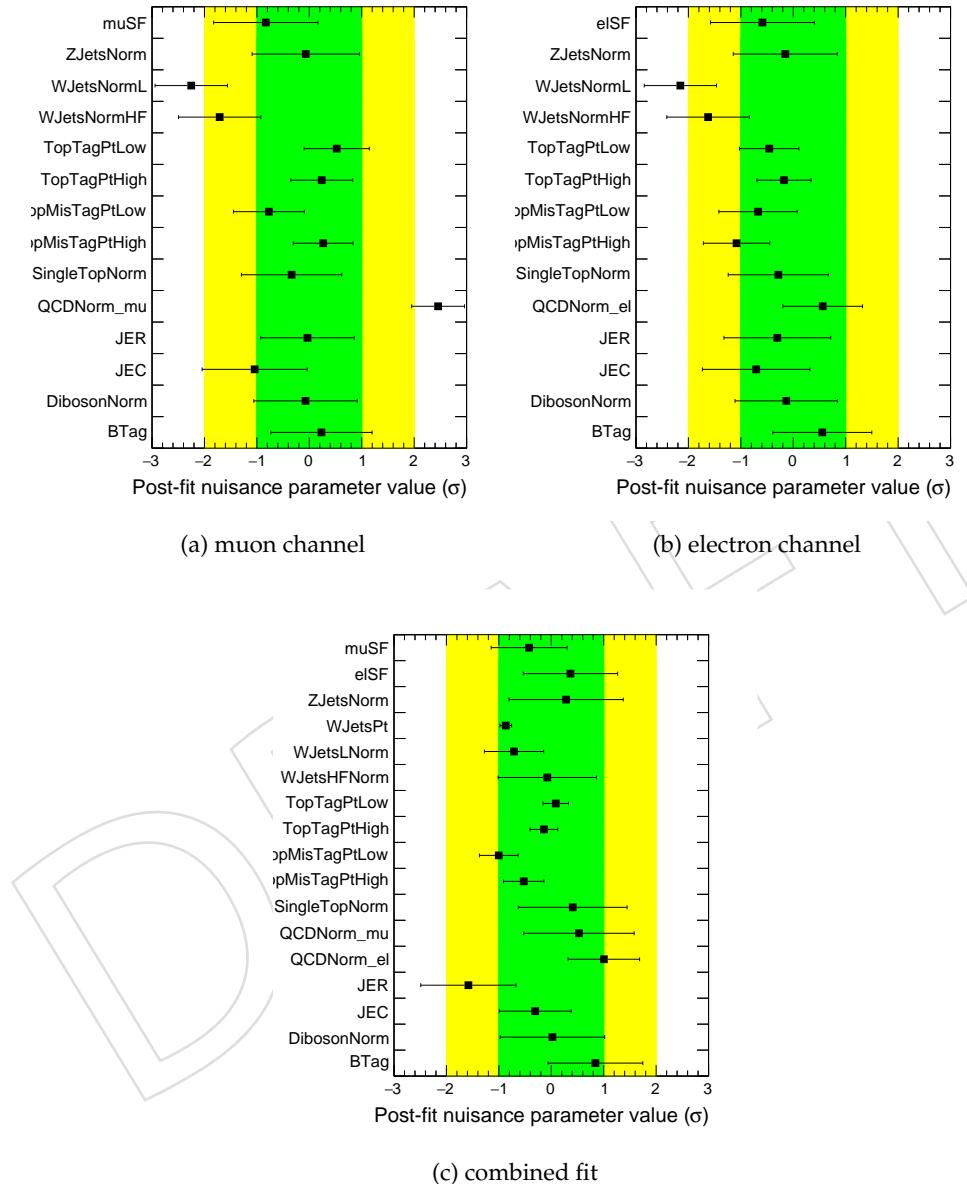


Figure 109: Post-fit nuisance parameters for the top jet p_T binned fit. Nuisance parameters compatible with zero indicate no modification to the systematic by the fit; nuisance parameter uncertainties of 1σ indicate no constraint on the systematic uncertainty due to the fit.

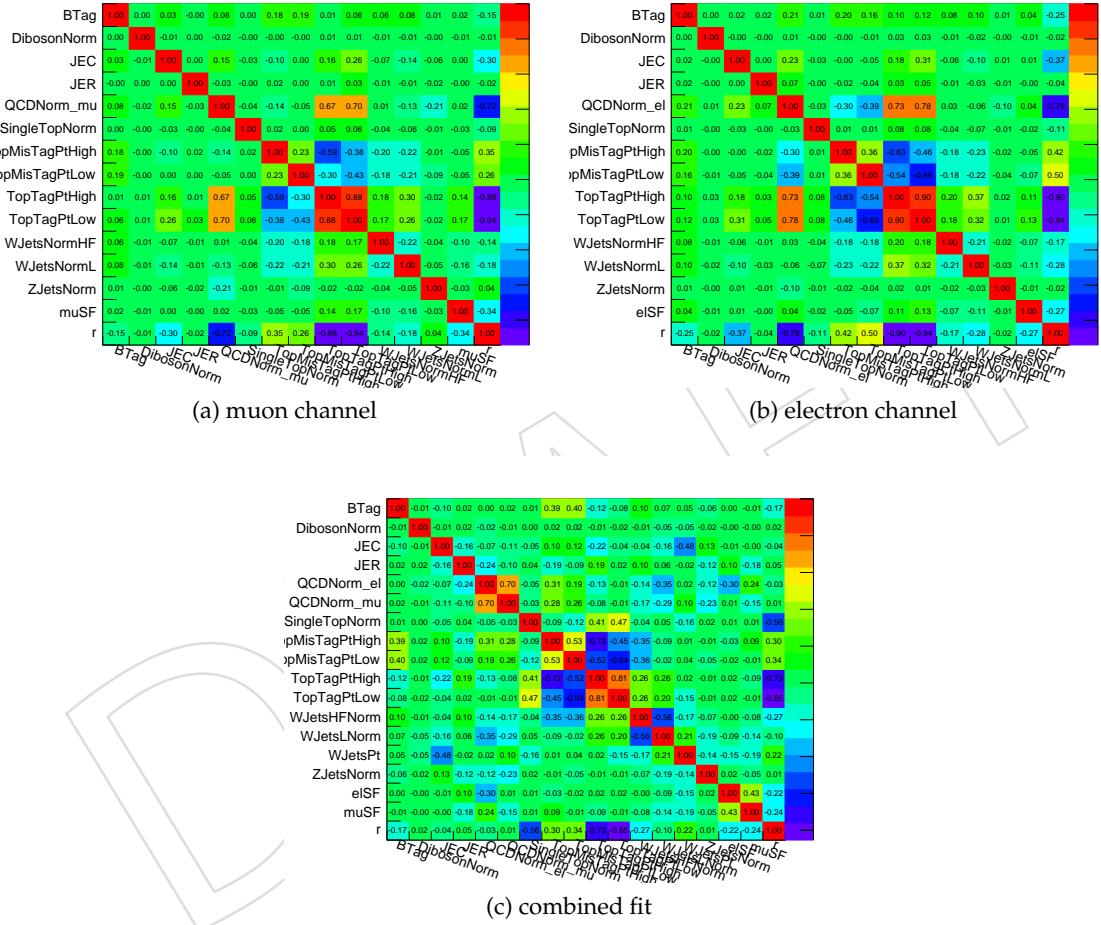


Figure 110: Correlations between nuisance parameters for the top jet p_T binned fit.

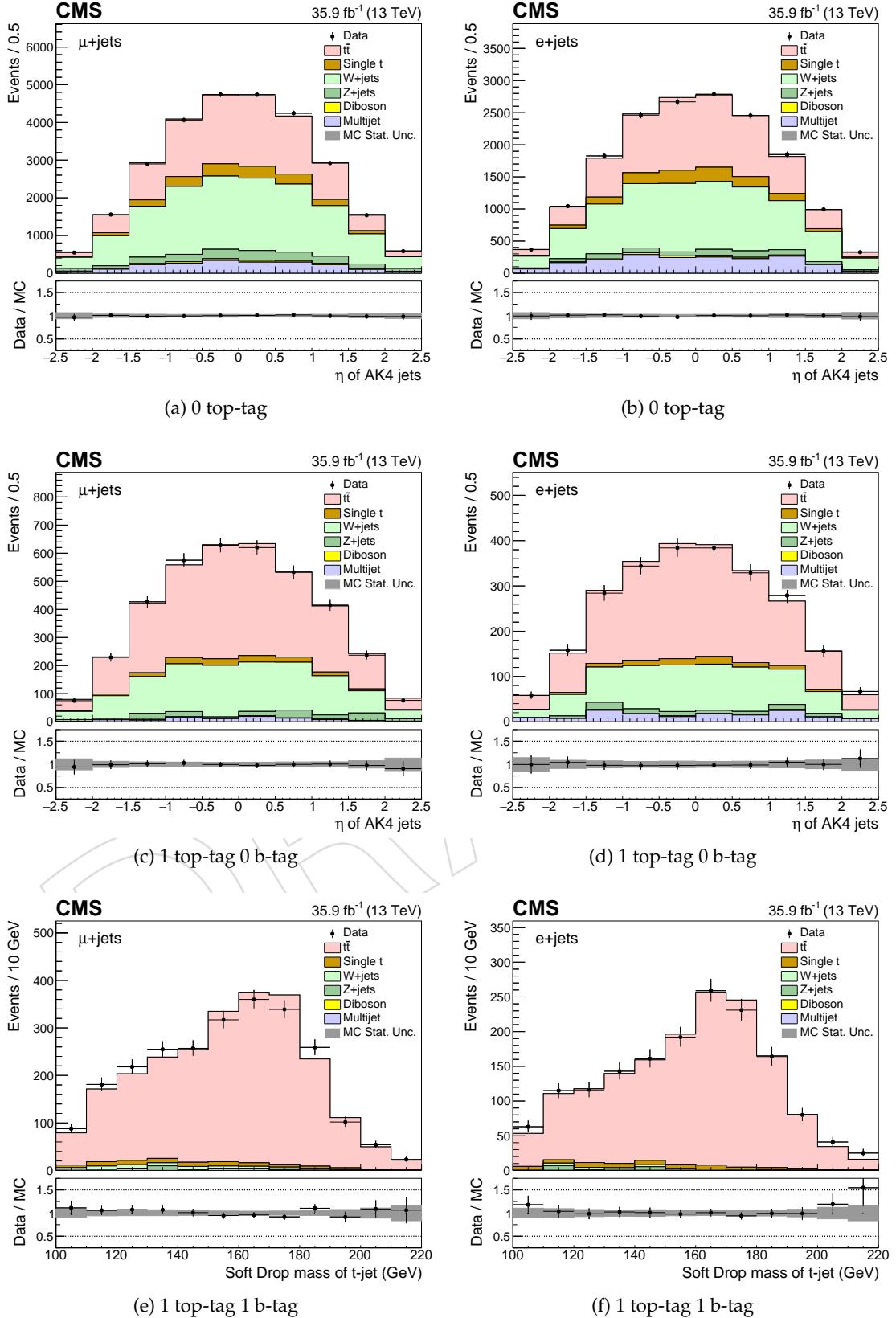


Figure 111: Post-fit kinematic distributions for top jet $p_T < 500 \text{ GeV}$ in the muon (left) and electron (right) channels for the top jet p_T binned fit combining both channels. For this fit, the QCD shape is taken from a data sideband.

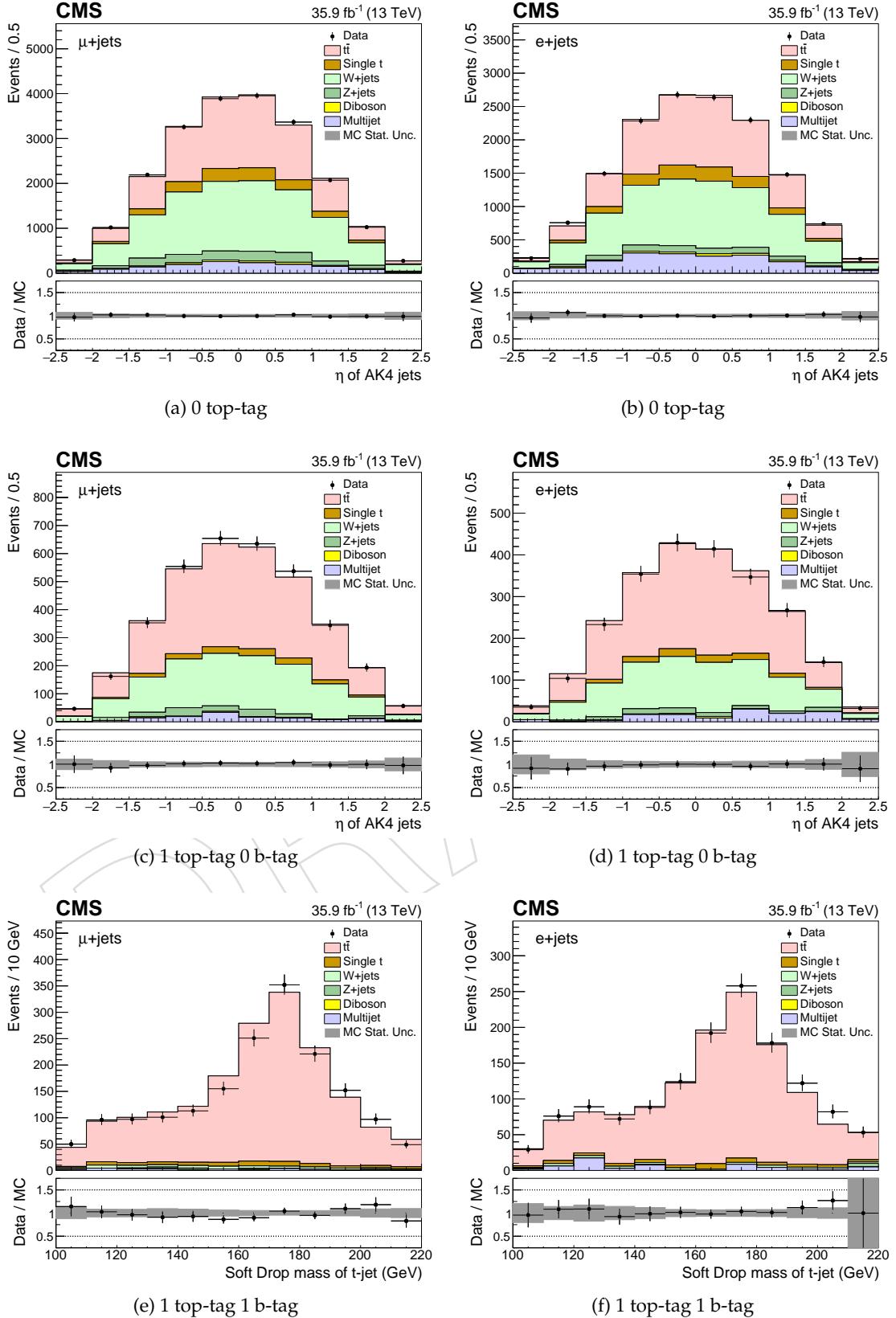


Figure 112: Post-fit kinematic distributions for top jet $p_T > 500$ GeV in the muon (left) and electron (right) channels for the top jet p_T binned fit combining both channels. For this fit, the QCD shape is taken from a data sideband.

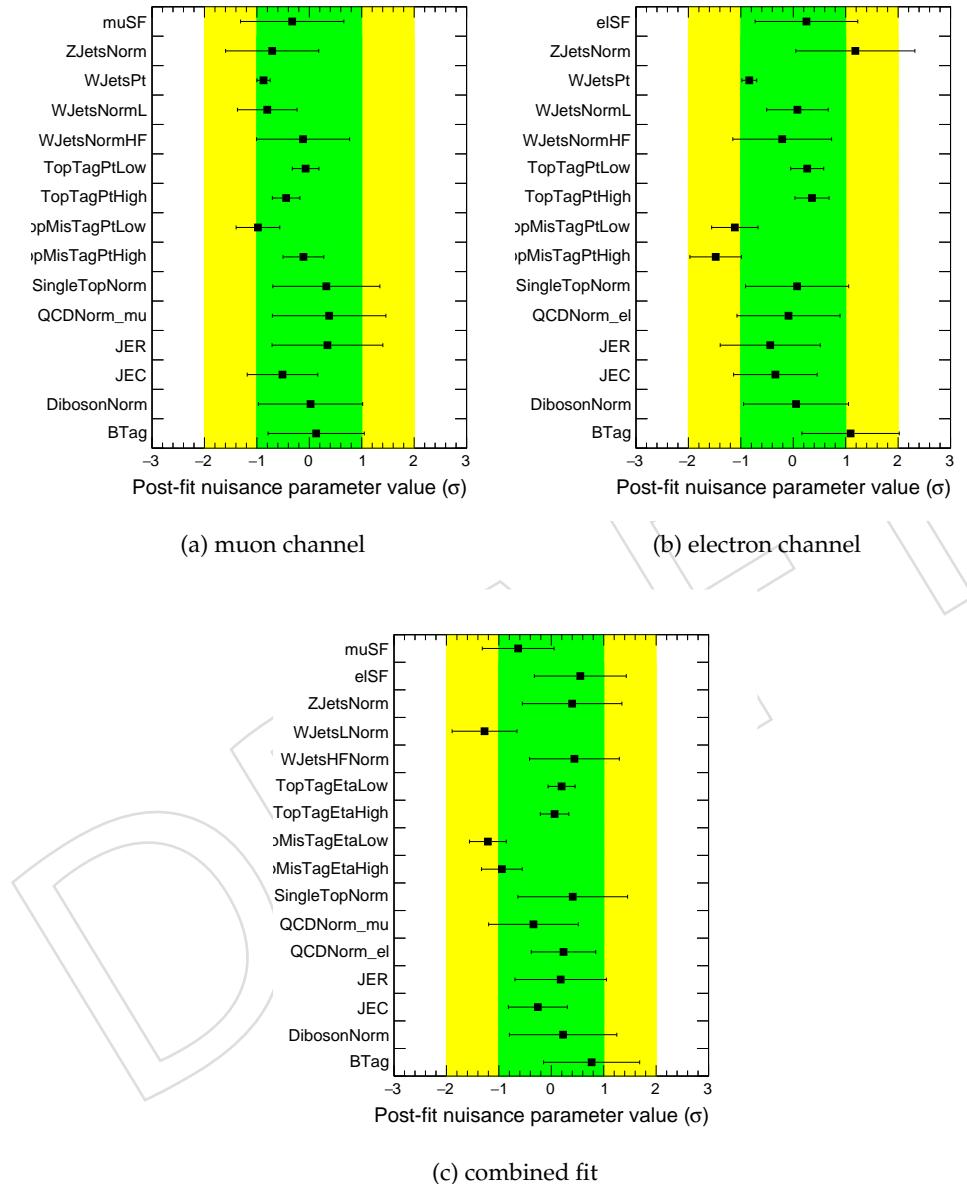


Figure 113: Post-fit nuisance parameters for the top jet $|\eta|$ binned fit. Nuisance parameters compatible with zero indicate no modification to the systematic by the fit; nuisance parameter uncertainties of 1σ indicate no constraint on the systematic uncertainty due to the fit.

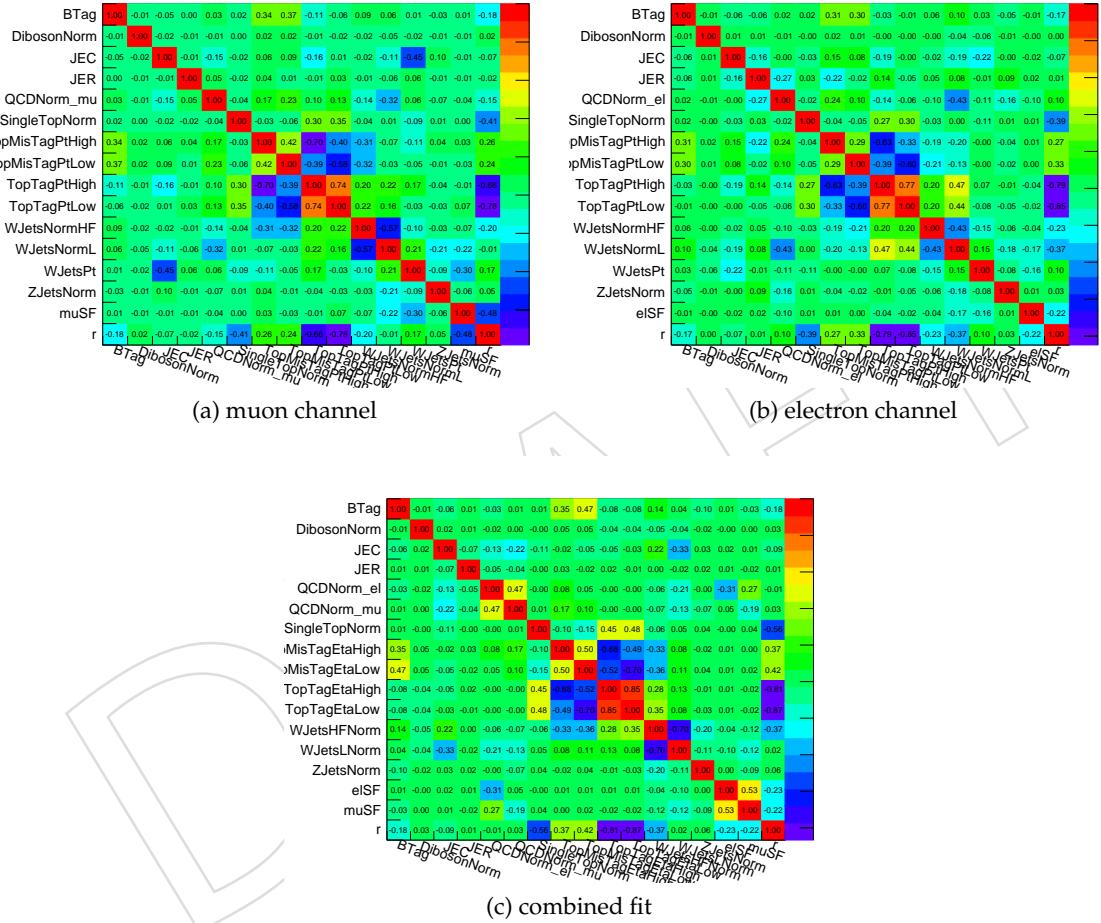


Figure 114: Correlations between nuisance parameters for the top jet $|\eta|$ binned fit.

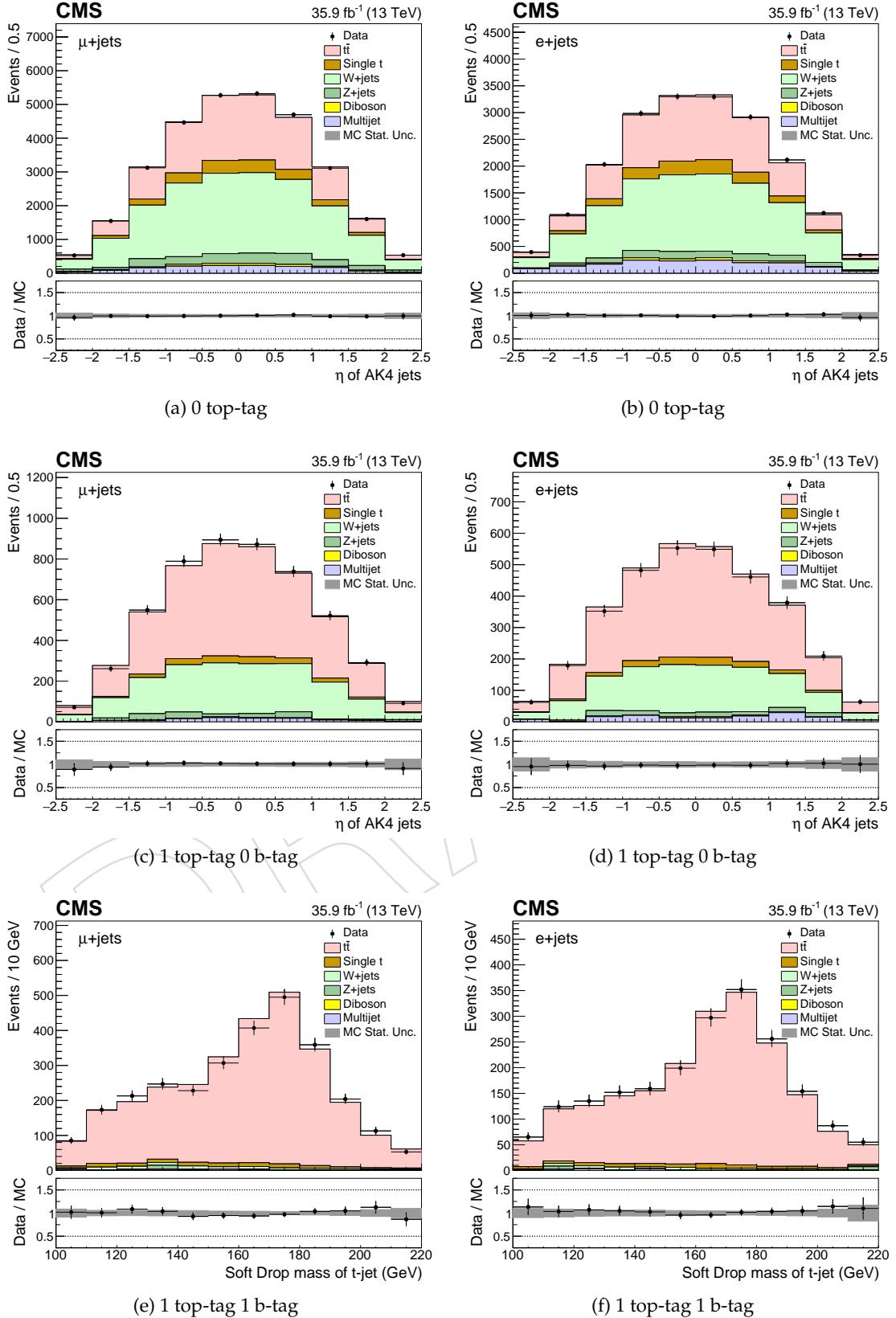


Figure 115: Post-fit kinematic distributions for top jet $|\eta| < 1.0$ GeV in the muon (left) and electron (right) channels for the top jet $|\eta|$ binned fit combining both channels. For this fit, the QCD shape is taken from a data sideband.

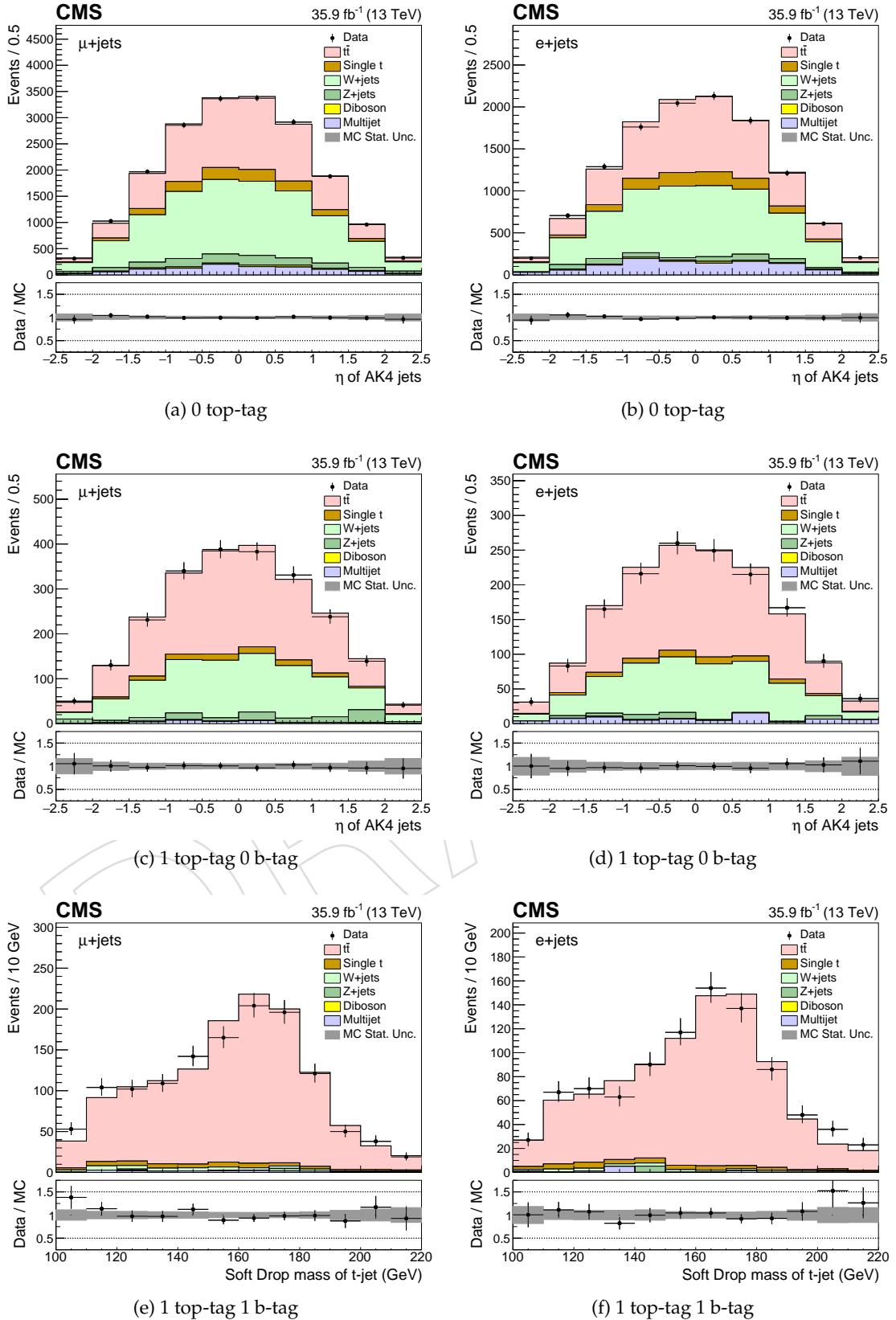


Figure 116: Post-fit kinematic distributions for top jet $|\eta| > 1.0$ GeV in the muon (left) and electron (right) channels for the top jet $|\eta|$ binned fit combining both channels. For this fit, the QCD shape is taken from a data sideband.

1108 **G.3 Results**

1109 Table 27 compares the posterior flat, p_T binned, and $|\eta|$ binned top-tag SFs. The top-tag SF
 1110 binned in top jet p_T does not bracket the flat SF, though it does agree within uncertainties. This
 1111 is likely due to the shift from introducing an additional uncertainty on the W+jets jet p_T . The
 1112 binned top-tag SF agrees better with the nominal at $p_T < 500$ GeV than at $p_T > 500$ GeV. The
 1113 top-tag SF binned in top jet $|\eta|$ brackets the flat SF, and overall has less variation than the SF
 1114 binned in top jet p_T . The uncertainties on the binned SFs are comparable to the uncertainty on
 1115 the flat SF, showing that the fit is indeed sensitive to the binned SFs. The difference between
 1116 the central values of the posterior binned SFs and the flat SF is taken as a shape uncertainty on
 1117 the top-tag and top top-mis-tag SFs.

	Top-tag SF	Top-mis-tag SF
Inclusive	0.18 ± 0.25	-1.06 ± 0.34
Top jet $p_T < 500$ GeV	0.09 ± 0.24	-1.00 ± 0.37
Top jet $p_T > 500$ GeV	-0.14 ± 0.26	-0.53 ± 0.39
Top jet $ \eta < 1.0$	0.20 ± 0.26	-1.21 ± 0.35
Top jet $ \eta > 1.0$	0.06 ± 0.27	-0.94 ± 0.39

Table 27: Posterior top(mis)tag SF nuisance parameters for the flat SF, as well as the SFs binned in top jet p_T and $|\eta|$, for the combined fit. While the results in the muon and electron channels are not shown, they agree with the combined result for all fits.

1118 H Unfolding Cross Checks

1119 This appendix describes a set of cross checks on the unfolding procedure used to extrapolate
 1120 the $t\bar{t}$ cross section as a function of the true top quark p_T or y . The cross checks generally
 1121 involve unfolding a measured top jet p_T or y distribution drawn from one half of a $t\bar{t}$ MC
 1122 sample with a response matrix drawn from the other half of the sample, and comparing the
 1123 result to the true top quark p_T or y distribution from the first half of the sample. The unfolded
 1124 and true distributions are expected to agree, meaning that any disagreements expose biases in
 1125 the unfolding procedure. For brevity, only the cross-checks involving the unfolding to parton
 1126 level are shown here. The same cross-checks were performed for the unfolding to particle level,
 1127 with very similar results.

1128 H.1 Unfolding Procedure

1129 The TUnfold algorithm is based on a least squares minimization with Tikhonov regularization.
 1130 Specifically, the algorithm finds the stationary point of

$$\mathcal{L}(x, \lambda) = \mathcal{L}_1 + \mathcal{L}_2 + \mathcal{L}_3 \quad (13)$$

$$\mathcal{L}_1 = (y - Ax)^T V_{yy}^{-1} (y - Ax) \quad (14)$$

$$\mathcal{L}_2 = \tau^2 (x - f_b x_0)^T (L^T L) (x - f_b x_0) \quad (15)$$

$$\mathcal{L}_3 = \lambda (Y - e^T x) \quad (16)$$

$$Y = \sum_i y_i \quad (17)$$

$$e_j = \sum_i A_{ij} \quad (18)$$

1131 Here, y is the measured result (with covariance matrix V_{yy}) and x is the unfolded measurement,
 1132 with A the response matrix linking them. The term \mathcal{L}_1 describes the least-squares minimiza-
 1133 tion, while \mathcal{L}_2 parameterizes the regularization. \mathcal{L}_3 is an optional area constraint, which is not
 1134 used in the following studies. The regularization is described by the regularization strength τ ,
 1135 the regularization mode L , the bias x_0 and its strength f_b .

1136 Two options are available for the regularization mode: regularization of the derivative and
 1137 regularization of the curvature. When regularizing the derivative, L is an $n \times n-1$ matrix with
 1138 two non-zero elements per row, $L_{i,i} = -1$ and $L_{i,i+1} = 1$. The derivative is thus approximated
 1139 by terms in the form of $(x_{i+1} - x_i)$. When regularizing the curvature, L is an $n \times n-2$ matrix
 1140 with three non-zero elements per row, namely $L_{i,i} = 1$, $L_{i,i+1} = -2$ and $L_{i,i+2} = 1$. Two options
 1141 are also available to determine the optimal regularization strength τ . The first, referred to as
 1142 the LCure method, involves finding a kink in the distribution $Y(X)$, where $Y(\tau) = \log \frac{\mathcal{L}_2}{\tau^2}$
 1143 and $X(\tau) = \log \mathcal{L}_1$. The second, referred to as the ScanTau method, relies on minimizing the
 1144 global bin correlations in the unfolded result. The global correlations include contributions
 1145 from systematic uncertainties.

1146 H.2 Unfolding Cross Checks for p_T unfolding

1147 H.2.1 Optimization Studies Without Systematics

1148 When first assessing the effect of different unfolding methods, the unfolding was performed
 1149 without including systematic uncertainties. As a result, the only uncertainties on the unfolded

1150 result were statistical uncertainties resulting from either the input measured top jet p_T distribution
 1151 or the response matrix. The optimal unfolding in this case can be considered to minimize
 1152 propagation of statistical uncertainties.

1153 **H.2.1.1 Curvature regularized unfolding** The unfolding was first performed using cur-
 1154 vature regularization. First, the unfolding was performed using the optimal τ determined from
 1155 ScanTau. The unfolding was then performed using the optimal τ determined from the LCure
 1156 method. In each case, 1000 toys were generated to measure the average unfolding bias, sta-
 1157 tistical uncertainty from the input distribution, and statistical uncertainty from the response
 1158 matrix. Figure 117 shows the results for the unfolding using τ from the ScanTau method, while
 1159 Figure 118 shows the results of the unfolding using τ from the LCure method. The LCure
 1160 method gives a smaller optimal τ , and a correspondingly smaller bias.

1161 **H.2.1.2 Derivative regularized unfolding** The unfolding was next performed using deriva-
 1162 tive regularization. First, the unfolding was performed using the optimal τ determined from
 1163 ScanTau. The unfolding was then performed using the optimal τ determined from the LCure
 1164 method. In each case, 1000 toys were generated to measure the average unfolding bias, input
 1165 statistical uncertainty, and response matrix statistical uncertainty. Figure 119 shows the results
 1166 for the unfolding using τ from the ScanTau method, while Figure 120 shows the results of the
 1167 unfolding using τ from the LCure method. The derivative regularization has very similar
 1168 bias to the curvature regularization when using the LCure method, and higher bias than the
 1169 curvature regularization when using the ScanTau method.

1170 **H.2.1.3 Curvature regularized unfolding, p_T range [350,2000] GeV** The unfolding was
 1171 then performed using a wider p_T range than previously. For previous iterations of the un-
 1172 folding, the parton level and measured phase spaces required the top p_T to be in the range
 1173 [400,1200] GeV. This is the p_T range covered by the differential cross section measurement.
 1174 Events outside of this range were treated as overflow, either as fakes (in measured phase space
 1175 but not parton-level) or misses (in parton-level phase space but not measured). With the wider
 1176 p_T range, two additional bins were added to cover the range [350,2000] GeV. The unfolding was
 1177 performed using the expanded response matrix and expanded measured distribution. How-
 1178 ever, the first and last bins were then dropped for the differential cross section result, and
 1179 thus were not included in the bias calculations. The expanded response matrices for the range
 1180 [350,2000] GeV are shown in Figure 121.

1181 The unfolding check was performed in the same way as previously. First, the unfolding was
 1182 performed using the optimal τ determined from ScanTau. The unfolding was then performed
 1183 using the optimal τ determined from the LCure method. In each case, 1000 toys were gener-
 1184 ated to measure the average unfolding bias, input statistical uncertainty, and response matrix
 1185 statistical uncertainty. Figure 122 shows the results for the unfolding using τ from the ScanTau
 1186 method, while Figure 123 shows the results of the unfolding using τ from the LCure method.
 1187 The unfolding using the p_T range [350,2000] GeV gives a roughly comparable bias in the range
 1188 [400,1200] GeV to the unfolding just using the range [400,1200] GeV when using the ScanTau
 1189 method, and is slightly worse when using the LCure method.

1190 **H.2.1.4 Non-regularized unfolding** From the above variants, the optimal regularization
 1191 method appears to be curvature regularization using the value of τ given by the LCure
 1192 method. In general, the smaller regularization strengths preferred by the LCure method yield
 1193 smaller biases without noticeable increase in statistical uncertainty. To pursue this further, a
 1194 non-regularized unfolding ($\tau = 0$) is performed. The results are shown in Figure 124. The

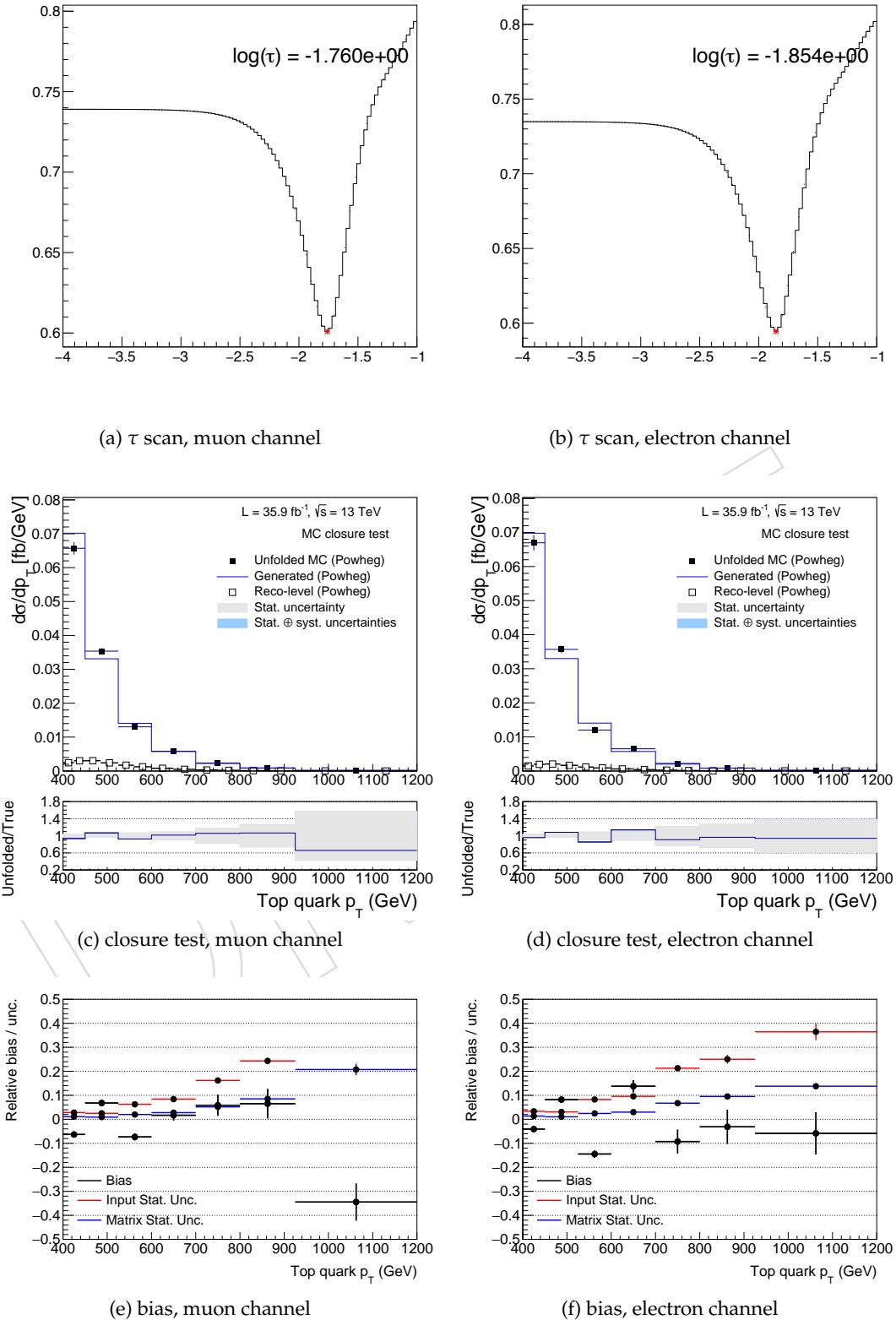


Figure 117: Results for the p_T unfolding using curvature regularization, without including systematic uncertainties, using the ScanTau method to determine τ . The top plots show the average global correlation per bin as a function of τ , which is minimized to determine the optimal τ . The middle plots show the results of the closure test unfolding one half-sample with the other. The bottom plots show the average bias, input statistical uncertainty, and response matrix statistical uncertainty when unfolding 1000 toys. Results are shown for the muon channel on the left and the electron channel on the right.

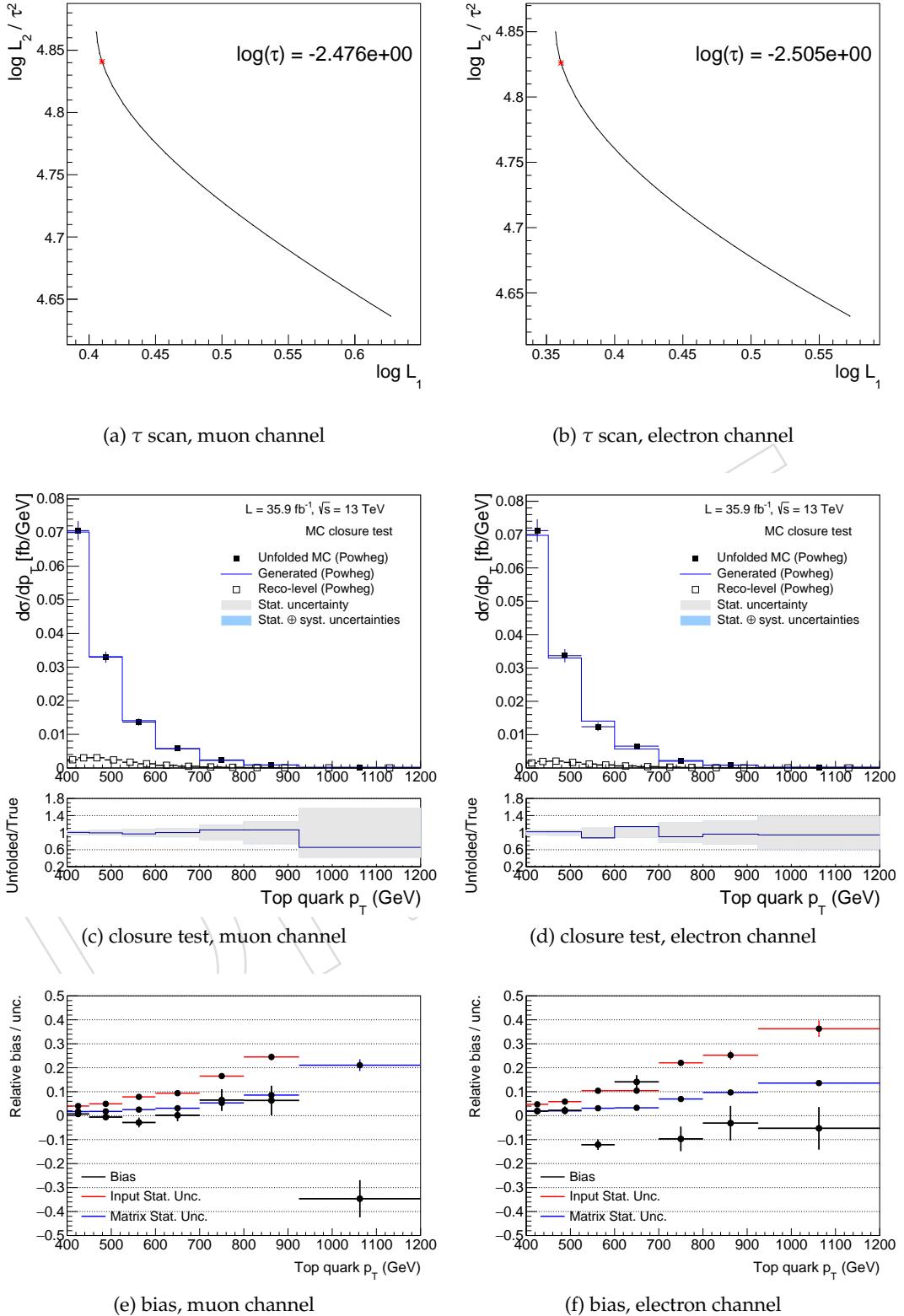


Figure 118: Results for the p_T unfolding using curvature regularization, without including systematic uncertainties, using the LCcurve method to determine τ . The top plots show the LCcurve, with the optimal τ determined from the point of maximum curvature. The middle plots show the results of the closure test unfolding one half-sample with the other. The bottom plots show the average bias, input statistical uncertainty, and response matrix statistical uncertainty when unfolding 1000 toys. Results are shown for the muon channel on the left and the electron channel on the right.

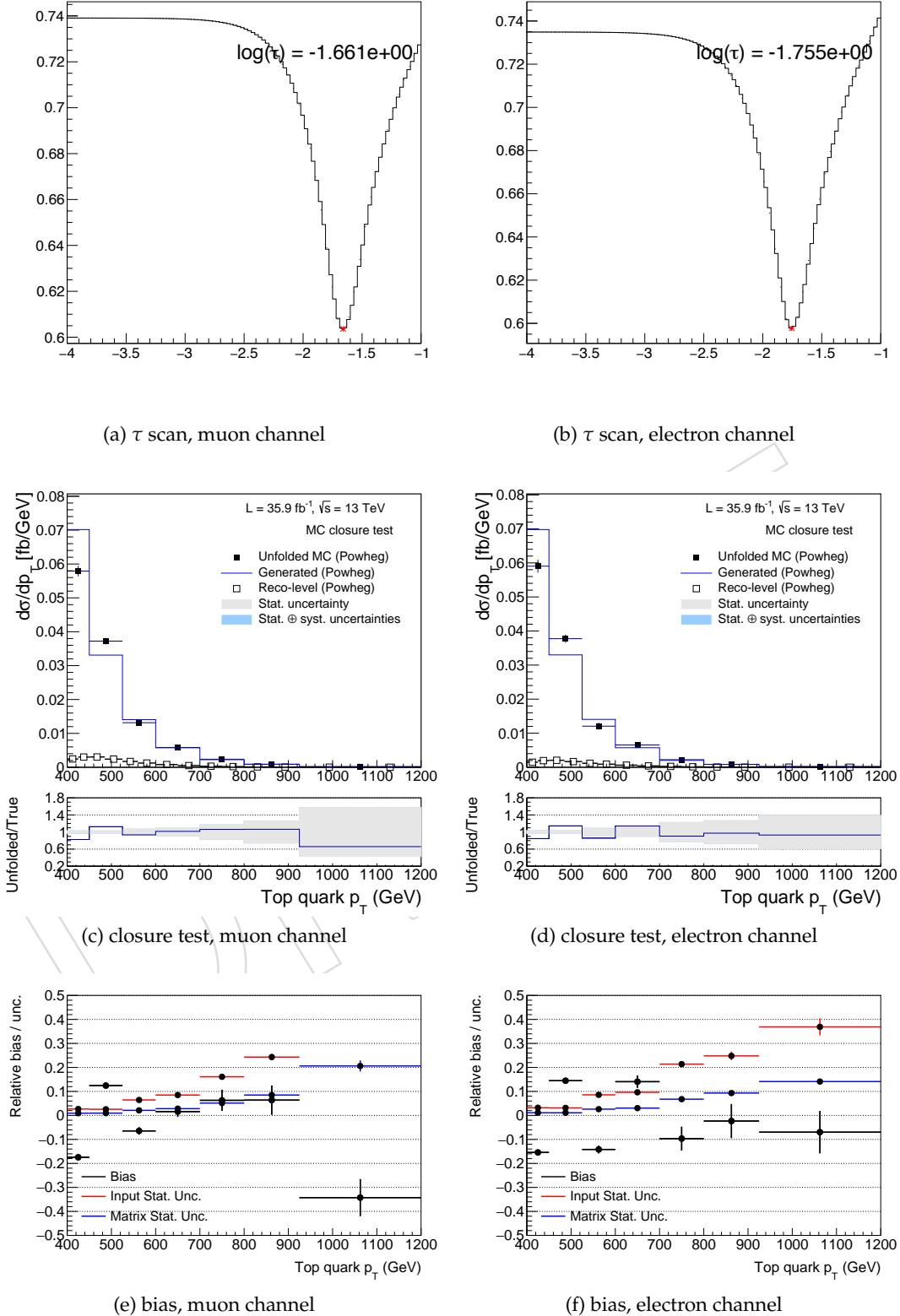


Figure 119: Results for the p_T unfolding using derivative regularization, without including systematic uncertainties, using the ScanTau method to determine τ . The top plots show the average global correlation per bin as a function of τ , which is minimized to determine the optimal τ . The middle plots show the results of the closure test unfolding one half-sample with the other. The bottom plots show the average bias, input statistical uncertainty, and response matrix statistical uncertainty when unfolding 1000 toys. Results are shown for the muon channel on the left and the electron channel on the right.

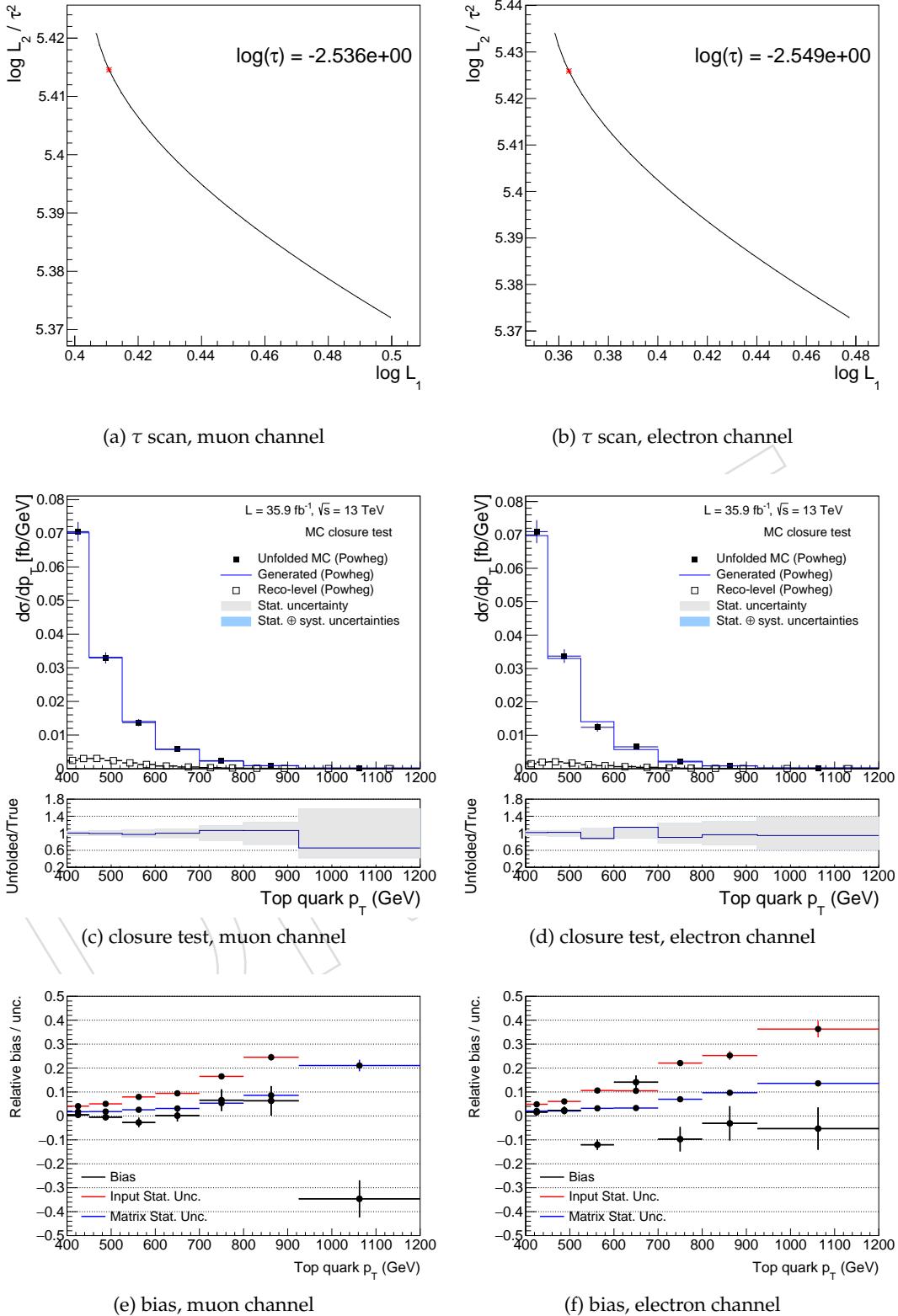


Figure 120: Results for the p_T unfolding using derivative regularization, without including systematic uncertainties, using the LCcurve method to determine τ . The top plots show the LCcurve, with the optimal τ determined from the point of maximum curvature. The middle plots show the results of the closure test unfolding one half-sample with the other. The bottom plots show the average bias, input statistical uncertainty, and response matrix statistical uncertainty when unfolding 1000 toys. Results are shown for the muon channel on the left and the electron channel on the right.

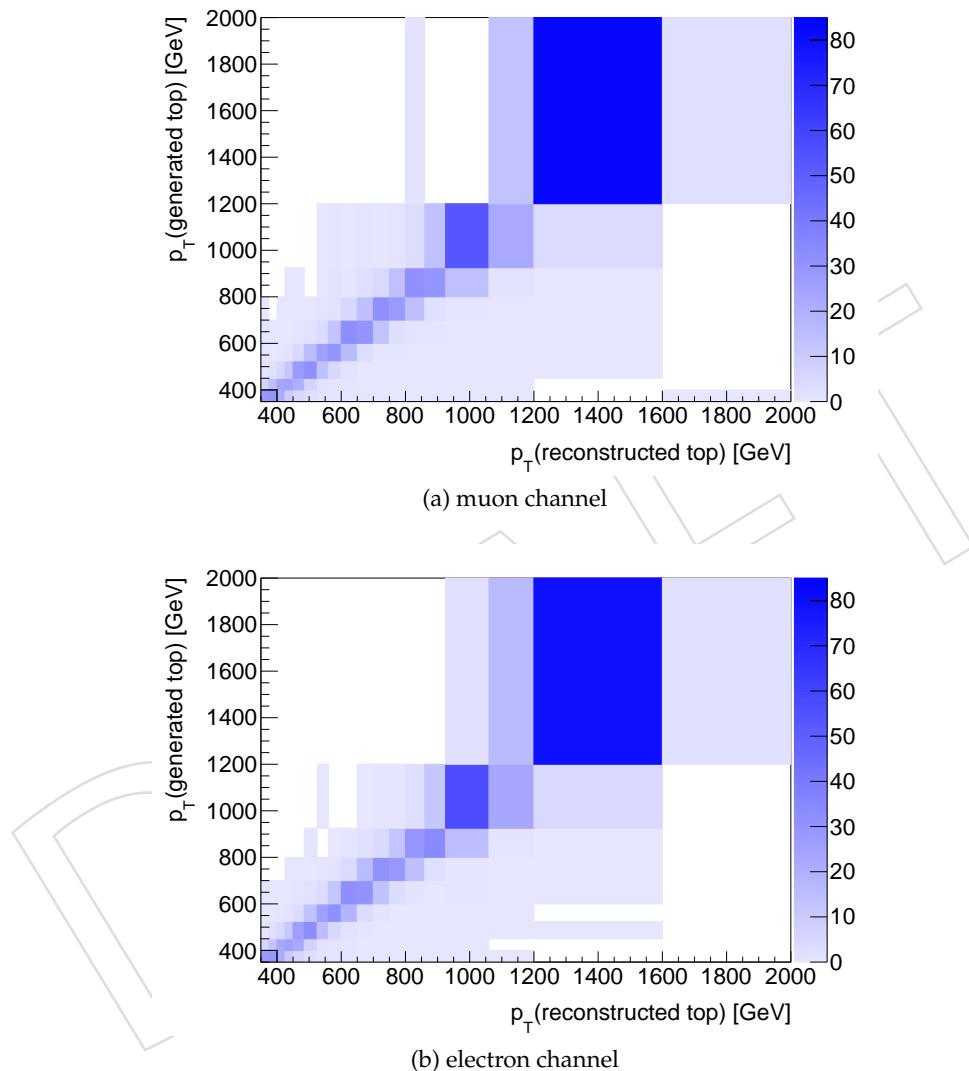


Figure 121: Response matrices for the p_T range [350,2000] GeV, for the muon and electron channels. The response matrices are normalized to 100% in each bin of generated top quark p_T .

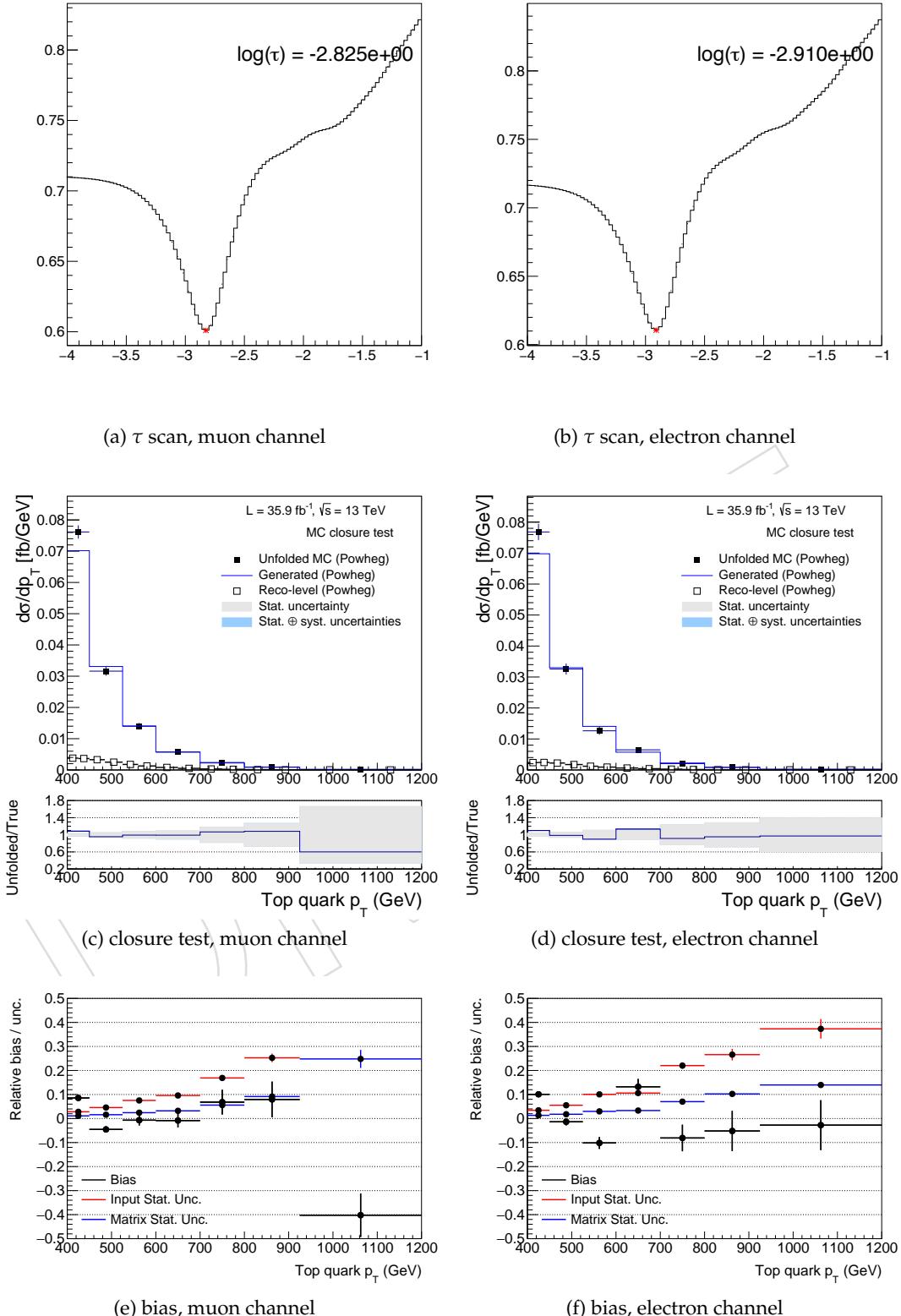


Figure 122: Results for the p_T unfolding using curvature regularization, without including systematic uncertainties, using the ScanTau method to determine τ . The unfolding is done for the extended p_T range [350,2000] GeV, though results are still only shown for [400,1200] GeV. The top plots show the average global correlation per bin as a function of τ , which is minimized to determine the optimal τ . The middle plots show the results of the closure test unfolding one half-sample with the other. The bottom plots show the average bias, input statistical uncertainty, and response matrix statistical uncertainty when unfolding 1000 toys. Results are shown for the muon channel on the left and the electron channel on the right.

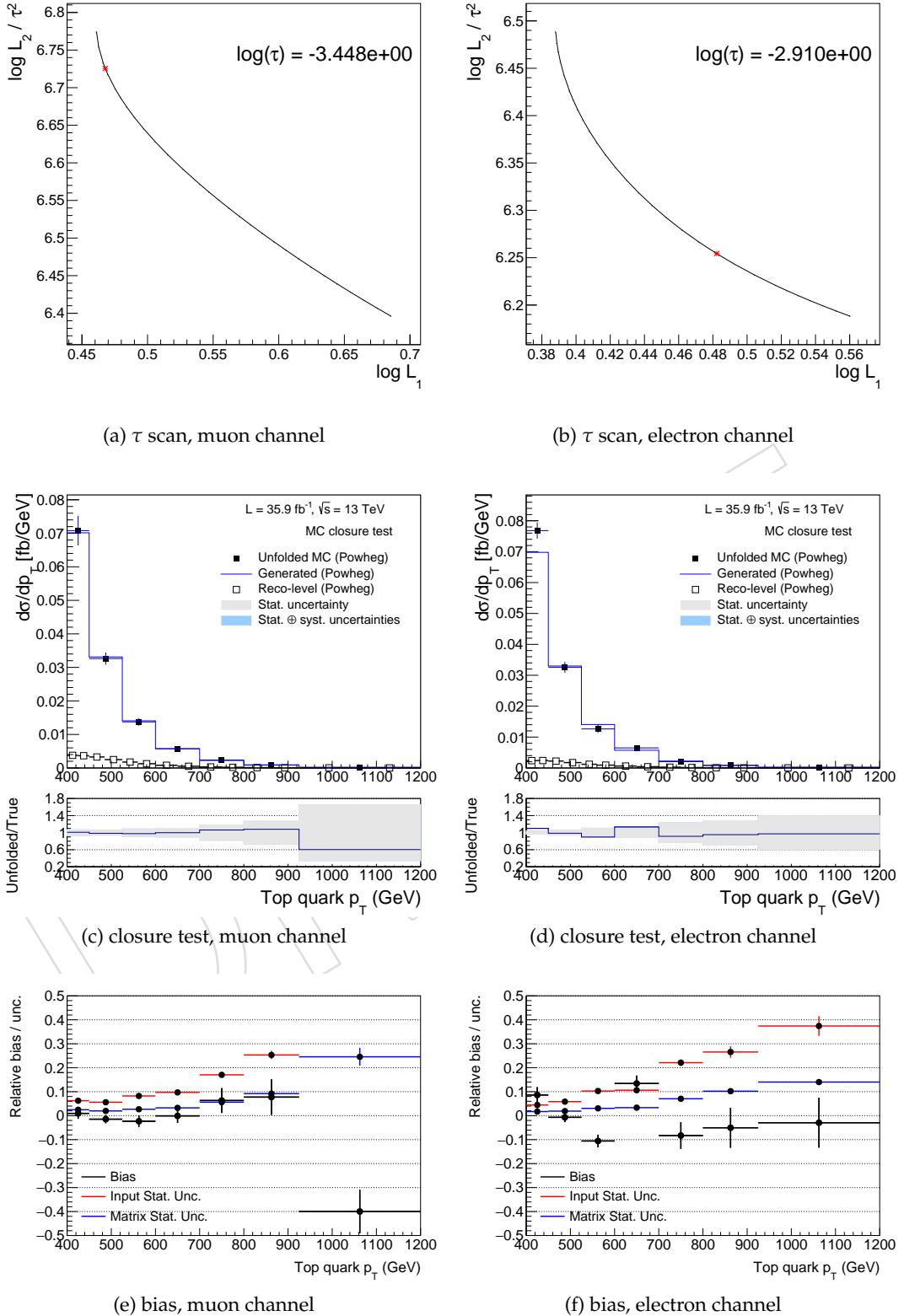


Figure 123: Results for the p_T unfolding using curvature regularization, without including systematic uncertainties, using the LCure method to determine τ . The unfolding is done for the extended p_T range [350,2000] GeV, though results are still only shown for [400,1200] GeV. The top plots show the LCure, with the optimal τ determined from the point of maximum curvature. The middle plots show the results of the closure test unfolding one half-sample with the other. The bottom plots show the average bias, input statistical uncertainty, and response matrix statistical uncertainty when unfolding 1000 toys. Results are shown for the muon channel on the left and the electron channel on the right.

1195 results seem comparable to the curvature regularization using LCurve. Finally, to cross-check
 1196 the overall effect of the regularization on the uncertainty and bias in the unfolding, a coarse τ
 1197 scan was performed. For each value of τ , the unfolding was performed and the relative input
 1198 and matrix statistical uncertainties in all bins averaged. The relative bias in each bin was also
 1199 averaged. The average input statistical uncertainty, matrix statistical uncertainty, and bias are
 1200 plotted as a function of τ in Figure 125 for the muon and electron channels. The bias and un-
 1201 certainty appear to be stable for $\log(\tau)$ below -2.5. Therefore, non-regularized unfolding seems
 1202 to be optimal, offering the lowest bias as well as the most stable unfolding approach.

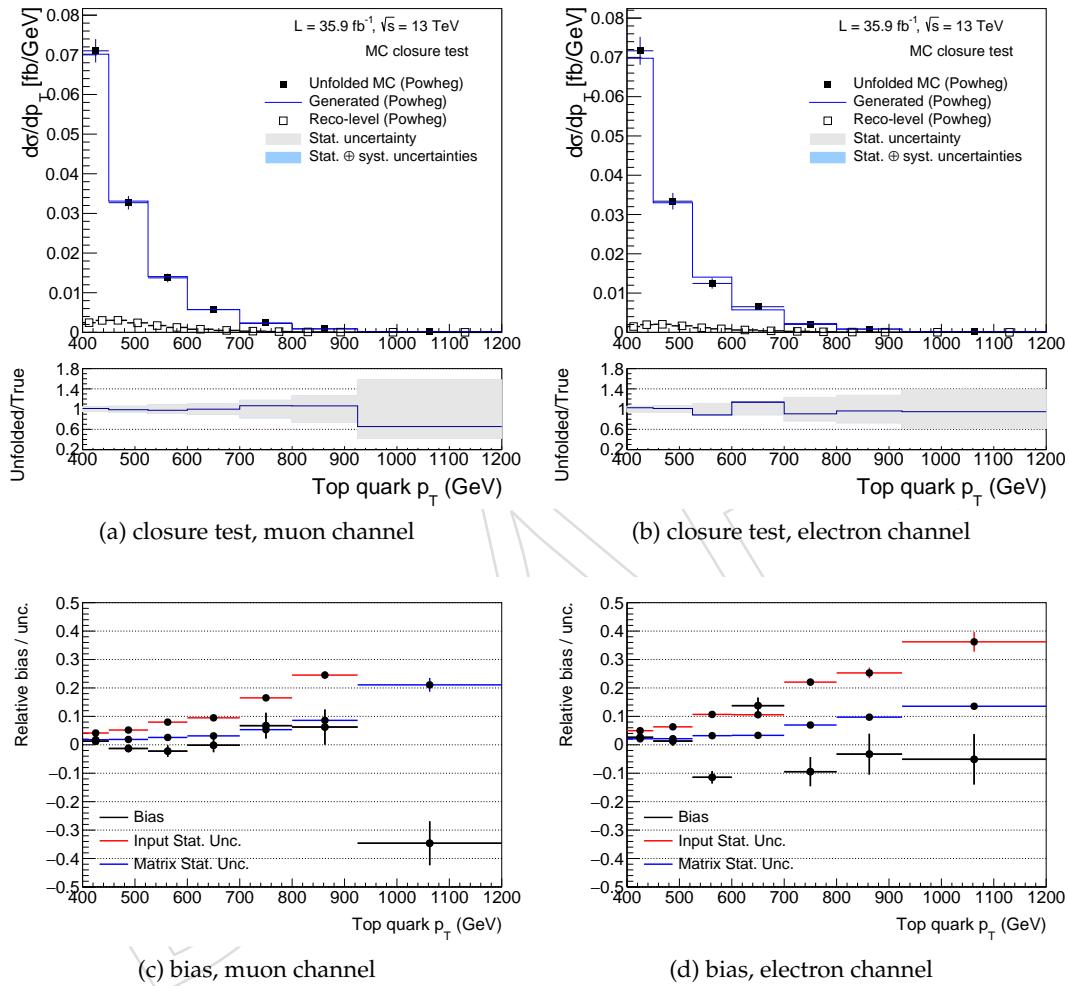


Figure 124: Results for the non-regularized ($\tau = 0$) p_T unfolding. The top plots show the results of the closure test unfolding one half-sample with the other, and the bottom plots show the bias, input statistical uncertainty, and response matrix statistical uncertainty from unfolding 1000 toys. Results are shown for the muon channel on the left and the electron channel on the right.

1203 H.2.2 Optimization Studies With Systematics

1204 To ensure that the non-regularized unfolding remains optimal in the presence of systematic
 1205 uncertainties, the unfolding checks were also performed including systematic uncertainties.
 1206 All systematic uncertainties affecting the response matrix were included. The experimental

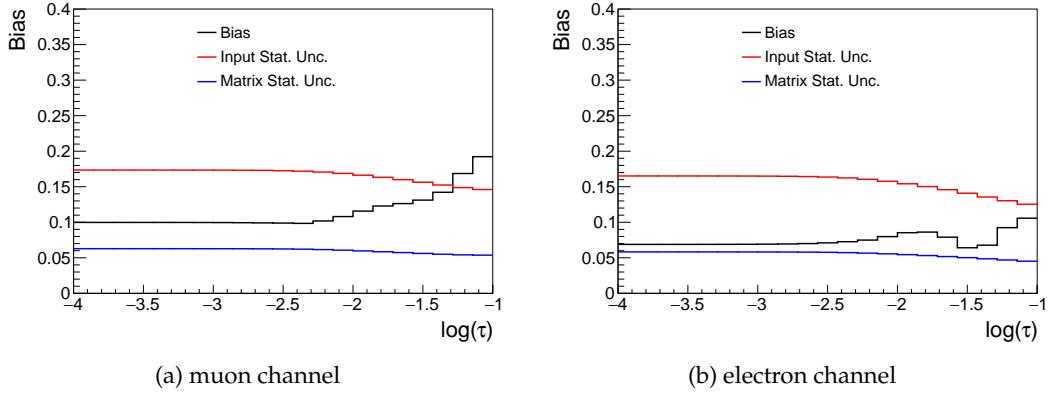


Figure 125: Average over all bins of relative input statistical uncertainty, matrix statistical uncertainty, and bias as a function of regularization strength τ for the p_T unfolding. The uncertainty and bias do not appear to depend on τ below $\log(\tau) = -2.5$. Above $\log(\tau) = -2.5$, the increase in bias is larger than the decrease in statistical uncertainty, favoring smaller regularization strengths.

1207 uncertainties, described in Section 7.1, include the lepton ID / trigger / reconstruction effi-
 1208 ciency, pileup corrections, jet energy scale, jet energy resolution, b tagging efficiency, and top
 1209 tagging efficiency. The theoretical uncertainties due to PDF, renormalization and factorization
 1210 scales, shower scales, underlying event, matching scale, color reconnection, and parton shower
 1211 are also considered. Uncertainties due to background normalizations or luminosity were not
 1212 considered, as the measured top jet p_T distribution is taken from $t\bar{t}$ MC rather than data.

1213 **H.2.2.1 Curvature regularized unfolding** The unfolding was first performed using cur-
 1214 vature regularization. First, the unfolding was performed using the optimal τ determined from
 1215 ScanTau. The unfolding was then performed using the optimal τ determined from the LCure
 1216 method. In each case, 1000 toys were generated to measure the average unfolding bias, statis-
 1217 tical uncertainty, and total uncertainty. Figure 126 shows the results for the unfolding using τ
 1218 from the ScanTau method, while Figure 127 shows the results of the unfolding using τ from
 1219 the LCure method. The optimal τ given by the ScanTau method is much lower than the value
 1220 given for the unfolding without systematic uncertainties. The LCure method yields slightly
 1221 smaller bias than ScanTau.

1222 **H.2.2.2 Derivative regularized unfolding** The unfolding was next performed using deriva-
 1223 tive regularization. First, the unfolding was performed using the optimal τ determined from
 1224 ScanTau. The unfolding was then performed using the optimal τ determined from the LCure
 1225 method. In each case, 1000 toys were generated to measure the average unfolding bias, statis-
 1226 tical uncertainty, and total uncertainty. Figure 128 shows the results for the unfolding using τ
 1227 from the ScanTau method, while Figure 129 shows the results of the unfolding using τ from
 1228 the LCure method. The derivative regularization has very similar bias to the curvature regu-
 1229 larization when using the LCure method, and higher bias than the curvature regularization
 1230 when using the ScanTau method.

1231 **H.2.2.3 Curvature regularized unfolding, p_T range [350,2000] GeV** The unfolding was
 1232 then performed using the expanded p_T range [350,2000] GeV. The unfolding was performed us-
 1233 ing the expanded response matrix and expanded measured distribution. Again, the first and
 1234 last bins were not considered when looking at the final bias, as they would not be included

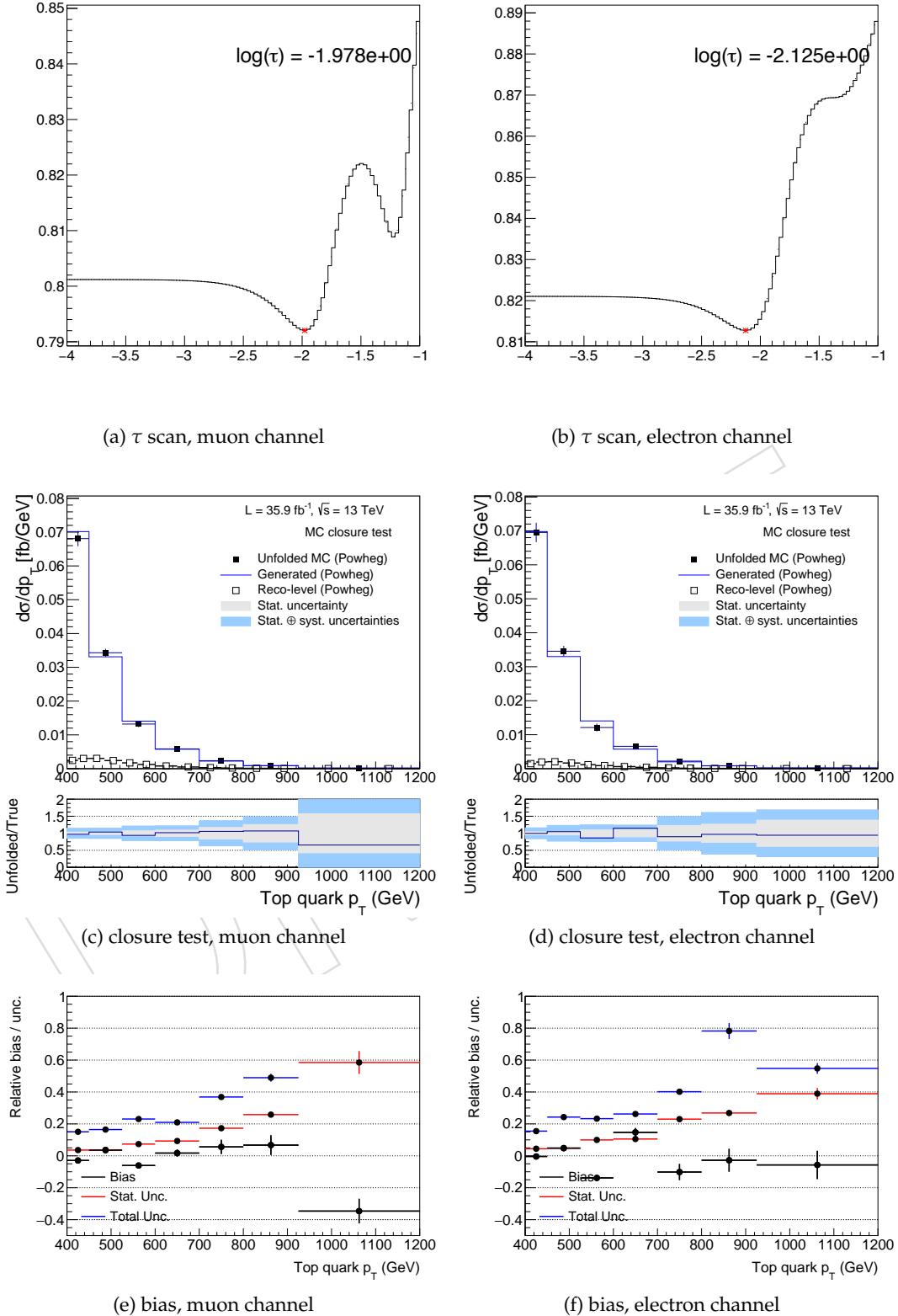


Figure 126: Results for the p_T unfolding using curvature regularization, including systematic uncertainties, using the ScanTau method to determine τ . The top plots show the average global correlation per bin as a function of τ , which is minimized to determine the optimal τ . The middle plots show the results of the closure test unfolding one half-sample with the other. The bottom plots show the average bias, statistical uncertainty, and total uncertainty when unfolding 1000 toys. Results are shown for the muon channel on the left and the electron channel on the right.

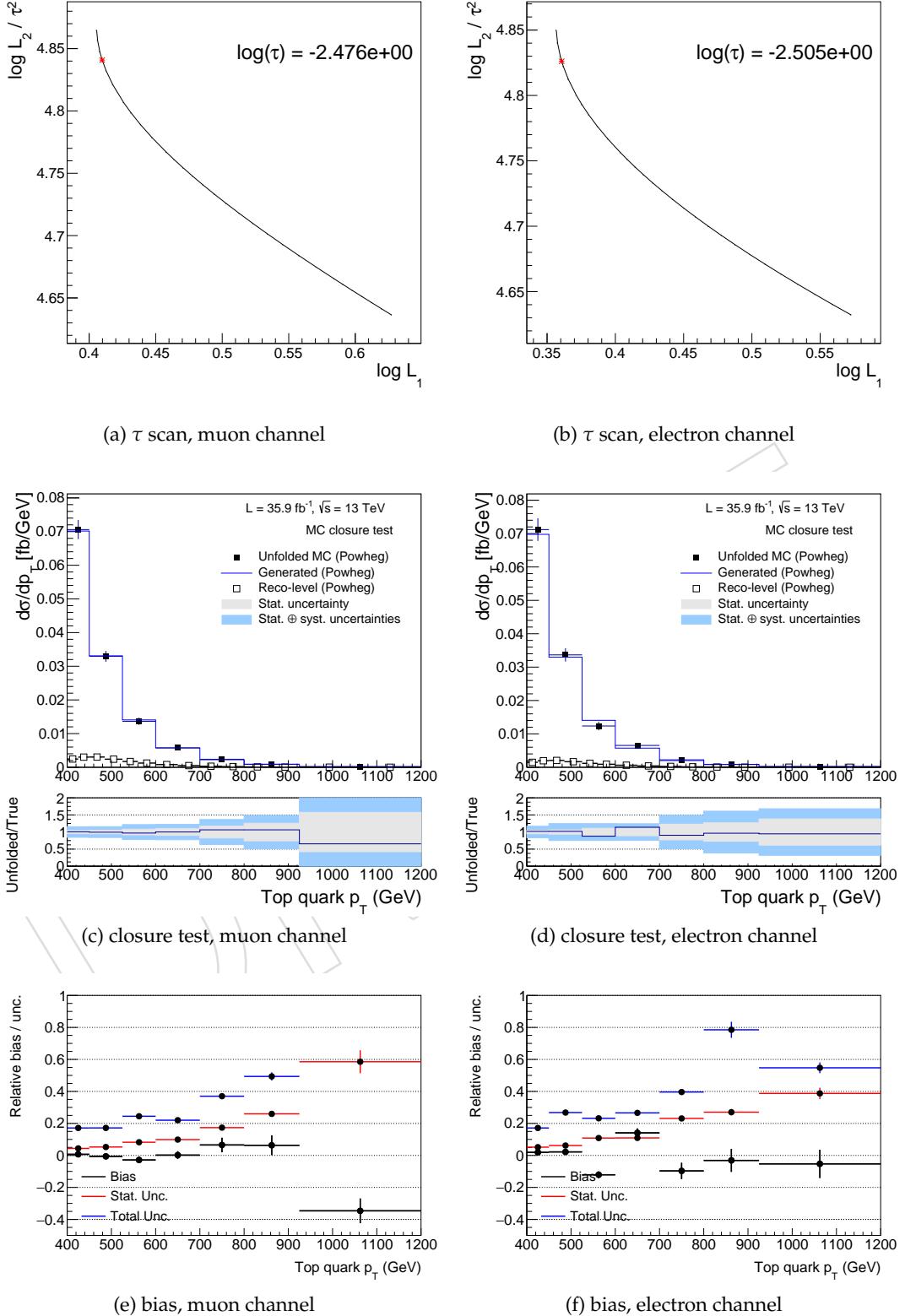


Figure 127: Results for the p_T unfolding using curvature regularization, including systematic uncertainties, using the LCcurve method to determine τ . The top plots show the LCcurve, with the optimal τ determined from the point of maximum curvature. The middle plots show the results of the closure test unfolding one half-sample with the other. The bottom plots show the average bias, statistical uncertainty, and total uncertainty when unfolding 1000 toys. Results are shown for the muon channel on the left and the electron channel on the right.

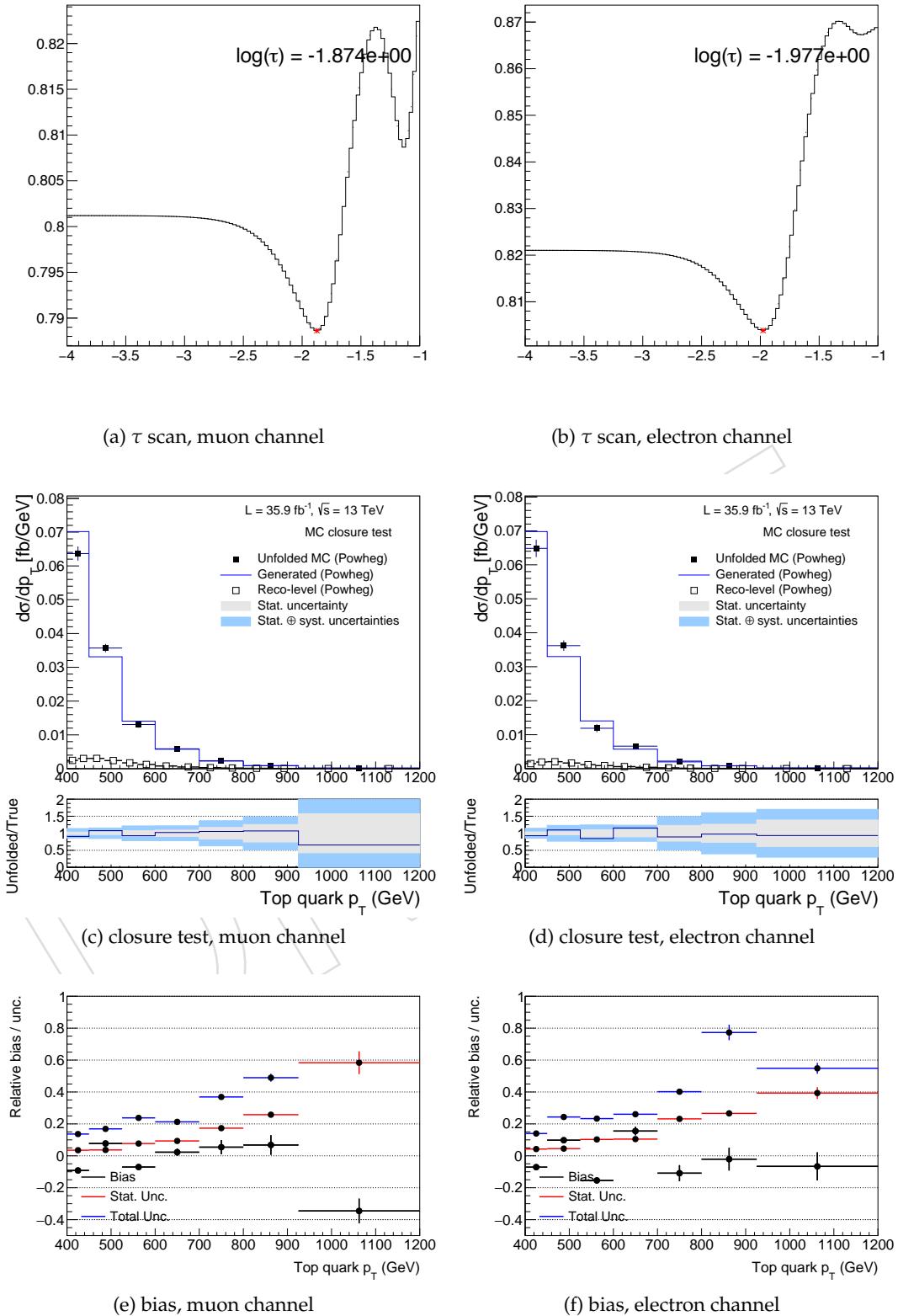


Figure 128: Results for the p_T unfolding using derivative regularization, including systematic uncertainties, using the ScanTau method to determine τ . The top plots show the average global correlation per bin as a function of τ , which is minimized to determine the optimal τ . The middle plots show the results of the closure test unfolding one half-sample with the other. The bottom plots show the average bias, statistical uncertainty, and total uncertainty when unfolding 1000 toys. Results are shown for the muon channel on the left and the electron channel on the right.

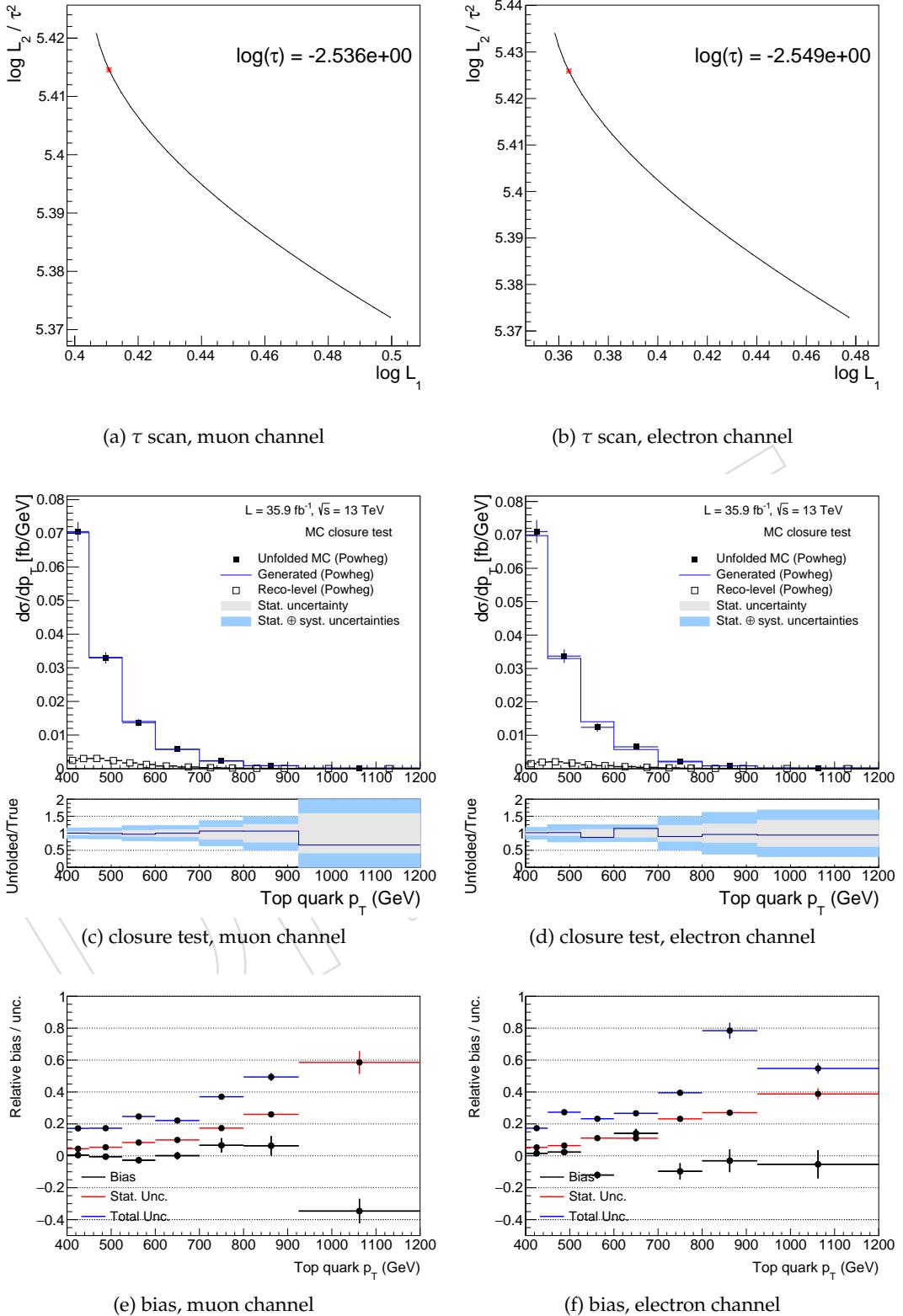


Figure 129: Results for the p_T unfolding using derivative regularization, including systematic uncertainties, using the LCcurve method to determine τ . The top plots show the LCcurve, with the optimal τ determined from the point of maximum curvature. The middle plots show the results of the closure test unfolding one half-sample with the other. The bottom plots show the average bias, statistical uncertainty, and total uncertainty when unfolding 1000 toys. Results are shown for the muon channel on the left and the electron channel on the right.

in the differential cross section result. First, the unfolding was performed using the optimal τ determined from ScanTau. The unfolding was then performed using the optimal τ determined from the LCure method. In each case, 1000 toys were generated to measure the average unfolding bias, statistical uncertainty, and total uncertainty. Figure 130 shows the results for the unfolding using τ from the ScanTau method, while Figure 131 shows the results of the unfolding using τ from the LCure method. The unfolding using the p_T range [350,2000] GeV gives a slightly better bias in the range [400,1200] GeV to the unfolding just using the range [400,1200] GeV when using the ScanTau τ , and slightly worse bias when using the LCure τ .

H.2.2.4 Non-regularized unfolding The optimal τ determined when including systematic uncertainties in the unfolding procedure is generally smaller than the optimal τ determined when neglecting systematic uncertainties. A non-regularized unfolding is therefore expected to give optimal performance for the unfolding when including systematic uncertainties. The results of the non-regularized unfolding are shown in Figure 132. The results seem comparable to the previous optimized unfolding results including systematic uncertainties. To determine the overall effect of the regularization on the uncertainty and bias in the unfolding, a coarse τ scan was performed. For each value of τ , the unfolding was performed and the relative statistical and total uncertainties in all bins averaged. The relative bias in each bin was also averaged. The statistical uncertainty, total uncertainty, and bias are plotted as a function of τ in Figure 133 for the muon and electron channels. Values of $\log(\tau)$ lower than -2.5 seem to yield comparable results to $\tau = 0$. Therefore, non-regularized unfolding seems to be optimal, offering the lowest bias as well as the most stable unfolding approach. This agrees with the results of unfolding without including systematic uncertainties. Therefore, non-regularized unfolding appears to offer the best overall performance when unfolding the p_T distribution.

H.3 Unfolding Cross Checks for rapidity unfolding

The same checks as performed on the top quark p_T unfolding are performed on the rapidity unfolding, with the exception of the extended unfolding range. The unfolding checks compare derivative and curvature regularization, the LCure and ScanTau methods for determining regularization strength, and non-regularized unfolding. Comparisons are made including and excluding systematic uncertainties. As with p_T unfolding, non-regularized unfolding is found to be optimal.

H.3.1 Optimization Studies Without Systematics

When first assessing the effect of different unfolding methods, the unfolding was performed without including systematic uncertainties. As a result, the only uncertainties on the unfolded result were statistical uncertainties resulting from either the input measured top jet rapidity distribution or the response matrix. The optimal unfolding in this case can be considered to minimize propagation of statistical uncertainties.

H.3.1.1 Curvature regularized unfolding The unfolding was first performed using curvature regularization. First, the unfolding was performed using the optimal τ determined from ScanTau. The unfolding was then performed using the optimal τ determined from the LCure method. In each case, 1000 toys were generated to measure the average unfolding bias, statistical uncertainty from the input distribution, and statistical uncertainty from the response matrix. Figure 134 shows the results for the unfolding using τ from the ScanTau method, while Figure 135 shows the results of the unfolding using τ from the LCure method. Both methods prefer similarly small values of τ , yielding similar biases and uncertainties.

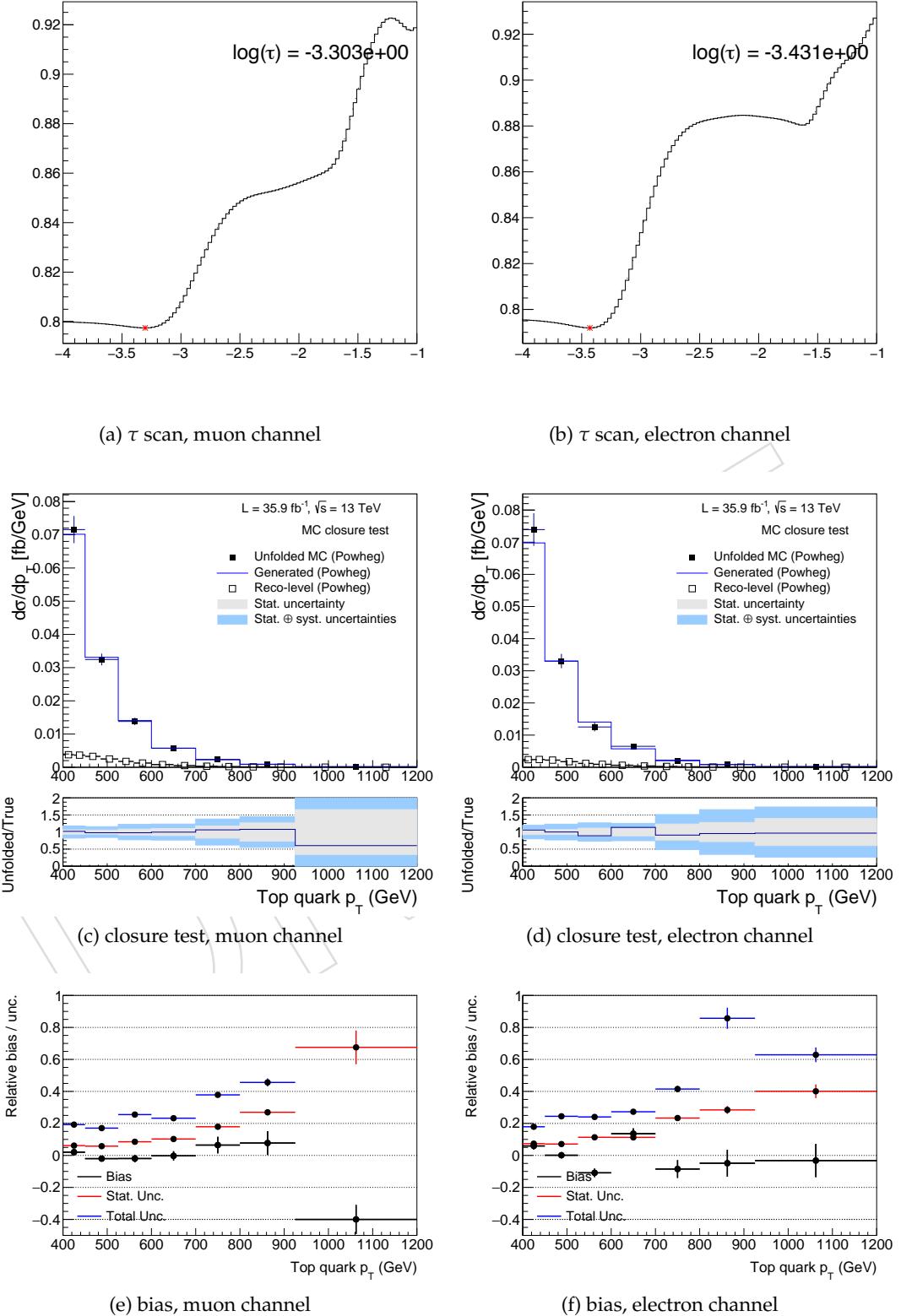


Figure 130: Results for the p_T unfolding using curvature regularization, including systematic uncertainties, using the ScanTau method to determine τ . The unfolding is done for the extended p_T range [350,2000] GeV, though results are still only shown for [400,1200] GeV. The top plots show the average global correlation per bin as a function of τ , which is minimized to determine the optimal τ . The middle plots show the results of the closure test unfolding one half-sample with the other. The bottom plots show the average bias, statistical uncertainty, and total uncertainty when unfolding 1000 toys. Results are shown for the muon channel on the left and the electron channel on the right.

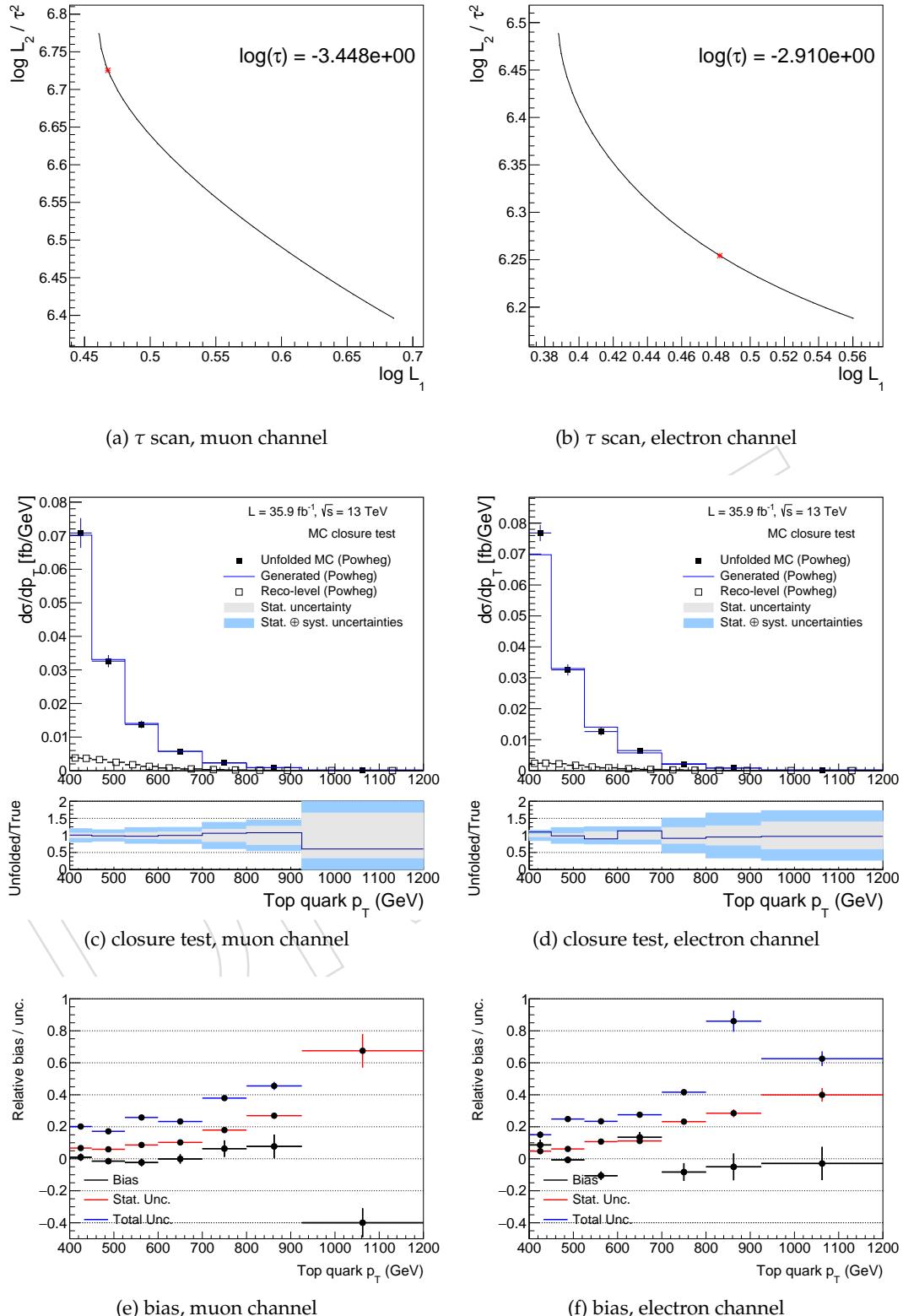


Figure 131: Results for the p_T unfolding using curvature regularization, including systematic uncertainties, using the LCure method to determine τ . The unfolding is done for the extended p_T range [350,2000] GeV, though results are still only shown for [400,1200] GeV. The top plots show the LCure, with the optimal τ determined from the point of maximum curvature. The middle plots show the results of the closure test unfolding one half-sample with the other. The bottom plots show the average bias, statistical uncertainty, and total uncertainty when unfolding 1000 toys. Results are shown for the muon channel on the left and the electron channel on the right.

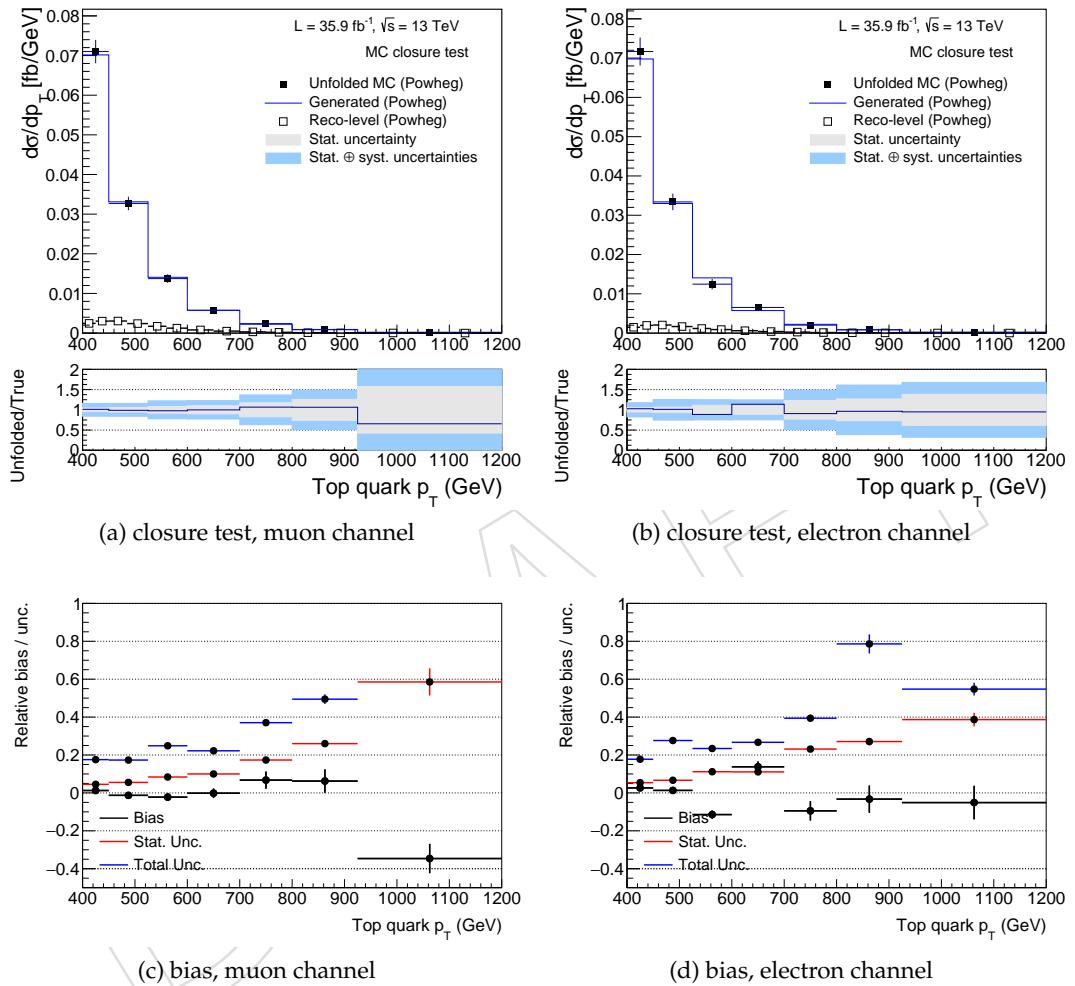


Figure 132: Results for the non-regularized ($\tau = 0$) p_T unfolding. The top plots show the results of the closure test unfolding one half-sample with the other, and the bottom plots show the bias, statistical uncertainty, and total uncertainty from unfolding 1000 toys. Results are shown for the muon channel on the left and the electron channel on the right.

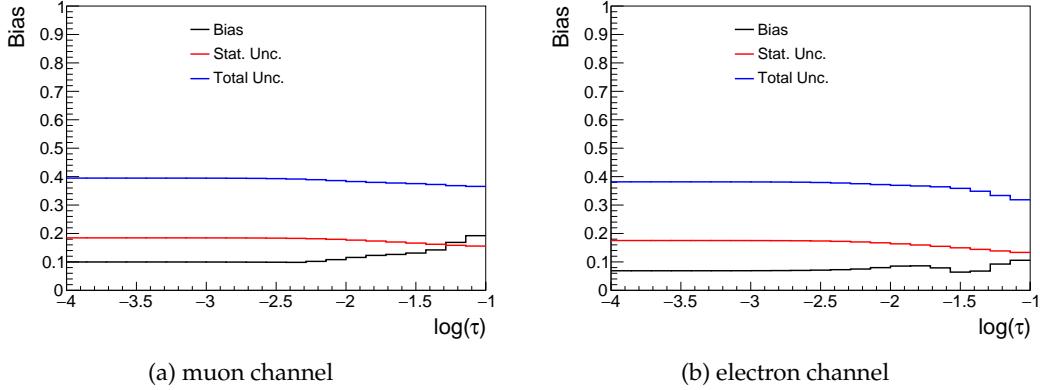


Figure 133: Sum over all bins of relative statistical uncertainty, total uncertainty, and bias as a function of regularization strength τ for the p_T unfolding. The uncertainty and bias do not appear to depend on τ below $\log(\tau) = -2.5$. Above $\log(\tau) = -2.5$, the small decrease in statistical and total uncertainty is offset by a larger increase in bias.

1279 **H.3.1.2 Derivative regularized unfolding** The unfolding was next performed using derivative regularization.
 1280 First, the unfolding was performed using the optimal τ determined from ScanTau.
 1281 The unfolding was then performed using the optimal τ determined from the LCurve
 1282 method. In each case, 1000 toys were generated to measure the average unfolding bias, input
 1283 statistical uncertainty, and response matrix statistical uncertainty. Figure 136 shows the results
 1284 for the unfolding using τ from the ScanTau method, while Figure 137 shows the results of the
 1285 unfolding using τ from the LCurve method. Both the LCurve and ScanTau methods prefer
 1286 small τ , yielding similar bias to the curvature regularization.

1287 **H.3.1.3 Non-regularized unfolding** From the above variants, the optimal regularization
 1288 uses a small τ . As such, a non-regularized unfolding is performed as a cross check. The results
 1289 are shown in Figure 138. The results are comparable to the previous unfolding variations.
 1290 Finally, to cross-check the overall effect of the regularization on the uncertainty and bias in the
 1291 unfolding, a coarse τ scan was performed. For each value of τ , the unfolding was performed
 1292 and the relative input and matrix statistical uncertainties in all bins averaged. The relative
 1293 bias in each bin was also averaged. The average input statistical uncertainty, response matrix
 1294 statistical uncertainty, and bias are plotted as a function of τ in Figure 139 for the muon and
 1295 electron channels. The coarse τ scan seems to indicate minimal bias and uncertainty at
 1296 $\log(\tau) = -2.8$. However, smaller values of τ yield only slightly larger bias and uncertainty.
 1297 Therefore, non-regularized unfolding seems to be optimal, offering a low bias as well as the
 1298 most stable unfolding approach.

1299 H.3.2 Optimization Studies With Systematics

1300 To ensure that the non-regularized unfolding remains optimal in the presence of systematic
 1301 uncertainties, the unfolding checks were also performed including systematic uncertainties.
 1302 The same systematic uncertainties as described in the p_T unfolding are included.

1303 **H.3.2.1 Curvature regularized unfolding** The unfolding was first performed using cur-
 1304 vature regularization. First, the unfolding was performed using the optimal τ determined from
 1305 ScanTau. The unfolding was then performed using the optimal τ determined from the LCurve
 1306 method. In each case, 1000 toys were generated to measure the average unfolding bias, statis-

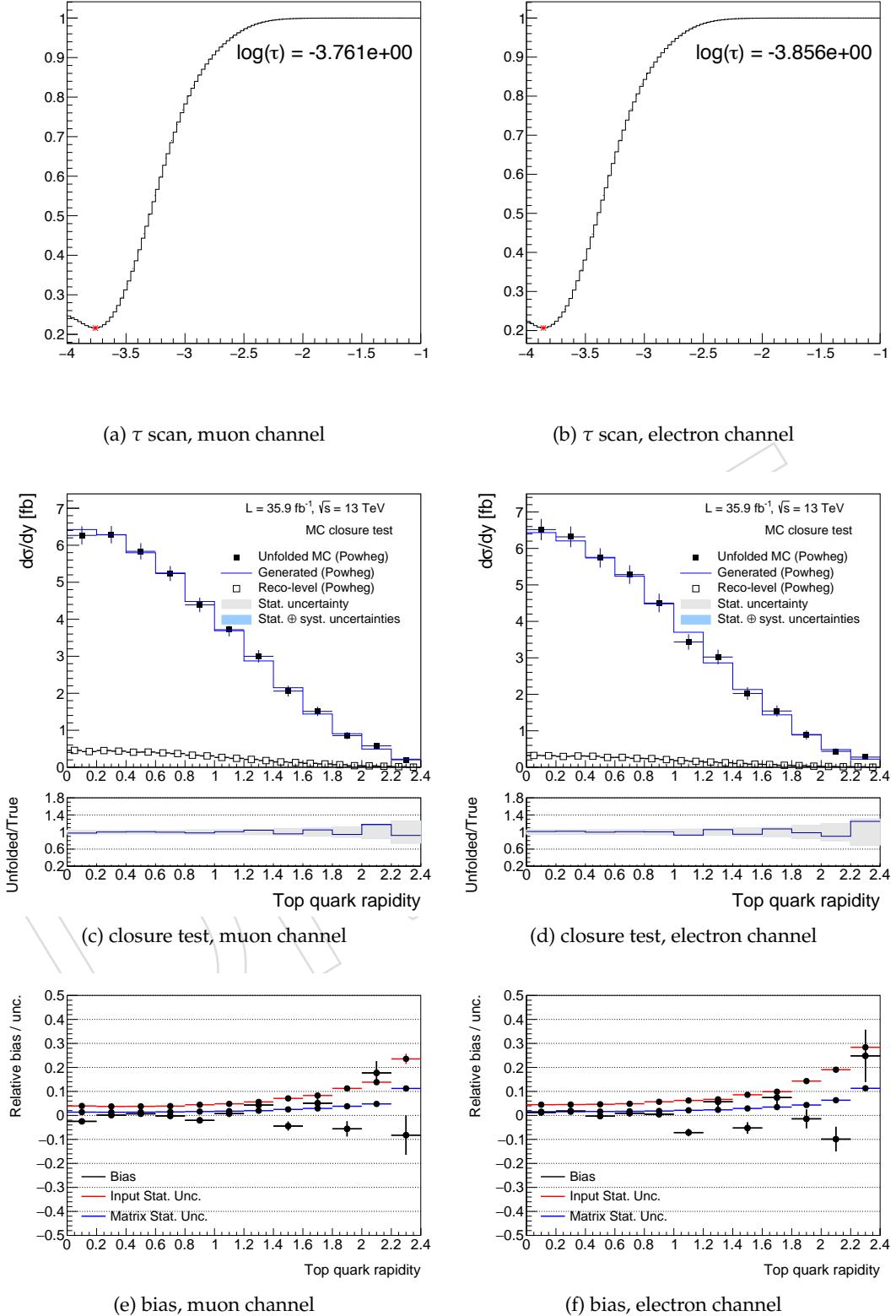


Figure 134: Results for the rapidity unfolding using curvature regularization, without including systematic uncertainties, using the ScanTau method to determine τ . The top plots show the average global correlation per bin as a function of τ , which is minimized to determine the optimal τ . The middle plots show the results of the closure test unfolding one half-sample with the other. The bottom plots show the average bias, input statistical uncertainty, and response matrix statistical uncertainty when unfolding 1000 toys. Results are shown for the muon channel on the left and the electron channel on the right.

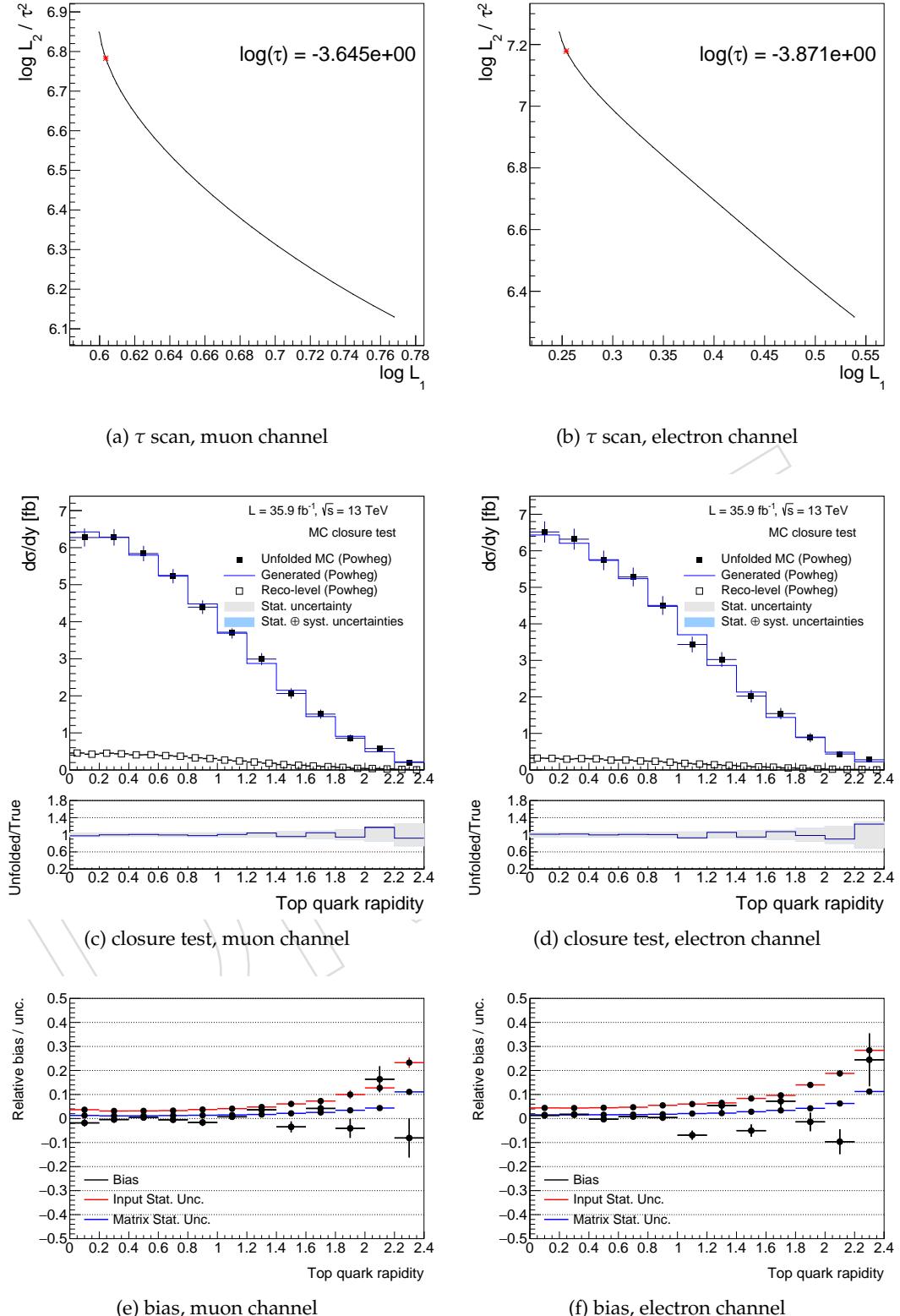


Figure 135: Results for the rapidity unfolding using curvature regularization, without including systematic uncertainties, using the LCure method to determine τ . The top plots show the LCure, with the optimal τ determined from the point of maximum curvature. The middle plots show the results of the closure test unfolding one half-sample with the other. The bottom plots show the average bias, input statistical uncertainty, and response matrix statistical uncertainty when unfolding 1000 toys. Results are shown for the muon channel on the left and the electron channel on the right.

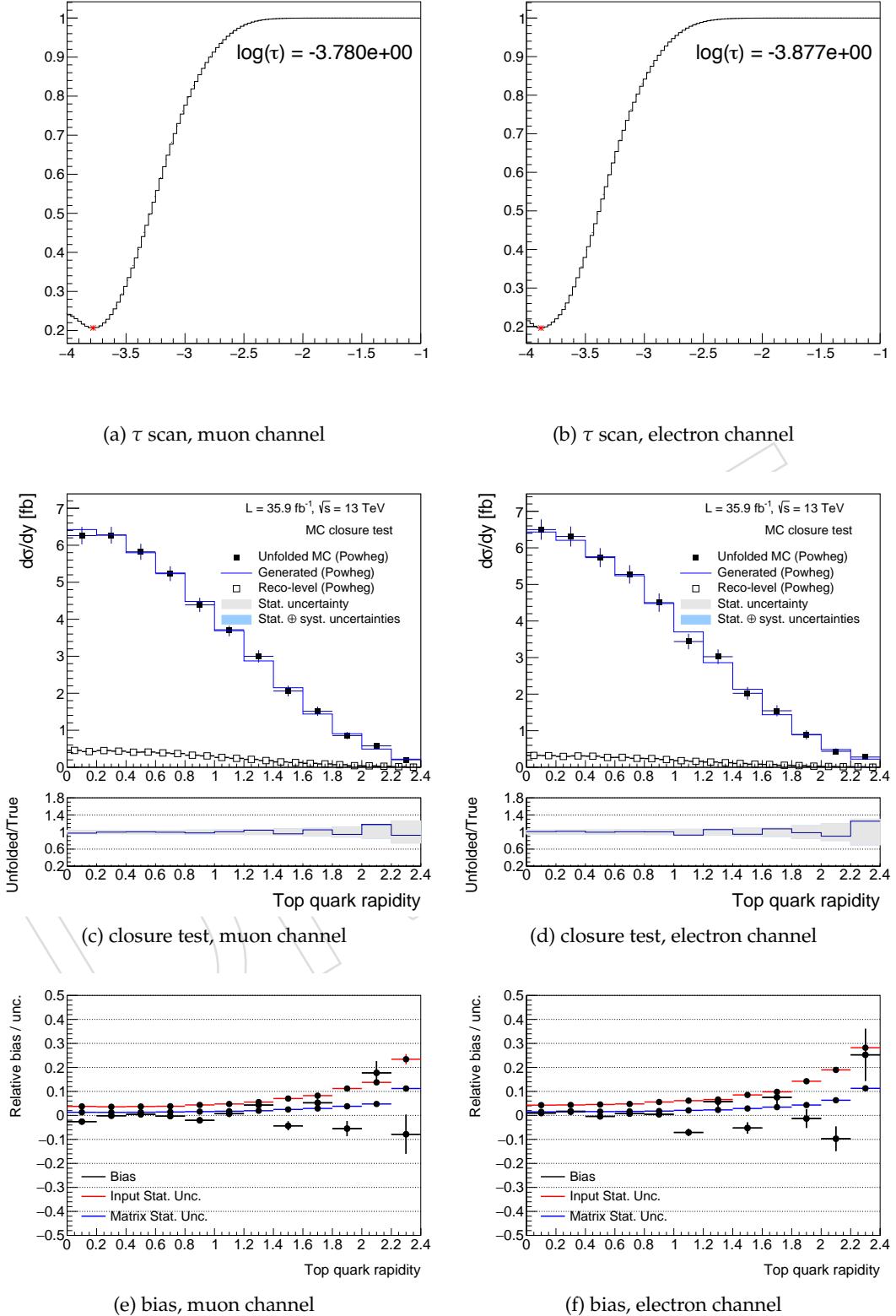


Figure 136: Results for the unfolding using derivative regularization, without including systematic uncertainties, using the ScanTau method to determine τ . The top plots show the average global correlation per bin as a function of τ , which is minimized to determine the optimal τ . The middle plots show the results of the closure test unfolding one half-sample with the other. The bottom plots show the average bias, input statistical uncertainty, and response matrix statistical uncertainty when unfolding 1000 toys. Results are shown for the muon channel on the left and the electron channel on the right.

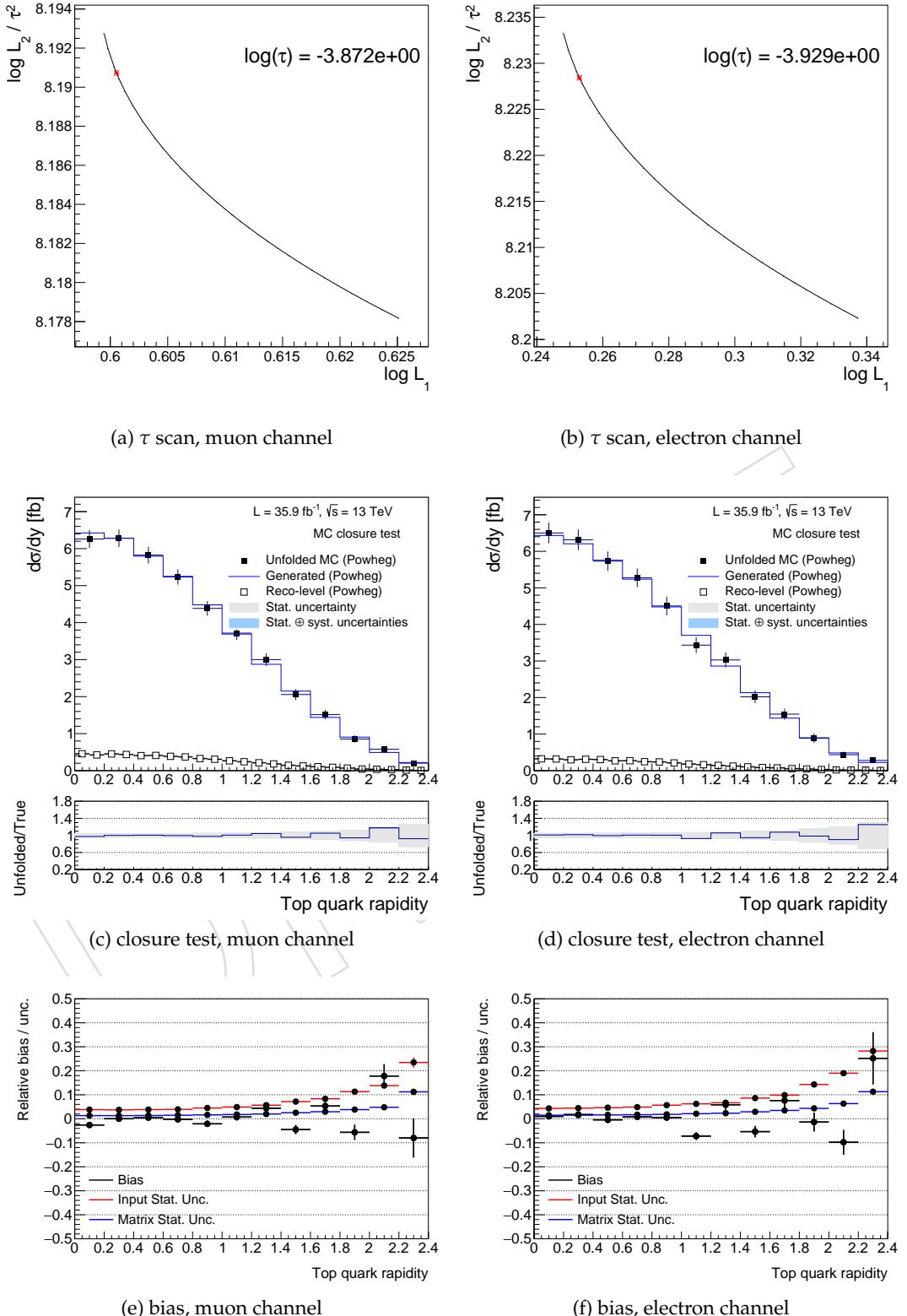


Figure 137: Results for the unfolding using derivative regularization, without including systematic uncertainties, using the LCcurve method to determine τ . The top plots show the LCcurve, with the optimal τ determined from the point of maximum curvature. The middle plots show the results of the closure test unfolding one half-sample with the other. The bottom plots show the average bias, input statistical uncertainty, and response matrix statistical uncertainty when unfolding 1000 toys. Results are shown for the muon channel on the left and the electron channel on the right.

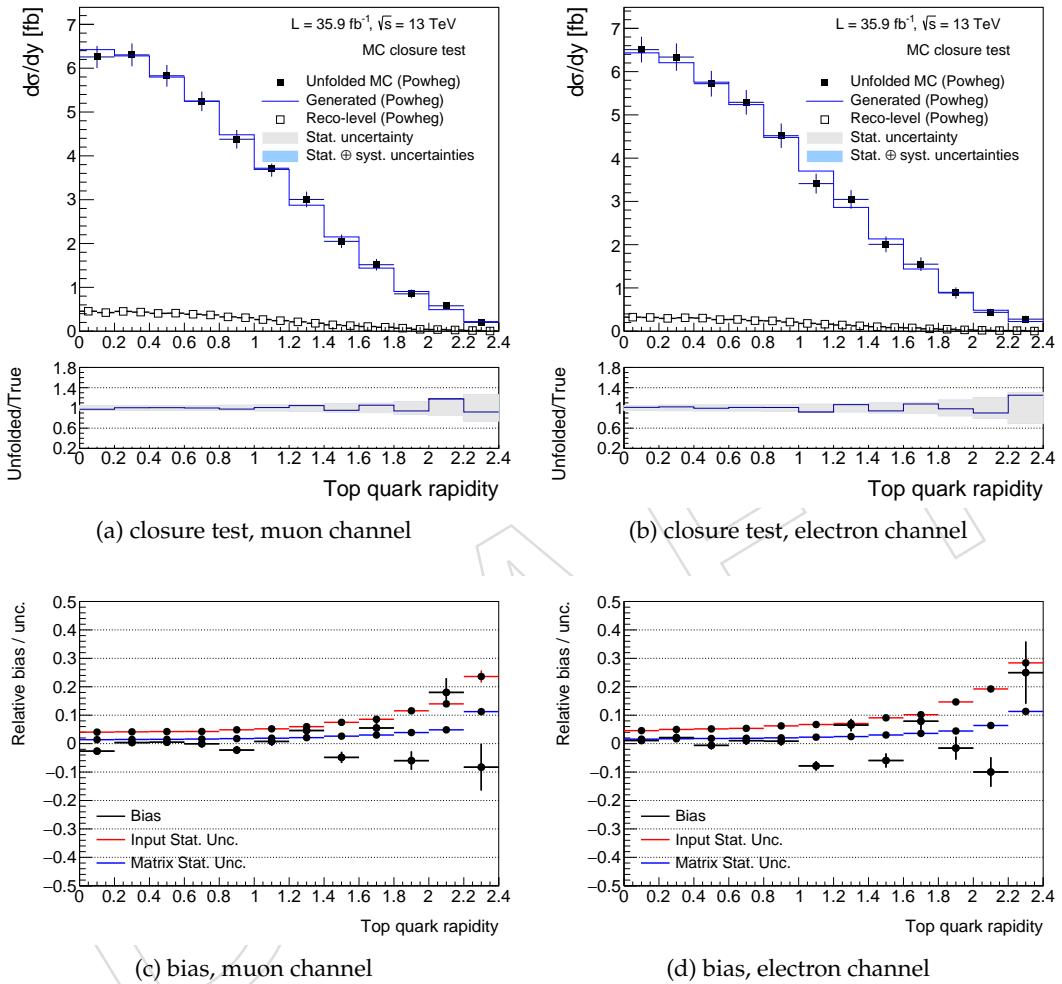


Figure 138: Results for the non-regularized ($\tau = 0$) unfolding. The top plots show the results of the closure test unfolding one half-sample with the other, and the bottom plots show the bias, input statistical uncertainty, and response matrix statistical uncertainty from unfolding 1000 toys. Results are shown for the muon channel on the left and the electron channel on the right.

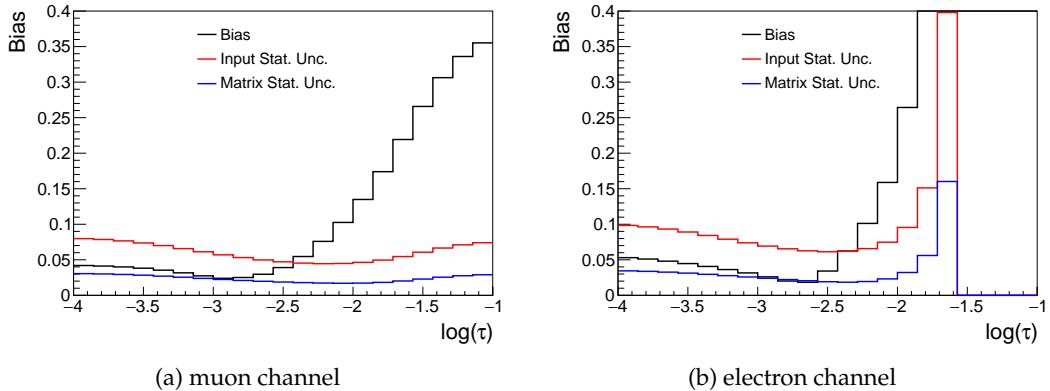


Figure 139: Average over all bins of relative input statistical uncertainty, matrix statistical uncertainty, and bias as a function of regularization strength τ . The uncertainty and bias appear to be minimized at $\log(\tau) = -2.8$.

tical uncertainty, and total uncertainty. Figure 140 shows the results for the unfolding using τ from the ScanTau method, while Figure 141 shows the results of the unfolding using τ from the LCure method. As in the unfolding without systematic uncertainties, both the LCure and ScanTau methods prefer small τ . Both methods yield similar biases and uncertainties.

H.3.2.2 Derivative regularized unfolding The unfolding was next performed using derivative regularization. First, the unfolding was performed using the optimal τ determined from ScanTau. The unfolding was then performed using the optimal τ determined from the LCure method. In each case, 1000 toys were generated to measure the average unfolding bias, statistical uncertainty, and total uncertainty. Figure 142 shows the results for the unfolding using τ from the ScanTau method, while Figure 143 shows the results of the unfolding using τ from the LCure method. The derivative regularization has comparable bias and uncertainty to the curvature regularization.

H.3.2.3 Non-regularized unfolding Similarly to the unfolding without systematic uncertainties, the optimal τ when including systematic uncertainties in the unfolding procedure is small. A non-regularized unfolding is therefore expected to give optimal performance for the unfolding when including systematic uncertainties. The results of the non-regularized unfolding are shown in Figure 144. The results seem comparable to the previous optimized unfolding results including systematic uncertainties. To determine the overall effect of the regularization on the uncertainty and bias in the unfolding, a coarse τ scan was performed. For each value of τ , the unfolding was performed and the relative statistical and total uncertainties in all bins averaged. The relative bias in each bin was also averaged. The statistical uncertainty, total uncertainty, and bias are plotted as a function of τ in Figure 145 for the muon and electron channels. In the coarse τ scan, bias and uncertainty are minimized for $\log(\tau) = -2.6$. However, smaller τ have only slightly higher bias and uncertainty. Therefore, non-regularized unfolding seems to be optimal, offering the lowest bias as well as the most stable unfolding approach. This agrees with the results of unfolding without including systematic uncertainties. Therefore, non-regularized unfolding appears to offer the best overall performance for rapidity unfolding. Using non-regularized unfolding for both the p_T and rapidity distributions maintains a consistent analysis.

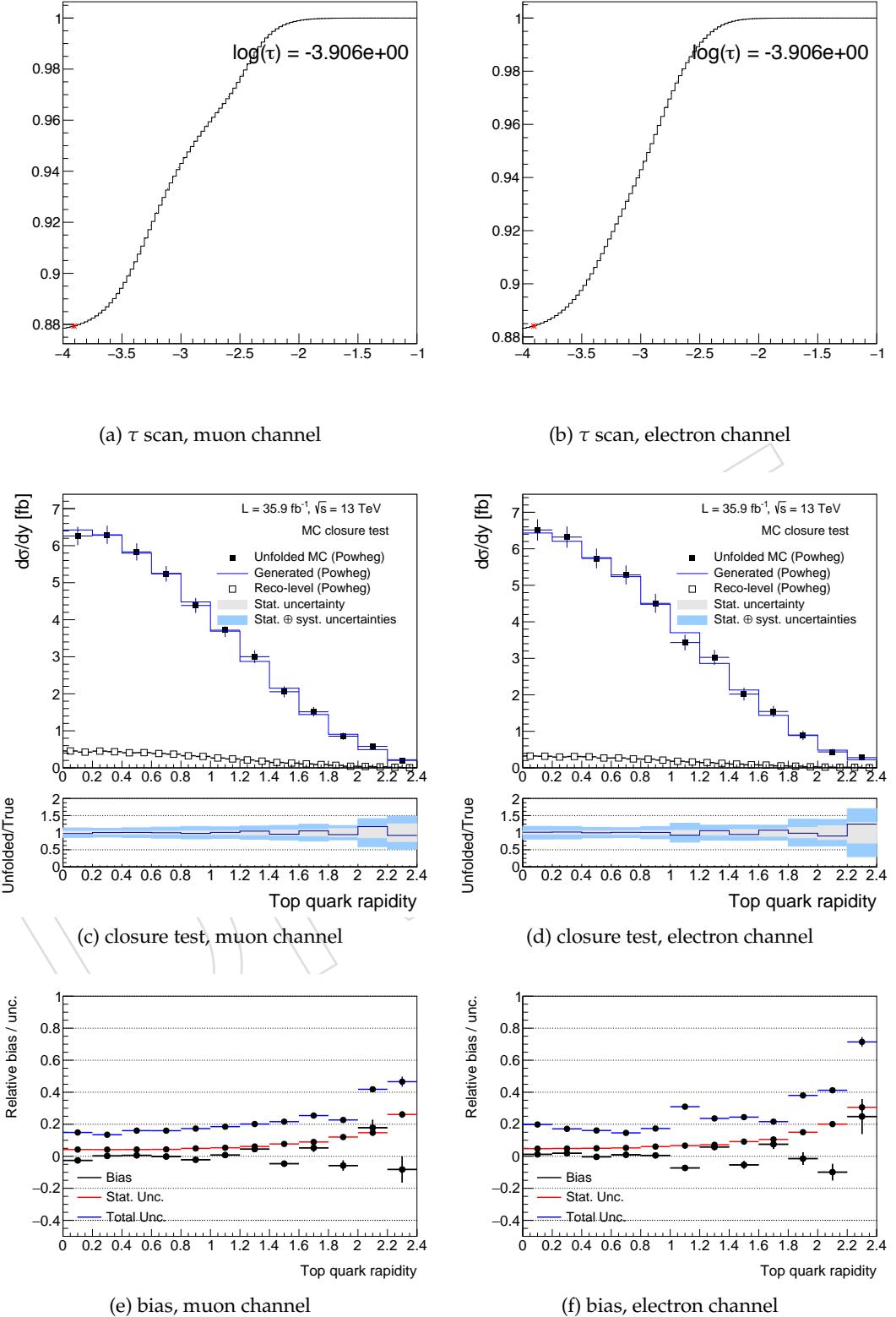


Figure 140: Results for the unfolding using curvature regularization, including systematic uncertainties, using the ScanTau method to determine τ . The top plots show the average global correlation per bin as a function of τ , which is minimized to determine the optimal τ . The middle plots show the results of the closure test unfolding one half-sample with the other. The bottom plots show the average bias, statistical uncertainty, and total uncertainty when unfolding 1000 toys. Results are shown for the muon channel on the left and the electron channel on the right.

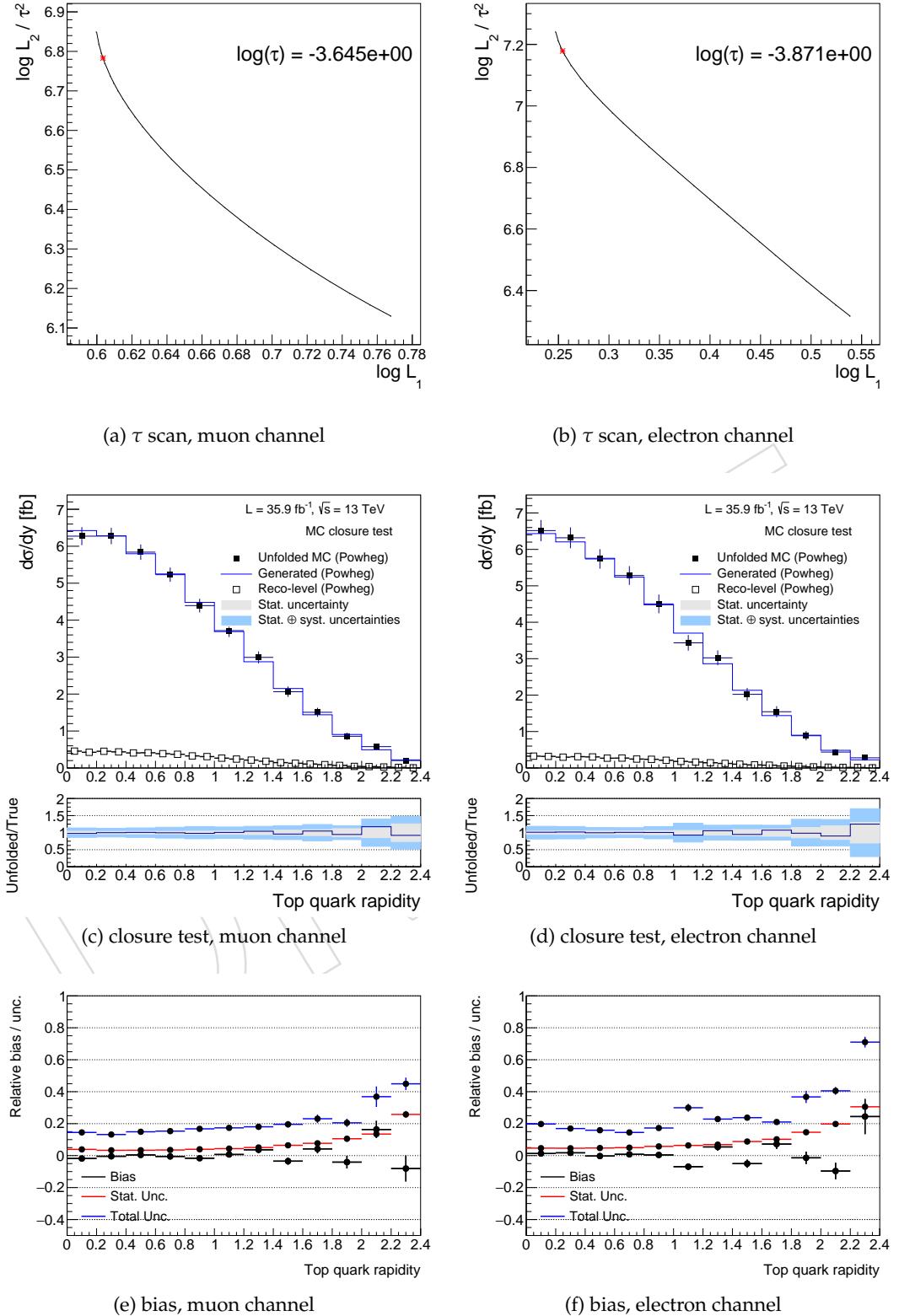


Figure 141: Results for the unfolding using curvature regularization, including systematic uncertainties, using the LCcurve method to determine τ . The top plots show the LCcurve, with the optimal τ determined from the point of maximum curvature. The middle plots show the results of the closure test unfolding one half-sample with the other. The bottom plots show the average bias, statistical uncertainty, and total uncertainty when unfolding 1000 toys. Results are shown for the muon channel on the left and the electron channel on the right.

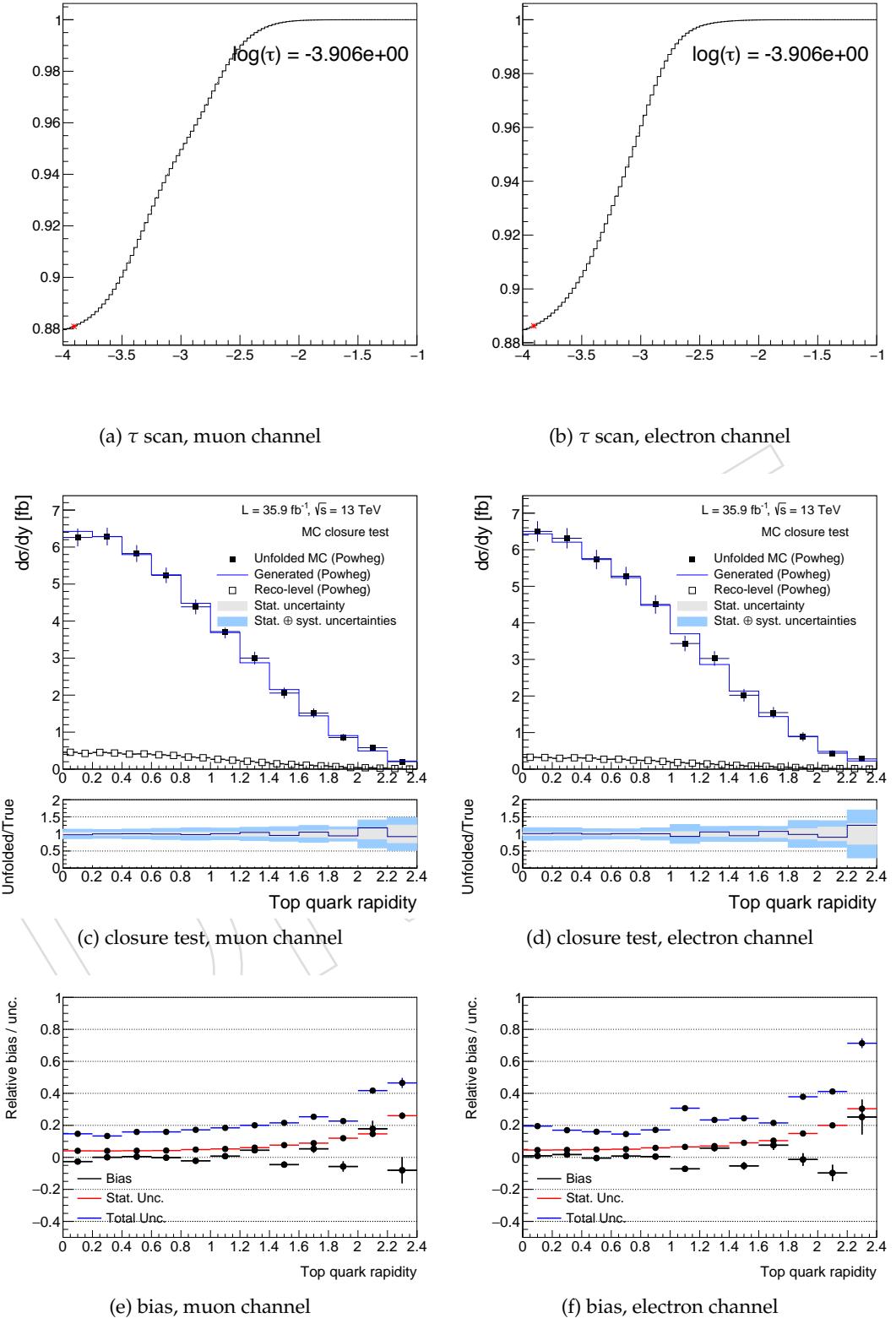


Figure 142: Results for the unfolding using derivative regularization, including systematic uncertainties, using the ScanTau method to determine τ . The top plots show the average global correlation per bin as a function of τ , which is minimized to determine the optimal τ . The middle plots show the results of the closure test unfolding one half-sample with the other. The bottom plots show the average bias, statistical uncertainty, and total uncertainty when unfolding 1000 toys. Results are shown for the muon channel on the left and the electron channel on the right.

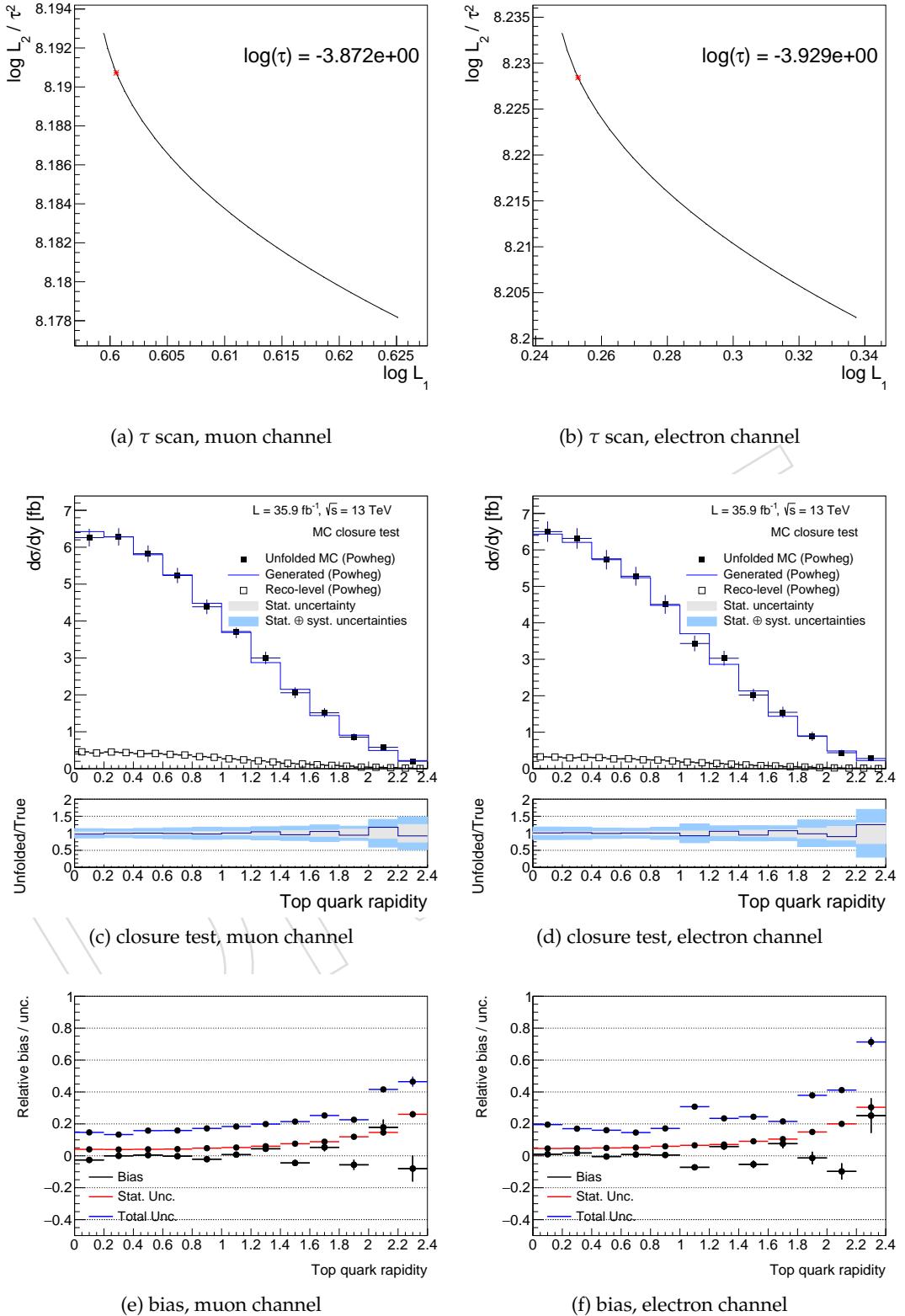


Figure 143: Results for the unfolding using derivative regularization, including systematic uncertainties, using the LCurve method to determine τ . The top plots show the LCurve, with the optimal τ determined from the point of maximum curvature. The middle plots show the results of the closure test unfolding one half-sample with the other. The bottom plots show the average bias, statistical uncertainty, and total uncertainty when unfolding 1000 toys. Results are shown for the muon channel on the left and the electron channel on the right.

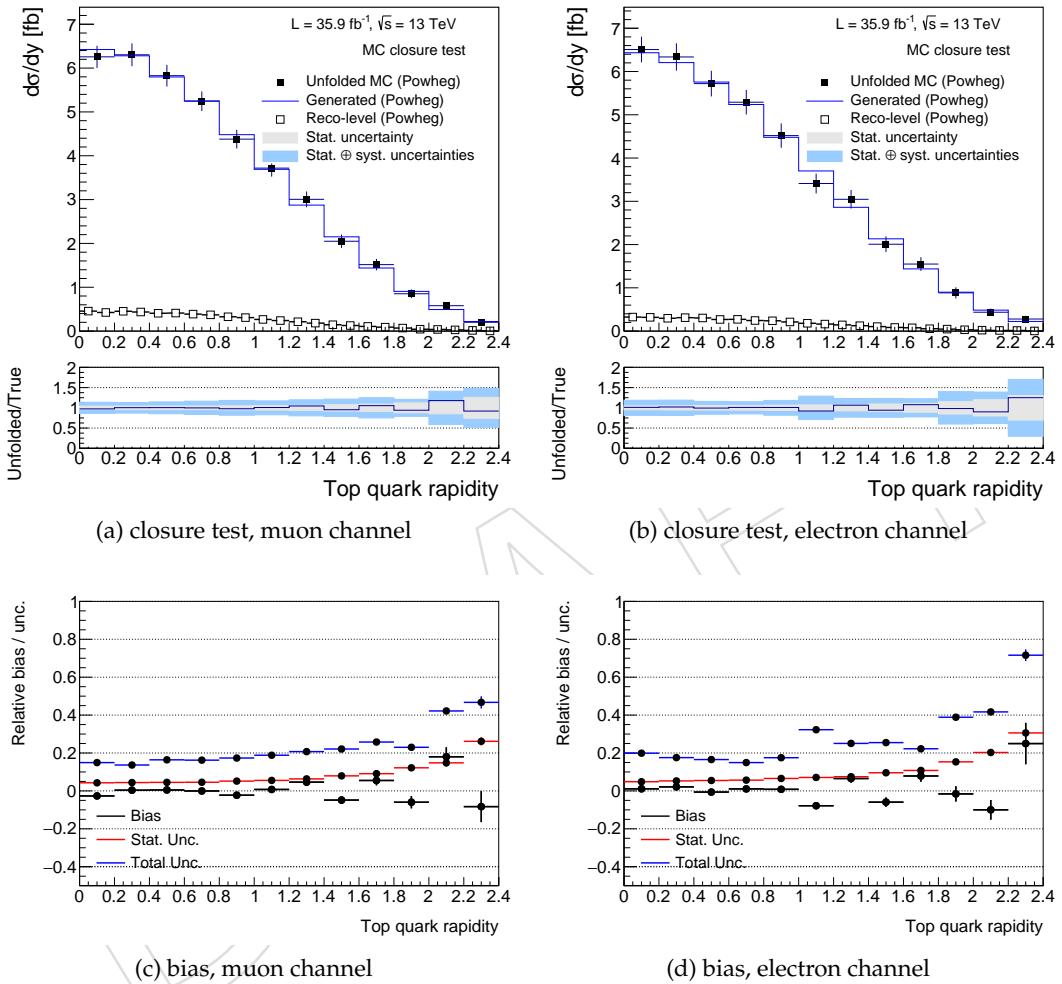


Figure 144: Results for the non-regularized ($\tau = 0$) unfolding. The top plots show the results of the closure test unfolding one half-sample with the other, and the bottom plots show the bias, statistical uncertainty, and total uncertainty from unfolding 1000 toys. Results are shown for the muon channel on the left and the electron channel on the right.

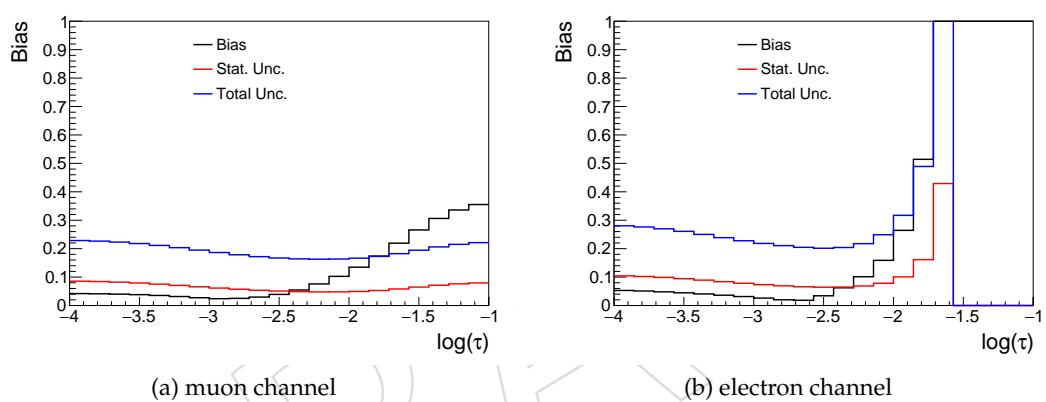


Figure 145: Sum over all bins of relative statistical uncertainty, total uncertainty, and bias as a function of regularization strength τ . The uncertainty and bias appear to be minimized for $\log(\tau) = -2.6$.