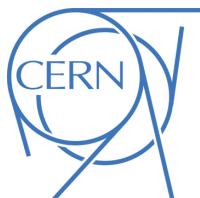




ATLAS NOTE

29th February 2016



Draft version 1.0

Estimation of fake lepton background for top analyses using the Matrix Method with the 2015 dataset at $\sqrt{s} = 13$ TeV with AnalysisTop-2.3.41

F. Derue^a

^a*Laboratoire de Physique Nucléaire et de Hautes Energies de Paris (LPNHE), UPMC, Université Paris-Diderot, CNRS/IN2P3, Institut Lagrange de Paris, Paris, France*

Abstract

This note presents the matrix method for estimating non-prompt and fake lepton backgrounds developed in the context of top analyses using the ATLAS detector. The analysis is performed on the ATLAS 2015 proton-proton collision data sample, collected at the LHC, corresponding to a luminosity of 3.2 fb^{-1} at $\sqrt{s} = 13$ TeV. The matrix method is based on the measurement of efficiencies of leptons with relaxed identification criteria. It uses the software AnalysisTop-2.3.41 and corresponds to the default setup of the ToFakes-00-00-08 package. Final states with lepton+jets and dilepton events are considered. For final states with one lepton, the systematic uncertainties on the fake estimates in the signal region are 90-140% (50-140%) for e +jets (μ +jets) events, depending on the jet and b -jet multiplicity. For final states with two leptons, the systematic uncertainties on the fake estimates in the $e\mu$ opposite-sign signal region are 100-300%.

21 **Contents**

22	1 Introduction	5
23	2 Data and simulation samples	5
24	2.1 Data sample	6
25	2.2 Simulation samples	6
26	3 Software	7
27	4 Object selection	7
28	4.1 Electrons	7
29	4.2 Muons	8
30	4.3 Jets	9
31	4.4 $E_{\text{T}}^{\text{miss}}$	9
32	4.5 Overlap between objects	9
33	5 Event selection	9
34	5.1 Event quality preselection	10
35	5.2 Top quark pair production - lepton+jets	10
36	5.3 Top quark pair production - dilepton	11
37	5.4 Resume of the different selections	11
38	6 Matrix method	13
39	6.1 Overview	13
40	6.2 Definition of tight and loose leptons	14
41	6.3 Measurement and parametrisation of the efficiencies	14
42	6.3.1 Measurement of real efficiencies	15
43	6.3.2 Measurement of fake efficiencies	17
44	6.3.3 Parametrisation of efficiencies	20
45	6.4 Results in $l+jets$ validation regions	25
46	7 Estimates in lepton+jets $t\bar{t}$ events	37
47	7.1 Estimates for $e+jets$ events	37
48	7.2 Estimates for $\mu+jets$ events	41
49	7.3 Summary of the results	45
50	8 Estimates in dilepton $t\bar{t}$ events	47
51	8.1 Measurement and parametrisation of the efficiencies	47
52	8.2 Estimates in the opposite-sign sample	48
53	8.3 Estimates in the same-sign sample	49
54	8.4 Summary of the results	51
55	9 To go further	53
56	10 Conclusion	55

57	Appendix	56
58	A Additional studies on real electron efficiencies	56
59	A.1 Variables used for the parametrisation of the efficiencies	56
60	A.2 Studies with true electrons	56
61	A.2.1 Modeling	56
62	A.2.2 True efficiencies	58
63	A.2.3 Correction factor	59
64	A.3 The tag-and-probe method	60
65	A.3.1 Background subtraction	60
66	A.3.2 Results obtained on Z data	61
67	A.3.3 Results obtained on Z data corrected by Monte Carlo	62
68	A.4 The high- E_T^{miss} method	64
69	A.4.1 Definition of the control region	64
70	A.4.2 Results obtained	65
71	A.5 Comparison of efficiencies from both methods	68
72	A.6 Binning used to measure real electron efficiencies	69
73	B Additional studies on real muon efficiencies	70
74	B.1 Variables used for the parametrisation of the efficiencies	70
75	B.2 Studies with true muons	70
76	B.2.1 Modeling	70
77	B.2.2 True efficiencies	71
78	B.2.3 Correction factor	73
79	B.3 The tag-and-probe method	75
80	B.3.1 Background subtraction	75
81	B.3.2 Results obtained on Z data	75
82	B.4 The high- $m_T(\text{lept}, E_T^{\text{miss}})$ method	78
83	B.4.1 Definition of the control region	78
84	B.4.2 Results obtained	80
85	B.5 Comparison of efficiencies from both methods	87
86	B.6 Binning used to measure real muon efficiencies	90
87	C Additional studies on fake electron efficiencies	91
88	C.1 Definition of the control regions	91
89	C.2 Results obtained	91
90	C.3 Impact of the control region definition	93
91	C.4 Impact of the MC subtraction	95
92	C.5 Impact of using different Z/W+jets simulated data	97
93	C.6 Composition of non-prompt and fake electrons	99
94	C.7 Binning used to measure fake electron efficiencies	100
95	D Additional studies on fake muon efficiencies	101
96	D.1 Definition of the control regions	101
97	D.2 Results obtained	102
98	D.3 Impact of the control region definition	106
99	D.4 Impact of the MC subtraction	112

100	D.5 Impact of using different Z/W+jets simulated data	118
101	D.6 Binning used to measure fake muon efficiencies	124
102	E Additional studies for $l+$jets $t\bar{t}$ events	125
103	E.1 Estimates for $e+$ jets events	125
104	E.1.1 Only fake leptons	125
105	E.1.2 Details on the systematic effects	128
106	E.1.3 Estimates without cuts on E_T^{miss} and $m_T(\text{lept}, E_T^{\text{miss}})$	131
107	E.1.4 Estimates in the signal region	132
108	E.1.5 Estimates in the control regions	136
109	E.1.6 Estimates with other MC samples	139
110	E.2 Estimates for $\mu+$ jets events	141
111	E.2.1 Only fake leptons	141
112	E.2.2 Details on the systematic effects	144
113	E.2.3 Estimates without cuts on E_T^{miss} and $m_T(\text{lept}, E_T^{\text{miss}})$	147
114	E.2.4 Estimates in the signal region	148
115	E.2.5 Estimates in the control regions	152
116	E.2.6 Estimates with other MC samples	155
117	F Additional studies for dilepton $t\bar{t}$ events	157
118	F.1 Estimates for opposite-sign dilepton events	157
119	F.2 Estimates for opposite-sign “not TT” dilepton events	166
120	F.3 Estimates for same-sign dilepton events	166
121	F.4 Estimates for same-sign “not TT” dilepton events	175

1. Introduction

The selection of events with top quarks is often based on the identification of one or more charged isolated leptons coming from the decay of W bosons, referred to as ‘prompt’ or ‘real’ leptons in the following. Acceptance, quality and isolation requirements are applied to select these leptons.

Non-prompt leptons and non-leptonic particles may satisfy these selection criteria, giving rise to so called ‘non-prompt and fake’ lepton backgrounds. In the case of electrons, these include contributions from semileptonic decays of b - and c -quarks, photon conversions and jets with large electromagnetic energy (from the hadronisation to π^0 ’s or from early showering in the calorimeter). Non-prompt or fake muons can originate from semileptonic decays of b - and c -quarks, from charged hadron decays in the tracking volume or in hadronic showers, or from punch-through particles emerging from high-energy hadronic showers. For analyses based on events with one lepton, this background stems from multi-jet events, characterised by a cross-section several orders of magnitude larger than for W boson or top events. In events with two leptons the non-prompt and fake lepton backgrounds are dominated by $W + \text{jets}$ and semileptonic $t\bar{t}$ events, with a fake lepton in addition to the real one, and more rarely events with two fake leptons.

These backgrounds are estimated using data-driven techniques. The most common method is called matrix method and has already been used for Run 1 ATLAS top quark studies [1, 2]. The same method is applied using the software AnalysisTop-2.3.41 and corresponds to the software implementation available in the default setup of the ToFakes-00-00-08 package. Results are presented on typical top selections such as the $t\bar{t}$ semileptonic and the dileptonic selections. The analysis is performed in the ATLAS 2015 proton-proton collision data sample, corresponding to an integrated luminosity of 3.2 fb^{-1} at $\sqrt{s} = 13 \text{ TeV}$.

Section 2.2 presents the simulated samples used in this note. Object reconstruction and identification is presented in section 4 whereas event selection is described in section 5. The matrix method is then described in section 6. In section 7 the results obtained in $t\bar{t}$ semileptonic events are presented. In section 8 estimates in the dilepton channel are provided, obtained either from the matrix method or from a Monte Carlo based method.

2. Data and simulation samples

The ATLAS detector [3] is built from a set of cylindrical subdetectors, which cover almost the full solid angle¹ around the interaction point. It is composed of an inner tracking system close to the interaction point, surrounded by a superconducting solenoid, electromagnetic and hadronic calorimeters, and a muon spectrometer with three superconducting toroid magnetic systems.

¹ ATLAS uses a right-handed coordinate system with its origin at the nominal interaction point (IP) in the centre of the detector and the z -axis along the beam pipe. The x -axis points from the IP to the centre of the LHC ring, and the y -axis points upward. Cylindrical coordinates (r, ϕ) are used in the transverse plane, ϕ being the azimuthal angle around the beam pipe. The pseudorapidity is defined in terms of the polar angle θ as $\eta = -\ln \tan(\theta/2)$ and the transverse momentum p_T is defined as $p_T = p \sin \theta$. The $\eta - \phi$ distance between two particles is defined as $\Delta R = \sqrt{\Delta\eta^2 + \Delta\phi^2}$.

153 2.1. Data sample

154 The analysis is performed using the complete pp collision data recorded at a centre-of-mass energy of \sqrt{s}
 155 =13 TeV with the ATLAS detector between 16th and 3rd November 2015, when the LHC was operating
 156 with 25 ns bunch spacing. Only the periods in which all the sub-detectors were operational are considered
 157 ('AllGood' good run lists or GRL [4]), resulting in a data sample with a total integrated luminosity
 158 3.209 fb⁻¹. The uncertainty on the integrated luminosity is 5.0%. It was estimated using the techniques
 159 described in Ref. [5].

160 2.2. Simulation samples

161 Monte Carlo simulated event samples are used in this analysis, to compare to the data and to evaluate
 162 signal and background efficiencies and uncertainties. Samples were processed either through the full
 163 ATLAS detector simulation [6] based on GEANT4 [7], or through a faster simulation making use of
 164 parameterised showers in the calorimeters [8]. Additional simulated pp collisions generated with Py-
 165 THIA 8.186 [9] were overlaid to model the effects of both in- and out-of-time pileup, from additional pp
 166 collisions in the same and nearby bunch crossings. All simulated events were processed using the same
 167 reconstruction algorithms and analysis chain as the data. Small corrections were applied to lepton trigger
 168 and reconstruction efficiencies to better model the response observed in data.

169 The baseline $t\bar{t}$ full simulation sample (DS 410000) was produced using the next-to-leading-order (NLO)
 170 matrix element generator PowHEG-Box v2 [10–12] using the CT10 set of parton distribution functions [13],
 171 and interfaced to PYTHIA 6.428 [14] for the parton showering and fragmentation. The Perugia 2012
 172 (P2012) [15] parameter set (tune) with the CTEQ6L PDF set [16] was used for the underlying event
 173 (UE) description. All $t\bar{t}$ final states involving at least one lepton were included. The EvtGen v1.2.0 [17]
 174 package was used to model the decays of heavy flavour hadrons. The Powheg model parameter h_{damp} ,
 175 which controls matrix element to parton shower matching in PowHEG and effectively regulates the high-
 176 p_T radiation, was set to the top-quark mass, 172.5 GeV, a setting which was found to best describe the $t\bar{t}$
 177 system p_T at $\sqrt{s} = 7$ TeV [18]. The Standard Model expectation of 0.1082 was assumed for the $W \rightarrow l\nu$
 178 branching ratio [19]. The top-quark mass was set to 172.5 GeV in all the simulation samples.

179 Events originating from $W+jets$ production, in which the W boson decays to a charged lepton and neutrino,
 180 were generated using SHERPA 2.1.1 [20]. Matrix elements were calculated for up to two additional partons
 181 at NLO and four additional partons at LO using the Comix [21] and OpenLoops [22] matrix element
 182 generators and merged with the SHERPA parton shower [23] using the ME+PS@NLO prescription [24].
 183 The CT10 PDF set was used in conjunction with a dedicated parton shower tuning developed by the
 184 SHERPA authors. The $Z+jets$ background was modelled using SHERPA samples, which were generated in an
 185 analogous way to the $W+jets$ samples. Additional $W+jets$ and $Z+jets$ samples are used for cross checks
 186 and additional studies, based either on POWHEG +PYTHIA 8 or MADGRAPH +PYTHIA 8.

187 Diboson production in association with jets was modelled using SHERPA 2.1.1 and CT10 PDFs, including
 188 matrix elements calculated for up to one ($llvv$ and $llll$) or zero ($lllv$) additional partons at NLO and up
 189 to three additional partons at leading order using the Comix and OpenLoops matrix element generators,
 190 merged with the SHERPA parton shower using the ME+PS@NLO prescription. Single top Wt and $t-$
 191 channel production were modelled using POWHEG-Box v2+PYTHIA 6 using the CT10 PDFs and the P2012
 192 UE tune. In order to remove the overlap with $t\bar{t}$ production, the Wt sample was produced using the
 193 'diagram removal' generation scheme [25].

Table 1: List of simulated samples used in this note. For each process, the dataset number is shown, as well as basic generator parameters used to simulate them.

Process	Dataset	Generator	PDF	Tune	Shower	Normalisation	Cross section [pb]
$t\bar{t}$	410000	POWHEG +EVTGEN	CT10	P2012	PYTHIA 6.428	NNLO+NNLL	451.6
single top							
t -channel	410011-12						44.1
s -channel	410025-26	POWHEG +EVTGEN	CT10	P2012	PYTHIA 6.428	NLO+NNLL	3.4
Wt -channel	410013-14						35.85
SHERPA samples							
W +jets							
$W(e\nu) + \text{jets}$	361300-23					NNLO	
$W(\mu\nu) + \text{jets}$	361324-47	SHERPA	CT10				
$W(\tau\nu) + \text{jets}$	361348-71						
Z/γ^*+jets						NNLO	
$Z/\gamma^*(\ell\ell) + \text{jets}, m_{ll} > 60 \text{ GeV}$	361372-395 (ee) 361396-419 ($\mu\mu$) 361420-443 ($\tau\tau$)	SHERPA	CT10			NNLO	
$Z/\gamma^*(\ell\ell) + \text{jets}, 10 < m_{ll} < 40 \text{ GeV}$	361468-475 (ee) 361476-483 ($\mu\mu$) 361484-491 ($\tau\tau$)	SHERPA	CT10			NNLO	
diboson							
WW	361081-82	SHERPA					
WZ	361083-85						
ZZ	361086-87						
POWHEG +PYTHIA 8 samples							
W +jets	361100-05	POWHEG +PYTHIA 8					
Z +jets	361106-08	POWHEG +PYTHIA 8					
MADGRAPH +PYTHIA 8 samples							
W +jets	361520-534	MADGRAPH +PYTHIA 8					
Z +jets	361500-514 361628-642	MADGRAPH +PYTHIA 8 MADGRAPH +PYTHIA 8					

194 3. Software

195 This analysis is done using AnalysisTop-2.3.41. The TopFakesUtils package, version TopFakesUtils-00-
 196 00-08, is used to create flat tuples from the TOPQ1 derivation. The efficiencies derived in this analysis are
 197 stored in the TopFakes package, version TopFakes-00-00-08, which is used to get the fake estimates.

198 No W +jets SF are applied in this analysis.

199 4. Object selection

200 All definitions of these objects will be discussed in the following and follow the recommendation of the
 201 ATLAS top working group for the $\sqrt{s} = 13$ TeV analyses.

202 4.1. Electrons

203 Electron candidates [26] are reconstructed from isolated electromagnetic calorimeter energy deposits
 204 matched to inner detector tracks and passing identification requirements. They are selected as follows:

- Candidates are selected with $|\eta_{\text{cl}}| < 2.47$ (where η_{cl} is the pseudorapidity of the calorimeter cluster associated with the electron candidate). Those within the transition region between the barrel and end-cap electromagnetic calorimeters, $1.37 < |\eta_{\text{cl}}| < 1.52$, are removed.
 - A cut on transverse energy $E_{\text{T}} > 25$ GeV is set, with $E_{\text{T}} = E / \cosh(\eta)$, where the energy is taken from the cluster E_{cl} and the direction from the associated track η_{track} .
 - Requirements on the transverse impact parameter significance of $|d_0/\sigma_{d_0}| < 3$ and on the longitudinal impact parameter relative to the primary vertex² $|\Delta z_0 \sin(\theta)| < 0.5$ mm are imposed.
 - Electrons are required to be isolated using requirements on the energy of calorimeter topological clusters in a cone of $\Delta R < 0.2$ around the electron (excluding the deposit associated to the electron) divided by the electron p_{T} , and on the sum of track p_{T} in a variable-sized cone around the electron direction (again excluding the track associated to the electron). The track isolation cone radius is given by the smaller of $\Delta R = 10 \text{ GeV}/p_{\text{T}}(e)$, where $p_{\text{T}}(e)$ is the p_{T} of the electron, and $\Delta R = 0.2$, *i.e.* a cone which increases in size at low p_{T} up to a maximum of 0.2 radians. Selection criteria (named ‘gradient isolation’), dependent on p_{T} and η , are designed to produce a nominal efficiency of 90% for electrons from $Z \rightarrow ee$ decays with p_{T} of 25 GeV which rises to 99% at 60 GeV. The efficiencies in $t\bar{t}$ events are somewhat smaller, due to the increased jet activity.
- Electron candidates passing tight likelihood-based [26] selection criteria and the isolation requirements are referred to as ‘tight electrons’. Loose electrons are electrons satisfying medium likelihood-based selection criteria and no requests on the isolation are made.

4.2. Muons

Muon candidates [27] are reconstructed by combining matching tracks reconstructed in both the inner detector and muon spectrometer. They are selected as follows:

- Candidates are selected with $|\eta| < 2.4$.
- A cut of $p_{\text{T}} > 25$ GeV is set to be on the plateau of the single muon trigger efficiency.
- Requirements on the transverse impact parameter significance of $|d_0/\sigma_{d_0}| < 3$ and on the longitudinal impact parameter relative to the primary vertex $|\Delta z_0 \sin(\theta)| < 0.5$ mm are imposed.
- Muons must pass the MCP ID track quality cuts.
- Muons are also required to be isolated, using the same variables as for electrons, with the selection criteria tuned to give similar efficiencies on $Z \rightarrow \mu\mu$ events.

These muons are referred to as ‘tight muons’. For loose muons, no request on the isolation is made but all other selection requirements are applied.

² The primary vertex is defined as the reconstructed vertex with the highest sum of associated track p_{T}^2 .

236 4.3. Jets

237 Jets are reconstructed using the anti- k_t algorithm [28, 29] with radius parameter $R = 0.4$, starting from
 238 topological clusters in the calorimeter [30]. The effects of pileup on jet energies are accounted for by
 239 a jet-area-based correction [31] and the resolution of the jets is improved by using global sequential
 240 corrections [32]. Jets are then calibrated to the hadronic energy scale using E- and η -dependent calibration
 241 factors based on MC simulations, with in-situ corrections based on Run 1 data [33, 34] and checked with
 242 early Run 2 data [35]. Corrections for semileptonic b -hadron decays are not applied. Jets are accepted
 243 within the fiducial region $p_T > 25$ GeV and $|\eta| < 2.5$. To reduce the contribution from jets associated
 244 with pileup, jets with $p_T < 50$ GeV and $|\eta| < 2.4$ are required to pass pileup rejection criteria [36].
 245 Reconstructed jets within $R < 0.2$ of a selected electron are removed, as discussed earlier.

246 Jets are b -tagged as likely to contain b hadrons using the MV2c20 algorithm, a multivariate discrim-
 247 inant making use of the long lifetime, high decay multiplicity, hard fragmentation and high mass of b
 248 hadrons [37–39]. Jets are defined as being b -tagged if the MV2c20 weight is larger than a cut value
 249 corresponding to approximately 70% b -tagging efficiency for b -jets in $t\bar{t}$ events, although the exact effi-
 250 ciency varies with p_T . In simulation, the tagging algorithm gives a rejection factor of about 440 against
 251 light-quark and gluon jets, and about eight against jets originating from charm quarks.

252 4.4. E_T^{miss}

253 The missing transverse momentum (E_T^{miss}) is defined as the negative of the global vector sum p_T of all
 254 selected physics objects (electrons, muons, jets) as well as specific “soft terms” accounting for unclassi-
 255 fied soft tracks and calorimeter clusters. In this way, the missing transverse momentum is adjusted to take
 256 into account the best calibration of the identified physics objects above [40].

257 4.5. Overlap between objects

258 The overlap between objects is resolved in the following order: (1) if an electron and a muon fulfilling the
 259 respective object selections share the same track (matching is done within $\theta = 0.0005$ and $\varphi = 0.0005$),
 260 the event is rejected. Next, (2), remove muons within jets, by requiring $\Delta(\mu, j) < 0.4$, and then, (3),
 261 remove the jet closest to a good electron within $\Delta R(e, j) < 0.2$. Finally, (4), electrons close to the
 262 remaining jets within $\Delta R < 0.4$ are removed from the event. When amount of fakes is estimated using
 263 the matrix method (see later on section 6), the analysis will be done with a looser definition of leptons
 264 (electrons and muons). In this case, the overlap removal procedure is applied in the same way but using
 265 these loose leptons.

266 5. Event selection

267 Events are required to pass either a single electron or single muon trigger, with thresholds that are fully
 268 efficient for leptons passing offline selections of $p_T > 25$ GeV.

269 For electrons, the OR of the HLT_E24_LHMEDIUM_L1EM20VH, HLT_E60_LHMEDIUM and HLT_E120_LHLOOSE
 270 trigger chains is used. For Monte Carlo, HLT_E24_LHMEDIUM_L1EM18VH has to be used instead of

271 HLT_E24_LHMEDIUM_L1EM20VH since the latter is not included in the simulation. This difference will
 272 be accounted for in the trigger scale factors and their corresponding uncertainties. For muons, the OR
 273 of the HLT_MU20_ILOOSE_L1MU15 and HLT_MU50 trigger chains is used. The p_T thresholds are 24, 60
 274 and 120 GeV for electrons (labelled e24, e60 and e120) and 20 or 50 GeV for muons (labelled mu20i
 275 and mu50). For muons, the triggers with the lower p_T threshold include isolation requirements on the
 276 candidate lepton. The trigger uses criteria that are looser than those used in analysis lepton selection, but
 277 are similar in nature. Additional muon pre-scaled trigger without isolation requirements (mu20) are also
 278 used. The “i” in the name of the trigger indicates the requirement of lepton isolation.

279 The events selected to study top quark pair and single top production in the lepton+jets and dilepton
 280 channels consist of one or two leptons (electrons or muons), a number of jets, including b -jets, and a
 281 significant amount of missing transverse energy. They serve as reference for evaluating the fake lepton
 282 estimates throughout the note.

283 5.1. Event quality preselection

284 To ensure that the selected event corresponds to a good collision event, the following selection is applied
 285 in all cases:

- 286 - In data, the event has to be selected by the appropriate GRL (see sec. 2.1).
- 287 - The event passes the (single lepton) trigger of choice and the lepton (l+jets, single top) or one of
 288 the leptons (dilepton) should match the trigger object.
- 289 - Reject events with jets classified as *bad*.
- 290 - The first primary vertex has to be of type *primary vertex* or *PileUp* and has to have at least four
 291 tracks with $p_T > 400$ MeV associated.

292 5.2. Top quark pair production - lepton+jets

293 To select events in the lepton+jets channel, the following logic is typically applied:

- 294 - Presence of exactly 1 selected charged lepton.
- 295 - presence of ≥ 4 selected jets.
- 296 - To further reduce the background from fake leptons: $E_T^{\text{miss}} > 30$ GeV and $m_T(\text{lept}, E_T^{\text{miss}}) >$
 297 30 GeV^3 .
- 298 - If one or more b -tagged jets are required, the MV2c20 algorithm at 70% efficiency is used.

³ The quantity $m_T(\text{lept}, E_T^{\text{miss}})$ is built as the invariant mass of the $\ell-\nu$ pair, neglecting the z component: $m_T(\text{lept}, E_T^{\text{miss}}) = \sqrt{2p_T^\ell p_T^\nu (1 - \cos(\phi^\ell ll - \phi^\nu))}$.

²⁹⁹ **5.3. Top quark pair production - dilepton**

³⁰⁰ Events in the dilepton channel are selected following the logic below:

- ³⁰¹ - To avoid overlap between samples, a truth selection of dilepton events is applied for MC simulated events.
- ³⁰³ - Before the event cleaning and overlap removal, events with cosmic muons are rejected by removing events with a pair of muons with opposite sign d_0 , $|d_0| > 0.5$ mm and $\Delta\varphi(\mu_1, \mu_2) > 3.1$.
- ³⁰⁵ - Exactly two charged leptons, with opposite sign (OS).
- ³⁰⁶ - At least two selected jets.
- ³⁰⁷ - In the ee and $\mu\mu$ channels $E_T^{\text{miss}} > 60$ GeV.
- ³⁰⁸ - In the $e\mu$ channel $H_T > 130$ GeV, with H_T being the sum of transverse momenta of all selected objects, including missing transverse energy.
- ³¹⁰ - In the ee and $\mu\mu$ channel to suppress low mass resonances and Z boson production, $m_{ll} > 15$ GeV and $|m_{ll} - 90.1| > 10$ GeV are required.
- ³¹² - b -tagged jets can be required using the MV2c20 tagger at 70% efficiency.

³¹³ **5.4. Resume of the different selections**

³¹⁴ For each of the considered lepton+jets or dilepton channels, different regions are defined by the requirements summarised in Table 2.

Table 2: Summary of the different regions used in this analysis. The term ‘pretag’ is used to indicate that no requirements on the number of b -jets are applied, while ‘OS’ (‘SS’) stands for opposite- (same-) sign charged leptons. MM stands for matrix method and FM for fitting method.

Channel	$n_{\text{jet}}/n_{b\text{-jet}}$ cuts	Other cuts	Used for
$l + \text{jets}$	≥ 1 jet, pretag	1) $m_T(\text{lept}, E_T^{\text{miss}}) < 20 \text{ GeV} \&$ $E_T^{\text{miss}} + m_T(\text{lept}, E_T^{\text{miss}}) < 60 \text{ GeV}$	$\varepsilon_{\text{fake}}$ extraction
	2 jets, pretag	2) $E_T^{\text{miss}} < 20 \text{ GeV}$	$\varepsilon_{\text{fake}}$ extraction
	2 jets, ≥ 1 b -tags	3) no cuts on $E_T^{\text{miss}}, m_T(\text{lept}, E_T^{\text{miss}})$	control region
	≥ 4 jets, pretag		
	≥ 4 jets, ≥ 1 b -tags		
	2 jets, pretag	$E_T^{\text{miss}} > 30 \text{ GeV}, m_T(\text{lept}, E_T^{\text{miss}}) > 30 \text{ GeV}$	results ($W + \text{jets}$ control region)
ee	2 jets, ≥ 1 b -tags		Results ($t\bar{t}$ signal region)
	≥ 4 jets, pretag		
	≥ 4 jets, ≥ 1 b -tags		
	≥ 1 jet, pretag	OS, $80 \text{ GeV} < m_{ee} < 100 \text{ GeV}$	$\varepsilon_{\text{real}}$ extraction
	≥ 1 jet, pretag	OS, $80 \text{ GeV} < m_{\mu\mu} < 100 \text{ GeV}$	$\varepsilon_{\text{real}}$ extraction
$\mu\mu$	≥ 2 jets, pretag	SS, $H_T > 100 \text{ GeV}$	control region
	≥ 2 jets, ≥ 1 b -tags		
	≥ 2 jets, pretag	OS, !TT, $H_T > 100 \text{ GeV}$	results ($t\bar{t}$ signal region, fake-enriched)
$e\mu$	≥ 2 jets, ≥ 1 b -tags		

316 6. Matrix method

317 The matrix method was already applied for the Run 1 top quark analyses in ATLAS [1].

318 6.1. Overview

319 In a data sample containing events with a single lepton, the number of events with one tight lepton (N^{tight})
 320 and the number of events with one loose lepton (N^{loose}) can be expressed as linear combinations of the
 321 number of events with a real or a non-prompt or fake lepton:

$$\begin{aligned} N^{\text{loose}} &= N_{\text{real}}^{\text{loose}} + N_{\text{fake}}^{\text{loose}}, \\ N^{\text{tight}} &= \varepsilon_{\text{real}} N_{\text{real}}^{\text{loose}} + \varepsilon_{\text{fake}} N_{\text{fake}}^{\text{loose}}, \end{aligned} \quad (1)$$

322 where $\varepsilon_{\text{real}}$ is the fraction of real leptons in the loose selection that also pass the tight one and $\varepsilon_{\text{fake}}$ is
 323 the fraction of non-prompt and fake lepton backgrounds in the loose selection that also pass the tight
 324 selection.

325 If $\varepsilon_{\text{real}}$ and $\varepsilon_{\text{fake}}$ are known, the number of events with a non-prompt or fake lepton can be calculated from
 326 Eq. 1 given the measured N^{loose} and N^{tight} . The relative efficiencies $\varepsilon_{\text{real}}$ and $\varepsilon_{\text{fake}}$ are measured in data in
 327 control samples enriched in either real or non-prompt or fake lepton. The number of tight events coming
 328 from non-prompt or fake lepton backgrounds can be expressed as:

$$N_{\text{fake}}^{\text{tight}} = \frac{\varepsilon_{\text{fake}}}{\varepsilon_{\text{real}} - \varepsilon_{\text{fake}}} (\varepsilon_{\text{real}} N^{\text{loose}} - N^{\text{tight}}). \quad (2)$$

329 The matrix method efficiencies $\varepsilon_{\text{real}}$ and $\varepsilon_{\text{fake}}$ depend on lepton kinematics and event characteristics, such
 330 as and the number of jets or b -jets. To correctly account for this, an event weight is computed from the
 331 efficiencies, which are parametrised as a function of the various object kinematics (as detailed Section ??):
 332

$$w_i = \frac{\varepsilon_{\text{fake}}}{\varepsilon_{\text{real}} - \varepsilon_{\text{fake}}} (\varepsilon_{\text{real}} - \delta_i), \quad (3)$$

333 where δ_i equals unity if the loose event i passes the tight event selection and 0 otherwise. The background
 334 estimate in a given bin of the final observable is given by the sum of w_i over all events in that bin.

335 In the case of a dilepton selection, the numbers of observed events with two tight leptons (denoted as
 336 N_{tt}), one loose and one tight lepton (N_{tl} and N_{lt}) or two loose leptons (N_{ll}) are counted. Here and in what
 337 follows, the leptons are ordered by p_T in the indexes, such that the leading lepton in N_{tl} region is tight
 338 and the leading lepton in N_{lt} is loose. Using $\varepsilon_{\text{real}}$ and $\varepsilon_{\text{fake}}$, already defined for the single lepton case,
 339 linear equations are obtained for the observed yields as a function on the number of events with zero, one
 340 and two real leptons together with two, one and zero non-prompt or fake leptons (N_{ff} , N_{rf} , N_{fr} and N_{ff}
 341 respectively):

$$\begin{pmatrix} N_{\text{rr}} \\ N_{\text{fr}} \\ N_{\text{rf}} \\ N_{\text{ff}} \end{pmatrix} = \mathbf{M}^{-1} \begin{pmatrix} N_{\text{tt}} \\ N_{\text{tl}} \\ N_{\text{lt}} \\ N_{\text{ll}} \end{pmatrix}, \quad (4)$$

342 where \mathbf{M} is a 4×4 matrix written in terms of $\varepsilon_{\text{real}}$ and $\varepsilon_{\text{fake}}$. It is calculated as:

$$\mathbf{M} = \begin{pmatrix} \varepsilon_{r,1} \varepsilon_{r,2} & \varepsilon_{r,1} \varepsilon_{f,2} & \varepsilon_{f,1} \varepsilon_{r,2} & \varepsilon_{f,1} \varepsilon_{f,2} \\ \varepsilon_{r,1} \bar{\varepsilon}_{r,2} & \varepsilon_{r,1} \bar{\varepsilon}_{f,2} & \varepsilon_{f,1} \bar{\varepsilon}_{r,2} & \varepsilon_{f,1} \bar{\varepsilon}_{f,2} \\ \bar{\varepsilon}_{r,1} \varepsilon_{r,2} & \bar{\varepsilon}_{r,1} \varepsilon_{f,2} & \bar{\varepsilon}_{f,1} \varepsilon_{r,2} & \bar{\varepsilon}_{f,1} \varepsilon_{f,2} \\ \bar{\varepsilon}_{r,1} \bar{\varepsilon}_{r,2} & \bar{\varepsilon}_{r,1} \bar{\varepsilon}_{f,2} & \bar{\varepsilon}_{f,1} \bar{\varepsilon}_{r,2} & \bar{\varepsilon}_{f,1} \bar{\varepsilon}_{f,2} \end{pmatrix}, \quad (5)$$

343 where the index on $\varepsilon_{\text{real}}$ and $\varepsilon_{\text{fake}}$ refers to the first (1) or second (2) lepton in the event, and $\bar{\varepsilon}$ stands for
 344 $(1 - \varepsilon)$. Similarly to the single lepton case, four weights, w_{rr} , w_{rf} , w_{fr} and w_{ff} are calculated on event-
 345 by-event basis. The probability that an event with two loose leptons contains at least one non-prompt or
 346 fake lepton is then given by $w_{\text{rf}} + w_{\text{fr}} + w_{\text{ff}}$. Finally, the estimated background contribution in a sample of
 347 events with two tight leptons is given by the event weight:

$$w_{\text{tt}} = \varepsilon_{r,1} \varepsilon_{f,2} w_{\text{rf}} + \varepsilon_{f,1} \varepsilon_{r,2} w_{\text{fr}} + \varepsilon_{f,1} \varepsilon_{f,2} w_{\text{ff}}. \quad (6)$$

348 6.2. Definition of tight and loose leptons

349 As discussed above, the tight sample is built by using the standard selection for electrons and muons,
 350 described in section 4.

351 The definition of the loose sample is an important ingredient in the success of the matrix method. Indeed,
 352 in Eq. 2 a factor $1/(\varepsilon_{\text{real}} - \varepsilon_{\text{fake}})$ enters in the calculation for single lepton case. Thus, for an object, $\varepsilon_{\text{real}}$
 353 and $\varepsilon_{\text{fake}}$ must be sufficiently different, otherwise the denominator in the weight expressions will be very
 354 small and the weights will be large.

355 Loose electrons are electrons satisfying medium likelihood-based selection criteria and no requests on
 356 the isolation are made. Given this loose electron definition, a loose $e+\text{jets}$ event is defined by requiring
 357 exactly one loose electron. The tight sample is defined considering the subset of these events where
 358 the single selected electron is also a tight electron⁴. For loose $e+\text{jets}$ events, the $E_{\text{T}}^{\text{miss}}$ definition (see
 359 section 4.4) is changed in order to consider medium likelihood-based quality electrons in the calibration
 360 (otherwise they would have been treated as jets).

361 Loose muons are defined as the tight ones (see section 4.2), except that the isolation requirement is
 362 dropped. Given this loose muon definition, a loose $\mu+\text{jets}$ event is defined by requiring exactly one loose
 363 muon. The tight sample is defined considering the subset of these events where the single selected muon
 364 is also a tight muon.

365 6.3. Measurement and parametrisation of the efficiencies

366 Real and fake efficiencies $\varepsilon_{\text{real}}$ and $\varepsilon_{\text{fake}}$ are measured in control regions which are representative of the
 367 signal regions in terms of kinematics and, in the case of the fake efficiency, non-prompt and fake lepton
 368 background composition. Table 3 summarises the definition of the different control regions used to extract
 369 the real and fake efficiencies, as explained in the following.

⁴ Note that this implies that the tight $e+\text{jets}$ sample is not exactly the same as the nominal $e+\text{jets}$ selection, where instead a single tight electron is required, regardless of the number of additional loose non-tight electrons. The consequence of this approximation is that the matrix method in the $e+\text{jets}$ channel will estimate the fake leptons background neglecting events where two fake electrons are reconstructed, one passing the tight selection and the second one only the loose. The amount of neglected background has been evaluated to be negligible.

Table 3: Summary of the different control regions used to extract the matrix method efficiencies. The term ‘pretag’ is used to indicate that no requirements on the number of b -jets are applied, while ‘OS’ stays for opposite-sign charge leptons.

Channel	$n_{\text{jet}}/n_{b\text{-jet}}$ cuts	Other cuts	Used for
$l+\text{jets}$	≥ 1 jet, pretag	$m_T(\text{lept}, E_T^{\text{miss}}) < 20 \text{ GeV}$, $E_T^{\text{miss}} + m_T(\text{lept}, E_T^{\text{miss}}) < 60 \text{ GeV}$	$\varepsilon_{\text{fake}}(l)$ extraction
ll	≥ 1 jet, pretag	OS, $80 \text{ GeV} < m_{ll} < 100 \text{ GeV}$	$\varepsilon_{\text{real}}(l)$ extraction

Not reviewed, for internal circulation only

370 6.3.1. Measurement of real efficiencies

371 The real efficiencies $\varepsilon_{\text{real}}$ are measured using the tag-and-probe method from the $Z \rightarrow ee$ and $Z \rightarrow \mu\mu$
 372 control regions. This method selects an unbiased sample of loose leptons (probes) from the Z boson
 373 decay by using a tight selection requirement on the other object produced from the particle’s decay (tags).
 374 The efficiency is determined by applying the tight selection to the probe lepton.

375 Additional selection criteria on the event properties are applied to further reject background: only events
 376 passing data-quality criteria are considered. The tag-and-probe pairs must also pass requirements on their
 377 reconstructed invariant mass. In order to not bias the selected probe sample, each valid combination of
 378 electron (muon) pairs in the event is considered; an electron (muon) can be the tag in one pair and the
 379 probe in another. Selection of events for the tag-and-probe method can be summarised as:

- 380 - Event quality preselection as described in section 5.1.
- 381 - Typical dilepton selection as described in section 5.3.
- 382 - At least one jet in the event.
- 383 - At least two electrons or two muons with for each pair:
 - 384 - A tight electron (muon), matched to the trigger elements firing the data-taking, is used as tag.
 - 385 - A loose electron (muon) used as probe.
 - 386 - The invariant mass of the pair is at least 50 GeV.

387 A total of nearly 3×10^5 electron (2×10^5 muon) probes is finally available. For each pair of leptons, the
 388 invariant mass is calculated and shown in Fig. 1.

389 After this selection, the sample still contains non-prompt and fake lepton backgrounds. The invariant
 390 mass of the tag-and-probe pair is used to discriminate signal electrons (muons) against background. The
 391 most important contribution to the background, even after requiring the tag-and-probe pair to be have
 392 opposite-sign (OS) charges, comes from random combinations of two particles which do not originate
 393 from a resonance decay.

394 The best method, in terms of simplicity and stability, is the side-band method, and is used to extract $\varepsilon_{\text{real}}$
 395 in each considered bin. The method relies on the background having a linear shape over the considered
 396 invariant mass range. This is the case in particular as we consider leptons with $p_T > 25 \text{ GeV}$. Other back-
 397 ground subtraction methods, including non linear shape for the background are reported in sections A.3.1
 398 and B.3.1 and all give similar results. The invariant mass distributions for opposite-sign and same-sign
 399 pairs at the denominator and numerator levels are divided in three regions A, B and C. The number of

Not reviewed, for internal circulation only

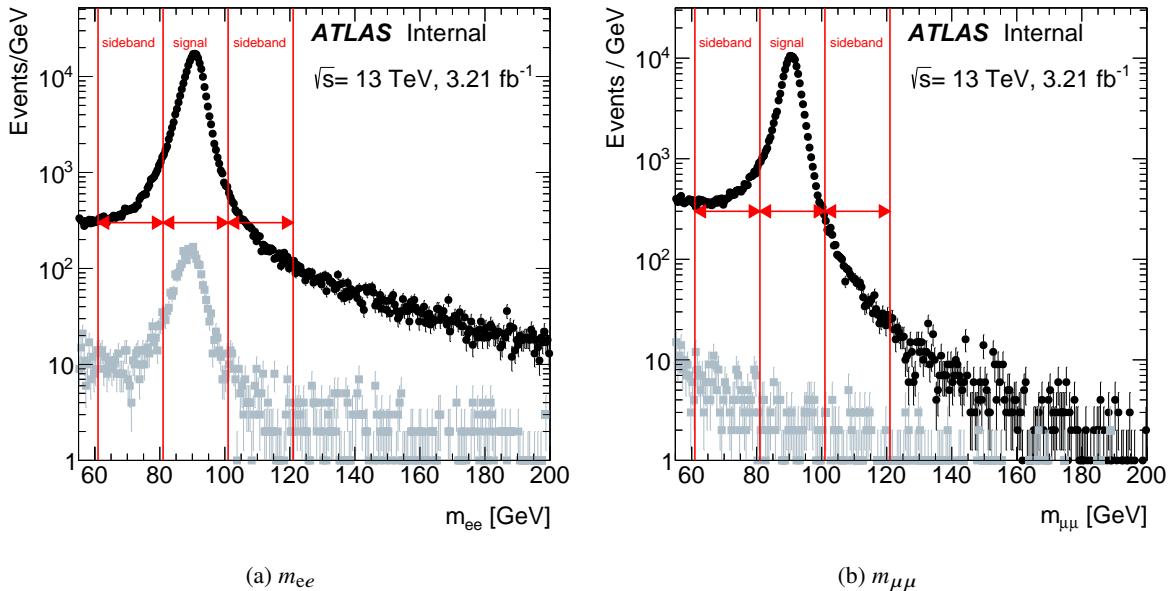


Figure 1: Distribution of the invariant mass of opposite- (black) and same-(grey) sign charge loose (a) electron and (b) muon pairs. Lines show the signal and sideband regions where the yields are calculated.

background events in region B and its uncertainty are estimated from the extrapolation of the side-bands A and C of the same-sign distribution.

402 This background subtraction is performed with three different side-band region definitions: 1) A=[61-
 403 81], B=[81-101], C=[101-121], 2) A=[66-76], B=[76-106], C=[106-116], 3) A=[61-81], B=[86-96],
 404 C=[101-121]. The central value of $\varepsilon_{\text{real}}$ in each considered bin is then determined as the mean value
 405 obtained in the three cases. The spread is also considered as potential intrinsic uncertainty on the meas-
 406 urement, but is in the end neglected, being considerably smaller than the difference between the tag-and-
 407 probe method and the alternative method results.

408 The electron and muon real efficiencies are also studied in MC simulation, as shown in sections A.2 and
 409 B.2. In case of electrons, a non-negligible difference between the different MC processes is observed,
 410 especially between $Z + \text{jets}$ and $t\bar{t}$ events. Since with the tag-and-probe method the $\varepsilon_{\text{real}}$ is extracted from
 411 $Z + \text{jets}$ events and then assumed to be valid for any real lepton process, a correction ρ is derived from
 412 MC simulation and then applied to the measured efficiency $\varepsilon_{\text{real}}$:

$$\varepsilon'_{\text{real}} = \rho \cdot \varepsilon_{\text{real}} = \left| \frac{\varepsilon_{\text{real}}(\text{av.})}{\varepsilon_{\text{real}}(Z+\text{jets})} \right|_{MC} \cdot \varepsilon_{\text{real}}, \quad (7)$$

where $\varepsilon_{\text{real}}(Z+\text{jets})$ is the real efficiency from MC simulation in $Z + \text{jets}$ events and $\varepsilon_{\text{real}}(\text{av.})$ is the average real efficiency between the two extremes $Z + \text{jets}$ and $t\bar{t}$. In this way, a sort of extrapolation from a $Z + \text{jets}$ region to an ‘unknown real lepton source composition’ region is performed. This correction is derived and applied separately in every bin of each of the considered distributions (see below). See section A.3 for more details on the correction factor. In case of muons, no correction is instead applied to the measured efficiency, since the difference between $Z + \text{jets}$ and other real lepton processes such as $t\bar{t}$ is

419 smaller and properly modeled by considering the various parametrisations used in the application of the
 420 efficiencies (see below).

421 The electron and muon real efficiencies have been also derived with the alternative method of selection
 422 $e+jets$ events with high E_T^{miss} or $\mu+jets$ events with high $m_T(\text{lept}, E_T^{\text{miss}})$. Results are presented in sec-
 423 tions A.4 and B.4 and are compared to the tag-and-probe method results in sections A.5 and B.5. Results
 424 are in good agreement and the small observed differences are used to estimate an associated systematic
 425 uncertainty, labelled in the following as “CRreal”.

426 6.3.2. Measurement of fake efficiencies

427 The fake efficiencies $\varepsilon_{\text{fake}}$ are measured in data samples dominated by non-prompt and fake lepton back-
 428 ground events. Selection of events can be summarised as :

- 429 - Event quality preselection as described in section 5.1.
- 430 - At least one jet in the event.
- 431 - Exactly one loose electron or muon.

432 These control regions, denoted CR_f, contain only one loose lepton, at least one jet and have low E_T^{miss}
 433 and/or $m_T(\text{lept}, E_T^{\text{miss}})$ or high lepton impact parameter. For electrons and muons, CR₁ is defined as a
 434 low E_T^{miss} and/or low $m_T(\text{lept}, E_T^{\text{miss}})$ region, by $m_T(\text{lept}, E_T^{\text{miss}}) < 20 \text{ GeV} \& m_T(\text{lept}, E_T^{\text{miss}}) + E_T^{\text{miss}} <$
 435 60 GeV ⁵. Fig. 2 and 3 show the distributions for electrons and muons of the missing transverse energy
 436 and $m_T(\text{lept}, E_T^{\text{miss}})$ for the data and the simulated samples. As can be seen in the plots and in Table 4,
 437 the contamination from real electrons from W and Z decays is about 11% (32%) at loose (tight) level.
 438 For muons, this contamination is 9% (17%) at loose (tight) level for the low- p_T trigger but reaches 44%
 (52%) at loose (tight) level for the high- p_T one. In both channels, to get a pure non-prompt and fake

Selection	Electron		Muon mu20i		Muon mu50	
	W (%)	Z (%)	W (%)	Z (%)	W (%)	Z (%)
Tight	18.9	13.8	11.9	5.3	44.3	8.2
Loose	4.5	6.2	6.5	2.9	37.7	7.0

Table 4: Contamination (in %) of real leptons in the control region defined by $m_T(\text{lept}, E_T^{\text{miss}}) < 20 \text{ GeV} \& m_T(\text{lept}, E_T^{\text{miss}}) + E_T^{\text{miss}} < 60 \text{ GeV}$, after requiring exactly one lepton and at least one jet. Results for muons are shown for objects which match either the low- p_T (mu20i) or the high- p_T (mu50) trigger.

439 lepton sample, the contamination from the real leptons is subtracted based on the MC simulation pre-
 440 dictions. This leads to a systematic uncertainty (labelled in the following as “MCup/down”) due to the
 441 normalisation of the real lepton samples, which is more important for electrons and less important for
 442 muons.

⁵ In the 8 TeV analysis [2] the control region for muons was defined by the significance of the transverse impact parameter. This is not directly applicable given that a cut on $|d_0/\sigma_{d_0}| < 3$ is part of the definition of muons (cf. section 4.2).

Not reviewed, for internal circulation only

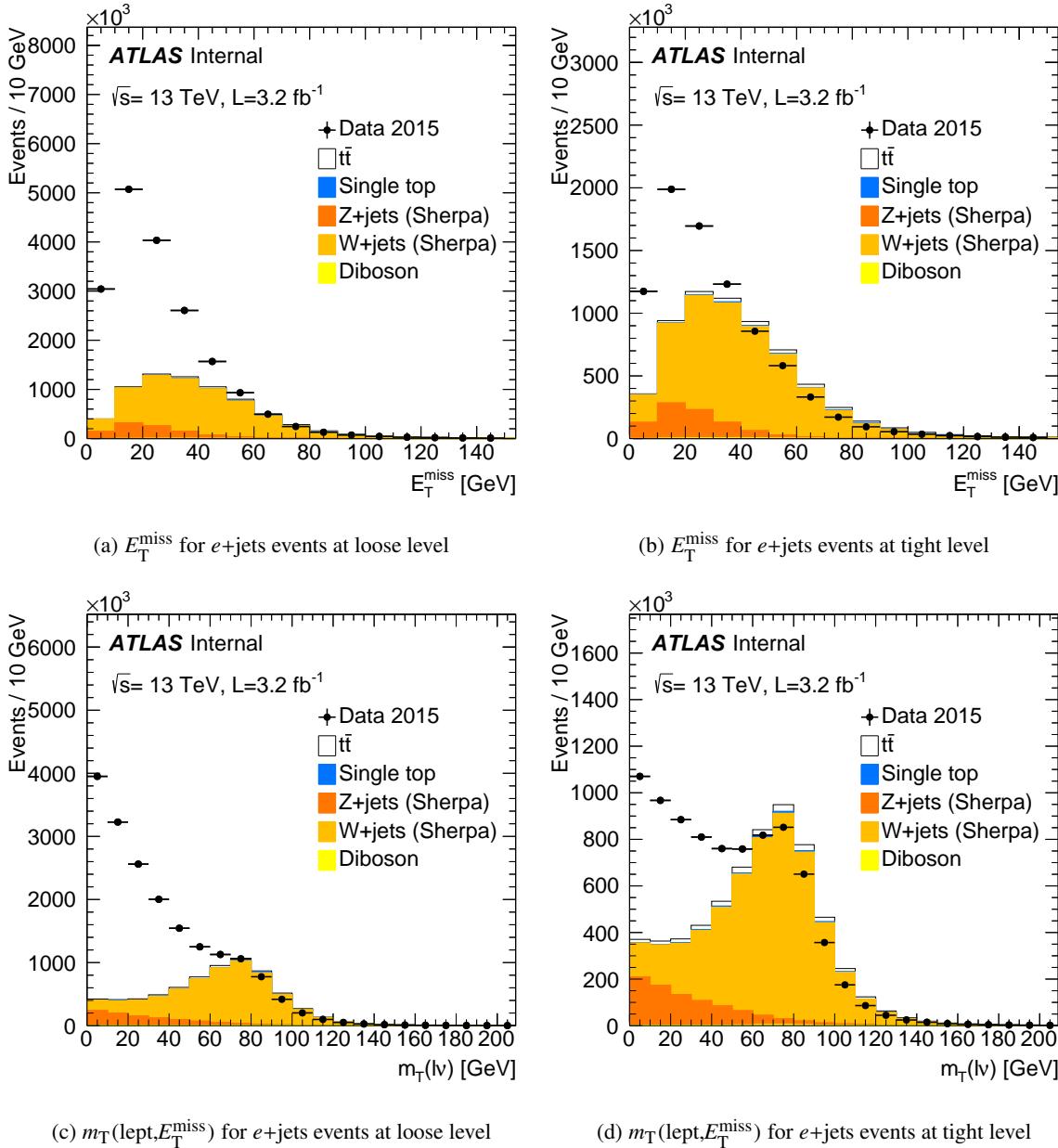


Figure 2: Distribution of (a,b) the missing transverse energy E_T^{miss} and (c,d) the $m_T(\text{lept}, E_T^{\text{miss}})$ for data and the different MC contributions at (a,c) loose and (b,d) tight levels for $e+\text{jets}$ events with one electron and at least one jet.

444 The use of other control regions is studied in sections C and D. The composition of the non-prompt and
 445 fake electrons in the control region is studied with simulated data in section ?? but the low available
 446 statistics makes the interpretation difficult.

447 After a CR_{fake} is chosen, $\varepsilon_{\text{fake}}$ can be simply determined as the ratio between the number of tight and
 448 loose events in this region. The contribution of real lepton events in the CR_{fake} is not negligible, it has to

Not reviewed, for internal circulation only

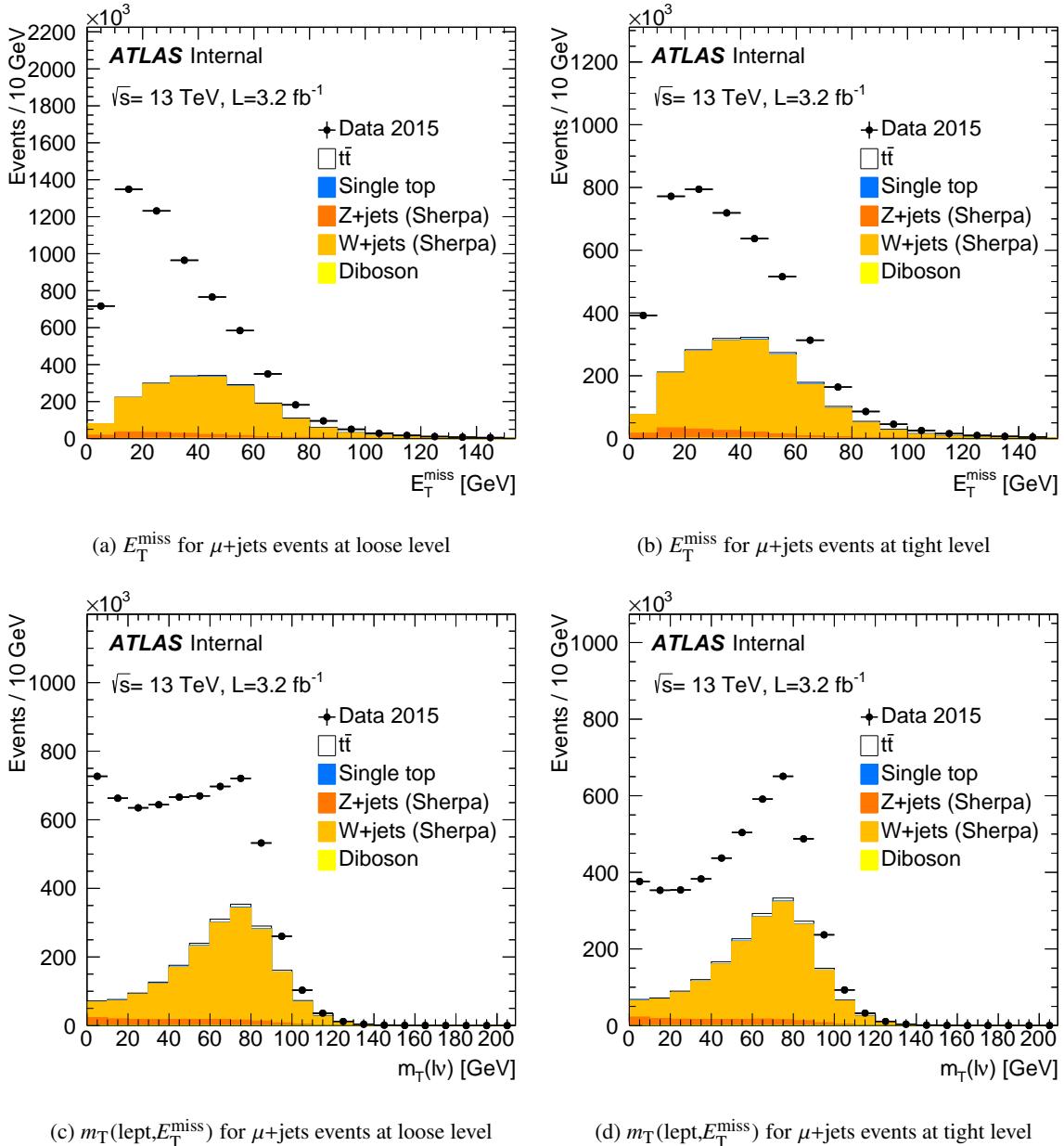


Figure 3: Distribution of (a,b) the missing transverse energy E_T^{miss} and (c,d) the $m_T(\text{lept}, E_T^{\text{miss}})$ for data and the different MC contributions at (a,c) loose and (b,d) tight levels for $\mu + \text{jets}$ events with one muon, trigger by HLT_MU20_ILOOSE_L1MU15, and at least one jet.

449 be subtracted from both the loose and tight samples:

$$\varepsilon_{\text{fake}} = \left(\frac{N_{\text{fake}}^{\text{tight}}}{N_{\text{fake}}^{\text{loose}}} \right)_{CR_{\text{fake}}} = \left(\frac{N_{\text{real}}^{\text{tight}} - N_{\text{real}}^{\text{loose}}}{N_{\text{real}}^{\text{loose}} - N_{\text{real}}^{\text{tight}}} \right)_{CR_{\text{fake}}}, \quad (8)$$

450 where $N_{\text{real}}^{\text{tight}}$ and $N_{\text{real}}^{\text{loose}}$ are the real lepton contributions to the tight and loose samples in the fake control region. These numbers can be evaluated from MC simulation. In case a MC simulation is used, a specific systematic uncertainty is introduced in the measurement of $\varepsilon_{\text{fake}}$, due to the uncertainty on the normalisation of the real lepton events contribution.

454 6.3.3. Parametrisation of efficiencies

455 One of the two muon triggers used to select events has an isolation requirement, while loose leptons are
 456 defined without any isolation cut. Efficiencies are therefore expected to be different for leptons matched
 457 to the trigger with or without isolation. Efficiencies are thus derived and applied depending on the trigger
 458 being fired by the lepton (see section 5) and on the lepton p_T being below or above the high- p_T trigger
 459 threshold. Efficiencies extracted in the case of the mu20 trigger are used in the dilepton channel for muons
 460 below the high- p_T trigger threshold not matched to the mu20i trigger.

461 Beside the dependence on the fired trigger described above (for muons), the values of $\varepsilon_{\text{real}}$ and $\varepsilon_{\text{fake}}$ are
 462 measured as a function of different variables, including: the lepton η and p_T , the angular distance ΔR
 463 between the lepton and its nearest jet, the angle in the transverse plane between the lepton and the E_T^{miss}
 464 ($\Delta\phi(\ell, E_T^{\text{miss}})$), the p_T of the leading jet, the jet and b -jet multiplicity in the event. Fig. 4 to 7 show $\varepsilon_{\text{real}}$
 465 and $\varepsilon_{\text{fake}}$, as a function of the different variables used for the parametrisation. Efficiencies are shown
 466 inclusively for electrons and muons in events with at least one jet and any number of b -jets, but separately
 467 for leptons firing each of the triggers, and in the relative lepton p_T regions. The significant dependency of
 468 the muon real and fake efficiencies on the muon p_T originates from the isolation requirements imposed to
 469 define a tight muon.

470 These efficiencies are used to compute the weights in Eq. 3 as a function of the different combinations of
 471 the variables listed above through:

$$\varepsilon_k(x_1, \dots, x_N; y_1, \dots, y_M) = \frac{1}{\varepsilon_k(x_1, \dots, x_N)^{M-1}} \cdot \prod_{j=1}^M \varepsilon_k(x_1, \dots, x_N; y_j). \quad (9)$$

472 Here the expression $\varepsilon_k(x_1, \dots, x_N)$ represents the efficiency measured as a function of all the x variables.
 473 The expression $\varepsilon_k(x_1, \dots, x_N; y_j)$ represents instead the efficiency measured as a function of all the x
 474 variables and of the variable y_j . Equation 9 implies that the full correlation between the variables x (typ-
 475 ically discrete variables, where no more than three bins are used) and each of the variables y (typically
 476 continuous variables, with a relatively large number of bins) is taken into account, while the correlation
 477 between the y variables is neglected. For each of the efficiencies ε_k , only a sub-set of the variables in
 478 each category, x or y , is used, as summarised in Table 5. This choice is driven by the observed depend-
 479 encies, the correlations between the variables and the stability of the estimates. In particular, for each of
 480 the efficiencies, the assumption of no correlation between the variables y is checked by comparing the
 481 observed dependency on the variable y_j , *i.e.* $\varepsilon_k(x_1, \dots, x_N; y_j)$, and the efficiency $\varepsilon_k(x_1, \dots, x_N; y_1, \dots, y_M)$
 482 averaged over all the other $\{y_{j'}\}_{j' \neq j}$ variables.

483 The main sources of systematic uncertainties on the non-prompt and fake lepton background deter-
 484 mination with the matrix method originate from the determination of the real efficiency, the use of MC
 485 simulation to correct the efficiency measurements, differences in the non-prompt and fake background
 486 composition in the signal regions and in the regions used to measure the efficiencies, and the treatment of
 487 the dependence of the efficiencies on lepton and event properties.

Not reviewed, for internal circulation only

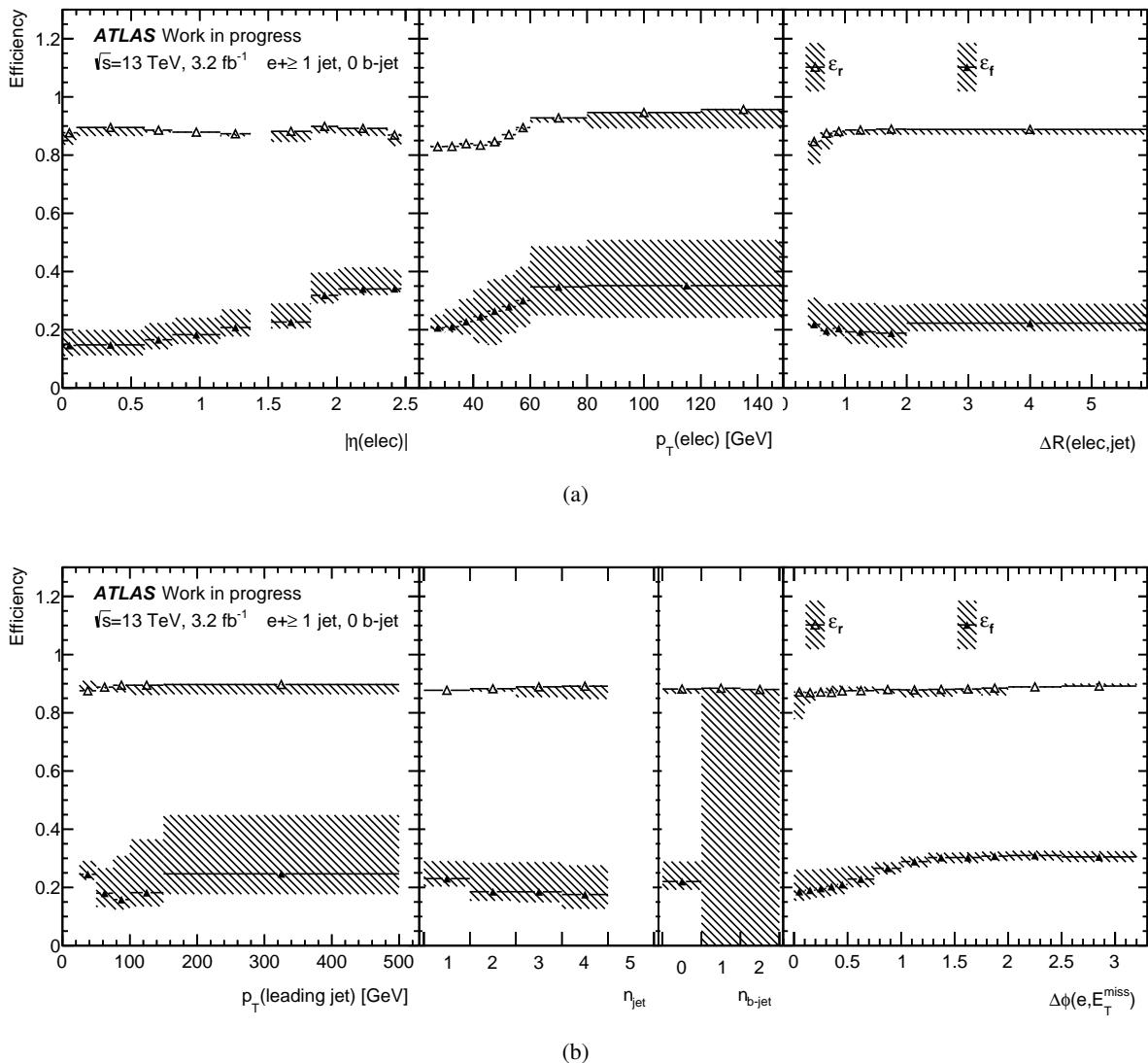


Figure 4: Efficiencies $\varepsilon_{\text{real}}$ and $\varepsilon_{\text{fake}}$ for electrons, as measured in data as a function of (a) the electron $|\eta_{\text{cl}}|$ and p_T its distance ΔR to the nearest jet, (b) the p_T of the leading jet, the jet and b -jet multiplicity and the angle in the transverse plane between the electron and the E_T^{miss} ($\Delta\phi(\text{elec}, E_T^{\text{miss}})$), for events with least one jet and 0 b -jet. The shaded area represents in each bin the combination of the statistical and systematic uncertainties on the efficiency measurements. The systematic uncertainties include the effect of using the alternative control regions (for both $\varepsilon_{\text{real}}$ and $\varepsilon_{\text{fake}}$), and the variations on the amount of real lepton events (for $\varepsilon_{\text{fake}}$).

488 The uncertainty on the real efficiency measurement method is assessed by measuring the efficiency in an
 489 independent way, by counting the fraction of tight leptons after selecting events with one loose electron
 490 (muon) in a region where the contamination from non-prompt and fake lepton events is expected to be
 491 negligible, *i.e.* by asking $E_T^{\text{miss}} > 150 \text{ GeV}$ ($m_T(\text{lept}, E_T^{\text{miss}}) > 100 \text{ GeV}$). It is found to be between
 492 1 and 7% (1-10%) in the case of electrons (muons) and to be comparable to the uncertainties on the
 493 measurement using the tag-and-probe method (see sections A.5 and B.5 for more details). The latter
 494 uncertainties, found to be around 3% for electrons and between 1 and 2% for muons, are dominated by

Not reviewed, for internal circulation only

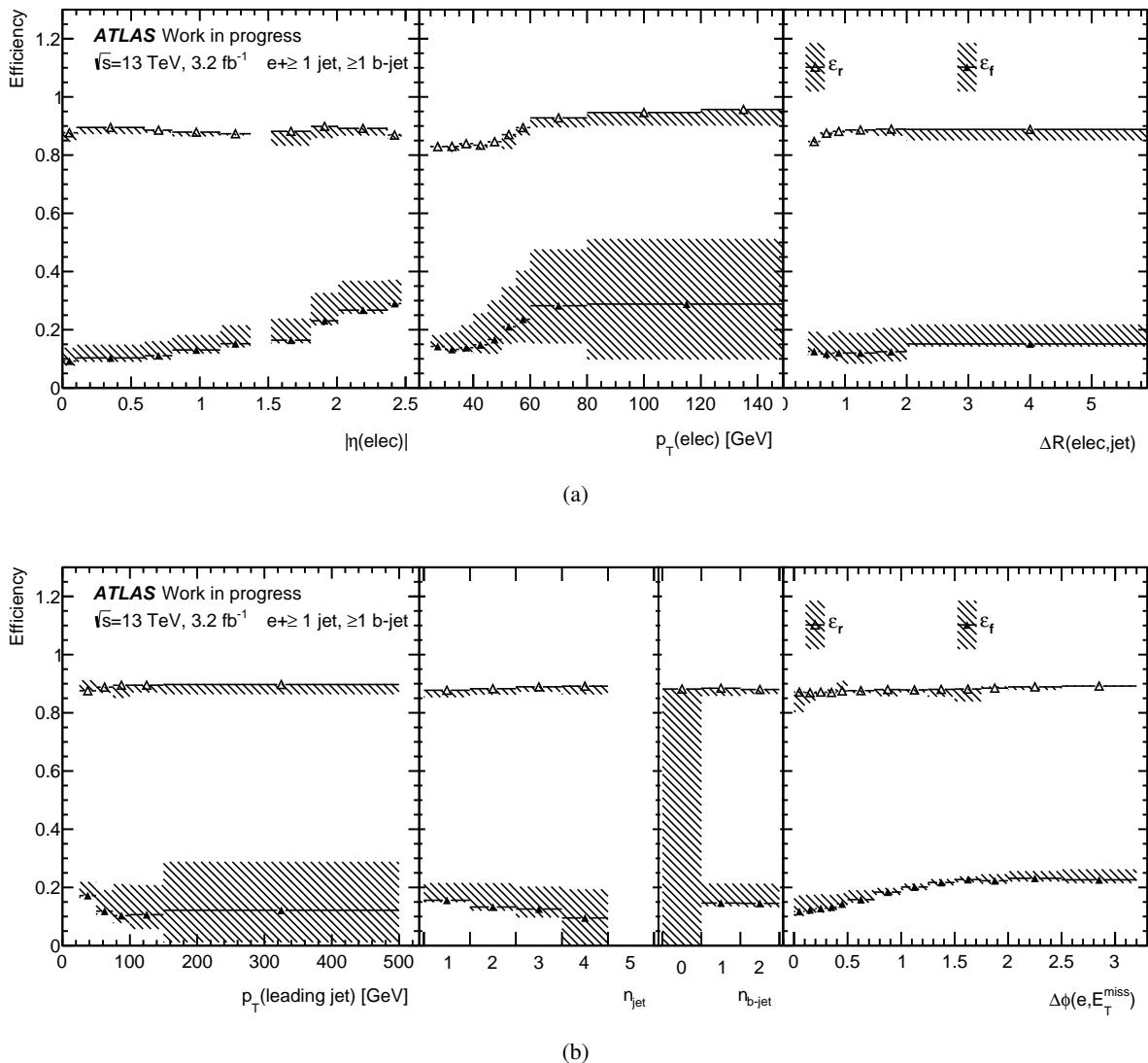


Figure 5: Efficiencies $\varepsilon_{\text{real}}$ and $\varepsilon_{\text{fake}}$ for electrons, as measured in data as a function of (a) the electron $|\eta_{\text{cl}}|$ and p_T its distance ΔR to the nearest jet, (b) the p_T of the leading jet, the jet and b -jet multiplicity and the angle in the transverse plane between the electron and the E_T^{miss} ($\Delta\phi(\text{elec}, E_T^{\text{miss}})$), for events with least one jet and at least one b -jet. The shaded area represents in each bin the combination of the statistical and systematic uncertainties on the efficiency measurements. The systematic uncertainties include the effect of using the alternative control regions (for both $\varepsilon_{\text{real}}$ and $\varepsilon_{\text{fake}}$), and the variations on the amount of real lepton events (for $\varepsilon_{\text{fake}}$).

495 the modeling of the background and the uncertainty on the correction based on MC simulation applied in
 496 the case of electrons.

497 The dominant source of systematic uncertainty on the fake efficiency measurement is that originating
 498 from the uncertainty on the normalisation of the processes determined from MC simulation in the control
 499 regions, mainly $Z + \text{jets}$ and $W + \text{jets}$. The uncertainty of their normalisation is $\sim 30\%$ and corresponds to
 500 an uncertainty of 3-13% (3-6%) on the fake efficiency in the case of electrons (muons) See sections C.4
 501 and D.4 for more details.

Not reviewed, for internal circulation only

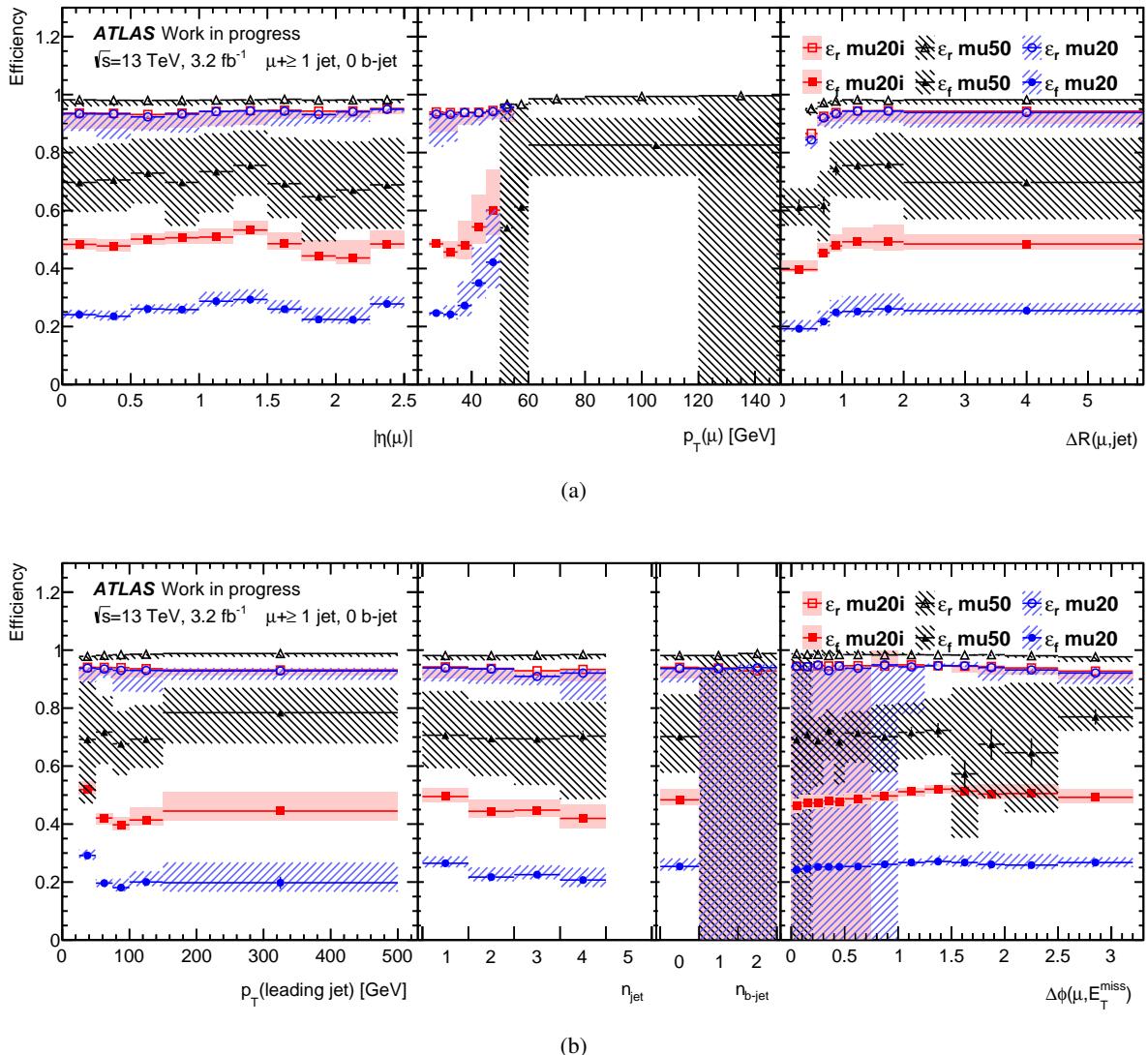


Figure 6: Efficiencies $\varepsilon_{\text{real}}$ and $\varepsilon_{\text{fake}}$ for muons, as measured in data as a function of (a) the muon $|\eta|$ and p_T its distance ΔR to its nearest jet, (b) the p_T of the leading jet, the jet and b -jet multiplicity and the angle in the transverse plane between the muon and the E_T^{miss} ($\Delta\phi(\mu, E_T^{\text{miss}})$), for events with at least one jet and 0 b -jet. The efficiencies are shown separately for probes which match specifically one of the triggers used to selected data (mu20i or mu50) or the low- p_T trigger with no isolation requirement (mu20). The shaded area represents the combination in each bin of the statistical and systematic uncertainties on the efficiency measurements. The systematic uncertainties include the effect of using the alternative control regions (for both $\varepsilon_{\text{real}}$ and $\varepsilon_{\text{fake}}$), and the variations on the amount of real lepton events (for $\varepsilon_{\text{fake}}$).

Another significant source of uncertainty is assessed through the use of alternative control regions to measure the efficiencies, defined by different combinations of cuts on E_T^{miss} and $m_T(\text{lept}, E_T^{\text{miss}})$, i.e. $m_T(\text{lept}, E_T^{\text{miss}}) < 20$ GeV. This approach allows to partially assess the uncertainty coming from the relative composition of the non-prompt and fake lepton samples in the control and signal regions. The uncertainty is found to be between 2 and 15% (2-7%) on the fake efficiency in the case of electrons

Not reviewed, for internal circulation only

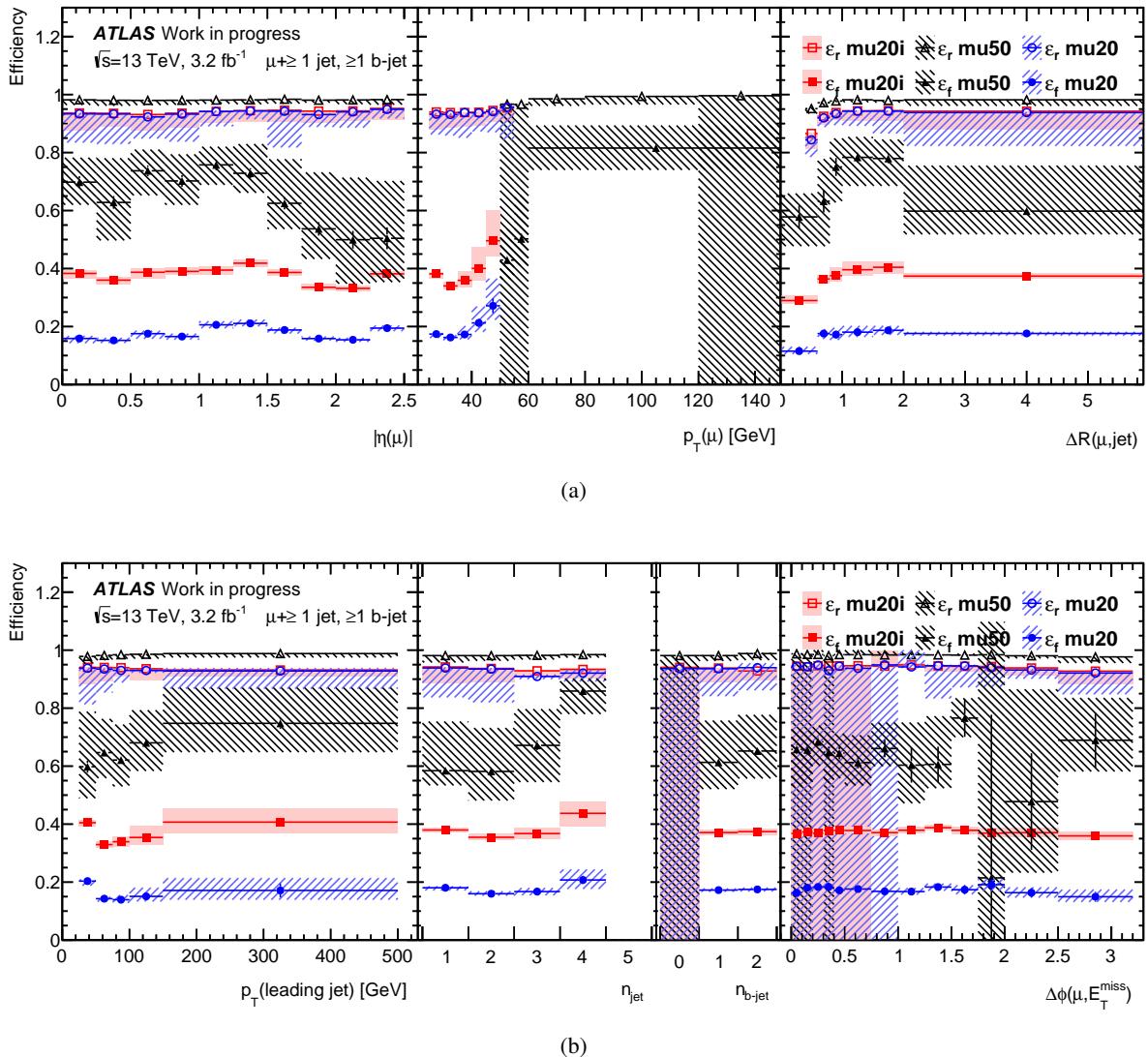


Figure 7: Efficiencies $\varepsilon_{\text{real}}$ and $\varepsilon_{\text{fake}}$ for muons, as measured in data as a function of (a) the muon $|\eta|$ and p_T its distance ΔR to its nearest jet, (b) the p_T of the leading jet, the jet and b -jet multiplicity and the angle in the transverse plane between the muon and the E_T^{miss} ($\Delta\phi(\mu, E_T^{\text{miss}})$), for events with at least one jet and at least one b -jet. The efficiencies are shown separately for probes which match specifically one of the triggers used to selected data (mu20i or mu50) or the low- p_T trigger with no isolation requirement (mu20). The shaded area represents the combination in each bin of the statistical and systematic uncertainties on the efficiency measurements. The systematic uncertainties include the effect of using the alternative control regions (for both $\varepsilon_{\text{real}}$ and $\varepsilon_{\text{fake}}$), and the variations on the amount of real lepton events (for $\varepsilon_{\text{fake}}$).

507 (muons). See sections C.3 and D.3 for more details.

508 Finally, different choices for the combinations of variables used in the efficiency parametrisation are
509 compared. In particular, the most relevant variation is found to come from the use of $\Delta\phi(\text{lept}, E_T^{\text{miss}})$ in
510 the $\varepsilon_{\text{fake}}$ parametrisation, and is used to assess the uncertainty related to the treatment of the efficiency
511 dependencies on lepton and event properties.

Table 5: Summary of the variables used to parametrise the real and fake lepton efficiencies in the matrix method. The column ‘Trigger’ refers to the specific trigger the lepton matches, $p_T^{\text{lead,jet}}$ stays for p_T of the leading jet in the event, $\Delta R(\ell,\text{jet})$ is the angular distance between the lepton and the closest jets, $\Delta\phi(\ell,E_T^{\text{miss}})$ is the angular distance in the transverse plane between the lepton and the missing energy in the event. For each of the efficiencies, the variables for which the explicit dependence is used are indicated. The variables are divided in two categories, x and y , depending the specific treatment in terms of correlation. See text for details.

	x variables		y variables					
	Trigger	n_{jet}	$n_{b-\text{jet}}$	$ \eta^\ell $	p_T^ℓ	$p_T^{\text{lead,jet}}$	$\Delta R(\ell,\text{jet})$	$\Delta\phi(\ell,E_T^{\text{miss}})$
$\varepsilon_{\text{real}}(e)$		✓		✓	✓	✓	✓	✓
$\varepsilon_{\text{real}}(\mu)$	✓	✓		✓	✓	✓	✓	✓
$\varepsilon_{\text{fake}}(e)$		✓	✓	✓	✓	✓	✓	✓
$\varepsilon_{\text{fake}}(\mu)$	✓	✓	✓	✓	✓	✓	✓	✓

512 To evaluate the uncertainty on the non-prompt and fake and background contribution, the matrix method
 513 input efficiencies are varied as described above, and the background distributions and yields are then re-
 514 derived. The observed deviation of the yields measured where lepton efficiencies are varied is assigned
 515 as an uncertainty. The total systematic uncertainty on the estimate is taken as the quadratic sum of the
 516 symmetrised individual variations.

517 6.4. Results in $l+\text{jets}$ validation regions

518 The estimates of the non-prompt and fake lepton backgrounds for semileptonic $t\bar{t}$ events is done using the
 519 selection criteria of such events, with different selection cuts on E_T^{miss} and $m_T(\text{lept},E_T^{\text{miss}})$. Fig. 8 and 12
 520 show the distributions of E_T^{miss} , $m_T(\text{lept},E_T^{\text{miss}})$, lepton p_T , lepton η , p_T of the leading jet and distance
 521 ΔR between the lepton and the closest jet, for events with only two jets, in order to keep the contribution
 522 from top events small. Distributions are shown without any cuts on E_T^{miss} and $m_T(\text{lept},E_T^{\text{miss}})$, to enrich
 523 the sample in non-prompt and fake lepton events. The contribution from the different simulated data
 524 are shown with the full histogram. The open histograms shown with different colors and line styles
 525 indicate the background estimate obtained with the different systematic variations of both $\varepsilon_{\text{real}}$ and $\varepsilon_{\text{fake}}$.
 526 The default estimates are shown in black. The estimates “CRreal”, using the real efficiencies from the
 527 alternative method, are shown in red. The effects of the variations of the level of MC to subtract to
 528 measure fake efficiencies (“MCup” and “MCdown”) are shown in blue and purple. The estimates using
 529 the fake efficiencies estimated in the alternative control region (“CRfake”) are shown in green. Finally
 530 the estimates “EffPar”, using the alternative parametrisation of efficiencies, are shown in cyan.

531 Fig. 9 and 13 show the same distributions in the same region, but focusing on the non-prompt and fake
 532 lepton predictions only, so that the absolute normalisation and shape variations are visible, also in regions
 533 where the fake leptons prediction is small with respect to the real leptons one.

534 Fig. 10 and 14 show the default estimates with their total uncertainties, together with the real lepton
 535 contributions broken down into physics processes, for the same distributions with the same selection as
 536 above. Fig. 11 and 15 show the same but requiring at least one b -tagged jet. The agreement is rather good
 537 for all variables and discrepancies are mostly covered by the systematics. Additional results are given in
 538 sections E.1 and E.2 in other control regions.

539 Tables 6 and 7 report the number of selected data events compared to the predicted non-prompt and fake
540 lepton events from matrix method and real lepton events from MC simulation, after selecting exactly two
541 or at least four jets, for pretag and ≥ 1 b -tag selection, without cuts on E_T^{miss} and $m_T(\text{lept}, E_T^{\text{miss}})$.

542 Tables 8 and 9 show the systematic variation effects on the non-prompt and fake lepton estimation in the
543 same regions.

544 The total expected number of events, obtained from the different MC simulated samples for real lepton
545 and non-prompt and fake lepton backgrounds, always exceed the number of events in the data, as could
546 already been noticed in most of plots. Results using other MC samples for the $W + \text{jets}$ and $Z + \text{jets}$
547 backgrounds can be found in section E.1.6 and E.2.6 but none of them is satisfactory.

548 In the $l + \text{jets}$ control regions, the systematic uncertainties on the fake estimates are between 70 and 120%
549 (30-70%) for $e + \text{jets}$ ($\mu + \text{jets}$) events, depending on the jet and b -jet multiplicity. The use of the real
550 lepton subtraction from CR_f is the dominant sources in $e + \text{jets}$ events, effecting between 60 and 80%.
551 In $\mu + \text{jets}$ events “CRreal”, the alternative estimate for $\varepsilon_{\text{real}}$, and “MCdown” produce relatively large
552 deviations, from 10 to 50%.

Not reviewed, for internal circulation only

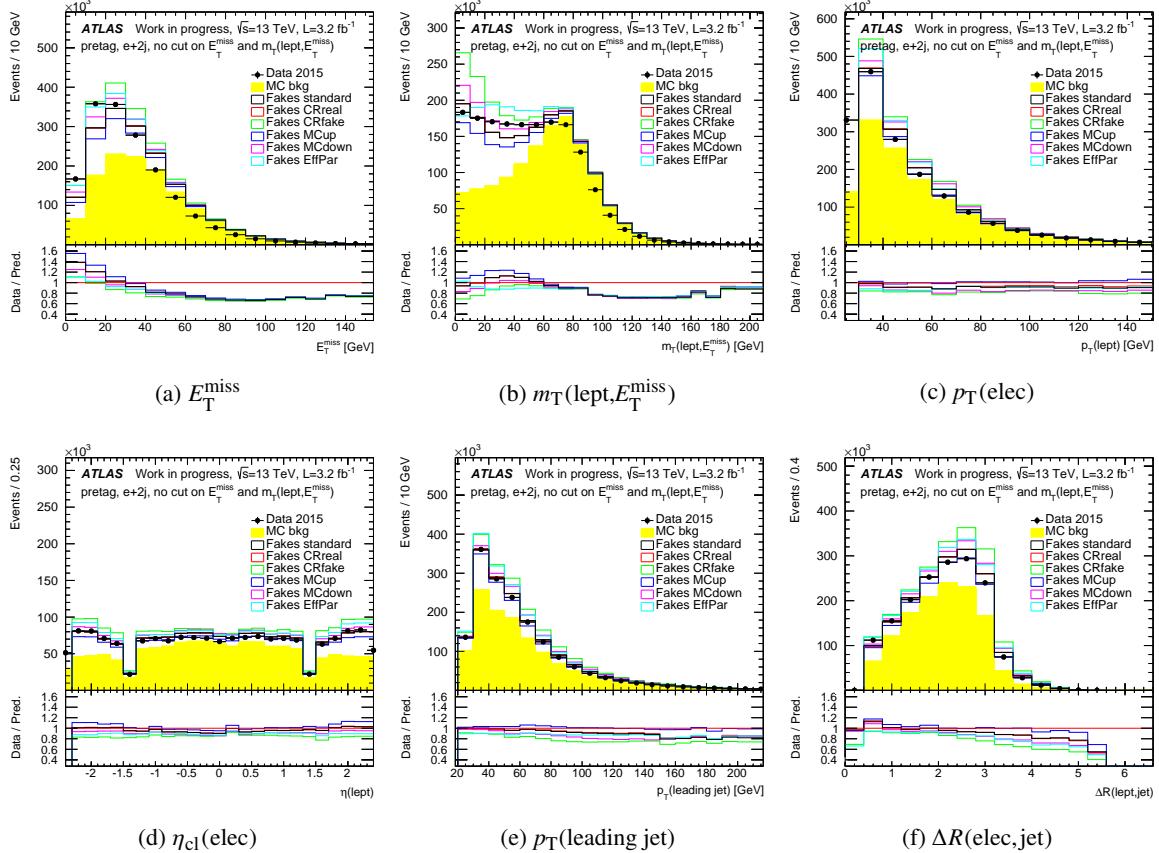


Figure 8: Distributions of (a) E_T^{miss} , (b) $m_T(\text{lept}, E_T^{\text{miss}})$, (c) p_T of the electron, (d) η of the electron, (e) p_T of the leading jet and (f) the distance ΔR of the electron to its nearest jet, for $e+jets$ events with two jets, without any b -tagging requirement, without any cuts on E_T^{miss} and $m_T(\text{lept}, E_T^{\text{miss}})$. The data is compared to the real lepton expectation from simulation, showing the contributions from $t\bar{t}$, single top, $W+jets$, $Z+jets$ and dibosons normalised to their cross-sections (Background), and non-prompt and fake lepton backgrounds (referred to as ‘NP & Fake Lep.’) estimated with the matrix method obtained with the different variations of both $\varepsilon_{\text{real}}$ and $\varepsilon_{\text{fake}}$.

Not reviewed, for internal circulation only

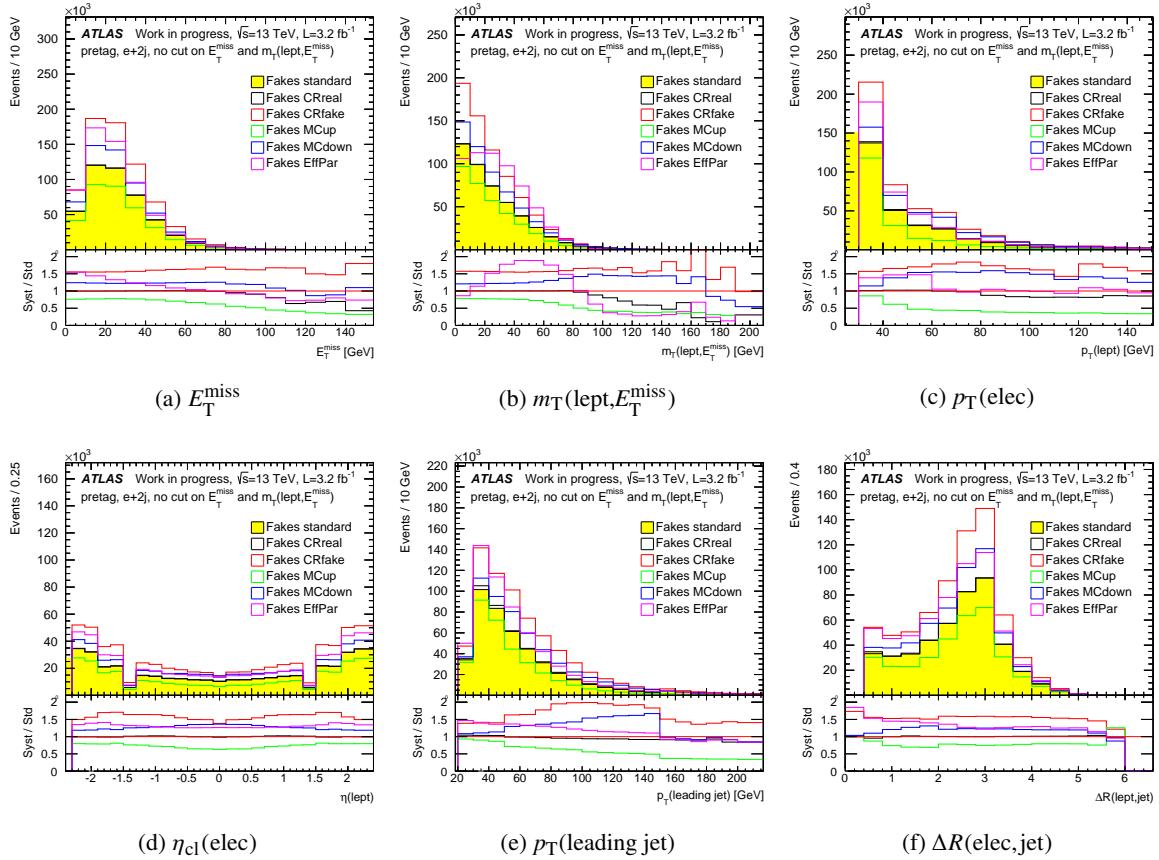


Figure 9: Distributions of (a) E_T^{miss} , (b) $m_T(\text{lept}, E_T^{\text{miss}})$, (c) p_T of the electron, (d) η of the electron, (e) p_T of the leading jet and (f) the distance ΔR of the electron to its nearest jet, for $e+jets$ events with two jets, without any b -tagging requirement, without any cuts on E_T^{miss} and $m_T(\text{lept}, E_T^{\text{miss}})$. Only the non-prompt and fake lepton backgrounds estimated with the matrix method, obtained with the different variations of both $\varepsilon_{\text{real}}$ and $\varepsilon_{\text{fake}}$, are shown.

Not reviewed, for internal circulation only

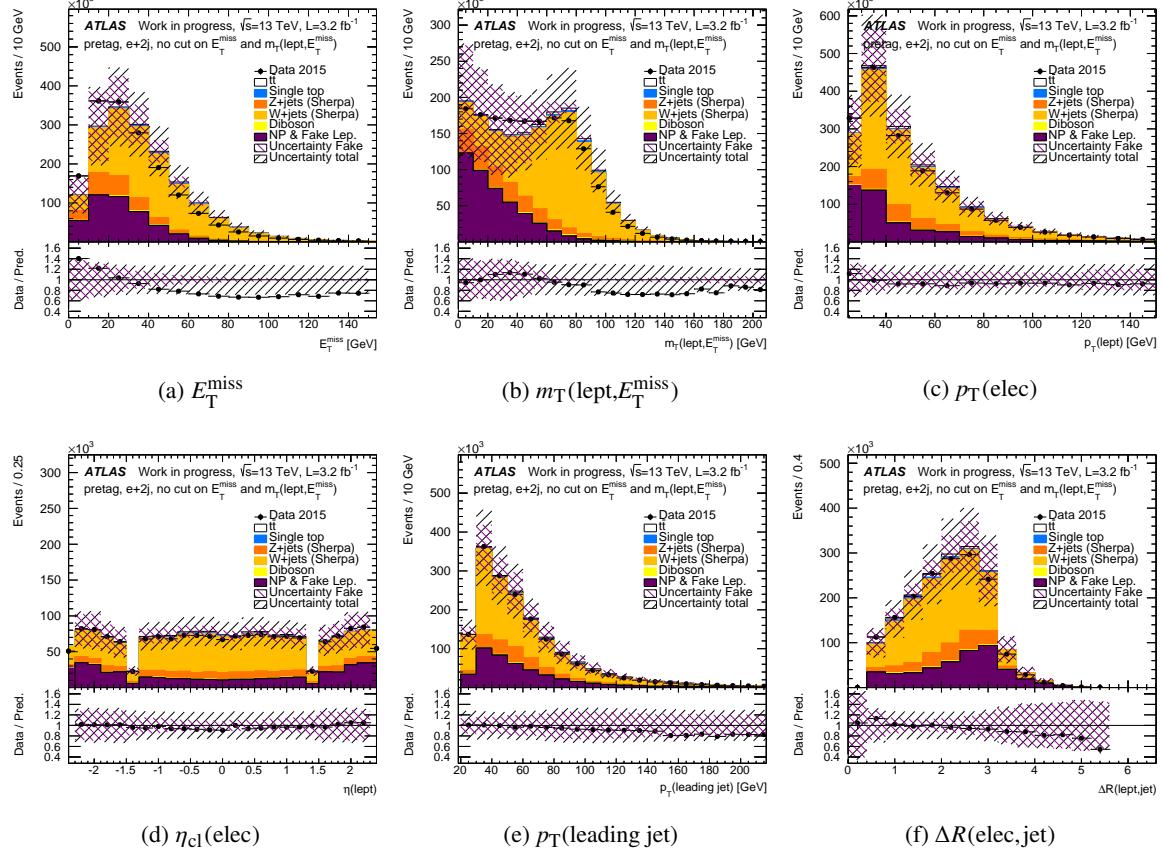


Figure 10: Distributions of (a) E_T^{miss} , (b) $m_T(\text{lept}, E_T^{\text{miss}})$, (c) p_T of the electron, (d) η_{cl} of the electron, (e) p_T of the leading jet and (f) the distance ΔR of the electron to its nearest jet, for $e+jets$ events with two jets, without any b -tagging requirement, without any cuts on E_T^{miss} and $m_T(\text{lept}, E_T^{\text{miss}})$. The data is compared to the real lepton expectation from simulation, showing separately the contributions from $t\bar{t}$, single top, $W+jets$, $Z+jets$ and dibosons normalised to their cross-sections, and non-prompt and fake lepton backgrounds (referred to as ‘NP & Fake Lep.’) estimated with the matrix method. The purple shaded area represents the combination of the statistical and the systematic uncertainties on the matrix method estimate in each bin. The black shaded area represents the total uncertainty including also the uncertainties on the processes predicted by the MC simulation and the one on the luminosity.

Not reviewed, for internal circulation only

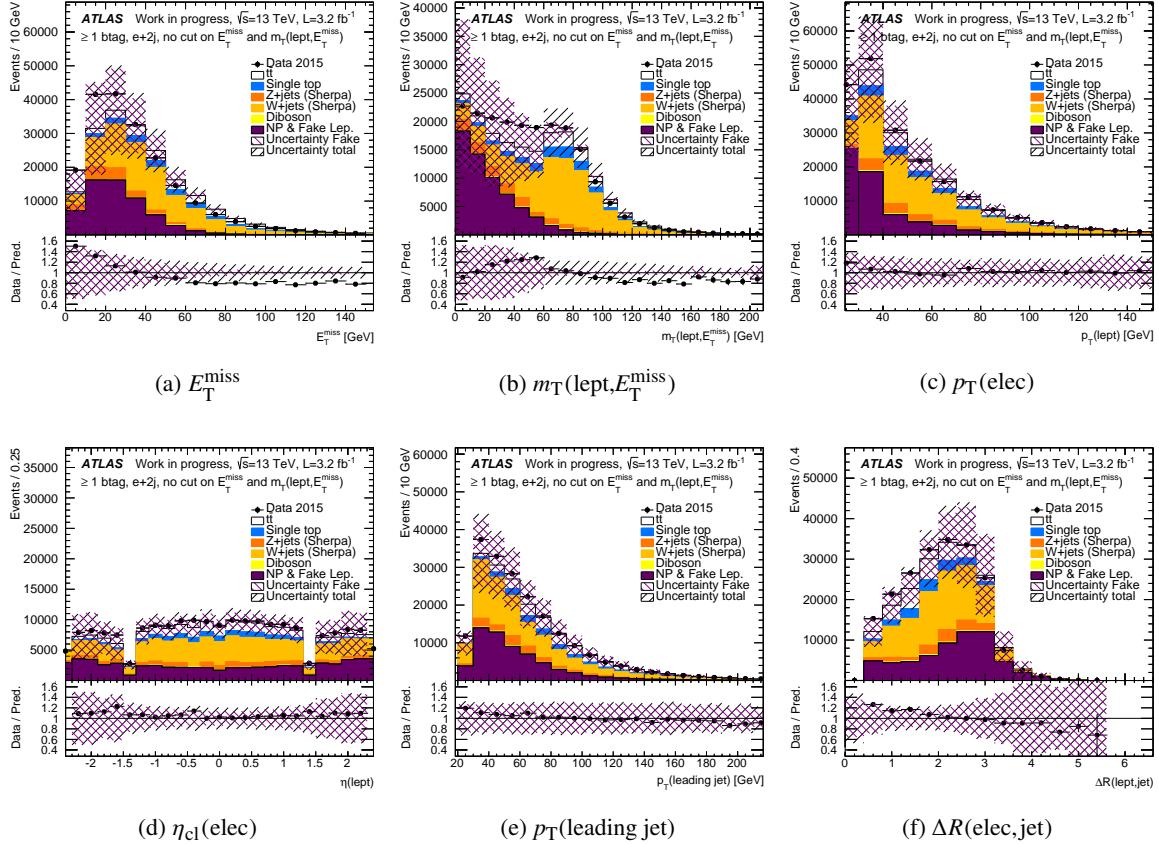


Figure 11: Distributions of (a) E_T^{miss} , (b) $m_T(\text{lept}, E_T^{\text{miss}})$, (c) p_T of the electron, (d) η_{cl} of the electron, (e) p_T of the leading jet and (f) the distance ΔR of the electron to its nearest jet, for $e+jets$ events with two jets, requiring at least one b -jet, without any cuts on E_T^{miss} and $m_T(\text{lept}, E_T^{\text{miss}})$. The data is compared to the real lepton expectation from simulation, showing separately the contributions from $t\bar{t}$, single top, $W+jets$, $Z+jets$ and dibosons normalised to their cross-sections, and non-prompt and fake lepton backgrounds (referred to as ‘NP & Fake Lep.’) estimated with the matrix method. The purple shaded area represents the combination of the statistical and the systematic uncertainties on the matrix method estimate in each bin. The black shaded area represents the total uncertainty including also the uncertainties on the processes predicted by the MC simulation and the one on the luminosity.

Not reviewed, for internal circulation only

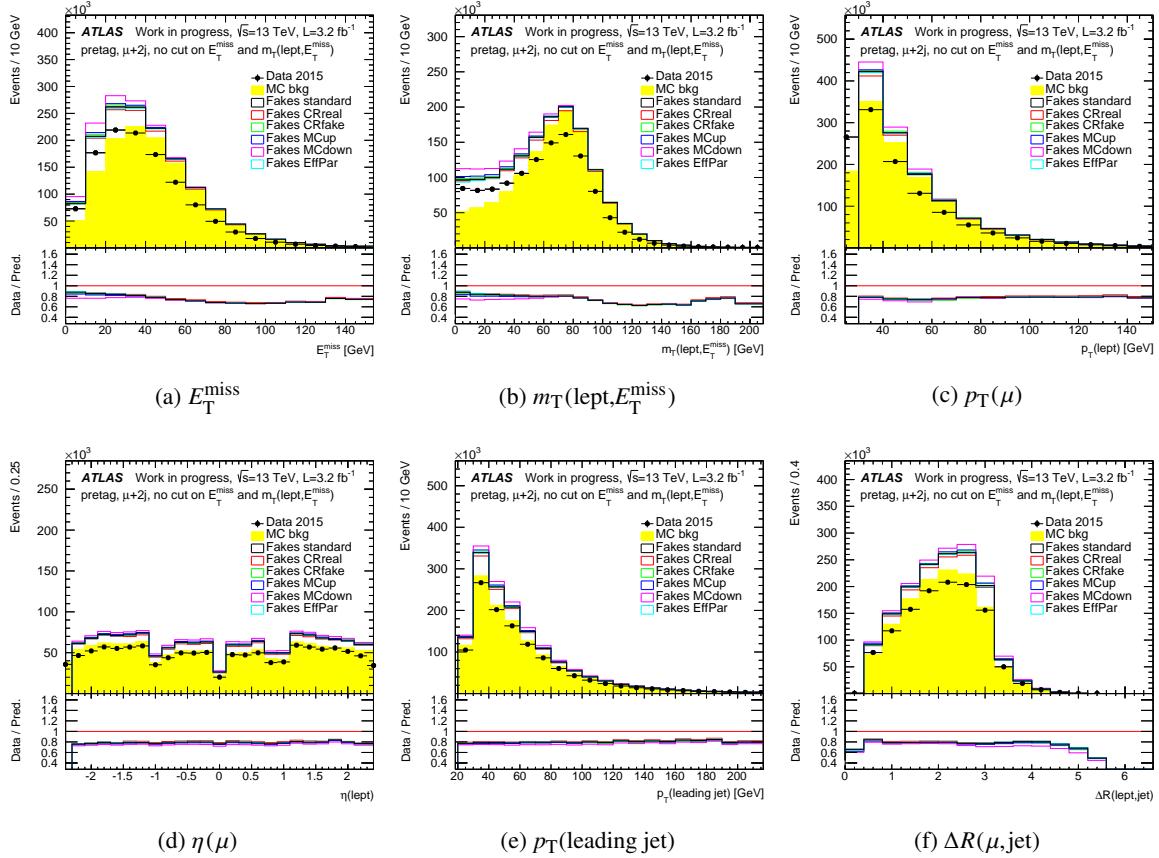


Figure 12: Distributions of (a) E_T^{miss} , (b) $m_T(\text{lept}, E_T^{\text{miss}})$, (c) p_T of the muon, (d) η of the muon, (e) p_T of the leading jet and (f) the distance ΔR of the muon to its nearest jet, for $\mu+\text{jets}$ events with two jets, without any b -tagging requirement, without any cuts on E_T^{miss} and $m_T(\text{lept}, E_T^{\text{miss}})$. The data is compared to the real lepton expectation from simulation, showing the contributions from $t\bar{t}$, single top, $W+\text{jets}$, $Z+\text{jets}$ and dibosons normalised to their cross-sections (Background), and non-prompt and fake lepton backgrounds (referred to as ‘NP & Fake Lep.’) estimated with the matrix method obtained with the different variations of both $\varepsilon_{\text{real}}$ and $\varepsilon_{\text{fake}}$.

Not reviewed, for internal circulation only

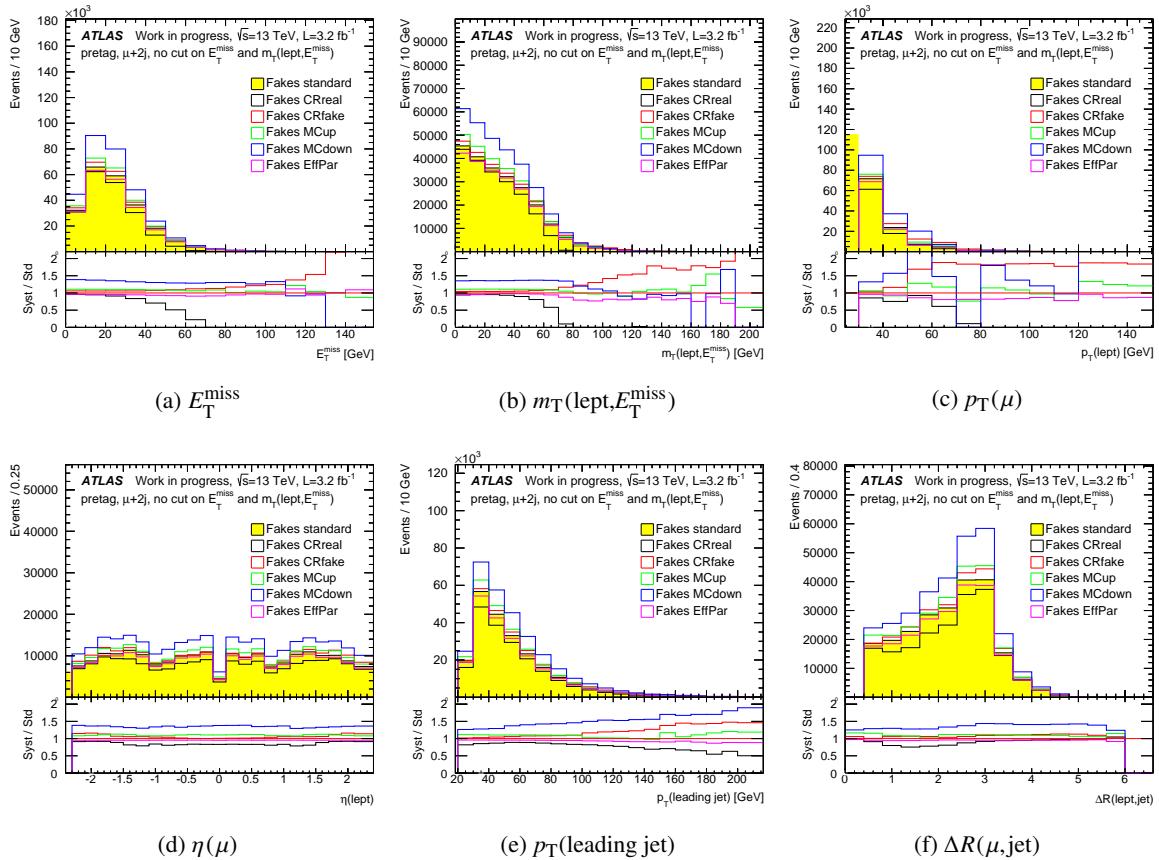


Figure 13: Distributions of (a) E_T^{miss} , (b) $m_T(\text{lept}, E_T^{\text{miss}})$, (c) p_T of the muon, (d) η of the muon, (e) p_T of the leading jet and (f) the distance ΔR of the muon to its nearest jet, for $\mu + \text{jets}$ events with two jets, without any b -tagging requirement, without any cuts on E_T^{miss} and $m_T(\text{lept}, E_T^{\text{miss}})$. Only the non-prompt and fake lepton backgrounds estimated with the matrix method, obtained with the different variations of both $\varepsilon_{\text{real}}$ and $\varepsilon_{\text{fake}}$, are shown.

Not reviewed, for internal circulation only

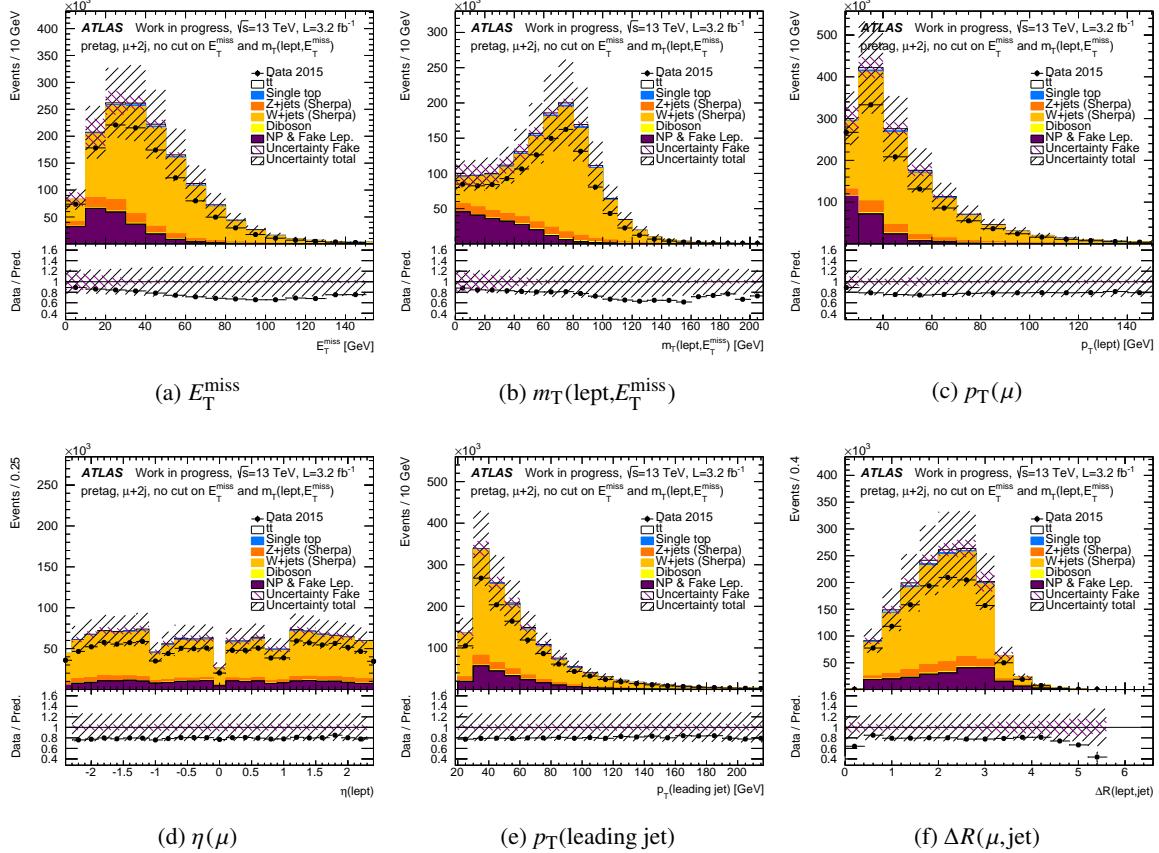


Figure 14: Distributions of (a) E_T^{miss} , (b) $m_T(\text{lept}, E_T^{\text{miss}})$, (c) p_T of the muon, (d) η of the muon, (e) p_T of the leading jet and (f) the distance ΔR of the muon to its nearest jet, for $\mu + \text{jets}$ events with two jets, without any b -tagging requirement, without any cuts on E_T^{miss} and $m_T(\text{lept}, E_T^{\text{miss}})$. The data is compared to the real lepton expectation from simulation, showing separately the contributions from $t\bar{t}$, single top, $W + \text{jets}$, $Z + \text{jets}$ and dibosons normalised to their cross-sections, and non-prompt and fake lepton backgrounds (referred to as ‘NP & Fake Lep.’) estimated with the matrix method. The purple shaded area represents the combination of the statistical and the systematic uncertainties on the matrix method estimate in each bin. The black shaded area represents the total uncertainty including also the uncertainties on the processes predicted by the MC simulation and the one on the luminosity.

Not reviewed, for internal circulation only

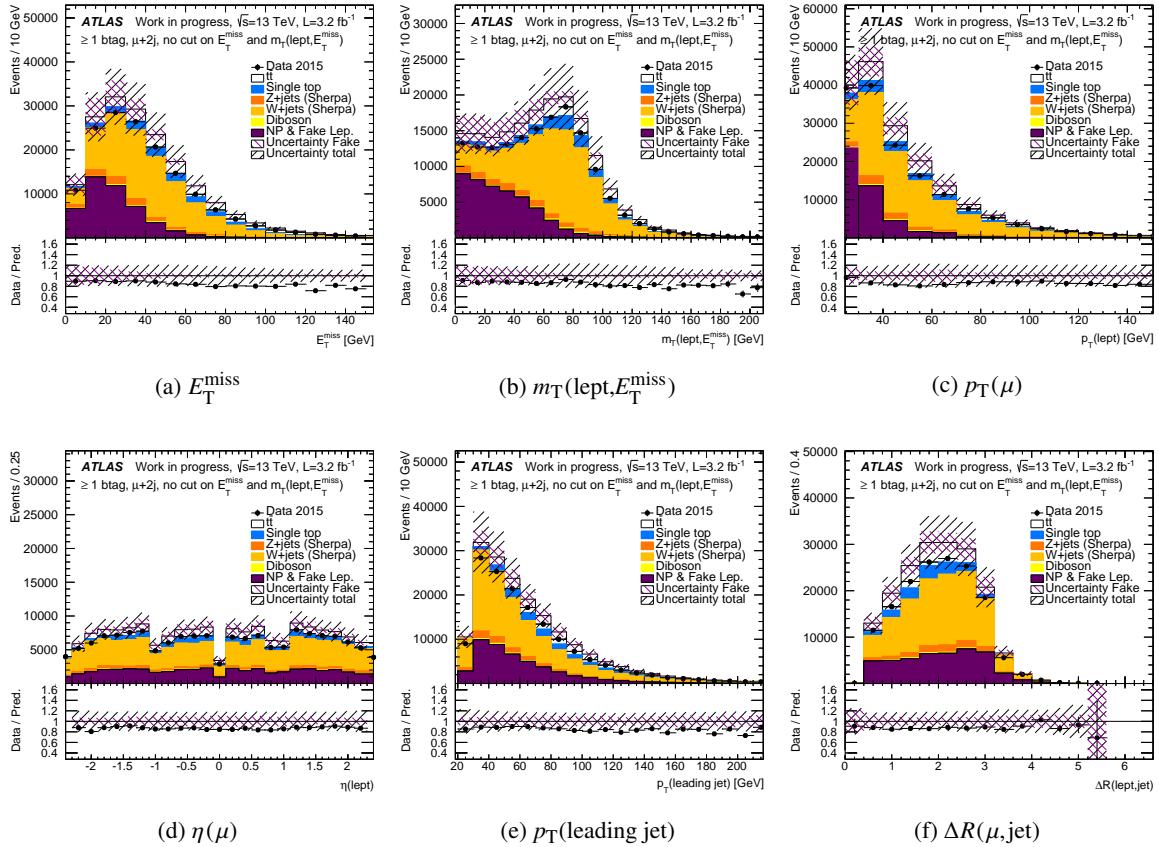


Figure 15: Distributions of (a) E_T^{miss} , (b) $m_T(\text{lept}, E_T^{\text{miss}})$, (c) p_T of the muon, (d) η of the muon, (e) p_T of the leading jet and (f) the distance ΔR of the muon to its nearest jet, for $\mu + \text{jets}$ events with two jets, requiring at least one b -jet, without any cuts on E_T^{miss} and $m_T(\text{lept}, E_T^{\text{miss}})$. The data is compared to the real lepton expectation from simulation, showing separately the contributions from $t\bar{t}$, single top, $W + \text{jets}$, $Z + \text{jets}$ and dibosons normalised to their cross-sections, and non-prompt and fake lepton backgrounds (referred to as ‘NP & Fake Lep.’) estimated with the matrix method. The purple shaded area represents the combination of the statistical and the systematic uncertainties on the matrix method estimate in each bin. The black shaded area represents the total uncertainty including also the uncertainties on the processes predicted by the MC simulation and the one on the luminosity.

Table 6: Number of events expected from the different MC simulated samples for real lepton and non-prompt and fake lepton backgrounds estimated with the matrix method in e +jets events, before any cut on E_T^{miss} and $m_T(\text{lept}, E_T^{\text{miss}})$, compared to the number of events observed in data.

e +jets no cuts on E_T^{miss} and $m_T(\text{lept}, E_T^{\text{miss}})$	2-jets pretag	2-jets $\geq 1 b$ -tags	≥ 4 -jets pretag	≥ 4 -jets $\geq 1 b$ -tags
Data 2015	1671710	200730	264541	134200
Fake Leptons	449637	61324	48215	9629
(stat+syst)	± 314344	± 48209	± 38416	± 10994
Fake Fraction	26.9%	30.6%	18.2%	7.2%
$t\bar{t}$	340901	25525	124741	112301
Single Top	17975	13275	9171	7773
W +jets	1008810	75481	99434	16373
Z +jets	201680	13123	21476	3646
Diboson	9520	1136	4127	978
Tot. MC	1272160	128543	258949	141072
Tot. Expected	1720750	189867	307165	150701

Not reviewed, for internal circulation only

Table 7: Number of events expected from the different MC simulated samples for real lepton and non-prompt and fake lepton backgrounds estimated with the matrix method in μ +jets events, before any cut on E_T^{miss} and $m_T(\text{lept}, E_T^{\text{miss}})$, compared to the number of events observed in data.

μ +jets no cuts on E_T^{miss} and $m_T(\text{lept}, E_T^{\text{miss}})$	2-jets pretag	2-jets $\geq 1 b$ -tags	≥ 4 -jets pretag	≥ 4 -jets $\geq 1 b$ -tags
Data 2015	1197890	155481	208632	116835
Fake Leptons	227214	45847	21953	8168
(stat+syst)	± 88201	± 14960	± 11708	± 5825
Fake Fraction	20.0%	32.6%	10.5%	7.0%
$t\bar{t}$	33096	24777	117571	105961
Single Top	17799	13144	8588	7279
W +jets	1118130	84964	106308	18008
Z +jets	100764	8311	11060	2025
Diboson	10026	1165	3892	901
Tot. MC	1279860	132361	247419	134175
Tot. Expected	1507070	178208	269372	142343

Table 8: Effect (in %) of the various sources of systematic uncertainty on the fake lepton estimation with the matrix method in $e+\text{jets}$ events, before any cut on E_T^{miss} and $m_T(\text{lept}, E_T^{\text{miss}})$.

$e+\text{jets}$ no cuts on E_T^{miss} and $m_T(\text{lept}, E_T^{\text{miss}})$	2-jets pretag	2-jets $\geq 1 b\text{-tags}$	$\geq 4\text{-jets}$ pretag	$\geq 4\text{-jets}$ $\geq 1 b\text{-tags}$
MC up	-24	-15	-40	-46
MC down	23	23	31	68
CR real	0	-2	-8	-11
CR fake	57	66	64	83
Alt.Par.	33	34	24	35
Tot.Syst.	70	80	80	115

Table 9: Effect (in %) of the various sources of systematic uncertainty on the fake lepton estimation with the matrix method in $\mu+\text{jets}$ events, before any cut on E_T^{miss} and $m_T(\text{lept}, E_T^{\text{miss}})$.

$\mu+\text{jets}$ no cuts on E_T^{miss} and $m_T(\text{lept}, E_T^{\text{miss}})$	2-jets pretag	2-jets $\geq 1 b\text{-tags}$	$\geq 4\text{-jets}$ pretag	$\geq 4\text{-jets}$ $\geq 1 b\text{-tags}$
MC up	10	7	-3	-23
MC down	35	29	42	46
CR real	-15	-9	-30	-53
CR fake	6	-7	13	6
Alt.Par.	-5	-9	-6	-8
Tot.Syst.	39	30	53	70

553 7. Estimates in lepton+jets $t\bar{t}$ events

554 The fake lepton background estimation techniques described in the previous sections are applied to different jet- and b -jet multiplicity regions, applying the ' $t\bar{t} \ell + \text{jets}$ ' event selection described in section 5.2.

556 7.1. Estimates for $e+\text{jets}$ events

557 Fig. 16 to 19 show data compared to the fake lepton estimate plus the various real lepton contributions from simulation results for $e+\text{jets}$ events. Additional plots are shown in section E.1.

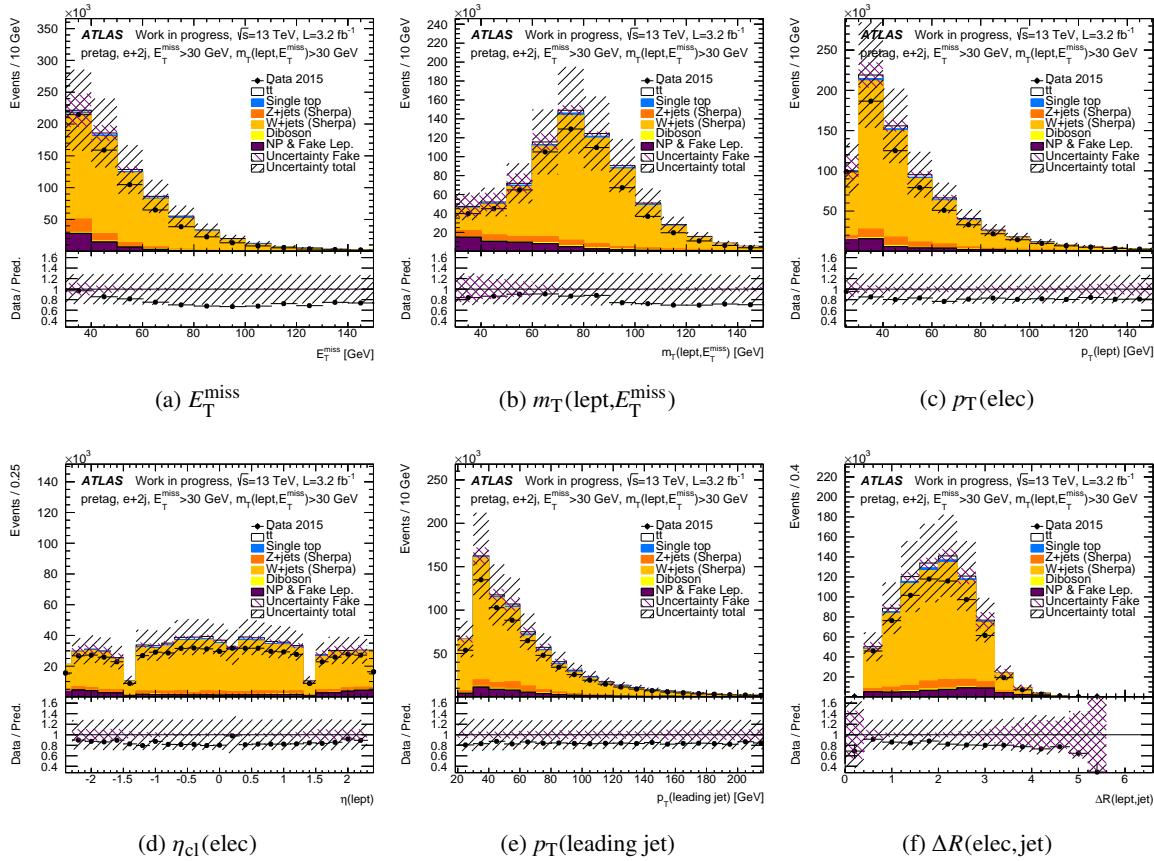


Figure 16: Distributions of (a) E_T^{miss} , (b) $m_T(\text{lept}, E_T^{\text{miss}})$, (c) p_T of the electron, (d) η_{cl} of the electron, (e) p_T of the leading jet and (f) the distance $\Delta R(\text{elec}, \text{jet})$ of the electron to its nearest jet, for $e+\text{jets}$ events with two jets, without any b -tagging requirement, requiring $E_T^{\text{miss}} > 30 \text{ GeV}$ and $m_T(\text{lept}, E_T^{\text{miss}}) > 30 \text{ GeV}$. The data is compared to the real lepton expectation from simulation, showing separately the contributions from $t\bar{t}$, single top, $W+\text{jets}$, $Z+\text{jets}$ and dibosons normalised to their cross-sections, and non-prompt and fake lepton backgrounds (referred to as ‘NP & Fake Lep.’) estimated with the matrix method. The purple shaded area represents the combination of the statistical and the systematic uncertainties on the matrix method estimate in each bin. The black shaded area represents the total uncertainty including also the uncertainties on the processes predicted by the MC simulation and the one on the luminosity. The lower parts of the figure show the ratios of expectation to data.

Not reviewed, for internal circulation only

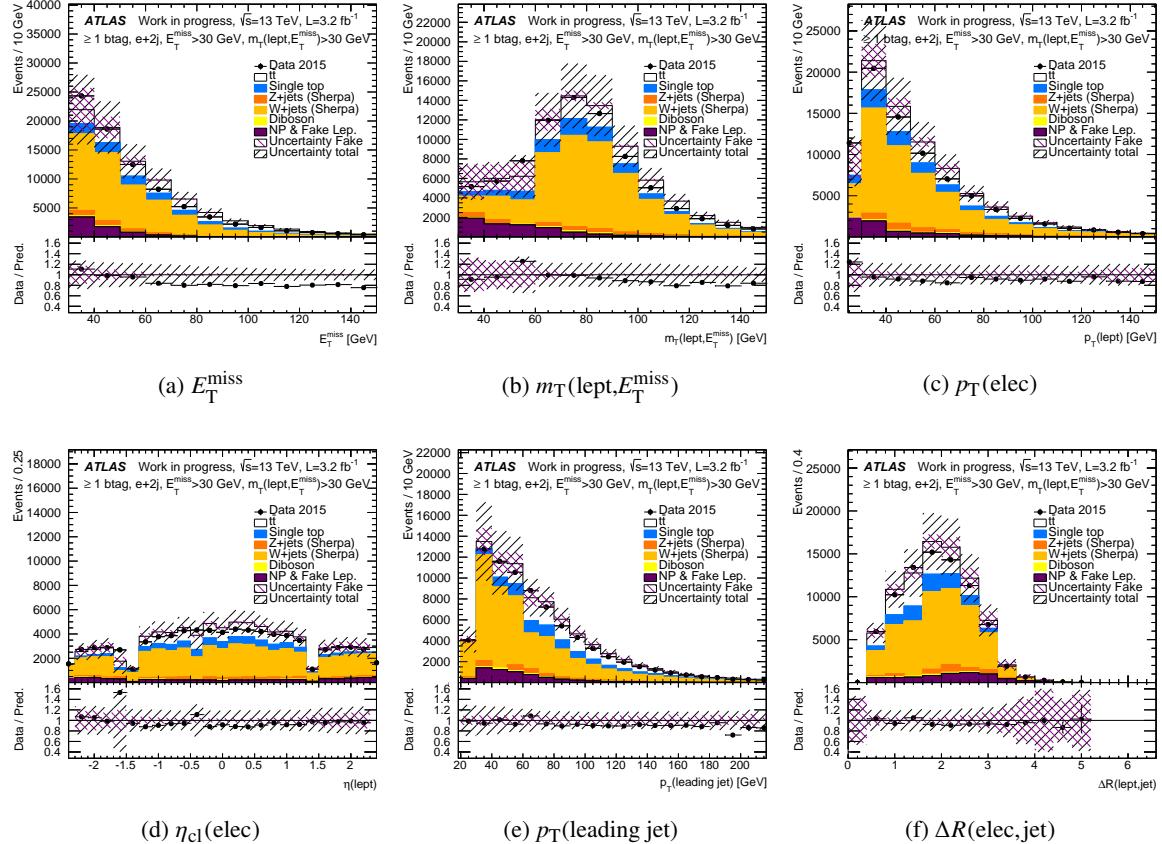


Figure 17: Distributions of (a) E_T^{miss} , (b) $m_T(\text{lept}, E_T^{\text{miss}})$, (c) p_T of the electron, (d) η_{cl} of the electron, (e) p_T of the leading jet and (f) the distance ΔR of the electron to its nearest jet for $e+\text{jets}$ events with two jets, requiring at least one b -jet, requiring $E_T^{\text{miss}} > 30 \text{ GeV}$ and $m_T(\text{lept}, E_T^{\text{miss}}) > 30 \text{ GeV}$. The data is compared to the real lepton expectation from simulation, showing separately the contributions from $t\bar{t}$, single top, $W+\text{jets}$, $Z+\text{jets}$ and dibosons normalised to their cross-sections, and non-prompt and fake lepton backgrounds (referred to as ‘NP & Fake Lep.’) estimated with the matrix method. The purple shaded area represents the combination of the statistical and the systematic uncertainties on the matrix method estimate in each bin. The black shaded area represents the total uncertainty including also the uncertainties on the processes predicted by the MC simulation and the one on the luminosity. The lower parts of the figure show the ratios of expectation to data.

Not reviewed, for internal circulation only

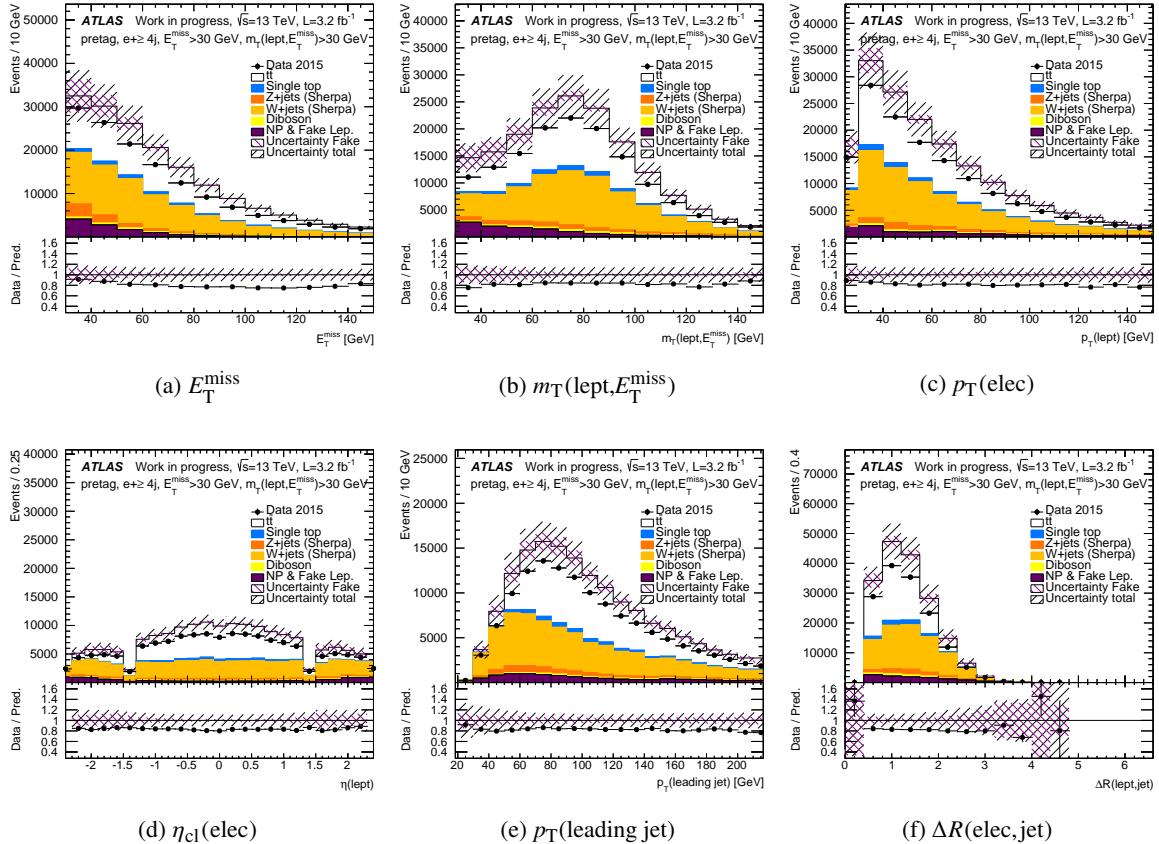


Figure 18: Distributions of (a) E_T^{miss} , (b) $m_T(\text{lept}, E_T^{\text{miss}})$, (c) p_T of the electron, (d) η_{cl} of the electron, (e) p_T of the leading jet and (f) the distance ΔR of the electron to its nearest jet for $e+\text{jets}$ events with at least four jets without any b -tagging requirement, requiring $E_T^{\text{miss}} > 30 \text{ GeV}$ and $m_T(\text{lept}, E_T^{\text{miss}}) > 30 \text{ GeV}$. The data is compared to the real lepton expectation from simulation, showing separately the contributions from $t\bar{t}$, single top, $W+\text{jets}$, $Z+\text{jets}$ and dibosons normalised to their cross-sections, and non-prompt and fake lepton backgrounds (referred to as ‘NP & Fake Lep.’) estimated with the matrix method. The purple shaded area represents the combination of the statistical and the systematic uncertainties on the matrix method estimate in each bin. The black shaded area represents the total uncertainty including also the uncertainties on the processes predicted by the MC simulation and the one on the luminosity. The lower parts of the figure show the ratios of expectation to data.

Not reviewed, for internal circulation only

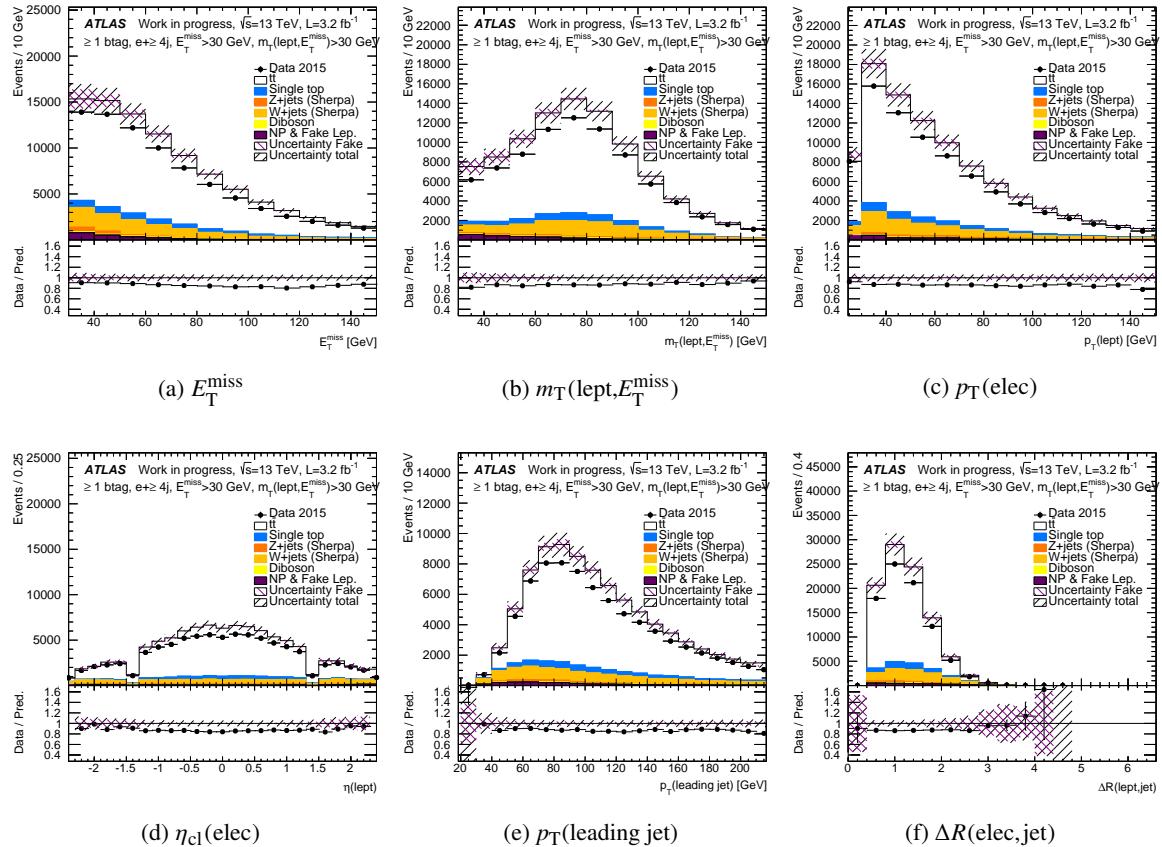


Figure 19: Distributions of (a) E_T^{miss} , (b) $m_T(\text{lept}, E_T^{\text{miss}})$, (c) p_T of the electron, (d) η_{cl} of the electron, (e) p_T of the leading jet and (f) the distance ΔR of the electron to its nearest jet for $e+\text{jets}$ events with at least four jets, requiring at least one b -jet, requiring $E_T^{\text{miss}} > 30 \text{ GeV}$ and $m_T(\text{lept}, E_T^{\text{miss}}) > 30 \text{ GeV}$. The data is compared to the real lepton expectation from simulation, showing separately the contributions from $t\bar{t}$, single top, $W + \text{jets}$, $Z + \text{jets}$ and dibosons normalised to their cross-sections, and non-prompt and fake lepton backgrounds (referred to as ‘NP & Fake Lep.’) estimated with the matrix method. The purple shaded area represents the combination of the statistical and the systematic uncertainties on the matrix method estimate in each bin. The black shaded area represents the total uncertainty including also the uncertainties on the processes predicted by the MC simulation and the one on the luminosity. The lower parts of the figure show the ratios of expectation to data.

Not reviewed, for internal circulation only

559 7.2. Estimates for $\mu+jets$ events

560 Fig. 20 to 23 show, for $\mu+jets$ events, data compared to the fake lepton estimate plus the various real
 561 lepton contributions from simulation, for the following distributions: E_T^{miss} , $m_T(\text{lept}, E_T^{\text{miss}})$, muon p_T and
 562 η , leading jet p_T and distance ΔR between the muon and its nearest jet. Additional plots are shown in
 section E.2.

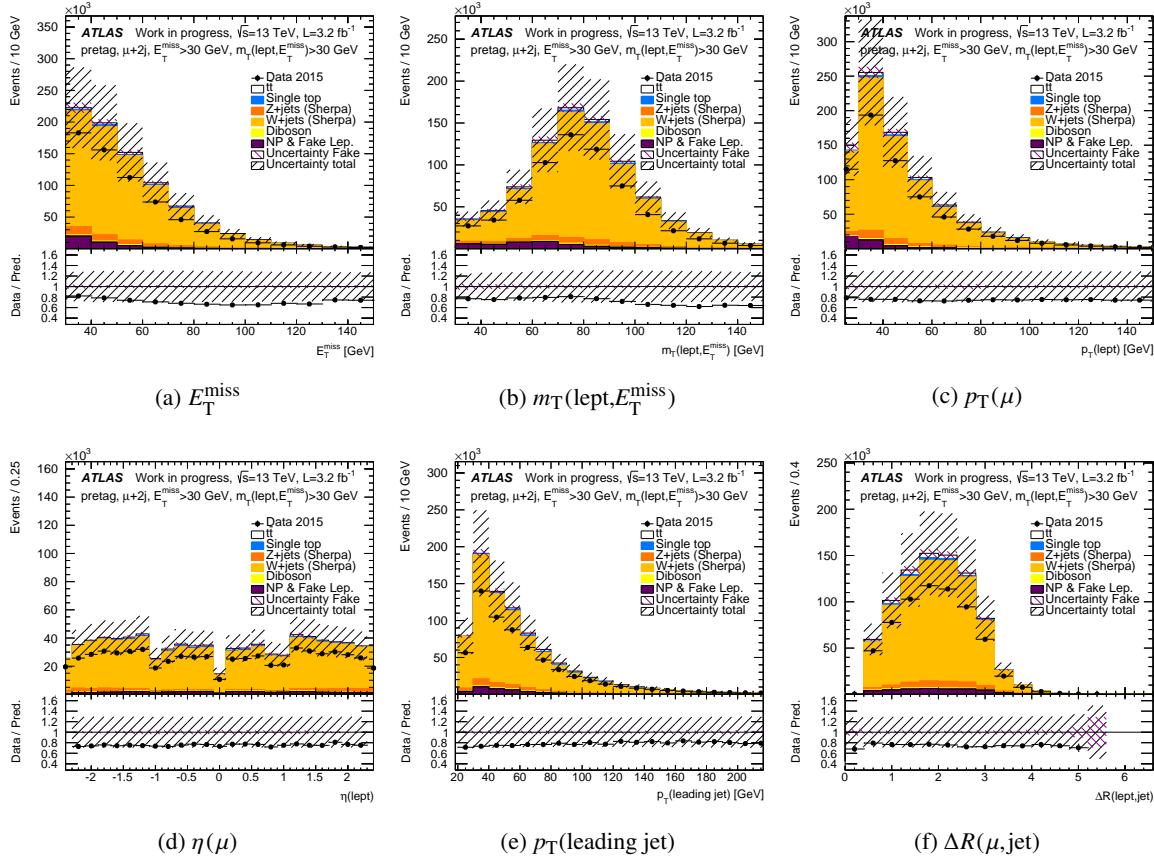


Figure 20: Distributions of (a) E_T^{miss} , (b) $m_T(\text{lept}, E_T^{\text{miss}})$, (c) p_T of the muon, (d) η of the muon, (e) p_T of the leading jet and (f) the distance ΔR of the muon to its nearest jet, for $\mu+jets$ events with two jets without any b -tagging requirement, requiring $E_T^{\text{miss}} > 30 \text{ GeV}$ and $m_T(\text{lept}, E_T^{\text{miss}}) > 30 \text{ GeV}$. The data is compared to the real lepton expectation from simulation, showing separately the contributions from $t\bar{t}$, single top, $W+jets$, $Z+jets$ and dibosons normalised to their cross-sections, and non-prompt and fake lepton backgrounds (referred to as ‘NP & Fake Lep.’) estimated with the matrix method. The purple shaded area represents the combination of the statistical and the systematic uncertainties on the matrix method estimate in each bin. The black shaded area represents the total uncertainty including also the uncertainties on the processes predicted by the MC simulation and the one on the luminosity. The lower parts of the figure show the ratios of expectation to data.

563

Not reviewed, for internal circulation only

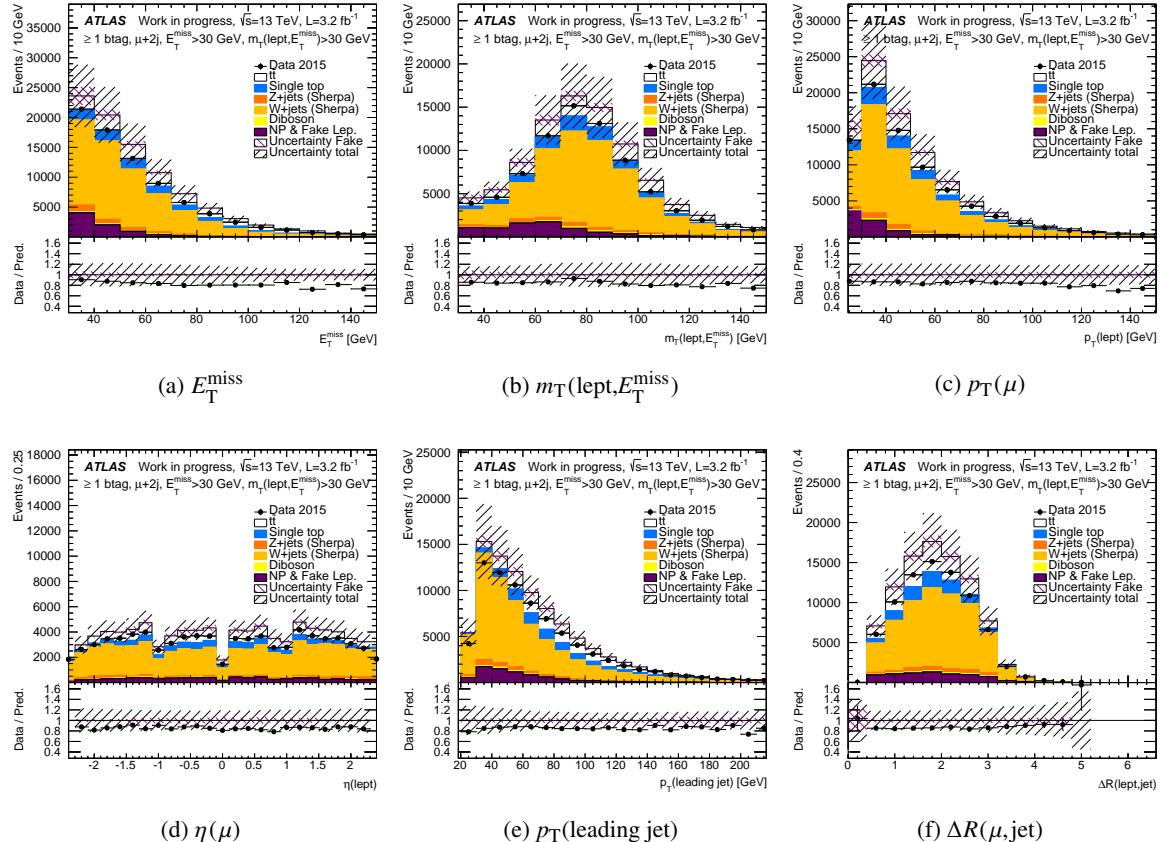


Figure 21: Distributions of (a) E_T^{miss} , (b) $m_T(\text{lept}, E_T^{\text{miss}})$, (c) p_T of the muon, (d) η of the muon, (e) p_T of the leading jet and (f) the distance ΔR of the muon to its nearest jet, for $\mu+\text{jets}$ events with two jets, requiring at least one b -jet, requiring $E_T^{\text{miss}} > 30 \text{ GeV}$ and $m_T(\text{lept}, E_T^{\text{miss}}) > 30 \text{ GeV}$. The data is compared to the real lepton expectation from simulation, showing separately the contributions from $t\bar{t}$, single top, $W+\text{jets}$, $Z+\text{jets}$ and dibosons normalised to their cross-sections, and non-prompt and fake lepton backgrounds (referred to as ‘NP & Fake Lep.’) estimated with the matrix method. The purple shaded area represents the combination of the statistical and the systematic uncertainties on the matrix method estimate in each bin. The black shaded area represents the total uncertainty including also the uncertainties on the processes predicted by the MC simulation and the one on the luminosity. The lower parts of the figure show the ratios of expectation to data.

Not reviewed, for internal circulation only

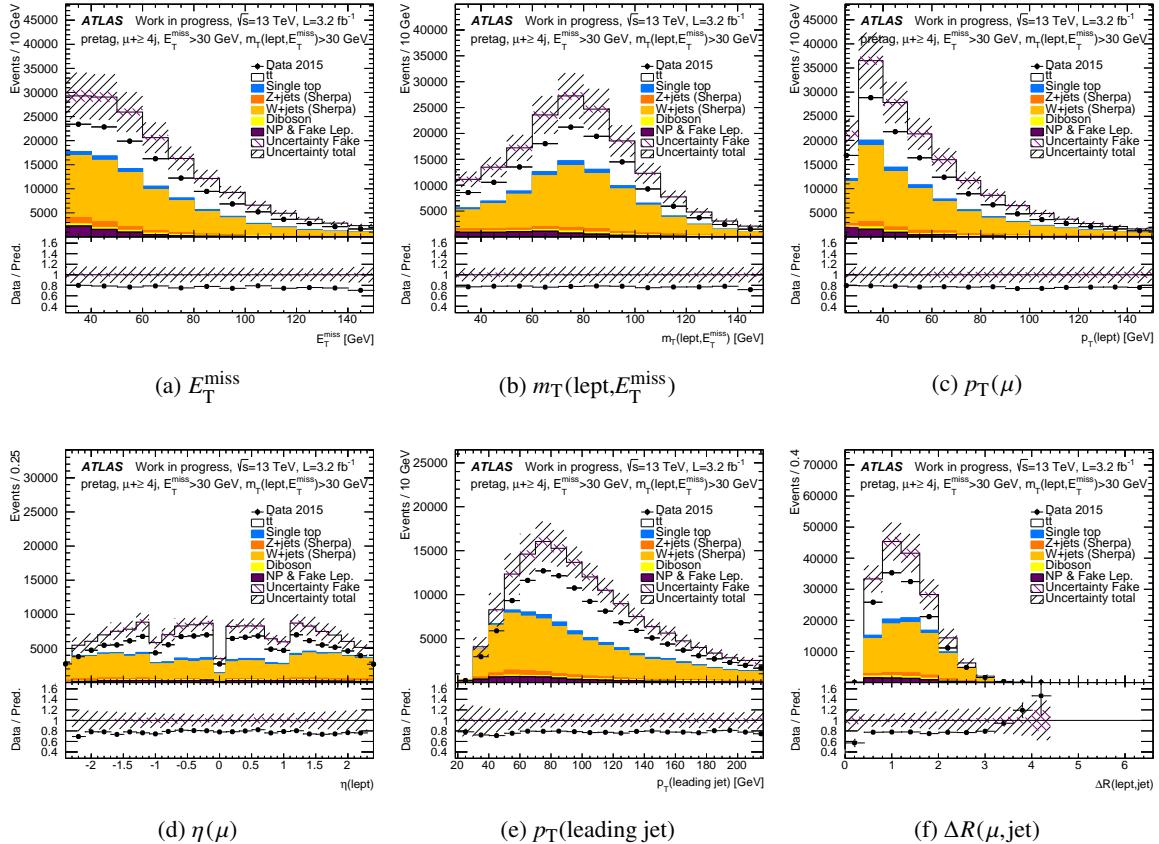


Figure 22: Distributions of (a) E_T^{miss} , (b) $m_T(\text{lept}, E_T^{\text{miss}})$, (c) p_T of the muon, (d) η of the muon, (e) p_T of the leading jet and (f) the distance ΔR of the muon to its nearest jet, for $\mu+\text{jets}$ events with at least four jets without any b -tagging requirement, requiring $E_T^{\text{miss}} > 30$ GeV and $m_T(\text{lept}, E_T^{\text{miss}}) > 30$ GeV. The data is compared to the real lepton expectation from simulation, showing separately the contributions from $t\bar{t}$, single top, $W+\text{jets}$, $Z+\text{jets}$ and dibosons normalised to their cross-sections, and non-prompt and fake lepton backgrounds (referred to as ‘NP & Fake Lep.’) estimated with the matrix method. The purple shaded area represents the combination of the statistical and the systematic uncertainties on the matrix method estimate in each bin. The black shaded area represents the total uncertainty including also the uncertainties on the processes predicted by the MC simulation and the one on the luminosity. The lower parts of the figure show the ratios of expectation to data.

Not reviewed, for internal circulation only

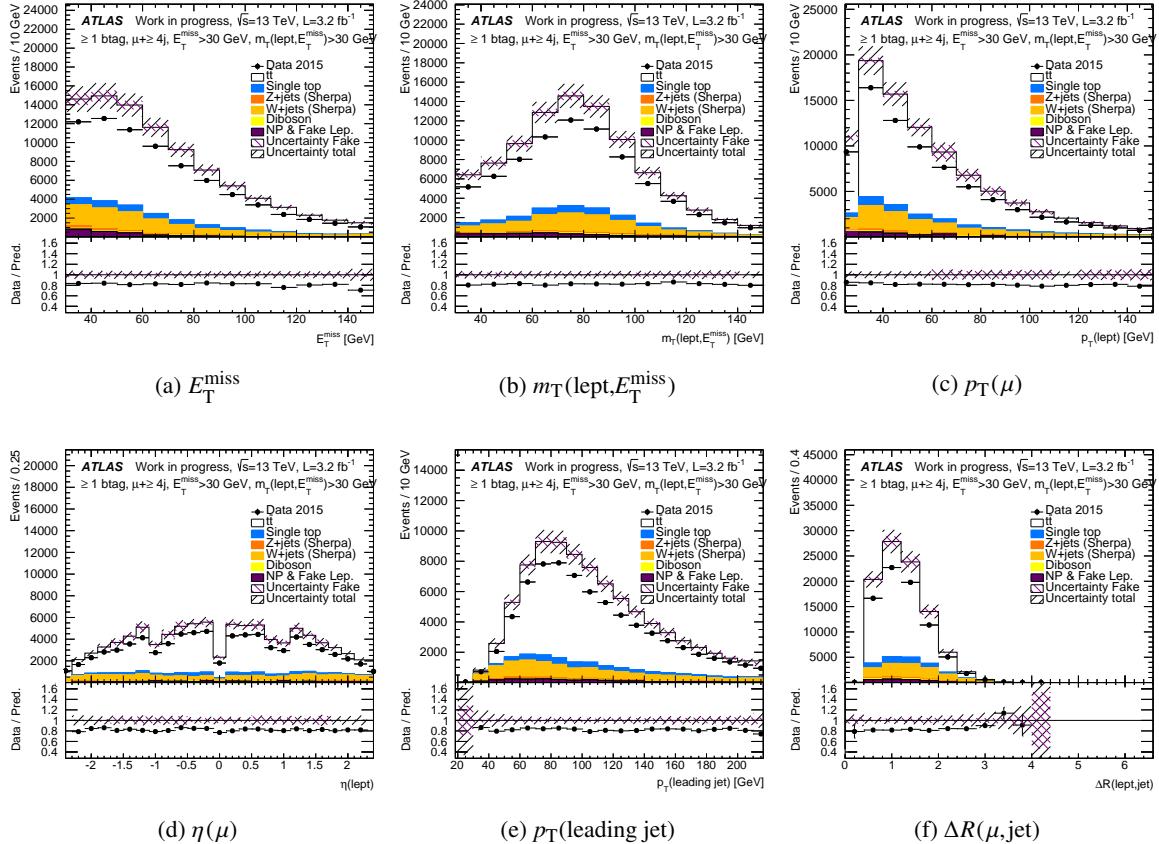


Figure 23: Distributions of (a) E_T^{miss} , (b) $m_T(\text{lept}, E_T^{\text{miss}})$, (c) p_T of the muon, (d) η of the muon, (e) p_T of the leading jet and (f) the distance ΔR of the muon to its nearest jet, for $\mu+\text{jets}$ events with at least four jets, requiring at least one b -jet, requiring $E_T^{\text{miss}} > 30 \text{ GeV}$ and $m_T(\text{lept}, E_T^{\text{miss}}) > 30 \text{ GeV}$. The data is compared to the real lepton expectation from simulation, showing separately the contributions from $t\bar{t}$, single top, $W+\text{jets}$, $Z+\text{jets}$ and dibosons normalised to their cross-sections, and non-prompt and fake lepton backgrounds (referred to as ‘NP & Fake Lep.’) estimated with the matrix method. The purple shaded area represents the combination of the statistical and the systematic uncertainties on the matrix method estimate in each bin. The black shaded area represents the total uncertainty including also the uncertainties on the processes predicted by the MC simulation and the one on the luminosity. The lower parts of the figure show the ratios of expectation to data.

564 7.3. Summary of the results

565 Tables 10 and 11 report the number of selected data events compared to the predicted non-prompt and
 566 fake lepton events from matrix method and real lepton events from MC simulation, after selecting exactly
 567 two or at least four jets, for pretag and ≥ 1 b -tag selection, after the $\ell + \text{jets } t\bar{t}$ selection.

568 Tables 12 and 13 show the systematic variation effects on the non-prompt and fake lepton estimation in
 569 the same regions.

570 The total expected number of events, obtained from the different MC simulated samples for real lepton
 571 and non-prompt and fake lepton backgrounds, always exceed the number of events in the data, as could
 572 already been noticed in most of plots. Results using other MC samples for the $W + \text{jets}$ and $Z + \text{jets}$
 573 backgrounds can be found in section E.1.6 and E.2.6 but none of them is satisfactory.

574 In the $\ell + \text{jets}$ signal regions, the systematic uncertainties on the fake estimates are between 90 and 140%
 575 (50-140%) for $e + \text{jets}$ ($\mu + \text{jets}$) events, depending on the jet and b -jet multiplicity. The use of the real
 576 lepton subtraction from CR_f and the alternative parametrisation "EffPar" are the dominant sources in
 577 $e + \text{jets}$ events, effecting between 40 and 90%. In $\mu + \text{jets}$ events "CRreal", the alternative estimate for $\varepsilon_{\text{real}}$,
 produces a relatively large deviation, from 30 to 120%.

Table 10: Number of events expected from the different MC simulated samples for real lepton and non-prompt and fake lepton backgrounds estimated with the matrix method in $e + \text{jets}$ events, after the $\ell + \text{jets } t\bar{t}$ selection, compared to the number of events observed in data.

$E_T^{\text{miss}} > 30 \text{ GeV} \& m_T(\text{lept}, E_T^{\text{miss}}) > 30 \text{ GeV}$	$e + \text{jets}$	2-jets	2-jets	≥ 4 -jets	≥ 4 -jets
		pretag	≥ 1 b -tags	pretag	≥ 1 b -tags
Data 2015	646526	80064	145927	83949	
Fake Leptons	54890	6655	11153	2232	
(stat+syst)	± 48069	± 7000	± 10030	± 3149	
Fake Fraction	8.5%	8.3%	7.6%	2.7%	
<hr/>					
$t\bar{t}$	24544	18479	85010	76559	
Single Top	12182	8980	6178	5234	
$W + \text{jets}$	629278	46577	64367	10740	
$Z + \text{jets}$	43383	2739	7337	1244	
Diboson	5848	690	2620	622	
<hr/>					
Tot. MC	715235	77466	165512	94402	
Tot. Expected	770125	84121	176666	96635	

578

Table 11: Number of events expected from the different MC simulated samples for real lepton and non-prompt and fake lepton backgrounds estimated with the matrix method in μ +jets events, after the ℓ +jets $t\bar{t}$ selection, compared to the number of events observed in data.

$E_T^{\text{miss}} > 30 \text{ GeV} \& m_T(\text{lept}, E_T^{\text{miss}}) > 30 \text{ GeV}$	μ +jets	2-jets pretag	2-jets ≥ 1 b-tags	≥ 4 -jets pretag	≥ 4 -jets ≥ 1 b-tags
Data 2015	644102	79003	132952	78080	
Fake Leptons	39696	8591	6168	2440	
(stat+syst)	± 24039	± 2769	± 3111	± 3364	
Fake Fraction	6.2%	10.9%	4.6%	3.2%	
$t\bar{t}$	24222	18225	81912	73800	
Single Top	12368	9111	5918	5018	
W +jets	729555	53334	70480	12114	
Z +jets	40718	3098	4929	955	
Diboson	6514	747	2601	603	
Tot. MC	813377	84517	165847	92487	
Tot. Expected	853073	93108	171870	94927	

Table 12: Effect (in %) of the various sources of systematic uncertainty on the fake lepton estimation with the matrix method in e +jets events, after the ℓ +jets $t\bar{t}$ selection.

$E_T^{\text{miss}} > 30 \text{ GeV} \& m_T(\text{lept}, E_T^{\text{miss}}) > 30 \text{ GeV}$	e +jets	2-jets pretag	2-jets ≥ 1 b-tags	≥ 4 -jets pretag	≥ 4 -jets ≥ 1 b-tags
MC up	-28	-24	-47	-55	
MC down	24	34	31	86	
CR real	-1	-8	-19	-31	
CR fake	57	65	63	87	
Alt.Par.	60	74	40	62	
Tot.Syst.	88	105	90	140	

Table 13: Effect (in %) of the various sources of systematic uncertainty on the fake lepton estimation with the matrix method in μ +jets events, after the ℓ +jets $t\bar{t}$ selection.

$E_T^{\text{miss}} > 30 \text{ GeV} \& m_T(\text{lept}, E_T^{\text{miss}}) > 30 \text{ GeV}$	μ +jets	2-jets pretag	2-jets ≥ 1 b-tags	≥ 4 -jets pretag	≥ 4 -jets ≥ 1 b-tags
MC up	8	4	-18	-49	
MC down	28	32	42	64	
CR real	-53	-33	-74	-120	
CR fake	9	-5	17	15	
Alt.Par.	-6	-13	-6	-10	
Tot.Syst.	60	48	88	140	

579 8. Estimates in dilepton $t\bar{t}$ events

580 In this section we consider the non-prompt and fake lepton backgrounds estimates in the dilepton channels
 581 for the same-sign and opposite-sign samples. The method considered here adapts the matrix method to
 582 the case of two leptons.

583 8.1. Measurement and parametrisation of the efficiencies

584 The matrix method applied on $t\bar{t}$ dileptonic events relies on real and fake efficiencies measured in the
 585 same way as in the method for single leptons. The same efficiencies used in the single lepton case can be
 586 used in the dilepton case.

587 In the dilepton analyses, one of these leptons does not necessarily fire one of the main analysis triggers. In
 588 this analysis, the two leptons have $p_T > 25$ GeV, so well above the lowest- p_T trigger threshold. Electron
 589 efficiencies derived in section 6.3 are not matched to a specific trigger and can be applied even for an
 590 electron which does not match one of these triggers. For muons, efficiencies are measured independently
 591 for leptons matched to the triggers used for the selection of events and for leptons which do not match such
 592 trigger. In practice, for this second case, both $\varepsilon_{\text{real}}$ and $\varepsilon_{\text{fake}}$ are measured separately for muons matching
 593 the non-isolated low- p_T trigger HLT_MU20_L1MU15 (labelled as mu20). This trigger is prescaled with
 594 a factor ~ 10 in most of the considered data periods. For this reason, when subtracting it in the control
 595 regions where $\varepsilon_{\text{fake}}$ is measured, the real muon contamination is normalised to the live luminosity of the
 596 trigger.

597 The strategy to apply efficiencies derived on muons matching one of these triggers is summarised in
 598 Table 14 and is the following. If any lepton matches the unbiased high- p_T trigger, then the event selection
 599 is not biased by isolation, so the efficiencies measured with respect to the unbiased triggers are applied. If
 600 no muons match the high- p_T trigger, the biased efficiencies are applied to the leading muon that matched
 601 the low- p_T trigger, and the unbiased efficiencies are applied to all other muons. When unbiased efficien-
 602 cies are used, the choice of what efficiencies to apply depends on the p_T of the muon. For high- p_T muons,
 603 the efficiencies measured with the high- p_T triggers are used. For low- p_T muons, the efficiencies measured
 604 with the prescaled unbiased triggers are used. This is done to allow the use of the efficiencies derived
 from high-statistics unprescaled single-muon samples whenever possible.

Table 14: Efficiencies $\varepsilon_{\text{real}}$ and $\varepsilon_{\text{fake}}$ applied to each muon, depending on the characteristics of the muon and the event trigger. High- and low- p_T are defined here by the boundary between the p_T thresholds of the triggers that require isolation and those that do not. For muons, low- p_T means the range $20 < p_T < 51$ GeV.

Event trigger	p_T rank among trigger-matched leptons	muon p_T range	efficiency applied
≥ 1 muon matches high- p_T trigger	any	high	high- p_T
≥ 1 muon matches high- p_T trigger	any	low	low- p_T unbiased
no muon matches high- p_T trigger	leading	any	low- p_T unbiased
no muon matches high- p_T trigger	subleading	high	high- p_T unbiased
no muon matches high- p_T trigger	subleading	low	low- p_T unbiased

605

606 Efficiencies were shown in the single lepton sections in Fig. 4 and 5 for electrons and Fig. 6 and 7 for
 607 muons. The efficiencies derived with this new muon trigger, with respect to the p_T of the lepton, are

similar to the ones obtained with the high-threshold trigger, above the threshold of this trigger and allows to extend the measurement below it. Due to this extension to lower p_T the efficiencies measured with respect to other variables are then lower.

Compared to the single lepton case, some variables used to parametrise efficiencies are ill-defined. It's the case in particular of $\Delta\phi$ defined for the single lepton case as the difference in azimuth between the lepton and the E_T^{miss} . If in the single lepton case the E_T^{miss} corresponds to the direction of emission of the neutrino produced from the W boson, in the dilepton case two neutrinos are emitted but only one E_T^{miss} direction is available. The default parametrisation of efficiencies is the same as in the single lepton case. Several systematic effects are studied, in a similar way as described for the single lepton case:

- $\varepsilon_{\text{real}}$ is measured also in an alternative region with the high- E_T^{miss} and high- $m_T(\text{lept}, E_T^{\text{miss}})$ methods. This is taken as a source of systematic uncertainties (labelled “CRreal”);
- $\varepsilon_{\text{fake}}$ is measured in another control region. This is taken as a source of systematic uncertainties (labelled as “CRfake”).
- The uncertainty on the amount of real leptons to subtract from the fake control region is changed by $\pm 30\%$. This is taken as a source of systematic uncertainties (labelled “MCup” and “MCdown”).

Compared to the single lepton case, the “EffPar” variation is not considered as based on the $\Delta\phi$ variable.

On the contrary of the 8 TeV analysis [2] no study of the composition of fake leptons was performed. In particular, if the composition was found to be very different to the one of the regions where efficiencies are measured it could lead to a systematic uncertainty, or an extrapolation of efficiencies to this composition.

8.2. Estimates in the opposite-sign sample

The opposite-sign sample is the one used for example in the measurements of the $t\bar{t}$ cross section or top quark mass measurement in the dilepton channels. All selection criteria, as described in section 5.3, are applied.

Fig. 25 shows the distributions of jet multiplicity and E_T^{miss} obtained in events with an opposite-sign lepton pair with at least two jets and without requirement on the number of b -jets. The agreement between data and expectation is not very good even if the low level of non-prompt and fake lepton background reflects essentially the agreement between data and simulation. Additional results are shown in annexe F.1.

636 8.3. Estimates in the same-sign sample

637 For the dilepton same-sign analysis, besides the standard samples, the associated production of $t\bar{t}$ and
 638 bosons, $t\bar{t}W$ +jets, $t\bar{t}Z$ +jets, $t\bar{t}W^+W^-$ +jets, which are described in section 2.2, are also considered. An-
 639 other background which has to be taken into account for this sample is the charge mis-identification. It is
 640 estimated from the different simulated samples with opposite-sign charge leptons such as $t\bar{t}$, single top,
 641 diboson and mostly Z +jets events. As being dominated by this latest, its uncertainty is taken as the one
 642 on the cross section of this process, 34% as reported in section 2.2.

643 Fig. 24 shows the distributions of jet multiplicity and E_T^{miss} obtained in the same-sign samples, for events
 644 with at least one jet at pretag level. The agreement between data and expectation is not very good even if
 645 the low level of non-prompt and fake lepton background reflects essentially the agreement between data
 and simulation. Additional results are shown in annexe F.3.

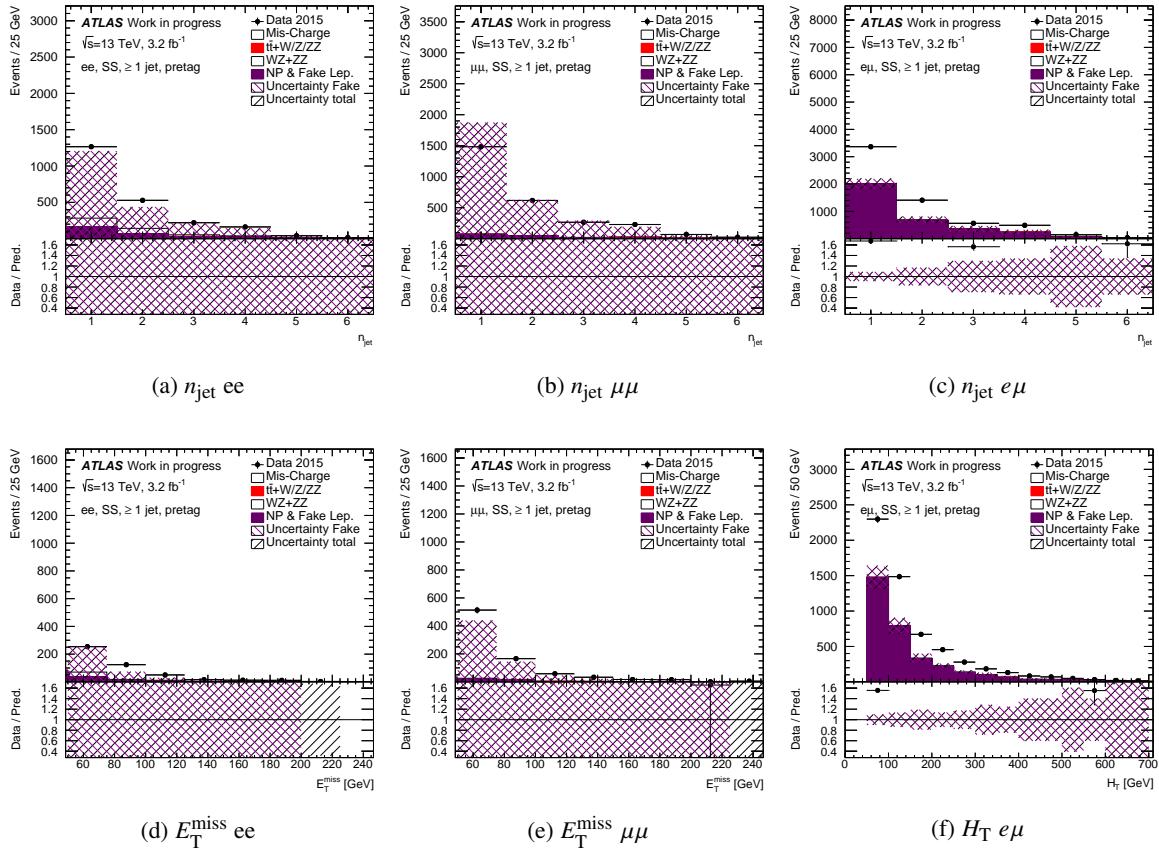


Figure 24: Distributions of (a,b,c) jet multiplicity and (d,e,f) E_T^{miss} in events with a same-sign (a,d) ee , (b,e) $\mu\mu$ and (c,f) $e\mu$ pair with at least two jets and without requirement on the number of b -jets. For $e\mu$ events the E_T^{miss} distribution is replaced by the H_T one. The data is compared to the real lepton expectation from simulation, showing separately the contributions from $t\bar{t}$, single top, Z +jets and dibosons normalised to their cross-sections, and non-prompt and fake lepton backgrounds (referred to as ‘NP & Fake Lep.’) estimated with the matrix method. The purple shaded area represents the combination of the statistical and the systematic uncertainties on the matrix method estimate in each bin. The black shaded area represents the total uncertainty including also the uncertainties on the processes predicted by the MC simulation and the one on the luminosity. The lower parts of the figure show the ratios of expectation to data.

646 8.4. Summary of the results

647 Table 15 shows the fake dilepton yield and the fraction (in %) with respect to the total yield, for the
 648 three channels ee , $\mu\mu$ and $e\mu$, in the opposite-sign charge samples. The yields are given for different jet
 649 and b -jet multiplicities. The yields obtained from the expectations are always much larger than the ones
 650 obtained on data. The fractions of fake leptons are of the order of the percent. The level of systematics
 651 (see Table 16) for the ee channel is $\sim 200\%$ and is dominated by the uncertainty on the level of real
 652 lepton to remove from the fake efficiency (in particular “MCdown”). For the $\mu\mu$ channel the systematic
 653 uncertainty is of the order of 200 to 1000% ! It is also largely dominated by the “MCdown” contribution.
 654 For the $e\mu$ channel the systematic uncertainty is 100-300%, resulting from the ones mentionned above
 for electrons and muons.

Table 15: Number of events expected from the different MC simulated samples for real lepton and non-prompt and fake lepton backgrounds estimated with the matrix method compared to the number of events observed in data. The full dilepton $t\bar{t}$ selection is applied on events with opposite-sign leptons and at least two jets.

$n_{b\text{-jet}}$	ee			$\mu\mu$			$e\mu$		
	pretag	≥ 1 b-tags	≥ 2 b-tags	pretag	≥ 1 b-tags	≥ 2 b-tags	pretag	≥ 1 b-tags	≥ 2 b-tags
Data 2015	5348	4247	2023	8557	6561	3170	25228	20463	10111
Fake Leptons	90.6	46	13	71	33	15	1315	590	200
(stat+syst)	± 180	± 90	± 29	± 370	± 330	± 175	± 1424	± 1176	± 580
Fake Fraction	1.7%	1.1%	0.6%	0.8%	0.5%	0.5%	5.2%	2.9%	2.0%
$t\bar{t}$	5179	4580	2261	8242	7332	3660	25085	22365	11260
Single Top	312	252	80	483	388	122	1458	1184	368
$Z + \text{jets}$	841	132	19	1833	232	38	1205	169	22
Diboson	7	1	1	18	4	1	14	4	1
Tot. MC	6352	4976	2366	10595	7972	3829	27809	23763	11673
Tot. Expected	6442	5022	2379	10667	8005	3843	29124	24354	11873

Table 16: Effect (in %) of the various sources of systematic uncertainty on the fake lepton estimation with the matrix method. The full dilepton $t\bar{t}$ selection is applied on events with opposite-sign leptons and at least two jets.

$n_{b\text{-jet}}$	ee			$\mu\mu$			$e\mu$		
	pretag	≥ 1 b-tags	≥ 2 b-tags	pretag	≥ 1 b-tags	≥ 2 b-tags	pretag	≥ 1 b-tags	≥ 2 b-tags
MC up	-58	-60	-60	-130	-310	-380	-22	-50	-80
MC down	180	190	220	104	31	-240	-26	-12	-30
CR real	50	1	57	420	920	1130	104	190	275
CR fake	58	50	50	280	140	52	14	25	40
Tot.Syst.	200	200	230	500	1000	110	200	290	

Not reviewed, for internal circulation only

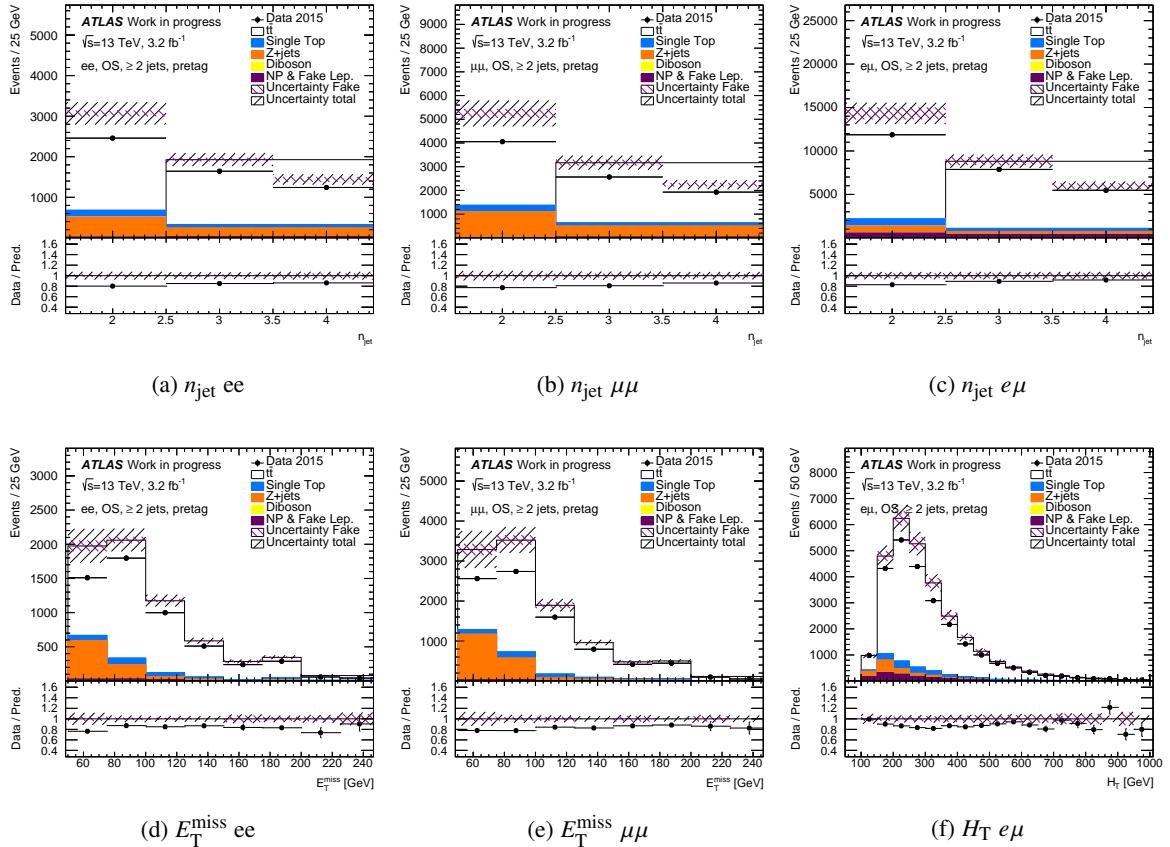


Figure 25: Distributions of (a,b,c) jet multiplicity and (d,e,f) E_T^{miss} in events with an opposite-sign (a,d) ee , (b,e) $\mu\mu$ and (c,f) $e\mu$ pair with at least two jets and without requirement on the number of b -jets. For $e\mu$ events the E_T^{miss} distribution is replaced by the H_T one. The data is compared to the real lepton expectation from simulation, showing separately the contributions from $t\bar{t}$, single top, $Z + \text{jets}$ and dibosons normalised to their cross-sections, and non-prompt and fake lepton backgrounds (referred to as ‘NP & Fake Lep.’) estimated with the matrix method. The purple shaded area represents the combination of the statistical and the systematic uncertainties on the matrix method estimate in each bin. The black shaded area represents the total uncertainty including also the uncertainties on the processes predicted by the MC simulation and the one on the luminosity. The lower parts of the figure show the ratios of expectation to data.

656 9. To go further

657 The implementation of the matrix method as available in the package TopFakes-00-00-08 aimed to re-
 658 produce the one developped for the Run 1 analysis [1, 2]. Still, the accuracy obtained on the data over
 659 expectation is much worse in the 13 TeV analysis than in the 8 TeV one. Here are few ideas to go further.

660 In general, on all results the ratio data over expectation is below one, often by 10-20%. The following
 661 remarks can be done :

- 662 - Different choices of $W+jets$ and $Z+jets$ samples were tested, based on **SHERPA** (the default), **POWHEG**
 663 +**PYTHIA** or **MADGRAPH** +**PYTHIA** and control plots can indeed differ. But the problem exists also in
 664 regions dominated by $t\bar{t}$ events. If a normalization issue appears in the analysis, or better samples
 665 are available, this could ameliorate the control plots. It could also affect the fake efficiency meas-
 666 urements, even if it was found little dependence on the choice of the $W+jets$ and $Z+jets$ samples.
- 667 - In this analysis no $W+jets$ scale factor are used.
- 668 - In a “real” analysis some backgrounds taken here from simulation, as the $Z+jets$ or $W+jets$, are
 669 estimated directly from the data. Doing something similar in the efficiency measurements or in the
 670 control plots could ameliorate the results.

671 The object definition and the loose/tight definitions could be changed :

- 672 - The lepton definition includes requirements on the transverse impact parameter significance. This
 673 affects in particular the fake muon definition which cannot be measured anymore in a region with
 674 high such value as done in the 8 TeV analysis⁶.
- 675 - The Loose definition could be looser. In particular for muons, the loose muon definition could be
 676 defined based on **LooseID** and not **MediumID**. It could enlarge the sample statistics, which is low
 677 up to now, and lower the fake efficiencies, which are high, in particular for the mu50 trigger. Lower
 678 fake efficiencies could help the matrix inversion even if the composition of the loose sample could
 679 start to be very different from the tight one, which is a problem to apply the derived efficiencies.

680 Concerning the real efficiency measurements, the following remarks can be done :

- 681 - The default method is a tag-and-probe technique. Such methods exist in the context of the Com-
 682 bined Performance Working Group. We never did careful comparison of the different methods, or
 683 tried to use directly their measurements. It might be possible to rely on their measurements on data
 684 and only apply some scale factors on simulation to extrapolate to the composition of our samples.
- 685 - The systematic uncertainties are based using an alternative method, measuring efficiencies in high-
 686 E_T^{miss} or $m_T(\text{lept}, E_T^{\text{miss}})$ regions. Differences seen with the efficiencies derived with the tag-and-
 687 probe method were in large part (but not entirely) explained in the 8 TeV analysis. In the 13 TeV
 688 analysis the available statistics is much lower and differences are higher. It is the case in particular
 689 in the $\mu+jets$ events which can be as large as 10% at low p_T . This could be due to a contamination
 690 of the sample with fake muons. Another control region, with higher cut on $m_T(\text{lept}, E_T^{\text{miss}})$, or based
 691 on another variable should be considered. All this lead to very high systematic uncertainties due to
 692 the “CRreal” in the $\mu+jets$ events, at a level that one should consider if they are relevant at the time
 693 being.

⁶ Some developments are done inside the working group to measure the efficiencies without this requirement and then to apply it, after extrapolation, to the leptons which include this definition.

- No optimisation of the cuts applied to define the “CRreal” regions was redone.
- Compared to the 8 TeV analysis, due to the lower statistics, the binning in n_{jet} and $n_{b\text{-jet}}$ has been reduced.
- One could consider also measurements of real efficiencies based only on simulated data. This is partially studied to derived a mean scale factor to be applied on the tag-and-probe results.

Concerning the fake efficiency measurements, the following remarks can be done :

- The fake electron efficiencies were not too hard to estimate. On the contrary the fake muon caused problems. They have to be defined in a similar control region at low- E_T^{miss} and $m_T(\text{lept}, E_T^{\text{miss}})$ as for electrons . Such definition was avoided in the Run 1 as it lead to high fake efficiencies (which is a problem for the method) and uncertainties. The same issue appear here. The normalization problem mentionned above (there seem to be too much MC) was a problem; for example at high jet multiplicities there were more MC than data and efficiencies were not measurable.
- The “MCup/down” systematic uncertainties are still based on a 30% uncertainty on the $W + \text{jets}$ and $Z + \text{jets}$. This lead to the relatively high systematic uncertainties on the fake estimates in the $l + \text{jets}$ events. It is not clear why this lead to extremely high uncertainties in the dilepton channels, in particular the “MCdown” one. Probably here also the normalization issue with MC seen in this analysis is a problem.
- This analysis does not use scale factors for $W + \text{jets}$ events. Once such factors are available they will enter in the measurement of fake efficiencies which need to remove the real lepton contamination due to such events. The issue is that the method to derive the scale factors need a fake measurement first. Some studies are needed either to derive both at same time or one after each other in an interative process.
- As for the 8 TeV analysis, and on the contrary of 7 TeV analyses, there is no uncertainty based on the fake composition. This is important in particular for the $e + \text{jets}$ events. The study of the composition was not even performed, in large part because of the difficulty to determine at reconstruction level if an electron is originated from a conversion or not - such variable is not easely available.

-

Concerning the dilepton results :

- Results obtained in the opposite-sign dilepton events show too much MC once again.
- The results in the $\mu\mu$ channel, in particular, show little amount of fakes and negative yields exist in many bins. Better results will need first to be obtained in the $\mu + \text{jets}$ events.
- Systematic uncertainties are very high, much higher than in the 8 TeV analysis. Here also the results should be ameliorated first in the $l + \text{jets}$ events.
- Results obtained in the same-sign dilepton events are shown for completeness but are not good at all.
- Efficiencies are applied considering that efficiencies (and so for example) of the two leptons are not correlated. This could be studied and may be improved.

Besides all these remarks, the method developped for the 8 TeV analysis could also be ameliorated as already often detailed in many presentations.

733 10. Conclusion

734 The matrix method for estimating non-prompt and fake lepton backgrounds for top analyses is presented
735 using 3.2 fb^{-1} of data collected in 2015 by the ATLAS detector with proton-proton collisions at the LHC.
736 It uses the software AnalysisTop-2.3.41 and corresponds to the default setup of the ToFakes-00-00-08
737 package. Final states with lepton+jets and dilepton events are considered. The matrix method is based
738 on the measurement of efficiencies of leptons with relaxed identification criteria. For final states with one
739 lepton, the systematic uncertainties on the fake estimates in the signal region are 90-140% (50-140%) for
740 $e+\text{jets}$ ($\mu+\text{jets}$) events, depending on the jet and b -jet multiplicity. For final states with two leptons, the
741 systematic uncertainties on the fake estimates in the $e\mu$ opposite-sign signal region are 100-300%.

742 Appendix

743 A. Additional studies on real electron efficiencies

744 In this section we show additional studies done on the electron real efficiency $\varepsilon_{\text{real}}$ measurement.

745 A.1. Variables used for the parametrisation of the efficiencies

746 The main part of the note is describing results obtained with the final parametrisation, but in order to take
747 into account the dependence of the efficiencies, more variables are considered:

- 748 - The electron transverse energy E_T .
- 749 - The electron pseudorapidity η_{cl} as obtained from the cluster information.
- 750 - The number of jets in the event.
- 751 - The number of b -jets in the event.
- 752 - The transverse momentum p_T of the leading jet;
- 753 - The distance ΔR between the electron and the nearest jet⁷.
- 754 - The distance in azimuth $\Delta\phi$ between the electron and the E_T^{miss} .
- 755 - The missing transverse energy E_T^{miss} .
- 756 - The transverse mass $m_T(\text{lept}, E_T^{\text{miss}})$.

757 A.2. Studies with true electrons

758 Electron identification efficiency is studied with simulated data in $t\bar{t}$ events but also single top, Z +jets and
759 W +jets events. Prompt electrons are selected using the MCTruthClassifier tool.

760 A.2.1. Modeling

761 Fig. 26 shows the kinematics of electrons in $t\bar{t}$, single top Z and W simulated data. Electrons from $t\bar{t}$ have
762 harder p_T spectrum and are more central. It was shown in the 8 TeV analysis [1] that electrons from Z
763 events with high number of partons in final states, tend to have similar kinematics as electrons from $t\bar{t}$.
764 Identification efficiency being dependent on the kinematic of the electrons, these differences can easily
765 be taken into account with a good parametrisation of the efficiencies.

766 Fig. 27 shows the isolation variables $E_T^{0.2}$ entering in the definition of the tight objects, the p_T of the
767 leading jet and the distance ΔR between the electron and its nearest jet. Electrons from Z +jets decays are
768 more isolated than the ones from $t\bar{t}$, due to the smaller jet activities in the event. They also have a softer

⁷ This variable is an indicator of the hadronic activity near the electron. Still one has to remind that this does not capture activity below the jet reconstruction p_T threshold, and is distorted by the use of electron-jet overlap vetoes in the electron selection.

Not reviewed, for internal circulation only

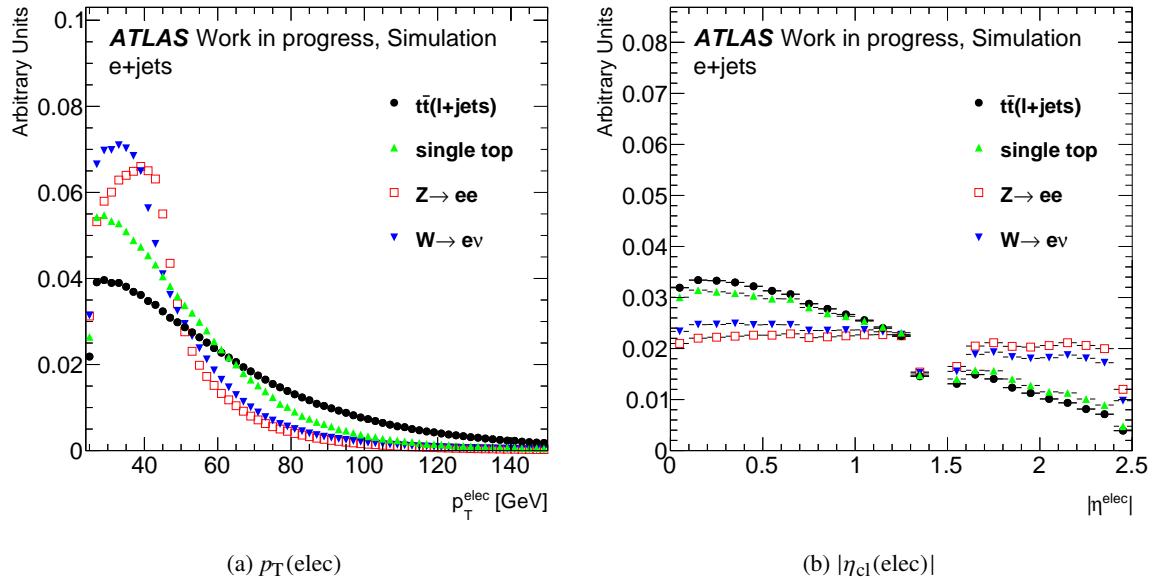


Figure 26: Distributions of (a) transverse momentum p_T and (b) pseudorapidity $|\eta_{\text{cl}}|$ for prompt electrons in $t\bar{t}$, single top, $Z + \text{jets}$ and $W + \text{jets}$ simulated data.

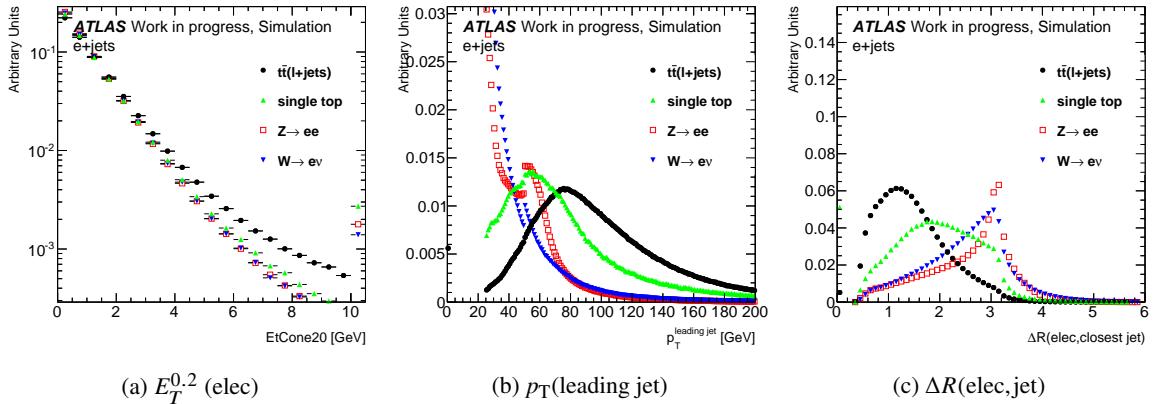


Figure 27: Distributions of (a) $E_T^{0.2}$, (b) p_T of the leading jet and (c) the distance ΔR between the electron and its nearest jet, for prompt electrons in $t\bar{t}$, single top, $Z + \text{jets}$ and $W + \text{jets}$ simulated data.

769 p_T spectrum of their leading jet. Finally, electrons from $Z + \text{jets}$ events are mostly back to back to this
 770 jet, except in events with large parton number in which case the distribution peak towards one, with very
 771 close jets (see Ref. [1]).

772 Finally, Fig. 28 shows the distribution of the distance in azimuth between the electron and the E_T^{miss} , the
 773 E_T^{miss} and the $m_T(\text{lept}, E_T^{\text{miss}})$ for the different simulated samples.

Not reviewed, for internal circulation only

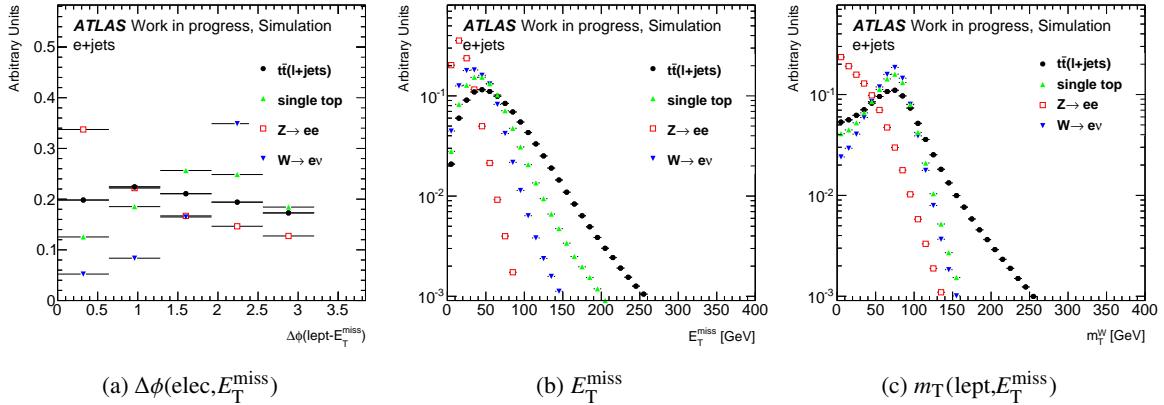


Figure 28: Distributions of (a) the distance $\Delta\phi(\text{elec}, E_{\text{T}}^{\text{miss}})$ in azimuth between the electron and the $E_{\text{T}}^{\text{miss}}$, (b) the $E_{\text{T}}^{\text{miss}}$ and (c) the $m_{\text{T}}(\text{lept}, E_{\text{T}}^{\text{miss}})$, for prompt electrons in $t\bar{t}$, single top, $Z + \text{jets}$ and $W + \text{jets}$ simulated data.

774 A.2.2. True efficiencies

775 Electrons from $t\bar{t}$ events have different kinematics and jet activity around them than the ones in $Z + \text{jets}$
 776 events which turn in different real efficiencies. True efficiencies are measured in the different simulated samples as the ratio of tight over loose distributions.

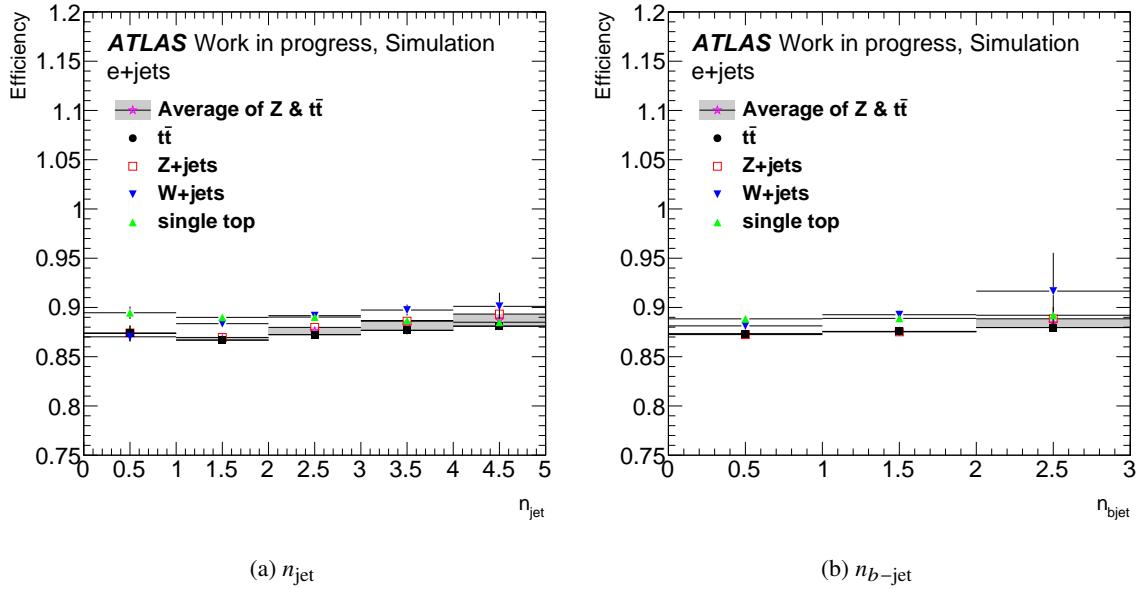


Figure 29: True electron real efficiency as a function of (a) the number of jets n_{jet} and (b) the number of b -jets $n_{b\text{-jet}}$. Efficiencies are shown for events with at least one jet in $t\bar{t}$, single top, $Z + \text{jets}$ and $W + \text{jets}$ simulated data. The purple star and the shaded area correspond respectively to the average and to the spread of the efficiencies measured from Z and $t\bar{t}$ events.

777

Not reviewed, for internal circulation only

778 Fig. 29 shows the true real efficiency as a function of the jet and b -jet multiplicities. Differences between
 779 efficiencies from $t\bar{t}$ and $Z + \text{jets}$ are small, but increase with n_{jet} and $n_{b-\text{jet}}$. Larger differences are seen
 780 in particular with electrons from single top events. Fig. 30 shows the true electron real efficiency as a
 781 function of the different variables. Differences of efficiencies between the $t\bar{t}$ and the $Z + \text{jets}$ samples can
 reach few %.

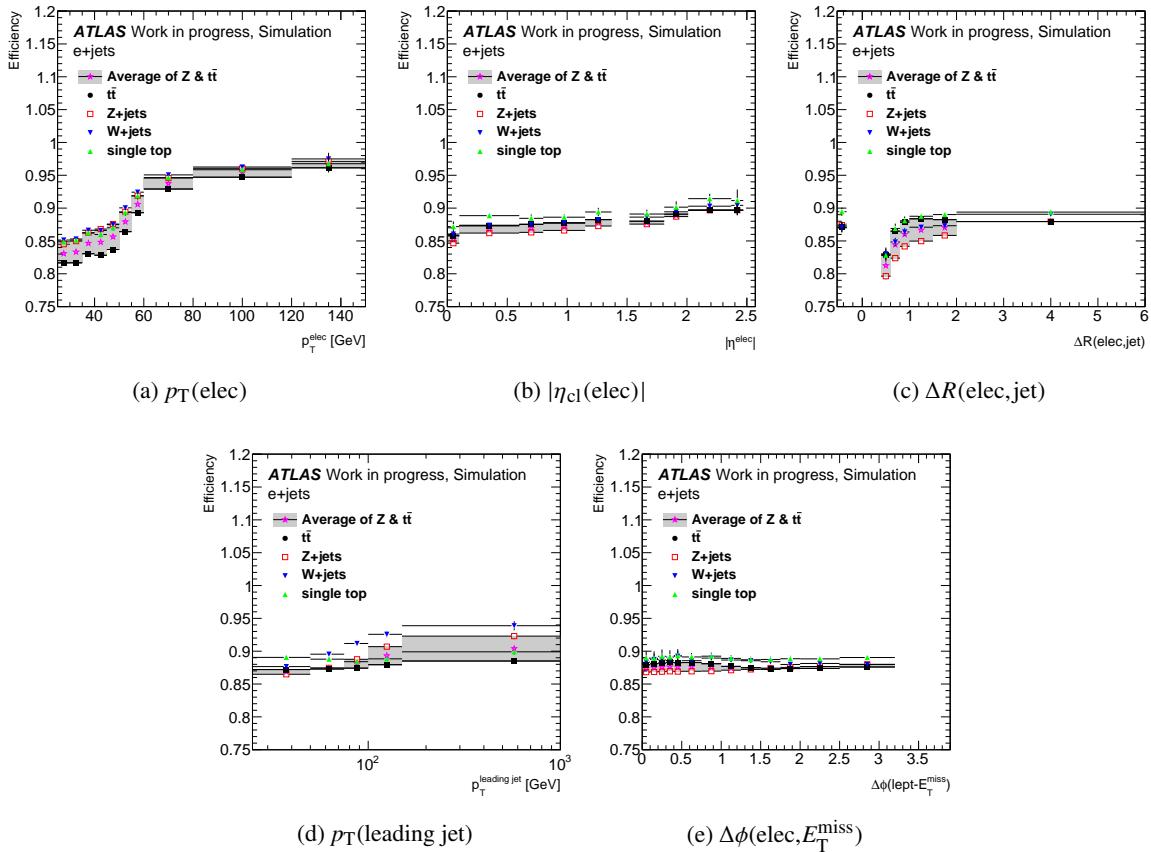


Figure 30: True electron real efficiency as a function of (a) the electron p_T , (b) the electron $|\eta_{\text{el}}|$, (c) the distance ΔR of the electron to its nearest jet, (d) the p_T of the leading jet and (f) the difference $\Delta\phi(\text{elec}, E_T^{\text{miss}})$ in azimuth between the electron and the E_T^{miss} . Efficiencies are shown for events with at least one jet in $t\bar{t}$, single top, $Z + \text{jets}$ and $W + \text{jets}$ simulated data. The purple star and the shaded area correspond respectively to the average and to the spread of the efficiencies measured from Z and $t\bar{t}$ events.

782

783 A.2.3. Correction factor

784 To correct from the previously seen differences, correction factors are derived from the MC samples
 785 which will be applied later to the data measured efficiencies. The central value for the correction is taken
 786 as the average of the $Z + \text{jets}$ ($\varepsilon(Z)$) and $t\bar{t}$ ($\varepsilon(t\bar{t})$) efficiencies, divided by $\varepsilon(Z)$. These corrections factors
 787 are shown in Fig. 31 as a function of the different variables. The spread between $\varepsilon(Z)$ and $\varepsilon(t\bar{t})$ is taken
 788 as a systematic uncertainty attributed to the difference of isolation and the unknown real electron sample

Not reviewed, for internal circulation only

789 composition. These correction factors are much smaller than the ones obtained in the 8 TeV analysis [1].

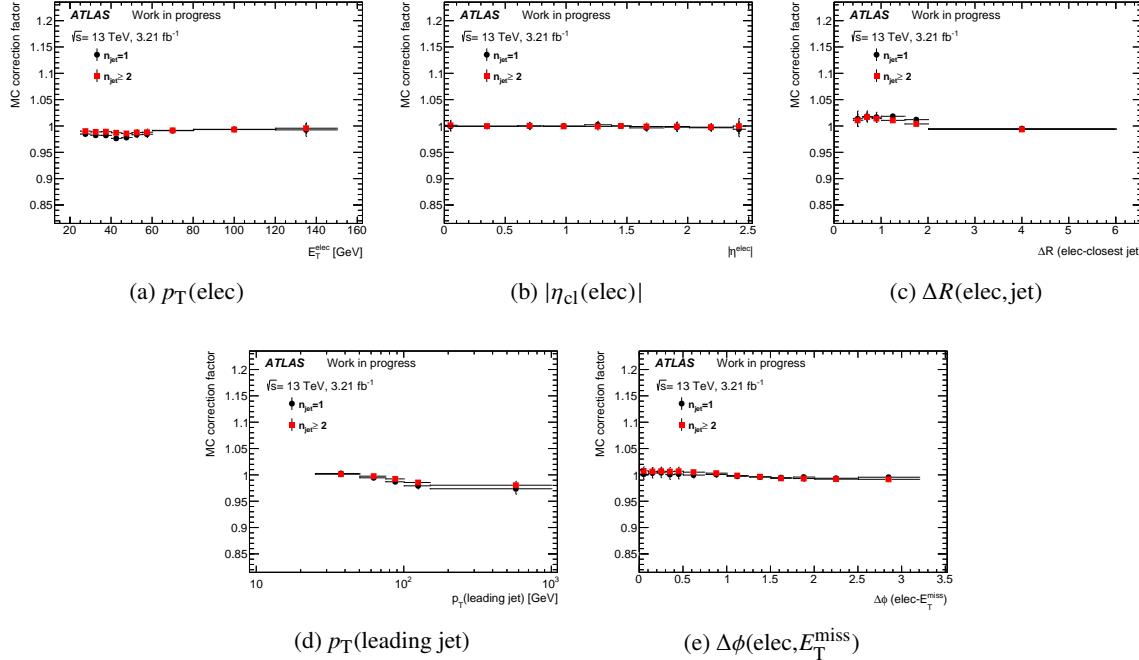


Figure 31: Monte Carlo correction factor to be applied on real electron efficiencies as a function of (a) the electron E_T , (b) the electron $|\eta_{\text{cl}}|$, (c) the distance ΔR between the electron and its nearest jet, (d) the p_T of the leading jet and (e) the difference $\Delta\phi(\text{elec}, E_T^{\text{miss}})$ in azimuth between the electron and the E_T^{miss} . The correction factor is shown for events with at least one jet.

790

791 A.3. The tag-and-probe method

792 The tag-and-probe method makes use of the characteristic signature of $Z \rightarrow ee$ decays. Main results
 793 obtained with this method were already presented in section 6.3.1. Additional material is presented
 794 here.

795 A.3.1. Background subtraction

796 After this selection, the sample still contains non-prompt and fake lepton backgrounds. The level of
 797 background in these Z samples is below the percent level. Measurement of efficiencies is not sensitive to
 798 the details of background subtraction methods and it is estimated. Still, several methods are considered:

- 799 - Side-band method which is the default one. This method relies on the background having a linear
 800 shape over the considered invariant mass range. The invariant mass distributions for opposite-sign
 801 and same-sign pairs at the denominator and numerator levels are divided in three regions A, B
 802 and C. The number of background events in region B and its uncertainty are estimated from the
 803 extrapolation of the side-bands A and C of the same-sign distribution.

Not reviewed, for internal circulation only

- Removal of same-sign (SS) events from opposite-sign ones (OS) in the signal region.
- A two-component fit with a signal contribution plus a background contribution is performed in each bin to the invariant mass m_{ee} distribution. The signal contribution is modeled by a Breit-Wigner distribution convolved with a parametrisation of the low-mass tail, arising mostly from material effects, by a Crystal Ball function. For the background contribution a variety of fit functions are considered such as an exponential and a single-sided exponential convolved with a Gaussian.
- The main uncertainty on the efficiency measurements are linked to the contamination of the sample of probes by background. To assess the precision of its subtraction a total of 18 variations is considered :
- Three different sideband regions (in GeV) : 1) A=[61-81], B=[81-101], C=[101-121], 2) A=[66-76], B=[76-106], C=[106-116], 3) A=[61-81], B=[86-96], C=[101-121].
 - Three different regions for the same sign method (in GeV) : [81-101], [76-106], [86-96].
 - For the fit method there is only one model for the signal function and two for the background one. For each there are two fit ranges (in GeV) : [60-120], [55-200] and three windows to measure the yields : [81-101], [76-106], [86-96].
- The central value and the spread (rms) are determined for all these measurements. Typically, in E_T bins, whereas the statistical uncertainty is 0.2%, the spread of these measurements is $\sim 0.5\%$ in most bins, except for $25 < E_T < 30$ GeV bin, which has the highest level of background, where it reaches 1%.

821 A.3.2. Results obtained on Z data

822 Fig. 32 shows $\varepsilon_{\text{real}}$ as a function of the jet and b -jet multiplicities. It was shown in section A.2.2 that such
 823 dependence with the jet multiplicity is expected in Z events. On the contrary, efficiencies is not expected
 to depend on the number of b -jets. Fig. 33 shows $\varepsilon_{\text{real}}$ as a function of the different variables for all probe

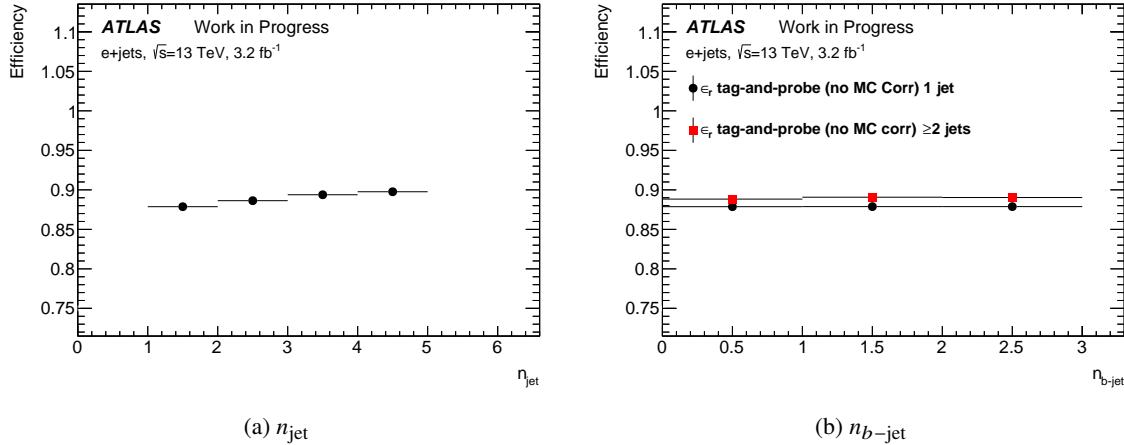


Figure 32: Electron real efficiency $\varepsilon_{\text{real}}$ obtained from the tag-and-probe method on data as a function of (a) the number of jets n_{jet} and (b) the number of b -jets $n_{b\text{-jet}}$. $\varepsilon_{\text{real}}$ is shown for all probes which match one of the single electron trigger (see section 5).

Not reviewed, for internal circulation only

825 electrons which match specifically one of the single electron triggers. Efficiencies are derived for two jet
 826 multiplicities, $n_{\text{jet}} = 1$ and $n_{\text{jet}} \geq 2$ ⁸. A strong dependency is visible with the E_T of the electron. The
 827 dependency with respect to $|\eta_{\text{cl}}|$ of the electron is very small. Some dependency is seen also with respect
 828 to the transverse momentum of the leading jet and the distance $\Delta R(\text{elec}, \text{jet})$. Finally, a small dependence
 is seen with respect to the distance in azimuth between the electron and the E_T^{miss} .

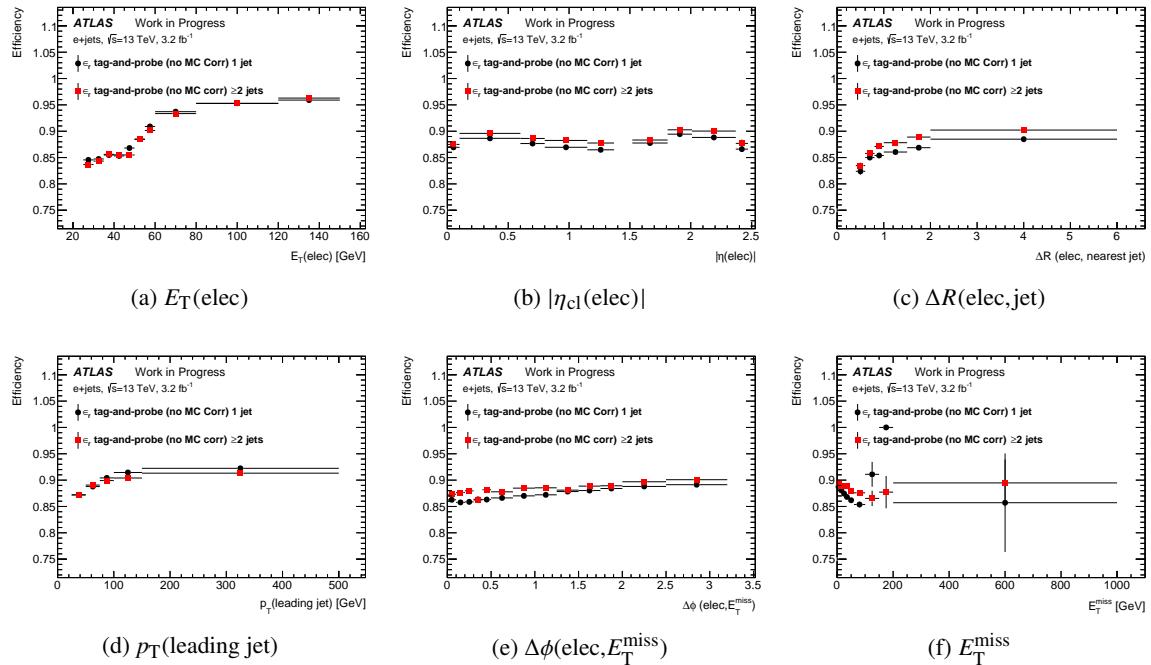


Figure 33: Electron real efficiency $\varepsilon_{\text{real}}$ obtained from the tag-and-probe method on data as a function of (a) the electron E_T , (b) the electron $|\eta_{\text{cl}}|$, (c) the distance ΔR between the electron and its nearest jet, (d) the p_T of the leading jet and (e) the difference $\Delta\phi(\text{elec}, E_T^{\text{miss}})$ in azimuth between the electron and the E_T^{miss} . $\varepsilon_{\text{real}}$ is shown for events with different jet multiplicities, for all probes which match one of the single electron trigger (see section 5).

829

830 A.3.3. Results obtained on Z data corrected by Monte Carlo

831 The final estimates are done from the $\varepsilon_{\text{real}}$ obtained from the tag-and-probe method on $Z \rightarrow e^+ e^-$ data
 832 multiplied by the correction obtained from simulation to correct for the differences between electrons
 833 from $Z + \text{jets}$ and $t\bar{t}$ events. Fig. 32 shows $\varepsilon_{\text{real}}$ as a function of the jet and b -jet multiplicities. It was
 834 shown in section A.2.2 that such dependence with the jet multiplicity is expected in Z events. On the
 835 contrary, efficiencies is not expected to depend on the number of b -jets. Fig. 35 shows these final $\varepsilon_{\text{real}}$ as
 836 a function of different variables for all probe electrons which match one of the single electron trigger (see
 837 section 5).

⁸ For 8 TeV analyses one was using three jet multiplicity bins, $n_{\text{jet}} = 1$, $n_{\text{jet}} = 2 - 3$ and $n_{\text{jet}} \geq 4$, but it was found to be difficult to do for 13 TeV analysis due to lack of statistics for the $n_{\text{jet}} \geq 4$ bin.

Not reviewed, for internal circulation only

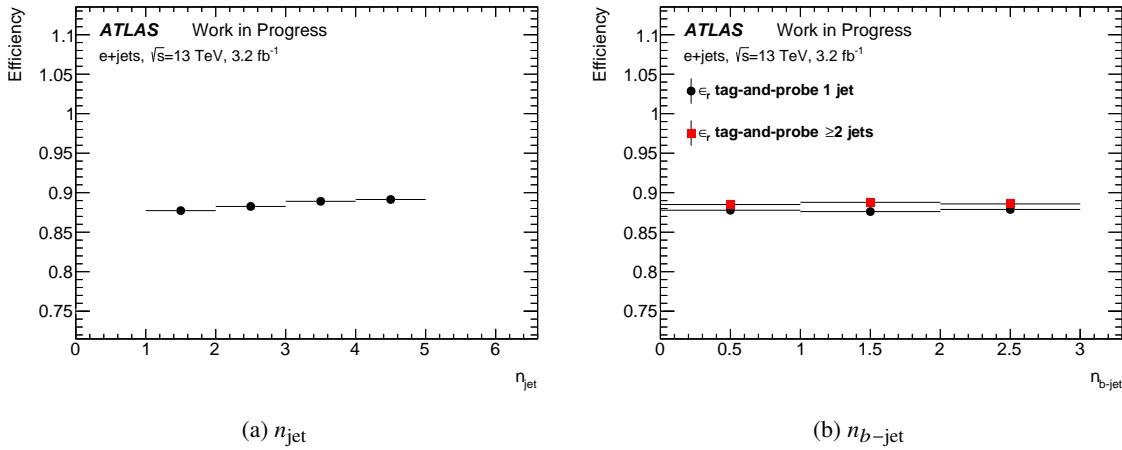


Figure 34: Electron real efficiency $\varepsilon_{\text{real}}$ obtained from the tag-and-probe method on data, multiplied by the MC correction factor, as a function of (a) the number of jets n_{jet} and (b) the number of b -jets $n_{b\text{-jet}}$. $\varepsilon_{\text{real}}$ is shown for events with only one jet, for all probes which match one of the single electron trigger (see section 5).

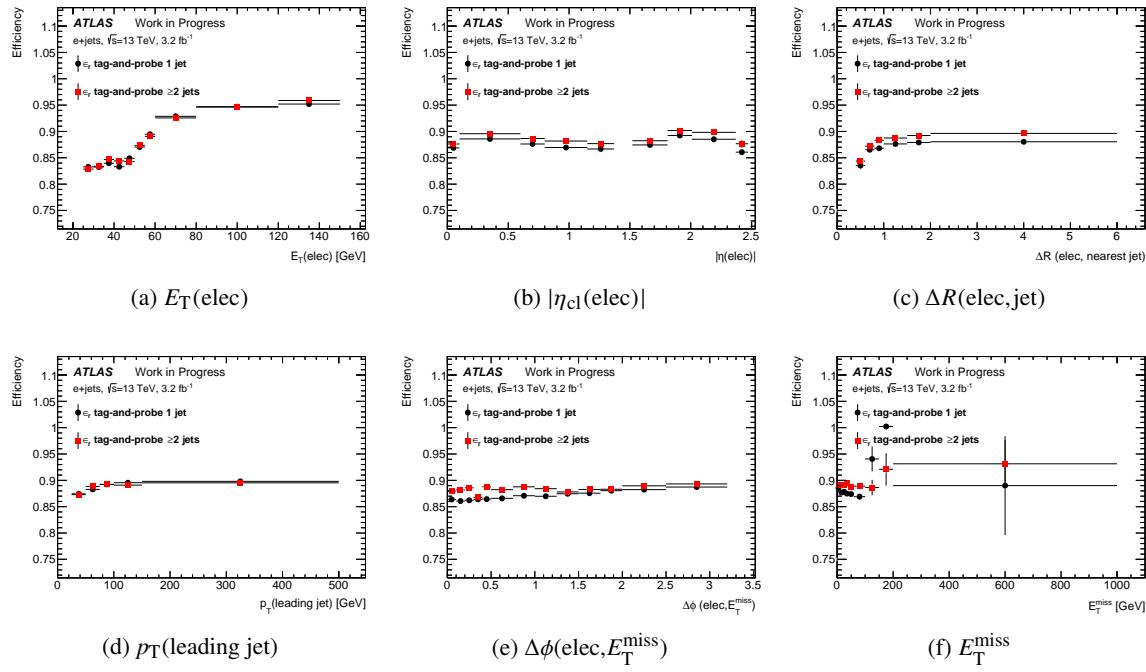


Figure 35: Electron real efficiency $\varepsilon_{\text{real}}$ obtained from the tag-and-probe method on data, multiplied by the MC correction factor, as a function of (a) the electron E_T , (b) the electron $|\eta_{\text{cl}}|$, (c) the distance ΔR between the electron and its nearest jet, (d) the p_T of the leading jet, (e) the difference $\Delta\phi(\text{elec}, E_T^{\text{miss}})$ in azimuth between the electron and the E_T^{miss} and (f) the E_T^{miss} . $\varepsilon_{\text{real}}$ is shown for events with only one jet, for all probes which match one of the single electron trigger (see section 5).

838 A.4. The high- E_T^{miss} method

839 To estimate electron real efficiencies $\varepsilon_{\text{real}}$ one selects a single electron sample with $E_T^{\text{miss}} > 150 \text{ GeV}$,
 840 corresponding to a pure sample of prompt isolated electrons. The real efficiency $\varepsilon_{\text{real}}$ is extracted as the
 841 ratio between tight and loose $e+\text{jets}$ events in the selected region.

842 Selection of events can be summarised as:

- 843 - Event quality preselection as described in section 5.1.
- 844 - Exactly one loose electron.
- 845 - At least one jet in the event.

846 A.4.1. Definition of the control region

847 Fig. 36 shows the E_T^{miss} distribution in the region for the loose and tight samples. The data are a bit overestimated by the sum of prompt leptons originated from different Monte Carlo samples. A total of

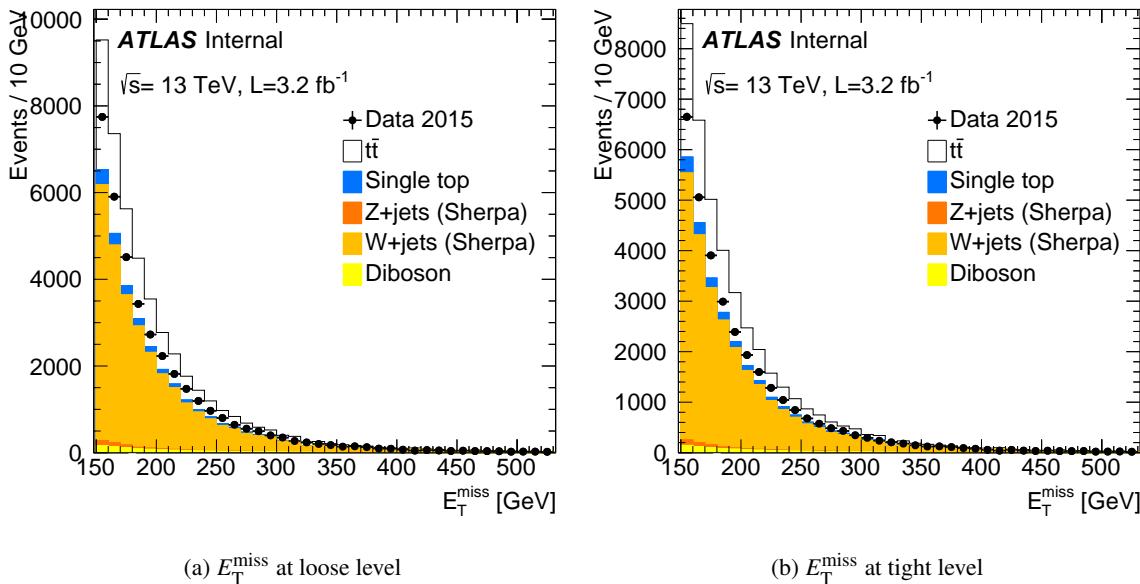


Figure 36: Distributions of the missing transverse energy E_T^{miss} , in the high values region ($E_T^{\text{miss}} > 150 \text{ GeV}$), for data and the different MC contributions at (a) loose and (b) tight levels for events with one electron, which matches one of the single electron trigger (see section 5), and at least one jet.

848 nearly 3.6×10^4 events are selected at loose level with one electron, at least one jet and without any
 849 b -tagging requirement. Table 17 shows the composition of these events. The dominant amount of real
 850 electrons in these regions stems from W boson events. Going to higher jet and b -jet multiplicities the
 851 picture changes and one gets dominated by $t\bar{t}$.
 852

Table 17: Sample composition (in %) of $e+jets$ events in the high $E_T^{\text{miss}} > 150 \text{ GeV}$ region for events with at least one jet. The sample is split into events with 0 b -jet and $\geq 1 b$ -jet.

	W	Z	$t\bar{t}$	Diboson	SingleTop	QCD
$\geq 1 j$	62.4	1.0	30.2	2.4	3.9	xx
$\geq 1 j, 0 b$	86.5	1.4	7.6	3.0	1.5	xx
$\geq 1 j, \geq 1 b$	16.3	0.4	73.4	1.2	1.2	xx

853 A.4.2. Results obtained

854 Fig. 37 shows $\varepsilon_{\text{real}}$, obtained from the high E_T^{miss} method, as a function of the jet and b -jet multiplicities.
855 Small dependencies are visible. The sample with at least one b -jet is enriched in $t\bar{t}$ events, on the contrary
856 of the sample with no b -jet which is enriched in $W+jets$ events. Electrons from $t\bar{t}$ and W events have
857 different isolation and thus efficiencies, as detailed in section A.2.2. Efficiencies are thus derived for two
jet multiplicities, $n_{\text{jet}} = 1$ and $n_{\text{jet}} \geq 2$ ⁹.

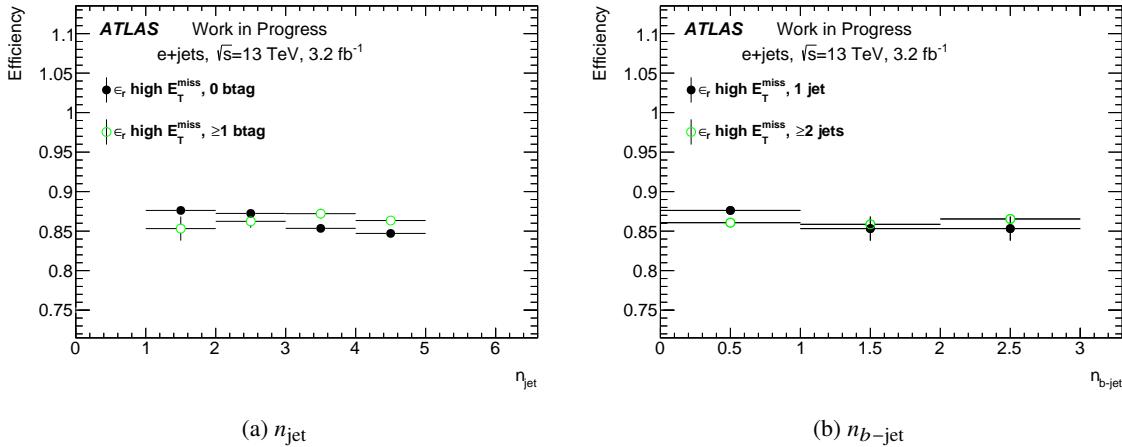


Figure 37: Electron real efficiency $\varepsilon_{\text{real}}$, obtained from the high E_T^{miss} method, as a function of (a) the number of jets n_{jet} and (b) the number of b -jets $n_{b-\text{jet}}$. $\varepsilon_{\text{real}}$ is shown for all electrons which match one of the single electron trigger (see section 5).

858

859 Fig. 38 shows $\varepsilon_{\text{real}}$, obtained from the high E_T^{miss} method, as a function of the different variables for events
860 with different jet and b -jet multiplicities. $\varepsilon_{\text{real}}$ has a dependency with the E_T of the electron, $\Delta R(\text{elec}, \text{jet})$
861 and $\Delta\phi(\text{elec}, E_T^{\text{miss}})$. It is almost flat with respect to $|\eta_{\text{cl}}|$ of the electron and the p_T of the leading jet. As
862 can be seen, on the contrary of Run 1 analyses, the low available statistics does not allow to use so many
863 bins. For the final parametrisation, no binning in n_{jet} and $n_{b-\text{jet}}$ is used. Fig. 39 shows $\varepsilon_{\text{real}}$, obtained from
864 the high E_T^{miss} method, as a function of the different variables for events with at least one jet and without
any requirement on the number of b -jet.

⁹ As for 8 TeV analyses there is not enough statistics to have more bins.

Not reviewed, for internal circulation only

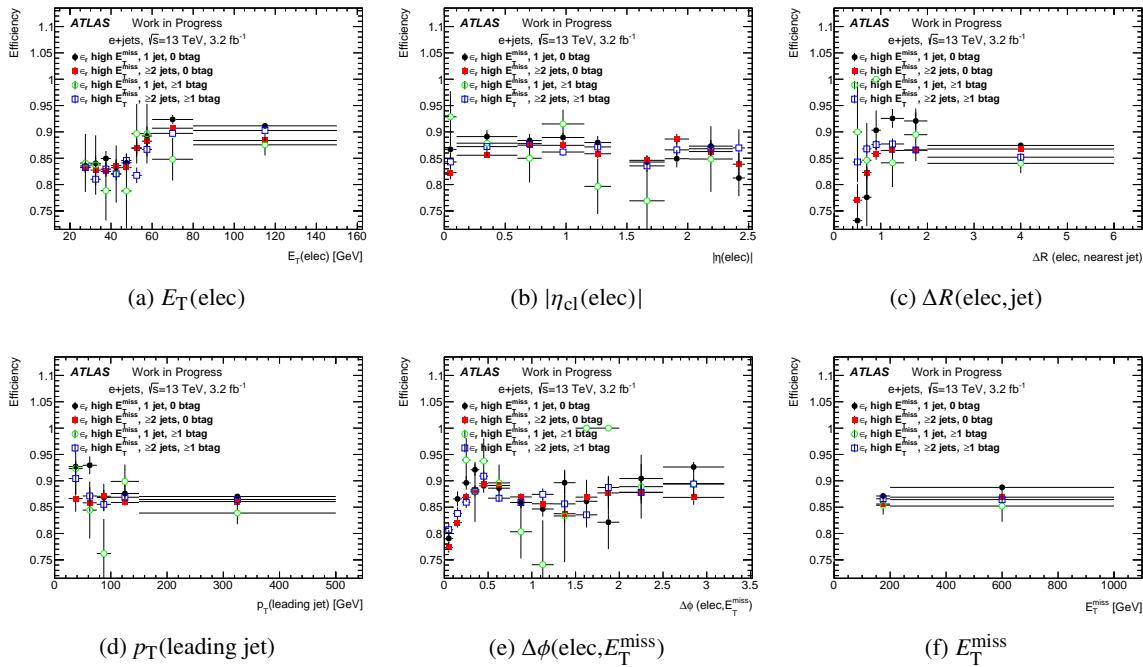


Figure 38: Electron real efficiency $\varepsilon_{\text{real}}$, obtained from the high E_T^{miss} method, as a function of (a) the electron E_T , (b) the electron $|\eta_{\text{cl}}|$, (c) the distance ΔR between the electron and its nearest jet, (d) the p_T of the leading jet, (e) the difference $\Delta\phi(\text{elec}, E_T^{\text{miss}})$ in azimuth between the electron and the E_T^{miss} and (f) the E_T^{miss} . $\varepsilon_{\text{real}}$ is shown for events with different jet and b -jet multiplicities, for all electrons which match one of the single electron trigger (see section 5).

Not reviewed, for internal circulation only

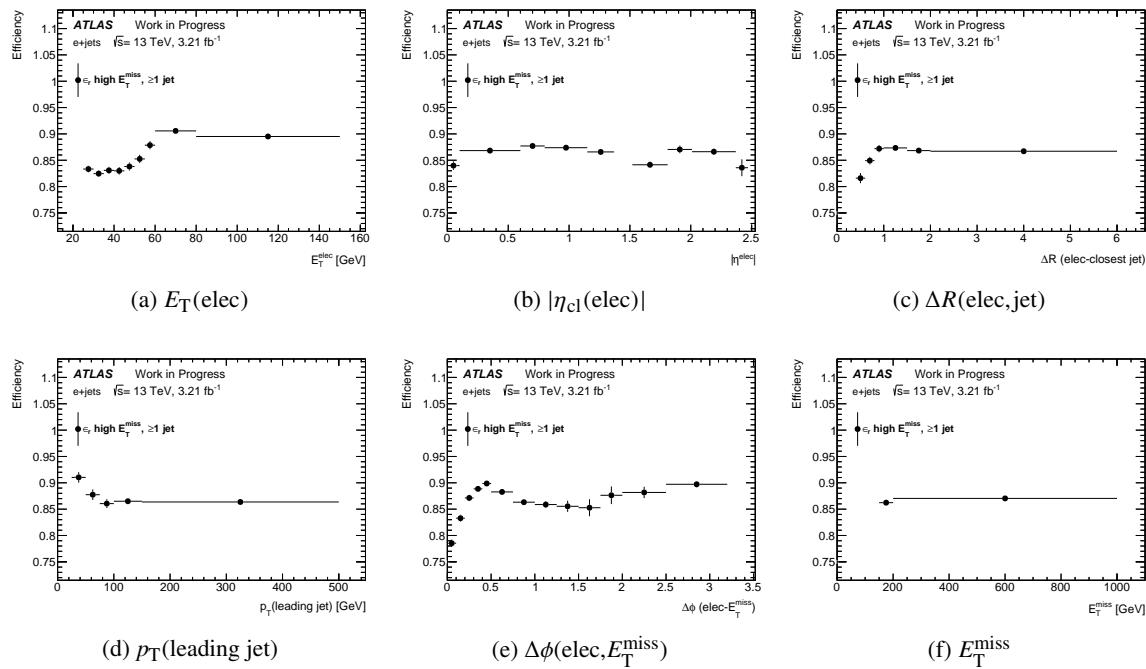


Figure 39: Electron real efficiency $\varepsilon_{\text{real}}$, obtained from the high E_T^{miss} method, as a function of (a) the electron E_T , (b) the electron $|\eta_{\text{cl}}|$, (c) the distance ΔR between the electron and its nearest jet, (d) the p_T of the leading jet, (e) the difference $\Delta\phi(\text{elec}, E_T^{\text{miss}})$ in azimuth between the electron and the E_T^{miss} and (f) the E_T^{miss} . $\varepsilon_{\text{real}}$ is shown for events with at least one jet and without any requirement on the number of b -jet, for all electrons which match one of the single electron trigger (see section 5).

Not reviewed, for internal circulation only

866 A.5. Comparison of efficiencies from both methods

867 Fig. 40 shows $\varepsilon_{\text{real}}$ as obtained by the two methods described above as a function of the electron kinematics and the variables describing the jet environment. Results are given for electrons which match one of 868 the single electron trigger (see section 5) and for different jet and b -jet multiplicities for the tag-and-probe 869 case. The tag-and-probe estimates tend to give higher efficiencies, in particular for large E_T and $|\eta_{\text{cl}}|$ of 870 the electron as well as p_T of the leading jet. It is also visible that the high- E_T^{miss} results lack statistics in 871 the case of low $\Delta R(\text{elec}, \text{jet})$, p_T of the leading jet and $\Delta\phi(\text{elec}, E_T^{\text{miss}})$.

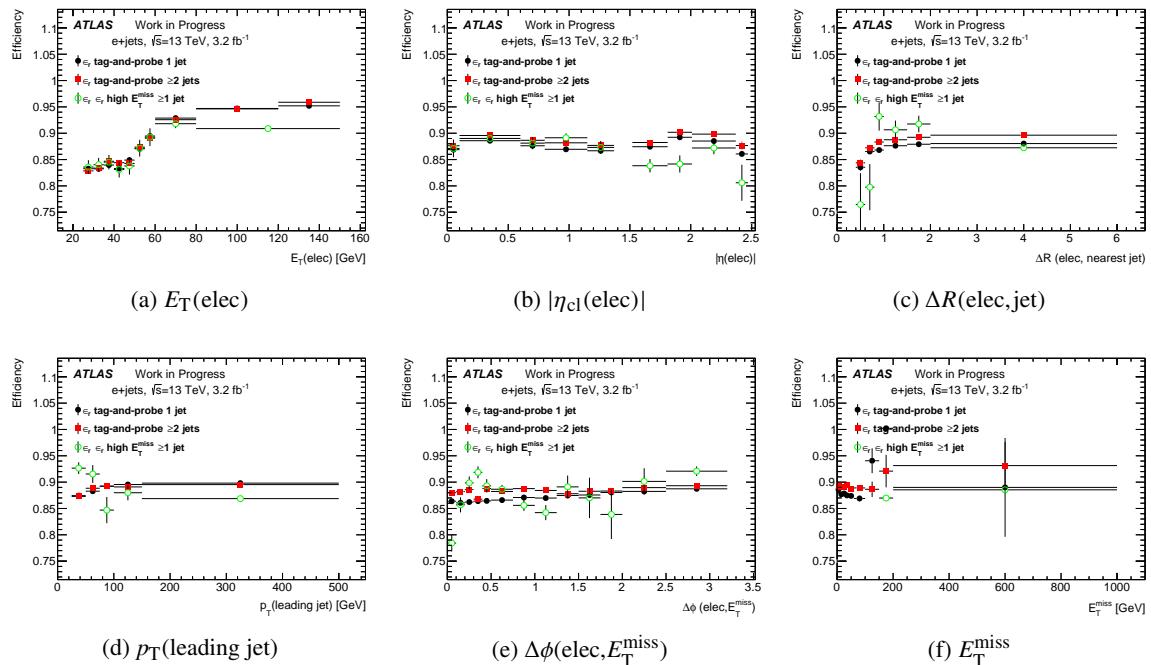


Figure 40: Electron real efficiency $\varepsilon_{\text{real}}$, obtained from the tag-and-probe and the high E_T^{miss} methods, as a function of (a) the electron E_T , (b) the electron $|\eta_{\text{cl}}|$, (c) the distance ΔR between the electron and its nearest jet, (d) the p_T of the leading jet, (e) the difference $\Delta\phi(\text{elec}, E_T^{\text{miss}})$ in azimuth between the electron and the E_T^{miss} and (f) the E_T^{miss} . $\varepsilon_{\text{real}}$ is shown for events with at least one jet and without any requirement on the number of b -jet, for all electrons which match one of the single electron trigger (see section 5).

872

873 From this comparison it can be concluded that :

- 874 - The results from the tag-and-probe method can be used as the default ones. The efficiencies can be 875 derived for two n_{jet} bins ($n_{\text{jet}} = 1$ and $n_{\text{jet}} \geq 2$ jets) but without $n_{b-\text{jet}}$ binning, and all variables can 876 be used.
- 877 - The High- E_T^{miss} results can be used to get a systematic uncertainty. The efficiencies can be derived 878 without using binning in n_{jet} or $n_{b-\text{jet}}$. Fewer variables can be used, only E_T and $|\eta_{\text{cl}}|$ of the electron 879 to avoid to get too large systematic uncertainties.

⁸⁸⁰ **A.6. Binning used to measure real electron efficiencies**

The binning used for the parametrisation of the real electron efficiencies are given in Tables 18 to 21.

Table 18: Measurement bins in electron transverse energy E_T used for the electron real efficiency parametrisation.

Method	Bin boundaries in E_T (GeV)									
tag-and-probe	25	30	35	40	45	50	60	80	100	∞
high- MET	25	30	35	40	45	50	60	80	∞	

Table 19: Measurement bins in electron pseudorapidity $|\eta_{\text{cl}}|$ used for the electron real efficiency parametrisation.

Bin boundaries in $ \eta_{\text{cl}} $										
0	0.1	0.6	0.8	1.15	1.37	1.52	1.81	2.01	2.37	2.47

Table 20: Measurement bins in distance to the nearest jet $\Delta R(\text{elec}, \text{jet})$ used for the electron real efficiency parametrisation.

Bin boundaries in $\Delta R(\text{elec}, \text{jet})$ to the nearest jet						
0.4	0.6	0.8	1	1.5	2	6

Table 21: Measurement bins in p_T of the leading jet used for the electron real efficiency parametrisation.

Bin boundaries in p_T of the leading jet (GeV)					
25	50	75	100	150	∞

⁸⁸¹

Table 22: Measurement bins in $|\Delta\phi|(\text{elec}, E_T^{\text{miss}})$ the distance in azimuth between the electron and the E_T^{miss} used for the electron real efficiency parametrisation.

Bin boundaries in $ \Delta\phi (\text{elec}, E_T^{\text{miss}})$													
0	0.1	0.2	0.3	0.4	0.5	0.75	1.0	1.25	1.5	1.75	2.0	2.5	3.2

⁸⁸²

883 B. Additional studies on real muon efficiencies

884 In this section we show additional studies done on the muon real efficiency $\varepsilon_{\text{real}}$ measurement.

885 B.1. Variables used for the parametrisation of the efficiencies

886 The main part of the note is describing results obtained with the final parametrisation, but in order to take
 887 into account the dependence of the efficiencies, more variables are considered:

- 888 - The muon transverse momentum p_T .
- 889 - The muon pseudorapidity η .
- 890 - The number of jets in the event.
- 891 - The number of b -jets in the event.
- 892 - The transverse momentum p_T of the leading jet;
- 893 - The distance ΔR between the muon and the nearest jet¹⁰.
- 894 - The distance in azimuth $\Delta\phi$ between the muon and the E_T^{miss} .
- 895 - The missing transverse energy E_T^{miss} .
- 896 - The transverse mass $m_T(\text{lept}, E_T^{\text{miss}})$.

897 B.2. Studies with true muons

898 Muon identification efficiency is studied with simulated data in $t\bar{t}$ events but also single top, $Z + \text{jets}$ and
 899 $W + \text{jets}$ events. Prompt muons are selected using the MCTruthClassifier tool.

900 B.2.1. Modeling

901 Fig. 41 shows the kinematics of muons in $t\bar{t}$, single top Z and W simulated data. Muons from $t\bar{t}$ have
 902 harder p_T spectrum and are more central. It was shown in the 8 TeV analysis [1] that muons from Z
 903 events with high number of partons in final states, tend to have similar kinematics as muons from $t\bar{t}$.
 904 Identification efficiency being dependent on the kinematic of the muons, these differences can easily be
 905 taken into account with a good parametrisation of the efficiencies. Fig. 42 shows the isolation variables
 906 $p_T^{0.3}$ entering in the definition of the tight objects, the p_T of the leading jet and the distance ΔR between
 907 the muon and its nearest jet. Muons from $Z + \text{jets}$ decays are more isolated than the ones from $t\bar{t}$, due to
 908 the smaller jet activities in the event. They also have a softer p_T spectrum of their leading jet. Finally,
 909 muons from $Z + \text{jets}$ events are mostly back to back to this jet, except in events with large parton number
 910 in which case the distribution peak towards one, with very close jets (see Ref. [1]).

911 Finally, Fig. 43 shows the distribution of the distance in azimuth between the muon and the E_T^{miss} , the
 912 E_T^{miss} and the $m_T(\text{lept}, E_T^{\text{miss}})$ for the different simulated samples.

¹⁰ This variable is an indicator of the hadronic activity near the muon. Still one has to remind that this does not capture activity below the jet reconstruction p_T threshold, and is distorted by the use of muon-jet overlap vetoes in the muon selection.

Not reviewed, for internal circulation only

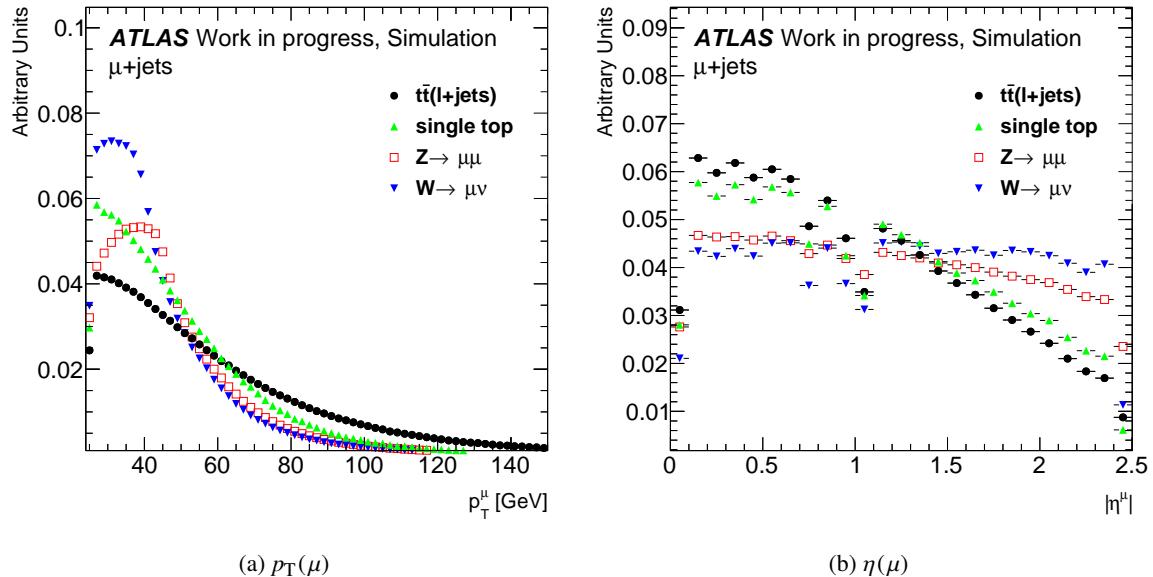


Figure 41: Distributions of (a) transverse momentum p_T and (b) pseudorapidity $|\eta|$ for prompt muons in $t\bar{t}$, single top, $Z+jets$ and $W+jets$ simulated data.

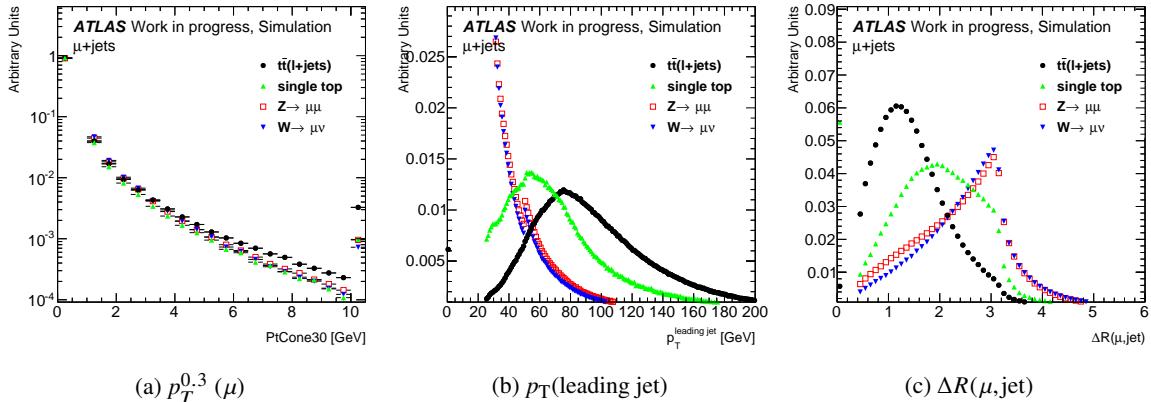


Figure 42: Distributions of (a) $p_T^{0.3}$, (b) p_T of the leading jet and (c) the distance ΔR between the muon and its nearest jet, for prompt muons in $t\bar{t}$, single top, $Z+jets$ and $W+jets$ simulated data.

913 B.2.2. True efficiencies

914 Muons from $t\bar{t}$ events have different kinematics and jet activity around them than the ones in $Z+jets$ events
 915 which turn in different real efficiencies. True efficiencies are measured in the different simulated samples
 916 as the ratio of tight over loose distributions.

917 Fig. 44 shows the true real efficiency as a function of the jet and b -jet multiplicities. Differences between
 918 efficiencies from $t\bar{t}$ and $Z+jets$ are small, but increase with n_{jet} and $n_{b-\text{jet}}$. Larger differences are seen in

Not reviewed, for internal circulation only

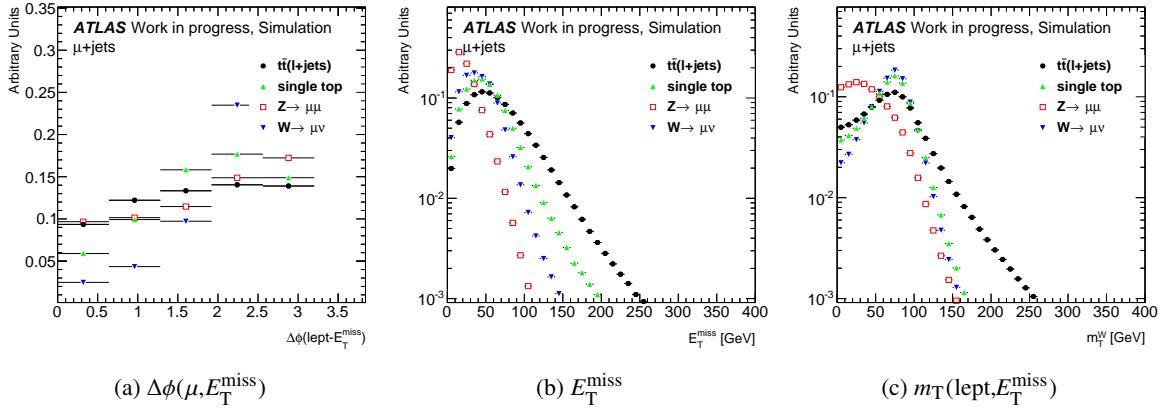


Figure 43: Distributions of (a) the distance $\Delta\phi(\mu, E_T^{\text{miss}})$ in azimuth between the muon and the E_T^{miss} , (b) the E_T^{miss} and (c) the $m_T(\text{lept}, E_T^{\text{miss}})$, for prompt muons in $t\bar{t}$, single top, $Z+jets$ and $W+jets$ simulated data.

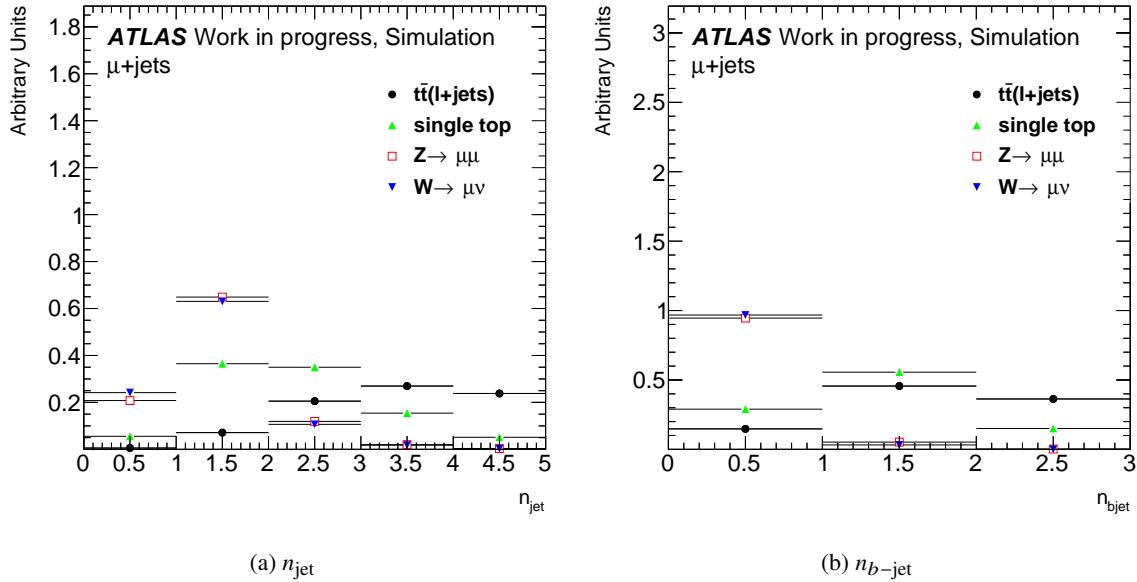


Figure 44: True muon real efficiency as a function of (a) the number of jets n_{jet} and (b) the number of b -jets $n_{b-\text{jet}}$. Efficiencies are shown for events with at least one jet in $t\bar{t}$, single top, $Z+jets$ and $W+jets$ simulated data. The purple star and the shaded area correspond respectively to the average and to the spread of the efficiencies measured from Z and $t\bar{t}$ events.

particular with muons from single top events. Fig. 45 shows the true muon real efficiency as a function of the different variables. Differences of efficiencies between the $t\bar{t}$ and the $Z+jets$ samples can reach few %.

Not reviewed, for internal circulation only

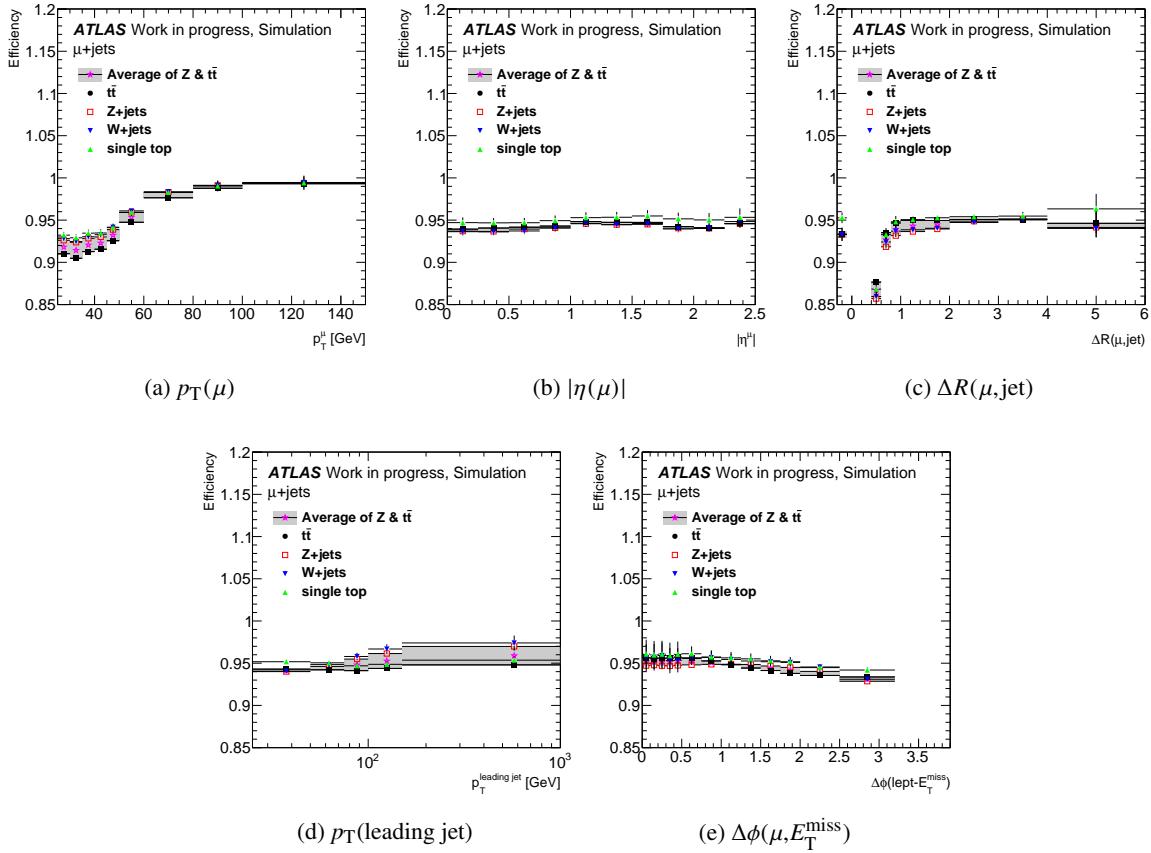


Figure 45: True muon real efficiency as a function of (a) the muon p_T , (b) the muon $|\eta|$, (c) the distance ΔR of the muon to its nearest jet, (d) the p_T of the leading jet and (f) the difference $\Delta\phi(\mu, E_T^{\text{miss}})$ in azimuth between the muon and the E_T^{miss} . Efficiencies are shown for events with at least one jet in $t\bar{t}$, single top, Z +jets and W +jets simulated data. The purple star and the shaded area correspond respectively to the average and to the spread of the efficiencies measured from Z and $t\bar{t}$ events.

922 B.2.3. Correction factor

923 To correct from the previously seen differences, correction factors are derived from the MC samples. The
 924 central value for the correction is taken as the average of the Z +jets ($\varepsilon(Z)$) and $t\bar{t}$ ($\varepsilon(t\bar{t})$) efficiencies,
 925 divided by $\varepsilon(Z)$. These corrections factors are shown in Fig. 46 as a function of the different variables.
 926 The spread between $\varepsilon(Z)$ and $\varepsilon(t\bar{t})$ is taken as a systematic uncertainty attributed to the difference of
 927 isolation and the unknown real muon sample composition. These correction factors are compatible with
 928 1 and, as for the 8 TeV analysis [1], no correction will be applied to the measurement based on data.

Not reviewed, for internal circulation only

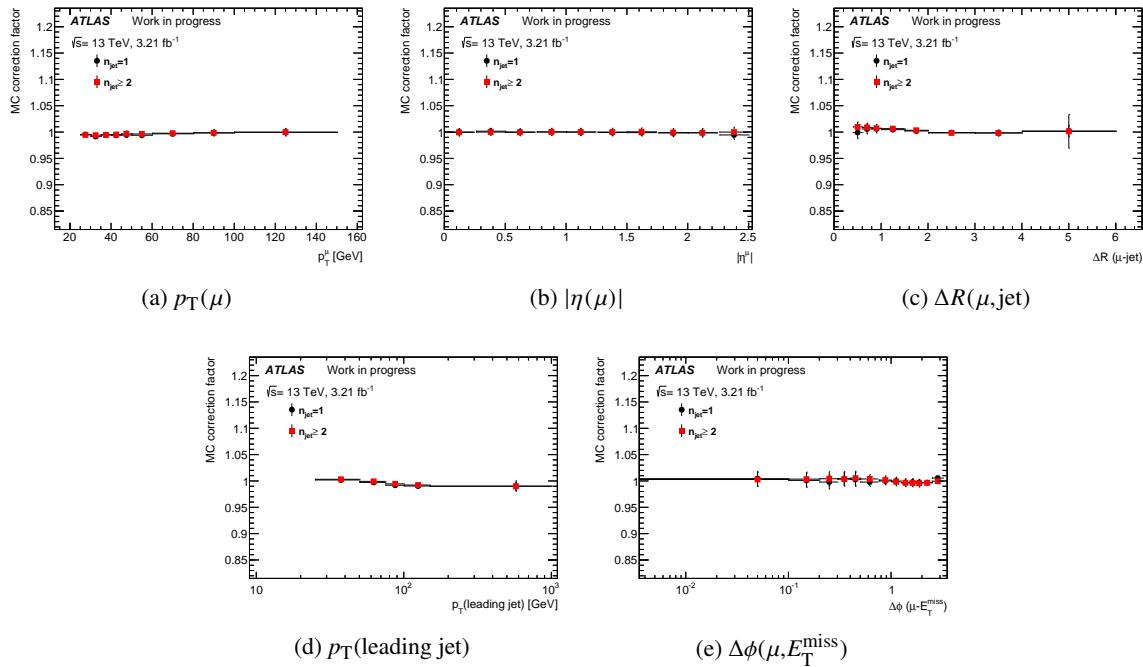


Figure 46: Monte Carlo correction factor to be applied on real muon efficiencies as a function of (a) the muon p_T , (b) the muon $|\eta|$, (c) the distance ΔR between the muon and its nearest jet, (d) the p_T of the leading jet and (e) the difference $\Delta\phi(\mu, E_T^{\text{miss}})$ in azimuth between the muon and the E_T^{miss} . The correction factor is shown for events with at least one jet.

929 **B.3. The tag-and-probe method**

930 The tag-and-probe method makes use of the characteristic signature of $Z \rightarrow \mu\mu$ decays. Main results
 931 obtained with this method were already presented in section 6.3.1. Additional material is presented
 932 here.

933 **B.3.1. Background subtraction**

934 After this selection, the sample still contains non-prompt and fake lepton backgrounds. The level of
 935 background in these Z samples is below the percent level. Measurement of efficiencies is not sensitive to
 936 the details of background subtraction methods and it is estimated. Still, several methods are considered:

- 937 - Side-band method which is the default one. This method relies on the background having a linear
 938 shape over the considered invariant mass range. The invariant mass distributions for opposite-sign
 939 and same-sign pairs at the denominator and numerator levels are divided in three regions A, B
 940 and C. The number of background events in region B and its uncertainty are estimated from the
 941 extrapolation of the side-bands A and C of the same-sign distribution.
- 942 - Removal of same-sign (SS) events from opposite-sign ones (OS) in the signal region.
- 943 - A two-component fit with a signal contribution plus a background contribution is performed in each
 944 bin to the invariant mass $m_{\mu\mu}$ distribution. The signal contribution is modeled by a Breit-Wigner
 945 distribution convolved with a parametrisation of the low-mass tail, arising mostly from material
 946 effects, by a Crystal Ball function. For the background contribution a variety of fit functions are
 947 considered such as an exponential and a single-sided exponential convolved with a Gaussian.

948 The main uncertainty on the efficiency measurements are linked to the contamination of the sample of
 949 probes by background. To assess the precision of its subtraction a total of 18 variations is considered :

- 950 - Three different sideband regions (in GeV) : 1) A=[61-81], B=[81-101], C=[101-121], 2) A=[66-
 951 76], B=[76-106], C=[106-116], 3) A=[61-81], B=[86-96], C=[101-121].
- 952 - Three different regions for the same sign method (in GeV) : [81-101], [76-106], [86-96].
- 953 - For the fit method there is only one model for the signal function and two for the background one.
 954 For each there are two fit ranges (in GeV) : [60-120], [55-200] and three windows to measure the
 955 yields : [81-101], [76-106], [86-96].

956 The central value and the spread (rms) are determined for all these measurements.

957 **B.3.2. Results obtained on Z data**

958 Fig. 47 shows $\varepsilon_{\text{real}}$ as a function of the jet and b -jet multiplicities. It was shown in section B.2.2 that such
 959 dependence with the jet multiplicity is expected in Z events. On the contrary, efficiencies is not expected
 960 to depend on the number of b -jets. Efficiencies are derived for each of the three triggers mu20i, mu20
 961 an mu50. The ones obtained for mu50 are much higher than the other ones, due to the higher p_T of the
 962 muons. Efficiencies obtained for the two low- p_T triggers are similar. Fig. 48 shows $\varepsilon_{\text{real}}$ as a function
 963 of the different variables for all probe muons which match specifically one of the single muon triggers.

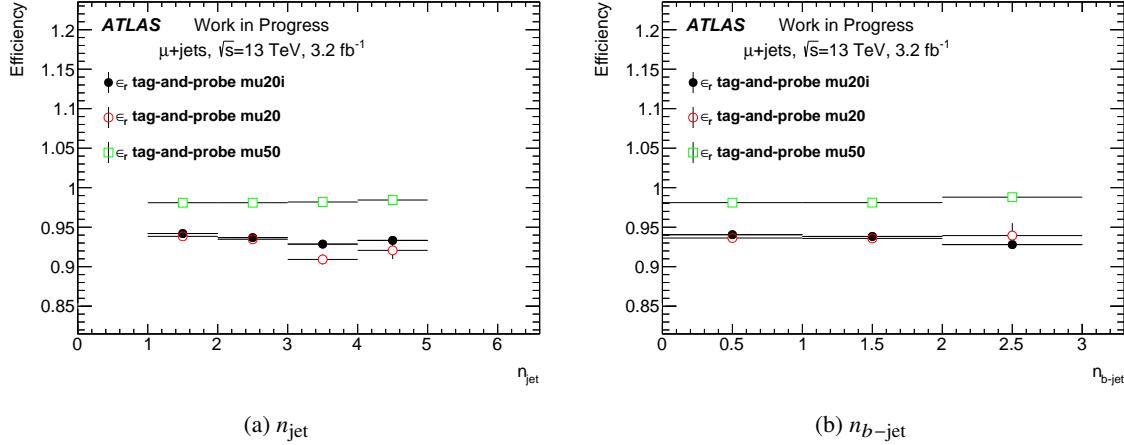


Figure 47: Muon real efficiency $\varepsilon_{\text{real}}$ obtained from the tag-and-probe method on data as a function of (a) the number of jets n_{jet} and (b) the number of b -jets $n_{b\text{-jet}}$. $\varepsilon_{\text{real}}$ is shown for all probes which match specifically one of the single muon trigger (see section 5).

Efficiencies are derived for two jet multiplicities, $n_{\text{jet}} = 1$ and $n_{\text{jet}} \geq 2^{11}$. A strong dependency is visible with the p_{T} of the muon. The dependency with respect to $|\eta|$ of the muon is very small. Some dependency is seen also with respect to the transverse momentum of the leading jet and the distance ΔR . Finally, a small dependence is seen with respect to the distance in azimuth between the muon and the $E_{\text{T}}^{\text{miss}}$.

¹¹ For 8 TeV analyses one was using three jet multiplicity bins, $n_{\text{jet}} = 1$, $n_{\text{jet}} = 2 - 3$ and $n_{\text{jet}} \geq 4$, but it was found to be difficult to do for 13 TeV analysis due to lack of statistics for the $n_{\text{jet}} \geq 4$ bin.

Not reviewed, for internal circulation only

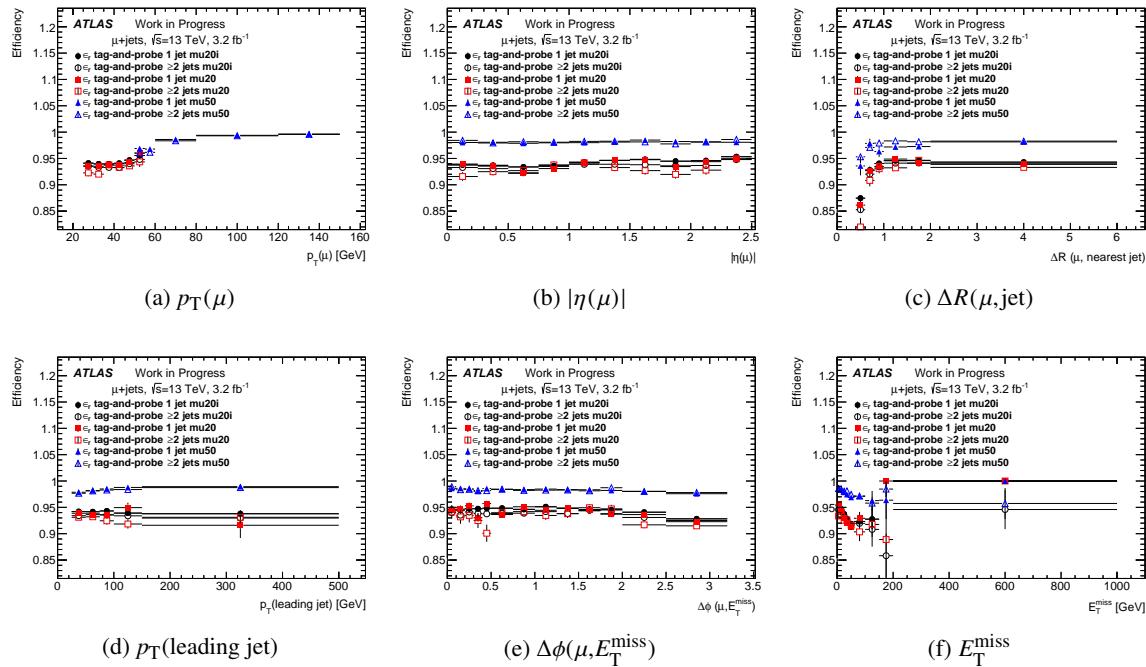


Figure 48: Muon real efficiency $\varepsilon_{\text{real}}$ obtained from the tag-and-probe method on data as a function of (a) the muon p_T , (b) the muon $|\eta|$, (c) the distance ΔR between the muon and its nearest jet, (d) the p_T of the leading jet and (e) the difference $\Delta\phi(\mu, E_T^{\text{miss}})$ in azimuth between the muon and the E_T^{miss} . $\varepsilon_{\text{real}}$ is shown for events with different jet multiplicities, for all probes which match one of the single muon trigger (see section 5), detailing results for each trigger.

968 B.4. The high- $m_T(\text{lept}, E_T^{\text{miss}})$ method

969 To estimate muon real efficiencies $\varepsilon_{\text{real}}$ one selects a single muon sample with $m_T(\text{lept}, E_T^{\text{miss}}) > 100 \text{ GeV}$,
 970 corresponding to a pure sample of prompt isolated muons. The real efficiency $\varepsilon_{\text{real}}$ is extracted as the ratio
 971 between tight and loose $e+\text{jets}$ events in the selected region.

972 Selection of events can be summarised as:

- 973 - Event quality preselection as described in section 5.1.
- 974 - Exactly one loose muon.
- 975 - At least one jet in the event.

976 B.4.1. Definition of the control region

977 A total of nearly 1.5×10^5 (2.8×10^5) muons are selected at loose level with the mu20i (mu50) trigger
 978 with one muon, at least one jet and without any b -tagging requirement.

979 Fig. 49 shows the $m_T(\text{lept}, E_T^{\text{miss}})$ distribution in the region for the loose and tight samples. The data
 980 are a bit underestimated by the sum of prompt leptons originated from different Monte Carlo samples, in
 981 particular in the bin 100-110 GeV. Table 23 shows the composition of these events. The dominant amount
 982 of real muons in these regions stems from W boson events. Going to higher jet and b -jet multiplicities the
 983 picture changes and one gets dominated by $t\bar{t}$.

984 The level of fake leptons is not estimated in this region but seems higher than what was obtained in the
 985 8 TeV analysis [2]. Further studies will be needed to check the impact of this contamination and to find
 986 possibly a better region to measure real efficiencies. The available statistics is 2-4 times lower, depending
 987 on the trigger used, than what was available for the 8 TeV analysis. It was not found possible to use a
 tighter selection, for example $m_T(\text{lept}, E_T^{\text{miss}}) > 120 \text{ GeV}$ due to this lack of statistics.

Table 23: Sample composition (in %) of $\mu+\text{jets}$ events in the high $m_T(\text{lept}, E_T^{\text{miss}}) > 100 \text{ GeV}$ region for events with
 at least one jet. The sample is split into events with 0 b -jet and $\geq 1 b$ -jet.

Selection	Trigger	W	Z	$t\bar{t}$	Diboson	SingleTop	QCD
$\geq 1 j$	mu20i	82.9	4.3	10.3	1.0	1.5	xx
$\geq 1 j, 0 b$		92.2	4.7	1.7	1.0	0.5	xx
$\geq 1 j, \geq 1 b$		24.7	1.6	64.8	0.7	8.2	xx
$\geq 1 j$	mu50	74.8	4.3	16.6	1.6	2.7	xx
$\geq 1 j, 0 b$		88.9	5.1	3.3	1.8	0.9	xx
$\geq 1 j, \geq 1 b$		23.8	1.6	64.8	0.9	8.9	xx

988

Not reviewed, for internal circulation only

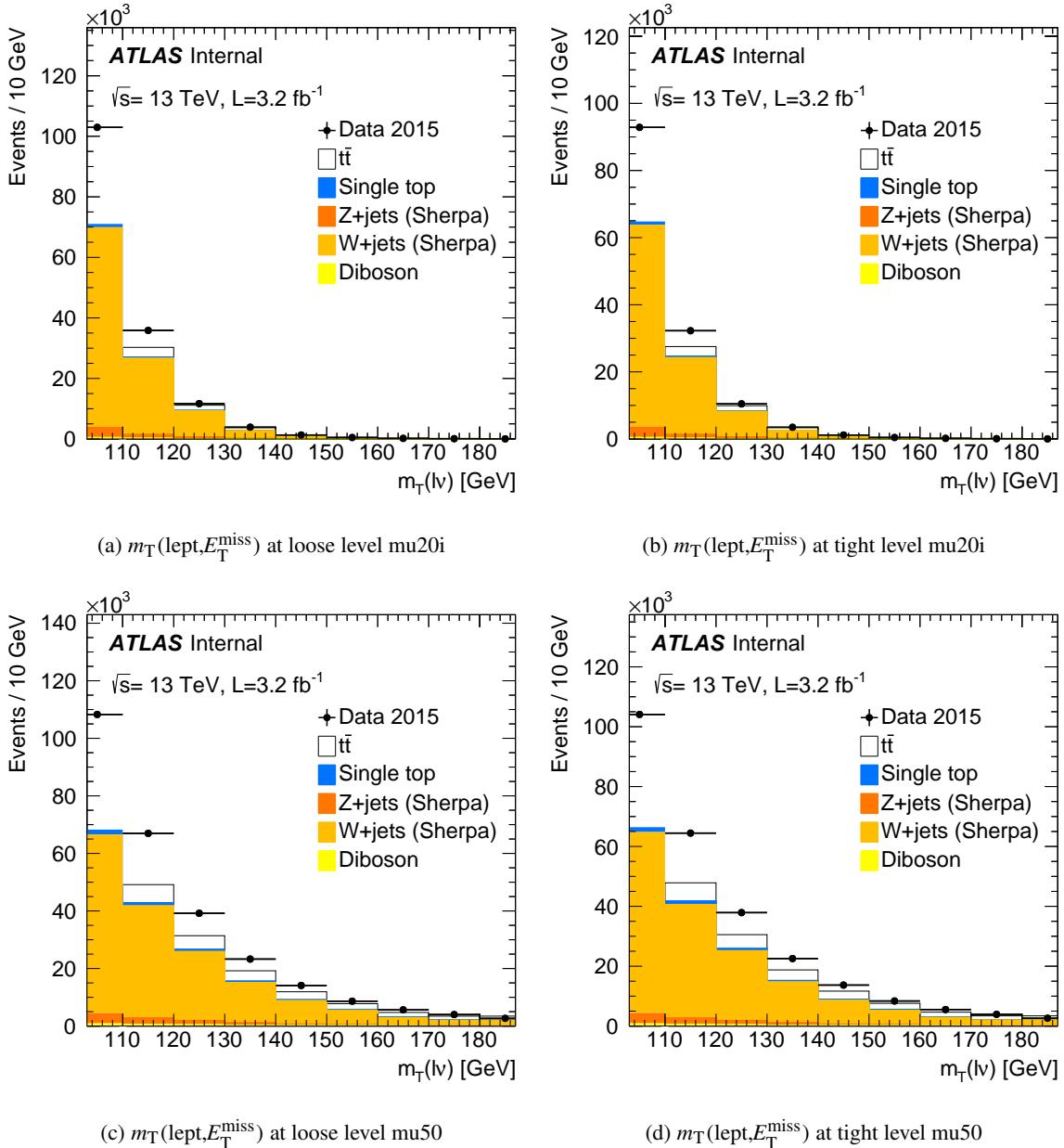


Figure 49: Distributions of $m_T(\text{lept}, E_T^{\text{miss}})$, in the high values region ($m_T(\text{lept}, E_T^{\text{miss}}) > 100 \text{ GeV}$), for data and the different MC contributions at (a,c) loose and (b,d) tight levels for events with one muon which matches the (a,b) mu20i or (c,d) mu50 triggers.

Not reviewed, for internal circulation only

989 B.4.2. Results obtained

990 Fig. 50 shows $\varepsilon_{\text{real}}$, obtained from the high $m_T(\text{lept}, E_T^{\text{miss}})$ method, as a function of the jet and b -jet
 991 multiplicities. Small dependencies are visible. The sample with at least one b -jet is enriched in $t\bar{t}$ events,
 992 on the contrary of the sample with no b -jet which is enriched in $W + \text{jets}$ events. Muons from $t\bar{t}$ and W
 993 events have different isolation and thus efficiencies, as detailed in section B.2.2. Efficiencies are thus
 derived for two jet multiplicities, $n_{\text{jet}} = 1$ and $n_{\text{jet}} \geq 2^{12}$.

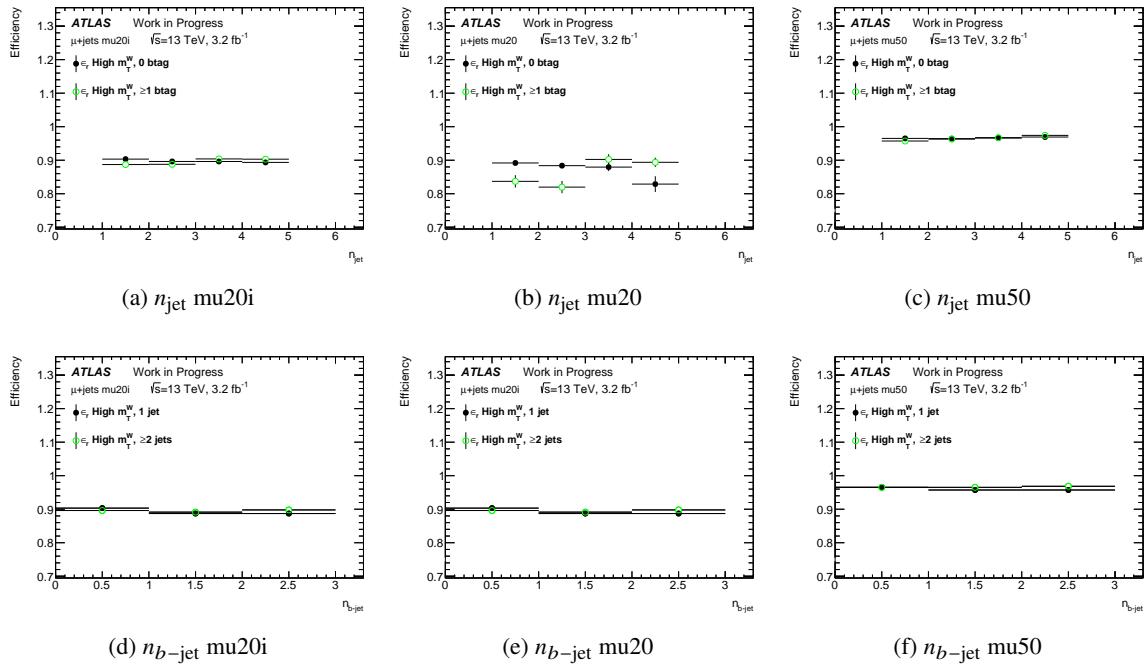


Figure 50: Muon real efficiency $\varepsilon_{\text{real}}$, obtained from the high $m_T(\text{lept}, E_T^{\text{miss}})$ method, as a function of (a,b,c) the number of jets n_{jet} and (d,e,f) the number of b -jets $n_{b\text{-jet}}$. $\varepsilon_{\text{real}}$ is shown for all muons which match one of the single muon trigger (a,d) mu20i, (b,e) mu20 and (c,f) mu50 (see section 5).

994

995 Fig. 51 to 53 show $\varepsilon_{\text{real}}$, obtained from the high $m_T(\text{lept}, E_T^{\text{miss}})$ method, as a function of the different
 996 variables for events with different jet and b -jet multiplicities. $\varepsilon_{\text{real}}$ has a strong dependency with the p_T of
 997 the muon, ΔR and $\Delta\phi(\mu, E_T^{\text{miss}})$. It is almost flat with respect to $|\eta|$ of the muon and the p_T of the leading
 998 jet.

¹² As for 8 TeV analyses there is not enough statistics to have more bins.

Not reviewed, for internal circulation only

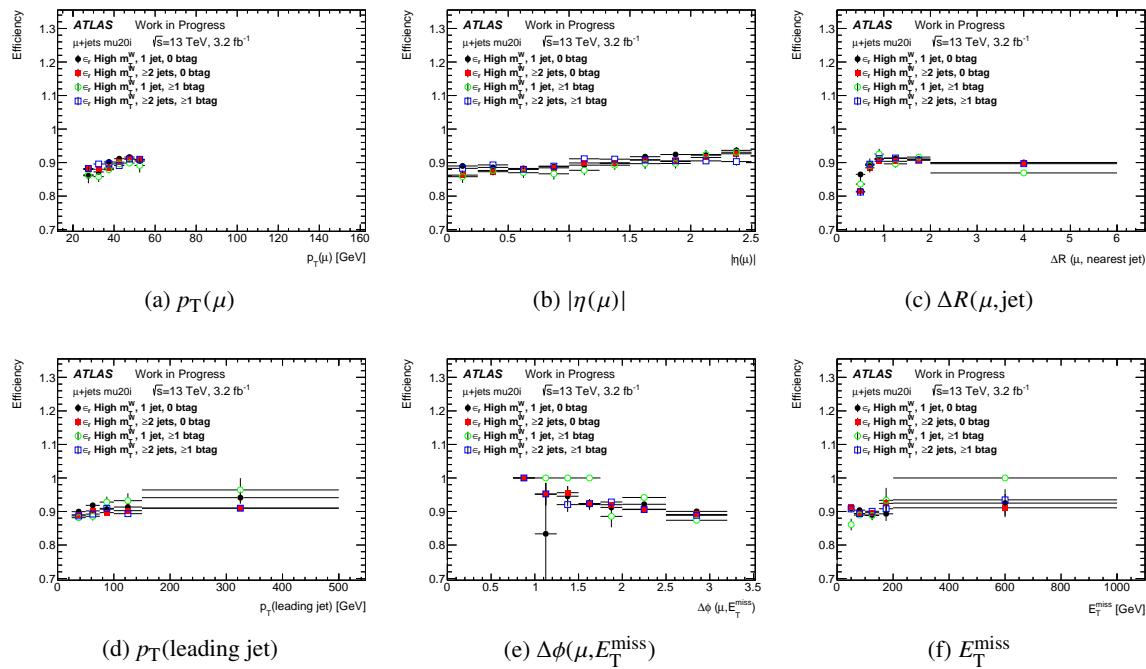


Figure 51: Muon real efficiency $\varepsilon_{\text{real}}$, obtained from the high $m_T(\text{lept}, E_T^{\text{miss}})$ method, as a function of (a) the muon p_T , (b) the muon $|\eta|$, (c) the distance ΔR between the muon and its nearest jet, (d) the p_T of the leading jet, (e) the difference $\Delta\phi(\mu, E_T^{\text{miss}})$ in azimuth between the muon and the E_T^{miss} and (f) the E_T^{miss} . $\varepsilon_{\text{real}}$ is shown for all muons which match the single muon trigger HLT_MU20_ILOOSE_L1MU15 (mu20i, see section 5).

Not reviewed, for internal circulation only

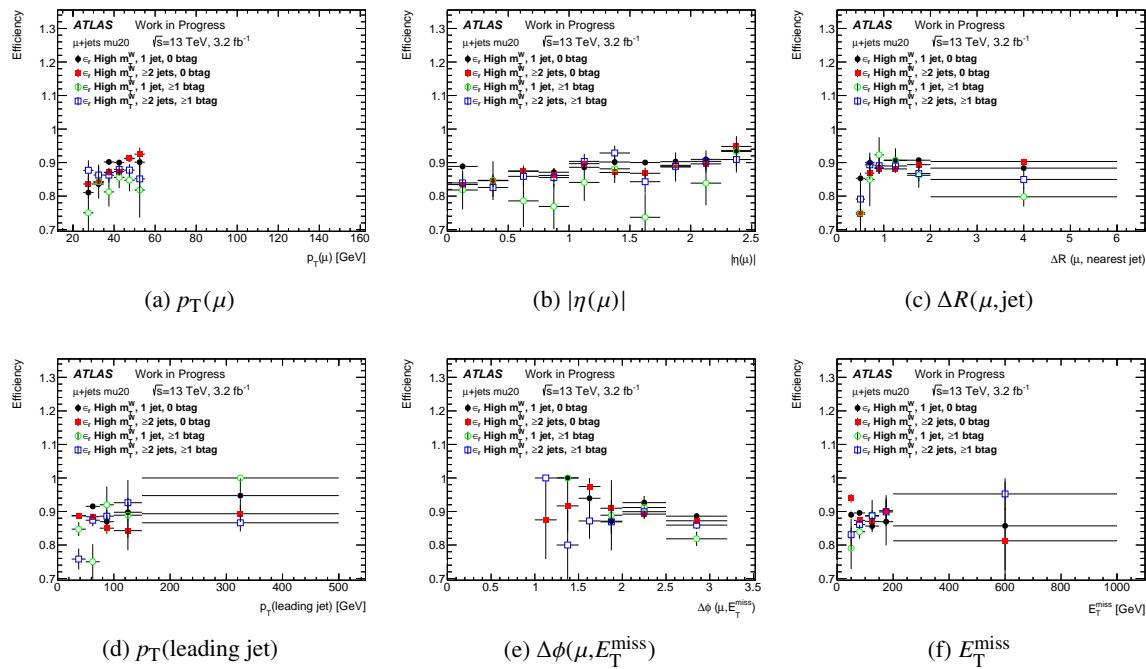


Figure 52: Muon real efficiency $\varepsilon_{\text{real}}$, obtained from the high $m_T(\text{lept}, E_T^{\text{miss}})$ method, as a function of (a) the muon p_T , (b) the muon $|\eta|$, (c) the distance ΔR between the muon and its nearest jet, (d) the p_T of the leading jet, (e) the difference $\Delta\phi(\mu, E_T^{\text{miss}})$ in azimuth between the muon and the E_T^{miss} and (f) the E_T^{miss} . $\varepsilon_{\text{real}}$ is shown for all muons which match the single muon trigger HLT_MU20_L1MU15 (mu20, see section 5).

Not reviewed, for internal circulation only

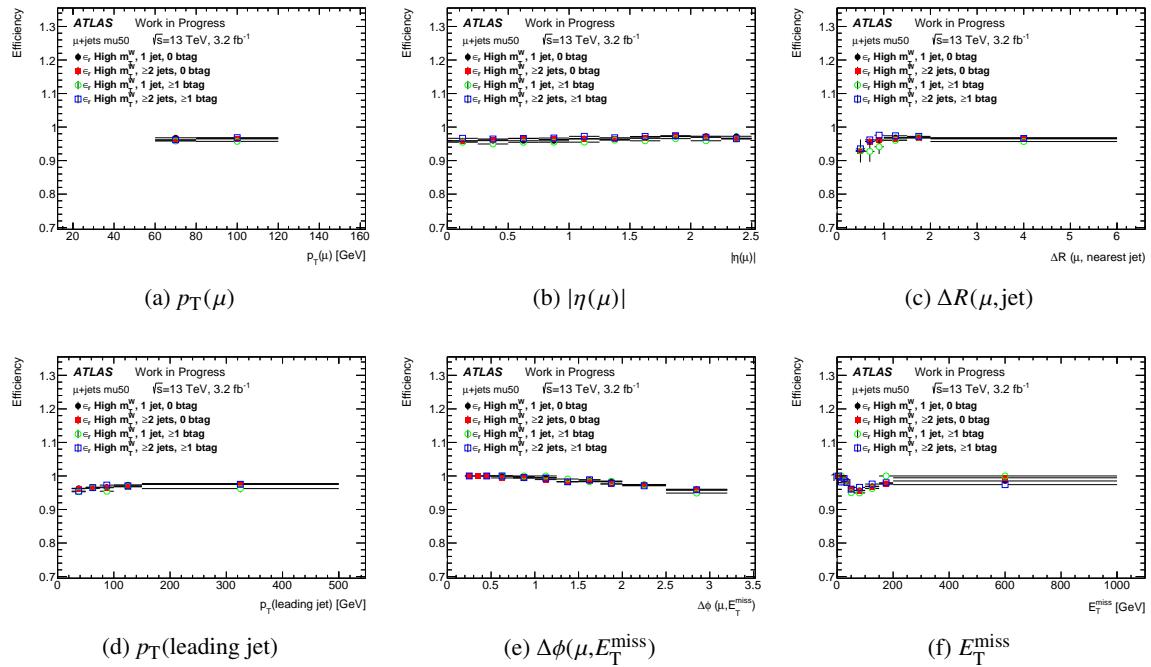


Figure 53: Muon real efficiency $\varepsilon_{\text{real}}$, obtained from the high $m_T(\text{lept}, E_T^{\text{miss}})$ method, as a function of (a) the muon p_T , (b) the muon $|\eta|$, (c) the distance ΔR between the muon and its nearest jet, (d) the p_T of the leading jet, (e) the difference $\Delta\phi(\mu, E_T^{\text{miss}})$ in azimuth between the muon and the E_T^{miss} and (f) the E_T^{miss} . $\varepsilon_{\text{real}}$ is shown for all muons which match the single muon trigger HLT_MU50 (mu50, see section 5).

- 999 On the contrary of Run 1 analyses, the low available statistics does not allow to use as many bins.
1000 For the final parametrisation, no binning in n_{jet} and $n_{b\text{-jet}}$ is used. Fig. 54 to 56 show $\varepsilon_{\text{real}}$, obtained from the high
1001 $m_T(\text{lept}, E_T^{\text{miss}})$ method, as a function of the different variables for events with at least one jet and without
any requirement on the number of b -jet.

Not reviewed, for internal circulation only

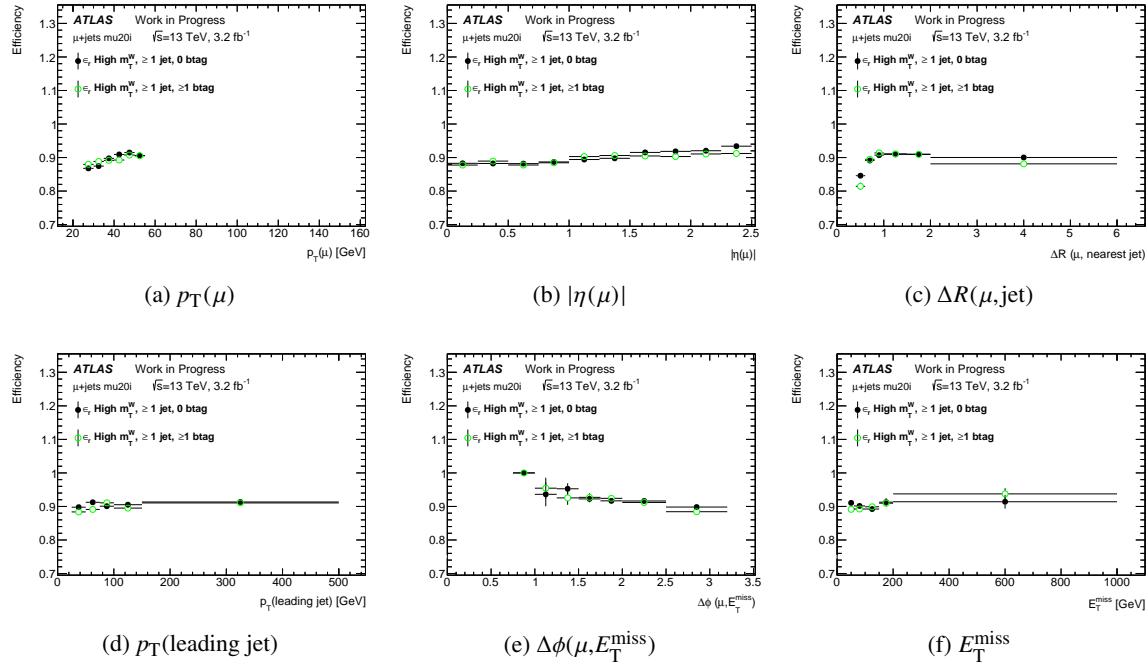


Figure 54: Muon real efficiency $\varepsilon_{\text{real}}$, obtained from the high $m_T(\text{lept}, E_T^{\text{miss}})$ method, as a function of (a) the muon E_T , (b) the muon $|\eta|$, (c) the distance ΔR between the muon and its nearest jet, (d) the p_T of the leading jet, (e) the difference $\Delta\phi(\mu, E_T^{\text{miss}})$ in azimuth between the muon and the E_T^{miss} and (f) the E_T^{miss} . $\varepsilon_{\text{real}}$ is shown for all muons which match the single muon trigger HLT_MU20_ILOOSE_L1MU15 (mu20i, see section 5) for events with at least one jet and without any requirement on the number of b -jet.

Not reviewed, for internal circulation only

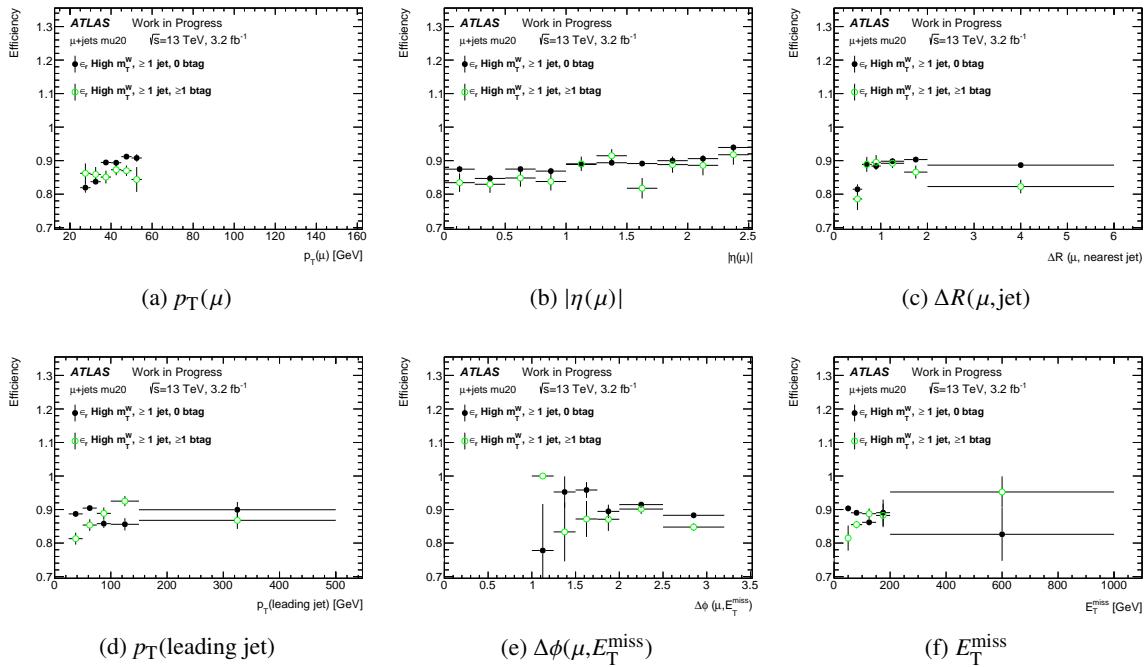


Figure 55: Muon real efficiency $\varepsilon_{\text{real}}$, obtained from the high $m_T(\text{lept}, E_T^{\text{miss}})$ method, as a function of (a) the muon E_T , (b) the muon $|\eta|$, (c) the distance ΔR between the muon and its nearest jet, (d) the p_T of the leading jet, (e) the difference $\Delta\phi(\mu, E_T^{\text{miss}})$ in azimuth between the muon and the E_T^{miss} and (f) the E_T^{miss} . $\varepsilon_{\text{real}}$ is shown for all muons which match the single muon trigger HLT_MU20_L1MU15 (mu20, see section 5) for events with at least one jet and without any requirement on the number of b -jet.

Not reviewed, for internal circulation only

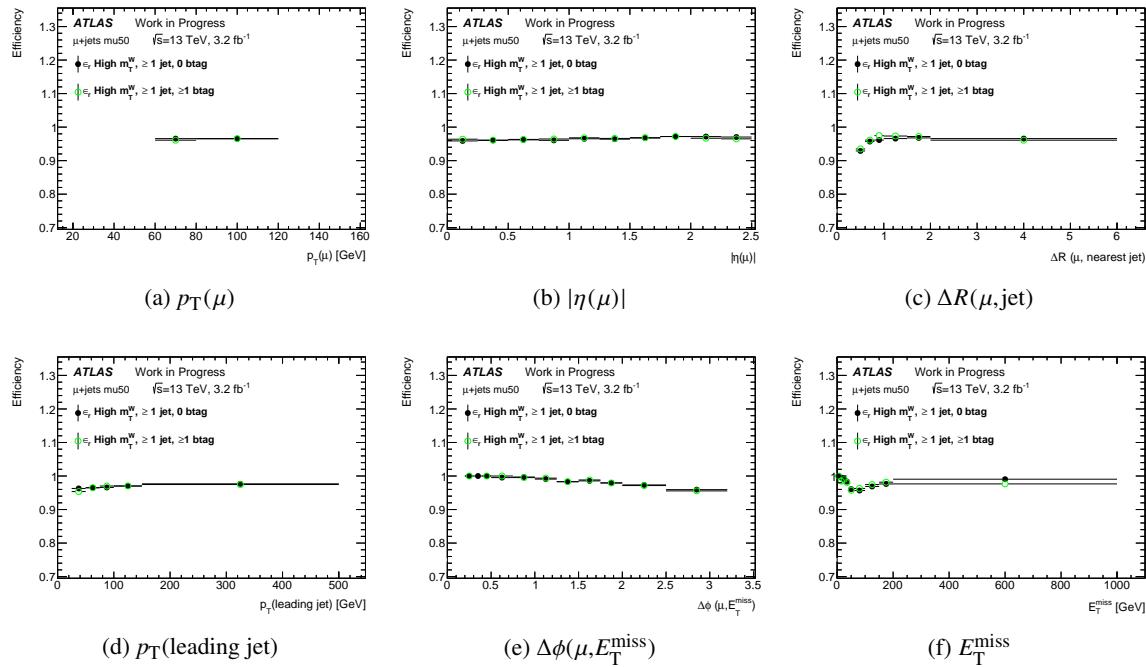


Figure 56: Muon real efficiency $\varepsilon_{\text{real}}$, obtained from the high $m_T(\text{lept}, E_T^{\text{miss}})$ method, as a function of (a) the muon E_T , (b) the muon $|\eta|$, (c) the distance ΔR between the muon and its nearest jet, (d) the p_T of the leading jet, (e) the difference $\Delta\phi(\mu, E_T^{\text{miss}})$ in azimuth between the muon and the E_T^{miss} and (f) the E_T^{miss} . $\varepsilon_{\text{real}}$ is shown for all muons which match the single muon trigger HLT_MU50 (mu50, see section 5) for events with at least one jet and without any requirement on the number of b -jet.

1003 **B.5. Comparison of efficiencies from both methods**

1004 Fig. 57 to 59 show $\varepsilon_{\text{real}}$ as obtained by the two methods described above as a function of the muon
 1005 kinematics and the variables describing the jet environment. Results are given separately for events
 1006 passing the triggers HLT_MU20_L1MU15, HLT_MU20_ILOOSE_L1MU15 and HLT_MU50. One can no-
 1007 tice that efficiencies from tag-and-probe method are much higher than the ones obtained from the high-
 1008 $m_T(\text{lept}, E_T^{\text{miss}})$ method, from 1 to 10%, in particular at low muon p_T . This is much more important than
 in the 8 TeV analysis [2] and could be due to a larger contamination of fake leptons in the sample. From

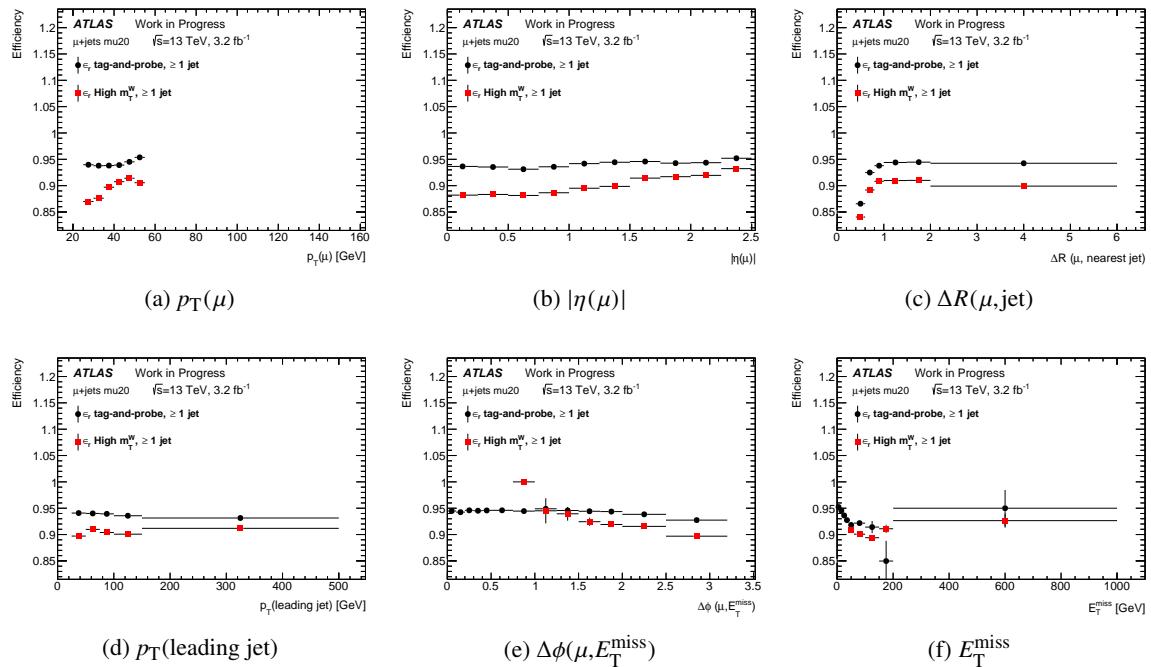


Figure 57: Muon real efficiency $\varepsilon_{\text{real}}$, obtained from the tag-and-probe and the high $m_T(\text{lept}, E_T^{\text{miss}})$ methods, as a function of (a) the muon p_T , (b) the muon $|\eta|$, (c) the distance ΔR between the muon and its nearest jet, (d) the p_T of the leading jet, (e) the difference $\Delta\phi(\mu, E_T^{\text{miss}})$ in azimuth between the muon and the E_T^{miss} and (f) the E_T^{miss} . $\varepsilon_{\text{real}}$ is shown for events with at least one jet and without any requirement on the number of b -jet, for all muons which match the single muon trigger HLT_MU20_ILOOSE_L1MU15 (mu20i, see section 5).

1009
1010 this comparison it can be concluded that :

- 1011 - The results from the tag-and-probe method can be used as the default ones. The efficiencies can be
 1012 derived for two n_{jet} bins ($n_{\text{jet}} = 1$ and $n_{\text{jet}} \geq 2$ jets) but without $n_{b-\text{jet}}$ binning, and all variables can
 1013 be used.
- 1014 - The High- $m_T(\text{lept}, E_T^{\text{miss}})$ results are used to get a systematic uncertainty. The efficiencies can be
 1015 derived without using binning in n_{jet} or $n_{b-\text{jet}}$.

Not reviewed, for internal circulation only

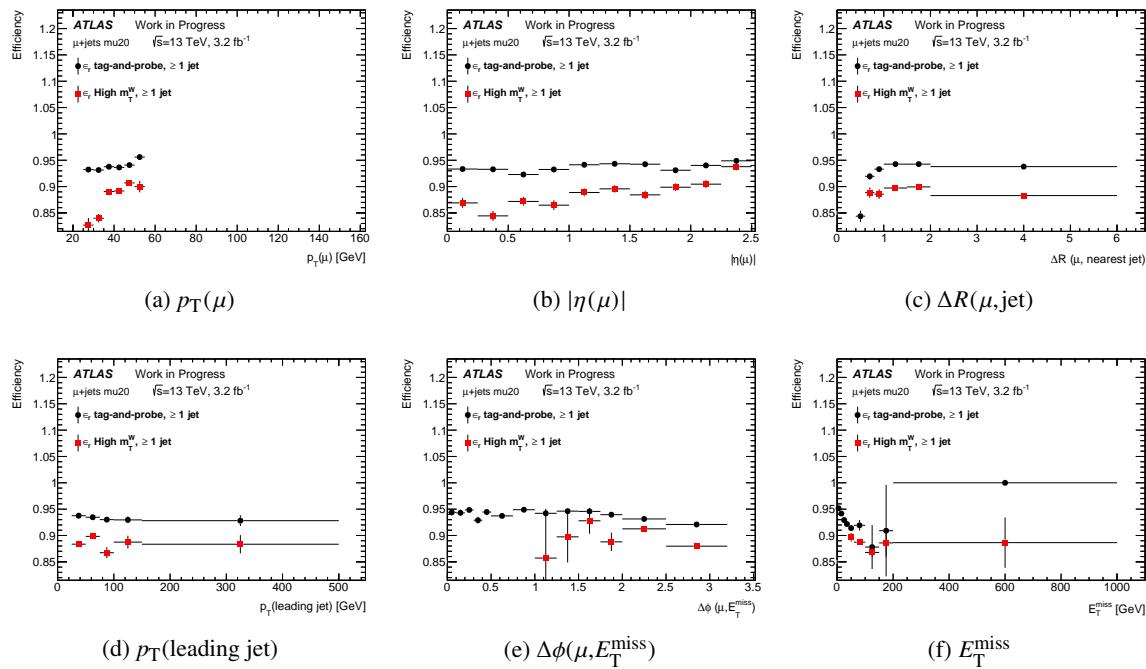


Figure 58: Muon real efficiency $\varepsilon_{\text{real}}$, obtained from the tag-and-probe and the high $m_T(\text{lept}, E_T^{\text{miss}})$ methods, as a function of (a) the muon p_T , (b) the muon $|\eta|$, (c) the distance ΔR between the muon and its nearest jet, (d) the p_T of the leading jet, (e) the difference $\Delta\phi(\mu, E_T^{\text{miss}})$ in azimuth between the muon and the E_T^{miss} and (f) the E_T^{miss} . $\varepsilon_{\text{real}}$ is shown for events with at least one jet and without any requirement on the number of b -jet, for all muons which match the single muon trigger HLT_MU20_L1MU15 (mu20, see section 5).

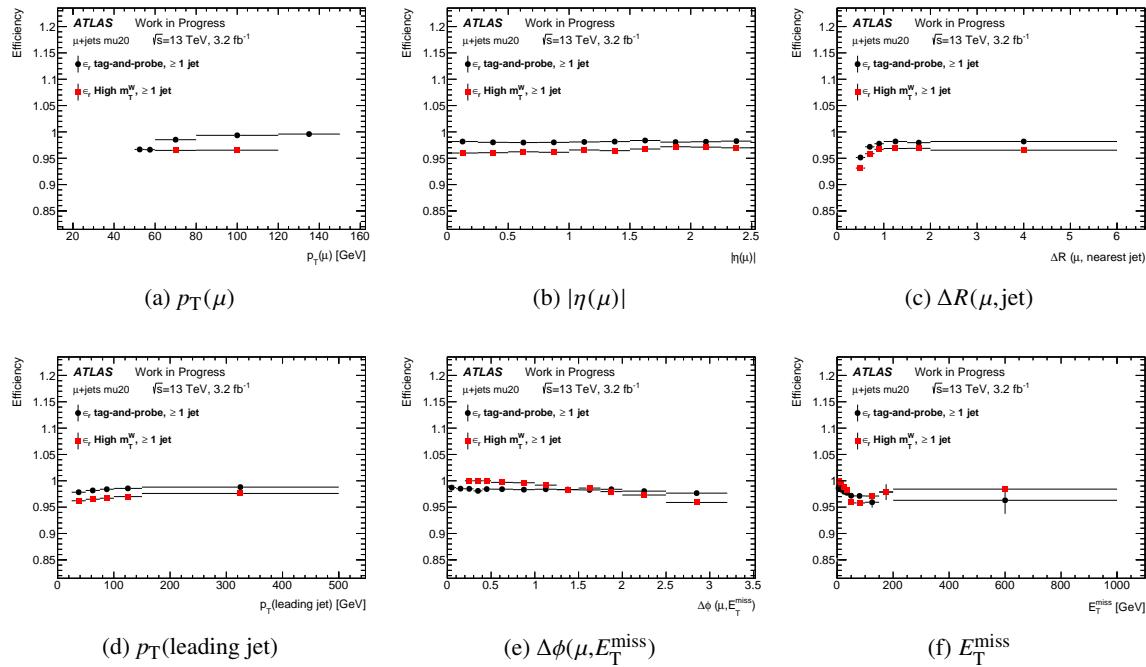


Figure 59: Muon real efficiency $\varepsilon_{\text{real}}$, obtained from the tag-and-probe and the high m_T^W (lept, E_T^{miss}) methods, as a function of (a) the muon p_T , (b) the muon $|\eta|$, (c) the distance ΔR between the muon and its nearest jet, (d) the p_T of the leading jet, (e) the difference $\Delta\phi(\mu, E_T^{\text{miss}})$ in azimuth between the muon and the E_T^{miss} and (f) the E_T^{miss} . $\varepsilon_{\text{real}}$ is shown for events with at least one jet and without any requirement on the number of b -jet, for all muons which match the single muon trigger HLT_MU50 (mu50, see section 5).

¹⁰¹⁶ **B.6. Binning used to measure real muon efficiencies**

The binnings used for the parametrisation of the real muon efficiencies are given in Tables 24 to 27.

Table 24: Measurement bins in muon transverse energy p_T used for the muon real efficiency parametrisation.

Method	Bin boundaries in p_T (GeV)									
tag-and-probe	25	30	35	40	45	50	60	80	100	∞
high- m_T (lept, E_T^{miss})	25	30	35	40	45	50	60	∞		

Table 25: Measurement bins in muon pseudorapidity $|\eta|$ used for the muon real efficiency parametrisation.

Bin boundaries in $ \eta $										
0	0.25	0.5	0.75	1.	1.25	1.5	1.75	2.	2.25	2.50

Table 26: Measurement bins in distance to the nearest jet $\Delta R(\mu, \text{jet})$ used for the muon real efficiency parametrisation.

Bin boundaries in $\Delta R(\mu, \text{jet})$ to the nearest jet								
0.4	0.6	0.8	1	1.5	2	3	4	6

Table 27: Measurement bins in p_T of the leading jet used for the real muon real efficiency parametrisation.

Bin boundaries in p_T of the leading jet (GeV)				
25	50	75	100	150
				∞

Table 28: Measurement bins in $|\Delta\phi|$ between the muon and of the E_T^{miss} used for the muon real efficiency parametrisation.

Bin boundaries in $ \Delta\phi $ between the muon and of the E_T^{miss}													
0	0.1	0.2	0.3	0.4	0.5	0.75	1	1.25	1.5	1.75	2	2.5	π

1018 C. Additional studies on fake electron efficiencies

1019 In this section we show additional studies done on the electron fake efficiency $\varepsilon_{\text{fake}}$ measurement.

1020 To estimate $\varepsilon_{\text{fake}}$ one selects a single electron sample with low $E_{\text{T}}^{\text{miss}}$ and/or low $m_{\text{T}}(\text{lept}, E_{\text{T}}^{\text{miss}})$. These
 1021 regions are contaminated by real leptons, coming in large part from W and Z decays. This contamination
 1022 is estimated, and removed, using simulation. The fake efficiency $\varepsilon_{\text{fake}}$ is extracted as the ratio between
 1023 tight and loose $e+\text{jets}$ events in the selected region.

1024 Selection of events can be summarised as:

- 1025 - Event quality preselection as described in section 5.1.
- 1026 - Exactly one loose electron.
- 1027 - At least one jet in the event.

1028 C.1. Definition of the control regions

1029 Efficiencies are measured in two regions of low $E_{\text{T}}^{\text{miss}}$ and/or low $m_{\text{T}}(\text{lept}, E_{\text{T}}^{\text{miss}})$:

- 1030 - CR_1 is defined by $m_{\text{T}}(\text{lept}, E_{\text{T}}^{\text{miss}}) < 20 \text{ GeV} \& m_{\text{T}}(\text{lept}, E_{\text{T}}^{\text{miss}}) + E_{\text{T}}^{\text{miss}} < 60 \text{ GeV}$.
- 1031 - CR_2 is defined by $E_{\text{T}}^{\text{miss}} < 20 \text{ GeV}$.

Table 29 shows this contamination in the different control regions. In CR_1 a total of nearly 7×10^6 (2×10^6)

Table 29: Contamination (in %) of real electrons in the different control regions, after requiring exactly one electron and at least one jet.

Selection	Control Region	W	Z
Tight	CR_1	13.8	18.9
	CR_2	27.5	13.1
Loose	CR_1	4.5	6.2
	CR_2	12.0	5.8

1032 events are selected in data at loose (tight) level. The contamination is made of nearly 7×10^5 (5×10^5)
 1033 events, estimated from the simulation.
 1034

1035 C.2. Results obtained

1036 Fig. 60 shows $\varepsilon_{\text{fake}}$ as a function of the jet and b -jet multiplicities. Efficiencies are thus derived for two jet
 1037 multiplicities, $n_{\text{jet}} = 1$ and $n_{\text{jet}} \geq 2$ and two b -jet multiplicities, $n_{b-\text{jet}} = 0$ and $n_{b-\text{jet}} \geq 1$ ¹³. Fig. 61 shows
 1038 $\varepsilon_{\text{fake}}$ as a function of the different variables for events with different jet and b -jet multiplicities. $\varepsilon_{\text{fake}}$ has
 1039 a strong dependency with the E_{T} and $|\eta_{\text{cl}}|$ of the electron and $\Delta\phi(\text{elec}, E_{\text{T}}^{\text{miss}})$, smaller with the p_{T} of the
 1040 leading jet. It is almost flat with respect to $\Delta R(\text{elec}, \text{jet})$.

¹³ As for 8 TeV analyses there is not enough statistics to have more bins.

Not reviewed, for internal circulation only

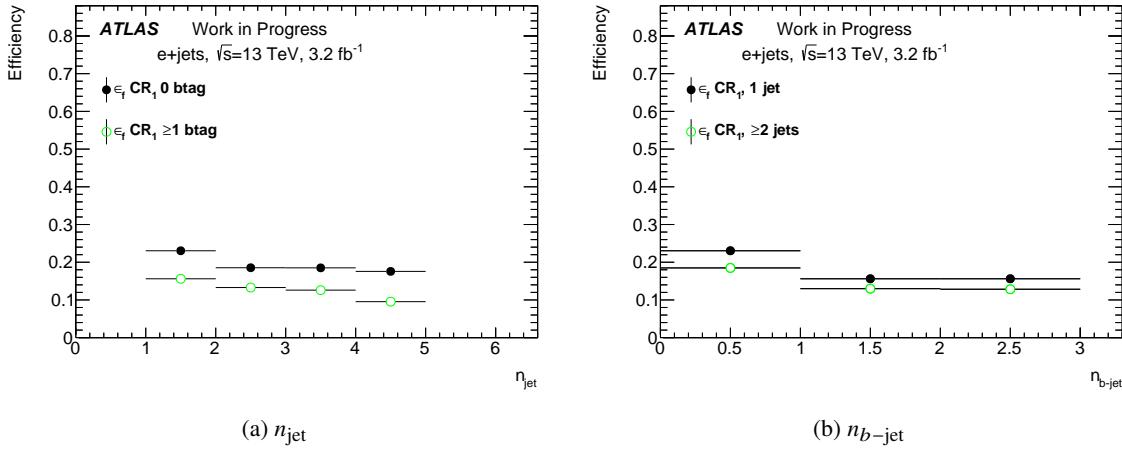


Figure 60: Electron fake efficiency $\varepsilon_{\text{fake}}$ as a function of (a) the number of jets n_{jet} and (b) the number of b -jets $n_{b\text{-jet}}$. $\varepsilon_{\text{fake}}$ is shown for all electrons which match one of the single electron trigger (see section 5).

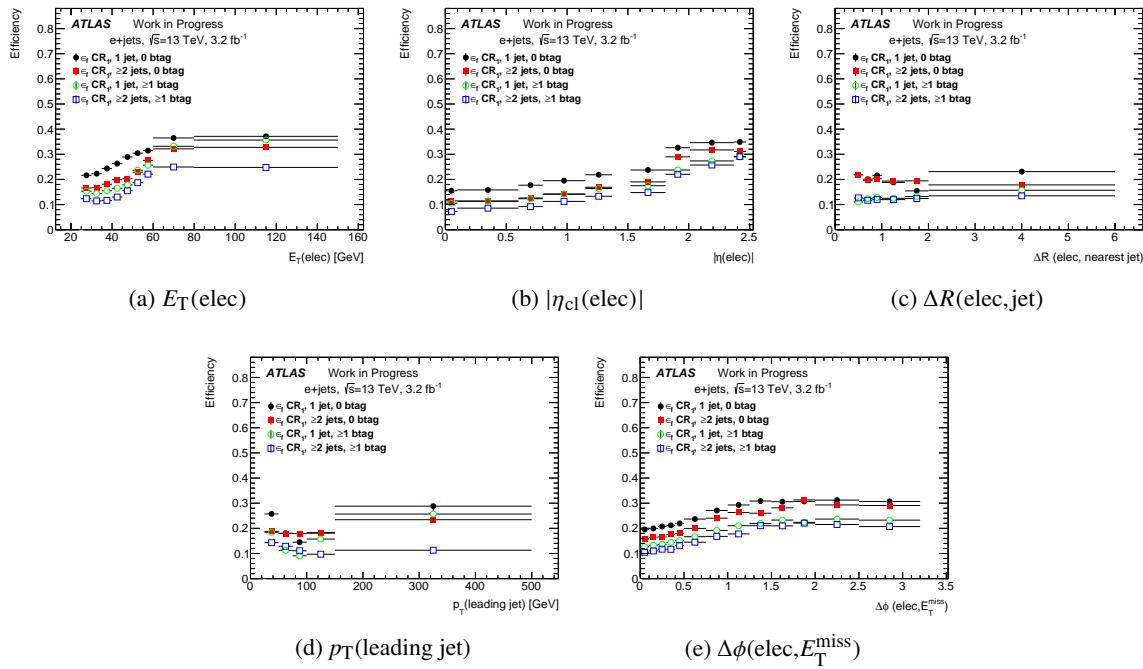


Figure 61: Electron fake efficiency $\varepsilon_{\text{fake}}$ as a function of (a) the electron E_T , (b) the electron $|\eta_{\text{cl}}|$, (c) the distance ΔR between the electron and its nearest jet, (d) the p_T of the leading jet and (e) the distance in azimuth $\Delta\phi(\text{elec}, E_T^{\text{miss}})$. $\varepsilon_{\text{fake}}$ is shown for all electrons which match one of the single electron trigger (see section 5).

1041 **C.3. Impact of the control region definition**

1042 Fig. 62 and 63 show a comparison of efficiencies obtained in the two control regions. One can notice
 1043 that efficiencies evaluated in the low- E_T^{miss} control region (CR_2) are significantly higher than the ones
 1044 derived from the low- $m_T(\text{lept}, E_T^{\text{miss}})$ & E_T^{miss} region (CR_1), especially when a zero-tag sample is selected.
 1045 In events with at least one b -tagged jet, $\varepsilon_{\text{fake}}$ tends to be smaller. In fact in these events, the fake lepton
 1046 is more likely coming from a heavy hadron decay, and the probability to pass the isolation cuts for this
 electron is in general smaller than for fake electrons from the other sources.

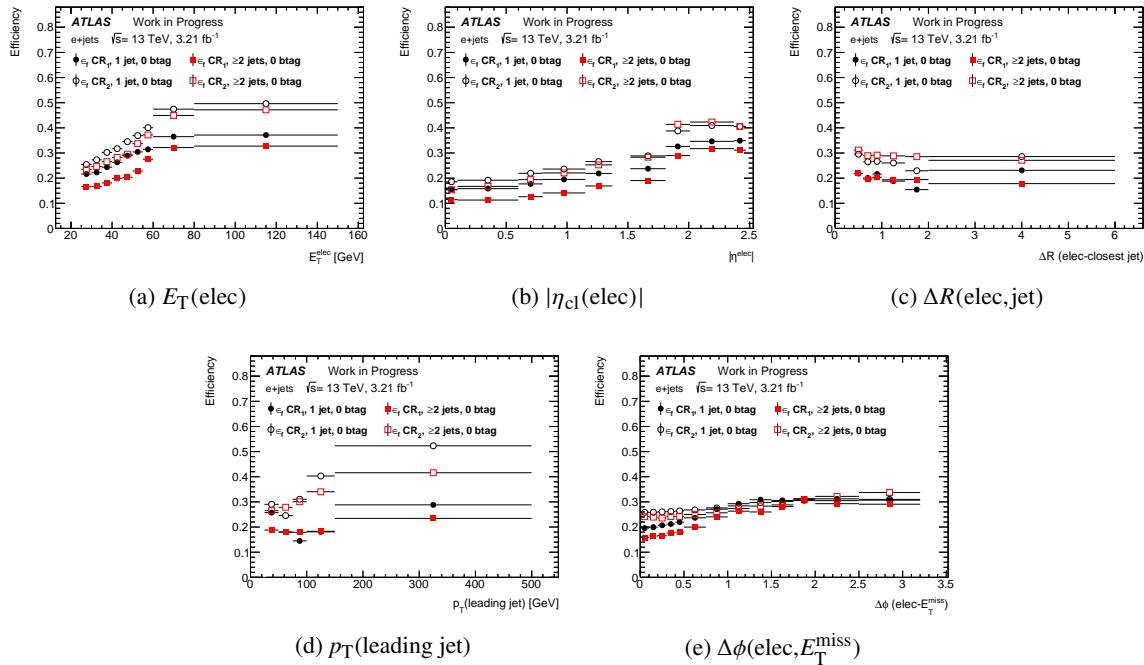


Figure 62: Electron fake efficiency $\varepsilon_{\text{fake}}$ as a function of (a) the electron E_T , (b) the electron $|\eta_{\text{cl}}|$, (c) the distance ΔR between the electron and its nearest jet, (d) the p_T of the leading jet and (e) the distance in azimuth $\Delta\phi(\text{elec}, E_T^{\text{miss}})$. $\varepsilon_{\text{fake}}$ is shown for events with different multiplicities and 0 b -jet, for all electrons which match one of the single electron trigger (see section 5). They are obtained from the different control regions CR_1 and CR_2 .

Not reviewed, for internal circulation only

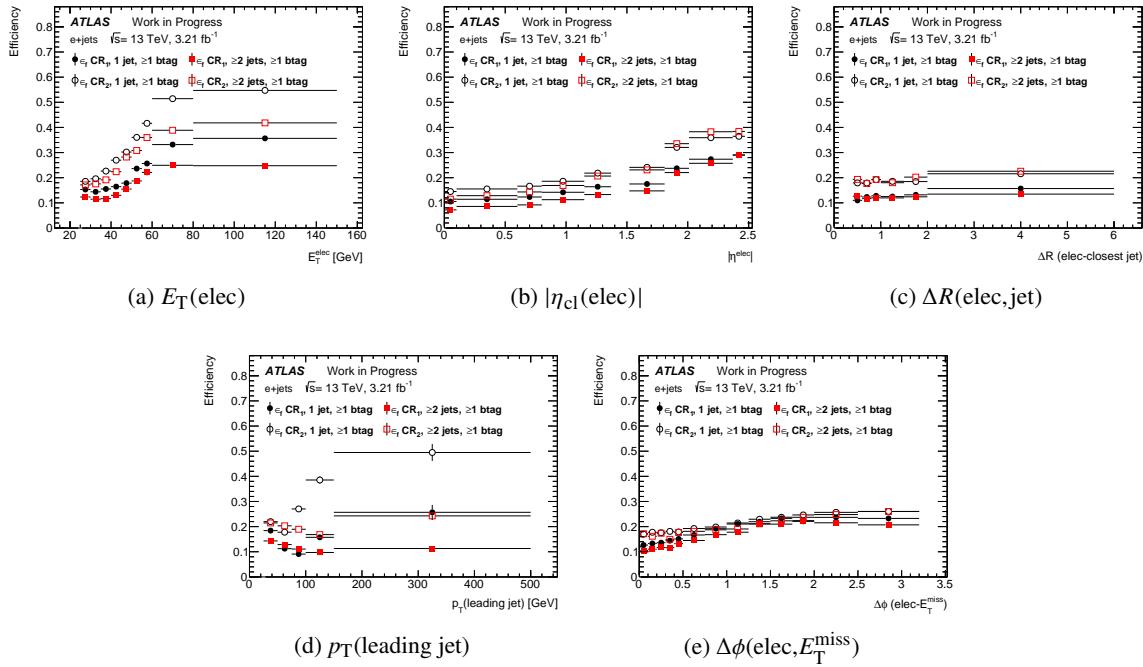


Figure 63: Electron fake efficiency $\varepsilon_{\text{fake}}$ as a function of (a) the electron E_T , (b) the electron $|\eta_{\text{cl}}|$, (c) the distance ΔR between the electron and its nearest jet, (d) the p_T of the leading jet and (e) the distance in azimuth $\Delta\phi(\text{elec}, E_T^{\text{miss}})$ between the electron and the E_T^{miss} . $\varepsilon_{\text{fake}}$ is shown for events with different multiplicities and at least one b -jet, for all electrons which match one of the single electron trigger (see section 5). They are obtained from the different control regions CR_1 and CR_2 .

1048 C.4. Impact of the MC subtraction

1049 In the process to derive the fake efficiencies one step is to subtract the real contamination. This contam-
 1050 ination is taken from MC simulation. Since there are uncertainties on the modeling of the MC and also
 1051 on the overall amount of real leptons the effect of MC subtraction needs to be studied. This is done in this
 1052 study, where the impact of a larger or a smaller MC subtraction on the fake efficiency is studied. For this
 1053 purpose 30% more MC and 30% less MC were subtracted in the fake efficiency calculation process and
 1054 the effect on the derived fake efficiency was studied. Fig. 64 and 65 show the effect of this over and under
 subtraction on the fake efficiency.

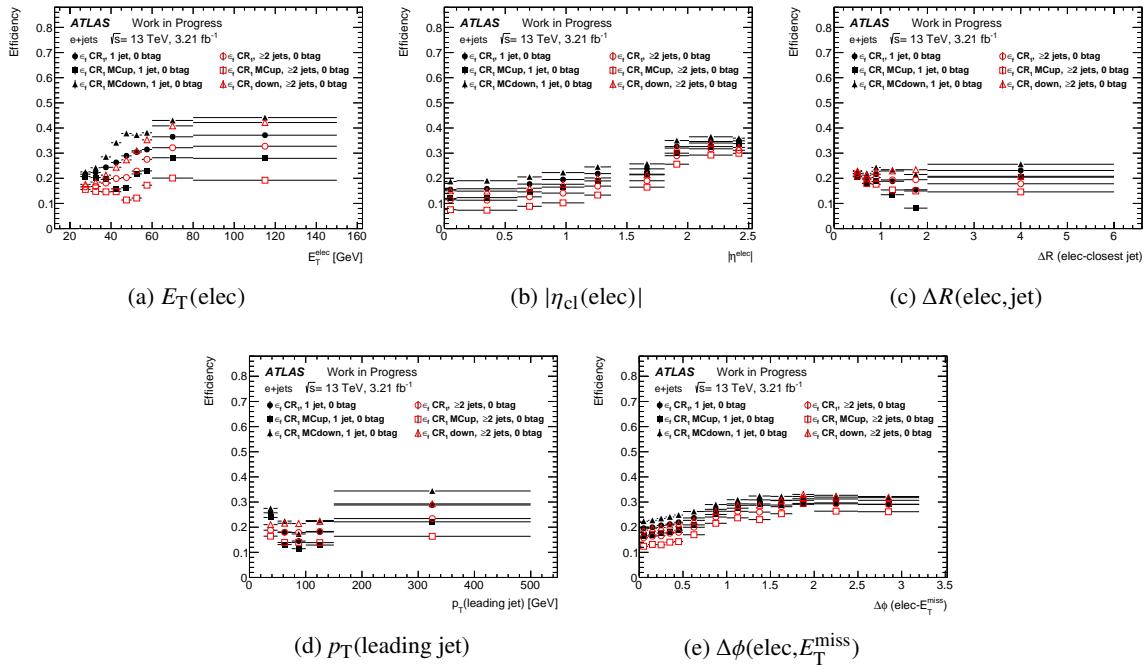


Figure 64: Electron fake efficiency $\varepsilon_{\text{fake}}$ as a function of (a) the electron E_T , (b) the electron $|\eta_{\text{cl}}|$, (c) the distance ΔR between the electron and its nearest jet, (d) the p_T of the leading jet and (e) the distance in azimuth $\Delta\phi(\text{elec}, E_T^{\text{miss}})$ between the electron and the E_T^{miss} . $\varepsilon_{\text{fake}}$ is shown for events with different multiplicities and 0 b-jet, for all electrons which match one of the single electron trigger (see section 5). They are obtained for different levels of MC to be subtracted (nominal, 30% more and 30% less).

Not reviewed, for internal circulation only

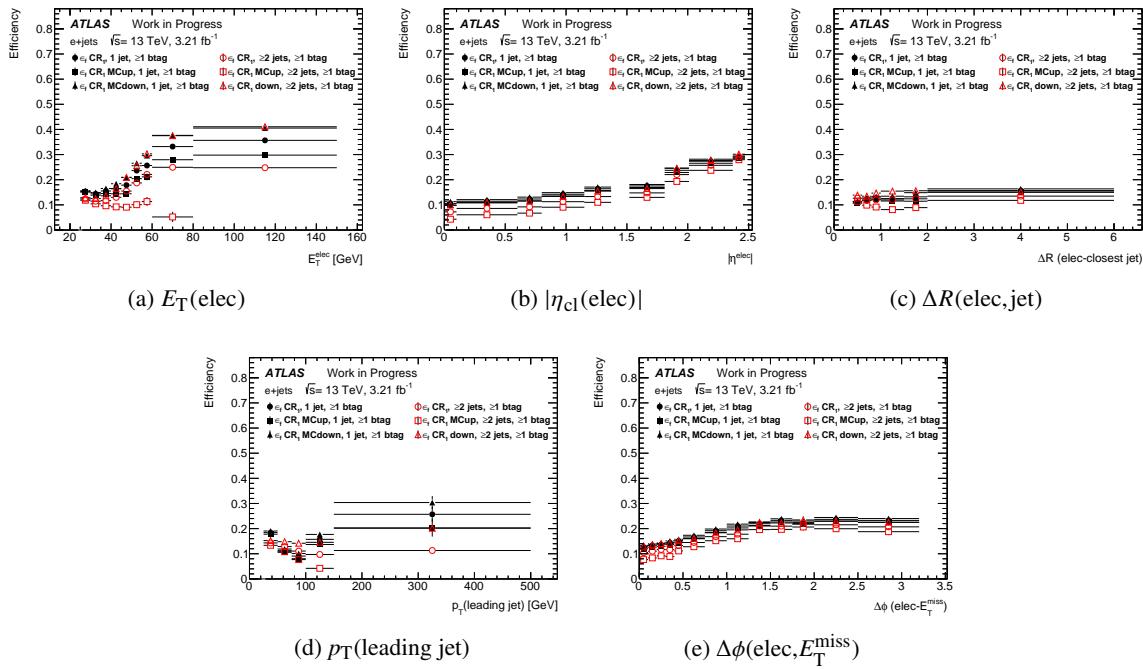


Figure 65: Electron fake efficiency $\varepsilon_{\text{fake}}$ as a function of (a) the electron E_T , (b) the electron $|\eta_{\text{cl}}|$, (c) the distance ΔR between the electron and its nearest jet, (d) the p_T of the leading jet and (e) the distance in azimuth $\Delta\phi(\text{elec}, E_T^{\text{miss}})$. $\varepsilon_{\text{fake}}$ is shown for events with different multiplicities and at least one b -jet, for all electrons which match one of the single electron trigger (see section 5). They are obtained for different levels of MC to be subtracted (nominal, 30% more and 30% less).

1056 C.5. Impact of using different Z/W+jets simulated data

1057 To get a pure fake electron sample, the contamination from real electrons has to be subtracted from the
 1058 selected sample. The biggest concerns are the W and Z contamination. Different samples of simulated
 1059 data are available for these two processes, SHERPA which are used by default but also based on POWHEG
 1060 +PYTHIA and MADGRAPH +PYTHIA. Fig. 66 and 67 show the effect of using these different simulations. The
 1061 efficiencies obtained from the different simulated data are close even if the ones derived using POWHEG
 are a bit higher.

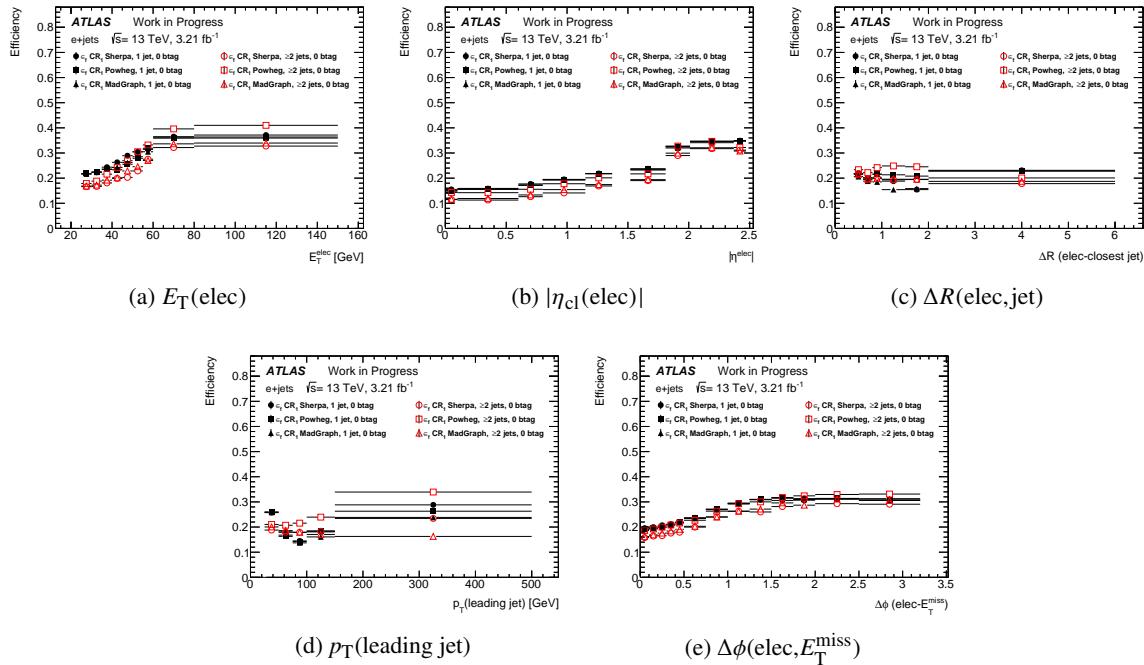


Figure 66: Electron fake efficiency $\varepsilon_{\text{fake}}$ as a function of (a) the electron E_T , (b) the electron $|\eta_{\text{cl}}|$, (c) the distance ΔR between the electron and its nearest jet, (d) the p_T of the leading jet and (e) the distance in azimuth $\Delta\phi(\text{elec}, E_T^{\text{miss}})$ between the electron and the E_T^{miss} . $\varepsilon_{\text{fake}}$ is shown for events with different multiplicities and 0 b -jet, for all electrons which match one of the single electron trigger (see section 5). They are obtained using different $W/Z + \text{jets}$ simulated data to estimate the real lepton contamination.

Not reviewed, for internal circulation only

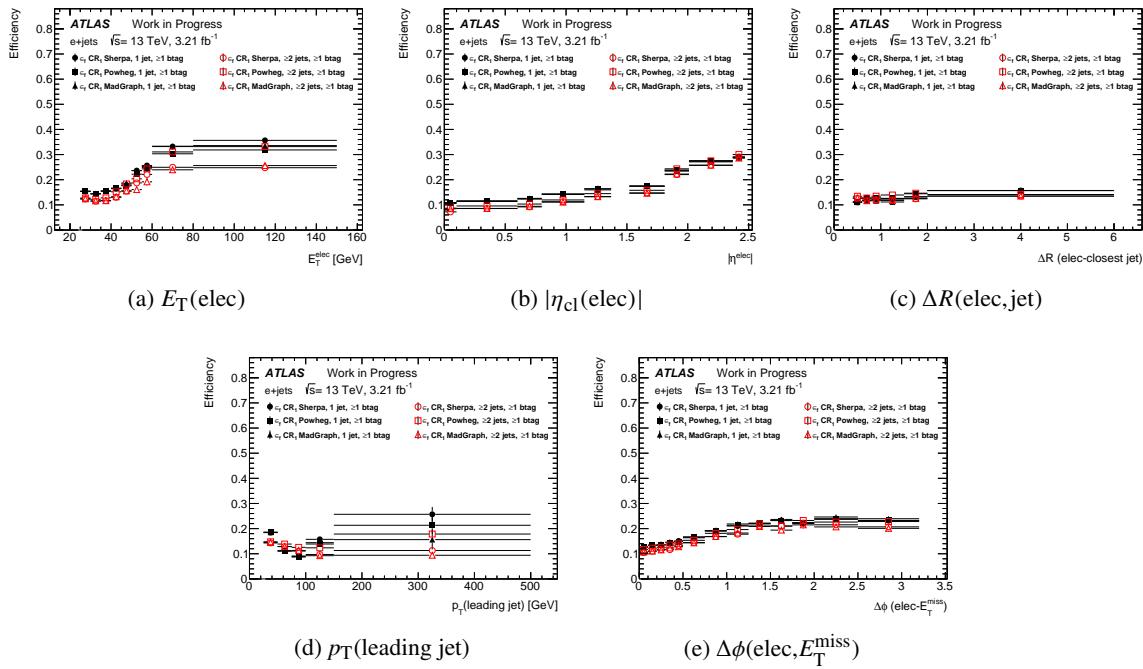


Figure 67: Electron fake efficiency $\varepsilon_{\text{fake}}$ as a function of (a) the electron E_T , (b) the electron $|\eta_{\text{cl}}|$, (c) the distance ΔR between the electron and its nearest jet, (d) the p_T of the leading jet and (e) the distance in azimuth $\Delta\phi(\text{elec}, E_T^{\text{miss}})$ between the electron and the E_T^{miss} . $\varepsilon_{\text{fake}}$ is shown for events with different multiplicities and at least one b -jet, for all electrons which match one of the single electron trigger (see section 5). They are obtained using different $W/Z + \text{jets}$ simulated data to estimate the real lepton contamination.

1063 C.6. Composition of non-prompt and fake electrons

1064 On the contrary of the 8 TeV analysis [2], the composition of non-prompt and fake electrons was not
1065 studied. In particular, there is no information in the flat tuples used for this study if the electron is
1066 originated from a conversion or not.

[Not reviewed, for internal circulation only]

¹⁰⁶⁷ **C.7. Binning used to measure fake electron efficiencies**

The binnings used for the parametrisation of the electron fake efficiencies are given in Tables 30 to 33.

Table 30: Measurement bins in electron transverse energy E_T used for the fake electron efficiency parametrisation.

Bin boundaries in E_T (GeV)									
25	30	35	40	45	50	60	80	∞	

Table 31: Measurement bins in electron pseudorapidity $|\eta_{\text{cl}}|$ used for the electron fake efficiency parametrisation.

Bin boundaries in $ \eta_{\text{cl}} $										
0	0.1	0.6	0.8	1.15	1.37	1.52	1.81	2.01	2.37	2.47

Table 32: Measurement bins in distance to the nearest jet $\Delta R(\text{elec}, \text{jet})$ used for the electron fake efficiency parametrisation.

Bin boundaries in $\Delta R(\text{elec}, \text{jet})$ to the nearest jet						
0.4	0.6	0.8	1	1.5	2.0	6

Table 33: Measurement bins in p_T of the leading jet used for the electron fake efficiency parametrisation.

Bin boundaries in p_T of the leading jet(GeV)					
25	50	75	100	150	∞

¹⁰⁶⁸

Table 34: Measurement bins in $|\Delta\phi|(\text{elec}, E_T^{\text{miss}})$ the distance in azimuth between the electron and of the E_T^{miss} used for the electron fake efficiency parametrisation.

Bin boundaries in $ \Delta\phi (\text{elec}, E_T^{\text{miss}})$													
0	0.1	0.2	0.3	0.4	0.5	0.75	1	1.25	1.5	1.75	2	2.5	π

¹⁰⁶⁹

1070 D. Additional studies on fake muon efficiencies

1071 In this section we show additional studies done on the muon fake efficiency $\varepsilon_{\text{fake}}$ measurement.

1072 To estimate $\varepsilon_{\text{fake}}$ one selects a single muon sample with low $E_{\text{T}}^{\text{miss}}$ and/or low $m_{\text{T}}(\text{lept}, E_{\text{T}}^{\text{miss}})$. These
 1073 regions are contaminated by real leptons, coming in large part from W and Z decays. This contamination
 1074 is estimated, and removed, using simulation. The fake efficiency $\varepsilon_{\text{fake}}$ is extracted as the ratio between
 1075 tight and loose $\mu + \text{jets}$ events in the selected region.

1076 Selection of events can be summarised as:

- 1077 - Event quality preselection as described in section 5.1.
- 1078 - Exactly one loose muon.
- 1079 - At least one jet in the event.

1080 D.1. Definition of the control regions

1081 Efficiencies are measured in two regions of low $E_{\text{T}}^{\text{miss}}$ and/or low $m_{\text{T}}(\text{lept}, E_{\text{T}}^{\text{miss}})$:

- 1082 - CR_1 is defined by $m_{\text{T}}(\text{lept}, E_{\text{T}}^{\text{miss}}) < 20 \text{ GeV} \& \& m_{\text{T}}(\text{lept}, E_{\text{T}}^{\text{miss}}) + E_{\text{T}}^{\text{miss}} < 60 \text{ GeV}$.
- 1083 - CR_2 is defined by $E_{\text{T}}^{\text{miss}} < 20 \text{ GeV}$.

1084 Table 35 shows this contamination in the different control regions. For the mu20i trigger, in CR_1 a total
 1085 of nearly 1.3×10^6 (7×10^5) events are selected in data at loose (tight) level. The contamination is made
 1086 of nearly 1.3×10^5 (1.2×10^5) events, estimated from the simulation. For the mu50 trigger, in CR_1 a total
 1087 of nearly 1×10^5 (8×10^5) events are selected in data at loose (tight) level. The contamination is made of
 nearly 5×10^5 (4×10^5) events, estimated from the simulation.

Table 35: Contamination (in %) of real muons in the different control regions, after requiring exactly one muon and at least one jet.

Selection	Control Region	mu20i		mu50	
		W	Z	W	Z
Tight	CR_1	11.9	5.3	44.3	8.2
	CR_2	20.4	4.2	48.0	5.1
Loose	CR_1	6.5	2.9	37.7	7.0
	CR_2	12.1	2.5	44.2	4.7

1088

1089 **D.2. Results obtained**

1090 Fig. 68 shows $\varepsilon_{\text{fake}}$ as a function of the jet and b -jet multiplicities. Efficiencies are thus derived for two jet multiplicities, $n_{\text{jet}} = 1$ and $n_{\text{jet}} \geq 2$ and two b -jet multiplicities, $n_{b\text{-jet}} = 0$ and $n_{b\text{-jet}} \geq 1$ ¹⁴.

Not reviewed, for internal circulation only

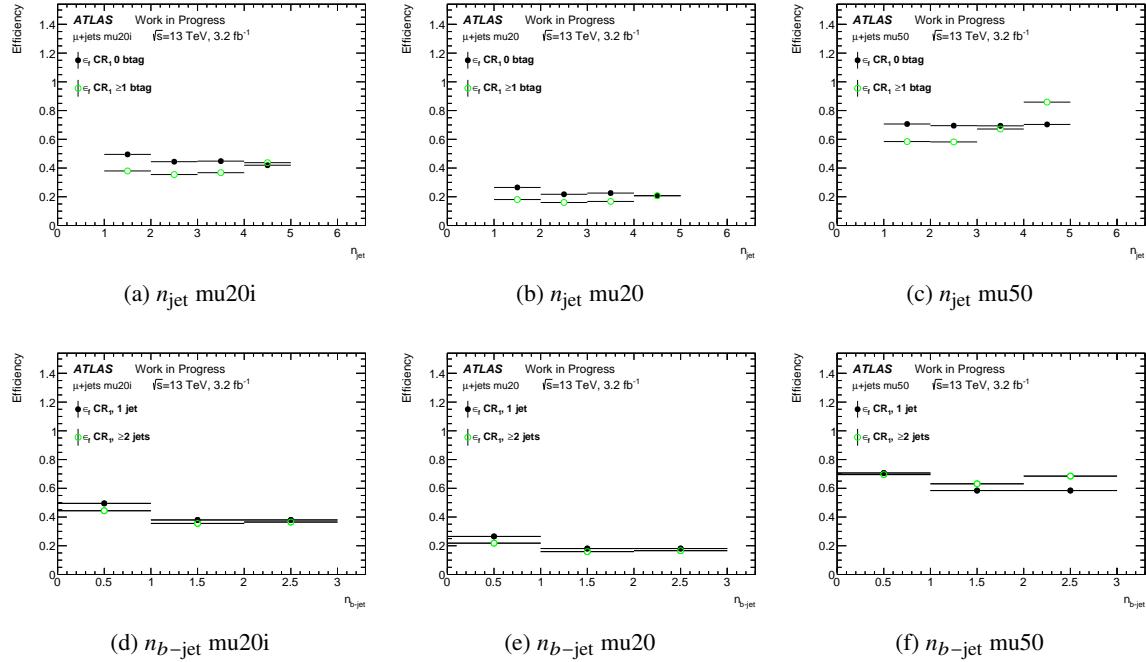


Figure 68: Muon fake efficiency $\varepsilon_{\text{fake}}$ as a function of (a,b,c) the number of jets n_{jet} and (d,e,f) the number of b -jets $n_{b\text{-jet}}$. $\varepsilon_{\text{fake}}$ is shown for all muons which match one of the single muon trigger (a,d) mu20i, (b,e) mu20 and (c,f) mu50 (see section 5).

1091

1092 Fig. 69 to 71 show $\varepsilon_{\text{fake}}$ as a function of the different variables for events with different jet and b -jet
 1093 multiplicities. $\varepsilon_{\text{fake}}$ has a strong dependency with the E_{T} and $|\eta|$ of the muon and $\Delta\phi(\mu, E_{\text{T}}^{\text{miss}})$, smaller
 1094 with the p_{T} of the leading jet. It is almost flat with respect to $\Delta R(\mu, \text{jet})$.

¹⁴ As for 8 TeV analyses there is not enough statistics to have more bins.

Not reviewed, for internal circulation only

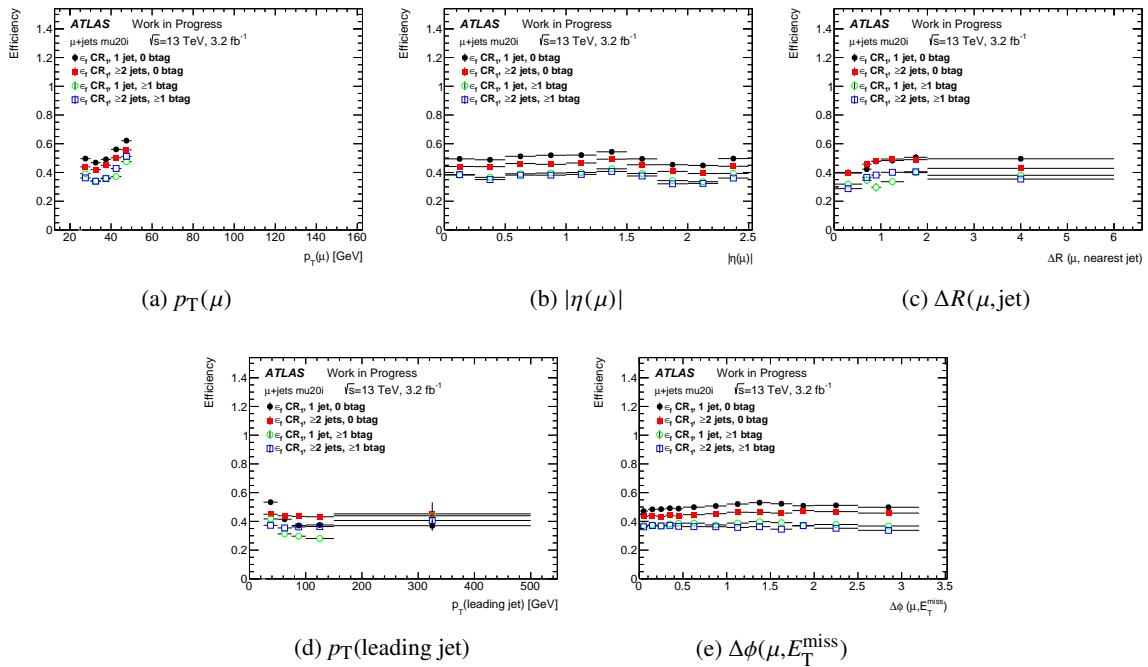


Figure 69: Muon fake efficiency $\varepsilon_{\text{fake}}$ as a function of (a) the muon p_T , (b) the muon $|\eta|$, (c) the distance ΔR between the muon and its nearest jet, (d) the p_T of the leading jet and (e) the distance in azimuth $\Delta\phi(\mu, E_T^{\text{miss}})$ between the muon and the E_T^{miss} . $\varepsilon_{\text{fake}}$ is shown for all muons which match the single muon trigger HLT_MU20_ILOOSE_L1MU15 (mu20i, see section 5).

Not reviewed, for internal circulation only

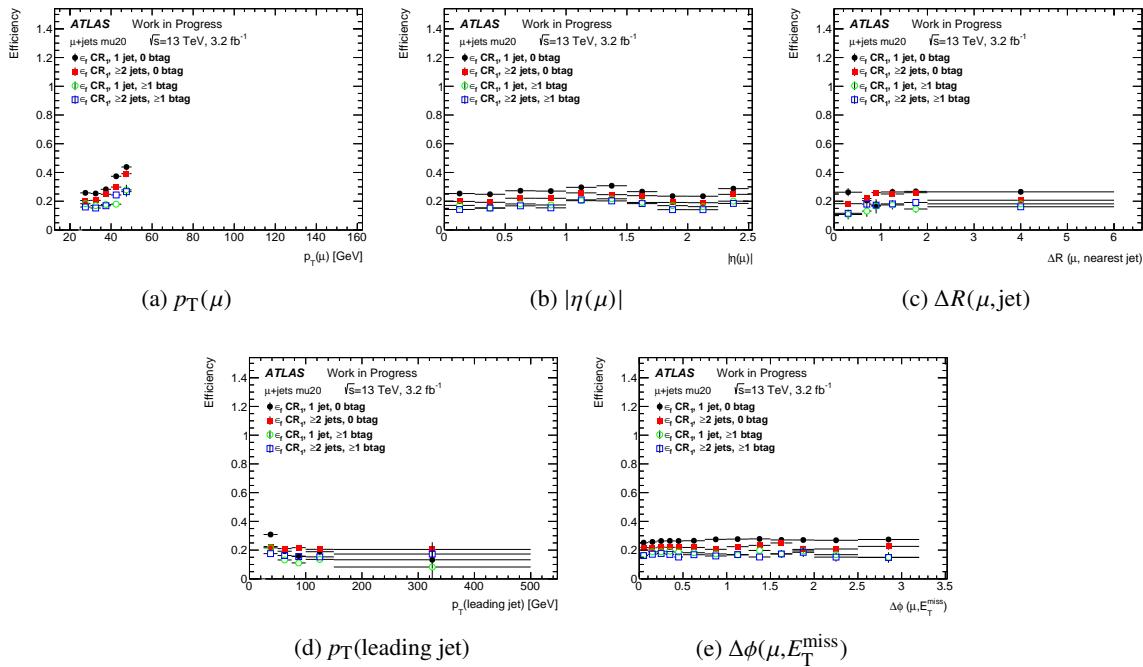


Figure 70: Muon fake efficiency $\varepsilon_{\text{fake}}$ as a function of (a) the muon p_T , (b) the muon $|\eta|$, (c) the distance ΔR between the muon and its nearest jet, (d) the p_T of the leading jet and (e) the distance in azimuth $\Delta\phi(\mu, E_T^{\text{miss}})$ between the muon and the E_T^{miss} . $\varepsilon_{\text{fake}}$ is shown for all muons which match the single muon trigger HLT_MU20_L1MU15 (mu20, see section 5).

Not reviewed, for internal circulation only

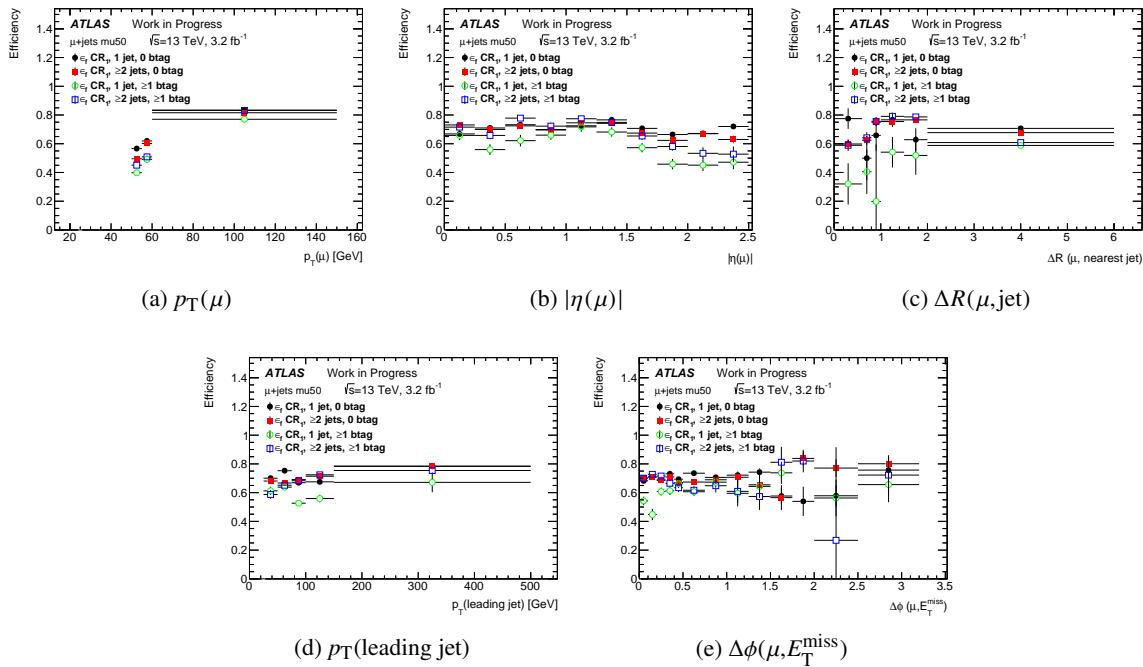


Figure 71: Muon fake efficiency $\varepsilon_{\text{fake}}$ as a function of (a) the muon p_T , (b) the muon $|\eta|$, (c) the distance ΔR between the muon and its nearest jet, (d) the p_T of the leading jet and (e) the distance in azimuth $\Delta\phi(\mu, E_T^{\text{miss}})$ between the muon and the E_T^{miss} . $\varepsilon_{\text{fake}}$ is shown for all muons which match the single muon trigger HLT_MU50 (mu50, see section 5).

Not reviewed, for internal circulation only

1095 D.3. Impact of the control region definition

1096 Fig. 72 to 77 show a comparison of efficiencies obtained in the two control regions. One can notice
 1097 that efficiencies evaluated in the low- E_T^{miss} control region (CR_2) are significantly higher than the ones
 1098 derived from the low- $m_T(\text{lept}, E_T^{\text{miss}})$ & E_T^{miss} region (CR_1), especially when a zero-tag sample is selected.
 1099 In events with at least one b -tagged jet, $\varepsilon_{\text{fake}}$ tends to be smaller. In fact in these events, the fake lepton
 1100 is more likely coming from a heavy hadron decay, and the probability to pass the isolation cuts for this
 muon is in general smaller than for fake muons from the other sources.

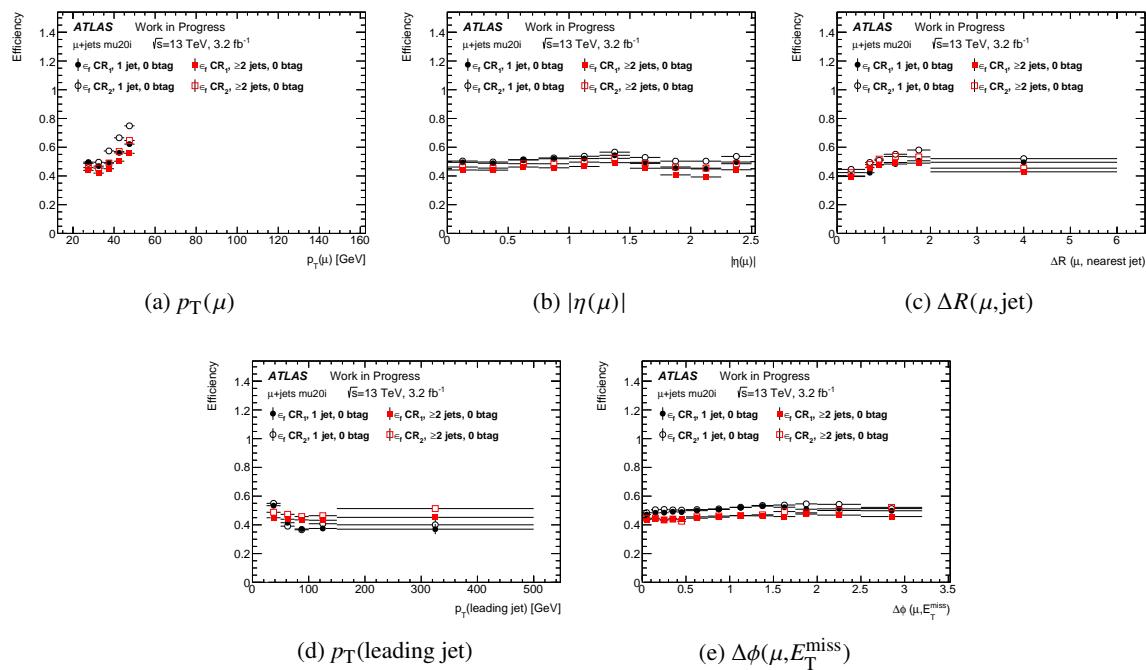


Figure 72: Muon fake efficiency $\varepsilon_{\text{fake}}$ as a function of (a) the muon p_T , (b) the muon $|\eta|$, (c) the distance ΔR between the muon and its nearest jet, (d) the p_T of the leading jet and (e) the distance in azimuth $\Delta\phi(\mu, E_T^{\text{miss}})$ between the muon and the E_T^{miss} . $\varepsilon_{\text{fake}}$ is shown for events with different multiplicities and 0 b -jet, for all muons which match the single muon trigger HLT_MU20_ILOOSE_L1MU15 (mu20i, see section 5). They are obtained from the different control regions CR_1 and CR_2 .

Not reviewed, for internal circulation only

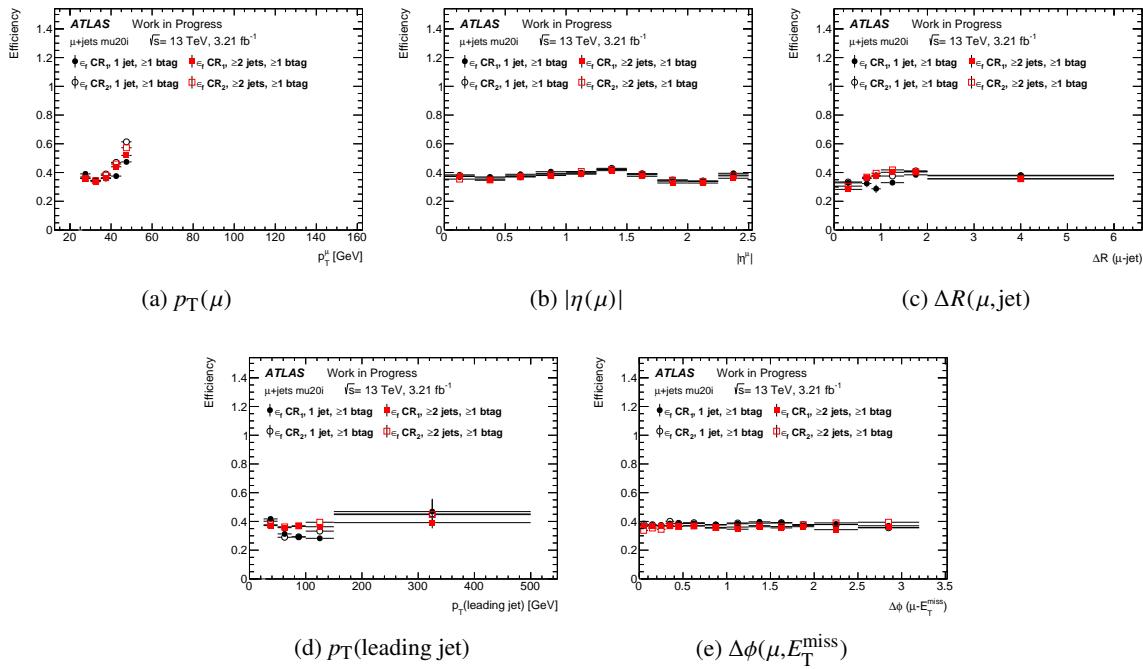


Figure 73: Muon fake efficiency $\varepsilon_{\text{fake}}$ as a function of (a) the muon p_T , (b) the muon $|\eta|$, (c) the distance ΔR between the muon and its nearest jet, (d) the p_T of the leading jet and (e) the distance in azimuth $\Delta\phi(\mu, E_T^{\text{miss}})$ between the muon and the E_T^{miss} . $\varepsilon_{\text{fake}}$ is shown for events with different multiplicities and at least one b -jet, for all muons which match the single muon trigger HLT_MU20_ILOOSE_L1MU15 (mu20i, see section 5). They are obtained from the different control regions CR_1 and CR_2 .

Not reviewed, for internal circulation only

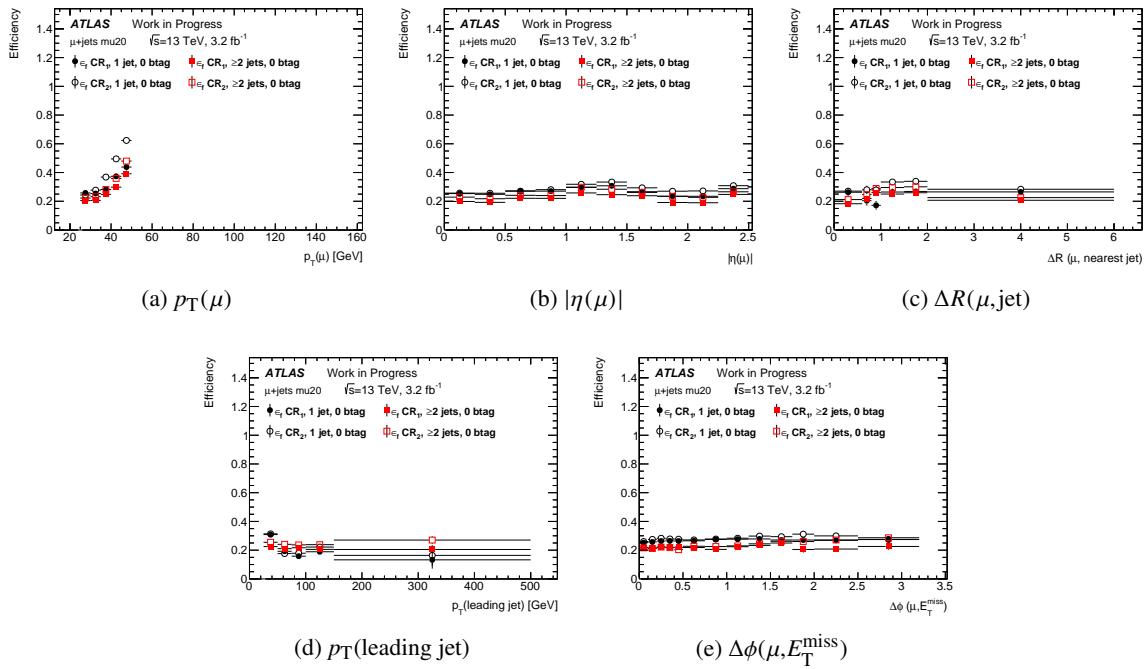


Figure 74: Muon fake efficiency $\varepsilon_{\text{fake}}$ as a function of (a) the muon p_T , (b) the muon $|\eta|$, (c) the distance ΔR between the muon and its nearest jet, (d) the p_T of the leading jet and (e) the distance in azimuth $\Delta\phi(\mu, E_T^{\text{miss}})$ between the muon and the E_T^{miss} . $\varepsilon_{\text{fake}}$ is shown for events with different multiplicities and 0 b -jet, for all muons which match the single muon trigger HLT_MU20_L1MU15 (mu20, see section 5). They are obtained from the different control regions CR_1 and CR_2 .

Not reviewed, for internal circulation only

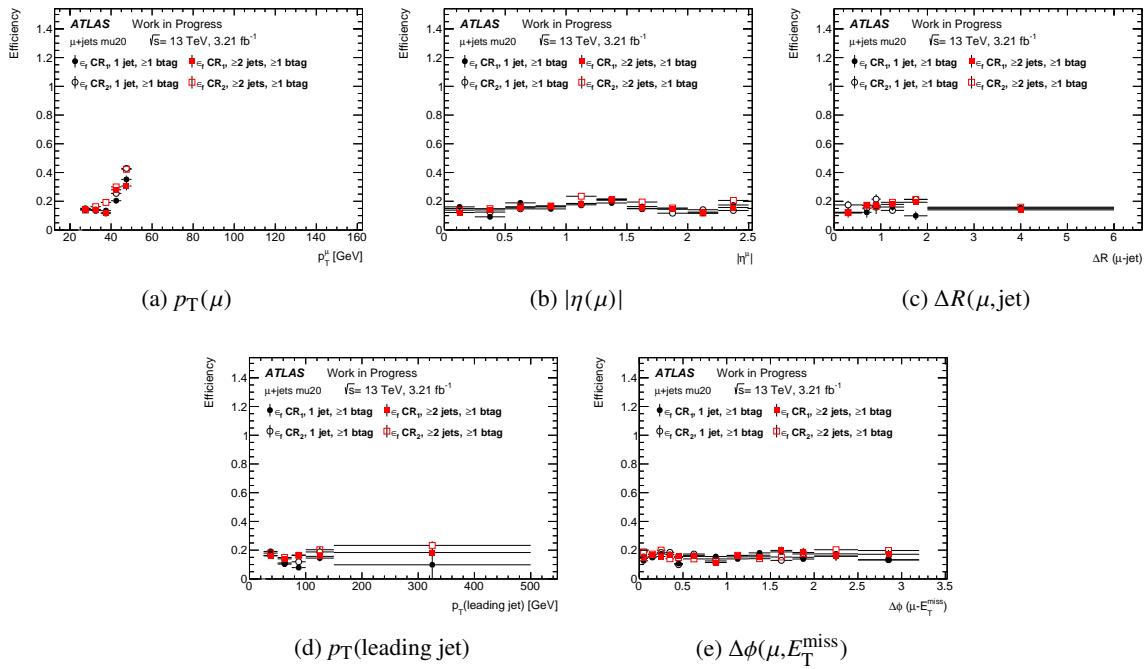


Figure 75: Muon fake efficiency $\varepsilon_{\text{fake}}$ as a function of (a) the muon p_T , (b) the muon $|\eta|$, (c) the distance ΔR between the muon and its nearest jet, (d) the p_T of the leading jet and (e) the distance in azimuth $\Delta\phi(\mu, E_T^{\text{miss}})$ between the muon and the E_T^{miss} . $\varepsilon_{\text{fake}}$ is shown for events with different multiplicities and at least one b -jet, for all muons which match the single muon trigger HLT_MU20_L1MU15 (mu20, see section 5). They are obtained from the different control regions CR_1 and CR_2 .

Not reviewed, for internal circulation only

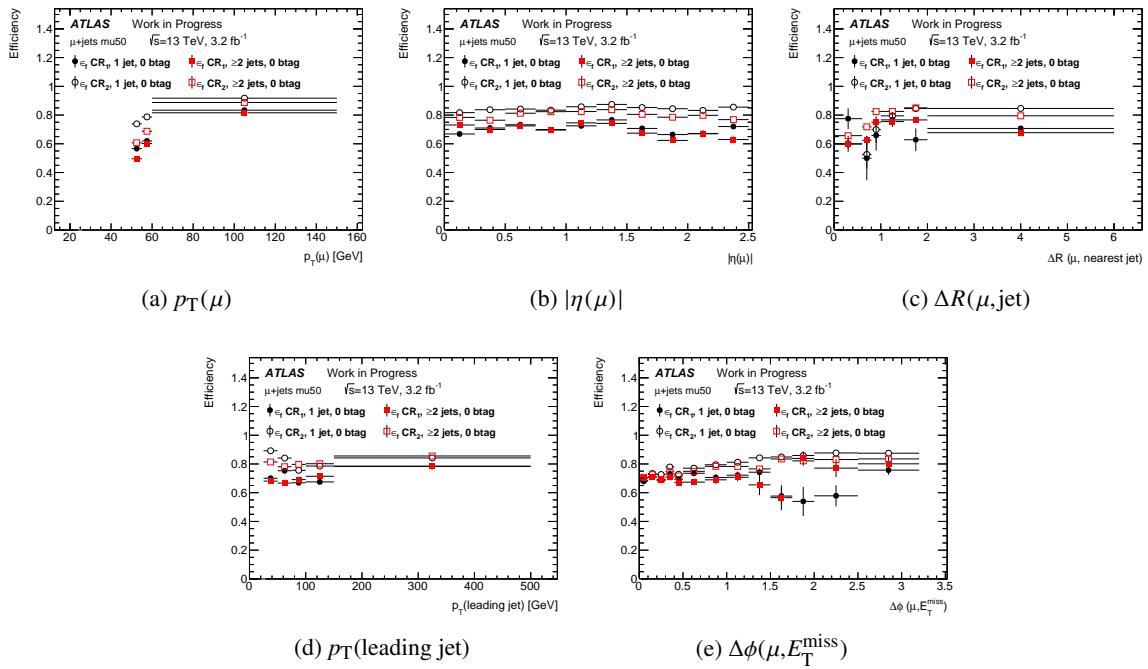


Figure 76: Muon fake efficiency $\varepsilon_{\text{fake}}$ as a function of (a) the muon p_T , (b) the muon $|\eta|$, (c) the distance ΔR between the muon and its nearest jet, (d) the p_T of the leading jet and (e) the distance in azimuth $\Delta\phi(\mu, E_T^{\text{miss}})$ between the muon and the E_T^{miss} . $\varepsilon_{\text{fake}}$ is shown for events with different multiplicities and 0 b -jet, for all muons which match the single muon trigger HLT_MU50 (mu50, see section 5). They are obtained from the different control regions CR_1 and CR_2 .

Not reviewed, for internal circulation only

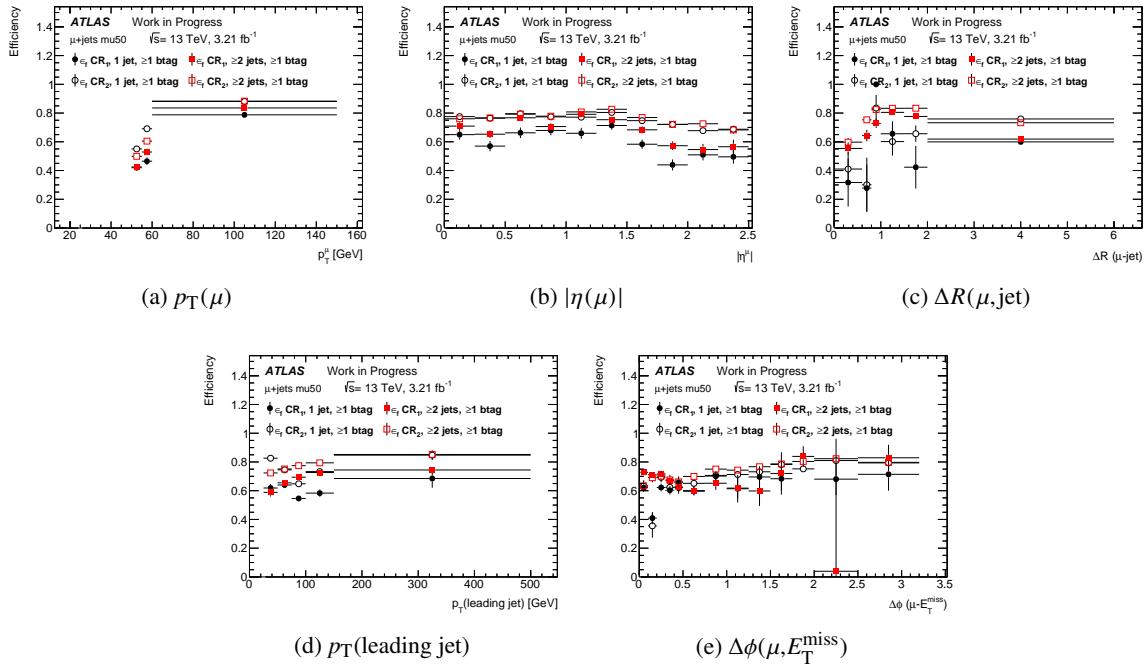


Figure 77: Muon fake efficiency $\varepsilon_{\text{fake}}$ as a function of (a) the muon p_T , (b) the muon $|\eta|$, (c) the distance ΔR between the muon and its nearest jet, (d) the p_T of the leading jet and (e) the distance in azimuth $\Delta\phi(\mu, E_T^{\text{miss}})$ between the muon and the E_T^{miss} . $\varepsilon_{\text{fake}}$ is shown for events with different multiplicities and at least one b -jet, for all muons which match the single muon trigger HLT_MU50 (mu50, see section 5). They are obtained from the different control regions CR_1 and CR_2 .

1102 D.4. Impact of the MC subtraction

1103 In the process to derive the fake efficiencies one step is to subtract the real contamination. This contam-
 1104 ination is taken from MC simulation. Since there are uncertainties on the modeling of the MC and also
 1105 on the overall amount of real leptons the effect of MC subtraction needs to be studied. This is done in this
 1106 study, where the impact of a larger or a smaller MC subtraction on the fake efficiency is studied. For this
 1107 purpose 30% more MC and 30% less MC were subtracted in the fake efficiency calculation process and
 1108 the effect on the derived fake efficiency was studied. Fig. 78 to 83 show the effect of this over and under
 subtraction on the fake efficiency.

Not reviewed, for internal circulation only

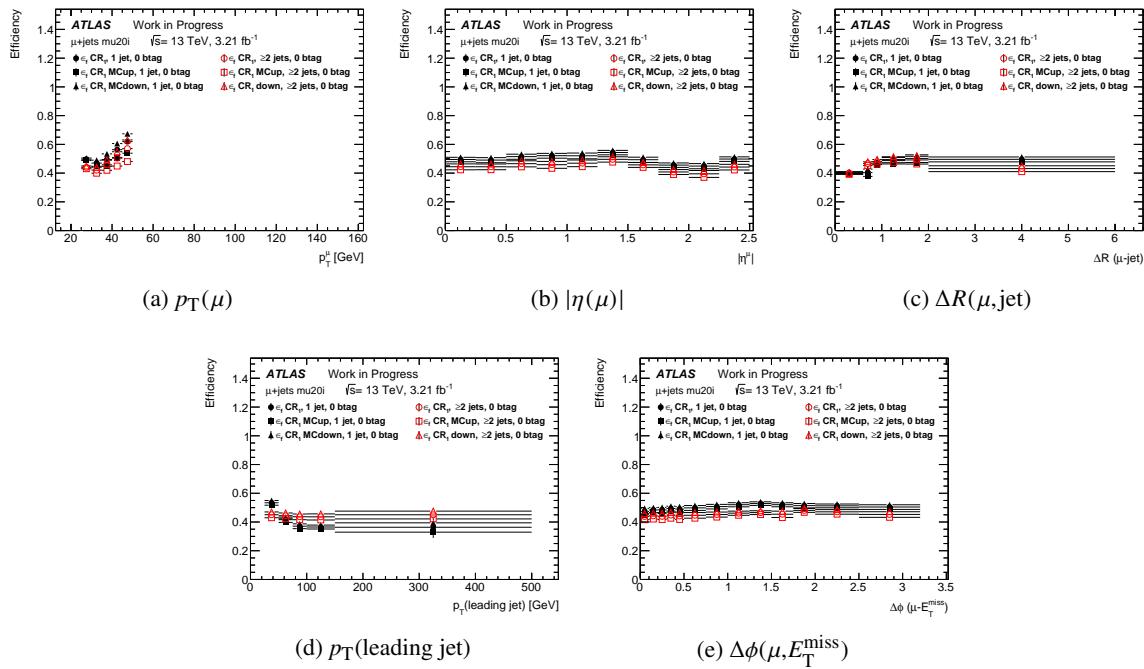


Figure 78: Muon fake efficiency $\varepsilon_{\text{fake}}$ as a function of (a) the muon p_T , (b) the muon $|\eta|$, (c) the distance ΔR between the muon and its nearest jet, (d) the p_T of the leading jet and (e) the distance in azimuth $\Delta\phi(\mu, E_T^{\text{miss}})$ between the muon and the E_T^{miss} . $\varepsilon_{\text{fake}}$ is shown for events with different multiplicities and 0 b -jet, for all muons which match the single muon trigger HLT_MU20_ILOOSE_L1MU15 (mu20i, see section 5). They are obtained for different levels of MC to be subtracted (nominal, 30% more and 30% less).

Not reviewed, for internal circulation only

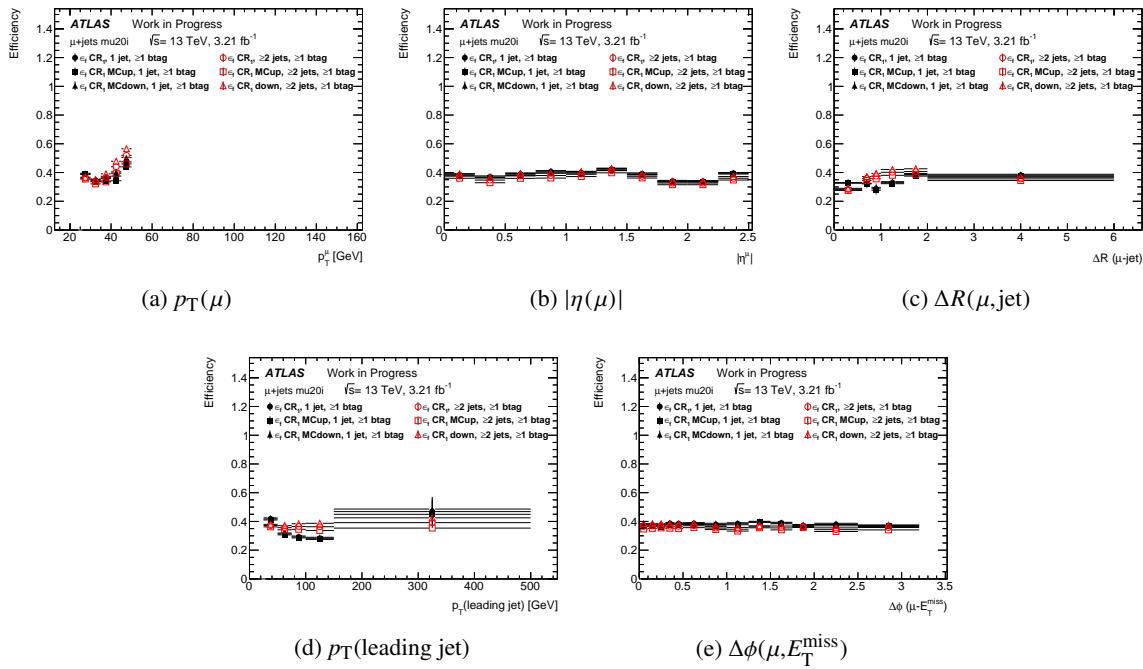


Figure 79: Muon fake efficiency $\varepsilon_{\text{fake}}$ as a function of (a) the muon p_T , (b) the muon $|\eta|$, (c) the distance ΔR between the muon and its nearest jet, (d) the p_T of the leading jet and (e) the distance in azimuth $\Delta\phi(\mu, E_T^{\text{miss}})$ between the muon and the E_T^{miss} . $\varepsilon_{\text{fake}}$ is shown for events with different multiplicities and at least one b -jet, for all muons which match the single muon trigger HLT_MU20_ILOOSE_L1MU15 (mu20i, see section 5). They are obtained for different levels of MC to be subtracted (nominal, 30% more and 30% less).

Not reviewed, for internal circulation only

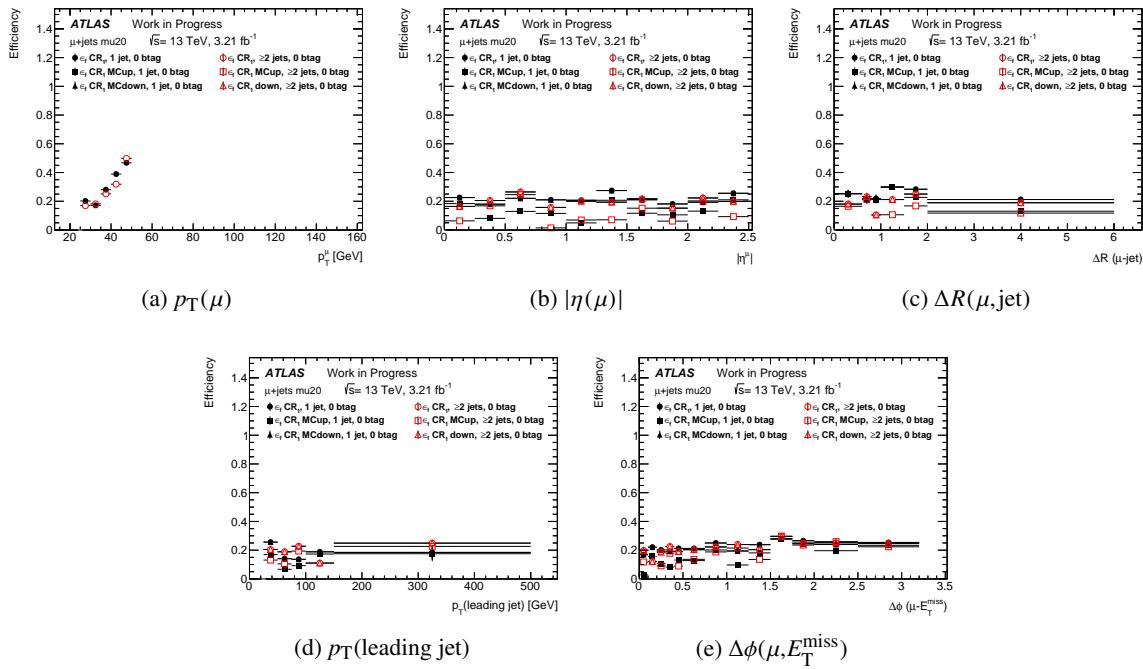


Figure 80: Muon fake efficiency $\varepsilon_{\text{fake}}$ as a function of (a) the muon p_T , (b) the muon $|\eta|$, (c) the distance ΔR between the muon and its nearest jet, (d) the p_T of the leading jet and (e) the distance in azimuth $\Delta\phi(\mu, E_T^{\text{miss}})$ between the muon and the E_T^{miss} . $\varepsilon_{\text{fake}}$ is shown for events with different multiplicities and 0 b -jet, for all muons which match the single muon trigger HLT_MU20_L1MU15 (mu20, see section 5). They are obtained for different levels of MC to be subtracted (nominal, 30% more and 30% less).

Not reviewed, for internal circulation only

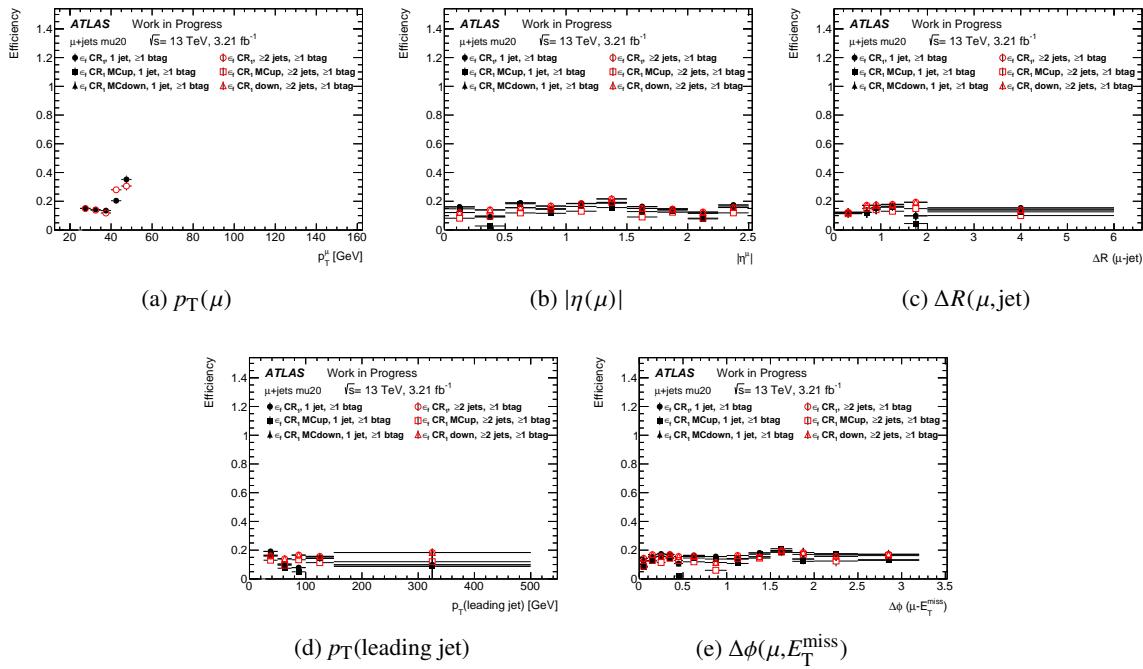


Figure 81: Muon fake efficiency $\varepsilon_{\text{fake}}$ as a function of (a) the muon p_T , (b) the muon $|\eta|$, (c) the distance ΔR between the muon and its nearest jet, (d) the p_T of the leading jet and (e) the distance in azimuth $\Delta\phi(\mu, E_T^{\text{miss}})$ between the muon and the E_T^{miss} . $\varepsilon_{\text{fake}}$ is shown for events with different multiplicities and at least one b -jet, for all muons which match the single muon trigger HLT_MU20_L1MU15 (mu20, see section 5). They are obtained for different levels of MC to be subtracted (nominal, 30% more and 30% less).

Not reviewed, for internal circulation only

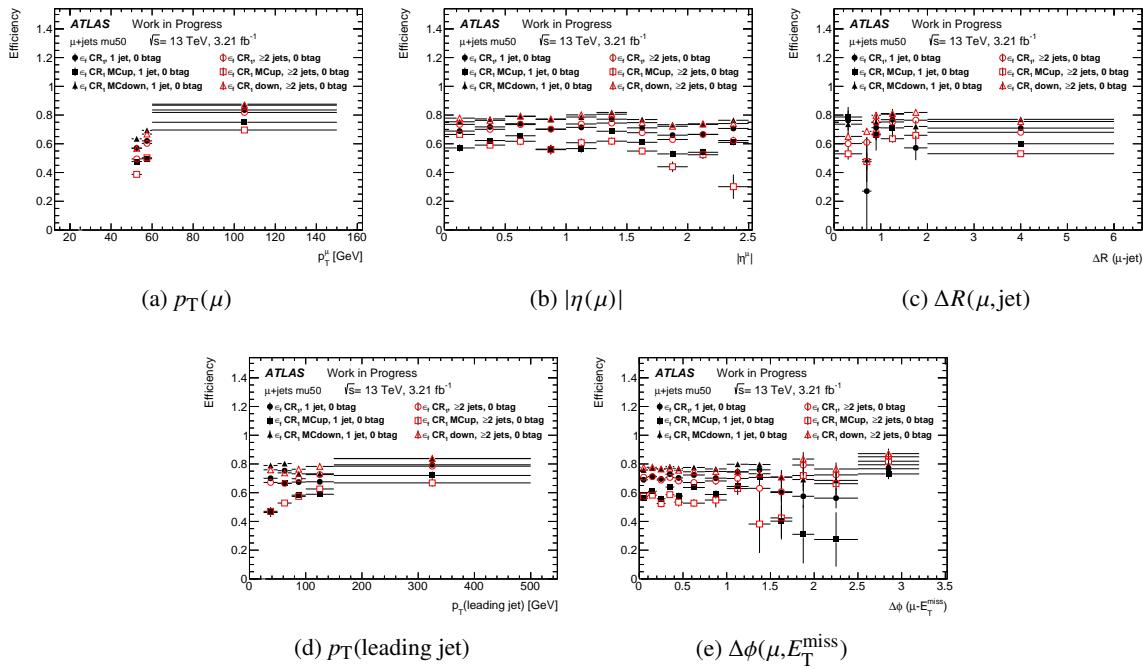


Figure 82: Muon fake efficiency $\varepsilon_{\text{fake}}$ as a function of (a) the muon p_T , (b) the muon $|\eta|$, (c) the distance ΔR between the muon and its nearest jet, (d) the p_T of the leading jet and (e) the distance in azimuth $\Delta\phi(\mu, E_T^{\text{miss}})$ between the muon and the E_T^{miss} . $\varepsilon_{\text{fake}}$ is shown for events with different multiplicities and 0 b -jet, for all muons which match the single muon trigger HLT_MU50 (mu50, see section 5). They are obtained for different levels of MC to be subtracted (nominal, 30% more and 30% less).

Not reviewed, for internal circulation only

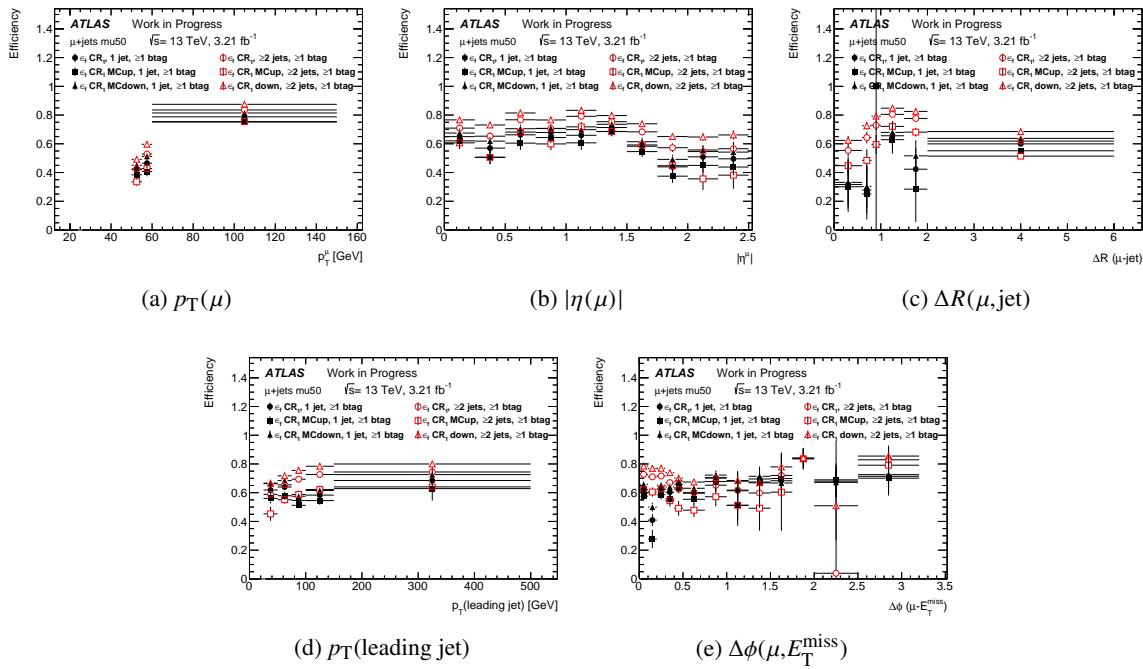


Figure 83: Muon fake efficiency $\varepsilon_{\text{fake}}$ as a function of (a) the muon p_T , (b) the muon $|\eta|$, (c) the distance ΔR between the muon and its nearest jet, (d) the p_T of the leading jet and (e) the distance in azimuth $\Delta\phi(\mu, E_T^{\text{miss}})$ between the muon and the E_T^{miss} . $\varepsilon_{\text{fake}}$ is shown for events with different multiplicities and at least one b -jet, for all muons which match the single muon trigger HLT_MU50 (mu50, see section 5). They are obtained for different levels of MC to be subtracted (nominal, 30% more and 30% less).

Not reviewed, for internal circulation only

1110 D.5. Impact of using different Z/W+jets simulated data

1111 To get a pure fake muon sample, the contamination from real muons has to be subtracted from the selected
 1112 sample. The biggest concerns are the W and Z contamination. Different samples of simulated data are
 1113 available for these two processes, SHERPA which are used by default but also based on Powheg +PYTHIA and
 1114 MADGRAPH +PYTHIA. Fig. ?? and ?? show the effect of using these different simulations. The efficiencies
 1115 obtained from the different simulated data are close even if the ones derived using Powheg are a bit higher.

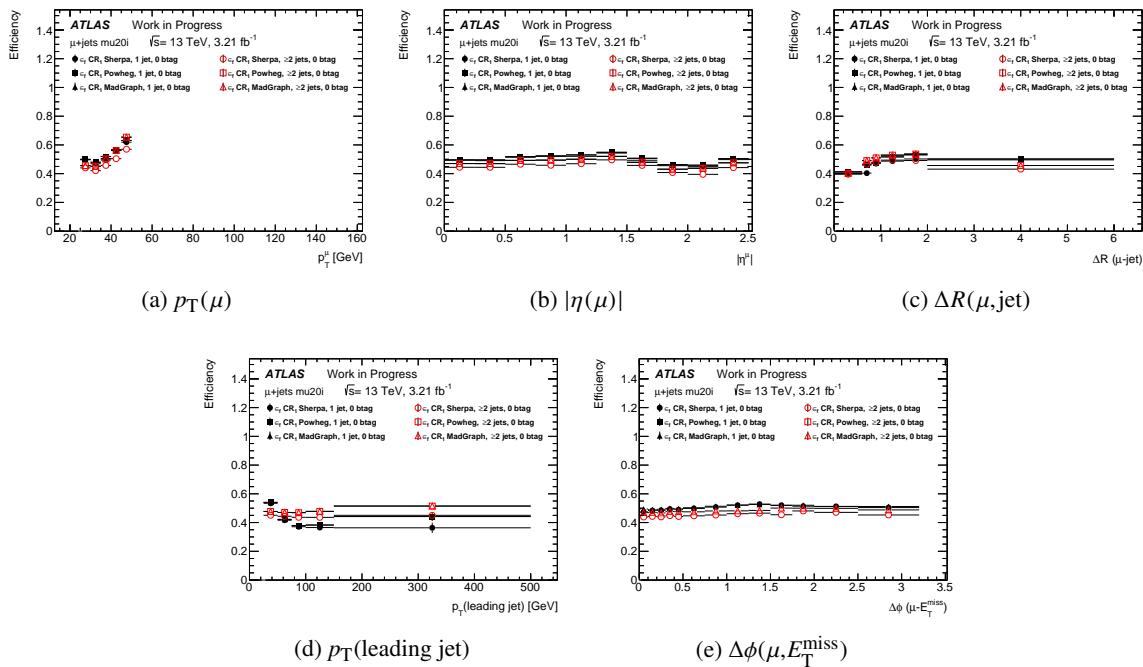


Figure 84: Muon fake efficiency $\varepsilon_{\text{fake}}$ as a function of (a) the muon p_T , (b) the muon $|\eta|$, (c) the distance ΔR between the muon and its nearest jet, (d) the p_T of the leading jet and (e) the distance in azimuth $\Delta\phi(\mu, E_T^{\text{miss}})$ between the muon and the E_T^{miss} . $\varepsilon_{\text{fake}}$ is shown for events with different multiplicities and 0 b -jet, for all muons which match the single muon trigger HLT_MU20_ILOOSE_L1MU15 (mu20i, see section 5). They are obtained using different $W/Z + \text{jets}$ simulated data to estimate the real lepton contamination.

Not reviewed, for internal circulation only

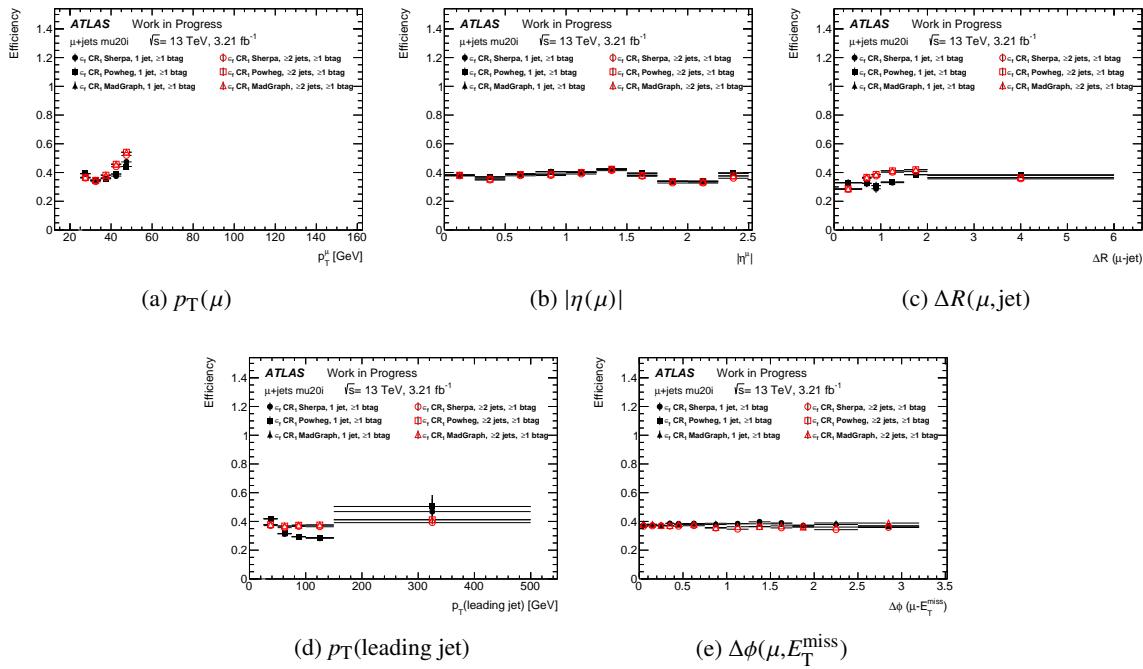


Figure 85: Muon fake efficiency $\varepsilon_{\text{fake}}$ as a function of (a) the muon p_T , (b) the muon $|\eta|$, (c) the distance ΔR between the muon and its nearest jet, (d) the p_T of the leading jet and (e) the distance in azimuth $\Delta\phi(\mu, E_T^{\text{miss}})$ between the muon and the E_T^{miss} . $\varepsilon_{\text{fake}}$ is shown for events with different multiplicities and at least one b -jet, for all muons which match the single muon trigger HLT_MU20_ILOOSE_L1MU15 (mu20i, see section 5). They are obtained using different $W/Z + \text{jets}$ simulated data to estimate the real lepton contamination.

Not reviewed, for internal circulation only

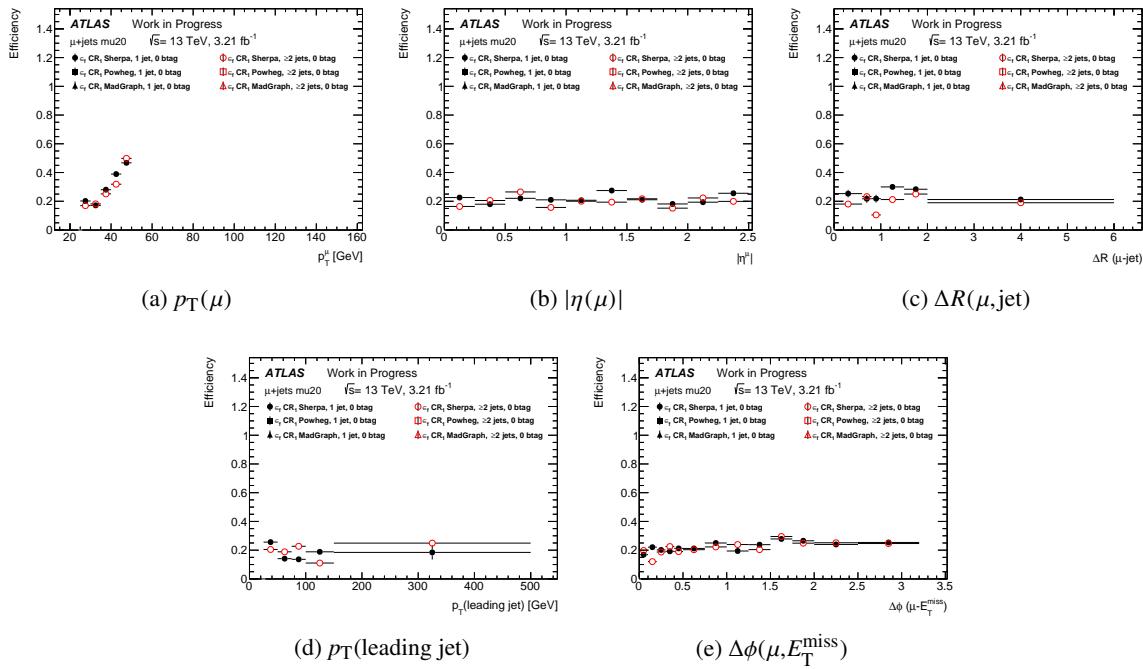


Figure 86: Muon fake efficiency $\varepsilon_{\text{fake}}$ as a function of (a) the muon p_T , (b) the muon $|\eta|$, (c) the distance ΔR between the muon and its nearest jet, (d) the p_T of the leading jet and (e) the distance in azimuth $\Delta\phi(\mu, E_T^{\text{miss}})$ between the muon and the E_T^{miss} . $\varepsilon_{\text{fake}}$ is shown for events with different multiplicities and 0 b -jet, for all muons which match the single muon trigger HLT_MU20_L1MU15 (mu20, see section 5). They are obtained using different $W/Z + \text{jets}$ simulated data to estimate the real lepton contamination.

Not reviewed, for internal circulation only

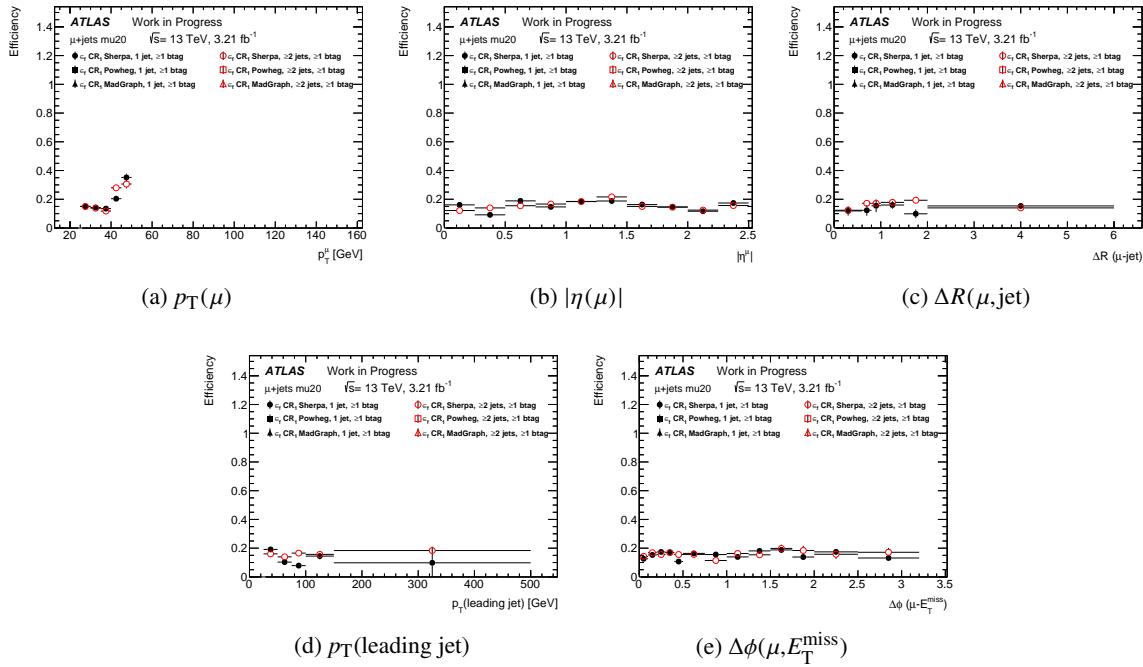


Figure 87: Muon fake efficiency $\varepsilon_{\text{fake}}$ as a function of (a) the muon p_T , (b) the muon $|\eta|$, (c) the distance ΔR between the muon and its nearest jet, (d) the p_T of the leading jet and (e) the distance in azimuth $\Delta\phi(\mu, E_T^{\text{miss}})$ between the muon and the E_T^{miss} . $\varepsilon_{\text{fake}}$ is shown for events with different multiplicities and at least one b -jet, for all muons which match the single muon trigger HLT_MU20_L1MU15 (mu20, see section 5). They are obtained using different $W/Z + \text{jets}$ simulated data to estimate the real lepton contamination.

Not reviewed, for internal circulation only

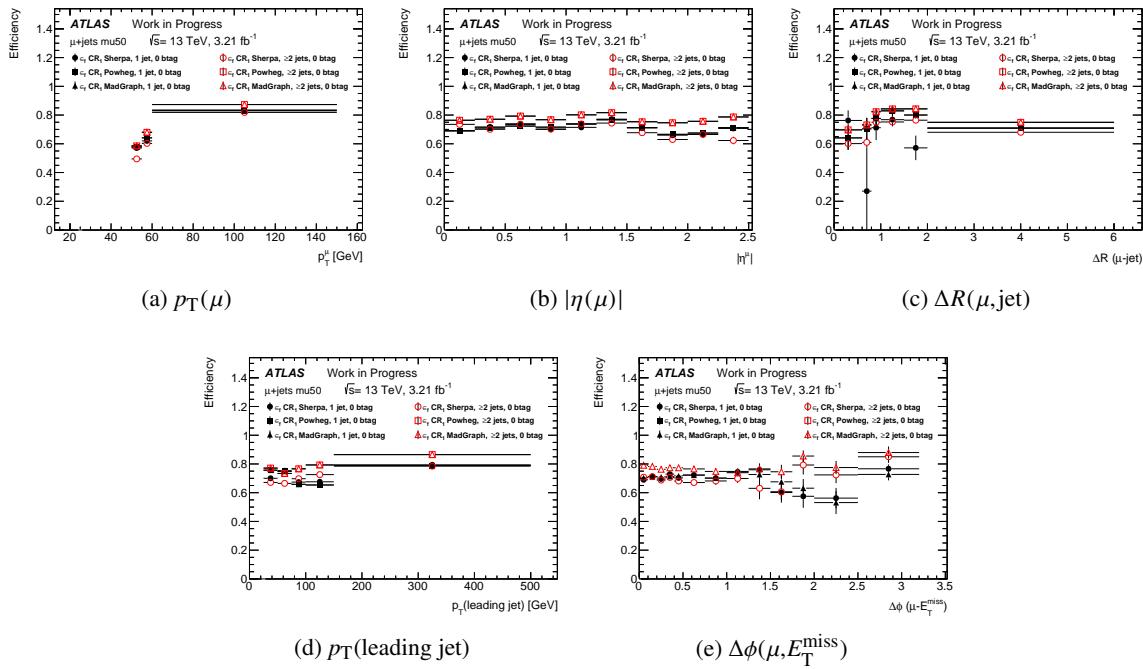


Figure 88: Muon fake efficiency $\varepsilon_{\text{fake}}$ as a function of (a) the muon p_T , (b) the muon $|\eta|$, (c) the distance ΔR between the muon and its nearest jet, (d) the p_T of the leading jet and (e) the distance in azimuth $\Delta\phi(\mu, E_T^{\text{miss}})$ between the muon and the E_T^{miss} . $\varepsilon_{\text{fake}}$ is shown for events with different multiplicities and 0 b -jet, for all muons which match the single muon trigger HLT_MU50 (mu50, see section 5). They are obtained using different $W/Z + \text{jets}$ simulated data to estimate the real lepton contamination.

Not reviewed, for internal circulation only

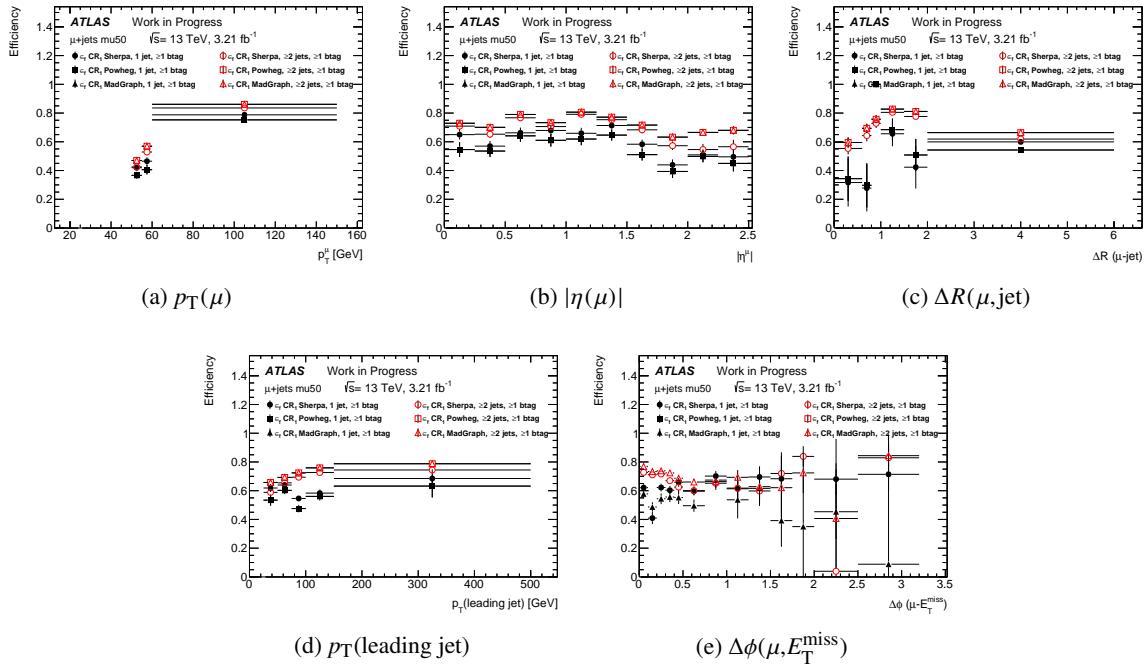


Figure 89: Muon fake efficiency $\varepsilon_{\text{fake}}$ as a function of (a) the muon p_T , (b) the muon $|\eta|$, (c) the distance ΔR between the muon and its nearest jet, (d) the p_T of the leading jet and (e) the distance in azimuth $\Delta\phi(\mu, E_T^{\text{miss}})$ between the muon and the E_T^{miss} . $\varepsilon_{\text{fake}}$ is shown for events with different multiplicities and at least one b -jet, for all muons which match the single muon trigger HLT_MU50 (mu50, see section 5). They are obtained using different $W/Z + \text{jets}$ simulated data to estimate the real lepton contamination.

¹¹¹⁷ **D.6. Binning used to measure fake muon efficiencies**

The binnings used for the parametrisation of the fake muon efficiencies are given in Tables 36 to 39.

Table 36: Measurement bins in muon transverse momentum p_T used for the muon fake efficiency parametrisation.

Bin boundaries in E_T (GeV)								
25	30	35	40	45	50	60	80	∞

Table 37: Measurement bins in muon pseudorapidity $|\eta|$ used for the muon fake efficiency parametrisation.

Bin boundaries in $ \eta $										
0	0.25	0.5	0.75	1.	1.25	1.5	1.75	2.	2.25	2.50

Table 38: Measurement bins in distance to the nearest jet $\Delta R(\mu, \text{jet})$ used for the muon fake efficiency parametrisation.

Bin boundaries in $\Delta R(\mu, \text{jet})$ to the nearest jet						
0.4	0.6	0.8	1	1.5	2.0	6

Table 39: Measurement bins in p_T of the leading jet used for the muon fake efficiency parametrisation.

Bin boundaries in p_T of the leading jet (GeV)					
25	50	75	100	150	∞

Table 40: Measurement bins in $|\Delta\phi|$ between the muon and of the E_T^{miss} used for the fake muon efficiency parametrisation.

Bin boundaries in $ \Delta\phi $ between the muon and of the E_T^{miss}													
0	0.1	0.2	0.3	0.4	0.5	0.75	1	1.25	1.5	1.75	2	2.5	π

1119 E. Additional studies for $l+jets t\bar{t}$ events

1120 E.1. Estimates for $e+jets$ events

1121 E.1.1. Only fake leptons

1122 Fig. 90 to 92 show the default non-prompt and fake leptons estimates for different distributions, for
 1123 different jet and b -jet multiplicities. They show also the different systematic variations for these estimates.

Not reviewed, for internal circulation only

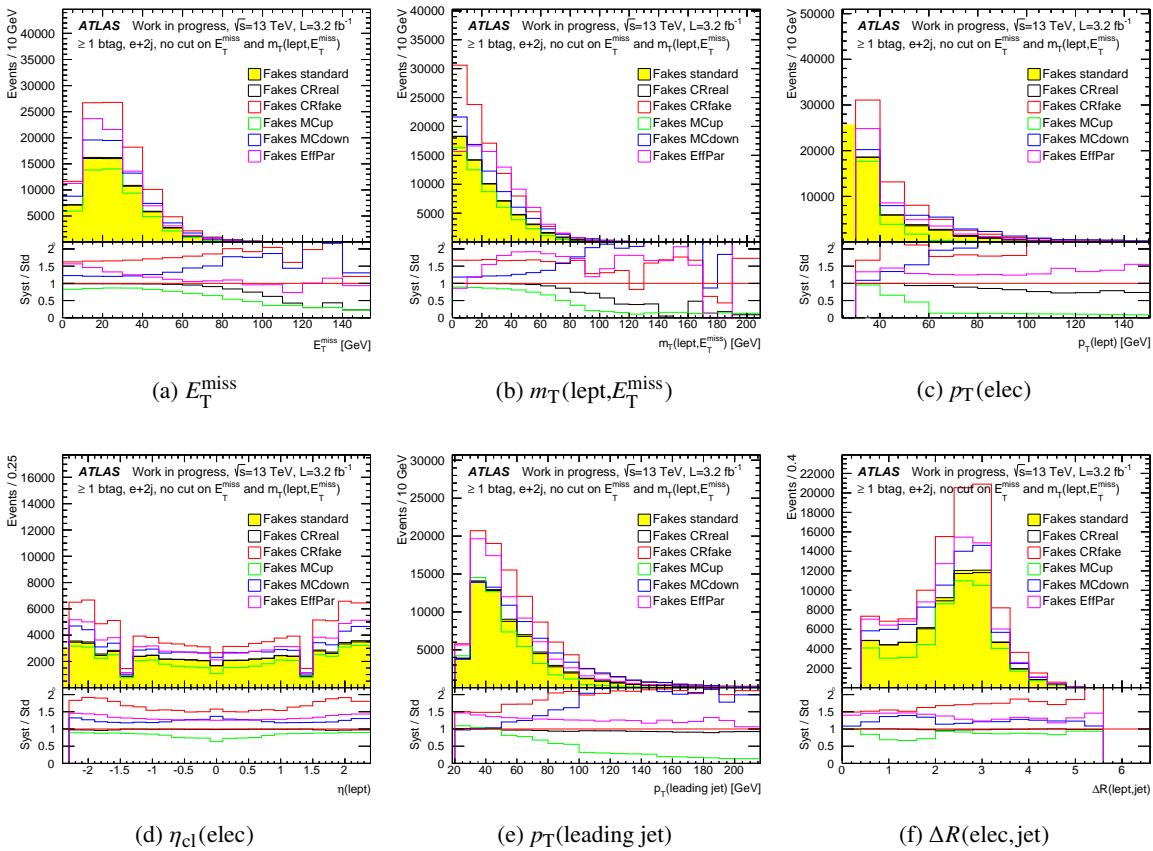


Figure 90: Distributions of (a) E_T^{miss} , (b) $m_T(\text{lept}, E_T^{\text{miss}})$, (c) p_T of the electron, (d) η of the electron, (e) p_T of the leading jet and (f) the distance ΔR of the electron to its nearest jet, for $e+jets$ events with two jets, with at least one b -jet, without any cuts on E_T^{miss} and $m_T(\text{lept}, E_T^{\text{miss}})$. Only the non-prompt and fake lepton backgrounds estimated with the matrix method, obtained with the different variations of both $\varepsilon_{\text{real}}$ and $\varepsilon_{\text{fake}}$, are shown.

Not reviewed, for internal circulation only

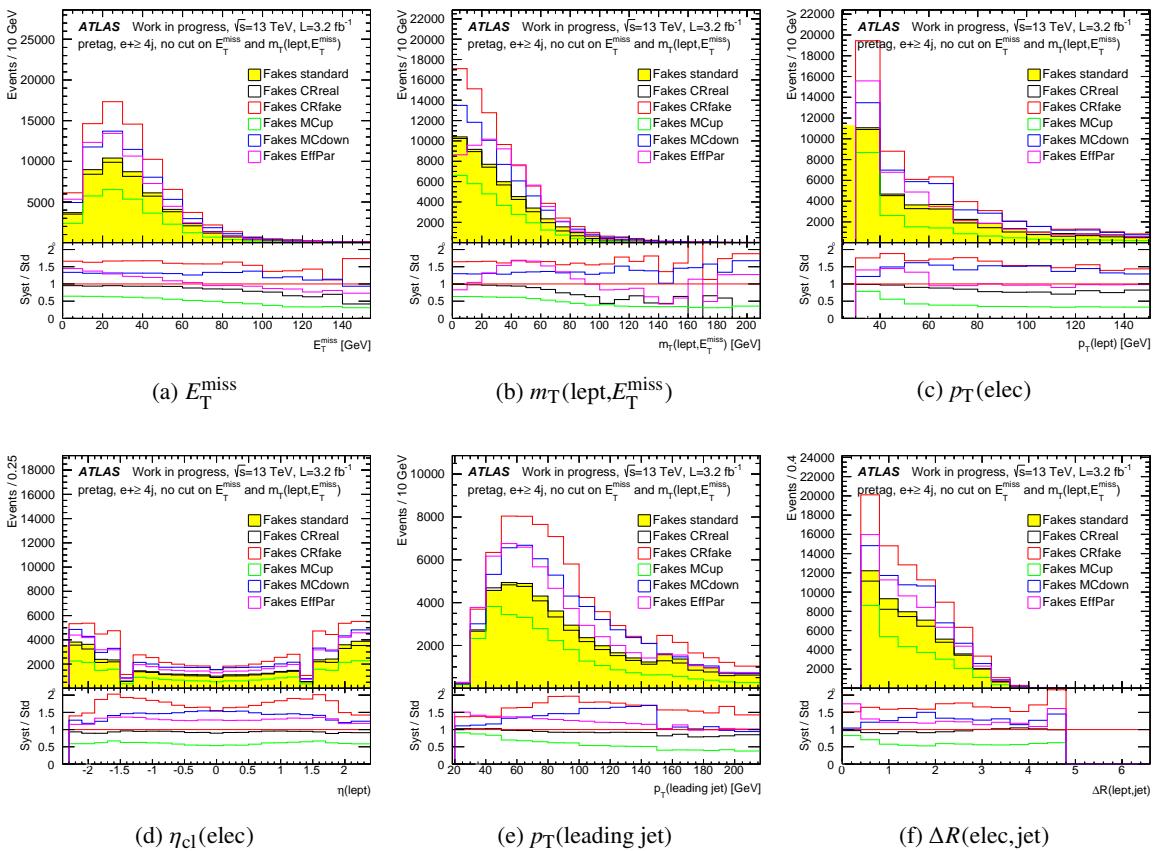


Figure 91: Distributions of (a) E_T^{miss} , (b) $m_T(\text{lept}, E_T^{\text{miss}})$, (c) p_T of the electron, (d) η of the electron, (e) p_T of the leading jet and (f) the distance ΔR of the electron to its nearest jet, for $e+jets$ events with at least four jets, with at least one b -jet, without any cuts on E_T^{miss} and $m_T(\text{lept}, E_T^{\text{miss}})$. Only the non-prompt and fake lepton backgrounds estimated with the matrix method, obtained with the different variations of both $\varepsilon_{\text{real}}$ and $\varepsilon_{\text{fake}}$, are shown.

Not reviewed, for internal circulation only

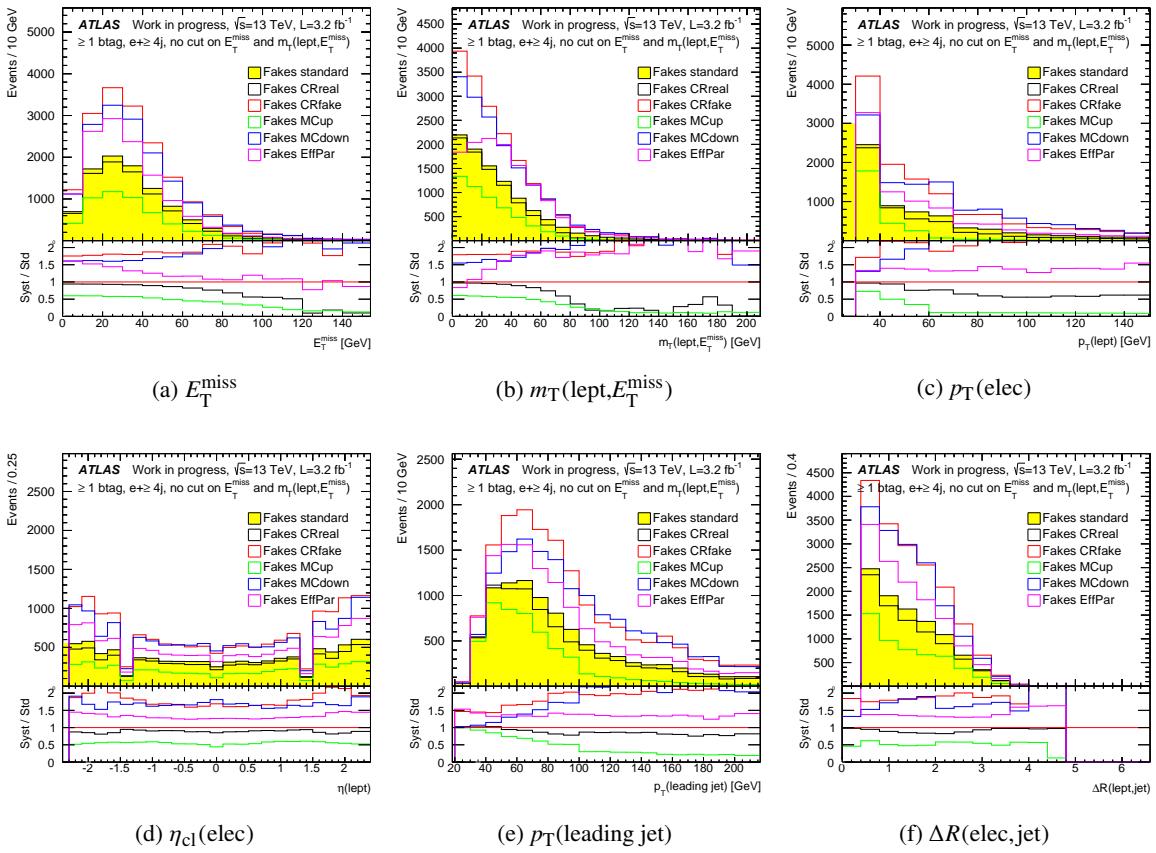


Figure 92: Distributions of (a) E_T^{miss} , (b) $m_T(\text{lept}, E_T^{\text{miss}})$, (c) p_T of the electron, (d) η of the electron, (e) p_T of the leading jet and (f) the distance ΔR of the electron to its nearest jet, for $e+jets$ events with at least four jets, with at least one b -jet, without any cuts on E_T^{miss} and $m_T(\text{lept}, E_T^{\text{miss}})$. Only the non-prompt and fake lepton backgrounds estimated with the matrix method, obtained with the different variations of both $\varepsilon_{\text{real}}$ and $\varepsilon_{\text{fake}}$, are shown.

1125 **E.1.2. Details on the systematic effects**

1126 Fig. 93 to 95 show the data compared to the real lepton expectation from simulation, showing the contri-
 1127 butions from $t\bar{t}$, single top, W +jets, Z +jets and dibosons normalised to their cross-sections (Background),
 1128 and non-prompt and fake lepton backgrounds (referred to as ‘NP & Fake Lep.’) estimated with the matrix
 1129 method for different distributions, for different jet and b -jet multiplicities. They show also the different
 systematic variations for the non-prompt and fake leptons estimates.

Not reviewed, for internal circulation only

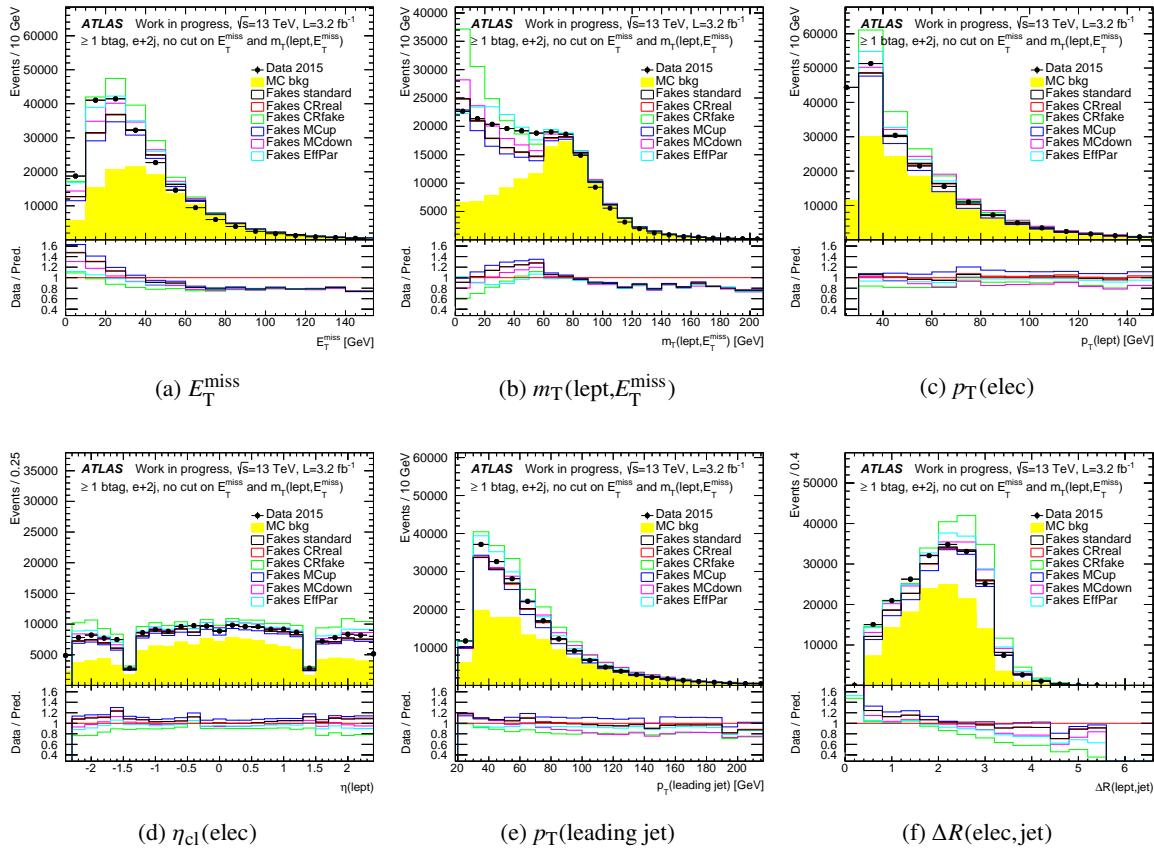


Figure 93: Distributions of (a) E_T^{miss} , (b) $m_T(\text{lept}, E_T^{\text{miss}})$, (c) p_T of the electron, (d) η of the electron, (e) p_T of the leading jet and (f) the distance ΔR of the electron to its nearest jet, for $e+jets$ events with two jets, with at least one b -jet, without any cuts on E_T^{miss} and $m_T(\text{lept}, E_T^{\text{miss}})$. The data is compared to the real lepton expectation from simulation, showing the contributions from $t\bar{t}$, single top, W +jets, Z +jets and dibosons normalised to their cross-sections (Background), and non-prompt and fake lepton backgrounds (referred to as ‘NP & Fake Lep.’) estimated with the matrix method obtained with the different variations of both $\varepsilon_{\text{real}}$ and $\varepsilon_{\text{fake}}$.

1130

Not reviewed, for internal circulation only

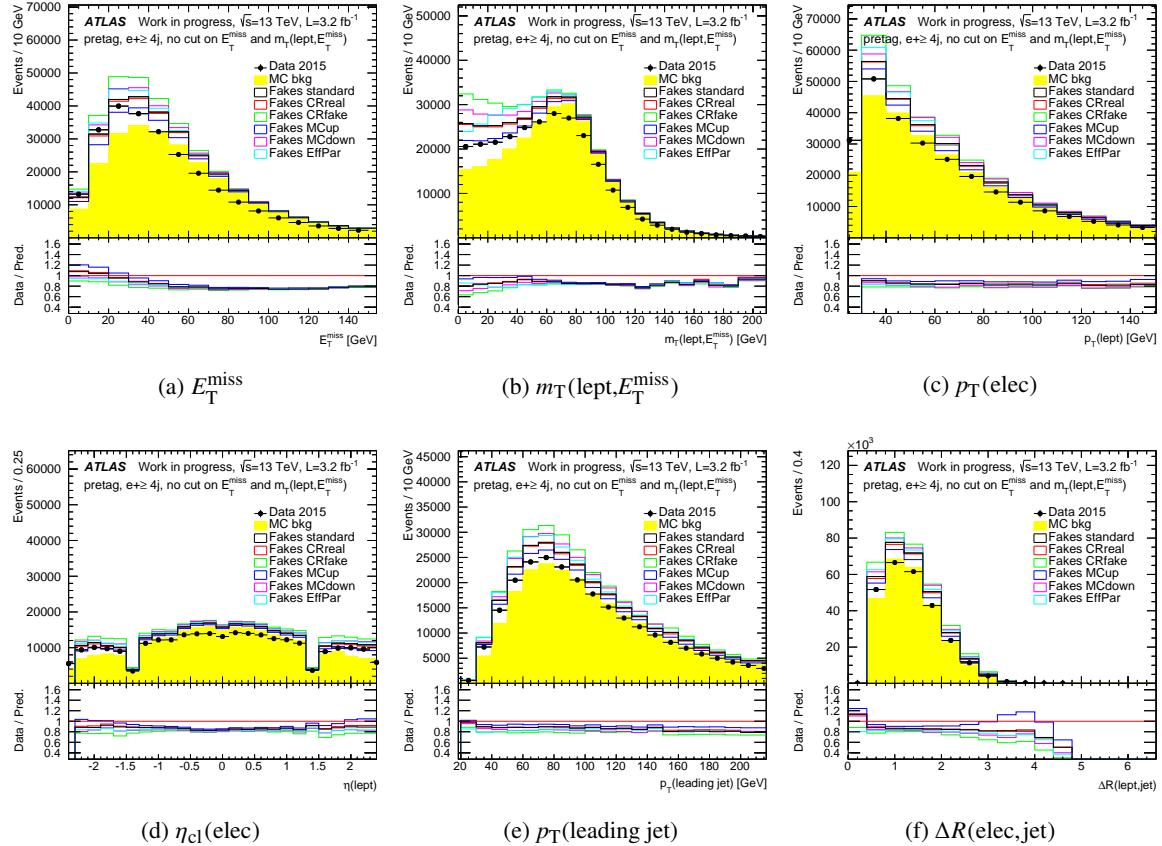


Figure 94: Distributions of (a) E_T^{miss} , (b) $m_T(\text{lept}, E_T^{\text{miss}})$, (c) p_T of the electron, (d) η of the electron, (e) p_T of the leading jet and (f) the distance ΔR of the electron to its nearest jet, for $e+jets$ events with at least four jets, without any b -tagging requirement, without any cuts on E_T^{miss} and $m_T(\text{lept}, E_T^{\text{miss}})$. The data is compared to the real lepton expectation from simulation, showing the contributions from $t\bar{t}$, single top, $W + \text{jets}$, $Z + \text{jets}$ and dibosons normalised to their cross-sections (Background), and non-prompt and fake lepton backgrounds (referred to as ‘NP & Fake Lep.’) estimated with the matrix method obtained with the different variations of both $\varepsilon_{\text{real}}$ and $\varepsilon_{\text{fake}}$.

Not reviewed, for internal circulation only

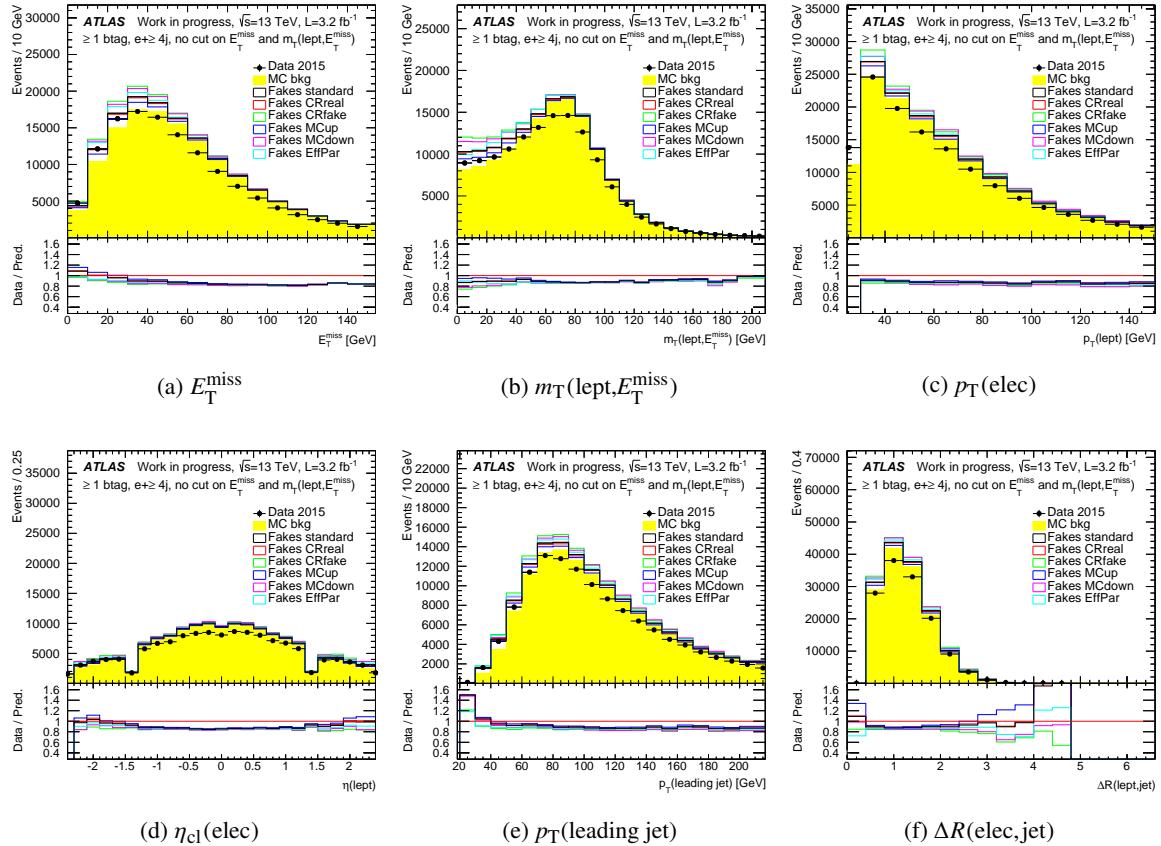


Figure 95: Distributions of (a) E_T^{miss} , (b) $m_T(\text{lept}, E_T^{\text{miss}})$, (c) p_T of the electron, (d) η of the electron, (e) p_T of the leading jet and (f) the distance ΔR of the electron to its nearest jet, for $e+jets$ events with at least four jets, without at least one b -jet, without any cuts on E_T^{miss} and $m_T(\text{lept}, E_T^{\text{miss}})$. The data is compared to the real lepton expectation from simulation, showing the contributions from $t\bar{t}$, single top, $W+jets$, $Z+jets$ and dibosons normalised to their cross-sections (Background), and non-prompt and fake lepton backgrounds (referred to as ‘NP & Fake Lep.’) estimated with the matrix method obtained with the different variations of both $\varepsilon_{\text{real}}$ and $\varepsilon_{\text{fake}}$.

1131 **E.1.3. Estimates without cuts on E_T^{miss} and $m_T(\text{lept}, E_T^{\text{miss}})$**

1132 Fig. 96 shows results without any cuts on E_T^{miss} and $m_T(\text{lept}, E_T^{\text{miss}})$ with two jets and at least two b -jets.

Not reviewed, for internal circulation only

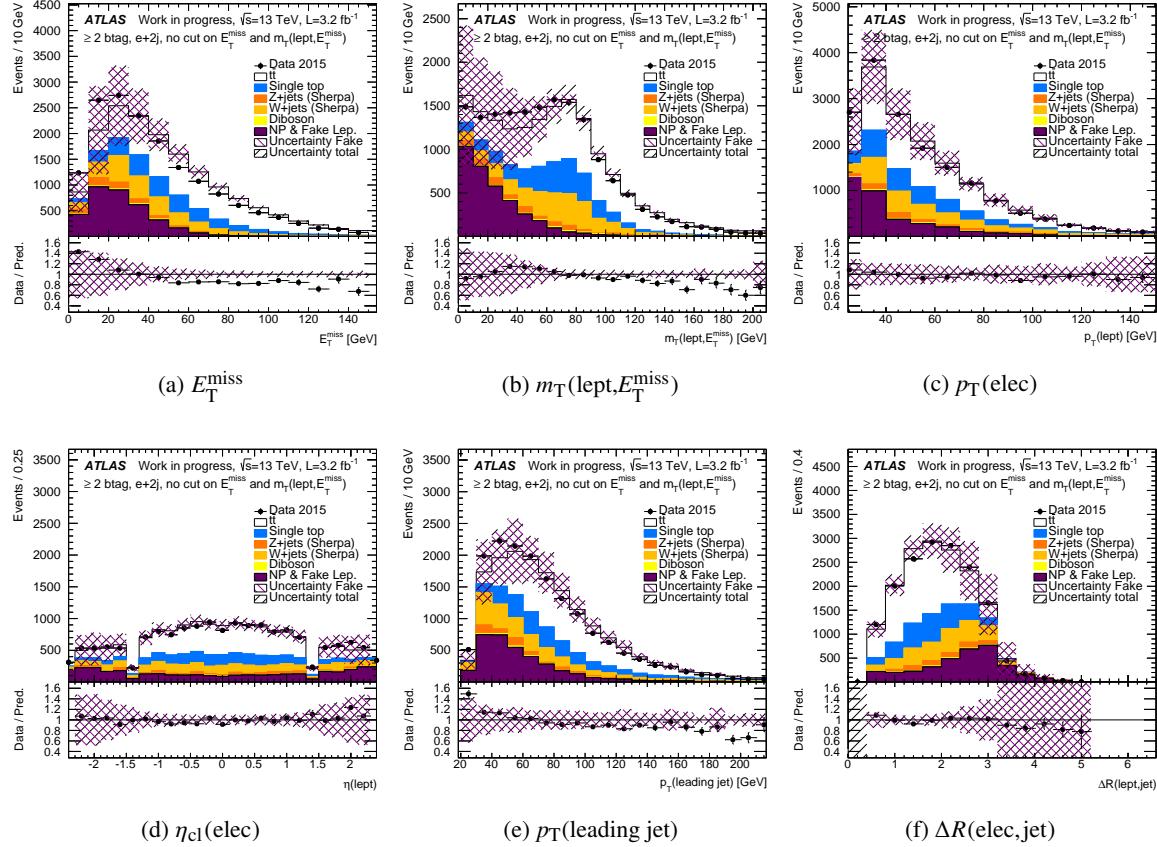


Figure 96: Distributions of (a) E_T^{miss} , (b) $m_T(\text{lept}, E_T^{\text{miss}})$, (c) p_T of the electron, (d) η_{cl} of the electron, (e) p_T of the leading jet and (f) the distance ΔR of the electron to its nearest jet, for $e+jets$ events with two jets, requiring at least two b -jets, without any cuts on E_T^{miss} and $m_T(\text{lept}, E_T^{\text{miss}})$. The data is compared to the real lepton expectation from simulation, showing separately the contributions from $t\bar{t}$, single top, $W+jets$, $Z+jets$ and dibosons normalised to their cross-sections, and non-prompt and fake lepton backgrounds (referred to as ‘NP & Fake Lep.’) estimated with the matrix method. The purple shaded area represents the combination of the statistical and the systematic uncertainties on the matrix method estimate in each bin. The black shaded area represents the total uncertainty including also the uncertainties on the processes predicted by the MC simulation and the one on the luminosity. The lower parts of the figure show the ratios of expectation to data.

1133

1134 **E.1.4. Estimates in the signal region**

1135 Fig. 97 shows results in the signal region with at least four jets and at least two b -jets.

1136 Fig. 98 to 100 shows results in the signal region with at least six jets and different b -jet multiplicities.

Not reviewed, for internal circulation only

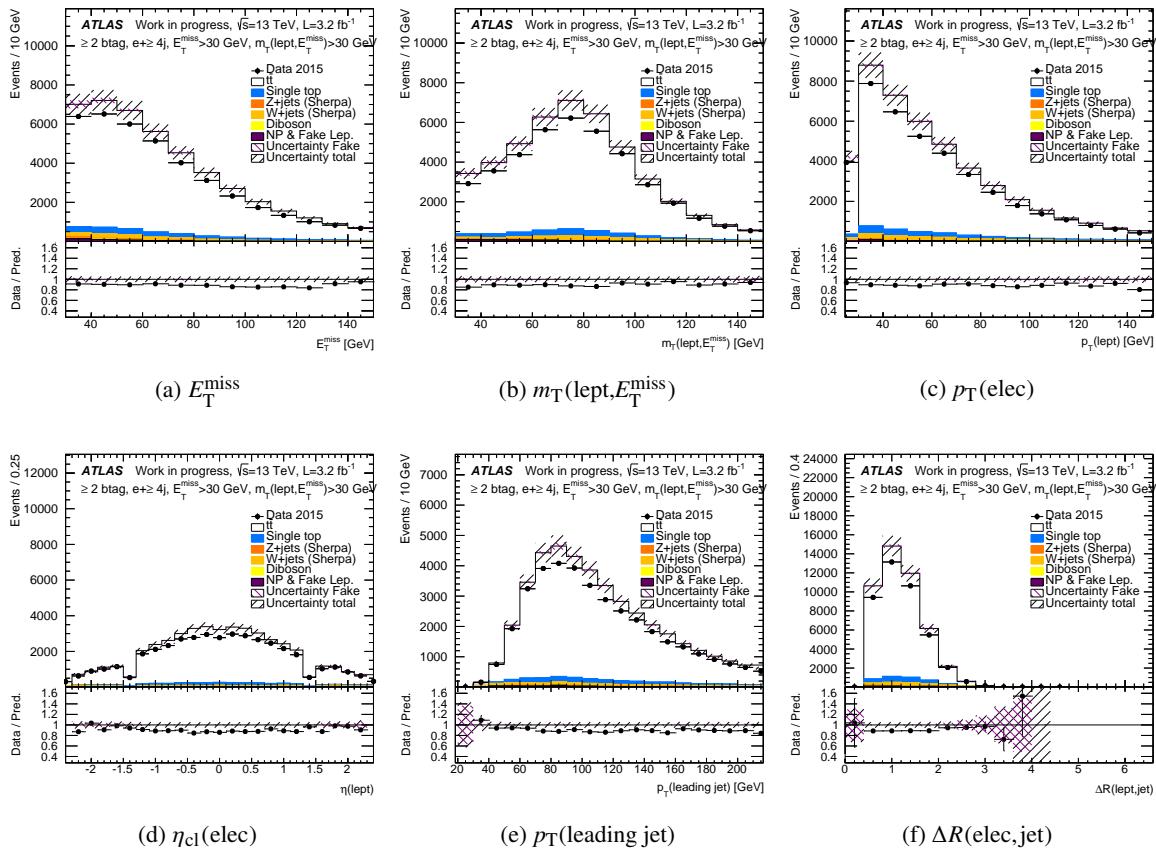


Figure 97: Distributions of (a) E_T^{miss} , (b) $m_T(\text{lept}, E_T^{\text{miss}})$, (c) p_T of the electron, (d) η_{cl} of the electron, (e) p_T of the leading jet and (f) the distance ΔR of the electron to its nearest jet for $e+jets$ events with at least four jets, requiring at least two b -jets, requiring $E_T^{\text{miss}} > 30 \text{ GeV}$ and $m_T(\text{lept}, E_T^{\text{miss}}) > 30 \text{ GeV}$. The data is compared to the real lepton expectation from simulation, showing separately the contributions from $t\bar{t}$, single top, $W+jets$, $Z+jets$ and dibosons normalised to their cross-sections, and non-prompt and fake lepton backgrounds (referred to as ‘NP & Fake Lep.’) estimated with the matrix method. The purple shaded area represents the combination of the statistical and the systematic uncertainties on the matrix method estimate in each bin. The black shaded area represents the total uncertainty including also the uncertainties on the processes predicted by the MC simulation and the one on the luminosity. The lower parts of the figure show the ratios of expectation to data.

Not reviewed, for internal circulation only

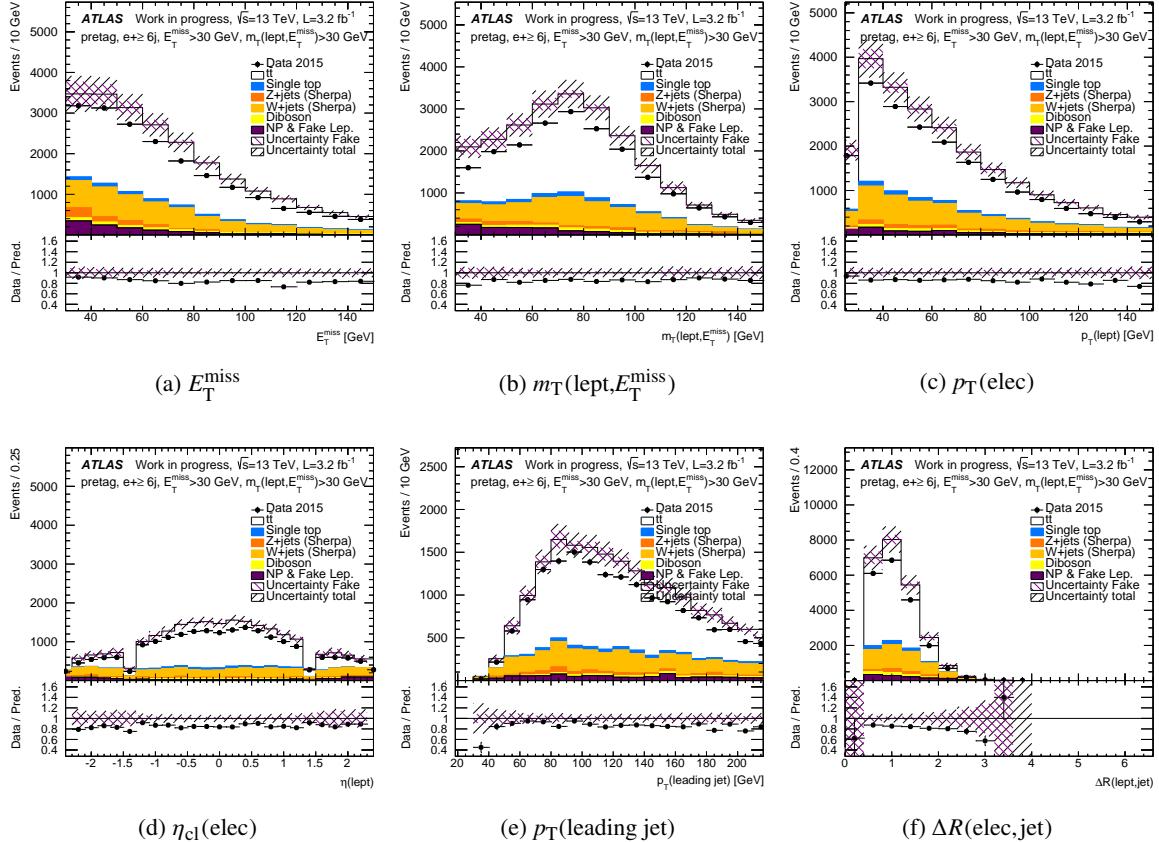


Figure 98: Distributions of (a) E_T^{miss} , (b) $m_T(\text{lept}, E_T^{\text{miss}})$, (c) p_T of the electron, (d) η_{cl} of the electron, (e) p_T of the leading jet and (f) the distance ΔR of the electron to its nearest jet, for $e+\text{jets}$ events with at least six jets without any b -tagging requirement, requiring $E_T^{\text{miss}} > 30 \text{ GeV}$ and $m_T(\text{lept}, E_T^{\text{miss}}) > 30 \text{ GeV}$. The data is compared to the real lepton expectation from simulation, showing separately the contributions from $t\bar{t}$, single top, $W+\text{jets}$, $Z+\text{jets}$ and dibosons normalised to their cross-sections, and non-prompt and fake lepton backgrounds (referred to as ‘NP & Fake Lep.’) estimated with the matrix method. The purple shaded area represents the combination of the statistical and the systematic uncertainties on the matrix method estimate in each bin. The black shaded area represents the total uncertainty including also the uncertainties on the processes predicted by the MC simulation and the one on the luminosity. The lower parts of the figure show the ratios of expectation to data.

Not reviewed, for internal circulation only

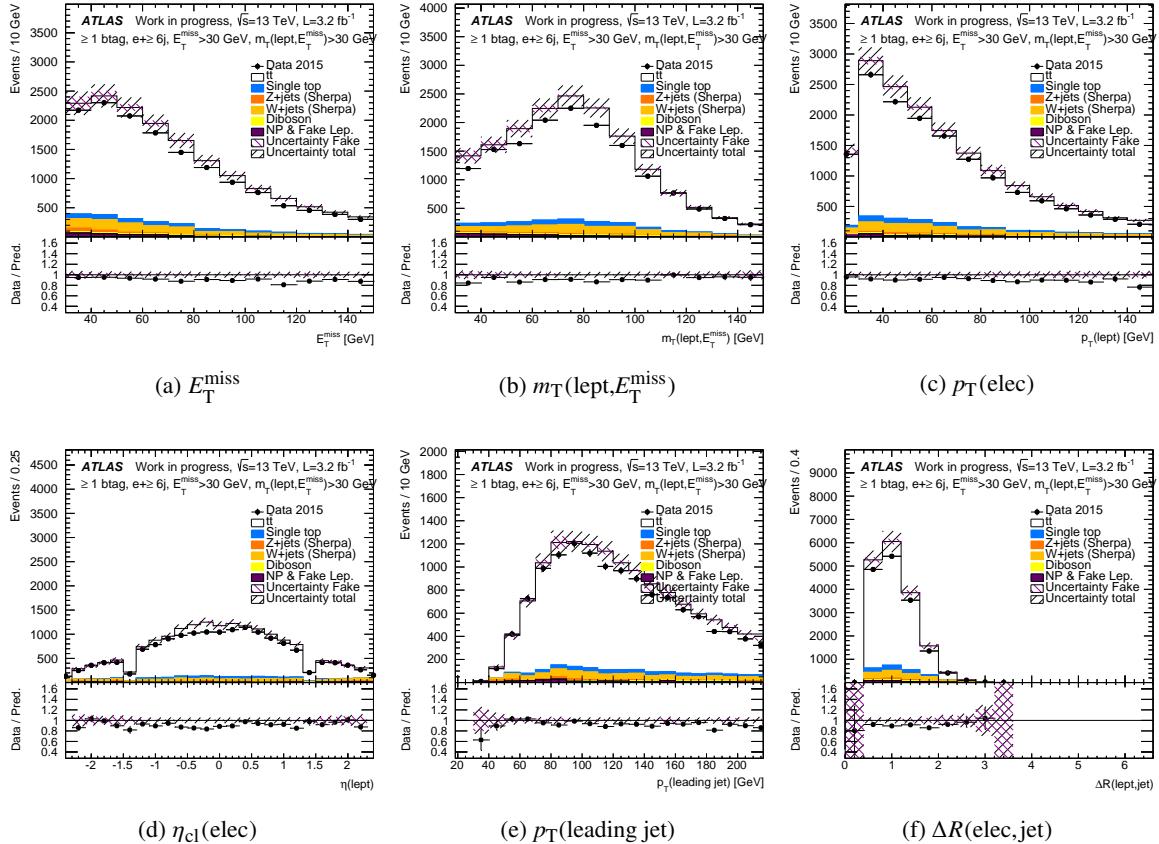


Figure 99: Distributions of (a) E_T^{miss} , (b) $m_T(\text{lept}, E_T^{\text{miss}})$, (c) p_T of the electron, (d) η_{cl} of the electron, (e) p_T of the leading jet and (f) the distance ΔR of the electron to its nearest jet, for $e+\text{jets}$ events with at least six jets, with at least one b -jet, requiring $E_T^{\text{miss}} > 30 \text{ GeV}$ and $m_T(\text{lept}, E_T^{\text{miss}}) > 30 \text{ GeV}$. The data is compared to the real lepton expectation from simulation, showing separately the contributions from $t\bar{t}$, single top, $W+\text{jets}$, $Z+\text{jets}$ and dibosons normalised to their cross-sections, and non-prompt and fake lepton backgrounds (referred to as ‘NP & Fake Lep.’) estimated with the matrix method. The purple shaded area represents the combination of the statistical and the systematic uncertainties on the matrix method estimate in each bin. The black shaded area represents the total uncertainty including also the uncertainties on the processes predicted by the MC simulation and the one on the luminosity. The lower parts of the figure show the ratios of expectation to data.

Not reviewed, for internal circulation only

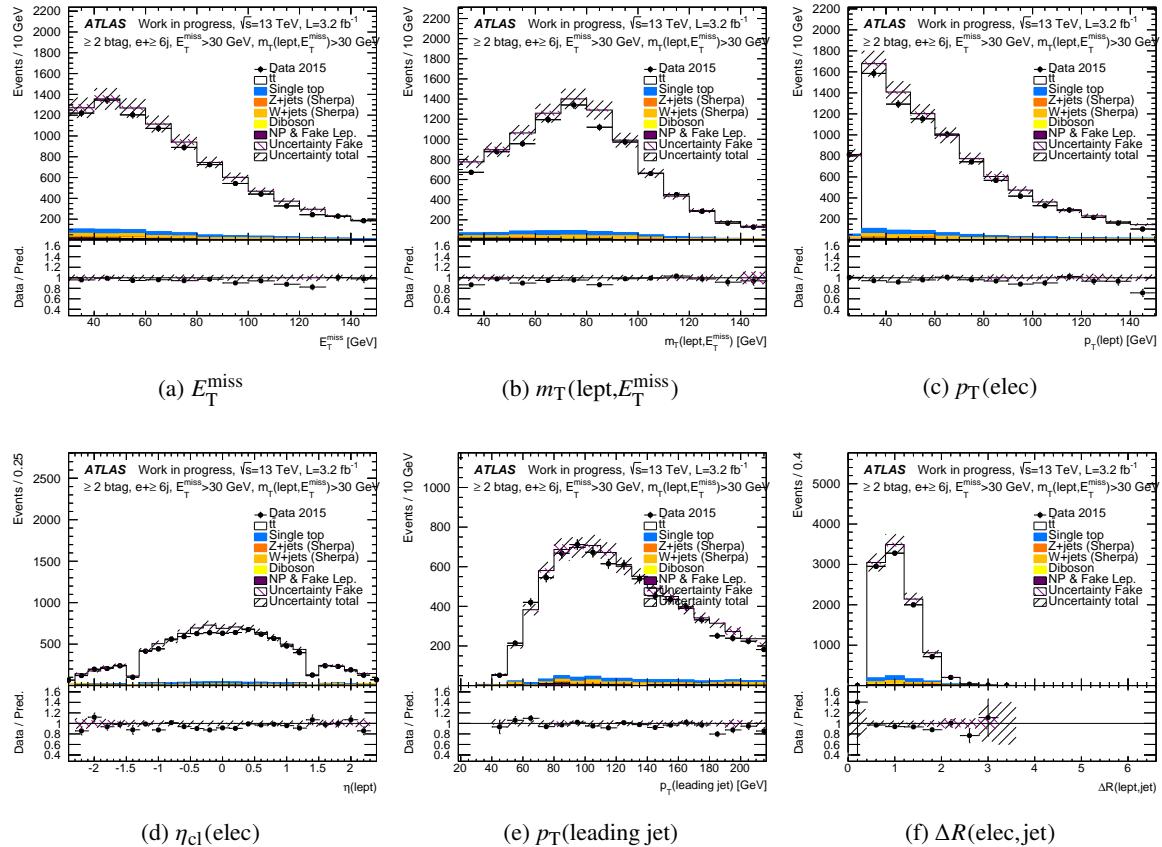


Figure 100: Distributions of (a) E_T^{miss} , (b) $m_T(\text{lept}, E_T^{\text{miss}})$, (c) p_T of the electron, (d) η_{cl} of the electron, (e) p_T of the leading jet and (f) the distance ΔR of the electron to its nearest jet, for $e+\text{jets}$ events with at least six jets, without at least two b -jets, requiring $E_T^{\text{miss}} > 30 \text{ GeV}$ and $m_T(\text{lept}, E_T^{\text{miss}}) > 30 \text{ GeV}$. The data is compared to the real lepton expectation from simulation, showing separately the contributions from $t\bar{t}$, single top, $W+\text{jets}$, $Z+\text{jets}$ and dibosons normalised to their cross-sections, and non-prompt and fake lepton backgrounds (referred to as ‘NP & Fake Lep.’) estimated with the matrix method. The purple shaded area represents the combination of the statistical and the systematic uncertainties on the matrix method estimate in each bin. The black shaded area represents the total uncertainty including also the uncertainties on the processes predicted by the MC simulation and the one on the luminosity. The lower parts of the figure show the ratios of expectation to data.

1137 **E.1.5. Estimates in the control regions**

1138 Fig. 101 to 103 show control plots for various distributions and different jet and b -tag multiplicities after applying the cuts $E_T^{\text{miss}} + m_T(\text{lept}, E_T^{\text{miss}}) < 60 \text{ GeV}$.

Not reviewed, for internal circulation only

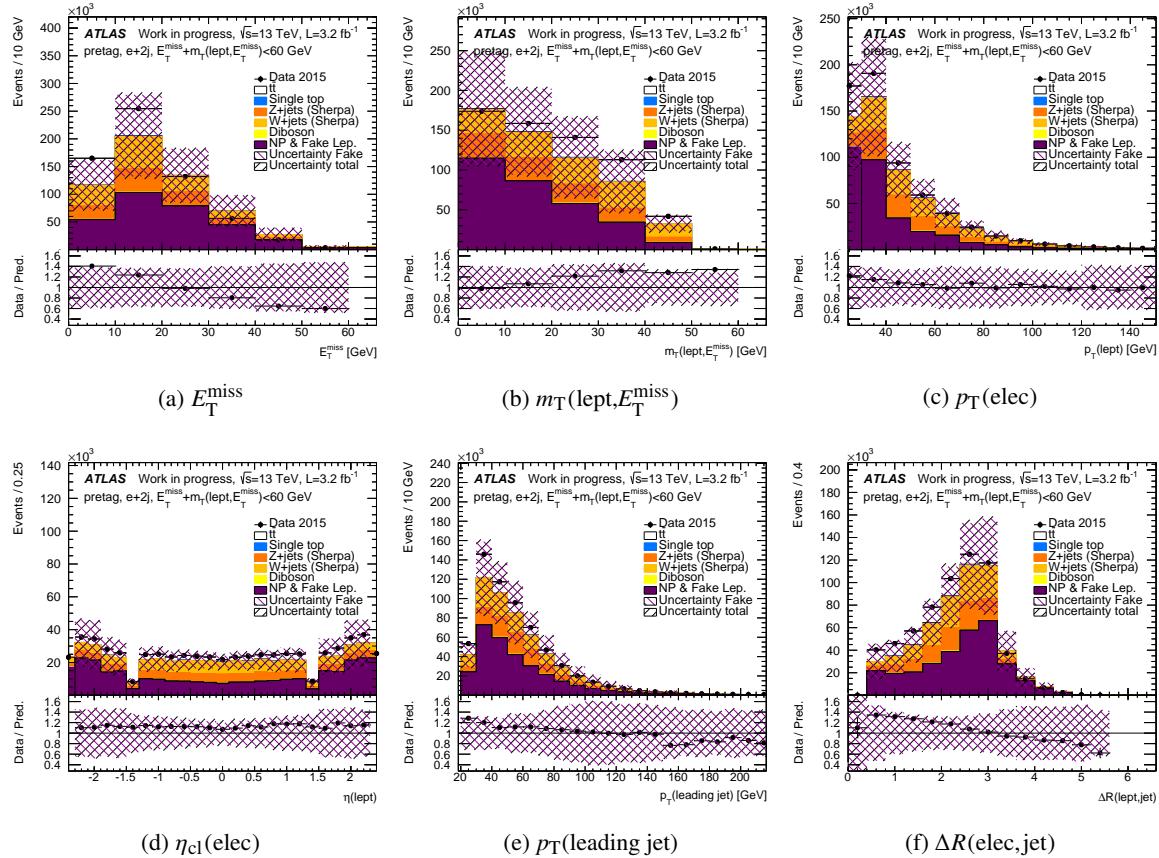


Figure 101: Distributions of (a) E_T^{miss} , (b) $m_T(\text{lept}, E_T^{\text{miss}})$, (c) p_T of the electron, (d) η_{cl} of the electron, (e) p_T of the leading jet and (f) the distance ΔR of the electron to its nearest jet, for $e+jets$ events with two jets, without any b -tagging requirement, requiring $E_T^{\text{miss}} + m_T(\text{lept}, E_T^{\text{miss}}) < 60 \text{ GeV}$. The data is compared to the real lepton expectation from simulation, showing separately the contributions from $t\bar{t}$, single top, $W+jets$, $Z+jets$ and dibosons normalised to their cross-sections, and non-prompt and fake lepton backgrounds (referred to as ‘NP & Fake Lep.’) estimated with the matrix method. The purple shaded area represents the combination of the statistical and the systematic uncertainties on the matrix method estimate in each bin. The black shaded area represents the total uncertainty including also the uncertainties on the processes predicted by the MC simulation and the one on the luminosity. The lower parts of the figure show the ratios of expectation to data.

1139

Not reviewed, for internal circulation only

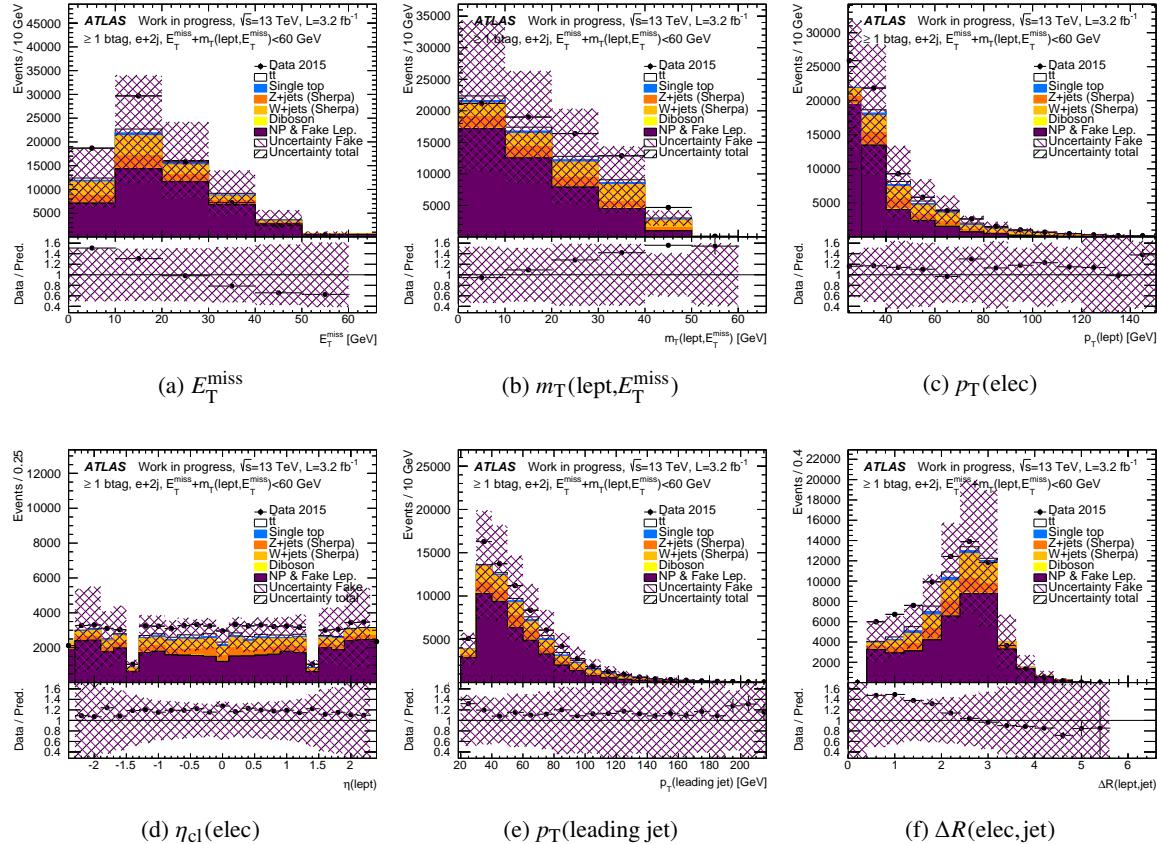


Figure 102: Distributions of (a) E_T^{miss} , (b) $m_T(\text{lept}, E_T^{\text{miss}})$, (c) p_T of the electron, (d) η_{cl} of the electron, (e) p_T of the leading jet and (f) the distance ΔR of the electron to its nearest jet, for $e+\text{jets}$ events with two jets, requiring at least one b -jet, requiring $E_T^{\text{miss}} + m_T(\text{lept}, E_T^{\text{miss}}) < 60 \text{ GeV}$. The data is compared to the real lepton expectation from simulation, showing separately the contributions from $t\bar{t}$, single top, $W+\text{jets}$, $Z+\text{jets}$ and dibosons normalised to their cross-sections, and non-prompt and fake lepton backgrounds (referred to as ‘NP & Fake Lep.’) estimated with the matrix method. The purple shaded area represents the combination of the statistical and the systematic uncertainties on the matrix method estimate in each bin. The black shaded area represents the total uncertainty including also the uncertainties on the processes predicted by the MC simulation and the one on the luminosity. The lower parts of the figure show the ratios of expectation to data.

Not reviewed, for internal circulation only

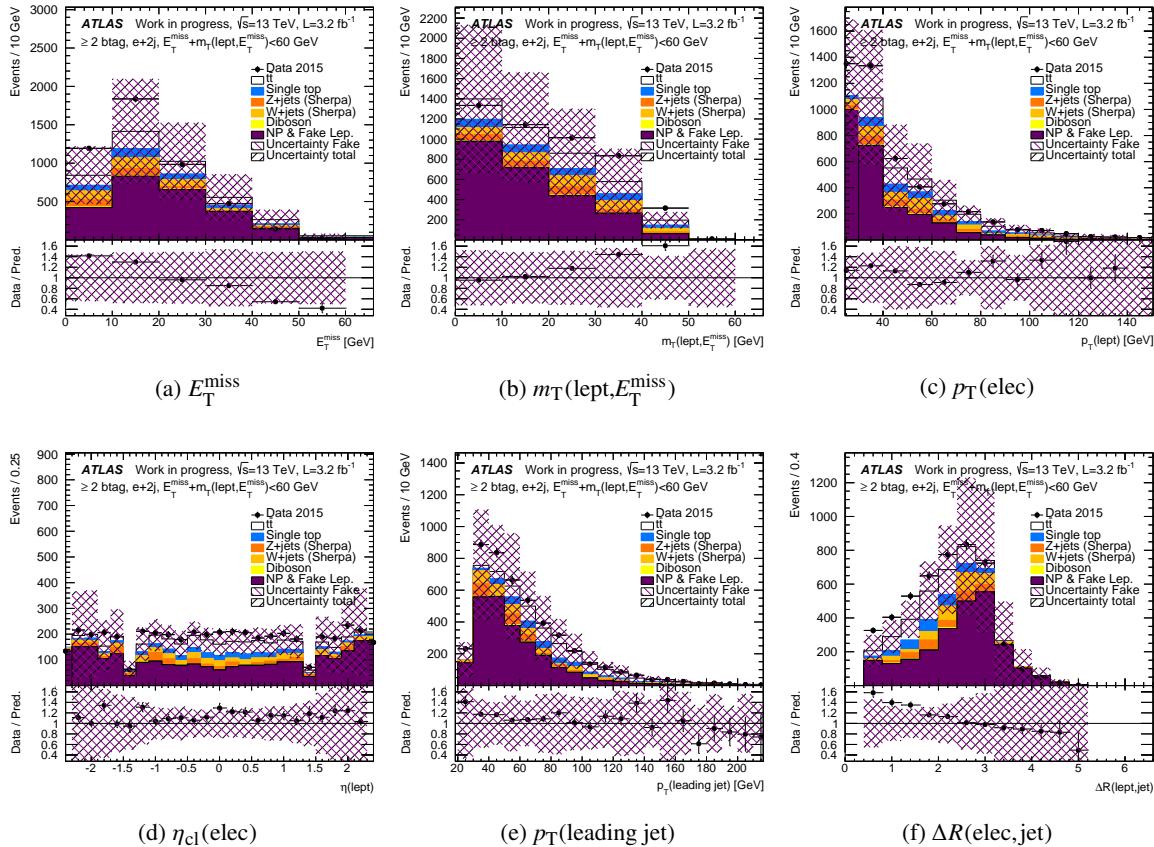


Figure 103: Distributions of (a) E_T^{miss} , (b) $m_T(\text{lept}, E_T^{\text{miss}})$, (c) p_T of the electron, (d) η_{cl} of the electron, (e) p_T of the leading jet and (f) the distance ΔR of the electron to its nearest jet, for $e+\text{jets}$ events with two jets, requiring at least two b -jets, requiring $E_T^{\text{miss}} + m_T(\text{lept}, E_T^{\text{miss}}) < 60 \text{ GeV}$. The data is compared to the real lepton expectation from simulation, showing separately the contributions from $t\bar{t}$, single top, $W+\text{jets}$, $Z+\text{jets}$ and dibosons normalised to their cross-sections, and non-prompt and fake lepton backgrounds (referred to as ‘NP & Fake Lep.’) estimated with the matrix method. The purple shaded area represents the combination of the statistical and the systematic uncertainties on the matrix method estimate in each bin. The black shaded area represents the total uncertainty including also the uncertainties on the processes predicted by the MC simulation and the one on the luminosity. The lower parts of the figure show the ratios of expectation to data.

1140 E.1.6. Estimates with other MC samples

1141 Fig. 104 and 105 show E_T^{miss} and $m_T(\text{lept}, E_T^{\text{miss}})$ distributions using three different MC samples to estimate the $W + \text{jets}$ and $Z + \text{jets}$ background, based on SHERPA, MADGRAPH +PYTHIA 8 and Powheg +PYTHIA 8.

Not reviewed, for internal circulation only

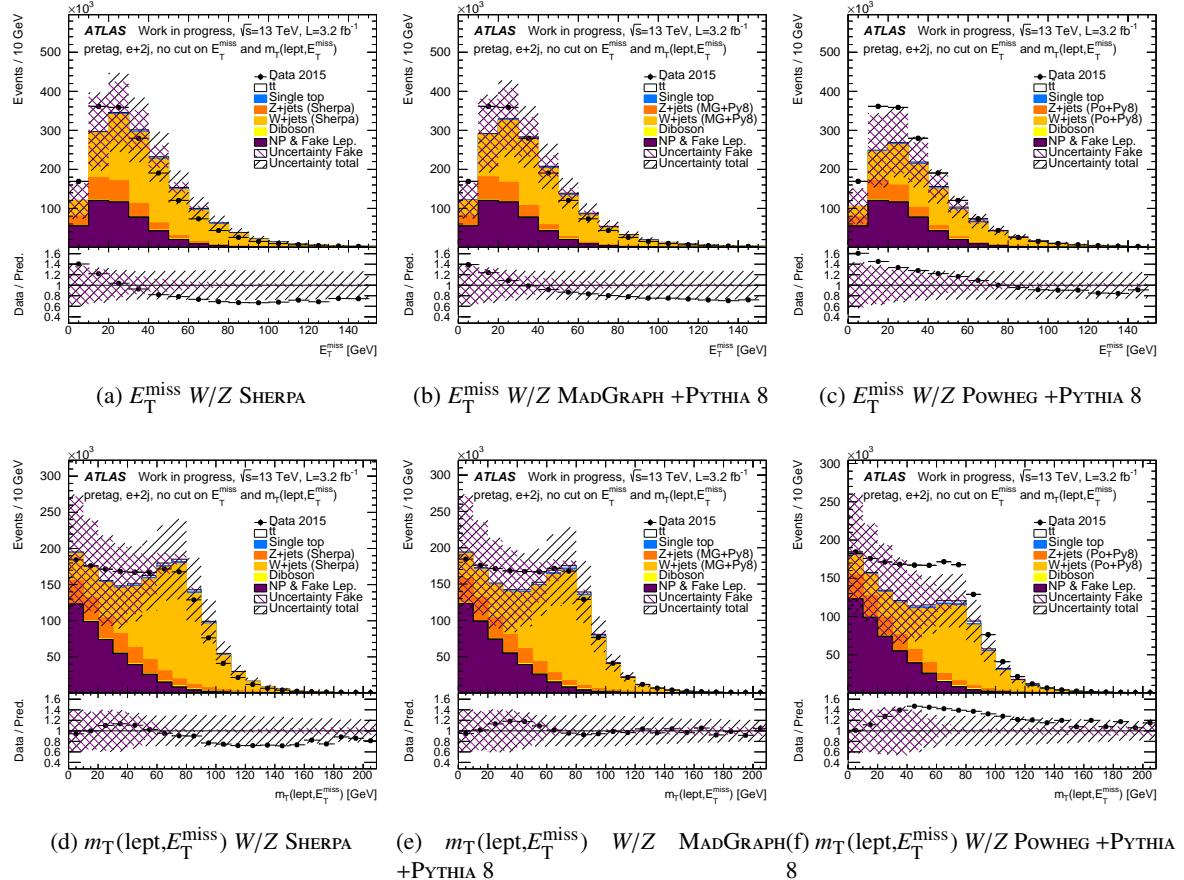


Figure 104: Distributions of (a,b,c) E_T^{miss} and (d,e,f) $m_T(\text{lept}, E_T^{\text{miss}})$ for $e + \text{jets}$ events with two jets, without any requirement on the b -jets, without any cuts on E_T^{miss} and $m_T(\text{lept}, E_T^{\text{miss}})$. The $W + \text{jets}$ and $Z + \text{jets}$ backgrounds are based on simulated samples using (a,d) SHERPA, (b,e) MADGRAPH +PYTHIA 8 and (c,f) Powheg +PYTHIA 8. The data is compared to the real lepton expectation from simulation, showing separately the contributions from $t\bar{t}$, single top, $W + \text{jets}$, $Z + \text{jets}$ and dibosons normalised to their cross-sections, and non-prompt and fake lepton backgrounds (referred to as ‘NP & Fake Lep.’) estimated with the matrix method. The purple shaded area represents the combination of the statistical and the systematic uncertainties on the matrix method estimate in each bin. The black shaded area represents the total uncertainty including also the uncertainties on the processes predicted by the MC simulation and the one on the luminosity. The lower parts of the figure show the ratios of expectation to data.

1142

Not reviewed, for internal circulation only

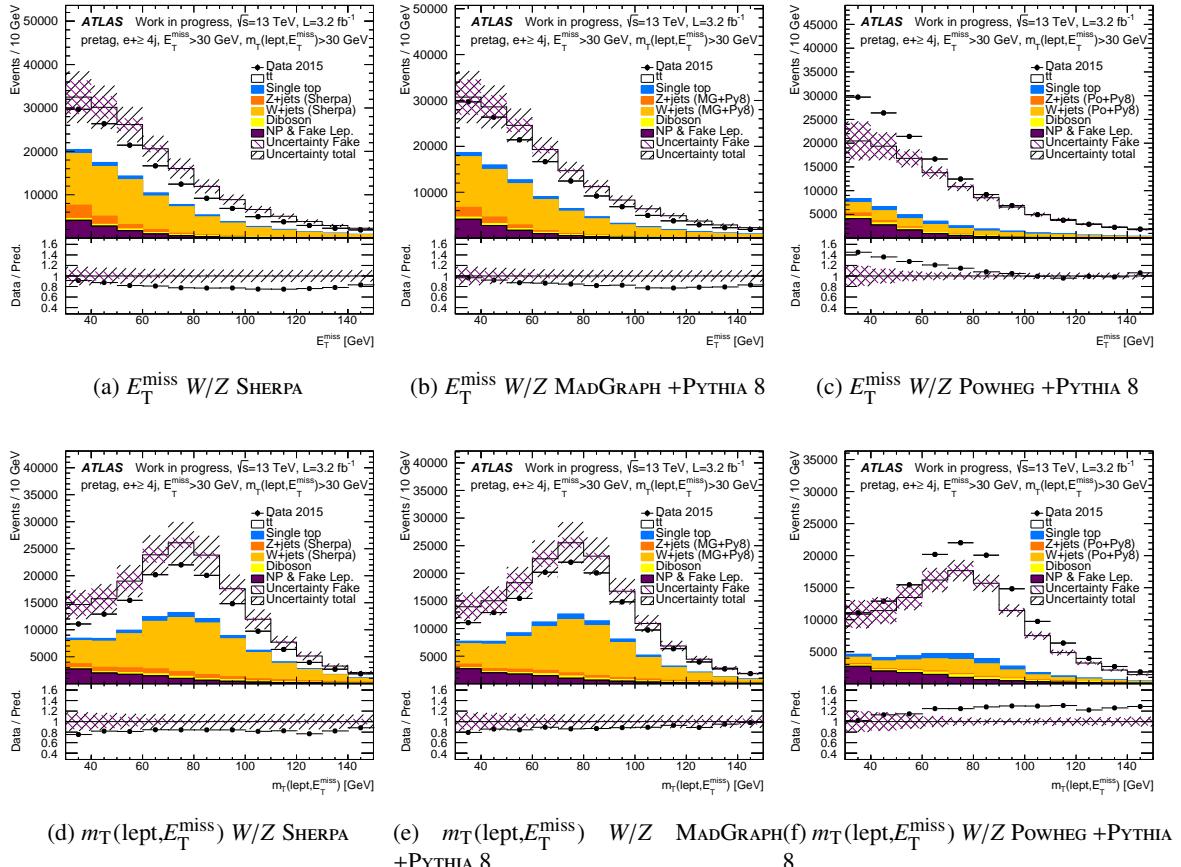


Figure 105: Distributions of (a,b,c) E_T^{miss} and (d,e,f) $m_T(\text{lept}, E_T^{\text{miss}})$ for $e+\text{jets}$ events with at least four jets, without any requirement on the b -jets, requiring $E_T^{\text{miss}} > 30 \text{ GeV}$ and $m_T(\text{lept}, E_T^{\text{miss}}) > 30 \text{ GeV}$. The $W+\text{jets}$ and $Z+\text{jets}$ backgrounds are based on simulated samples using (a,d) SHERPA, (b,e) MADGRAPH +PYTHIA 8 and (c,f) POWHEG +PYTHIA 8. The data is compared to the real lepton expectation from simulation, showing separately the contributions from $t\bar{t}$, single top, $W+\text{jets}$, $Z+\text{jets}$ and dibosons normalised to their cross-sections, and non-prompt and fake lepton backgrounds (referred to as ‘NP & Fake Lep.’) estimated with the matrix method. The purple shaded area represents the combination of the statistical and the systematic uncertainties on the matrix method estimate in each bin. The black shaded area represents the total uncertainty including also the uncertainties on the processes predicted by the MC simulation and the one on the luminosity. The lower parts of the figure show the ratios of expectation to data.

1143 E.2. Estimates for μ +jets events

1144 E.2.1. Only fake leptons

Fig. 106 to 108 show the default non-prompt and fake leptons estimates for different distributions, for different jet and b -jet multiplicities. They show also the different systematic variations for these estimates.

Not reviewed, for internal circulation only

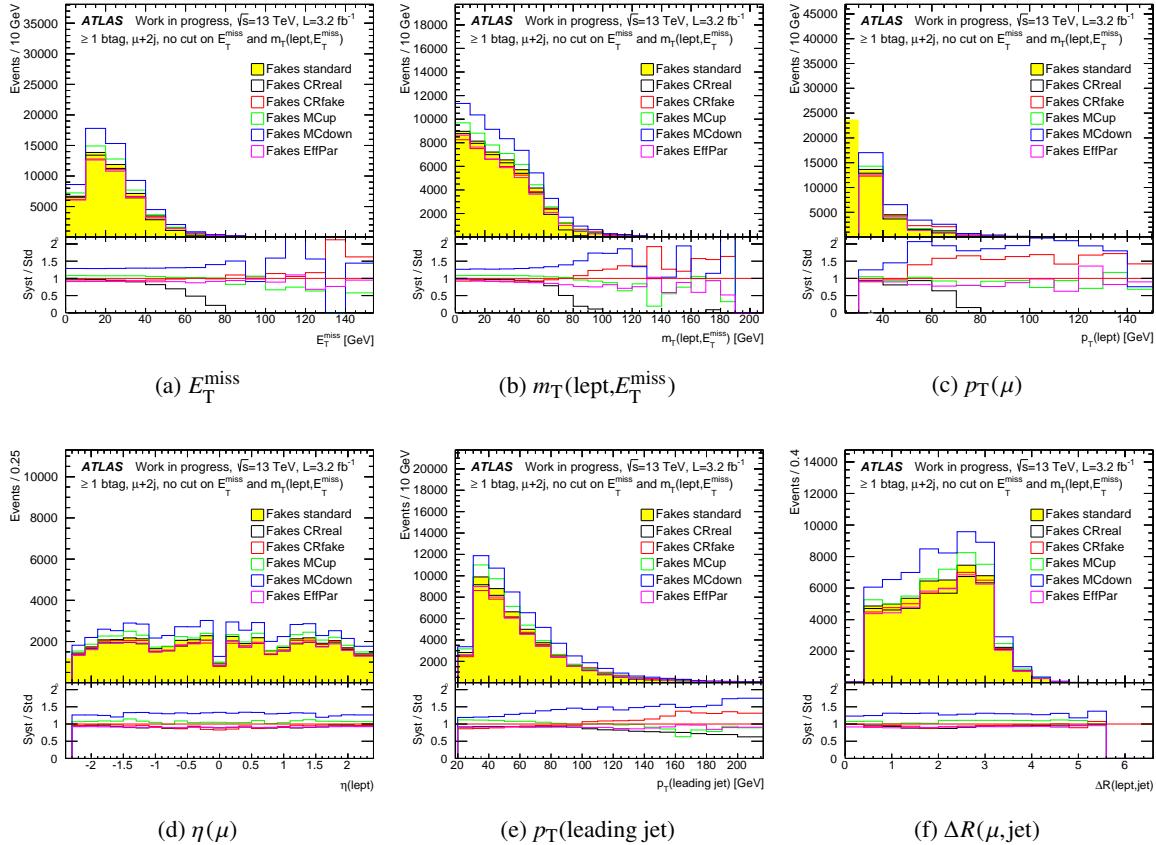


Figure 106: Distributions of (a) E_T^{miss} , (b) $m_T(\text{lept}, E_T^{\text{miss}})$, (c) p_T of the muon, (d) η of the muon, (e) p_T of the leading jet and (f) the distance ΔR of the muon to its nearest jet, for $\mu + \text{jets}$ events with two jets, with at least one b -jet, without any cuts on E_T^{miss} and $m_T(\text{lept}, E_T^{\text{miss}})$. Only the non-prompt and fake lepton backgrounds estimated with the matrix method, obtained with the different variations of both $\varepsilon_{\text{real}}$ and $\varepsilon_{\text{fake}}$, are shown.

1147

Not reviewed, for internal circulation only

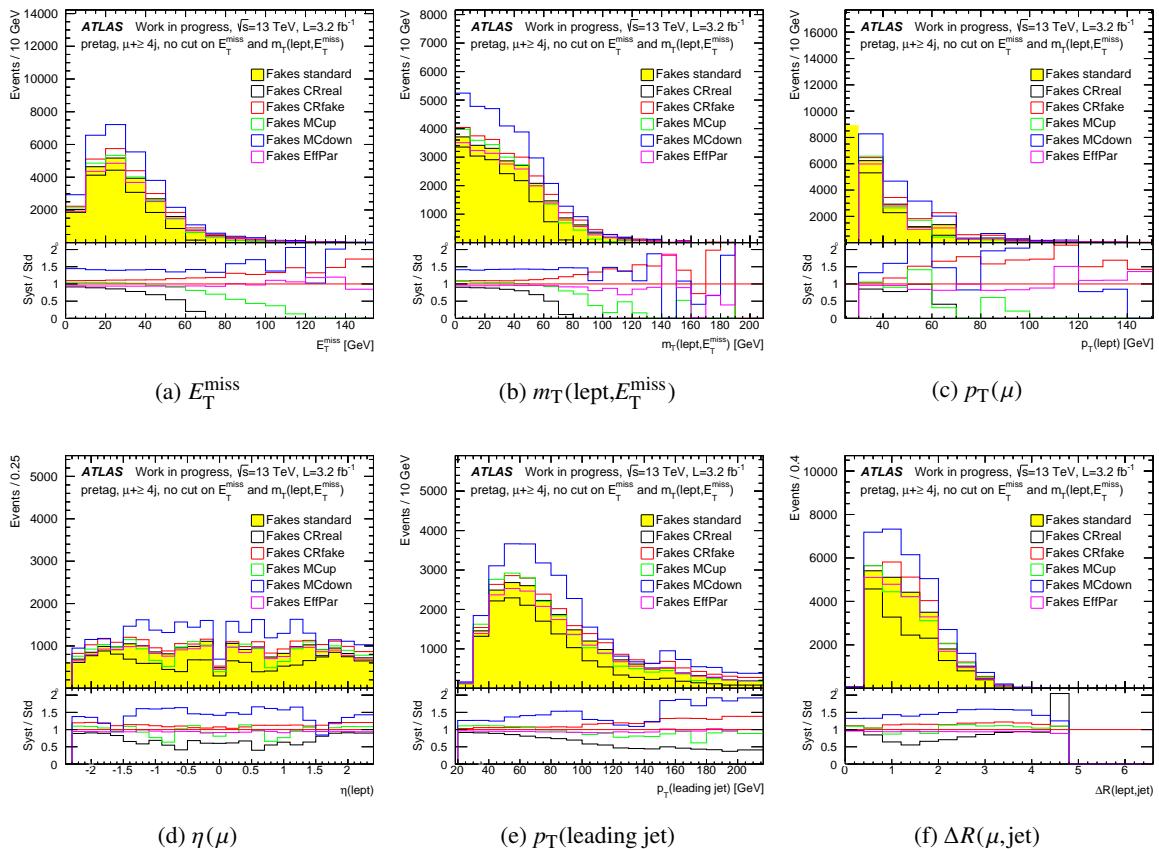


Figure 107: Distributions of (a) E_T^{miss} , (b) $m_T(\text{lept}, E_T^{\text{miss}})$, (c) p_T of the muon, (d) η of the muon, (e) p_T of the leading jet and (f) the distance ΔR of the muon to its nearest jet, for $\mu + \text{jets}$ events with at least four jets, with at least one b -jet, without any cuts on E_T^{miss} and $m_T(\text{lept}, E_T^{\text{miss}})$. Only the non-prompt and fake lepton backgrounds estimated with the matrix method, obtained with the different variations of both $\varepsilon_{\text{real}}$ and $\varepsilon_{\text{fake}}$, are shown.

Not reviewed, for internal circulation only

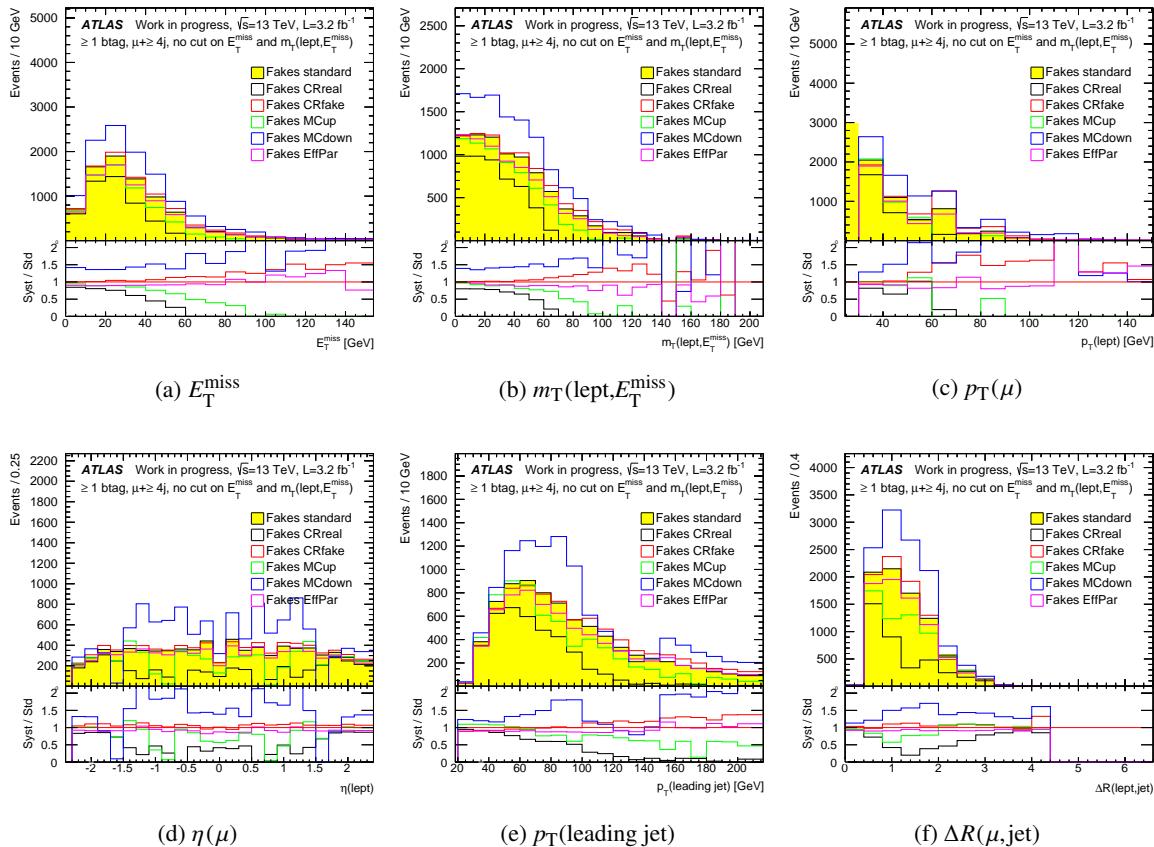


Figure 108: Distributions of (a) E_T^{miss} , (b) $m_T(\text{lept}, E_T^{\text{miss}})$, (c) p_T of the muon, (d) η of the muon, (e) p_T of the leading jet and (f) the distance ΔR of the muon to its nearest jet, for $\mu + \text{jets}$ events with at least four jets, with at least one b -jet, without any cuts on E_T^{miss} and $m_T(\text{lept}, E_T^{\text{miss}})$. Only the non-prompt and fake lepton backgrounds estimated with the matrix method, obtained with the different variations of both $\varepsilon_{\text{real}}$ and $\varepsilon_{\text{fake}}$, are shown.

1148 **E.2.2. Details on the systematic effects**

1149 Fig. 109 to 111 show the data compared to the real lepton expectation from simulation, showing the
 1150 contributions from $t\bar{t}$, single top, $W + \text{jets}$, $Z + \text{jets}$ and dibosons normalised to their cross-sections (Back-
 1151 ground), and non-prompt and fake lepton backgrounds (referred to as ‘NP & Fake Lep.’) estimated with
 1152 the matrix method for different distributions, for different jet and b -jet multiplicities. They show also the
 different systematic variations for the non-prompt and fake leptons estimates.

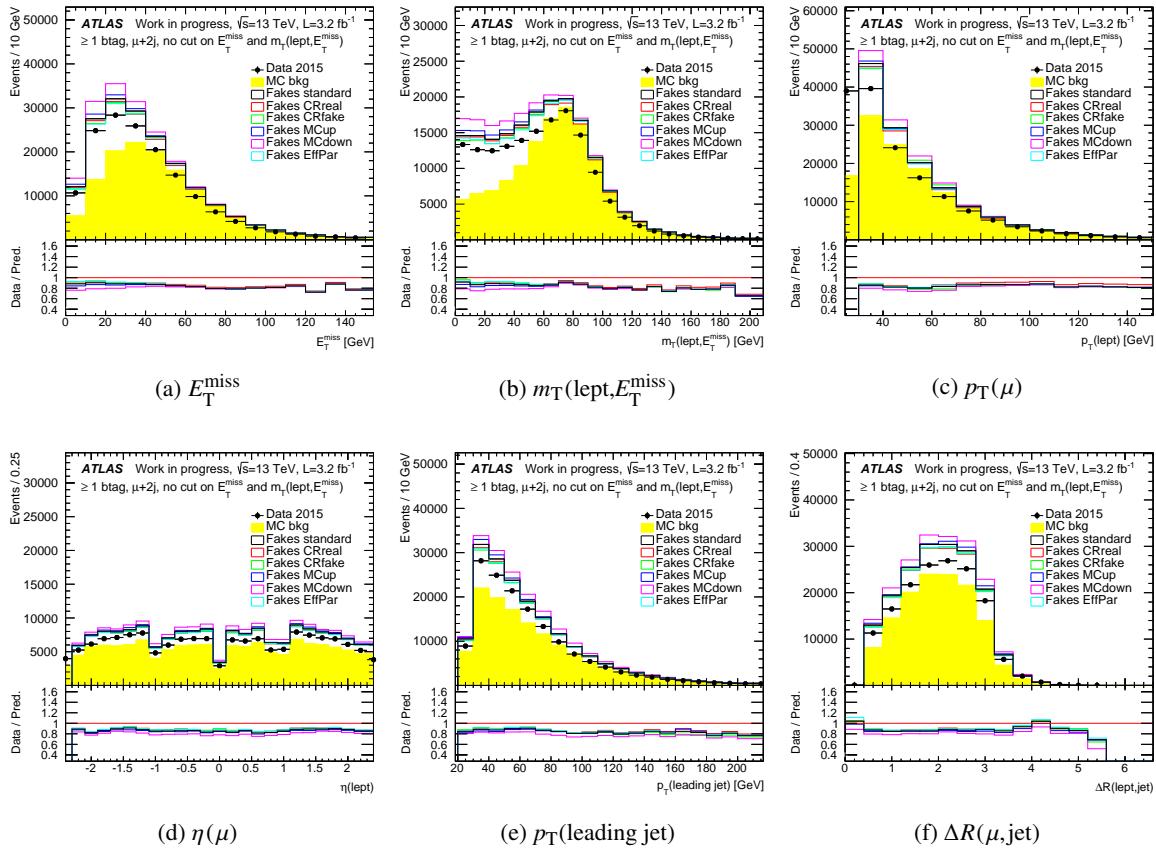


Figure 109: Distributions of (a) E_T^{miss} , (b) $m_T(\text{lept}, E_T^{\text{miss}})$, (c) p_T of the muon, (d) η of the muon, (e) p_T of the leading jet and (f) the distance ΔR of the muon to its nearest jet, for $\mu + \text{jets}$ events with two jets, with at least one b -jet, without any cuts on E_T^{miss} and $m_T(\text{lept}, E_T^{\text{miss}})$. The data is compared to the real lepton expectation from simulation, showing the contributions from $t\bar{t}$, single top, $W + \text{jets}$, $Z + \text{jets}$ and dibosons normalised to their cross-sections (Background), and non-prompt and fake lepton backgrounds (referred to as ‘NP & Fake Lep.’) estimated with the matrix method obtained with the different variations of both $\varepsilon_{\text{real}}$ and $\varepsilon_{\text{fake}}$.

Not reviewed, for internal circulation only

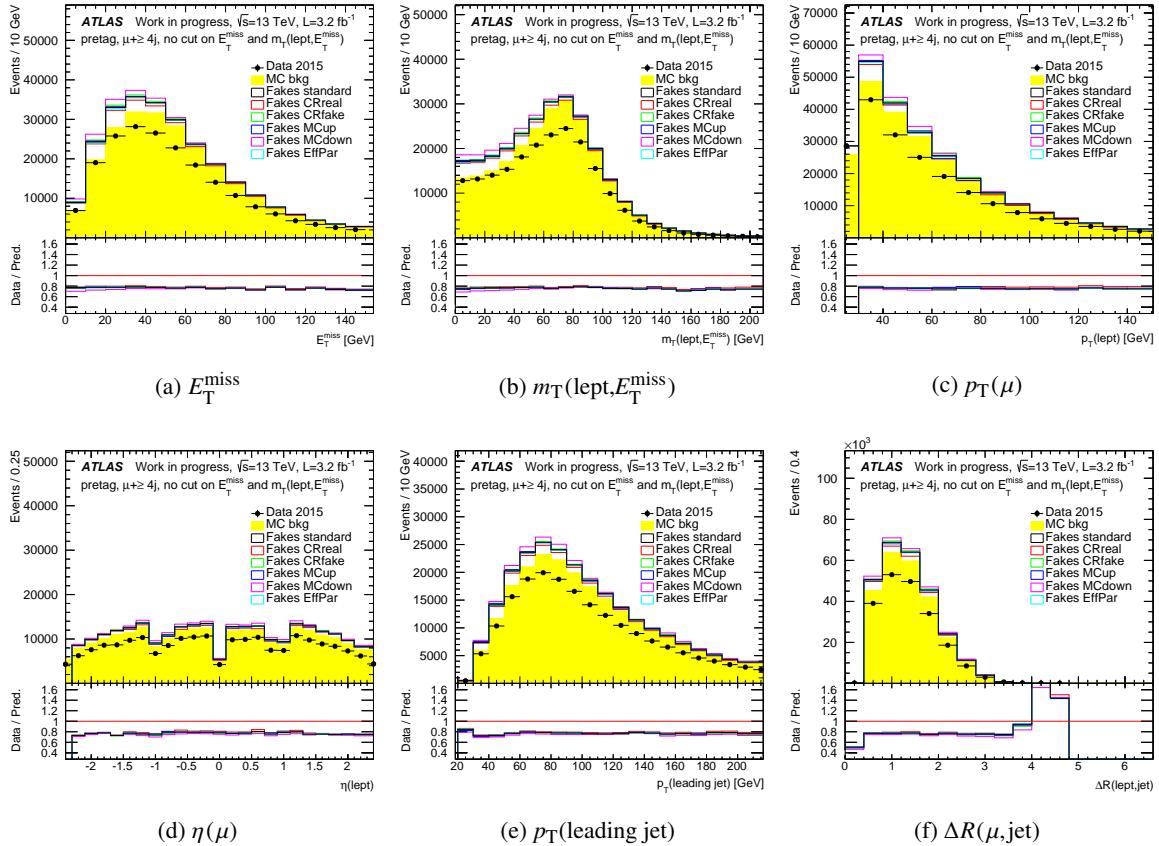


Figure 110: Distributions of (a) E_T^{miss} , (b) $m_T(\text{lept}, E_T^{\text{miss}})$, (c) p_T of the muon, (d) η of the muon, (e) p_T of the leading jet and (f) the distance ΔR of the muon to its nearest jet, for $\mu+jets$ events with at least four jets, without any b -tagging requirement, without any cuts on E_T^{miss} and $m_T(\text{lept}, E_T^{\text{miss}})$. The data is compared to the real lepton expectation from simulation, showing the contributions from $t\bar{t}$, single top, $W+jets$, $Z+jets$ and dibosons normalised to their cross-sections (Background), and non-prompt and fake lepton backgrounds (referred to as ‘NP & Fake Lep.’) estimated with the matrix method obtained with the different variations of both $\varepsilon_{\text{real}}$ and $\varepsilon_{\text{fake}}$.

Not reviewed, for internal circulation only

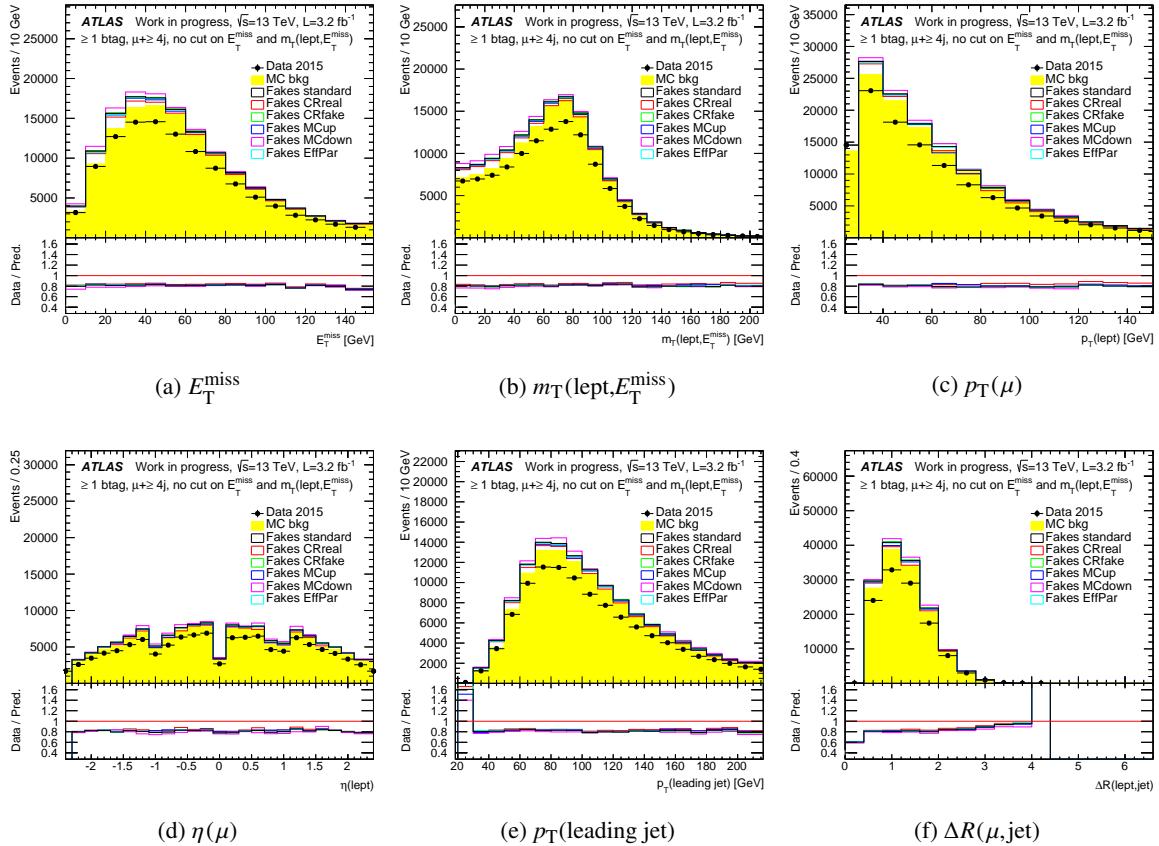


Figure 111: Distributions of (a) E_T^{miss} , (b) $m_T(\text{lept}, E_T^{\text{miss}})$, (c) p_T of the muon, (d) η of the muon, (e) p_T of the leading jet and (f) the distance ΔR of the muon to its nearest jet, for $\mu+\text{jets}$ events with at least four jets, without at least one b -jet, without any cuts on E_T^{miss} and $m_T(\text{lept}, E_T^{\text{miss}})$. The data is compared to the real lepton expectation from simulation, showing the contributions from $t\bar{t}$, single top, $W+\text{jets}$, $Z+\text{jets}$ and dibosons normalised to their cross-sections (Background), and non-prompt and fake lepton backgrounds (referred to as ‘NP & Fake Lep.’) estimated with the matrix method obtained with the different variations of both $\varepsilon_{\text{real}}$ and $\varepsilon_{\text{fake}}$.

1154 **E.2.3. Estimates without cuts on E_T^{miss} and $m_T(\text{lept}, E_T^{\text{miss}})$**

1155 Fig. 112 shows results without any cuts on E_T^{miss} and $m_T(\text{lept}, E_T^{\text{miss}})$ with two jets and at least two b -jets.

Not reviewed, for internal circulation only

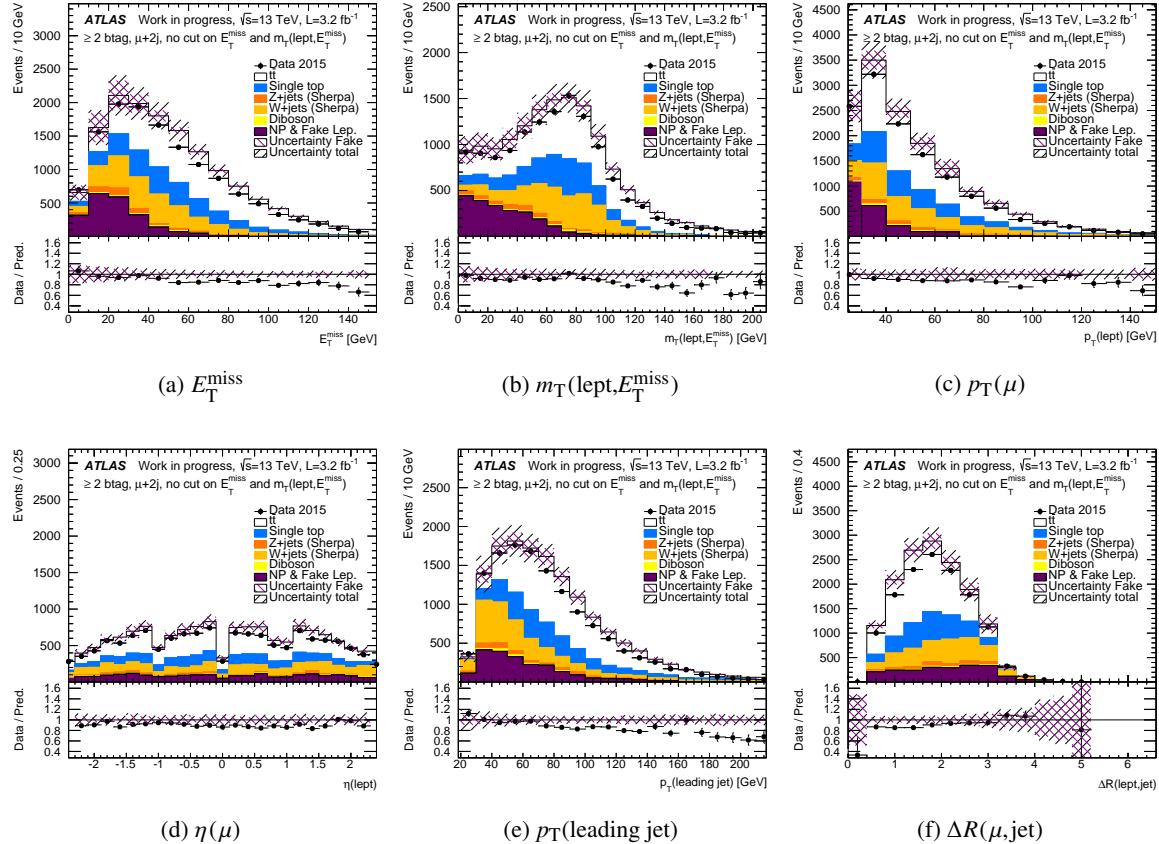


Figure 112: Distributions of (a) E_T^{miss} , (b) $m_T(\text{lept}, E_T^{\text{miss}})$, (c) p_T of the muon, (d) η of the muon, (e) p_T of the leading jet and (f) the distance ΔR of the muon to its nearest jet, for μ +jets events with two jets, requiring at least two b -jets, without any cuts on E_T^{miss} and $m_T(\text{lept}, E_T^{\text{miss}})$. The data is compared to the real lepton expectation from simulation, showing separately the contributions from $t\bar{t}$, single top, W +jets, Z +jets and dibosons normalised to their cross-sections, and non-prompt and fake lepton backgrounds (referred to as ‘NP & Fake Lep.’) estimated with the matrix method. The purple shaded area represents the combination of the statistical and the systematic uncertainties on the matrix method estimate in each bin. The black shaded area represents the total uncertainty including also the uncertainties on the processes predicted by the MC simulation and the one on the luminosity. The lower parts of the figure show the ratios of expectation to data.

1156

1157 **E.2.4. Estimates in the signal region**

1158 Fig. 113 shows results in the signal region with at least four jets and at least two b -jets.

1159 Fig. 114 to 116 shows results in the signal region with at least six jets and different b -jet multiplicities.

Not reviewed, for internal circulation only

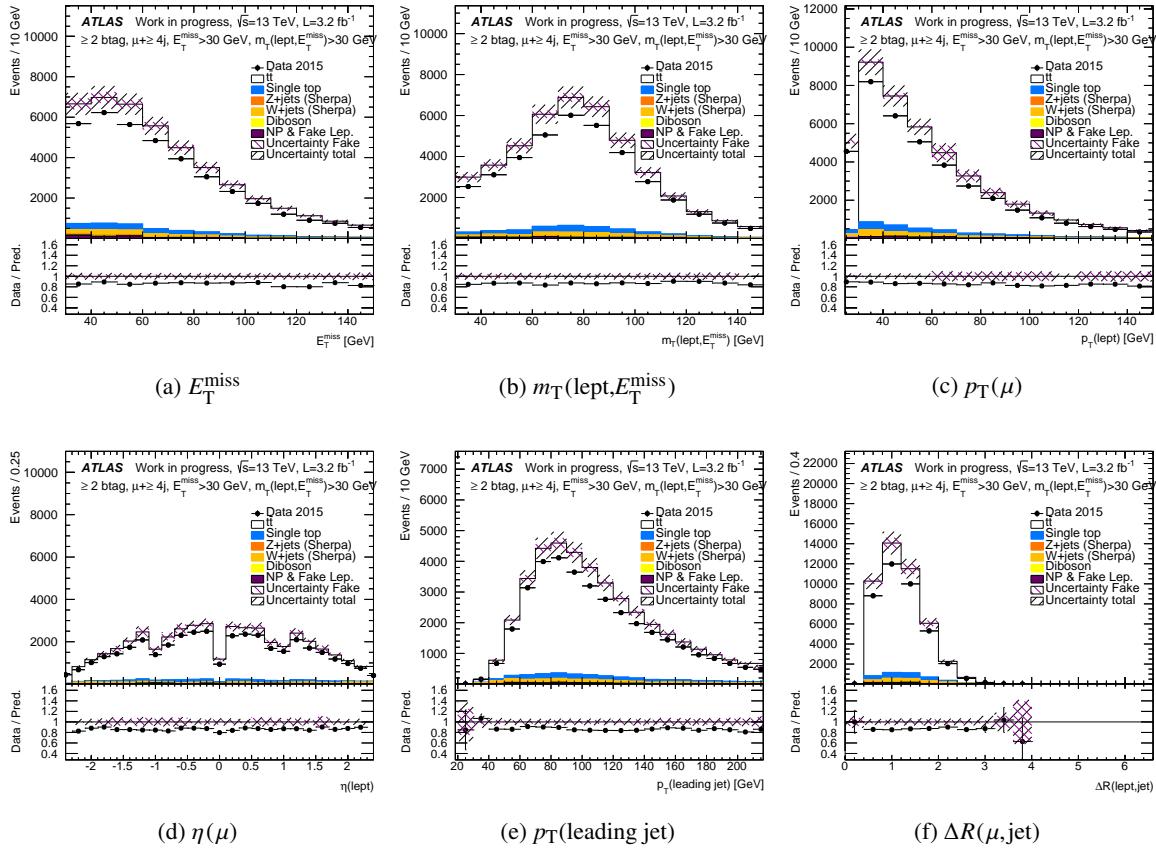


Figure 113: Distributions of (a) E_T^{miss} , (b) $m_T(\text{lept}, E_T^{\text{miss}})$, (c) p_T of the muon, (d) η of the muon, (e) p_T of the leading jet and (f) the distance ΔR of the muon to its nearest jet for $\mu+\text{jets}$ events with at least four jets, requiring at least two b -jets, requiring $E_T^{\text{miss}} > 30 \text{ GeV}$ and $m_T(\text{lept}, E_T^{\text{miss}}) > 30 \text{ GeV}$. The data is compared to the real lepton expectation from simulation, showing separately the contributions from $t\bar{t}$, single top, $W+\text{jets}$, $Z+\text{jets}$ and dibosons normalised to their cross-sections, and non-prompt and fake lepton backgrounds (referred to as ‘NP & Fake Lep.’) estimated with the matrix method. The purple shaded area represents the combination of the statistical and the systematic uncertainties on the matrix method estimate in each bin. The black shaded area represents the total uncertainty including also the uncertainties on the processes predicted by the MC simulation and the one on the luminosity. The lower parts of the figure show the ratios of expectation to data.

Not reviewed, for internal circulation only

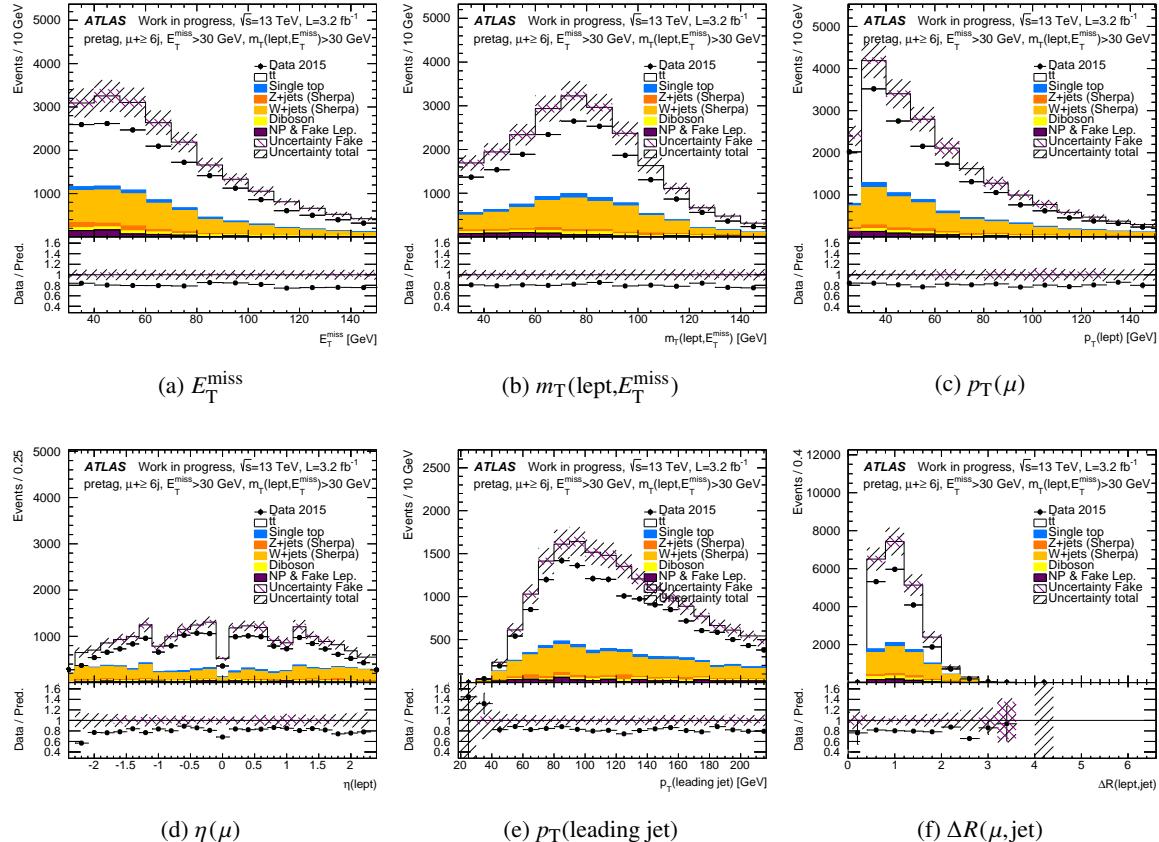


Figure 114: Distributions of (a) E_T^{miss} , (b) $m_T(\text{lept}, E_T^{\text{miss}})$, (c) p_T of the muon, (d) η of the muon, (e) p_T of the leading jet and (f) the distance ΔR of the muon to its nearest jet, for $\mu + \text{jets}$ events with at least six jets without any b -tagging requirement, requiring $E_T^{\text{miss}} > 30 \text{ GeV}$ and $m_T(\text{lept}, E_T^{\text{miss}}) > 30 \text{ GeV}$. The data is compared to the real lepton expectation from simulation, showing separately the contributions from $t\bar{t}$, single top, $W + \text{jets}$, $Z + \text{jets}$ and dibosons normalised to their cross-sections, and non-prompt and fake lepton backgrounds (referred to as ‘NP & Fake Lep.’) estimated with the matrix method. The purple shaded area represents the combination of the statistical and the systematic uncertainties on the matrix method estimate in each bin. The black shaded area represents the total uncertainty including also the uncertainties on the processes predicted by the MC simulation and the one on the luminosity. The lower parts of the figure show the ratios of expectation to data.

Not reviewed, for internal circulation only

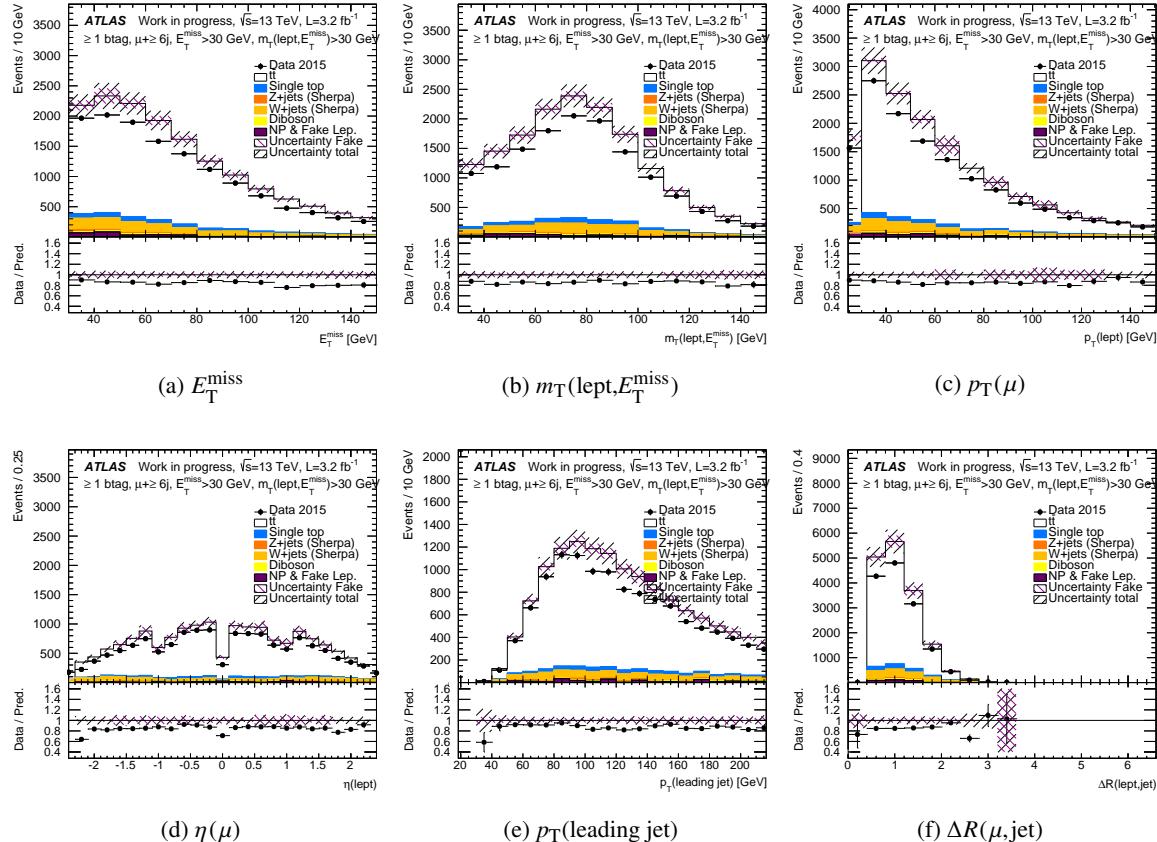


Figure 115: Distributions of (a) E_T^{miss} , (b) $m_T(\text{lept}, E_T^{\text{miss}})$, (c) p_T of the muon, (d) η of the muon, (e) p_T of the leading jet and (f) the distance ΔR of the muon to its nearest jet, for $\mu + \text{jets}$ events with at least six jets, with at least one b -jet, requiring $E_T^{\text{miss}} > 30 \text{ GeV}$ and $m_T(\text{lept}, E_T^{\text{miss}}) > 30 \text{ GeV}$. The data is compared to the real lepton expectation from simulation, showing separately the contributions from $t\bar{t}$, single top, $W + \text{jets}$, $Z + \text{jets}$ and dibosons normalised to their cross-sections, and non-prompt and fake lepton backgrounds (referred to as ‘NP & Fake Lep.’) estimated with the matrix method. The purple shaded area represents the combination of the statistical and the systematic uncertainties on the matrix method estimate in each bin. The black shaded area represents the total uncertainty including also the uncertainties on the processes predicted by the MC simulation and the one on the luminosity. The lower parts of the figure show the ratios of expectation to data.

Not reviewed, for internal circulation only

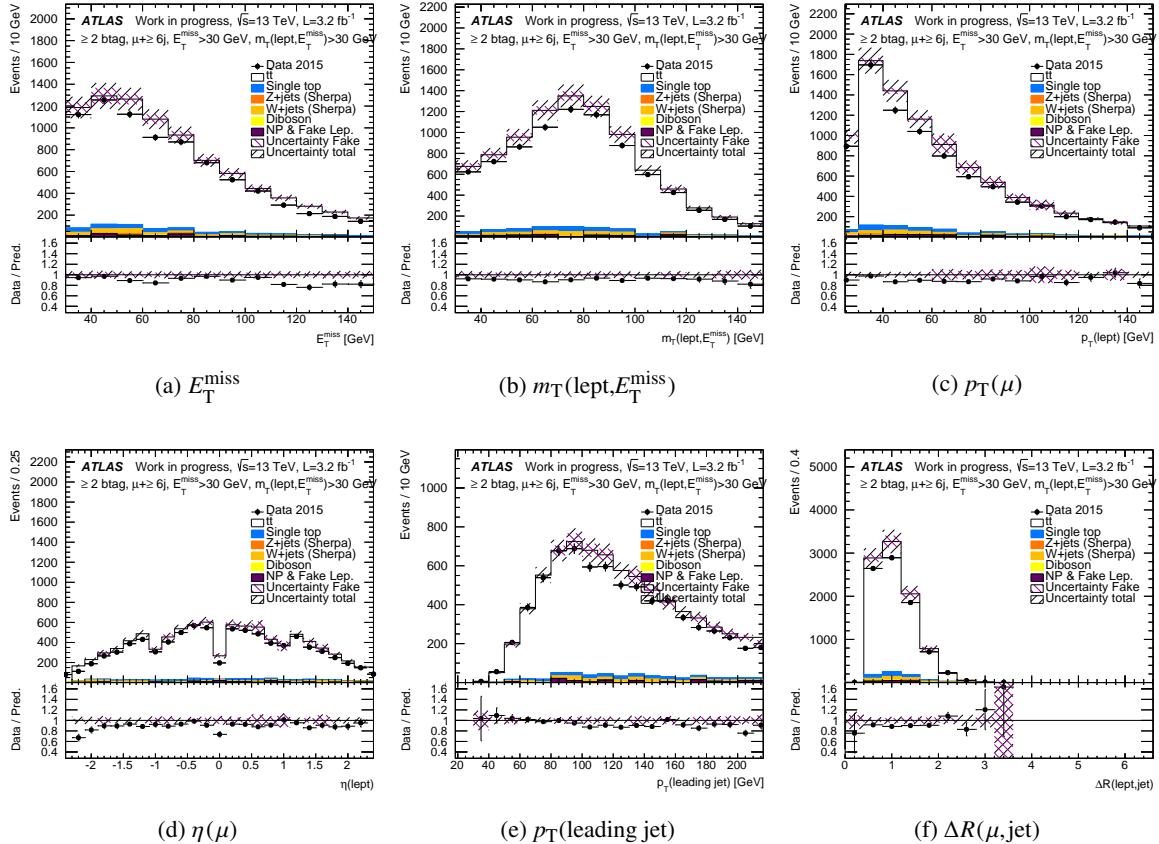


Figure 116: Distributions of (a) E_T^{miss} , (b) $m_T(\text{lept}, E_T^{\text{miss}})$, (c) p_T of the muon, (d) η of the muon, (e) p_T of the leading jet and (f) the distance ΔR of the muon to its nearest jet, for $\mu + \text{jets}$ events with at least six jets, without at least two b -jets, requiring $E_T^{\text{miss}} > 30 \text{ GeV}$ and $m_T(\text{lept}, E_T^{\text{miss}}) > 30 \text{ GeV}$. The data is compared to the real lepton expectation from simulation, showing separately the contributions from $t\bar{t}$, single top, $W + \text{jets}$, $Z + \text{jets}$ and dibosons normalised to their cross-sections, and non-prompt and fake lepton backgrounds (referred to as ‘NP & Fake Lep.’) estimated with the matrix method. The purple shaded area represents the combination of the statistical and the systematic uncertainties on the matrix method estimate in each bin. The black shaded area represents the total uncertainty including also the uncertainties on the processes predicted by the MC simulation and the one on the luminosity. The lower parts of the figure show the ratios of expectation to data.

1160 **E.2.5. Estimates in the control regions**

1161 Fig. 117 to 119 show control plots for various distributions and different jet and b -tag multiplicities after applying the cuts $E_T^{\text{miss}} + m_T(\text{lept}, E_T^{\text{miss}}) < 60 \text{ GeV}$.

Not reviewed, for internal circulation only

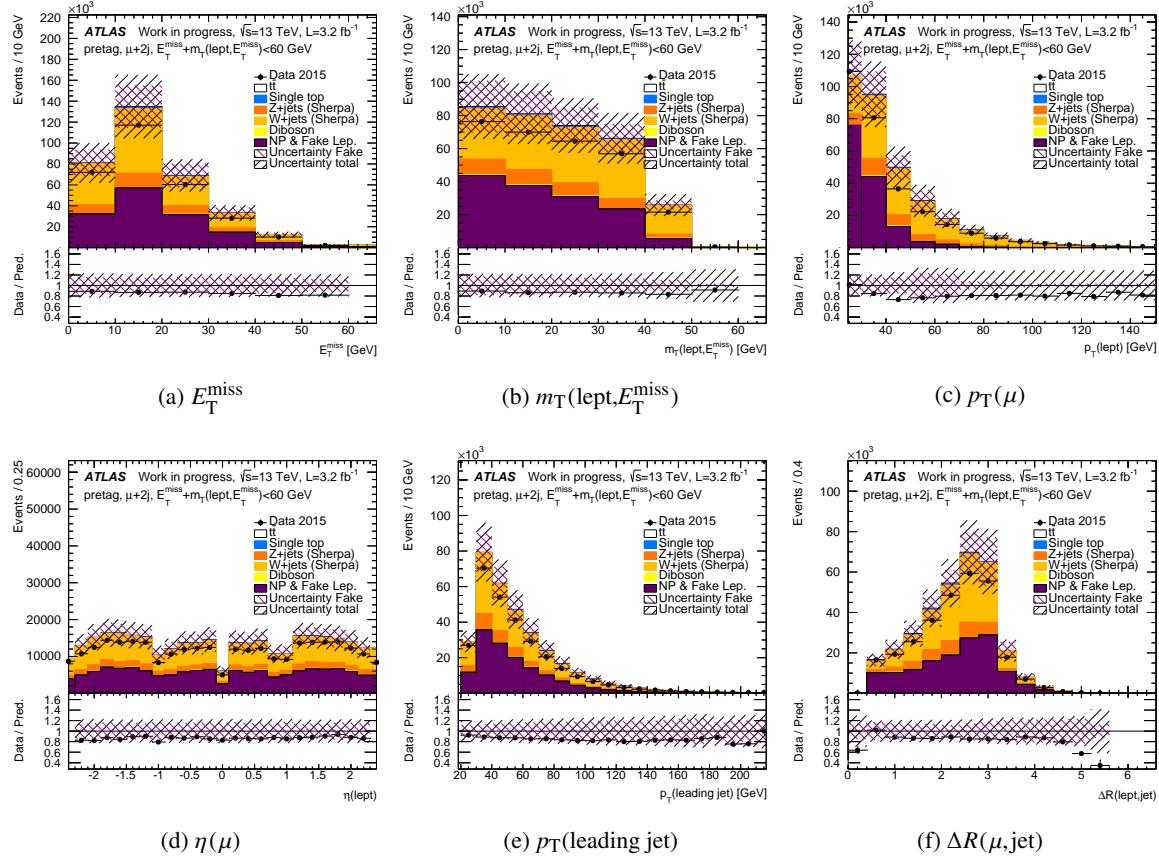


Figure 117: Distributions of (a) E_T^{miss} , (b) $m_T(\text{lept}, E_T^{\text{miss}})$, (c) p_T of the muon, (d) η of the muon, (e) p_T of the leading jet and (f) the distance ΔR of the muon to its nearest jet, for $\mu+jets$ events with two jets, without any b -tagging requirement, requiring $E_T^{\text{miss}} + m_T(\text{lept}, E_T^{\text{miss}}) < 60 \text{ GeV}$. The data is compared to the real lepton expectation from simulation, showing separately the contributions from $t\bar{t}$, single top, $W+jets$, $Z+jets$ and dibosons normalised to their cross-sections, and non-prompt and fake lepton backgrounds (referred to as ‘NP & Fake Lep.’) estimated with the matrix method. The purple shaded area represents the combination of the statistical and the systematic uncertainties on the matrix method estimate in each bin. The black shaded area represents the total uncertainty including also the uncertainties on the processes predicted by the MC simulation and the one on the luminosity. The lower parts of the figure show the ratios of expectation to data.

1162

Not reviewed, for internal circulation only

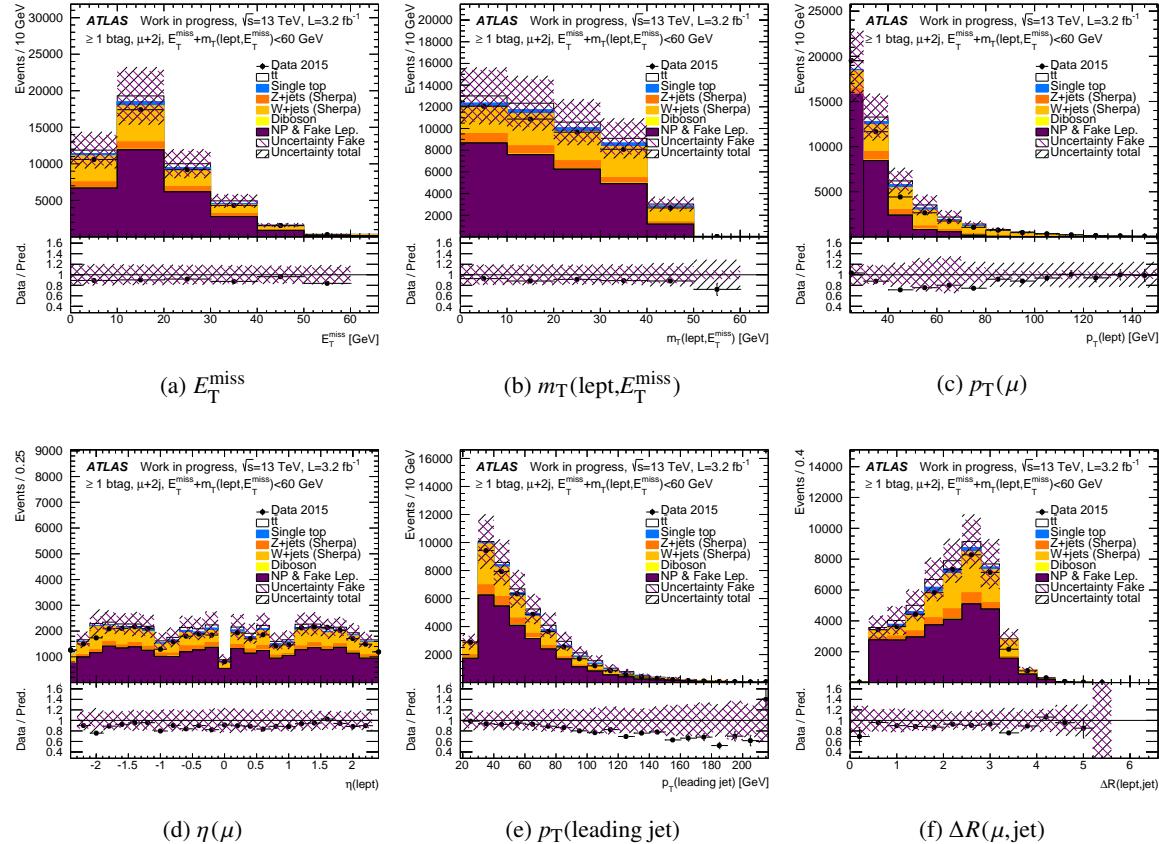


Figure 118: Distributions of (a) E_T^{miss} , (b) $m_T(\text{lept}, E_T^{\text{miss}})$, (c) p_T of the muon, (d) η of the muon, (e) p_T of the leading jet and (f) the distance ΔR of the muon to its nearest jet, for $\mu + \text{jets}$ events with two jets, requiring at least one b -jet, requiring $E_T^{\text{miss}} + m_T(\text{lept}, E_T^{\text{miss}}) < 60 \text{ GeV}$. The data is compared to the real lepton expectation from simulation, showing separately the contributions from $t\bar{t}$, single top, $W + \text{jets}$, $Z + \text{jets}$ and dibosons normalised to their cross-sections, and non-prompt and fake lepton backgrounds (referred to as ‘NP & Fake Lep.’) estimated with the matrix method. The purple shaded area represents the combination of the statistical and the systematic uncertainties on the matrix method estimate in each bin. The black shaded area represents the total uncertainty including also the uncertainties on the processes predicted by the MC simulation and the one on the luminosity. The lower parts of the figure show the ratios of expectation to data.

Not reviewed, for internal circulation only

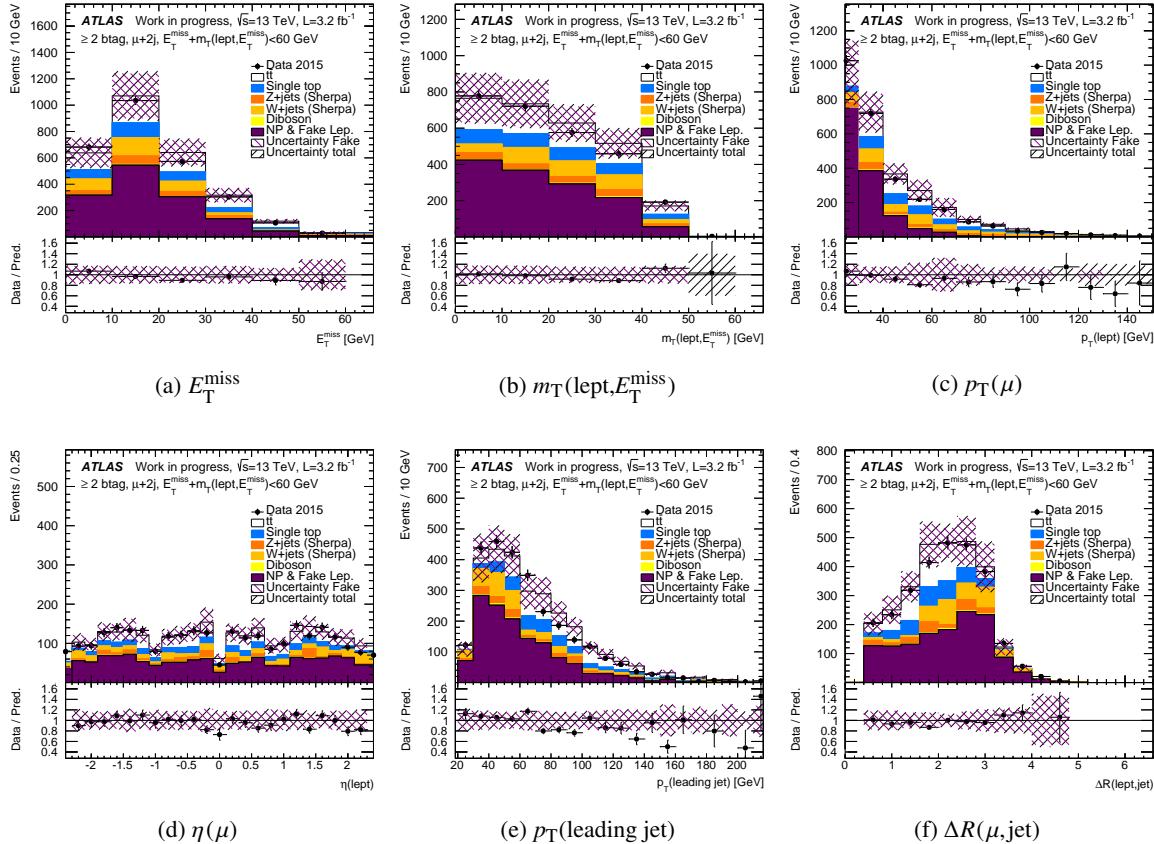


Figure 119: Distributions of (a) E_T^{miss} , (b) $m_T(\text{lept}, E_T^{\text{miss}})$, (c) p_T of the muon, (d) η of the muon, (e) p_T of the leading jet and (f) the distance ΔR of the muon to its nearest jet, for $\mu+\text{jets}$ events with two jets, requiring at least two b -jets, requiring $E_T^{\text{miss}} + m_T(\text{lept}, E_T^{\text{miss}}) < 60 \text{ GeV}$. The data is compared to the real lepton expectation from simulation, showing separately the contributions from $t\bar{t}$, single top, $W+\text{jets}$, $Z+\text{jets}$ and dibosons normalised to their cross-sections, and non-prompt and fake lepton backgrounds (referred to as ‘NP & Fake Lep.’) estimated with the matrix method. The purple shaded area represents the combination of the statistical and the systematic uncertainties on the matrix method estimate in each bin. The black shaded area represents the total uncertainty including also the uncertainties on the processes predicted by the MC simulation and the one on the luminosity. The lower parts of the figure show the ratios of expectation to data.

1163 **E.2.6. Estimates with other MC samples**

1164 Fig. 120 and 121 show E_T^{miss} and $m_T(\text{lept}, E_T^{\text{miss}})$ distributions using three different MC samples to estimate the $W + \text{jets}$ and $Z + \text{jets}$ background, based on SHERPA, MADGRAPH +PYTHIA 8 and Powheg +PYTHIA 8.

Not reviewed, for internal circulation only

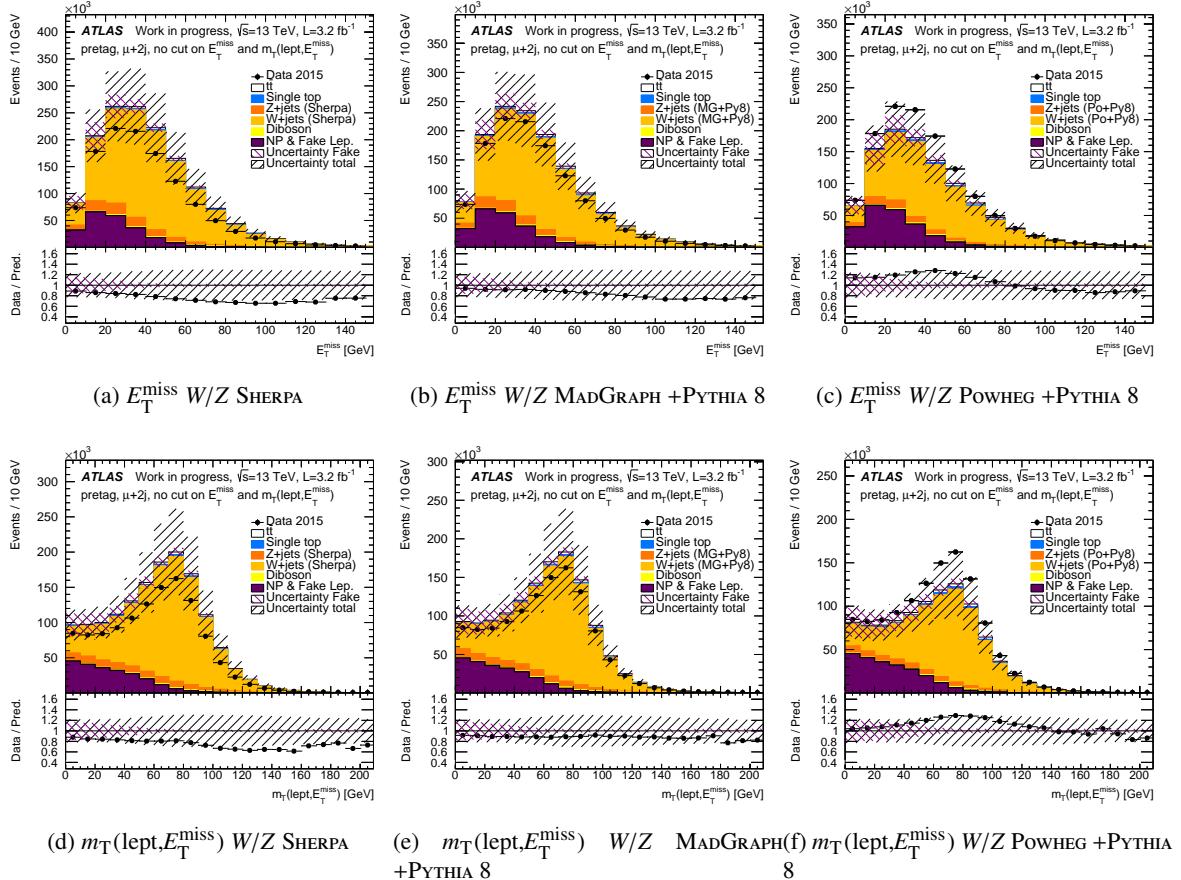


Figure 120: Distributions of (a,b,c) E_T^{miss} and (d,e,f) $m_T(\text{lept}, E_T^{\text{miss}})$ for $\mu + \text{jets}$ events with two jets, without any requirement on the b -jets, without any cuts on E_T^{miss} and $m_T(\text{lept}, E_T^{\text{miss}})$. The $W + \text{jets}$ and $Z + \text{jets}$ backgrounds are based on simulated samples using (a,d) SHERPA, (b,e) MADGRAPH +PYTHIA 8 and (c,f) POWHEG +PYTHIA 8. The data is compared to the real lepton expectation from simulation, showing separately the contributions from $t\bar{t}$, single top, $W + \text{jets}$, $Z + \text{jets}$ and dibosons normalised to their cross-sections, and non-prompt and fake lepton backgrounds (referred to as ‘NP & Fake Lep.’) estimated with the matrix method. The purple shaded area represents the combination of the statistical and the systematic uncertainties on the matrix method estimate in each bin. The black shaded area represents the total uncertainty including also the uncertainties on the processes predicted by the MC simulation and the one on the luminosity. The lower parts of the figure show the ratios of expectation to data.

1165

Not reviewed, for internal circulation only

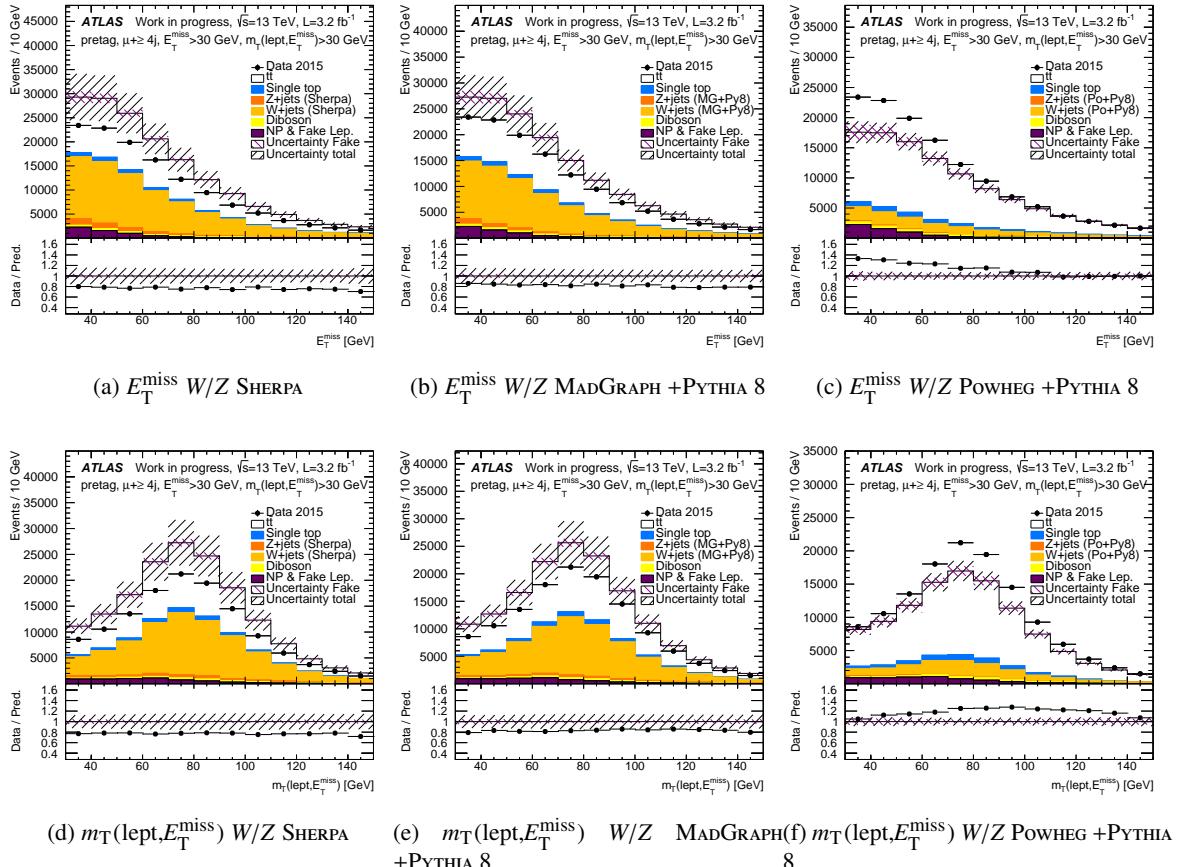


Figure 121: Distributions of (a,b,c) E_T^{miss} and (d,e,f) $m_T(\text{lept}, E_T^{\text{miss}})$ for $\mu+\text{jets}$ events with at least four jets, without any requirement on the b -jets, requiring $E_T^{\text{miss}} > 30 \text{ GeV}$ and $m_T(\text{lept}, E_T^{\text{miss}}) > 30 \text{ GeV}$. The $W+\text{jets}$ and $Z+\text{jets}$ backgrounds are based on simulated samples using (a,d) SHERPA, (b,e) MADGRAPH +PYTHIA 8 and (c,f) POWHEG +PYTHIA 8. The data is compared to the real lepton expectation from simulation, showing separately the contributions from $t\bar{t}$, single top, $W+\text{jets}$, $Z+\text{jets}$ and dibosons normalised to their cross-sections, and non-prompt and fake lepton backgrounds (referred to as ‘NP & Fake Lep.’) estimated with the matrix method. The purple shaded area represents the combination of the statistical and the systematic uncertainties on the matrix method estimate in each bin. The black shaded area represents the total uncertainty including also the uncertainties on the processes predicted by the MC simulation and the one on the luminosity. The lower parts of the figure show the ratios of expectation to data.

F. Additional studies for dilepton $t\bar{t}$ events

1166 In this section we show additional control plots obtained for the same-sign and the opposite-sign samples
1168 with the non-prompt and fake lepton background estimates obtained from the matrix method.

1169 **F.1. Estimates for opposite-sign dilepton events**

1170 Main results concerning the opposite-sign sample have been shown in section 8.2.

1171 Fig. 122 to 129 show control plots in the three channels ee , $\mu\mu$ and $e\mu$ in the pretag and b -tagged samples,
1172 for events with at least two jets. The samples in the ee and $\mu\mu$ channels are dominated by the $Z + \text{jets}$
1173 events in the pretag sample. But this is largely reduced in the b -tagged sample. As already seen the level
1174 of non-prompt and fake lepton backgrounds in these samples is very small, around or below the percent
1175 level.

Not reviewed, for internal circulation only

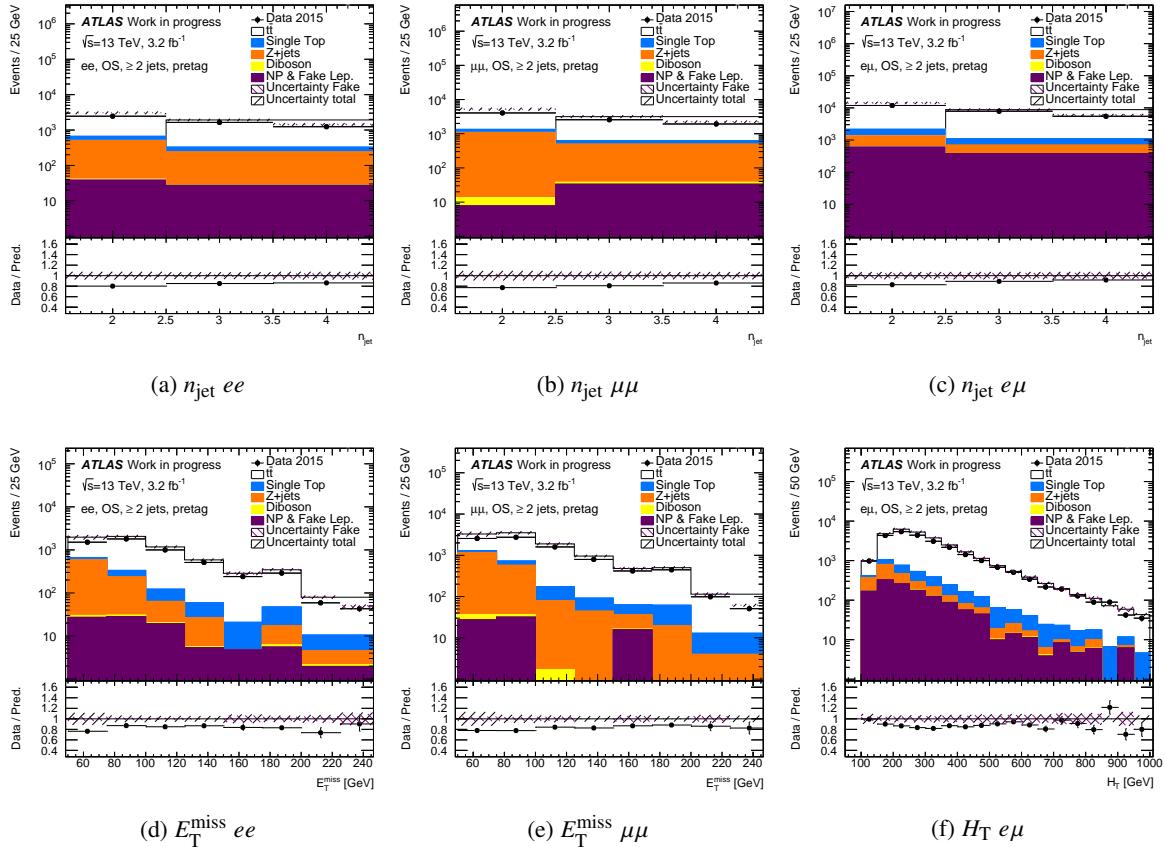


Figure 122: Distributions of (a,b,c) jet multiplicity n_{jet} and (d,e,f) $E_{\text{T}}^{\text{miss}}$ in events with opposite-sign (a,d) ee , (b,e) $\mu\mu$ and (d,f) $e\mu$ pairs with at least two jets, without requirement on the number of b -jets, and with the full selection except cuts on $E_{\text{T}}^{\text{miss}}$ and $m_{\text{T}}(\text{lept}, E_{\text{T}}^{\text{miss}})$. In the $e\mu$ channel the $E_{\text{T}}^{\text{miss}}$ is replaced by the H_{T} distribution. The data is compared to the real lepton expectation from simulation, showing separately the contributions from $t\bar{t}$, single top, Z +jets and dibosons normalised to their cross-sections, and non-prompt and fake lepton backgrounds (referred to as ‘NP & Fake Lep.’) estimated with the matrix method. The purple shaded area represents the combination of the statistical and the systematic uncertainties on the matrix method estimate in each bin. The black shaded area represents the total uncertainty including also the uncertainties on the processes predicted by the MC simulation and the one on the luminosity. The lower parts of the figure show the ratios of expectation to data.

Not reviewed, for internal circulation only

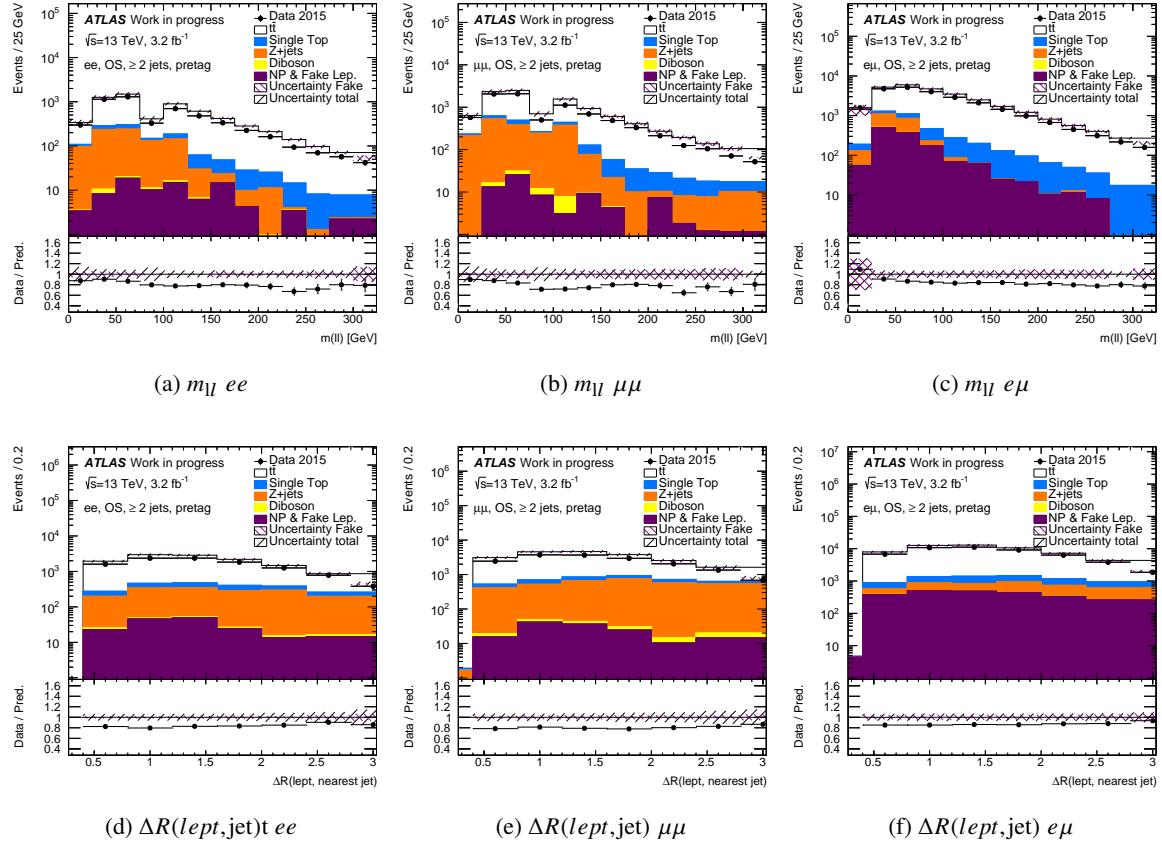


Figure 123: Distributions of (a,b,c) the invariant mass of the two leptons m_{ll} and (d,e,f) the distance $\Delta R(\text{lept}, \text{jet})$ between one of the lepton and its nearest jet, in events with opposite-sign (a,d) ee , (b,e) $\mu\mu$ and (d,f) $e\mu$ pairs with at least two jets, without requirement on the number of b -jets, and with the full selection. The data is compared to the real lepton expectation from simulation, showing separately the contributions from $t\bar{t}$, single top, $Z + \text{jets}$ and dibosons normalised to their cross-sections, and non-prompt and fake lepton backgrounds (referred to as ‘NP & Fake Lep.’) estimated with the matrix method. The purple shaded area represents the combination of the statistical and the systematic uncertainties on the matrix method estimate in each bin. The black shaded area represents the total uncertainty including also the uncertainties on the processes predicted by the MC simulation and the one on the luminosity. The lower parts of the figure show the ratios of expectation to data.

Not reviewed, for internal circulation only

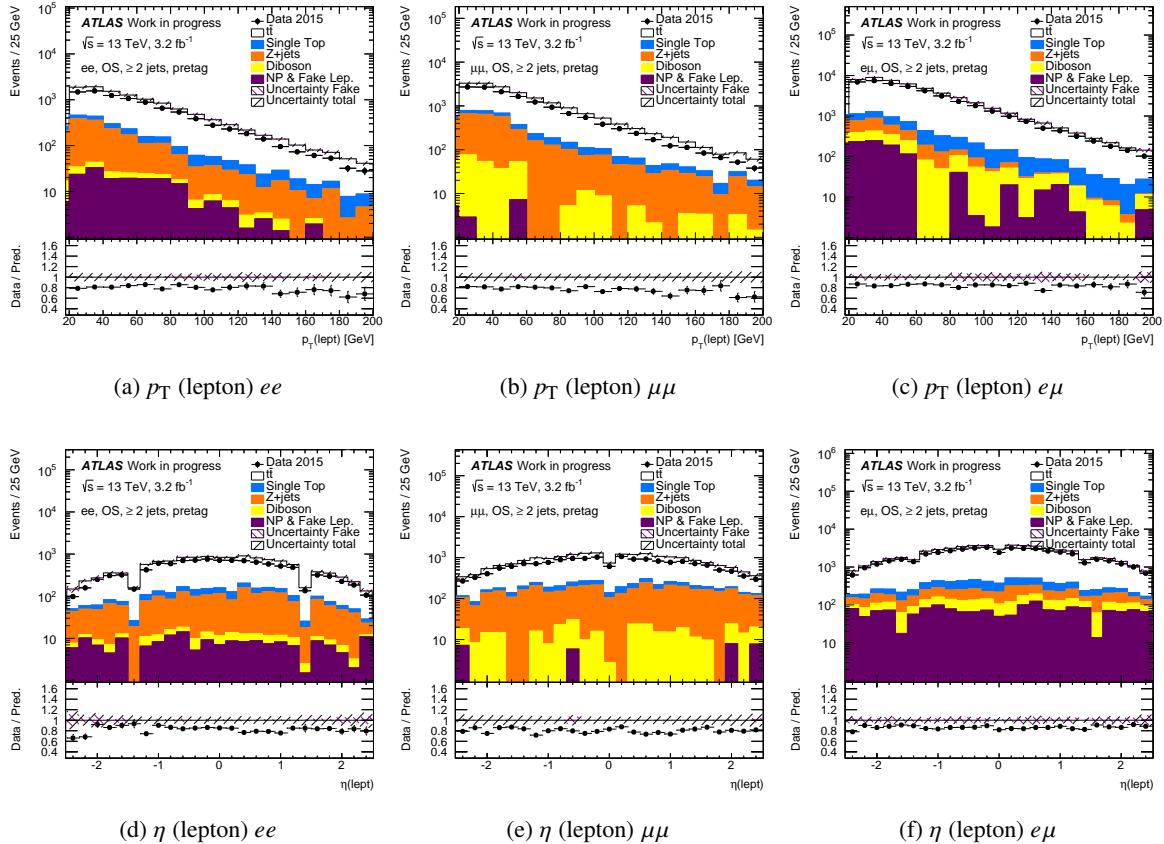


Figure 124: Distributions of (a,b,c) p_T and (d,e,f) η of the lepton, in events with opposite-sign (a,d) ee , (b,e) $\mu\mu$ and (d,f) $e\mu$ pairs with at least two jets, without requirement on the number of b -jets, and with the full selection. The data is compared to the real lepton expectation from simulation, showing separately the contributions from $t\bar{t}$, single top, Z +jets and dibosons normalised to their cross-sections, and non-prompt and fake lepton backgrounds (referred to as ‘NP & Fake Lep.’) estimated with the matrix method. The purple shaded area represents the combination of the statistical and the systematic uncertainties on the matrix method estimate in each bin. The black shaded area represents the total uncertainty including also the uncertainties on the processes predicted by the MC simulation and the one on the luminosity. The lower parts of the figure show the ratios of expectation to data.

Not reviewed, for internal circulation only

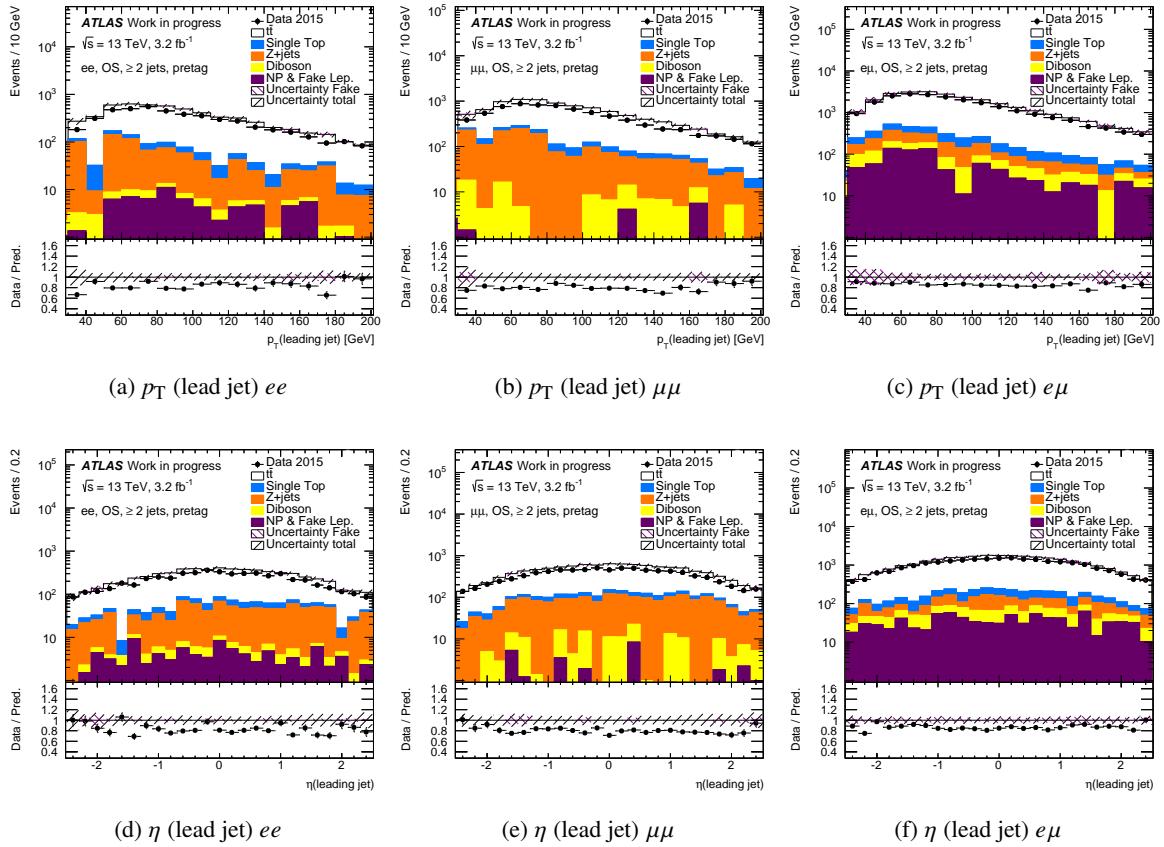


Figure 125: Distributions of (a,b,c) p_T and (d,e,f) η of the leading jet, in events with opposite-sign (a,d) ee , (b,e) $\mu\mu$ and (d,f) $e\mu$ pairs with at least two jets, without requirement on the number of b -jets, and with the full selection. The data is compared to the real lepton expectation from simulation, showing separately the contributions from $t\bar{t}$, single top, Z+jets and dibosons normalised to their cross-sections, and non-prompt and fake lepton backgrounds (referred to as ‘NP & Fake Lep.’) estimated with the matrix method. The purple shaded area represents the combination of the statistical and the systematic uncertainties on the matrix method estimate in each bin. The black shaded area represents the total uncertainty including also the uncertainties on the processes predicted by the MC simulation and the one on the luminosity. The lower parts of the figure show the ratios of expectation to data.

Not reviewed, for internal circulation only

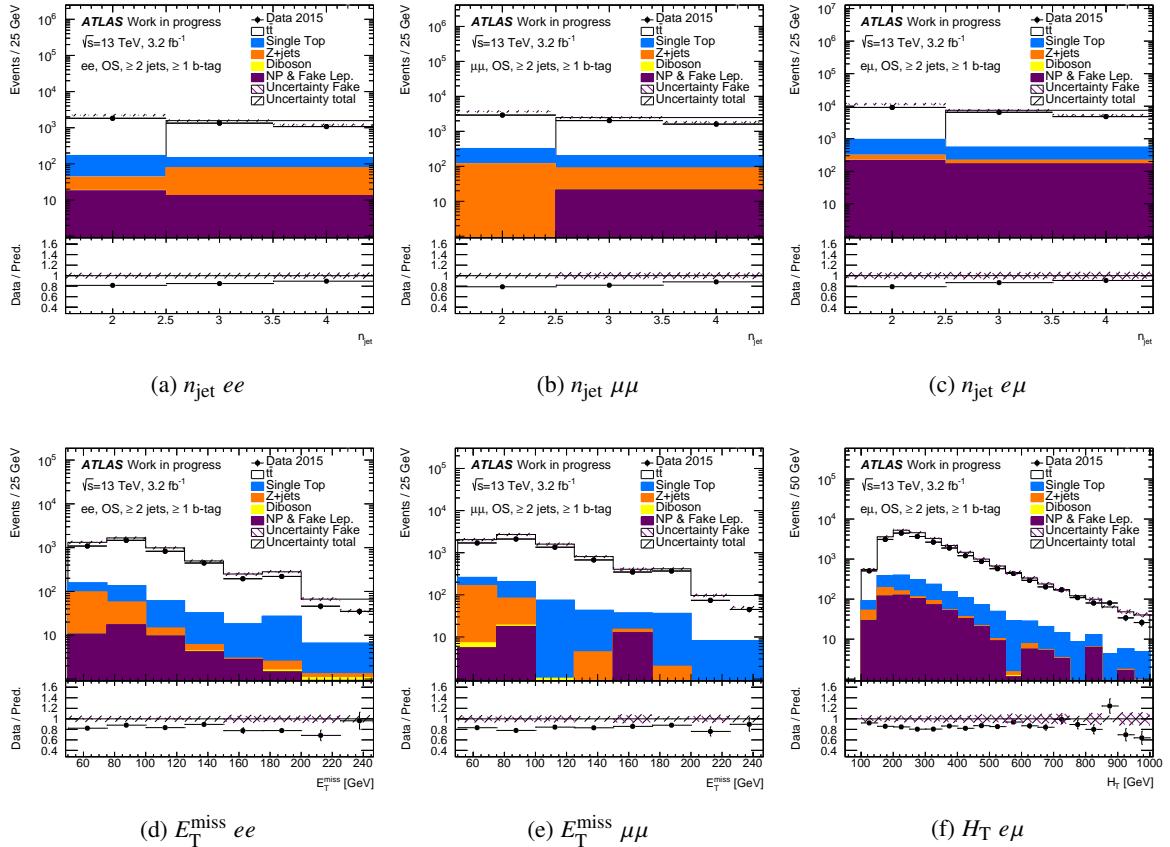


Figure 126: Distributions of (a,b,c) jet multiplicity n_{jet} and (d,e,f) $E_{\text{T}}^{\text{miss}}$ in events with opposite-sign (a,d) ee , (b,e) $\mu\mu$ and (d,f) $e\mu$ pairs with at least two jets, with at least one b -jet, and with the full selection. In the $e\mu$ channel the $E_{\text{T}}^{\text{miss}}$ is replaced by the H_{T} distribution. The data is compared to the real lepton expectation from simulation, showing separately the contributions from $t\bar{t}$, single top, $Z + \text{jets}$ and dibosons normalised to their cross-sections, and non-prompt and fake lepton backgrounds (referred to as ‘NP & Fake Lep.’) estimated with the matrix method. The purple shaded area represents the combination of the statistical and the systematic uncertainties on the matrix method estimate in each bin. The black shaded area represents the total uncertainty including also the uncertainties on the processes predicted by the MC simulation and the one on the luminosity. The lower parts of the figure show the ratios of expectation to data.

Not reviewed, for internal circulation only

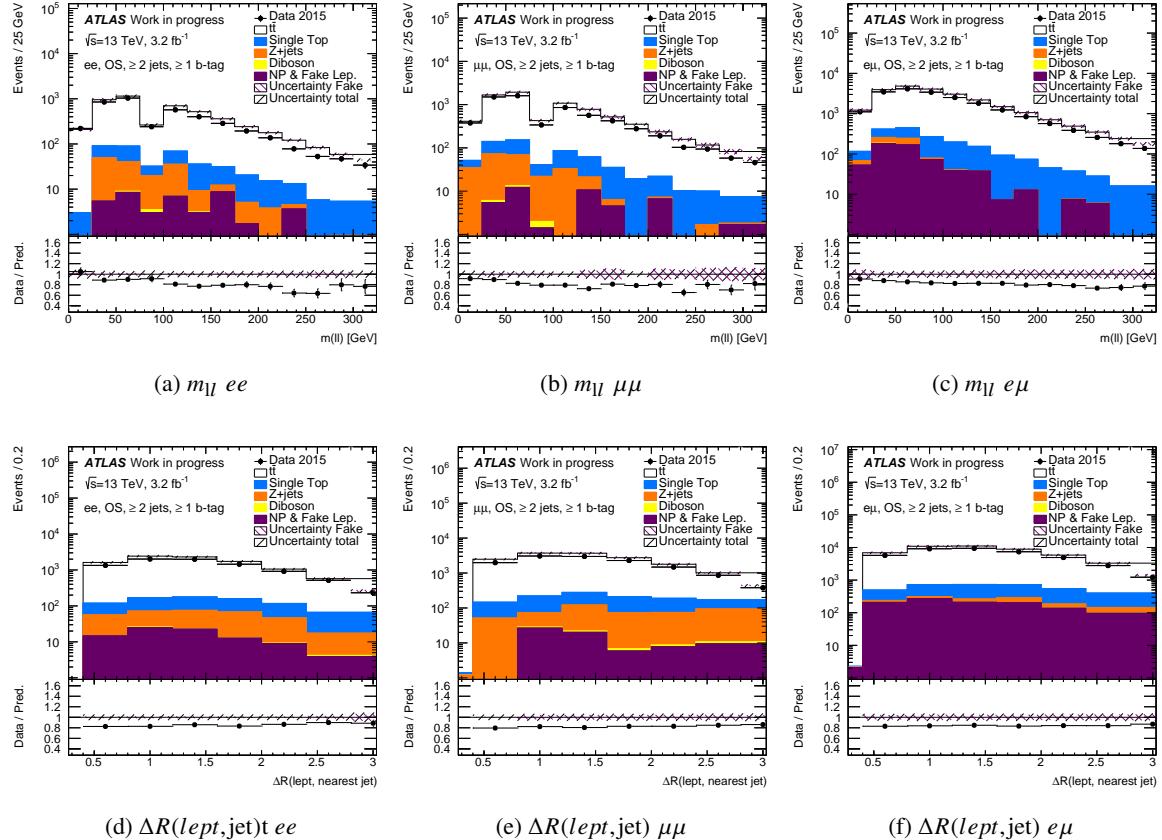


Figure 127: Distributions of (a,b,c) the invariant mass of the two leptons m_{ll} and (d,e,f) the distance $\Delta R(\text{lept}, \text{jet})$ between one of the lepton and its nearest jet, in events with opposite-sign (a,d) ee , (b,e) $\mu\mu$ and (d,f) $e\mu$ pairs with at least two jets, with at least one b -jet, and with the full selection. The data is compared to the real lepton expectation from simulation, showing separately the contributions from $t\bar{t}$, single top, Z +jets and dibosons normalised to their cross-sections, and non-prompt and fake lepton backgrounds (referred to as ‘NP & Fake Lep.’) estimated with the matrix method. The purple shaded area represents the combination of the statistical and the systematic uncertainties on the matrix method estimate in each bin. The black shaded area represents the total uncertainty including also the uncertainties on the processes predicted by the MC simulation and the one on the luminosity. The lower parts of the figure show the ratios of expectation to data.

Not reviewed, for internal circulation only

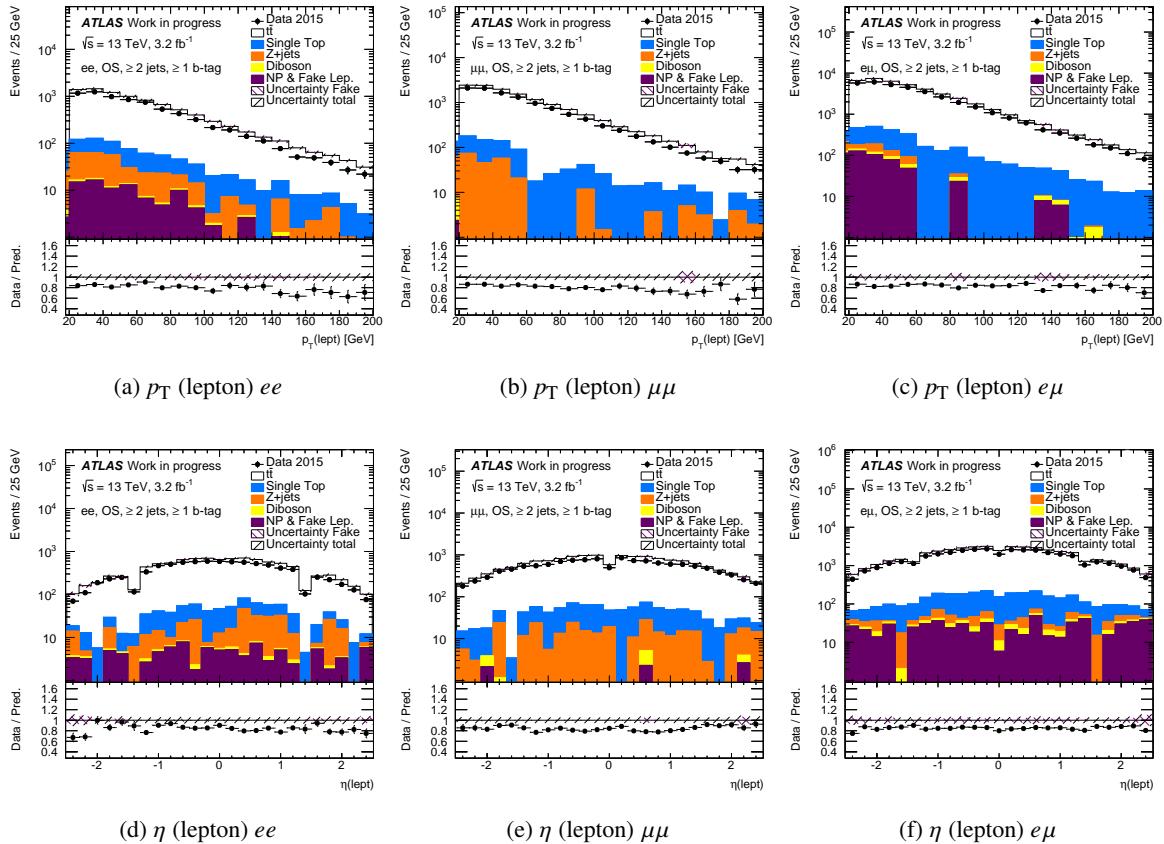


Figure 128: Distributions of (a,b,c) p_T and (d,e,f) η of the lepton, in events with opposite-sign (a,d) ee , (b,e) $\mu\mu$ and (d,f) $e\mu$ pairs with at least two jets, with at least one b -jet, and with the full selection. The data is compared to the real lepton expectation from simulation, showing separately the contributions from $t\bar{t}$, single top, $Z + \text{jets}$ and dibosons normalised to their cross-sections, and non-prompt and fake lepton backgrounds (referred to as ‘NP & Fake Lep.’) estimated with the matrix method. The purple shaded area represents the combination of the statistical and the systematic uncertainties on the matrix method estimate in each bin. The black shaded area represents the total uncertainty including also the uncertainties on the processes predicted by the MC simulation and the one on the luminosity. The lower parts of the figure show the ratios of expectation to data.

Not reviewed, for internal circulation only

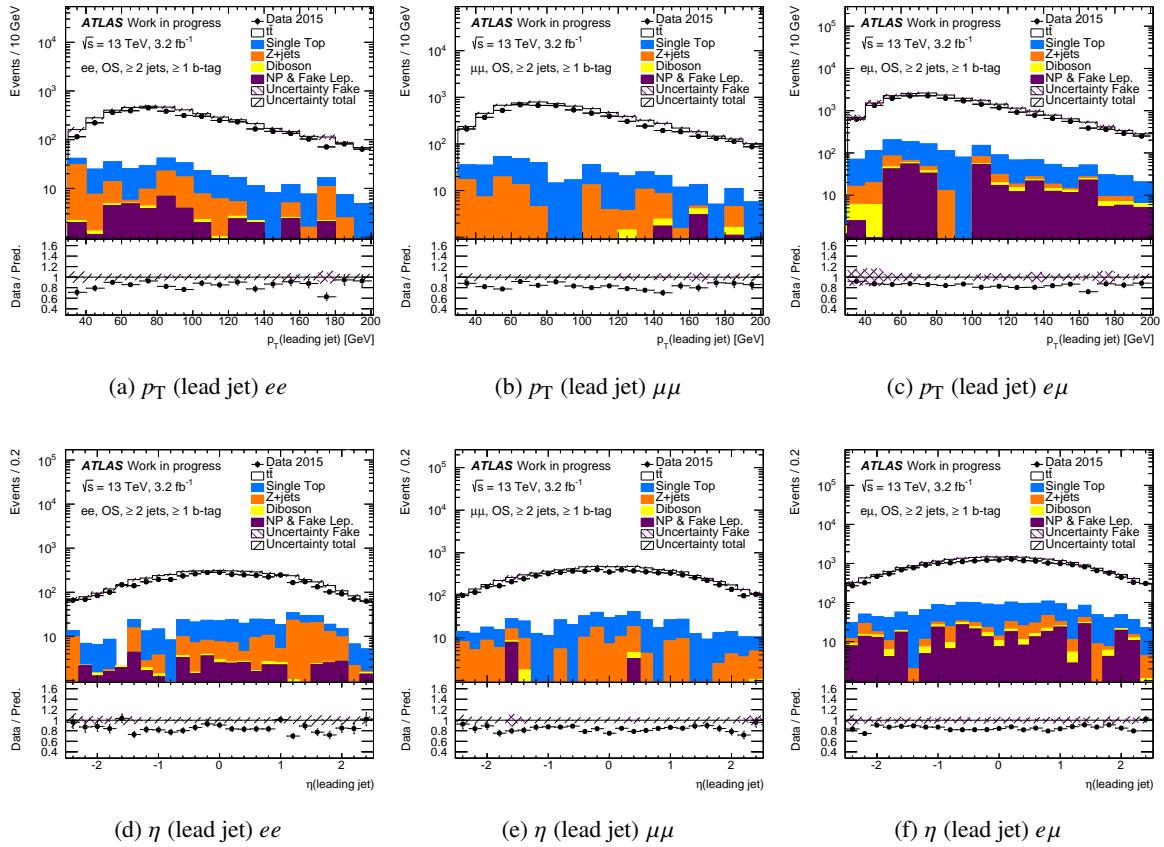


Figure 129: Distributions of (a,b,c) p_T and (d,e,f) η of the leading jet, in events with opposite-sign (a,d) ee , (b,e) $\mu\mu$ and (d,f) $e\mu$ pairs with at least two jets, with at least one b -jet, and with the full selection. The data is compared to the real lepton expectation from simulation, showing separately the contributions from $t\bar{t}$, single top, $Z + \text{jets}$ and dibosons normalised to their cross-sections, and non-prompt and fake lepton backgrounds (referred to as ‘NP & Fake Lep.’) estimated with the matrix method. The purple shaded area represents the combination of the statistical and the systematic uncertainties on the matrix method estimate in each bin. The black shaded area represents the total uncertainty including also the uncertainties on the processes predicted by the MC simulation and the one on the luminosity. The lower parts of the figure show the ratios of expectation to data.

1176 **F.2. Estimates for opposite-sign “not TT” dilepton events**

1177 To further enhance the fake contribution in the opposite-sign sample, at least one of the leptons can be
1178 required to fail the tight selection. On the contrary of the 8 TeV analysis [2] no attempt was done to redo
1179 this study.

1180 **F.3. Estimates for same-sign dilepton events**

1181 Main results concerning the same-sign sample have been shown in section 8.3.

1182 Fig. 130 to 137 show additional control plots in the three channels ee , $\mu\mu$ and $e\mu$ in the pretag and b -
1183 tagged samples, for events with at least one jet. No selection is done on the E_T^{miss} or the H_T . In the
1184 ee and $\mu\mu$ channels additional selection criteria are applied on the invariant mass of the two leptons:
1185 $m_{ll} > 15$ GeV and $|m_{ll} - 91| > 10$ GeV. In the ee channel the main background is due to the charge
1186 mis-identification, whereas it is much lower for the other channels. In Fig. 131 and 135 one can clearly
1187 see the selection done on m_{ll} on its distributions in the ee and $\mu\mu$ channels. In the pretag channel the data
1188 and expectation (i.e the sum of MC and fakes) are in good agreement. It is a bit less true for the case with
1189 events with at least one b -jet where a small deficit can be seen. Finally one can see some bins with lag
1190 uncertainties, in particular in the ee channel. This is due to the low statistics.

Not reviewed, for internal circulation only

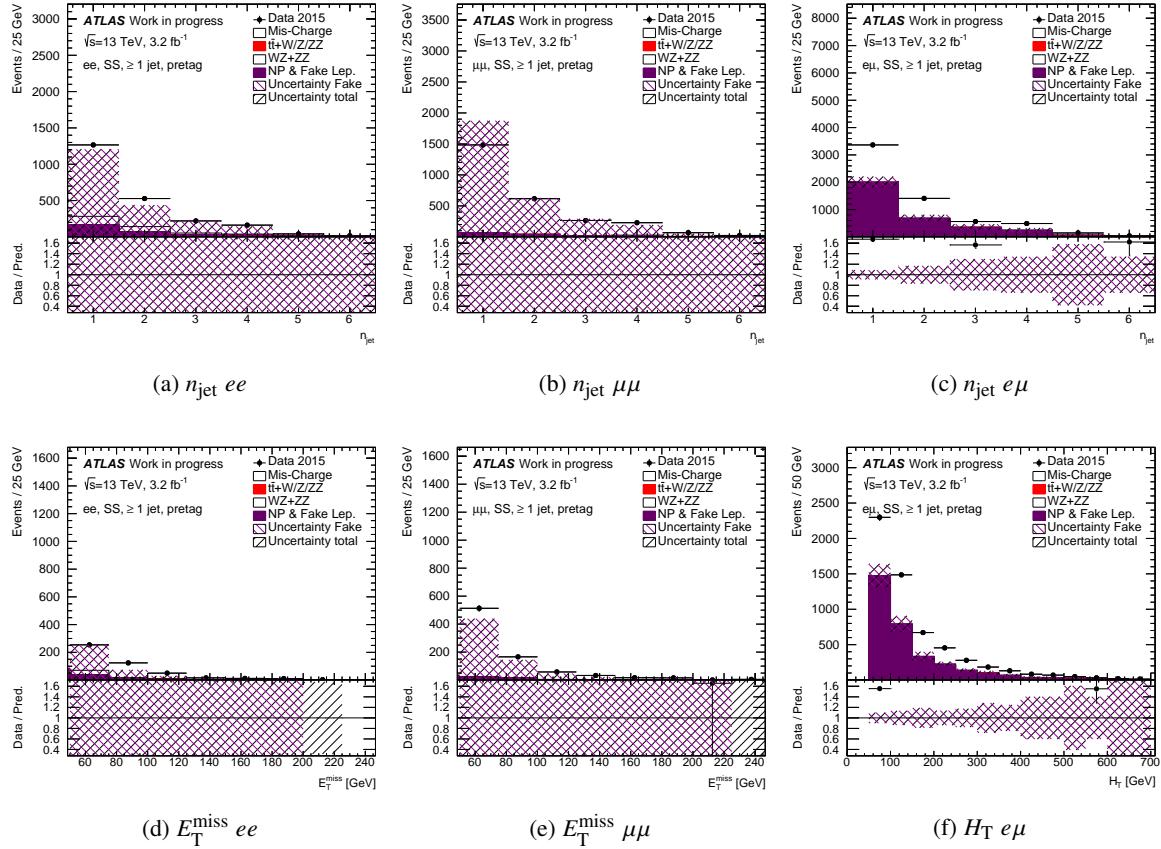


Figure 130: Distributions of (a,b,c) jet multiplicity n_{jet} and (d,e,f) $E_{\text{T}}^{\text{miss}}$ in events with same-sign (a,d) ee , (b,e) $\mu\mu$ and (d,f) $e\mu$ pairs with at least one jet, without requirement on the number of b -jets, and with the full selection except cuts on $E_{\text{T}}^{\text{miss}}$ and $m_{\text{T}}(\text{lept}, E_{\text{T}}^{\text{miss}})$. In the $e\mu$ channel the $E_{\text{T}}^{\text{miss}}$ is replaced by the H_{T} distribution. The data is compared to the real lepton expectation from simulation, showing separately the contributions from charge-misidentification (evaluated from $t\bar{t}$, single top and $Z + \text{jets}$) and dibosons normalised to their cross sections and non-prompt and fake lepton backgrounds (referred to as ‘NP & Fake Lep.’) estimated with the matrix method. The purple shaded area represents the combination of the statistical and the systematic uncertainties on the matrix method estimate in each bin. The black shaded area represents the total uncertainty including also the uncertainties on the processes predicted by the MC simulation and the one on the luminosity. The lower parts of the figure show the ratios of expectation to data.

Not reviewed, for internal circulation only

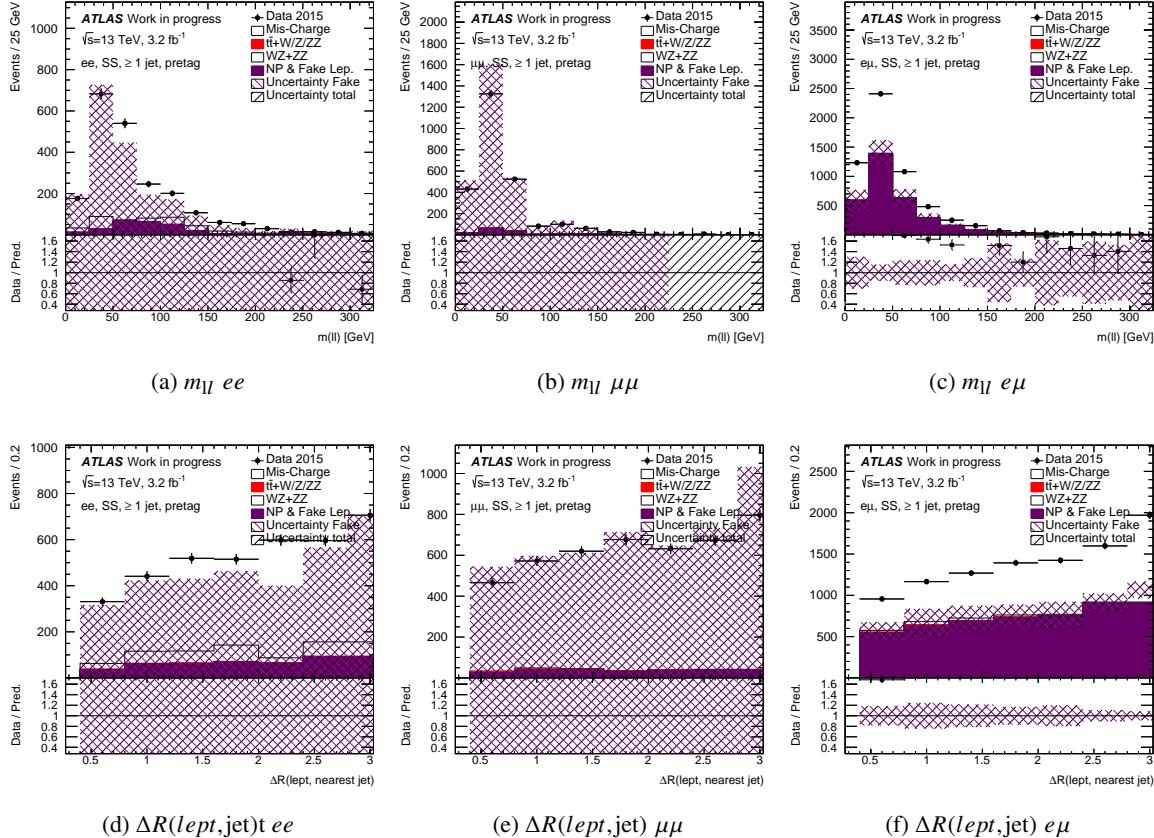


Figure 131: Distributions of (a,b,c) the invariant mass of the two leptons m_{ll} and (d,e,f) the distance $\Delta R(\text{lept}, \text{jet})$ between one of the lepton and its nearest jet, in events with same-sign (a,d) ee , (b,e) $\mu\mu$ and (d,f) $e\mu$ pairs with at least one jet, without requirement on the number of b -jets, and with the full selection except cuts on E_T^{miss} and $m_T(\text{lept}, E_T^{\text{miss}})$. The data is compared to the real lepton expectation from simulation, showing separately the contributions from charge-misidentification (evaluated from $t\bar{t}$, single top and $Z + \text{jets}$) and dibosons normalised to their cross sections and non-prompt and fake lepton backgrounds (referred to as ‘NP & Fake Lep.’) estimated with the matrix method. The purple shaded area represents the combination of the statistical and the systematic uncertainties on the matrix method estimate in each bin. The black shaded area represents the total uncertainty including also the uncertainties on the processes predicted by the MC simulation and the one on the luminosity. The lower parts of the figure show the ratios of expectation to data.

Not reviewed, for internal circulation only

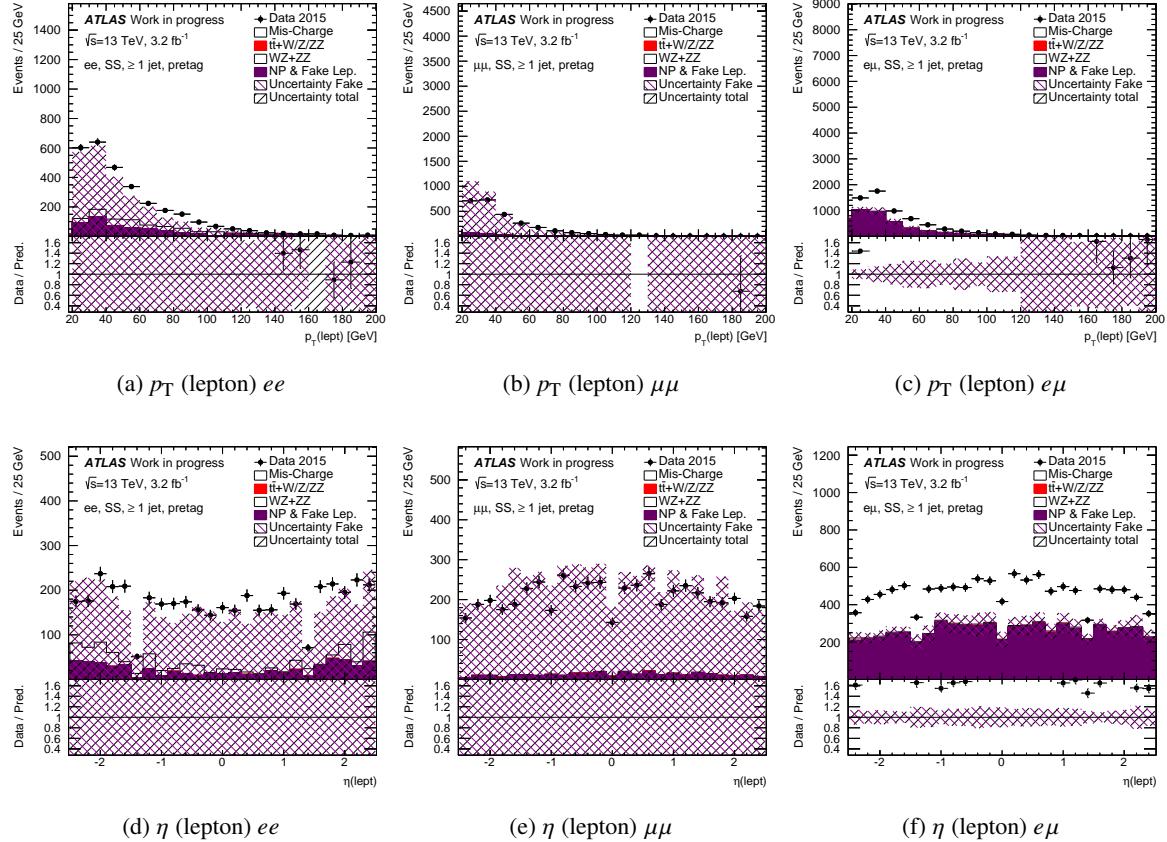


Figure 132: Distributions of (a,b,c) p_T and (d,e,f) η of the lepton, in events with same-sign (a,d) ee , (b,e) $\mu\mu$ and (d,f) $e\mu$ pairs with at least one jet, without requirement on the number of b -jets, and with the full selection except cuts on E_T^{miss} and $m_T(\text{lept}, E_T^{\text{miss}})$. The data is compared to the real lepton expectation from simulation, showing separately the contributions from charge-misidentification (evaluated from $t\bar{t}$, single top and $Z + \text{jets}$) and dibosons normalised to their cross sections and non-prompt and fake lepton backgrounds (referred to as ‘NP & Fake Lep.’) estimated with the matrix method. The purple shaded area represents the combination of the statistical and the systematic uncertainties on the matrix method estimate in each bin. The black shaded area represents the total uncertainty including also the uncertainties on the processes predicted by the MC simulation and the one on the luminosity. The lower parts of the figure show the ratios of expectation to data.

Not reviewed, for internal circulation only

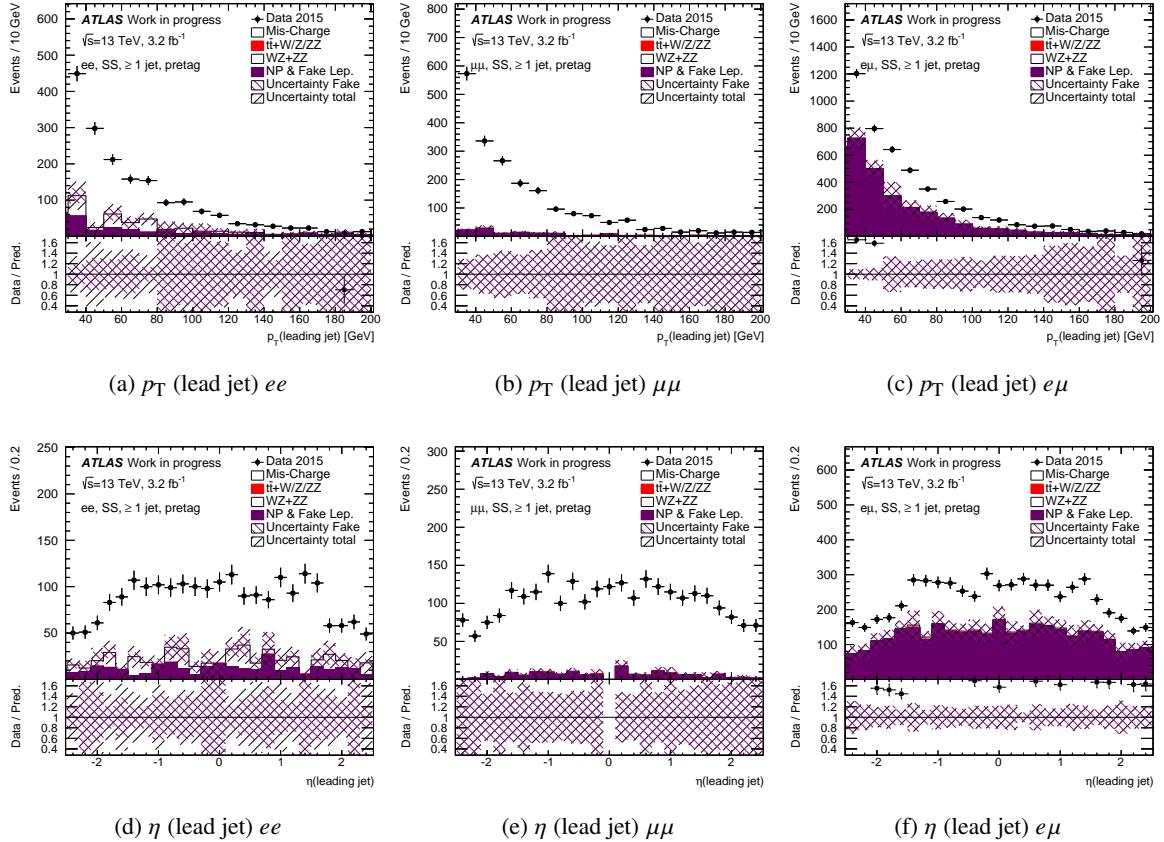


Figure 133: Distributions of (a,b,c) p_T and (d,e,f) η of the leading jet, in events with same-sign (a,d) ee , (b,e) $\mu\mu$ and (d,f) $e\mu$ pairs with at least one jet, without requirement on the number of b -jets, and with the full selection except cuts on E_T^{miss} and $m_T(\text{lept}, E_T^{\text{miss}})$. The data is compared to the real lepton expectation from simulation, showing separately the contributions from charge-misidentification (evaluated from $t\bar{t}$, single top and $Z + \text{jets}$) and dibosons normalised to their cross sections and non-prompt and fake lepton backgrounds (referred to as ‘NP & Fake Lep.’) estimated with the matrix method. The purple shaded area represents the combination of the statistical and the systematic uncertainties on the matrix method estimate in each bin. The black shaded area represents the total uncertainty including also the uncertainties on the processes predicted by the MC simulation and the one on the luminosity. The lower parts of the figure show the ratios of expectation to data.

Not reviewed, for internal circulation only

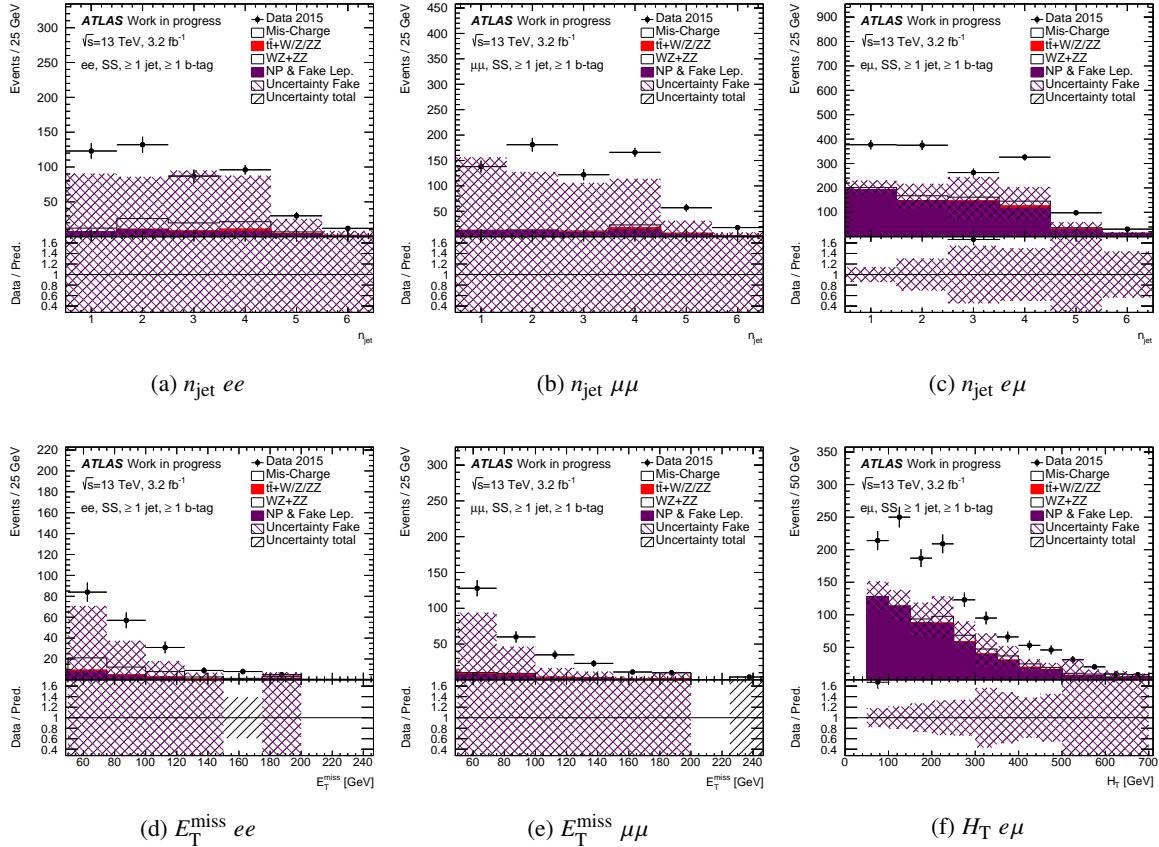


Figure 134: Distributions of (a,b,c) jet multiplicity n_{jet} and (d,e,f) E_T^{miss} in events with same-sign (a,d) ee , (b,e) $\mu\mu$ and (d,f) $e\mu$ pairs with at least one jet, with at least one b -jet, and with the full selection except cuts on E_T^{miss} and $m_T(\text{lept}, E_T^{\text{miss}})$. In the $e\mu$ channel the E_T^{miss} is replaced by the H_T distribution. The data is compared to the real lepton expectation from simulation, showing separately the contributions from charge-misidentification (evaluated from $t\bar{t}$, single top and $Z + \text{jets}$) and dibosons normalised to their cross sections and non-prompt and fake lepton backgrounds (referred to as ‘NP & Fake Lep.’) estimated with the matrix method. The purple shaded area represents the combination of the statistical and the systematic uncertainties on the matrix method estimate in each bin. The black shaded area represents the total uncertainty including also the uncertainties on the processes predicted by the MC simulation and the one on the luminosity. The lower parts of the figure show the ratios of expectation to data.

Not reviewed, for internal circulation only

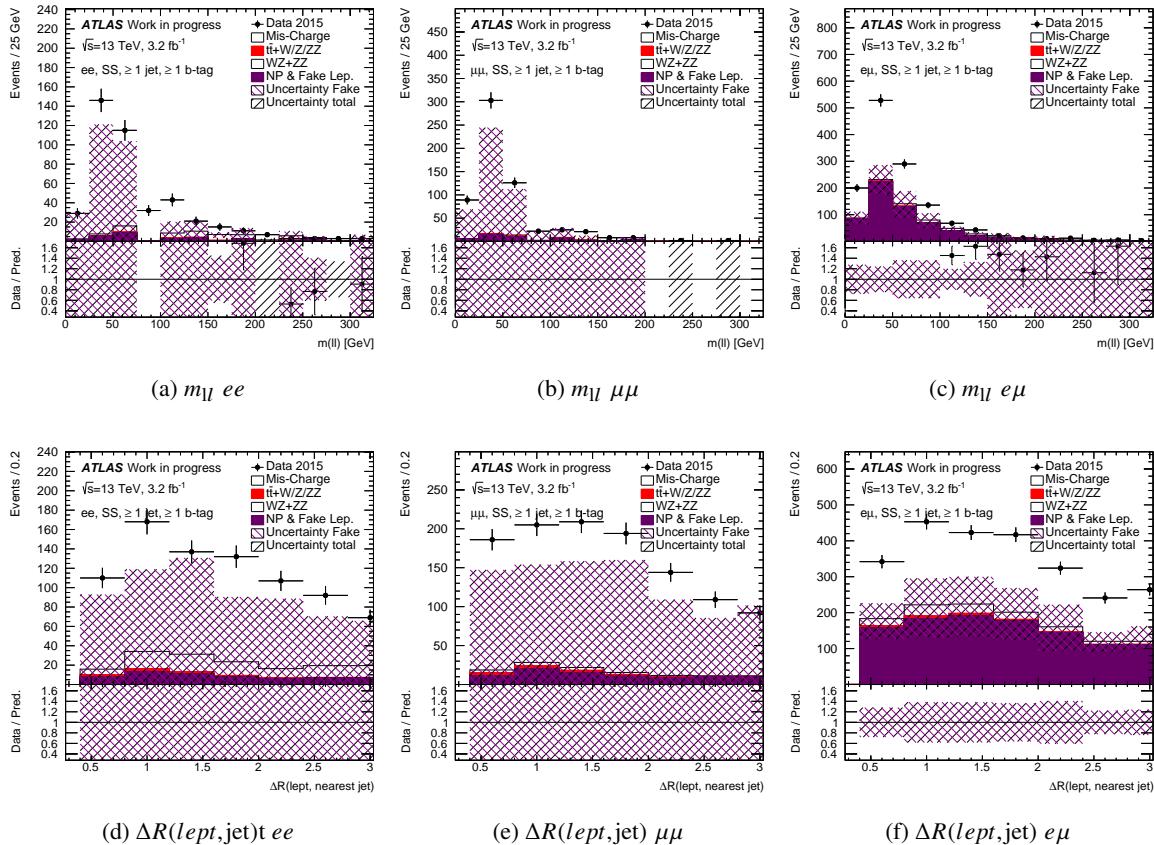


Figure 135: Distributions of (a,b,c) the invariant mass of the two leptons m_{ll} and (d,e,f) the distance $\Delta R(\text{lept}, \text{jet})$ between one of the lepton and its nearest jet, in events with same-sign (a,d) ee , (b,e) $\mu\mu$ and (d,f) $e\mu$ pairs with at least one jet, with at least one b -jet, and with the full selection except cuts on E_T^{miss} and $m_T(\text{lept}, E_T^{\text{miss}})$. The data is compared to the real lepton expectation from simulation, showing separately the contributions from charge-misidentification (evaluated from $t\bar{t}$, single top and $Z + \text{jets}$) and dibosons normalised to their cross sections and non-prompt and fake lepton backgrounds (referred to as ‘NP & Fake Lep.’) estimated with the matrix method. The purple shaded area represents the combination of the statistical and the systematic uncertainties on the matrix method estimate in each bin. The black shaded area represents the total uncertainty including also the uncertainties on the processes predicted by the MC simulation and the one on the luminosity. The lower parts of the figure show the ratios of expectation to data.

Not reviewed, for internal circulation only

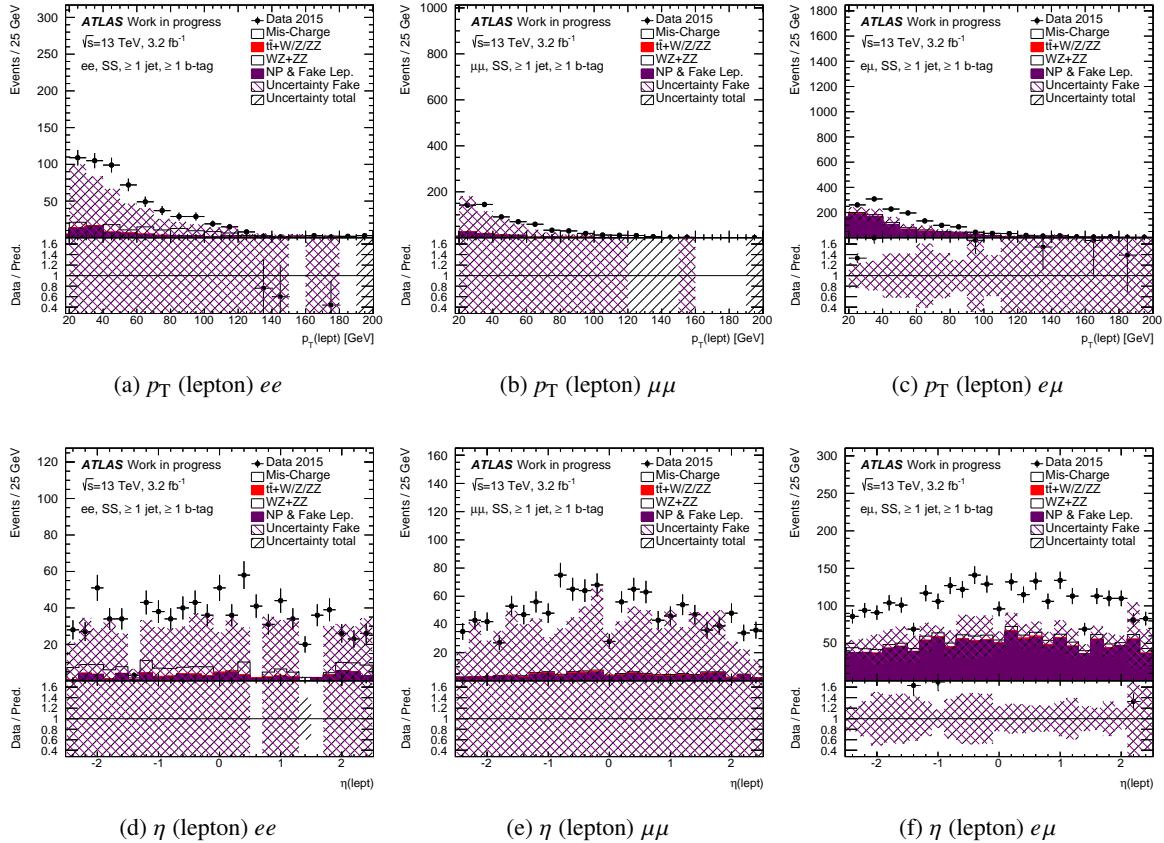


Figure 136: Distributions of (a,b,c) p_T and (d,e,f) η of the lepton, in events with same-sign (a,d) ee , (b,e) $\mu\mu$ and (d,f) $e\mu$ pairs with at least one jet, with at least one b -jet, and with the full selection except cuts on E_T^{miss} and $m_T(\text{lept}, E_T^{\text{miss}})$. The data is compared to the real lepton expectation from simulation, showing separately the contributions from charge-misidentification (evaluated from $t\bar{t}$, single top and $Z + \text{jets}$) and dibosons normalised to their cross sections and non-prompt and fake lepton backgrounds (referred to as ‘NP & Fake Lep.’) estimated with the matrix method. The purple shaded area represents the combination of the statistical and the systematic uncertainties on the matrix method estimate in each bin. The black shaded area represents the total uncertainty including also the uncertainties on the processes predicted by the MC simulation and the one on the luminosity. The lower parts of the figure show the ratios of expectation to data.

Not reviewed, for internal circulation only

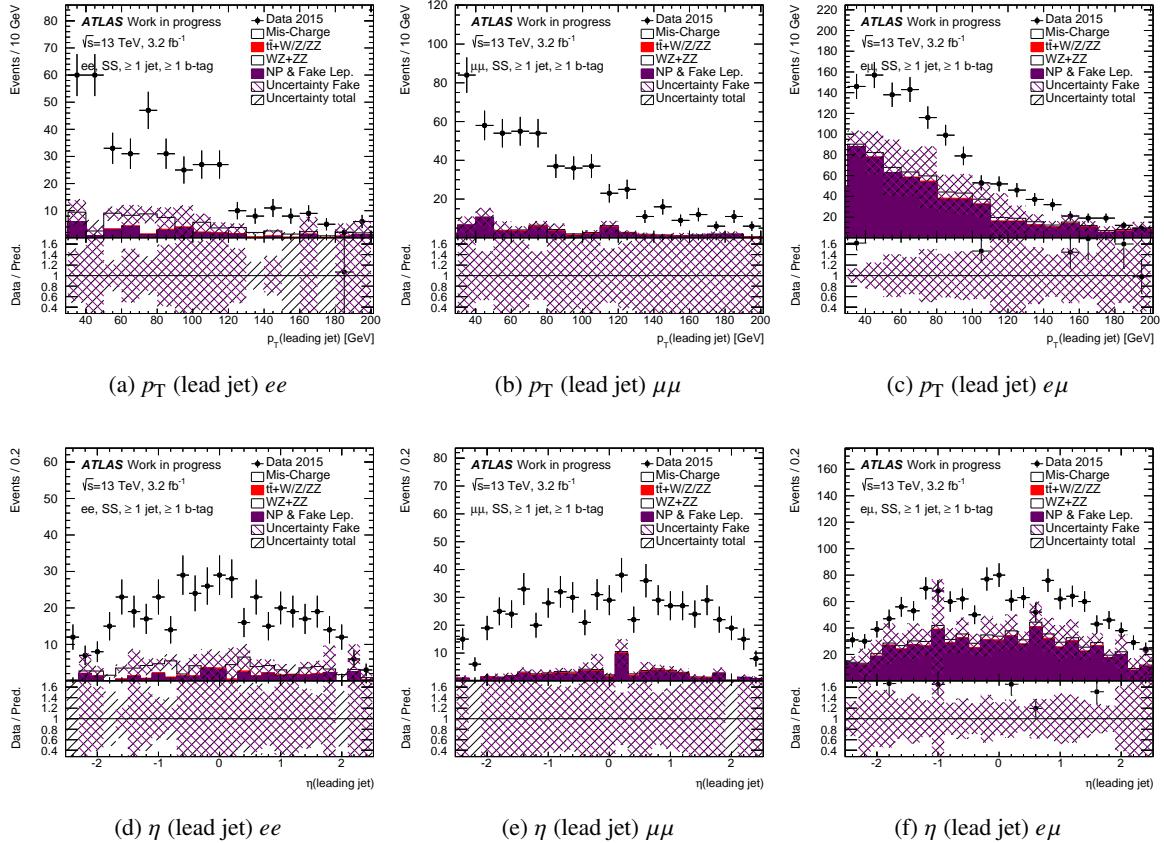


Figure 137: Distributions of (a,b,c) p_T and (d,e,f) η of the leading jet, in events with same-sign (a,d) ee , (b,e) $\mu\mu$ and (d,f) $e\mu$ pairs with at least one jet, with at least one b -jet, and with the full selection except cuts on E_T^{miss} and $m_T(\text{lept}, E_T^{\text{miss}})$. The data is compared to the real lepton expectation from simulation, showing separately the contributions from charge-misidentification (evaluated from $t\bar{t}$, single top and $Z + \text{jets}$) and dibosons normalised to their cross sections and non-prompt and fake lepton backgrounds (referred to as ‘NP & Fake Lep.’) estimated with the matrix method. The purple shaded area represents the combination of the statistical and the systematic uncertainties on the matrix method estimate in each bin. The black shaded area represents the total uncertainty including also the uncertainties on the processes predicted by the MC simulation and the one on the luminosity. The lower parts of the figure show the ratios of expectation to data.

1191 F.4. Estimates for same-sign “not TT” dilepton events

1192 To further enhance the fake contribution in the same-sign sample, at least one of the leptons can be
1193 required to fail the tight selection. On the contrary of the 8 TeV analysis [2] no attempt was done to redo
1194 this study.

1195 References

- [1] ATLAS Collaboration, *Estimation of non-prompt and fake lepton backgrounds in final states with top quarks produced in proton–proton collisions at $\sqrt{s} = 8$ TeV with the ATLAS Detector*, ATLAS-CONF-2014-058, 2014, URL: <http://cdsweb.cern.ch/record/1951336>.
- [2] Becker, K. et al.,
Estimation of Fake Lepton Background for Top Analyses Using the $\sqrt{s} = 8$ TeV Dataset, ATL-COM-PHYS-2013-1100, 2013, URL: <https://cds.cern.ch/record/1571043>.
- [3] ATLAS Collaboration, *The ATLAS Experiment at the CERN Large Hadron Collider*, JINST **3** (2008) S08003.
- [4] ‘The 2015 data selection used the good run list:
`data15_13TeV.periodAllYear_DetStatus - v73 - pro19 - 08DQDefects - 00 - 01 - 02_PHYS_Standard`
- [5] ATLAS Collaboration, *Improved luminosity determination in pp collisions at $\sqrt{s} = 7$ TeV using the ATLAS detector at the LHC*, Eur. Phys. J. C **73** (2013) 2518, arXiv: <1302.4393> [hep-ex].
- [6] ATLAS Collaboration, *The ATLAS Simulation Infrastructure*, Eur. Phys. J. C **70** (2010) 823, arXiv: <1005.4568> [hep-ex].
- [7] S. Agostinelli et al., *GEANT4 - A Simulation Toolkit*, Nucl. Instr. and Meth. **A506** (2003) 250.
- [8] ATLAS Collaboration,
The simulation principle and performance of the ATLAS fast calorimeter simulation FastCaloSim, ATL-PHYS-PUB-2010-013, 2010, URL: <http://cds.cern.ch/record/1300517>.
- [9] T. Sjöstrand, S. Mrenna, P. Skands, *A Brief Introduction to PYTHIA 8.1*, Comput. Phys. Commun. **178** (2008) 852, arXiv: <0710.3820> [hep-ph].
- [10] P. Nason, *A New method for combining NLO QCD with shower Monte Carlo algorithms*, JHEP **0411** (2004) 040, arXiv: <hep-ph/0409146> [hep-ph].
- [11] S. Frixione, P. Nason and C. Oleari,
Matching NLO QCD computations with parton shower simulations: the POWHEG method, JHEP **11** (2007) 070, arXiv: <0709.2092> [hep-ph].
- [12] S. Alioli, P. Nason, C. Oleari, and E. Re, *A general framework for implementing NLO calculations in shower Monte Carlo programs: the POWHEG BOX*, JHEP **1006** (2010) 043, arXiv: <1002.2581> [hep-ph].
- [13] H-L. Lai et al., *New parton distributions for collider physics*, Phys. Rev. **D82** (2010) 074024, arXiv: <1007.2241> [hep-ph].
- [14] T. Sjöstrand , S. Mrenna, and P. Skands, *Pythia 6.4 Physics and Manual*, JHEP **05** (2006) 026, arXiv: <hep-ph/0603175v2> [hep-ph].
- [15] P.Z. Skands, *Tuning Monte Carlo Generators: The Perugia Tunes*, Phys. Rev. D **82** (2010) 074018, arXiv: <1005.3457> [hep-ph].
- [16] J. Pumplin et al.,
New generation of parton distributions with uncertainties from global QCD analysis, JHEP **012** (2002) 0207, arXiv: <hep-ph/0201195> [hep-ph].
- [17] D. J. Lange, *The EvtGen particle decay simulation package*, Nucl. Instrum. Meth. A **462** (2001) 152.

- [18] ATLAS Collaboration, *Comparison of Monte Carlo generator predictions to ATLAS measurements of top pair production at 7 TeV*, ATL-PHYS-PUB-2015-002, 2015, URL: <http://cdsweb.cern.ch/record/1981319>.
- [19] K.A. Olive et al. (Particle Data Group), *The review of particle physics*, Chin. Phys. C **38** (2014) 090001, URL: <http://pdg.lbl.gov/>.
- [20] T. Gleisberg et al., *Event generation with SHERPA 1.1*, JHEP **02** (2009) 007, arXiv: [0811.4622 \[hep-ph\]](https://arxiv.org/abs/0811.4622).
- [21] T. Gleisberg and S. Höche, *Comix, a new matrix element generator*, JHEP **39** (2008) 0812, arXiv: [0808.3674 \[hep-ph\]](https://arxiv.org/abs/0808.3674).
- [22] F. Cascioli, P. Maierhofer, and S. Pozzorini, *Scattering Amplitudes with Open Loops*, Phys. Rev. Lett. **108** (2012) 111601, arXiv: [1111.5206 \[hep-ph\]](https://arxiv.org/abs/1111.5206).
- [23] S. Schumann and F. Krauss, *A Parton shower algorithm based on Catani-Seymour dipole factorisation*, JHEP **038** (2012) 0803, arXiv: [0709.1027 \[hep-ph\]](https://arxiv.org/abs/0709.1027).
- [24] S. Höche, F. Krauss, M. Schönher, and F. Siegert, *A Parton shower algorithm based on Catani-Seymour dipole factorisation*, JHEP **027** (2013) 1304, arXiv: [1207.5030 \[hep-ph\]](https://arxiv.org/abs/1207.5030).
- [25] S. Frixione, Frixione, E. Laenen, P. Motylinski, B. R. Webber, and C. D. White, *Single-top hadroproduction in association with a W boson*, JHEP **07** (2008) 029, arXiv: [0805.3067 \[hep-ph\]](https://arxiv.org/abs/0805.3067).
- [26] ATLAS Collaboration, *Electron efficiency measurements with the ATLAS detector using the 2012 LHC proton–proton collision data*, ATLAS-CONF-2014-032, 2014, URL: <http://cdsweb.cern.ch/record/1706245>.
- [27] ATLAS Collaboration, *Measurement of the muon reconstruction performance of the ATLAS detector using 2011 and 2012 LHC proton–proton collision data*, Eur. Phys. J. C **74** (2014) 3130, arXiv: [1407.3935 \[hep-ex\]](https://arxiv.org/abs/1407.3935).
- [28] M. Cacciari and G. P. Salam, *Dispelling the N^3 myth for the kt jet-finder*, Physics Letters B **641** (2006) 57, ISSN: 0370-2693, arXiv: [hep-ph/0512210 \[hep-ph\]](https://arxiv.org/abs/hep-ph/0512210).
- [29] M. Cacciari, G. Salam, G. Soyez, *The anti- k_t jet clustering algorithm*, JHEP **04** (2008) 063, arXiv: [0802.1189 \[hep-ph\]](https://arxiv.org/abs/0802.1189).
- [30] W. Lampl et al., *Calorimeter Clustering Algorithms: Description and Performance*, ATL-LARG-PUB-2008-002, 2008, URL: <http://cds.cern.ch/record/1099735>.
- [31] M. Cacciari, G. P. Salam, and G. Soyez, *The Catchment Area of Jets*, JHEP **04** (2008) 05, arXiv: [0802.1188 \[hep-ph\]](https://arxiv.org/abs/0802.1188).
- [32] ATLAS Collaboration, *Jet global sequential corrections with the ATLAS detector in proton–proton collisions at $\sqrt{s} = 8$ TeV*, ATLAS-CONF-2015-002, 2015, URL: <http://cdsweb.cern.ch/record/2001682>.
- [33] ATLAS Collaboration, *Jet energy measurement with the ATLAS detector in proton–proton collisions at $\sqrt{s} = 7$ TeV*, Eur. Phys. J. C **73** (2013) 2304, arXiv: [1112.6426 \[hep-ex\]](https://arxiv.org/abs/1112.6426).

- 1275 [34] ATLAS Collaboration, *Monte Carlo Calibration and Combination of In-situ Measurements of Jet*
1276 *Energy Scale, Jet Energy Resolution and Jet Mass in ATLAS*, ATLAS-CONF-2015-037, 2015,
1277 URL: <http://cdsweb.cern.ch/record/2044941>.
- 1278 [35] ATLAS Collaboration, *Jet Calibration and Systematic Uncertainties for Jets Reconstructed in the*
1279 *ATLAS Detector at $\sqrt{s} = 13 \text{ TeV}$* , ATL-PHYS-PUB-2015-015, 2015,
1280 URL: <https://cds.cern.ch/record/2037613>.
- 1281 [36] ATLAS Collaboration, *Tagging and suppression of pileup jets with the ATLAS detector*,
1282 ATLAS-CONF-2014-018, 2014, URL: <http://cdsweb.cern.ch/record/1700870>.
- 1283 [37] ATLAS Collaboration, *Expected performance of the ATLAS b-tagging algorithms in Run-2*,
1284 ATL-PHYS-PUB-2015-022, 2015, URL: <http://cdsweb.cern.ch/record/2037697>.
- 1285 [38] ATLAS Collaboration,
1286 *Commissioning of the ATLAS high performance b-tagging algorithms in the 7 TeV collision data*,
1287 ATLAS-CONF-2011-102, 2011, URL: <http://cdsweb.cern.ch/record/1369219>.
- 1288 [39] ATLAS Collaboration, *Measurement of the b-tag Efficiency in a Sample of Jets Containing Muons*
1289 *with 5 fb^{-1} of data from the ATLAS detector*, ATLAS-CONF-2012-043, 2012,
1290 URL: <http://cdsweb.cern.ch/record/1435197>.
- 1291 [40] ATLAS Collaboration, *Performance of missing transverse momentum reconstruction with the*
1292 *ATLAS detector in the first proton–proton collisions at $\sqrt{s} = 13 \text{ TeV}$* , ATL-PHYS-PUB-2015-027,
1293 2015, URL: <http://cdsweb.cern.ch/record/2037904>.