

2013/05/01

Head Id:

Archive Id: 145226

Archive Date: 2012/02/04

Archive Tag: trunk

Search for supersymmetry with the α_T variable in the 7 TeV dataset of 2011

R. Bainbridge³, O. Buchmüller³, D. Burton³, Y. Eshaq⁶, H. Flächer², E. Laird^{5,4}, C. Lucas², J. Marrouche³, B. Mathias³, T. Medvedeva⁵, Z. Meng², S. Rogerson³, and P. Sphicas¹

¹CERN, Geneva, Switzerland

²University of Bristol, UK

³Imperial College, London, UK

⁴Brown University, RI, US

⁵Princeton University, NJ, US

⁶University of Rochester, NY, US

Abstract

A search for events with jets and missing transverse energy is performed in a data sample of pp collisions collected at $\sqrt{s} = 7$ TeV by the CMS experiment at the LHC. The analyzed data sample corresponds to an integrated luminosity of 4.98 fb^{-1} . In this search, a kinematic variable, α_T , is used as the main discriminator between events with genuine and misreconstructed missing transverse energy. No excess of events over the standard model expectation is found. Exclusion limits in the parameter space of the constrained minimal supersymmetric extension of the standard model are set. In this model, squark masses below 1.3 TeV are excluded at 95% CL. Gluino masses below 1.3 TeV are also ruled out at 95% CL for values of the universal scalar mass parameter below 600 GeV. Exclusion limits are also set in various simplified models.

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

| | |
|--------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------|
| PDFAuthor: | R. Bainbridge, O. Buchmuller, D. Burton, Y. Eshaq, H. Flaecher, E. Laird, C. Lucas, J. Marrouche, B. Mathias, T. Medvedeva, Z. Meng, S. Rogerson, P. Sphicas |
| PDFTitle: | Search for supersymmetry with the AlphaT variable on the 7 TeV dataset of 2011 |
| PDFSubject: | CMS, physics, jets, missing energy, supersymmetry, SUSY, AlphaT |
| PDFKeywords: | CMS, physics, jets, missing energy, supersymmetry, SUSY, AlphaT |

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

1 Contents

| | | | |
|----|-----|------------------------------------------------------------------------|----|
| 1 | 1 | Introduction | 3 |
| 2 | 2 | Event reconstruction and selection | 4 |
| 3 | 2.1 | Hadronic pre-selection | 4 |
| 4 | 2.2 | Tagging jets from b-quarks | 4 |
| 5 | 2.3 | Suppressing the multi-jet background with the α_T variable | 5 |
| 6 | 2.4 | The hadronic signal region | 6 |
| 7 | 2.5 | Breakdown of SM backgrounds in the hadronic signal region | 6 |
| 8 | 2.6 | Data–MC control distributions | 7 |
| 9 | 3 | Triggers | 9 |
| 10 | 4 | Data and Monte Carlo Samples | 10 |
| 11 | 5 | Background estimation | 11 |
| 12 | 5.1 | Overview of the method | 11 |
| 13 | 5.2 | Definition of the control samples | 13 |
| 14 | 5.3 | Distributions, yields and translation factors from the control samples | 14 |
| 15 | 5.4 | Larger muon samples without an α_T requirement | 15 |
| 16 | 5.5 | Estimating the background from multi-jet events | 17 |
| 17 | 6 | Systematic uncertainties on the background estimates | 20 |
| 18 | 6.1 | Using closure tests to identify possible biases | 20 |
| 19 | 6.2 | MC modelling of the α_T distribution | 20 |
| 20 | 6.3 | Systematic uncertainties on translation factors | 21 |
| 21 | 6.4 | Motivating a H_T -dependent systematic for the translation factors | 22 |
| 22 | 6.5 | Achieving more accurate n_b estimates specifically for $n_b = 3$ | 22 |
| 23 | 6.6 | Estimating the impact of the b-tagging and mis-tagging scale factors | 24 |
| 24 | 6.7 | Procedure | 25 |
| 25 | 6.8 | Procedure for signal scans | 27 |
| 26 | 7 | Signal efficiency and uncertainty | 28 |
| 27 | 7.1 | SMS-dependent fastsim uncertainty | 33 |
| 28 | 8 | Likelihood model | 33 |
| 29 | 8.1 | Hadronic sample | 33 |
| 30 | 8.2 | H_T evolution models | 34 |
| 31 | 8.3 | Electroweak control samples | 35 |
| 32 | 8.4 | Contributions from signal | 37 |
| 33 | 8.5 | Total likelihood | 37 |
| 34 | 9 | Final Results | 37 |
| 35 | 9.1 | Standard Model | 37 |
| 36 | 9.2 | SUSY | 42 |
| 37 | 10 | Summary | 46 |
| 38 | A | Distributions for the hadronic signal region | 49 |
| 39 | B | Additional information on triggers | 51 |
| 40 | C | Additional information for the control samples | 55 |
| 41 | C.1 | Distributions of b-tag multiplicity | 55 |
| 42 | C.2 | Distributions and predictions from the $\mu + \text{jets}$ sample | 56 |

| | | | |
|----|-----|----------------------------------------------------------------------------------------------------|-----|
| 44 | C.3 | Distributions and predictions from the $\mu\mu + \text{jets}$ sample | 60 |
| 45 | C.4 | Distributions and predictions from for the $\gamma + \text{jets}$ sample | 64 |
| 46 | C.5 | Illustrative total predictions for SM backgrounds | 67 |
| 47 | D | Additional information on QCD multi-jet background estimation | 71 |
| 48 | E | Closure tests and systematic uncertainties | 72 |
| 49 | E.1 | Defining muon samples without an α_T requirement | 72 |
| 50 | E.2 | Closure tests for inclusive analysis | 73 |
| 51 | E.3 | Closure tests for b-tag analysis | 74 |
| 52 | E.4 | Closure tests concerning pile-up | 76 |
| 53 | F | Experimental systematic uncertainties on signal efficiency times acceptance . . . | 77 |
| 54 | F.1 | Systematics due to jet energy scale uncertainties | 77 |
| 55 | F.2 | Systematics due to the MHT/MET cut | 82 |
| 56 | F.3 | Systematics due to the dead ECAL cut | 86 |
| 57 | F.4 | Systematics due to the lepton and photon vetoes | 90 |
| 58 | G | Nuisance parameters in the likelihood model | 94 |
| 59 | H | Maximum likelihood yields | 96 |
| 60 | I | Validating the asymptotic CL_s method | 98 |
| 61 | J | Two-dimensional hadronic signal region | 99 |
| 62 | K | HT distributions for example signal models | 101 |
| 63 | K.1 | T1 topology: $m_{\tilde{g}} = 800 \text{ GeV}$, $m_{\text{LSP}} = 200 \text{ GeV}$ | 101 |
| 64 | K.2 | T2tt topology: $m_{\tilde{q}} = 400 \text{ GeV}$, $m_{\text{LSP}} = 50 \text{ GeV}$ | 102 |
| 65 | K.3 | T2bb topology: $m_{\tilde{q}} = 500 \text{ GeV}$, $m_{\text{LSP}} = 150 \text{ GeV}$ | 103 |
| 66 | K.4 | T1bbbb topology: $m_{\tilde{g}} = 1000 \text{ GeV}$, $m_{\text{LSP}} = 500 \text{ GeV}$ | 104 |
| 67 | L | Comparison of asymptotic CL_s vs $CL_s + \text{toys}$ for T1bbbb | 105 |
| 68 | M | Signal efficiency | 106 |
| 69 | M.1 | CMSSM | 106 |
| 70 | M.2 | T1 | 107 |
| 71 | M.3 | T2 | 108 |
| 72 | M.4 | T2tt | 109 |
| 73 | M.5 | T2bb | 110 |
| 74 | M.6 | T1tttt | 111 |
| 75 | M.7 | T1bbbb | 112 |
| 76 | M.8 | Signal contamination for T1tttt | 113 |

77 1 Introduction

78 In this note we present an update of the search for a missing energy signature in dijet and multi-
 79 jet events using the kinematic variable α_T , as first introduced in Refs. [1–3] and described in
 80 Sec. 2. These results are based on a data sample of pp collisions collected in 2011 at a centre-of-
 81 mass energy of 7 TeV, which corresponds to an integrated luminosity of 4.98 fb^{-1} . The search
 82 focuses on event topologies in which new heavy particles are pair-produced, each of which
 83 then decays to a weakly interacting massive particle (WIMP) that remains undetected, thus
 84 leading to a missing energy signature. In the case of SUSY, the candidate heavy particles are
 85 squarks and gluinos and the WIMP candidate is the lightest (and stable) neutralino χ_1^0 . Thus,
 86 this search requires at least two high- p_T jets and significant \cancel{E}_T in the final state. The results
 87 presented below are interpreted in the context of SUSY, although they are also applicable to
 88 other New Physics scenarios that are characterised by a missing transverse energy signature,
 89 such as Extra Dimensions and Little Higgs models.

90 The analysis follows closely Ref. [4], which in turn is based on Ref. [5] and two simulation-
 91 based studies [2, 3]. The main differences with respect to Ref. [4] are twofold¹: the search
 92 has been adapted to improve the sensitivity to final-state signatures rich in heavy quarks; and
 93 a new data control sample for background estimation is added. As before, a search for an
 94 excess of events in data over the Standard Model expectation is performed in multi-jet final-
 95 states with significant \cancel{E}_T . The dominant background is multi-jet production, a manifestation
 96 of quantum chromodynamics (QCD), which is suppressed by the α_T variable to a negligible
 97 level. To estimate the remaining significant backgrounds, we make use of three data control
 98 samples: a $\mu + \text{jets}$ sample to determine the background from $W + \text{jets}$, $t\bar{t}$ and single top events;
 99 a $\gamma + \text{jets}$ sample to determine the irreducible background from $Z \rightarrow \nu\bar{\nu} + \text{jets}$ events; and a $\mu\mu$
 100 + jets sample that is also used to determine the $Z \rightarrow \nu\bar{\nu} + \text{jets}$ background.

101 One of the primary motivations for the existence of SUSY involves a solution to stabilize di-
 102 vergent corrections to the Higgs boson mass. These corrections are proportional to the masses
 103 of the particles that couple to it. Hence, the largest correction will come from the top quark.
 104 SUSY must be a broken symmetry, but “natural SUSY” arguments dictate the level at which
 105 this breaking takes place, thereby setting a “natural” limit on the difference in mass between
 106 the top quark and a potential top squark. Thus, if a low-mass Higgs boson exists, pairs of top
 107 squarks could be copiously produced at the LHC, which would then subsequently decay via
 108 top quarks to bottom quarks.

109 Hence, “natural SUSY” scenarios would be characterised by final-state signatures rich in bot-
 110 tom quarks. A search relying on methods to identify (b-tag) jets originating from bottom quarks
 111 (b jets) would significantly improve the sensitivity to this class of signature. This is achieved
 112 by categorising the events found in the signal region according to the number of b-tags re-
 113 constructed in each event. Events are similarly categorised in the data control samples, and the
 114 information from all samples and b-tag categories is then used simultaneously in the likelihood
 115 model in order used to interpret the results in a coherent and powerful way.

116 An important aspect of this analysis is that, independently of the number of b-tags recon-
 117 structed in an event, the same conditions are used to record the events samples and the same
 118 methods are used to estimate the various SM backgrounds their associated systematic uncer-
 119 tainties.

¹A further change has also been investigated and was pre-approved by the SUSY group on 1st March 2012: the hadronic signal region is enlarged to cover more signal phase space, by extending the region to lower values of the α_T variable. This approach is explained further in Appendix J, although it is not used by the analysis presented in this note.

To interpret the results from the analysis, a simplified and practical model of SUSY-breaking, the constrained minimal supersymmetric extension of the standard model (CMSSM) [6, 7] is used. The CMSSM is described by five parameters: the universal scalar and gaugino mass parameters (m_0 and $m_{1/2}$, respectively), the universal trilinear soft SUSY breaking parameter A_0 , and two low-energy parameters, the ratio of the two vacuum expectation values of the two Higgs doublets, $\tan \beta$, and the sign of the Higgs mixing parameter, $\text{sign}(\mu)$. Throughout this note, two CMSSM parameter sets, referred to as RM1 and RM2, are used to illustrate possible CMSSM yields. The parameter values defining RM1 are $m_0 = 320\text{ GeV}$, $m_{1/2} = 520\text{ GeV}$, $A_0 = 0$, $\tan \beta = 10$, and $\text{sign}(\mu) > 0$. Identical values are used for RM2 except $m_0 = 1800\text{ GeV}$ and $m_{1/2} = 280\text{ GeV}$.

The results are also interpreted using simplified model spectra (SMS) [8–10]. A simplified model is defined by an effective Lagrangian describing the interactions of a small number of new particles, which can be equally well described by a small number of observables, such as masses and cross-sections. Simplified models are therefore particularly useful for evaluating phase-space coverage of both individual searches and experiment-wide search programs, as well as providing an excellent starting point for characterizing positive signals of new physics.

2 Event reconstruction and selection

2.1 Hadronic pre-selection

The offline event reconstruction and selection criteria described below are explained in more detail in Refs. [4, 5]. Jets are reconstructed from the energy deposits in the calorimeter towers, clustered by the anti- k_T algorithm [11] with a size parameter of 0.5. The raw jet energies measured by the calorimeter systems are corrected to establish a uniform relative response in η and a calibrated absolute response in transverse momentum p_T with an associated uncertainty between 2% and 4%, depending on the jet η and p_T [12]. Jets considered in the analysis are required to have transverse energy $E_T > 50\text{ GeV}$. The highest- E_T jet is required to be within the central tracker acceptance ($|\eta| < 2.5$) and the two highest- E_T jets must each have $E_T > 100\text{ GeV}$. Events are vetoed if any additional jet satisfies both $E_T > 50\text{ GeV}$ and $|\eta| > 3$, or rare spurious signals are identified in the calorimeters [13, 14]. To suppress SM processes with genuine E_T from neutrinos, events containing an isolated electron [15] or muon [16] with $p_T > 10\text{ GeV}$ are vetoed. To select a pure multi-jet topology, events are vetoed in which an isolated photon [17] with $p_T > 25\text{ GeV}$ is found.

The following two variables characterize the visible energy and missing momentum in the transverse plane: the scalar sum of the transverse energy E_T of jets, defined as $H_T = \sum_{i=1}^{N_{\text{jet}}} E_T$, and the magnitude of the vector sum of the transverse momenta \vec{p}_T of jets, defined as $\cancel{H}_T = |\sum_{i=1}^{N_{\text{jet}}} \vec{p}_T|$, where N_{jet} is the number of jets with $E_T > 50\text{ GeV}$. Significant hadronic activity in the event is ensured by requiring $H_T > 275\text{ GeV}$. The aforementioned criteria complete the definition of the hadronic “pre-selection”.

2.2 Tagging jets from b-quarks

Jets originating from bottom quarks are identified through vertices that are displaced with respect to the primary interaction [18]. The algorithm used to tag b jets is the Combined Secondary Vertex tagger, using the “Medium” working point, which is achieved by requiring a cut of >0.679 on the algorithm discriminator variable and results in a gluon/light-quark quark mis-tag rate of 1% (where “light” means u , d and s quarks) and an efficiency in the range 60 – 70% depending on the jet p_T . This tagger is supported by the b-POG [18] and its per-

164 performance is well understood: the efficiency with which jets from heavy quarks are identified
 165 and the purity of the selection have been studied extensively with data and simulation [19].
 166 The algorithm performance has been also independently verified using a $t\bar{t}$ MC sample and the
 167 measured efficiency and purity is in agreement with those provided by the b-POG [19]. The
 168 mis-tag rate of hadronically-decaying tau leptons has also been studied using MC. The effect
 169 on the final hadronic signal selection is found to be negligible. An additional b jet-enriched
 170 QCD multi-jet sample was also tested and no events were found to survive the signal selection
 171 criteria.

172 2.3 Suppressing the multi-jet background with the α_T variable

173 Following the pre-selection criteria, the multi-jet background from QCD is still several orders of
 174 magnitude larger than the typical signal expected from SUSY. While the vast majority of these
 175 multi-jet events do not exhibit significant E_T , large values can be observed due to stochastic
 176 fluctuations in the measurement of jet energies, or *mismeasurements* caused by nonuniformi-
 177 ties in the calibration of the calorimeters, or detector inefficiencies.

The α_T kinematic variable, first introduced in Refs. [1–3], is used in the selection to efficiently reject events either without significant E_T or with transverse energy mismeasurements, while retaining a large sensitivity to new physics with genuine E_T signatures. For dijet events, the variable is defined as:

$$\alpha_T = \frac{E_T^{j2}}{M_T} \quad (1)$$

where E_T^{j2} is the transverse energy of the least energetic jet of the two, and M_T is the transverse mass of the dijet system:

$$M_T = \sqrt{\left(\sum_{i=1}^2 E_T^{ji}\right)^2 - \left(\sum_{i=1}^2 p_x^{ji}\right)^2 - \left(\sum_{i=1}^2 p_y^{ji}\right)^2}. \quad (2)$$

178 For a perfectly measured dijet event with $E_T^{j1} = E_T^{j2}$ and jets back to back in ϕ , and in the
 179 limit of large jet momenta compared to their masses, the value of α_T is 0.5. In the case of
 180 an imbalance in the measured transverse energies of back-to-back jets, α_T is smaller than 0.5.
 181 Values significantly greater than 0.5 are observed when the two jets are not back to back and
 182 balance genuine E_T .

For events with three or more jets, a dijet system is formed by combining the jets in the event into two *pseudo-jets*. The E_T of each of the two pseudo-jets is calculated as the scalar sum of the measured E_T of contributing jets. The combination chosen is the one that minimizes the E_T difference (ΔH_T) between the two pseudo-jets. This simple clustering criterion provides the best separation between multi-jet events and events with genuine E_T . Thus, in the case of events with at least three jets, the α_T variable can be defined as:

$$\alpha_T = \frac{1}{2} \cdot \frac{H_T - \Delta H_T}{\sqrt{H_T^2 - \Delta H_T^2}} = \frac{1}{2} \cdot \frac{1 - (\Delta H_T / H_T)}{\sqrt{1 - (\Delta H_T / H_T)^2}} \quad (3)$$

183 Events with severe jet energy mis-measurements due to detector inefficiencies, or extremely
 184 rare but large stochastic fluctuations, or multiple jets below the $E_T = 50$ GeV threshold, can
 185 lead to values of α_T slightly above 0.5. Such events are efficiently rejected by requiring a slightly
 186 tighter requirement on α_T , such as $\alpha_T > 0.55$, and by applying dedicated vetoes, described
 187 further below in Sec. 2.4.

188 2.4 The hadronic signal region

189 As done in Ref. [4], the hadronic signal region is defined by $H_T > 275 \text{ GeV}$ and $\alpha_T > 0.55$, which
 190 is divided into eight bins in H_T : two bins of width 50 GeV in the range $275 < H_T < 375 \text{ GeV}$, five
 191 bins of width 100 GeV in the range $375 < H_T < 875 \text{ GeV}$, and a final open bin, $H_T > 875 \text{ GeV}$.
 192 This choice of H_T and α_T thresholds is driven mainly by trigger constraints (efficiencies, rates)
 193 and the requirement of negligible contamination from QCD multi-jet events in all H_T bins. As
 194 done in Refs. [4, 5], jet E_T thresholds are scaled down from their nominal values to 36.7 GeV and
 195 43.3 GeV for the two lowest H_T bins in the region $275 < H_T < 375 \text{ GeV}$. This is done in order
 196 to maintain a large ratio between H_T and the jet E_T threshold, thus maintaining comparable
 197 jet multiplicities, kinematics and background admixture as observed for the higher HT bins.
 198 The contribution from $t\bar{t}$ is particularly sensitive to the jet E_T thresholds in the lowest H_T bins.
 199 Consequently, acceptance for stop-stop production and decay is also improved.

200 Finally, some additional *cleaning filters*, described in more detail in Refs. [4, 5], are added fol-
 201 lowing the α_T requirement to protect against pathological effects such as reconstruction failures
 202 or severe energy losses due to detector inefficiencies. To protect against multiple jets failing the
 203 E_T threshold, the jet-based estimate of the missing transverse energy, \cancel{H}_T , is compared to the cal-
 204 orimeter tower-based estimate, $\cancel{E}_T^{\text{calo}}$, and events with $R_{\text{miss}} = \cancel{H}_T / \cancel{E}_T^{\text{calo}} > 1.25$ are rejected.
 205 To protect against severe energy losses, events with significant jet mismeasurements caused by
 206 masked regions in the ECAL (which amount to about 1% of the ECAL channel count), or by
 207 missing instrumentation in the barrel-endcap gap, are removed with the following procedure.
 208 The jet-based estimate of the missing transverse energy, \cancel{H}_T , is used to identify jets most likely
 209 to have given rise to the \cancel{H}_T as those whose momentum is closest in ϕ to the total $\vec{\cancel{H}}_T$ which
 210 results after removing them from the event. The azimuthal distance between this jet and the
 211 recomputed \cancel{H}_T is referred to as $\Delta\phi^*$ in what follows. Events with $\Delta\phi^* < 0.5$ are rejected if
 212 the distance in the (η, ϕ) plane between the selected jet and the closest masked ECAL region,
 213 ΔR_{ECAL} , is smaller than 0.3. Similarly, events are rejected if the jet points within 0.3 in η of
 214 the ECAL barrel-endcap gap at $|\eta| = 1.5$. These final selections complete the definition of the
 215 hadronic signal sample.

216 Events found in the hadronic signal region (and the three control regions, described in Sec. 5)
 217 are further categorised according to the number of b-tags reconstructed in each event. The re-
 218 sulting sub-samples comprise events containing exactly zero, one, two, or at least three b-tags.
 219 Possible systematic biases due to MC mismodelling of the efficiency with which jets originat-
 220 ing from bottom quarks are tagged, which may lead to bin migration between the different
 221 sub-samples, are addressed in Sec. 6.

222 A disjoint hadronic control sample consisting predominantly of multi-jet events is defined by
 223 inverting the α_T requirement for a given H_T region, which is used primarily in the estimation
 224 of any residual background from QCD multi-jet events, described in Sec. 5.5.

225 2.5 Breakdown of SM backgrounds in the hadronic signal region

226 The remaining significant backgrounds in the hadronic signal region are expected to stem from
 227 SM processes with genuine \cancel{E}_T in the final state. In the dijet case, the largest backgrounds with
 228 genuine \cancel{E}_T are the associated production of W or Z bosons with jets, followed by either the
 229 weak decays $Z \rightarrow \nu\bar{\nu}$ or $W \rightarrow \tau\nu$, where the τ decays hadronically and is identified as a jet; or
 230 by leptonic decays that are not rejected by the dedicated electron or muon vetoes. At higher jet
 231 multiplicities, top quark production followed by semileptonic weak top quark decay becomes
 232 important.

Figure 1 shows the expectation from MC in all bins of the hadronic signal region for different background processes and the CMSSM reference model RM1. No requirement is made on the number of b-tagged jets per event. Table 1 shows a study of the relative background composition (expressed as a percentage of the total SM background) in the hadronic signal region, again in the absence of a requirement on the number of b-tagged jets per event. The following conclusions can be drawn: background contributions from $Z \rightarrow \nu\bar{\nu} + \text{jets}$ depends on H_T , which give $\sim 43\%$ (of the total SM background) for $275 < H_T < 325 \text{ GeV}$ increasing to $\sim 53\%$ for $H_T > 875 \text{ GeV}$; events containing W or Z bosons decaying to electrons or muons that are missed by the lepton vetoes contribute between $\sim 25\%$ (low H_T) and $\sim 13\%$ (high H_T); background events containing hadronically decaying τ leptons, denoted τ_h , rise with slowly H_T from $\sim 22\%$ of the total background to $\sim 27\%$; the fraction of τ leptons that decay to leptons, denoted by τ_l , that are missed by the lepton vetoes is $\sim 10\%$; the fraction of τ_h matched to a selected jet in the analysis also rises with H_T .

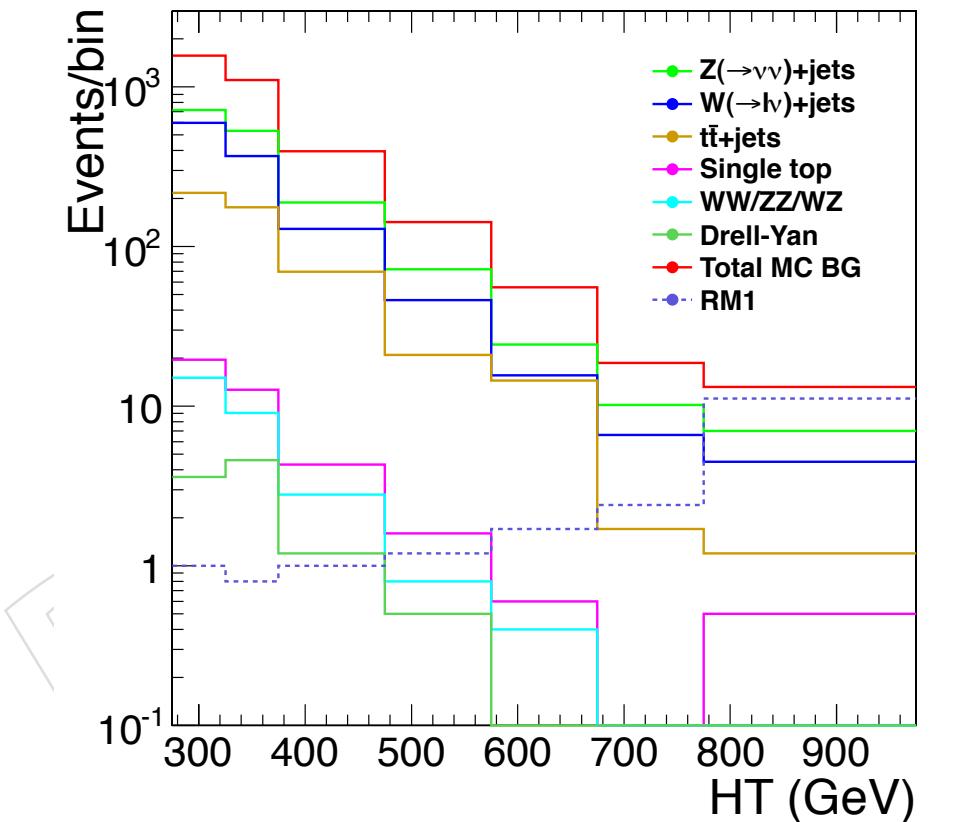


Figure 1: Expectation from MC in all bins of the hadronic signal region for the following different background processes: $Z \rightarrow \nu\bar{\nu} + \text{jets}$, $W + \text{jets}$, $t\bar{t}$, single top + jets, di-boson production ($WW/WZ/ZZ$), and Drell-Yan. The total SM expectation is also shown, along with that for the CMSSM benchmark model RM1.

2.6 Data–MC control distributions

Variables from the MC simulation are not used in absolute terms in order to estimate yields from background processes, but rather only ratios are used, in which any systematics arising from mismodelling will largely cancel. However, it is important to demonstrate the quality of the MC simulation, which is done with the data–MC comparison plots described below.

Table 1: Relative background composition as given by MC simulation in all bins of the hadronic signal region (expressed as a percentage of the total SM background).

| H_T (GeV) | N _{evts} | $Z \rightarrow \nu\bar{\nu} + \text{jets}$ (%) | W + jets, $t\bar{t}$, single top, DY and di-boson | | | |
|---------------|-------------------|---------------------------------------------------|----------------------------------------------------|--------------|--------------|--------------------------------|
| | | | missed e, μ from W/Z (%) | τ_h (%) | τ_l (%) | τ_h matched to jet (%) |
| | | | | | | |
| 275–325 | 3938.0 | 43 | 24 | 22 | 11 | 7 |
| 325–375 | 1569.9 | 46 | 25 | 22 | 9 | 7 |
| 375–475 | 1104.2 | 48 | 20 | 23 | 10 | 7 |
| 475–575 | 396.0 | 48 | 17 | 24 | 11 | 10 |
| 575–675 | 142.4 | 51 | 17 | 23 | 10 | 11 |
| 675–775 | 55.5 | 44 | 19 | 31 | 7 | 17 |
| 775–875 | 18.7 | 55 | 17 | 22 | 4 | 9 |
| 875– ∞ | 13.2 | 53 | 13 | 27 | 7 | 19 |

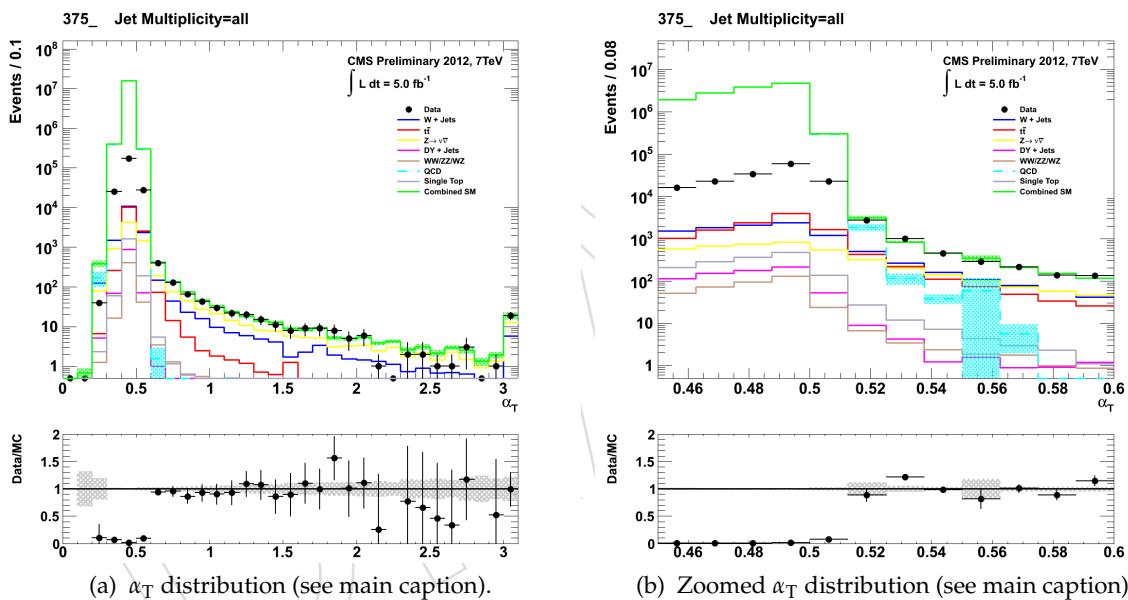


Figure 2: Data–MC comparison of the α_T distributions for the hadronic signal region, following all pre-selection and the requirements $H_T > 375$ GeV. No requirement is made on the number of b-tagged jets in the event. Bands represent the uncertainties due to the limited size of MC samples. *The discrepancy in the α_T distributions for values $\alpha_T < 0.55$ is due to the trigger not being applied in the MC simulation.*

Figures 2a and 2b show the α_T distribution (the latter is simply zoomed to the region $0.45 < \alpha_T < 0.6$) after all pre-selection requirements and $H_T > 375$ GeV. These plots demonstrate that the α_T variable is an excellent discriminator between multi-jet background from QCD and processes with genuine \cancel{E}_T . The number of expected multi-jet events falls quickly to zero with increasing α_T . The plots also demonstrate the excellent agreement between the data and MC expectation for the region $\alpha_T > 0.55$ (below this value, the discrepancy between data and MC is due to the trigger not being emulated in MC).

Figure 18 (Appendix A) shows the data–MC comparisons of key variables in the hadronic

signal region, following the full selection criteria, including α_T , and the requirement $H_T > 375\text{ GeV}$. Figures 18a and 18b show the comparisons between data and MC simulation for the H_T variable and the number of reconstructed jets per event, respectively. The following variable, known as $\Delta\phi^*$, is designed to identify jets with mismeasured transverse energy:

$$\Delta\phi^* = \min_k \left(\Delta\phi \left(\left(\sum_{i=0}^n -\vec{j}_i \right) + \vec{j}_k; \vec{j}_k \right) \right), \quad (4)$$

where n is the number of jets and \vec{j}_i (\vec{j}_k) is the momentum of the i^{th} (k^{th}) jet. The variable checks if there is at least one jet which, if rescaled by a certain factor, would be able to balance the event. For multi-jet events with one dominating mismeasurement, this angle tends to be small and so is a good indicator of contamination from multi-jet events. The $\Delta\phi^*$ distribution in Fig. 18c shows an approximately flat behaviour as expected from SM processes with real missing transverse energy such as $t\bar{t}$, $W + \text{jets}$ and $Z \rightarrow \nu\bar{\nu} + \text{jets}$ events. There is an apparent small contamination from QCD multi-jet events with a large uncertainty.² Figure 18d shows the distributions of effective mass M_{eff} , which is defined as the scalar sum $H_T + |\cancel{H}_T|$. Generally, the observed distributions in data show no significant deviation from the SM expectation.

Control distributions for the hadronic signal sample with the requirement of at least one b-tagged jet per event is shown in Fig. 19 (Appendix A). Again, the agreement between data and MC is good, giving confidence that the sample is well understood. The MC distributions highlight the composition of the sample, which comprises mainly $t\bar{t}$ events, plus some smaller, but non-negligible contributions from $W + \text{jets}$ and $Z \rightarrow \nu\bar{\nu} + \text{jets}$. The contribution from QCD multi-jet events is negligible in the region $\alpha_T > 0.55$.

3 Triggers

The trigger strategy has changed from that of the 1.1 fb^{-1} analysis [4], in which a cross trigger, applying thresholds simultaneously on the quantities H_T and \cancel{H}_T , was used. In this analysis, a cross trigger using the quantities H_T and α_T is used, labelled henceforth as `HT_AlphaT`. A single object (prescaled) H_T trigger, labelled `HT`, is also used to collect events for the hadronic control region. The H_T thresholds of the `HT` and `HT_AlphaT` triggers are chosen to match, with thresholds of 250, 300, 350 and 400 GeV. The α_T thresholds of the `HT_AlphaT` trigger are tuned according to the threshold on the H_T leg in order to fully suppress QCD multi-jet events (whilst simultaneously satisfying other criteria, such as sensitivity to trigger rates).

To ensure that the H_T leg of the `HT_AlphaT` cross trigger (and the `HT` single object trigger) is efficient with respect to the final event selection, the lower bounds of the offline H_T bins are offset by 25 GeV with respect to the online thresholds. The analysis therefore uses the following H_T binning: 275, 325, 375, 475, 575, 675, 775, and $>875\text{ GeV}$.

Data for the control samples of the analysis, described in Sec. 5, are collected using the `HT` and `HT_AlphaT` triggers, along with a single object photon trigger (`Photon`) and a cross trigger requiring some H_T and a muon above a p_T threshold (`Mu_HT`). The `Photon` trigger is measured to be fully efficient for the threshold $p_T^{\text{photon}} > 150\text{ GeV}$. The `Mu_HT` trigger efficiency is $(91.3 \pm 0.1)\%$ for the single muon control sample, which requires a single muon satisfying $p_T^{\mu\text{on}} > 45\text{ GeV}$. The efficiency is independent of H_T in the region $H_T > 375\text{ GeV}$. In the case of the di-muon control sample, the efficiency is 95% for the $375 < H_T < 475\text{ GeV}$ bin, rising to 97% for the highest H_T bin. The muon control samples are described in Sec. 5.2.

²Contamination from QCD multi-jet events is permitted in the final simultaneous fit, as described in Sections 8 and 9, which for the final result gives an expectation that is consistent with zero).

Table 2: Measured efficiencies of the H_T and α_T legs of the HT and HT_AlphaT triggers, for relevant analysis bins. The product of the efficiencies of the two legs gives the total efficiency of the HT_AlphaT cross trigger (rightmost column), in a given offline H_T and α_T bin.

| H_T range (GeV) | ϵ on H_T leg (%) | ϵ on α_T leg (%) |
|-------------------|-----------------------------|----------------------------------|
| 275–325 | $87.8^{+1.9}_{-1.9}$ | $82.8^{+1.0}_{-1.1}$ |
| 325–375 | $90.6^{+2.9}_{-2.9}$ | $95.9^{+0.7}_{-0.9}$ |
| 375–475 | $95.7^{+0.1}_{-0.1}$ | $98.5^{+0.5}_{-0.9}$ |
| 475– ∞ | $100.0^{+0.0}_{-0.0}$ | $100.0^{+0.0}_{-4.8}$ |

The efficiency of the HT_AlphaT trigger in bins of H_T and α_T is shown in Fig. 20a (Appendix B). The statistical uncertainties associated with these efficiencies can be found in Figs. 20b and 20c (Appendix B). Table 2 summarises the measured efficiencies for the HT and HT_AlphaT triggers in relevant bins of H_T , as used by the analysis and described in Sec. 2.4. Lists of the various thresholds used during 2011 for the aforementioned triggers can be found in Appendix B, along with distributions showing the efficiency turn-on curves for the HT_AlphaT (Fig. 21, Appendix B) and HT (Fig. 23, Appendix B) triggers at various thresholds. The trigger is 100% efficient for signal signatures with significant genuine E_T , as illustrated by Fig. 22 (Appendix B).

4 Data and Monte Carlo Samples

The following datasets are used to populate the hadronic signal and control samples. They correspond to the full data run of 2011 and an integrated luminosity of $4.98 \pm 0.11 \text{ fb}^{-1}$. The official JSON from the 30th October 2011 is used to filter only certified runs and luminosity sections.

```

307 /HT/Run2011A-May10ReReco-v1/AOD
308 /HT/Run2011A-PromptReco-v4/AOD
309 /HT/Run2011A-05Aug2011-v1/AOD
310 /HT/Run2011A-PromptReco-v6/AOD
311 /HT/Run2011B-PromptReco-v1/AOD
312 /Photon/Run2011A-May10ReReco-v1/AOD
313 /Photon/Run2011A-PromptReco-v4/AOD
314 /Photon/Run2011A-05Aug2011-v1/AOD
315 /Photon/Run2011A-PromptReco-v6/AOD
316 /Photon/Run2011B-PromptReco-v1/AOD
317 /MuHad/Run2011A-May10ReReco-v1/AOD
318 /MuHad/Run2011A-PromptReco-v4/AOD
319 /MuHad/Run2011A-05Aug2011-v1/AOD
320 /MuHad/Run2011A-PromptReco-v6/AOD
321 /MuHad/Run2011B-PromptReco-v1/AOD

```

The background Monte Carlo samples for this analysis are taken from the Summer11 simulation production for physics at 7 TeV with the PU_S4 scenario.

```

324 /QCD_TuneZ2_HT-*_7TeV-madgraph/Summer11-PU_S4_START42_V11-v*/AODSIM
325 /QCD_BLepEnriched_TuneZ2_7TeV-pythia6-evtgen/Summer11-PU_S3_START42_V11-v1/AODSIM
326 /GJets_TuneZ2_40_HT_100_7TeV-madgraph/Summer11-PU_S4_START42_V11-v1/AODSIM
327 /GJets_TuneZ2_100_HT_200_7TeV-madgraph/Summer11-PU_S4_START42_V11-v1/AODSIM
328 /GJets_TuneZ2_200_HT_inf_7TeV-madgraph/Summer11-PU_S4_START42_V11-v1/AODSIM
329 /TTJets_TuneZ2_7TeV-madgraph-tauola/Summer11-PU_S4_START42_V11-v1/AODSIM

```

```

330 /T_TuneZ2_s-channel_7TeV-powheg-tauola/Fall111-PU_S6_START42_V14B-v1/AODSIM
331 /T_TuneZ2_t-channel_7TeV-powheg-tauola/Fall111-PU_S6_START42_V14B-v1/AODSIM
332 /T_TuneZ2_tW-channel-DS_7TeV-powheg-tauola/Fall111-PU_S6_START42_V14B-v1/AODSIM
333 /Tbar_TuneZ2_s-channel_7TeV-powheg-tauola/Fall111-PU_S6_START42_V14B-v1/AODSIM
334 /Tbar_TuneZ2_t-channel_7TeV-powheg-tauola/Fall111-PU_S6_START42_V14B-v1/AODSIM
335 /Tbar_TuneZ2_tW-channel-DS_7TeV-powheg-tauola/Fall111-PU_S6_START42_V14B-v1/AODSIM
336 /WJetsToLNu_TuneZ2_7TeV-madgraph-tauola/Summer11-PU_S4_START42_V11-v1/AODSIM
337 /WJetsToLNu_*_HT_*_TuneZ2_7TeV-madgraph-tauola/Summer11-PU_S4_START42_V11-v1/AODSIM
338 /DYJetsToLL_TuneZ2_M-50_7TeV-madgraph-tauola/Fall111-PU_S6_START42_V14B-v1/AODSIM
339 /ZJetsToNuNu_*_HT_*_7TeV-madgraph/Summer11-PU_S4_START42_V11-v*/AODSIM
340 /WW_TuneZ2_7TeV_pythia6_tauola/Fall111-PU_S6_START42_V14B-v1/AODSIM
341 /WZ_TuneZ2_7TeV_pythia6_tauola/Fall111-PU_S6_START42_V14B-v1/AODSIM
342 /ZZ_TuneZ2_7TeV_pythia6_tauola/Fall111-PU_S6_START42_V14B-v1/AODSIM

343 The signal Monte Carlo samples for this analysis are taken from the Summer11 FastSim simu-
344 lation production for physics at 7 TeV.
345 /mSUGRA_m0-220to3000_m12-100to1000_tanb-10andA0-0_7TeV-Pythia6Z/StoreResults-PU_
346 /SMS-T1_Mgluino-100to1200_mLSP-50to1150_7TeV-Pythia6Z/Summer11-PU_START42_V11_Fa
347 /SMS-T2_Mgluino-100to1200_mLSP-50to1150_7TeV-Pythia6Z/Summer11-PU_START42_V11_Fa
348 /SMS-T2tt_Mstop-225to1200_mLSP-50to1025_7TeV-Pythia6Z/Summer11-PU_START42_V11_Fa
349 /SMS-T2bb_Msbottom-100to1200_mLSP-50to1150_7TeV-Pythia6Z/Summer11-PU_START42_V11
350 /SMS-T1tttt_Mgluino-450to1200_mLSP-50to800_7TeV-Pythia6Z/Summer11-PU_START42_V11
351 /SMS-T1bbbb_Mgluino-100to1200_mLSP-50to1150_7TeV-Pythia6Z/Summer11-PU_START42_V11

```

352 5 Background estimation

353 5.1 Overview of the method

354 The α_T variable is used to suppress QCD multi-jet events to a negligible level in the hadronic
355 signal region. In the absence of multi-jet events, the remaining background events result almost
356 entirely from the following three SM processes: $W + \text{jets}$, $Z \rightarrow \nu\bar{\nu} + \text{jets}$ and $t\bar{t}$ production. Small
357 contributions from single top, Drell-Yan and di-boson production are also expected. Three
358 independent control samples are used to predict the contributions from these SM processes.
359 The $Z \rightarrow \nu\bar{\nu} + \text{jets}$ background is estimated from two independent samples: a di-muon + (at
360 least two) jets sample, and a single photon + (at least two) jets sample. A single muon + (at
361 least two) jets sample is used to estimate the contribution from all remaining SM backgrounds,
362 which is dominated by $W + \text{jets}$ and $t\bar{t}$ production.

363 The method to estimate these background contributions in the hadronic signal region relies
364 on the use of a *translation factor* (TF) determined from MC samples to transform the observed
365 yield in a given H_T and b-tag multiplicity (n_b) bin of a control sample, $N_{\text{obs}}^{\text{control}}(H_T, n_b)$, into a
366 predicted yield for the corresponding bin of the hadronic signal region, $N_{\text{pred}}^{\text{signal}}(H_T, n_b)$. Each
367 translation factor is simply a ratio of the yields obtained from MC simulation for the same
368 H_T and b-tag multiplicity bins of each control sample, $N_{\text{MC}}^{\text{control}}(H_T, n_b)$, and the signal region,
369 $N_{\text{MC}}^{\text{signal}}(H_T, n_b)$:

$$N_{\text{pred}}^{\text{signal}}(H_T, n_b) = N_{\text{obs}}^{\text{control}}(H_T, n_b) \times \text{TF} \quad , \quad \text{TF} = \frac{N_{\text{MC}}^{\text{signal}}(H_T, n_b)}{N_{\text{MC}}^{\text{control}}(H_T, n_b)} \quad (5)$$

When constructing the translation factors, the MC yields from the following SM processes are considered: $W + \text{jets}$ (N_W), $t\bar{t} + \text{jets}$ ($N_{t\bar{t}}$), $Z \rightarrow \nu\bar{\nu} + \text{jets}$ ($N_{Z \rightarrow \nu\bar{\nu}}$), $DY + \text{jets}$ (N_{DY}), single top + jets production via the s , t , and tW -channels (N_{top}), and $WW + \text{jets}$, $WZ + \text{jets}$, and $ZZ + \text{jets}$ ($N_{\text{di-boson}}$). All yields from MC are normalised to the luminosity of the data samples, 3.9 fb^{-1} . The MC samples are listed in Sec. 4. The sum of expected yields from all MC samples, obtained for the relevant control sample selection, enter the denominator of each translation factor:

$$N_{\text{MC}}^{\text{control}}(H_T, n_b) = N_W + N_{t\bar{t}} + N_{Z \rightarrow \nu\bar{\nu}} + N_{DY} + N_{\text{top}} + N_{\text{di-boson}} \quad (6)$$

For the b jet multiplicity bins satisfying $n_b \leq 2$, the $\mu + \text{jets}$ control sample is used to predict primarily the $W + \text{jets}$ and $t\bar{t} + \text{jets}$ backgrounds, but all remaining residual SM backgrounds except $Z \rightarrow \nu\bar{\nu} + \text{jets}$ events are also considered. (The $Z \rightarrow \nu\bar{\nu} + \text{jets}$ background is instead accounted for through the $\mu\mu + \text{jets}$ and $\gamma + \text{jets}$ samples.) The sum of expected yields from all MC samples except $Z \rightarrow \nu\bar{\nu} + \text{jets}$, obtained for the hadronic signal region selection, enter the numerator of each translation factor:

$$N_{\text{MC}}^{\text{signal}}(H_T, n_b \leq 2) = N_W + N_{t\bar{t}} + N_{DY} + N_{\text{top}} + N_{\text{di-boson}} \quad (7)$$

For the same b jet multiplicity bins ($n_b \leq 2$), the $\mu\mu + \text{jets}$ and $\gamma + \text{jets}$ control samples are used to predict the $Z \rightarrow \nu\bar{\nu} + \text{jets}$ process only, and the expected yields in the bins of the signal region as obtained from the $Z \rightarrow \nu\bar{\nu}$ sample enter the numerator of each translation factor:

$$N_{\text{MC}}^{\text{signal}}(H_T, n_b \leq 2) = N_{Z \rightarrow \nu\bar{\nu}} \quad (8)$$

For the b jet multiplicity bins satisfying $n_b \geq 3$, the $\mu + \text{jets}$ control sample is again used to predict primarily the $W + \text{jets}$ and $t\bar{t} + \text{jets}$ backgrounds, but also to predict all other remaining residual SM backgrounds *including* $Z \rightarrow \nu\bar{\nu} + \text{jets}$ events. The sum of expected yields from all MC samples *including* $Z \rightarrow \nu\bar{\nu} + \text{jets}$, obtained for the hadronic signal region selection, enter the numerator of each translation factor:

$$N_{\text{MC}}^{\text{signal}}(H_T, n_b \geq 3) = N_W + N_{t\bar{t}} + N_{DY} + N_{Z \rightarrow \nu\bar{\nu}} + N_{\text{top}} + N_{\text{di-boson}} \quad (9)$$

In this case, the $\mu\mu + \text{jets}$ and $\gamma + \text{jets}$ control samples are not used, as the yields in the two data control samples are expected to be negligible due to the requirement of at least three b jets per event. The method of using a $W + \text{jets}$ sample to predict the $Z \rightarrow \nu\bar{\nu} + \text{jets}$ background has been used previously [4, 5], and this approach is addressed by a set of closure tests described in Sec. 6, in which a $\mu + \text{jets}$ sample (rich in $W + \text{jets}$ and $t\bar{t}$) is used to make predictions of yields in a $\mu\mu + \text{jets}$ sample (rich in $Z \rightarrow \ell\ell + \text{jets}$).

All of the above predictions are considered simultaneously in a fit defined formally by the likelihood model described in Sec. 8. “Vanilla” predictions for the total SM background can be made by considering separately the sum of the predictions from the $\mu + \text{jets}$ and $\gamma + \text{jets}$ samples or the $\mu + \text{jets}$ and $\mu\mu + \text{jets}$ samples.

The selection criteria for each of the three control samples closely resemble those for the signal region, and differ only in the use of a muon, di-muon, or photon *tag* and some minimal kinematic selections to enrich further the samples of W or Z bosons. Hence, the reliance on MC to extrapolate correctly from a control region to the signal region is minimised and many systematic effects are expected to cancel largely in the ratio (i.e. translation factor). However,

405 a systematic uncertainty is assigned to each translation factor to account for theoretical uncer-
 406 tainties and effects such as the mismodelling of kinematics (e.g. acceptances) and instrumental
 407 effects (e.g. reconstruction inefficiencies).

408 Kinematic cuts are applied to enrich as much as possible the $W + \text{jets}$, $t\bar{t}$, and $Z \rightarrow \nu\bar{\nu}$ com-
 409 ponents in the muon and di-muon control samples. The definition of the samples are geared
 410 towards efficiency rather than purity (even so, the purities are at the level >90%) and any
 411 contamination from "backgrounds" (e.g. $t\bar{t}$ in the case of the $\mu\mu + \text{jets}$ sample) are simply in-
 412 corporated into the translation factors. The assumption here is that all the control samples
 413 comprise processes (including "backgrounds") that can be well modelled by the simulation.
 414 These comprise primarily $W + \text{jets}$, $t\bar{t}$, $Z \rightarrow \nu\bar{\nu}$ and $DY + \text{jets}$. Residual contributions arise from
 415 single top and di-boson production.

416 Furthermore, while these samples could also contain potential contamination from a SUSY sig-
 417нал, every effort has been made to suppress signal contamination in these samples because
 418 their main purpose is to establish a data-driven estimate for the SM backgrounds with a gen-
 419 uine E_T final-state signature in the hadronic signal sample. Therefore, in the following, we
 420 refer to these samples as *control samples* although in the final simultaneous fit, any potential
 421 signal contamination is properly taken into account.

422 The magnitude of the systematic uncertainties on the translation factors are motivated by a
 423 comprehensive set of closure tests, in which yields from one control sample, along with the cor-
 424 responding translation factors from MC, are used to predict the yields in another. No assumed
 425 systematic uncertainties are used in any of the closure tests. Instead, the level of statistical
 426 consistency between predicted and observed yields from these various closure tests validates
 427 the use of MC to determine the translation factors and also motivates the magnitude of the
 428 systematics uncertainties assigned to these factors. The treatment for estimating the systematic
 429 uncertainties on the translation factors is described in Sec. 6.

430 5.2 Definition of the control samples

431 5.2.1 The $\mu + \text{jets}$ control sample

432 Events from the $W + \text{jets}$ and $t\bar{t}$ processes are found in the hadronic signal sample due to
 433 unidentified leptons (either out of acceptance or not reconstructed) and hadronic tau decays
 434 originating from high- p_T W bosons. An estimate of these background processes is obtained
 435 through the use of a $\mu + \text{jets}$ sample. The selection criteria for this sample are chosen to identify
 436 W bosons decaying to a muon and a neutrino in the phase-space of the signal. The muon is not
 437 considered in the calculation of event-level variables such as H_T , \cancel{H}_T and α_T . All cuts on such
 438 jet-based quantities are consistent with those applied in the hadronic search region, and the
 439 same H_T binning is used. In order to select events containing W bosons, exactly one tight iso-
 440 lated muon within an acceptance of $p_T > 10$ GeV and $|\eta| < 2.5$ is required, and the transverse
 441 mass of the W candidate must satisfy $M_T(\mu, E_T) > 30$ GeV (to suppress QCD multi-jet events).
 442 Events are vetoed if either of the following conditions are met: $\Delta R(\mu, \text{jet}_i) < 0.5$, running over
 443 all jets i ; or a second muon candidate exists that is either loose, non-isolated or outside accep-
 444 tance and the two muons have an invariant mass that satisfies $m_Z - 25 < M_{\mu_1\mu_2} < m_Z + 25$ (to
 445 suppress $Z \rightarrow \mu\mu$).

446 5.2.2 The $\mu\mu + \text{jets}$ control sample

447 The $Z \rightarrow \nu\bar{\nu} + \text{jets}$ process forms an irreducible background and can be estimated using the
 448 $Z \rightarrow \mu\mu + \text{jets}$ process, which has identical kinematic properties, but a different acceptance
 449 and a smaller branching ratio. A background estimate is obtained through the use of a $\mu\mu +$

450 jets sample. The selection criteria are tuned to identify Z bosons decaying to two muons in
 451 the phase-space of the signal. The muons are not considered in the calculation of event-level
 452 variables such as H_T , \cancel{H}_T and α_T . All cuts on such jet-based quantities are consistent with those
 453 applied in the hadronic search region, and the same H_T binning is used. In order to select an
 454 event samples containing Z bosons, exactly two tight isolated muons within an acceptance of
 455 $p_T > 10 \text{ GeV}$ and $|\eta| < 2.5$ are required, and the invariant mass of the two muons must satisfy
 456 $m_Z - 25 < M_{\mu_1\mu_2} < m_Z + 25$. Events are vetoed if $\Delta R(\mu_i, \text{jet}_j) < 0.5$ is satisfied, running over
 457 all muons i and all jets j . The $\mu\mu + \text{jets}$ sample can be used to make predictions in the two
 458 lowest H_T bins, providing coverage where the $\gamma + \text{jets}$ sample cannot.

459 5.2.3 The $\gamma + \text{jets}$ control sample

460 The $Z \rightarrow \nu\bar{\nu} + \text{jets}$ process can also be estimated using the $\gamma + \text{jets}$ process, which has a larger
 461 cross section and kinematic properties similar to those of $Z \rightarrow \nu\bar{\nu}$ events when the photon is
 462 ignored [20, 21]. The $\gamma + \text{jets}$ sample is defined by requiring exactly one photon satisfying tight
 463 isolation criteria and within an acceptance of $p_T > 150 \text{ GeV}$ and $|\eta| < 1.45$. Furthermore, events
 464 are vetoed if $\Delta R(\gamma, \text{jet}_j) < 1.0$ is satisfied, running over all jets j . As in the $\mu + \text{jets}$ sample, the
 465 photon is not considered in the calculation of event-level variables such as H_T , \cancel{H}_T and α_T . All
 466 cuts on jet-based quantities are consistent with those applied in the hadronic search region, and
 467 the same H_T binning is used. Given that the photon is ignored, the $\gamma + \text{jets}$ sample can only be
 468 used for the region $H_T > 375 \text{ GeV}$ due to the photon acceptance of $p_T > 150 \text{ GeV}$ (enforced by
 469 the trigger) and the cut $\alpha_T > 0.55$, which implies a H_T threshold of $\sim 350 \text{ GeV}$.

470 5.3 Distributions, yields and translation factors from the control samples

471 The number of events in the hadronic signal region from the $W + \text{jets}$, $t\bar{t}$, and $Z \rightarrow \nu\bar{\nu}$ processes
 472 is estimated using the method described by Equ. 5. The $W + \text{jets}$ and $t\bar{t}$ backgrounds are esti-
 473 mated using the $\mu + \text{jets}$ sample. Two independent samples are used to estimate the $Z \rightarrow \nu\bar{\nu}$
 474 + jets background, namely the $\gamma + \text{jets}$ and $\mu\mu + \text{jets}$ samples. Supplementary information
 475 concerning these samples can be found in Appendix C.

476 Figure 24 (Appendix C.1) shows the b-tag multiplicity distributions for the hadronic signal
 477 region and the $\mu + \text{jets}$ and $\mu\mu + \text{jets}$ control samples, as observed in data. Also shown is the
 478 expectation from MC. As expected, the relative contribution from $t\bar{t}$ to the total SM background
 479 is enhanced when at least one b-tag is required.

480 Control distributions for all three samples are shown in Appendices C.2 ($\mu + \text{jets}$ sample), C.3
 481 ($\mu\mu + \text{jets}$ sample), and C.4 ($\gamma + \text{jets}$ sample), which demonstrate the level of agreement between
 482 data and MC expectation. For each sample, two sets of distributions are shown: first, with no
 483 requirement made on the number of b-tagged jets per event; and second, with the requirement
 484 of at least one b-tagged jet per event. In general, the agreement between data and MC is good,
 485 giving confidence that the samples are well understood. The MC distributions highlight the
 486 composition of each sample, which comprise mainly of events from the aforementioned pro-
 487 cesses with genuine \cancel{E}_T . The contribution from QCD multi-jet events is negligible. The extra
 488 requirement of up to two b-tags per event suppresses further all backgrounds except for $t\bar{t}$ pro-
 489 duction. Requiring at least three b-tags suppresses also $t\bar{t}$ production. To ensure sufficiently
 490 large yields in b-tagged control samples, in order to guarantee statistically significant predic-
 491 tions, a different trigger strategy and selection criteria are used, as described below in Sec. 5.4.
 492 There is a negligible contribution from QCD multi-jet events in all cases.

493 Tables 20, 21, 22 and 23 (Appendix C.2) summarise the observed yields per H_T bin in the μ
 494 + jets control sample when requiring exactly zero, one, two or at least three b-tagged jets per

495 event, respectively. Also summarised are the expectations from MC for both the $\mu +$ jets con-
 496 trol sample and the hadronic signal region and the resulting translation factors, again binned
 497 according to H_T and the number of b-tagged jets per event. The errors associated with the
 498 translation factors reflect the uncertainty due to the finite size of the MC samples used to de-
 499 termine the factors. Any trigger inefficiency is also factored into the translation factors (i.e., the
 500 trigger is effectively emulated and yields from the MC samples are corrected to account for any
 501 inefficiency). Also, all MC expectations are corrected to account for any discrepancies between
 502 data and MC for the efficiency and mistag rate of the b-tagging algorithm used, as described
 503 further in Sec. 6.6. However, no systematic uncertainties on the translation factors are quoted
 504 in the tables.

505 The same tables also list the predicted yields for the $W +$ jets and $t\bar{t}$ processes in each of the H_T
 506 bins of the hadronic signal region. These predictions are given for illustrative purposes only.
 507 For the final result, the predictions for the total SM background are determined by a simulta-
 508 neous fit to the yields in the signal region and all three control samples, as described in Sec. 8
 509 onwards. In addition to observed yields, the simultaneous fit takes as input the translation
 510 factors with their associated statistical and systematic uncertainties.

511 Similarly, Tables 24, 25 and 26 (Appendix C.3) and 27, 28 and 29 (Appendix C.4) summarise the
 512 same information for the $\mu\mu +$ jets and $\gamma +$ jets control samples, respectively, when requiring
 513 exactly zero, one and two b-tags per event. These tables also list the predicted yields from the
 514 $Z \rightarrow \nu\bar{\nu} +$ jets process in each of the H_T bins of the hadronic signal region. When requiring
 515 at least three b-tagged jets per event, only the $\mu +$ jets sample is used to predict the total SM
 516 background. As noted previously, these predictions are listed for illustrative purposes only
 517 and are not used in the final result.

518 Illustrative predictions for the sum of events from all SM background processes can be made
 519 for each bin in the signal region, by combining the individual predictions quoted above. One
 520 such combination can be made by using the individual predictions from the $\mu +$ jets and $\mu\mu +$
 521 jets samples (or, alternatively, the predictions from the $\mu +$ jets and $\gamma +$ jets samples), the result
 522 of which can be compared with the observed yields in the bins of the signal region. Predictions
 523 in this way are made for bins of the three exclusive b-tag categories requiring exactly zero, one
 524 and two b-tags per event, as shown in Tables 30, 31 and 32 (Appendix C.5). When requiring at
 525 least three b-tagged jets per event, only the $\mu +$ jets sample has sufficiently large yields to predict
 526 accurately the total SM background. This method is summarised in Table 33. (Predictions for
 527 the hadronic signal samples containing the exactly zero, one and two b-tags per event using
 528 only the $\mu +$ jets sample can be found in Appendix C.5.2.) As previously stated, the errors on
 529 the predictions reflect the statistical uncertainties only. It is noted again that these predictions
 530 are for illustration only, with the final SM expectations for all signal region bins given by the
 531 simultaneous fit to all data samples.

532 5.4 Larger muon samples without an α_T requirement

533 As described above, the selection criteria for each of the three control samples is defined to
 534 resemble as closely as possible those for the signal region. This is done in order to minimise
 535 the reliance on MC to correctly model subtle differences in the background kinematics and
 536 topologies for the different phase spaces of the control and signal regions. In this way, the α_T
 537 cut is also used within the event selection for the three control samples, which greatly reduces
 538 the acceptance and reduces the statistical power of these samples to make predictions. The α_T
 539 requirement is also due to using the `HT_AlphaT` trigger to collect data for these samples.

540 In the case of the $\mu +$ jets and $\mu\mu +$ jets samples, the α_T cut is not strictly required, as the

- 541 preceeding kinematic selections ensure that the samples are rich in events from $W + \text{jets}$, $t\bar{t}$ and
 542 $Z \rightarrow \mu\mu$ processes, with negligible contamination from QCD multi-jet events. Therefore, the α_T
 543 requirement can be removed, thus significantly increasing the acceptance of these two control
 544 samples and improving their predictive power.
- 545 In the case of the $\gamma + \text{jets}$ sample, the requirement $\alpha_T > 0.55$ is necessary to suppress significant
 546 contamination from QCD multi-jet events, even after the substantial photon p_T cut in the offline
 547 selection.
- 548 Suppressing QCD multi-jet events to negligible levels is achieved by the selection criteria that
 549 define the $\mu + \text{jets}$ and $\mu\mu + \text{jets}$ samples, as described in Sec. 5.2: namely, the requirement of
 550 exactly one or two tight isolated muon(s); and additional kinematic cuts, such as acceptance
 551 windows (or the converse) of invariant and transverse mass, to select samples rich in W and
 552 Z bosons. The absence of QCD multi-jet events is demonstrated by the control distributions
 553 shown in Appendices C.2 and C.3.
- 554 Thus, in this way, the yields in the $\mu + \text{jets}$ and $\mu\mu + \text{jets}$ control samples can be significantly
 555 increased. However, this method requires using a sample that does not use an α_T cut in its
 556 selection to make predictions for the hadronic signal region, which certainly does use the α_T
 557 variable. The closure tests described below demonstrate that this is a valid approach and that
 558 the different α_T acceptances for control and signal regions have no significant systematic bias
 559 on the prediction.
- 560 In the absence of an α_T cut, events in the the $\mu + \text{jets}$ and $\mu\mu + \text{jets}$ samples cannot be collected
 561 with an `HT_AlphaT` trigger, as was the case in Sec. 5.2. Instead, a muon- H_T cross trigger, la-
 562 belled henceforth as `Mu_HT`, is used. Due to varying thresholds used during data-taking, as
 563 detailed in Table 18 in Appendix B, the muon p_T requirement is raised to 45 GeV to maintain a
 564 flat-top efficiency of $91.3 \pm 0.1\%$. Furthermore, the threshold on the H_T leg is raised as high as
 565 300 GeV, so these triggers and corresponding samples can only be used to make background
 566 predictions for the six highest H_T bins, covering the range $H_T > 375$ GeV. In the case of the
 567 di-muon control sample, the efficiency is found to be H_T -dependent in the range 95 – 97% due
 568 to the presence of events containing two muons above the threshold of 45 GeV, either of which
 569 could be responsible for the trigger.
- 570 Conversely, the $\mu + \text{jets}$ and $\mu\mu + \text{jets}$ samples defined *with* an α_T cut and collected by the
 571 `HT_AlphaT` triggers are used to make predictions in the two lowest H_T bins, in the range $275 <$
 572 $H_T < 375$ GeV. These details are summarized in Table 3.

Table 3: List of triggers used for the larger $\mu + \text{jets}$ and $\mu\mu + \text{jets}$ samples.

| H_T bin (GeV) | 275–325 | 325–375 | 375–475 | 475–575 | 575–675 | 675–775 | 775–875 | >875 |
|-----------------|------------------------|------------------------|--------------------|--------------------|--------------------|--------------------|--------------------|--------------------|
| α_T cut | 0.55 | 0.55 | None | None | None | None | None | None |
| Muon p_T cut | 10 | 10 | 45 | 45 | 45 | 45 | 45 | 45 |
| Trigger | <code>HT_AlphaT</code> | <code>HT_AlphaT</code> | <code>Mu_HT</code> | <code>Mu_HT</code> | <code>Mu_HT</code> | <code>Mu_HT</code> | <code>Mu_HT</code> | <code>Mu_HT</code> |
| Dataset | <code>HT</code> | <code>HT</code> | <code>MuHad</code> | <code>MuHad</code> | <code>MuHad</code> | <code>MuHad</code> | <code>MuHad</code> | <code>MuHad</code> |
| Thresholds | Table 17 | Table 17 | Table 18 |

- 573 That the α_T variable introduces no acceptance bias for processes with genuine E_T is due to the
 574 accurate modelling by the CMS simulation of such processes, namely $W + \text{jets}$, $t\bar{t}$, and $Z \rightarrow \nu\bar{\nu}$
 575 + jets. Background estimates for these processes are provided by the $\mu + \text{jets}$ and $\mu\mu + \text{jets}$
 576 samples (when the muon pair is ignored), as identified at the beginning of this Section. Im-
 577 portantly, the same cannot be said for QCD multi-jet events, as in this case the only events
 578 that survive the α_T cut are pathological cases in which a jet is severely mismeasured or even

“lost” due to detector inefficiencies. Such effects certainly do change the apparent event kinematics and, in these cases, an α_T cut will selectively choose events with particular kinematic features and topologies. For these pathological cases, one cannot rely on MC to model correctly their behaviour and therefore the α_T acceptance. This is a crucial distinction to be made between QCD multi-jet events and processes with significant genuine E_T . The assumption is that processes with genuine E_T are selected by the α_T variable based on the escaping invisible particle(s) rather than any pathological effects.

That removing the α_T requirement does not introduce any significant bias on the yield predictions of processes with genuine E_T is demonstrated by closure tests described in Sec. 6.

5.5 Estimating the background from multi-jet events

Negligible contamination from QCD multi-jet events in the signal region is expected due to the application of the cuts $H_T > 275\text{ GeV}$ and $\alpha_T > 0.55$. Any residual leakage is removed using various cleaning filters, as described in Sec. 2.4. These filters include the requirement $H_T/E_T < 1.25$, which ensures that jets below the E_T threshold do not contribute significantly to H_T .

Although the signal region is expected to be free from multi-jet events, a conservative approach is taken and the likelihood model is given the freedom to estimate any potential contamination from multi-jet events. This is achieved by considering the H_T -dependence of the variable R_{α_T} , which is defined as the ratio of events above and below the threshold value of $\alpha_T^{\text{cut}} = 0.55$ for a given H_T bin. This dependence is modelled accurately by a (falling) exponential function of H_T , which takes the form $R_{\alpha_T}(H_T) = A_{n_b} e^{-k H_T}$, where the parameters A_{n_b} and k (GeV^{-1}) are normalisation and exponential decay constants, respectively. A unique decay constant (k) is assumed for all b-tag multiplicity categories and an independent normalisation parameter (A_{n_b}) is provided for each b-tag multiplicity. The same exponential model has been used in 2010 and 2011 [4, 5] and has been validated using both MC simulation and data. Further data-driven studies are detailed below. The formal description of the likelihood model can be found in Sec. 8.

The exponential behaviour is a result of several features, namely: the improving jet energy resolutions with increasing H_T ; the reducing impact of pathological effects at increasing energy scales; and, for the region $H_T > 375\text{ GeV}$, the average jet multiplicity per bin increases slowly with H_T , which results in a narrower α_T distribution (peaking at 0.5) due to the increased combinatorics in the ΔH_T jet recombination scheme.

Due to the signal region definition and this exponentially falling behaviour, MC studies demonstrate that the expected contamination from multi-jet events falls quickly from a negligible contribution to zero in the region $400 - 500\text{ GeV}$. Therefore, the accuracy of the exponential modelling at high H_T (i.e., $H_T > 575\text{ GeV}$) is not of paramount importance, although the model has been proven valid for the full H_T range used in the analysis, as demonstrated in data below.

Maximum likelihood (ML) values of the parameters k and A_{n_b} are determined by the likelihood fit. However, the value of decay constant k is constrained first via measurements in multi-jet-enriched data side-bands. This is done in order to account correctly for the background composition, namely QCD multi-jets or backgrounds with genuine E_T . Cross-checks demonstrate that removing this constraint altogether has little effect on the total SM background expectations determined by the likelihood fit. The data samples in the side-band regions are collected using the set of prescaled H_T triggers described in Sec. 3.

Figure 3 depicts a simple cartoon that defines the data side-bands used to constrain the value

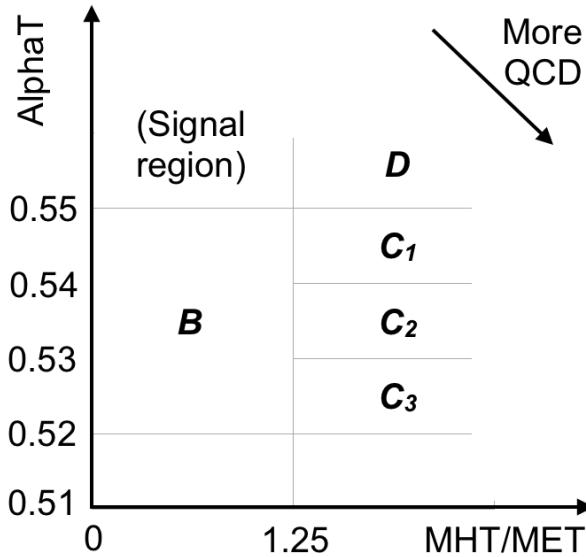


Figure 3: QCD side-bands.

of the parameter k . The side-bands are defined using the variables α_T and H_T/\cancel{E}_T . The signal region is labelled in the cartoon and is defined by $\alpha_T > 0.55$ and $H_T/\cancel{E}_T < 1.25$. Region B is a side-band defined by inverting the α_T cut (i.e. $\alpha_T < 0.55$). Region C is defined by inverting also the H_T/\cancel{E}_T cut (i.e. $H_T/\cancel{E}_T > 1.25$) and splitting into three slices in α_T : $0.52 < \alpha_T < 0.53$, $0.53 < \alpha_T < 0.54$, and $0.54 < \alpha_T < 0.55$ (labelled as C_1 , C_2 , and C_3 , respectively). By inverting one and then the other cut on these two variables, the sample is increasingly enriched in QCD multi-jet events (loosely in the direction of the arrow). A further multi-jet-enriched side-band, labelled D , is defined by inverting just the H_T/\cancel{E}_T cut; this region is not used to constrain the value of the parameter k , but instead is used in a further cross-check on the validity of the exponential model, described below.

Figures 30 shows the resulting behaviour of R_{α_T} as a function of H_T for the side-band regions B , C_1 , C_2 , and C_3 . The large uncertainties on the measurements are due to the large prescales applied to the HT triggers. Measurements are made in the region $275 < H_T < 575$ only.³ An exponential fit to the data is made for each side-band region. Table 4 summarises the best fit parameter values for k and the associated p -values.

Table 4: Best fit values for the parameters k as obtained from the regions B , C_1 , C_2 , and C_3 . The latter three measurements are used to calculate a weighted mean (identified as region C). Also quoted is the maximum likelihood value of the parameter k given by the simultaneous fit using the sample defined by region D . Quoted errors are statistical only.

| Side-band region | $k(\times 10^{-2} \text{ GeV}^{-1})$ | p -value |
|----------------------|--------------------------------------|------------|
| B | 2.96 ± 0.64 | 0.24 |
| C_1 | 1.19 ± 0.45 | 0.93 |
| C_2 | 1.47 ± 0.37 | 0.42 |
| C_3 | 1.17 ± 0.55 | 0.98 |
| C (weighted mean) | 1.31 ± 0.26 | - |
| D (likelihood fit) | 1.31 ± 0.09 | 0.57 |

³A possible extension to this analysis is to extend the signal region to include lower values of α_T , namely covering the regions $0.53 < \alpha_T < 0.55$ for $H_T > 575 \text{ GeV}$ and $0.52 < \alpha_T < 0.53$ for $H_T > 775 \text{ GeV}$, as described in Appendix J. This limits the data side-bands to the H_T region defined in the text.

639 The best fit value for the parameter k as obtained from region B is $(2.96 \pm 0.64) \times 10^{-2} \text{ GeV}^{-1}$,
 640 which is taken as the central value for the constraint to be used in the fit. The assumption
 641 that this approach gives an unbiased estimator for k is motivated by the fact that the event
 642 kinematics in the region $0.52 < \alpha_T < 0.55$ and $\alpha_T > 0.55$ are similar. This is validated below.

643 The best fit values for the parameters k as obtained from the regions C_1 , C_2 , and C_3 are used to
 644 estimate a systematic uncertainty on the central value. The best fit values exhibit no strong de-
 645 pendence on the α_T slice, supporting the assumption above that region B provides an unbiased
 646 estimator of k . However, the observed (albeit not significant) variations between the differ-
 647 ent α_T slices in region C are used to determine a systematic uncertainty for k . The weighted
 648 mean and standard deviation of these three measurements is calculated to be $(1.31 \pm 0.26) \times$
 649 10^{-2} GeV^{-1} . The relative error on this measurement is 20%, which is applied to the central
 650 value to give an estimate of the systematic uncertainty.

651 In summary, the aforementioned data side-bands are used to provide a constrained value of
 652 k as input to the likelihood fit, the value of which is determined to be $[-2.96 \pm 0.61(\text{stat.}) \pm$
 653 $0.46(\text{syst.})] \times 10^{-2} \text{ GeV}^{-1}$. These uncertainties are used as penalty terms in the likelihood
 654 model, as described in Sec. 8.

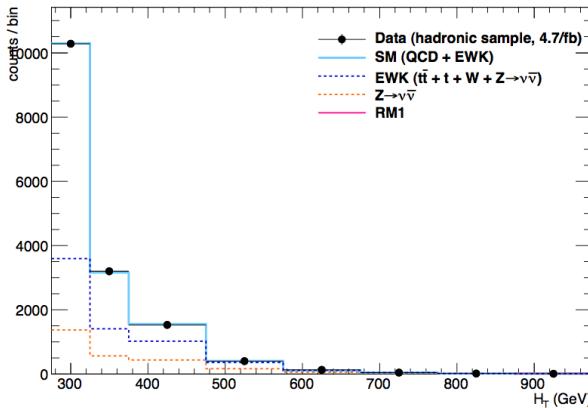


Figure 4: Comparison of the observed yields and SM expectations given by the simultaneous fit in bins of H_T for the side-band region D . No requirement on the number of b jets is made. Shown are the observed event yields in data (black dots with error bars representing the statistical uncertainties) and the expectations given by the simultaneous fit for the $Z \rightarrow \nu\bar{\nu} + \text{jets}$ process (orange dotted-dashed line); the sum of all processes with genuine E_T , which are primarily $t\bar{t}$, $W + \text{jets}$, and $Z \rightarrow \nu\bar{\nu} + \text{jets}$ (dark blue long-dashed line); and the sum of QCD and all aforementioned SM processes (light blue solid line).

655 One further cross-check is performed using the side-band D . In this case, the likelihood fit
 656 is performed for this multi-jet-enriched region, and no constraint is applied to k , which is
 657 allowed to be determined freely by the fit. The fit is performed over full H_T region, i.e. it is
 658 not limited to $275 < H_T < 575 \text{ GeV}$, and no requirement on the number of reconstructed b jets
 659 per event is made. Figure 4 shows the resulting fit, which gives a ML value of $(1.31 \pm 0.09) \times$
 660 10^{-2} GeV^{-1} . This final cross-check provides further supporting evidence for two crucial aspects
 661 of the method described above. First, the fit provides evidence that the exponential function
 662 used in the likelihood model is valid for the entire H_T regime. Second, the ML value for k is
 663 in excellent agreement with the weighted mean obtained from the region C (the two values
 664 can be compared directly in Table 4). This further supports the assumption that region B (with
 665 $0.52 < \alpha_T < 0.55$) can provide an unbiased estimator for k in the signal region ($\alpha_T > 0.55$).

6 Systematic uncertainties on the background estimates

6.1 Using closure tests to identify possible biases

A comprehensive set of closure tests have been performed with data in order to identify possible sources of systematic biases and uncertainties in the analysis. The procedure uses yields measured in one control sample, along with the corresponding translation factors obtained from simulation, to predict the yields observed in another control sample. The level of agreement between the predicted and observed yields is expressed as the ratio $(N_{\text{obs}} - N_{\text{pred}})/N_{\text{pred}}$ while only considering the statistical uncertainty on N_{pred} . No systematic uncertainty is assigned to the prediction. Therefore, the level of closure is defined by the statistical significance of a deviation in the ratio from zero. The ratio is measured for each H_T bin. In this way, these closure tests allow to establish the presence of significant biases or otherwise, and any possible dependence on H_T .

The closure tests described below are used: to determine a systematic uncertainty on the accuracy of the MC simulation to model correctly the α_T acceptance; to validate the use of translation factors determined from simulation and to assign a reasonable systematic uncertainty to each translation factor; and to identify and measure possible biases arising from the use of a b-tagging algorithm. An appropriate systematic uncertainty must be assigned to each translation factor to account for theoretical uncertainties (e.g. on ratios of cross-sections for different processes) and limitations in the accuracy of the MC simulation to model correctly kinematics and instrumental effects (e.g. reconstruction inefficiencies).

6.2 MC modelling of the α_T distribution

For the b-tag analysis, the $\mu + \text{jets}$ and $\mu\mu + \text{jets}$ control samples are defined by selection criteria that do not include an α_T requirement, as described in Sec. 5.4. This is done to increase yields in the control samples and provide a more accurate prediction for the hadronic signal region. This relies on the accurate modelling of the α_T distribution in MC, which is demonstrated by the following closure tests.

The first test uses the $\mu + \text{jets}$ sample defined without an α_T cut to make a prediction of the yields in the $\mu + \text{jets}$ sample defined with the requirement $\alpha_T > 0.55$. As usual, translation factors are determined from MC simulation to make the prediction. The closure test, which compares predicted and observed yields, allows us to understand if the MC is able to model accurately the α_T acceptance for the $W + \text{jets}$ and $t\bar{t}$ processes. The former data sample is recorded with `Mu-HT` triggers and the latter with `HT_AlphaT` triggers. One detail to note is that the closure test is performed on 3.9 fb^{-1} of data, during which time the threshold on the H_T leg of the `Mu-HT` trigger is not higher than 200 GeV, meaning that the closure test can be performed in all eight H_T bins. Figure 31a (Appendix E.1) shows the ratio $(N_{\text{obs}} - N_{\text{pred}})/N_{\text{pred}}$ as a function of H_T . All ratios are statistically compatible with zero and there is no apparent dependence on H_T . The red line represents the constant best fit value across all H_T bins and indicates no significant bias.

Three further closure tests are performed, albeit with smaller yields and less statistical precision, but the same conclusions can be drawn. The three tests are based on a $\mu\mu + \text{jets}$ sample (Fig. 31b), a $\mu + \text{jets}$ sample containing exactly one b-tagged jet (Fig. 31c), and a $\mu + \text{jets}$ sample containing at least two b-tagged jets (Fig. 31d). Again, all ratios are statistically compatible with zero, there is no apparent dependence on H_T , and the constant best fit values are compatible with zero bias.

These closure tests demonstrate that the MC simulation models accurately the α_T distribution,

711 and the use of the larger $\mu + \text{jets}$ and $\mu\mu + \text{jets}$ samples does not introduce any bias into the
 712 background predictions for the hadronic signal region of the b-tag analysis.

713 6.3 Systematic uncertainties on translation factors

714 The final background predictions are given by the simultaneous fit to the yields in the hadronic
 715 signal sample and the three control samples. The fitting procedure also uses the translation
 716 factors obtained from MC per H_T bin per control sample. The simultaneous fit is allowed some
 717 freedom via the (statistical and systematic) uncertainties assigned to each individual translation
 718 factor. Hence, estimating their systematic uncertainties is an important ingredient for the fitting
 719 procedure.

720 Previously, a conservative systematic uncertainty of 30%, constant in H_T , was assigned to the
 721 translation factors used in conjunction with the $\mu + \text{jets}$ samples [4, 5], which was dominated
 722 by the efficiency on the lepton veto (and therefore, indirectly, on the MC modelling of lepton
 723 acceptances and reconstruction efficiencies). Similarly, a conservative value of 40%, again con-
 724 stant in H_T , was assumed for the $\gamma + \text{jets}$ sample, which was dominated by MC modelling
 725 of photon reconstruction efficiency and selection purity, and the theoretical uncertainty on the
 726 ratio of cross sections for the $\gamma + \text{jets}$ and $Z \rightarrow \nu\bar{\nu} + \text{jets}$ processes. However, in this analysis,
 727 closure tests are used to motivate the magnitude of the systematic uncertainties assigned to the
 728 translation factors.

729 Figures 32a and 32b (Appendix E.2) both show the closure tests of using the $\mu + \text{jets}$ sample to
 730 make a prediction for the $\mu\mu + \text{jets}$ sample. The former plot uses “low-stats” $\mu + \text{jets}$ and $\mu\mu$
 731 + jets samples defined by selection criteria that include an α_T requirement and recorded with
 732 HT_AlphaT triggers, which allows the closure to be shown for all eight H_T bins. The latter
 733 plot uses “high-stats” samples defined without an α_T requirement and recorded using Mu_HT
 734 trigger with a H_T threshold of 300 GeV, which allows the closure to be shown only for the six
 735 highest H_T bins. Again, all ratios are statistically compatible with zero, there is no apparent
 736 dependence on H_T , and the red lines indicate constant best fit values that are compatible with
 737 zero bias.

738 The composition of the $\mu + \text{jets}$ and $\mu\mu + \text{jets}$ samples is very different, with the former com-
 739 prising largely $W + \text{jets}$ and $t\bar{t}$ events, and the latter dominated by $Z \rightarrow \nu\bar{\nu} + \text{jets}$ events.
 740 Hence, the closure demonstrated by these tests highlights the ability of the MC simulation to
 741 model accurately the different background compositions of the two samples, the sum of which
 742 provides an estimate for the total SM background found in the hadronic signal region.

743 Two further closure tests are performed, which both use the $\gamma + \text{jets}$ sample to predict yields
 744 in the $\mu\mu + \text{jets}$ sample. These tests are shown in Figs. 32c and 32d, which are the “low-stats”
 745 and “high-stats” tests, respectively. These two data samples provide independent predictions
 746 for the irreducible $Z \rightarrow \nu\bar{\nu} + \text{jets}$ background. Thus, this is an important check on the validity
 747 of using the $\gamma + \text{jets}$ process to predict the $Z \rightarrow \nu\bar{\nu} + \text{jets}$ process, for which a 20% theoretical
 748 uncertainty on the ratio of cross-sections is assumed [20, 21]. Again, all ratios are statistically
 749 compatible with zero, there is no apparent dependence on H_T , and the red lines indicate a
 750 constant best fit value of $\sim 15\%$ that is statistically significant at the 2σ level.

751 Given that the procedure for predicting the number of “control” events in the “signal” region
 752 will be the same in the b-tagged analysis, as a first step it is important to characterise the accu-
 753 racy of the translation factors within a closure test that steps between different b-tag multiplici-
 754 ties and different control samples (muon+jets and dimuon+jets). In order to do this, we split
 755 each control sample into three parts: one containing exactly 0-b-tagged jets, one containing ex-

756 actly 1-b-tagged jet, and the other containing at least 2 b-tagged-jets. We use the same 3.88 fb^{-1}
 757 sample to perform these closure tests as in the classic analysis.

758 Figures 33a, 33b and 33c show how well the translation factors are described within the $\mu + \text{jets}$
 759 sample, when going from zero to one b-tags, zero to more than one b-tags, and one to more
 760 than one b-tags, respectively. The red line shows the result of a one-parameter fit to the ratios.

761 Figures 34a and 34b show the effect of traversing between the $\mu + \text{jets}$ and $\mu\mu + \text{jets}$ samples for
 762 0 and 1 b-tags, respectively. Both these procedures close within the statistical uncertainties.

763 6.4 Motivating a H_T -dependent systematic for the translation factors

764 Figure 5 shows a set of relevant closure tests (open symbols) overlaid on top of grey bands that
 765 represent H_T -dependent systematic uncertainties. The chosen closure tests probe the MC mod-
 766 elling of the α_T distribution and the background composition. Three H_T regions are considered:
 767 $275 < H_T < 575 \text{ GeV}$, $575 < H_T < 775 \text{ GeV}$, and $H_T > 775 \text{ GeV}$. Within each region, all in-
 768 dividual closure tests are used to calculate a weighted mean and variance, using the statistical
 769 uncertainties associated with each closure test. The systematic is defined as 3σ , based on the
 770 weighted variance. A choice of 3σ is rather conservative, but necessary in order to adequately
 771 cover any small biases. The systematics for each region are taken to be fully uncorrelated in the
 772 simultaneous fit, which is a conservative approach given that one can expect large correlations
 773 between adjacent bins (due to comparable kinematics). This method gives values of 6%, 20%
 774 and 39%, which are rounded to 10%, 20% and 40% and used in the simultaneous fit.

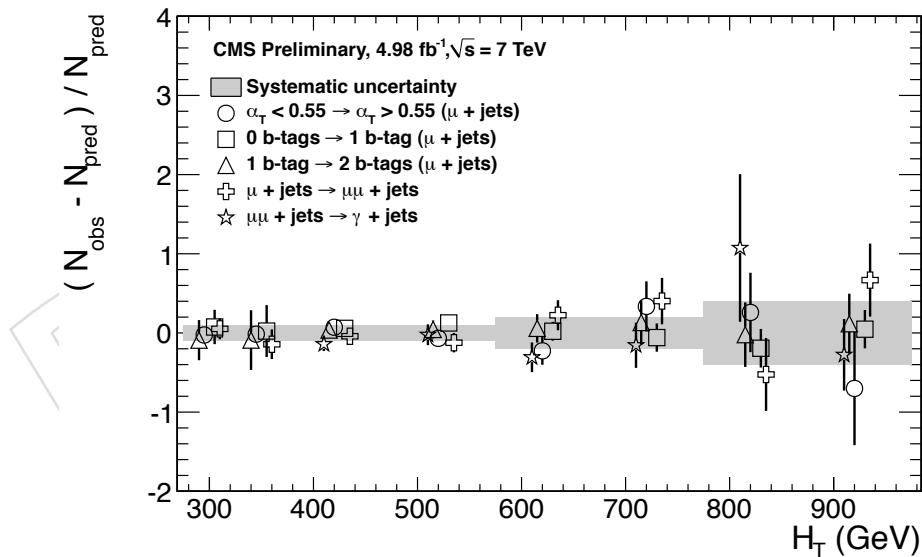


Figure 5: A set of closure tests (open symbols) overlaid on top of grey bands that represent the systematic uncertainties used for three H_T regions in the final simultaneous fit. The solid circles and their errors represent the weighted mean and standard deviation for the five closure tests of each individual H_T bin.

775 6.5 Achieving more accurate n_b estimates specifically for $n_b = 3$

776 6.5.1 The Principle

777 Rather than relying on the yields from the Monte Carlo samples to determine the value of the
 778 translation factors between the single-muon and hadronic sample for $n_b = 3$, it is possible to

779 estimate these yields (and hence the translation factors) to a higher degree of accuracy based
 780 on measurements of the efficiency, mis-tagging rates, and composition of the generator level
 781 b and non- b jet content of the Monte Carlo samples. This method is then used throughout all
 782 b -tag multiplicity bins for simplicity, and to take full advantage of the improved accuracies.

783 6.5.2 The Method

784 Let us factorise out the efficiency of matching reconstruction level jets to generator level jets.
 785 Let n_i represent the yield in MC of events with i reconstructed b jets. This is the quantity we
 786 would like to estimate to a higher degree of accuracy for $i = 3$. Let n_{bs} represent the yield
 787 in MC of events with b generated b jets which are all matched to reconstructed jets and also
 788 with s generated non- b jets which are all matched to reconstructed jets. The latter definition
 789 includes u, d, s, c, g and τ jets such that the mis-tagging is averaged over these flavours. Let ϵ
 790 represent the b jet efficiency averaged over jet- p_T and η , and m represent this flavour-averaged
 791 b jet mis-tagging rate, also averaged over jet- p_T and η . These three quantities are measured
 792 individually for each H_T bin used in this analysis, and separately for the hadronic and muon
 793 selections (primarily to ensure the correct admixture for the mis-tagging rate).

794 In order to measure n_{bs} , ϵ and m from the MC samples, truth level matching is required. For
 795 every event which is considered in building up these rates, we loop over all generator level jets
 796 and match to the closest reconstructed level jet in $\Delta R = \sqrt{\Delta\eta^2 + \Delta\phi^2}$ space which has both
 797 $\Delta R < 0.5$ and p_T^{gen} / p_T^{reco} within $\pm 50\%$.

798 In the syntax below, $(b \geq i, s \geq j)$ represents events with the pairing specified. Writing down
 799 the event level reconstructed b jet content in terms of the matched b jet and non- b jet content,
 800 we have:

$$n_0 = \sum_{(b \geq 0, s \geq 0)} n_{bs} \times (1 - \epsilon)^b \times (1 - m)^s \quad (10)$$

$$n_1 = \sum_{(b \geq 0, s \geq 1)} n_{bs} \times (1 - \epsilon)^b \times sm(1 - m)^{s-1} + \sum_{(b \geq 1, s \geq 0)} n_{bs} \times (1 - m)^s \times b\epsilon(1 - \epsilon)^{b-1} \quad (11)$$

$$\begin{aligned} n_2 = & \sum_{(b \geq 0, s \geq 2)} n_{bs} \times (1 - \epsilon)^b \times \frac{s(s-1)}{2} m^2 (1 - m)^{s-2} \\ & + \sum_{(b \geq 2, s \geq 0)} n_{bs} \times (1 - m)^s \times \frac{b(b-1)}{2} \epsilon^2 (1 - \epsilon)^{b-2} \\ & + \sum_{(b \geq 1, s \geq 1)} n_{bs} \times b\epsilon(1 - \epsilon)^{b-1} \times sm(1 - m)^{s-1} \end{aligned} \quad (12)$$

$$\begin{aligned}
n_3 = & \sum_{(b \geq 0, s \geq 3)} n_{bs} \times (1 - \epsilon)^b \times \frac{s(s-1)(s-2)}{6} m^3 (1 - m)^{s-3} \\
& + \sum_{(b \geq 3, s \geq 0)} n_{bs} \times (1 - m)^s \times \frac{b(b-1)(b-2)}{6} \epsilon^3 (1 - \epsilon)^{b-3} \\
& + \sum_{(b \geq 1, s \geq 2)} n_{bs} \times b\epsilon (1 - \epsilon)^{b-1} \times \frac{s(s-1)}{2} m^2 (1 - m)^{s-2} \\
& + \sum_{(b \geq 2, s \geq 1)} n_{bs} \times sm (1 - m)^{s-1} \times \frac{b(b-1)}{2} \epsilon^2 (1 - \epsilon)^{b-2}
\end{aligned} \tag{13}$$

801 If we look at Equation 13, we know that events with $b > 2$ in the Standard Model are relatively
 802 suppressed and so we would expect that the biggest contribution to this bin comes from events
 803 with $b = 2$ and one mis-tagged jet.

804 6.5.3 Closure Tests

805 In order to validate the procedure, we can use Equations 10 to 12 to predict the MC yields in
 806 these bins and compare the two numbers. These results for the muon selection are summarised
 807 in Table 5.

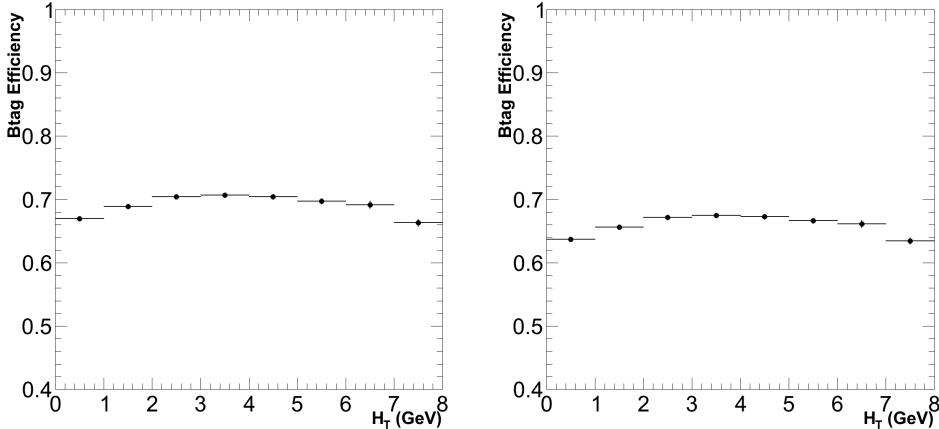
Table 5: Comparing MC yields in the muon selection with the estimate from Equation 13. The numbers are normalised to 100pb^{-1} and are without an alphaT cut. No SF_b , SF_c or SF_{light} corrections are applied.

| H_T Bin (GeV) | 275–325 | 325–375 | 375–475 | 475–575 |
|-----------------|-------------------|------------------|------------------|------------------|
| Pred Btag = 0 | 135.204 | 71.15 | 38.77 | 17.52 |
| Obs Btag = 0 | 137.49 ± 3.47 | 72.59 ± 2.53 | 39.55 ± 0.26 | 17.94 ± 0.18 |
| Pred Btag = 1 | 44.26 | 22.71 | 13.49 | 6.74 |
| Obs Btag = 1 | 40.78 ± 1.29 | 20.76 ± 0.77 | 13.37 ± 0.22 | 6.21 ± 0.16 |
| Pred Btag = 2 | 19.06 | 10.47 | 5.94 | 3.19 |
| Obs Btag = 2 | 19.83 ± 0.54 | 10.72 ± 0.36 | 6.03 ± 0.18 | 3.34 ± 0.13 |
| Pred Btag > 2 | 1.36 | 0.67 | 0.45 | 0.31 |
| Obs Btag > 2 | 1.44 ± 0.10 | 0.78 ± 0.07 | 0.58 ± 0.06 | 0.30 ± 0.06 |
| H_T Bin (GeV) | 575–675 | 675–775 | 775–875 | $875-\infty$ |
| Pred Btag = 0 | 7.64 | 3.71 | 1.79 | 2.14 |
| Obs Btag = 0 | 7.82 ± 0.12 | 3.77 ± 0.09 | 1.81 ± 0.05 | 2.19 ± 0.06 |
| Pred Btag = 1 | 3.11 | 1.34 | 0.63 | 0.74 |
| Obs Btag = 1 | 2.95 ± 0.11 | 1.27 ± 0.09 | 0.61 ± 0.05 | 0.64 ± 0.05 |
| Pred Btag = 2 | 1.58 | 0.63 | 0.300 | 0.29 |
| Obs Btag = 2 | 1.61 ± 0.12 | 0.67 ± 0.06 | 0.30 ± 0.04 | 0.32 ± 0.04 |
| Pred Btag > 2 | 0.19 | 0.07 | 0.04 | 0.05 |
| Obs Btag > 2 | 0.18 ± 0.03 | 0.06 ± 0.02 | 0.03 ± 0.01 | 0.03 ± 0.01 |

808 6.6 Estimating the impact of the b -tagging and mis-tagging scale factors

809 As provided by the b-tag POG [22], certain corrections need to be made to both the b -tagging
 810 efficiency and mis-tagging rates in order to match the factors underlying the Monte Carlo sam-
 811 ples used to the distributions seen in data. For the b -tagger used in this analysis (CSV), the
 812 relevant dependences are given by:

Figure 6: MC measured b -tagging efficiency in H_T bins: uncorrected (left) and corrected (right) using the b-POG scale factors. Note that the x -axis binning corresponds to the H_T ranges used in this analysis e.g. bin 1 = 275 - 325 GeV etc.



$$SF_b(p_T) = 0.6981 \times \frac{1.0 + 0.414063 \times p_T}{1.0 + 0.300155 \times p_T} \quad (14)$$

$$SF_c(p_T) = SF_b(p_T) \quad (15)$$

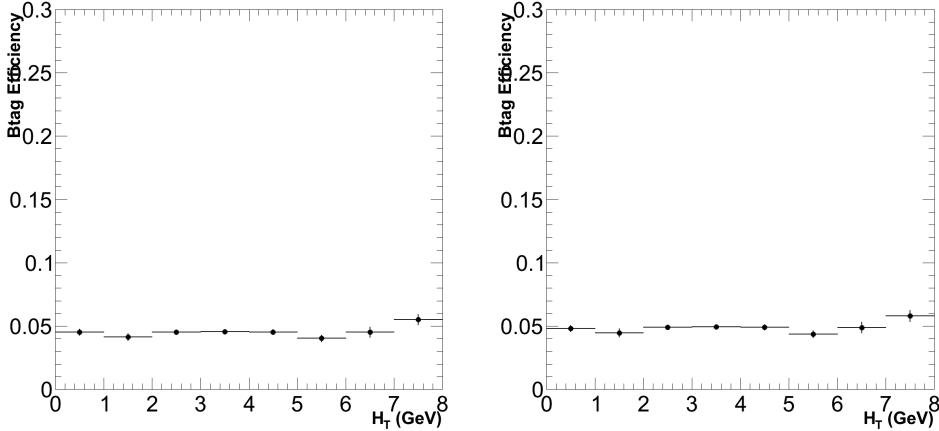
$$\begin{aligned} SF_{light}(p_T) = & 1.04318 \\ & + (8.48162e^{-4} \times p_T) \\ & + (-2.5795e^{-6} \times p_T^2) \\ & + (1.64156e^{-9} \times p_T^3) \end{aligned} \quad (16)$$

where $SF_b(p_T)$ is the b -tag efficiency scale factor, and $SF_c(p_T)$, $SF_{light}(p_T)$ are the mis-tagging scale factors. Note that these scale factors are only applied to the MC, in order to correct the relevant MC yields, and hence provide corrected translation factors. To estimate the impact of these corrections, the following procedure is followed:

6.7 Procedure

Since Equations 10 to 13 are used to predict the yields, corrected values of ϵ and m are required in order to provide corrected MC yields. These corrected efficiencies and mis-tagging rates are determined using a procedure which works on an event-by-event basis. Generator level b jets matched to reconstruction level b jets which fire the b -tagger are weighted by $SF_b(p_T)$ and similarly for the mis-tagging rates, separately for c -quarks and light quarks/gluons/ τ -jets. The efficiency and mis-tagging rates are then estimated in the respective H_T bins by dividing these reweighted numbers by the total number of generator level b jets and non- b jets matched to reconstruction level jets respectively, in the same H_T bin. Both uncorrected and corrected efficiency and mis-tagging rates are shown in Figures 6 and 7 respectively, as a function of H_T bin.

Figure 7: MC measured mis-tagging rates in H_T bins: uncorrected (left) and corrected (right) using the b-POG scale factors. Note that the x -axis binning corresponds to the H_T ranges used in this analysis e.g. bin 1 = 275 - 325 GeV etc.



6.7.1 Scale Factor Uncertainties

The b-POG also provide uncertainties on SF_b , SF_c and SF_{light} . In order to estimate the size of the systematic uncertainty on our translation factors (from the scale factors), we take the corrected translation factors as evaluated using the method highlighted above, and vary the scale factors up and down separately by their uncertainties (in a correlated fashion). The uncertainties provided for SF_b are summarised in Table 6. The uncertainties for SF_c are twice those for SF_b . The uncertainties for SF_{light} are parameterised in Equation 17

Table 6: Uncertainties on SF_b using the CSV tagger, with $30 < p_T < 670$ GeV $|\eta| < 2.4$. The uncertainty for $p_T > 670$ GeV is twice the value at 670 GeV.

| Jet p_T (GeV) | 30-40 | 40-50 | 50-60 | 60-70 | 70-80 | 80-100 | 100-120 |
|-----------------------------------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|
| SF _b uncertainty \pm | 0.029567 | 0.0295095 | 0.0210867 | 0.0219349 | 0.0227033 | 0.0204062 | 0.0185857 |
| Jet p_T (GeV) | 120-160 | 160-210 | 210-260 | 260-320 | 320-400 | 400-500 | 500-670 |
| SF _b uncertainty \pm | 0.0256242 | 0.0383341 | 0.0409675 | 0.0420284 | 0.0541299 | 0.0578761 | 0.0655432 |

$$\begin{aligned}
 SF_{light}- &= ((0.962627 + (0.000448344 \times p_T)) + (-1.25579 \times 10^{-6} \times (p_T^2))) \\
 &\quad + (4.82283 \times 10^{-10} \times (p_T^3)) \\
 SF_{light}+ &= ((1.12368 + (0.00124806 \times p_T)) + (-3.9032 \times 10^{-6} \times (p_T^2))) \\
 &\quad + (2.80083 \times 10^{-9} \times (p_T^3))
 \end{aligned} \tag{17}$$

6.7.2 Summary of results

Corrected translation factors for all b -tag multiplicity bins, and the uncertainties generated from varying the scale-factors up and down by their uncertainties are shown in Table 7. The changes in the yields for the $n_b = 3$ bin are also shown, in Table 8, to illustrate that the translation factors are relatively insensitive to the uncertainty, given it is a ratio of the hadronic and muon yields.

Table 7: Corrected translation factors and the impact of the uncertainties of the b-POG data/MC scale factors on these translation factors, in bins of H_T and b -tag multiplicity. (Also quoted are the statistical uncertainties.)

| $N_{b\text{-tag}}$ | 275–325 | 325–375 | 375–475 | 475–575 |
|--------------------|-----------------------------------------|-----------------------------------------|-----------------------------------------|-----------------------------------------|
| = 0 | $2.7410^{+0.0088}_{-0.0088} \pm 0.1624$ | $2.3320^{+0.0037}_{-0.0036} \pm 0.2135$ | $0.4463^{+0.0004}_{-0.0004} \pm 0.0050$ | $0.3387^{+0.0000}_{-0.0001} \pm 0.0063$ |
| = 1 | $1.7180^{+0.0322}_{-0.0304} \pm 0.0730$ | $1.6645^{+0.0320}_{-0.0303} \pm 0.1040$ | $0.3497^{+0.0066}_{-0.0063} \pm 0.0165$ | $0.2771^{+0.0056}_{-0.0054} \pm 0.0248$ |
| = 2 | $1.2897^{+0.0088}_{-0.0088} \pm 0.0838$ | $1.1195^{+0.0127}_{-0.0124} \pm 0.1123$ | $0.2226^{+0.0031}_{-0.0030} \pm 0.0124$ | $0.1819^{+0.0030}_{-0.0029} \pm 0.0154$ |
| ≥ 3 | $1.1755^{+0.0021}_{-0.0016} \pm 0.0519$ | $1.1064^{+0.0012}_{-0.0010} \pm 0.0738$ | $0.2366^{+0.0013}_{-0.0012} \pm 0.0105$ | $0.2116^{+0.0012}_{-0.0011} \pm 0.0126$ |
| $N_{b\text{-tag}}$ | 575–675 | 675–775 | 775–875 | 875– ∞ |
| = 0 | $0.2813^{+0.0000}_{-0.0000} \pm 0.0087$ | $0.2022^{+0.0011}_{-0.0011} \pm 0.0102$ | $0.1633^{+0.0002}_{-0.0002} \pm 0.0116$ | $0.0942^{+0.0002}_{-0.0002} \pm 0.0075$ |
| = 1 | $0.2129^{+0.0059}_{-0.0056} \pm 0.0277$ | $0.2137^{+0.0018}_{-0.0017} \pm 0.0505$ | $0.1324^{+0.0042}_{-0.0040} \pm 0.0298$ | $0.0831^{+0.0016}_{-0.0014} \pm 0.0156$ |
| = 2 | $0.1287^{+0.0025}_{-0.0024} \pm 0.0183$ | $0.2084^{+0.0015}_{-0.0014} \pm 0.0530$ | $0.0688^{+0.0017}_{-0.0016} \pm 0.0293$ | $0.0622^{+0.0010}_{-0.0010} \pm 0.0264$ |
| ≥ 3 | $0.1432^{+0.0000}_{-0.0000} \pm 0.0153$ | $0.2760^{+0.0003}_{-0.0003} \pm 0.0414$ | $0.0653^{+0.0008}_{-0.0008} \pm 0.0229$ | $0.0474^{+0.0002}_{-0.0001} \pm 0.0143$ |

Table 8: Corrected MC yields in the hadronic and muon selections, and the impact of the uncertainties of the b-POG data/MC scale factors on these yields, in bins of H_T for $n_b = 3$. (Also quoted are the statistical uncertainties.)

| Sample | 275–325 | 325–375 | 375–475 | 475–575 |
|----------|------------------------------------------|-----------------------------------------|------------------------------------------|------------------------------------------|
| Hadronic | $12.7839^{+1.7942}_{-1.9359} \pm 0.3400$ | $5.0463^{+0.7359}_{-0.7979} \pm 0.2217$ | $4.9250^{+0.7470}_{-0.8137} \pm 0.1930$ | $3.1325^{+0.4880}_{-0.5320} \pm 0.1666$ |
| Muon | $10.8750^{+1.5097}_{-1.6293} \pm 0.3809$ | $4.5610^{+0.6694}_{-0.7259} \pm 0.2289$ | $20.8193^{+3.0588}_{-3.3143} \pm 0.4424$ | $14.8043^{+2.2330}_{-2.4217} \pm 0.3926$ |
| Sample | 575–675 | 675–775 | 775–875 | 875– ∞ |
| Hadronic | $1.2912^{+0.1998}_{-0.2167} \pm 0.1256$ | $0.9751^{+0.1529}_{-0.1661} \pm 0.1371$ | $0.1308^{+0.0196}_{-0.0210} \pm 0.0448$ | $0.1177^{+0.0188}_{-0.0203} \pm 0.0344$ |
| Muon | $9.0189^{+1.3946}_{-1.5125} \pm 0.4033$ | $3.5331^{+0.5574}_{-0.6061} \pm 0.1837$ | $2.0037^{+0.3222}_{-0.3505} \pm 0.1467$ | $2.4859^{+0.4033}_{-0.4354} \pm 0.1891$ |

6.8 Procedure for signal scans

For the signal scans used in this analysis, the relatively small number of events generated per point (10,000) force the use of a different procedure to that outlined for the SM MC samples above. Furthermore, the use of fast-simulation in the reconstruction introduces an extra set of scale-factor corrections, to be applied simultaneously with those correcting the full-simulation to the data.

For these scenarios, an event-by-event weight is applied, as recommended by the b-POG. This weight depends on the generator level content, as well as the b -tagged status of the matched reconstruction level jets. As an example, consider an event with 2 generator b jets matched to 2 reconstruction level jets, of which only 1 fires the b -tagger, and 2 generator non- b jets, also matched to 2 reconstruction level jets, of which also only 1 fires the b -tagger. The probability for this to happen is given by:

$$\begin{aligned}
 P &= \epsilon(p_T^{jet1}, \eta^{jet1}) \\
 &\times (1 - \epsilon(p_T^{jet2}, \eta^{jet2})) \\
 &\times m(p_T^{jet3}, \eta^{jet3}, ID^{jet3}) \\
 &\times (1 - m(p_T^{jet4}, \eta^{jet4}, ID^{jet4}))
 \end{aligned} \tag{18}$$

where $\epsilon(p_T, \eta)$ and $m(p_T, \eta)$ are the b -tagging efficiency and mis-tagging rates respectively, as measured in the SM MC samples in a particular selection i.e. hadronic or muon, in a particular H_T bin. The weight to correct this event is therefore given by:

$$w = \frac{SF_b\epsilon \times (1 - SF_b\epsilon) \times SF_{c,light}m \times (1 - SF_{c,light}m)}{\epsilon \times (1 - \epsilon) \times m \times (1 - m)} \quad (19)$$

where the relevant dependences are implied. Once all events have been reweighted this way and the correct normalisation applied, the yields in each b -tag bin represent the corrected MC yields. The scale factors in this case represent the scale factors from Equations 14 to 16 multiplied by the fast-sim to full-sim correction factors. The fast-sim to full-sim correction factors are summarised in AN-2012-175. The central value for these corrections is the same for all SMS samples, and is taken as the ratio between the efficiency/mis-tagging rates of a ttbar full-sim to ttbar fastsim sample. Similar to the differences between fullsim and data, the b -tagging efficiency is higher in fastsim than in fullsim (thus needs a correction less than one), and the mistag rate is lower (thus needs a correction larger than one). To demonstrate the procedure above, the central value of these corrections is applied on the T1bbbb SMS sample. Figure 8 summarises the change in yields averaged over the SMS phase space (and for all H_T bins) in bins of b -tag multiplicity, comparing the fastsim to fullsim to data correction, to no correction at all.

7 Signal efficiency and uncertainty

Appendix M contains figures showing the signal efficiency of the event selection for the hadronic signal region in the plane of the CMSSM and various simplified models. The production and decay modes of the simplified models under consideration are summarised in Table 9. The signal efficiencies are integrated over all eight H_T bins, but are shown separately for each of the four b -tag multiplicity bins. Figure 61 shows the signal efficiencies in the plane of the CMSSM. Figures 62 (T1), 63 (T2), 64 (T2tt), 65 (T2bb), 66 (T1tttt), and 67 (T1bbbb) show the signal efficiencies for the simplified models. Figure 68 compares the signal efficiency of the event selection criteria for the hadronic signal region and single muon control samples, for events containing exactly one, two and at least three b -tags per event for the T1tttt simplified model.

Table 9: Production and decay modes for various simplified models.

| Model | Production and decay modes | Figure showing efficiency |
|--------|--------------------------------------------------------------------------------|---------------------------|
| T1 | $\tilde{g}\tilde{g} \rightarrow q\bar{q}\tilde{\chi}^0 q\bar{q}\tilde{\chi}^0$ | 62 |
| T2 | $\tilde{q}\bar{q} \rightarrow q\tilde{\chi}^0 q\tilde{\chi}^0$ | 63 |
| T2tt | $\tilde{t}\bar{t} \rightarrow t\tilde{\chi}^0 t\tilde{\chi}^0$ | 64 |
| T2bb | $\tilde{b}\bar{b} \rightarrow b\tilde{\chi}^0 b\tilde{\chi}^0$ | 65 |
| T1tttt | $\tilde{g}\tilde{g} \rightarrow t\bar{t}\tilde{\chi}^0 t\bar{t}\tilde{\chi}^0$ | 66 |
| T1bbbb | $\tilde{g}\tilde{g} \rightarrow b\bar{b}\tilde{\chi}^0 b\bar{b}\tilde{\chi}^0$ | 67 |

The systematic uncertainties on the signal event yield are estimated following in general the recipes discussed in [4, 5] and are split into two parts: theoretical uncertainties on the predicted cross section of the different production processes (squark-squark, squark-gluino, gluino-gluino) and experimental uncertainties on the integrated luminosity and on the selection efficiency.

Appendix F summarises the studies undertaken to establish the experimental systematic uncertainty on the expected signal acceptance times efficiency of the hadronic selection for the CMSSM and various simplified models. Uncertainties on the parton distribution functions [4], the luminosity measurement, the jet energy scale, and various cuts used in the selection are considered, namely: the MHT/MET cleaning filter, the dead ECAL cleaning filter, and the lepton/photon event vetoes.

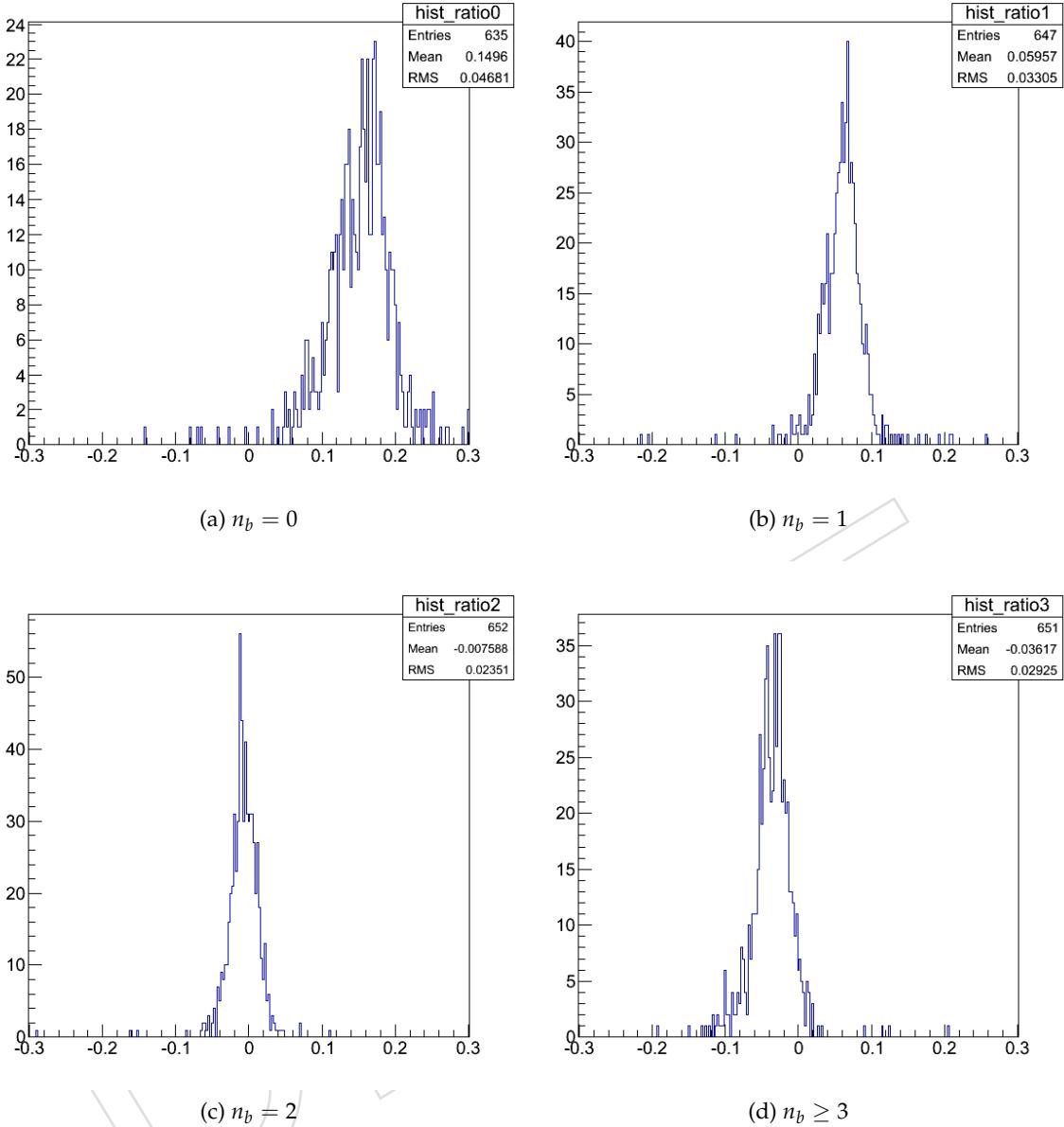


Figure 8: Comparison of the corrected (fastsim to fullsim to data) to uncorrected yields of the T1bbbb SMS sample for different b-tag multiplicities. The relative yields are shown in each case. Each b-tag multiplicity shown is averaged over the respective $H_T > 275$ GeV bins. Each entry in the histograms represents a point in the SMS signal scan phase space.

- 890 The aim is to make the interpretation of the analysis result easy to understand. Rather than
 891 trying to estimate the level of systematic that is applicable point-by-point in a model space,
 892 general behaviours are considered across the model space, and constant, conservative system-
 893 atics are estimated instead for appropriate *regions* in the model space. The same systematic
 894 value is used for all bins in the signal region, and the most conservative one is chosen.
 895 In the case of the CMSSM, the various systematic uncertainties considered can vary consider-
 896 ably in the $m_0 - m_{1/2}$ plane due to the dependence on the mass hierarchy of the squarks and
 897 gluinos (and charginos and neutralinos). The production and decay modes that are realised

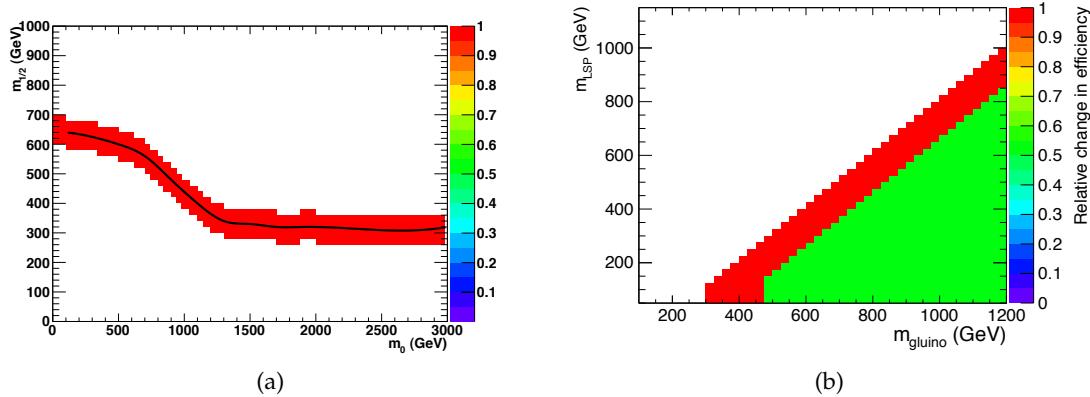


Figure 9: (a) Definition of a $m_{1/2} \pm 60$ GeV band around the observed limit in the CMSSM plane. (b) Definition of regions "near" to (red) and "far" from (green) the diagonal, to categorise models into ones with and without compressed spectra, respectively.

in the $m_0 - m_{1/2}$ plane are varied, and so only those models in the vicinity of the observed limit are considered when estimating a systematic uncertainty on the analysis efficiency times acceptance. Only models that fall within a band of $m_{1/2} \pm 60$ GeV around the observed limit are considered, as illustrated in Fig. 9a. This width is determined by considering the 68% CL band around the expected limit, which is of the order ~ 50 GeV.

In the case of the simplified models, the model space is divided into two regions near to and far from the diagonal, where "far" is realised by the condition:

$$m_{\text{sq}}(m_{\text{gl}}) - m_{\text{LSP}} > 350 \text{ GeV} \quad \& \& \quad m_{\text{sq}}(m_{\text{gl}}) > 475 \text{ GeV}$$

and all points failing this condition are classified as "near". This is done to categorise models into ones with compressed spectra and ones without, as illustrated in Fig. 9b. Generally, models with compressed spectra have larger uncertainties on the efficiency times acceptance of the signal selection, due to their final states having lower visible (H_T) and invisible (\cancel{H}_T) hadronic energies.

Figures 36a and 36b show the relative change in the signal efficiency times acceptance in the CMSSM plane when varying the energy of all jets in an event up or down according to a p_T - and η -dependent jet energy scale uncertainty (i.e. vary the event scale up and down), as recommended by the Jet POG. The "island" of larger relative changes at high values of m_0 and $m_{1/2}$ is due to the dominant process being direct electroweak production of charginos and neutralinos, resulting in short cascades and little visible hadronic energy. Figure 36c shows correspondingly the distribution of the relative change in the signal efficiency times acceptance for all model points that fall within the band $m_{1/2} \pm 60$ GeV around the observed limit. The large number of entries in the central bin around zero is due to points that have low efficiency with few MC events passing the full selection and no events moving in or out of acceptance.⁴ The 68th percentile for the distribution, as indicated by the upper bound on the red shaded area in Fig. 36c, is 5.2%. This is taken to be the systematic uncertainty and is listed in Table 10.

⁴The low efficiency occurs at high m_0 and can lead to very MC events passing all selection criteria for many points in the plane. A loose requirement of at least 25 events (20% statistical uncertainty) is required in order to have an entry in the 1D distribution. The low efficiency (and hence large statistical uncertainties) also explain the apparently large fluctuations observed between adjacent points. This is also an additional strong reason why conservative systematics estimated on many models with similar topologies is better than a point-by-point approach.

Similarly, Figs. 37 and 38 show the relative change in the signal efficiency times acceptance for various simplified models when varying the event scale up and down. Figs. 39 and 40 show correspondingly the distributions of the relative change in the signal efficiency times acceptance for all model points that fall within a band "near" to and "far" from the diagonal. The resulting systematic uncertainties for the various simplified models are determined by the 68th percentiles for these distributions, as indicated by the upper bounds on the red shaded areas in the figures. The uncertainties are listed in Tables 11 ("near") and 12 ("far"). The resulting systematic uncertainty is largest near the diagonal, where the models are defined by compressed spectra. Far from the diagonal, the mass splitting is sufficiently large that the probability of jets moving out of acceptance due to jet energy scale corrections is negligible.

Figure 41a compares the MHT/MET distributions obtained from data and MC. The distributions are produced following the selection criteria that define the $\mu + \text{jets}$ sample, up to and including the definition of the H_T of the event, which is required to satisfy $H_T > 375 \text{ GeV}$. No α_T cut or event cleaning filters (as described in Sec. 2.4) are applied. This sample largely comprises events from the $W + \text{jets}$ and $t\bar{t}$ processes, with negligible contamination from QCD multi-jet events. These backgrounds have significant, genuine \cancel{E}_T and so the efficiency of the MHT/MET cut is expected to be high, and is indeed observed to be $\sim 80\%$. Figure 41b shows the ratio of the two data and MC distributions, and so reflects the accuracy of the MC modelling of the cut. The observed discrepancy for a cut value of $\text{MHT}/\text{MET} < 1.25$ is 16%. This discrepancy is taken to represent the uncertainty on the MC modelling of this cut for processes with significant, genuine \cancel{E}_T , including the SUSY models presented in this note.

Figures 42a and 43 show the fraction of expected signal yield that is rejected (i.e. signal inefficiency) by the MHT/MET cleaning cut, for the CMSSM and various simplified models, respectively. Figs. 42b, 44 and 45 show correspondingly the distributions of the signal inefficiency for all model points that fall within a $\pm 50 \text{ GeV}$ band around the observed limit in the CMSSM plane, and "near" to and "far" from the diagonal in the simplified models, respectively. For the simplified models, larger inefficiencies are observed near the diagonal, where the models have compressed spectra. Thus, the visible and invisible hadronic energies (H_T and \cancel{H}_T) are small, jet energy resolutions are worse, and therefore the MHT/MET cut value becomes effectively tighter. Table 38 summarises conservative estimates of the signal inefficiencies for all six simplified models and both the near and far regions.⁵

Table 10: Sources of experimental systematic uncertainties on the signal efficiency time acceptance for the CMSSM.

| Name | Systematic (%) |
|-------------------------------|----------------|
| Parton distribution functions | 10.0 |
| Integrated luminosity [23] | 2.2 |
| FastSim acceptance | 5.0 |
| Jet energy scale | 5.2 |
| MHT/MET selection | 2.8 |
| Dead ECAL selection | 0.6 |
| Lepton/photon vetoes | 3.6 |
| b jet corrections | 8.5 |
| Total systematic uncertainty | 15.7 |

The systematic uncertainty is taken as the product of the uncertainty on the MC modelling of the MHT/MET cut and the fraction of expected signal yield that is rejected by the MHT/MET cleaning cut. The resulting values are quoted in Tables 10, 11 and 12.

⁵"Conservative" is defined as the mean + 2*RMS value of the distributions.

Table 11: Conservative estimates of systematic uncertainty (%) for jet energy scale and the listed selection cuts when considering model points in the region near to the diagonal (i.e. small mass splitting and compressed spectra) for various simplified models. The quoted uncertainties are added in quadrature with the following to obtain the total: pdf (10%), integrated luminosity (2.2%) [23] and FastSim acceptance (5%).

| Model | JES | MHT/MET | Dead ECAL | Lepton Veto | b jet corrections | Total |
|--------|------|---------|-----------|-------------|-------------------|-------|
| T1 | 10.5 | 1.8 | 0.2 | - | 8.5 | 17.8 |
| T2 | 8.4 | 0.6 | 0.1 | - | 8.5 | 16.5 |
| T2tt | 9.5 | 3.3 | 0.4 | 0.8 | 1.5 | 15.3 |
| T2bb | 9.2 | 0.6 | 0.1 | 0.3 | 1.5 | 14.7 |
| T1tttt | - | - | - | - | - | - |
| T1bbbb | 11.5 | 1.8 | 0.3 | 0.5 | 1.5 | 16.4 |

Table 12: Conservative estimates of systematic uncertainty (%) for jet energy scale and the listed selection cuts when considering model points in the region far from the diagonal (i.e. large mass splitting) for various simplified models. The quoted uncertainties are added in quadrature with the following to obtain the total: pdf (10%), integrated luminosity (2.2%) [23] and FastSim acceptance (5%).

| Model | JES | MHT/MET | Dead ECAL | Lepton Veto | b-tag corrections | Total |
|--------|-----|---------|-----------|-------------|-------------------|-------|
| T1 | 0.6 | 0.5 | 0.2 | - | 14.0 | 18.1 |
| T2 | 0.6 | 0.1 | 0.1 | - | 14.0 | 18.1 |
| T2tt | 2.3 | 0.5 | 0.2 | 0.6 | 3.5 | 12.2 |
| T2bb | 1.2 | 0.2 | 0.1 | 0.3 | 3.5 | 12.0 |
| T1tttt | 1.5 | 5.3 | 0.5 | 1.4 | 3.5 | 13.2 |
| T1bbbb | 1.2 | 0.2 | 0.2 | 0.4 | 3.5 | 12.0 |

Similarly, a study comparing the efficiency of the dead ECAL cut for data and MC was performed, using the same sample of $\mu + \text{jets}$ events. In this case, the data–MC discrepancy is 1.8%, which is again taken to represent the uncertainty on the MC modelling of this cut for processes with significant, genuine \cancel{E}_T .

Figures 46a and 47 show the fraction of expected signal yield that is rejected by the dead ECAL cleaning cut, for the CMSSM and various simplified models, respectively. Figs. 46b, 48 and 49 show correspondingly the distributions of the signal inefficiency for all model points that fall within a ± 50 GeV band around the observed limit in the CMSSM plane, and “near” to and “far” from the diagonal in the simplified models, respectively. Larger inefficiencies are observed for models with higher jet multiplicity final-states, such as those arising from gluino-gluino pair-production or simplified models with decays via top quarks, due to the higher probability of an event having at least mis-measured jet pointing to a region of dead ECAL cells or the transition between the ECAL barrel and end-cap regions. This effect is exaggerated further for models with compressed spectra.

Again, the systematic uncertainty is taken as the product of the uncertainty on the MC modelling of the dead ECAL filter and the fraction of expected signal yield that is rejected by the dead ECAL cut. The resulting values are quoted in Tables 10, 11 and 12.

Figures 50a and 51 show the fraction of expected signal yield that is rejected by the lepton and photon vetoes, for the CMSSM and various simplified models, respectively. The vetoes are applied immediately after a filter with identical logic but based on truth information. Hence, any observed inefficiency represents the fraction of signal events that should not be vetoed. Figs. 50b, 52 and 53 show correspondingly the distributions of the signal inefficiency for all

model points that fall within a ± 50 GeV band around the observed limit in the CMSSM plane, and "near" to and "far" from the diagonal in the simplified models, respectively. Generally, the inefficiencies are observed to be very small, and are taken directly as a systematic.

Tables 10, 11 and 12 summarise all the aforementioned systematic uncertainties on the signal efficiency times acceptance. Generally, the uncertainties are dominated by pdf and jet energy scale uncertainties. The total for the CMSSM is $\sim 14\%$; the totals for the simplified models in the region far from the diagonal (i.e. with large mass splittings) are $15 - 18\%$; and similarly in the region near to the diagonal (i.e. compressed spectra), the total uncertainties are $12 - 18\%$. These uncertainties are all included in the limit calculation.

7.1 SMS-dependent fastsim uncertainty

Whilst the central value of the fastsim to fullsim corrections are independent of the SMS sample used, the uncertainties are SMS sample dependent, taken as half of the difference between the ttbar fastsim efficiency/mis-tagging rate and the efficiency/mis-tagging rate as measured in the SMS sample. These uncertainties are added in quadrature to the fullsim to data uncertainties specified above, and are then used to modify the scale factors up and down (in a correlated fashion). In order to estimate a systematic uncertainty based on this, the absolute values of the relative changes in yields for each of the 32 (8 H_T bins and 4 n_b bins) bins for each SMS point are calculated. Based on these values, the 68% quantile for each SMS point is shown in Figure 10 for the T1bbbb SMS. The final systematic uncertainty is again determined in both a near and far region, where the 68% quantile is used averaged over all SMS points. The cumulative histograms for T1bbbb are shown in Figure 10 leading to a systematic uncertainty of 1.5% in the near region and 3.5% in the far region. A similar procedure is followed for the other SMS models, the results of which are also summarised in Tables 11 and 12 for the near and far regions respectively.

8 Likelihood model

8.1 Hadronic sample

Let N be the number of bins of H_T , which need not have equal width. Let n^i represent the number of events observed satisfying all selection requirements in each H_T bin i . Then the likelihood of the observations is written this way:

$$L_{hadronic} = \prod_i \text{Pois}(n^i | b^i + s^i) \quad (20)$$

where b^i represents the expected Standard Model background in bin i , s^i represents the expected number of signal events in bin i , and Pois represents the Poisson distribution. It is assumed that:

$$b^i \equiv \text{EWK}^i + \text{QCD}^i \quad (21)$$

where EWK^i is the expected yield of electroweak events in bin i , and QCD^i is the expected yield of QCD events in bin i .

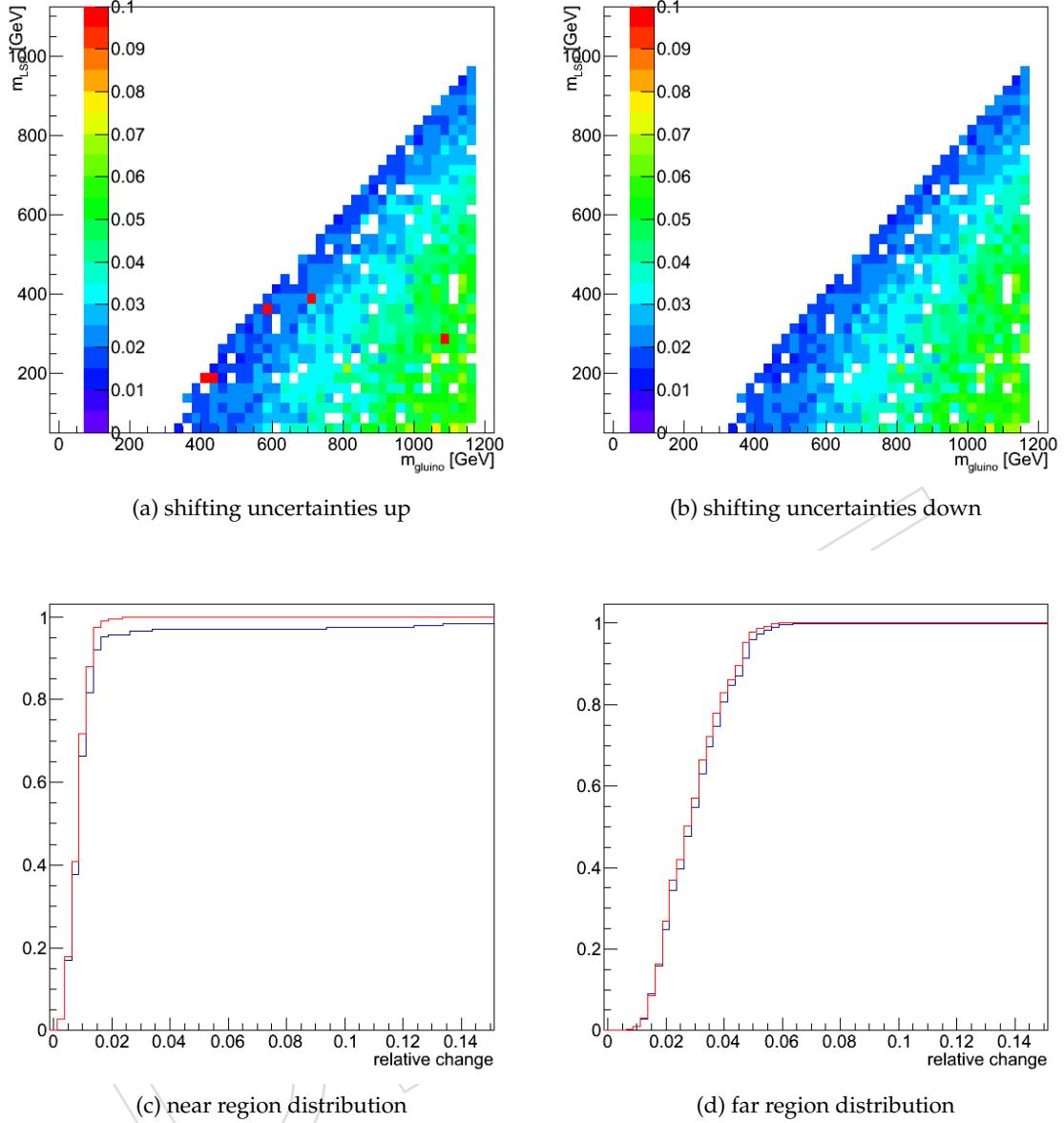


Figure 10: Top: The 68% quantiles over the 32 signal bins for each SMS point on the T1bbbb model. The quantiles shown are the absolute values of the relative changes in yields when varying the scale factors (a) up and (b) down by their uncertainties. Bottom: Averaging the absolute values of the relative changes in the (c) near and (d) far regions over all 32 signal bins for the T1bbbb model. The cumulative distributions are used to calculate the 68% quantile for varying the uncertainties up (black) and down (red) as 1.5% and 3.5% respectively.

1012 8.2 H_T evolution models

1013 The hypothesis that for a process p the α_T ratio falls exponentially in H_T can be written this
 1014 way:

$$R_{\alpha_T}(H_T) = Ae^{-kH_T} \quad (22)$$

1015 where A and k are parameters whose values will be determined. Let m_i represent the number
 1016 of events observed with $\alpha_T \leq 0.55$ in each H_T bin i , and let $\langle H_T \rangle^i$ represent the mean H_T of such
 1017 events. The expected background from the process is written thus:

$$b_p^i = \int_{x_i}^{x_{i+1}} \frac{dN}{dH_T} R_{\alpha_T} dH_T , \quad (23)$$

1018 where $\frac{dN}{dH_T}$ is the distribution of H_T for events with $\alpha_T \leq 0.55$, x_i is the lower edge of the bin,
 1019 and x_{i+1} is the upper edge of the bin (∞ for the final bin). It is assumed that

$$\frac{dN}{dH_T}(x) = \sum_i m^i \delta(x - \langle H_T \rangle^i) , \quad (24)$$

1020 i.e. within a bin the whole distribution occurs at the mean value of H_T in that bin. Then

$$b_p^i = \int_{x_i}^{x_{i+1}} m^i \delta(x - \langle H_T \rangle^i) A e^{-kx} dx = m^i A e^{-k \langle H_T \rangle^i} , \quad (25)$$

1021 and in particular,

$$QCD^i = m^i A_{QCD} e^{-k_{QCD} \langle H_T \rangle^i} . \quad (26)$$

1022 Let f_{Zinv}^i represent the expected yield from $Z \rightarrow \nu\bar{\nu}$ in bin i divided by the expected electroweak
 1023 background EWK^i . It is modeled as linear vs. H_T , i.e.

$$f_{Zinv}^i = f_{Zinv}^0 + \frac{\langle H_T \rangle^i - \langle H_T \rangle^0}{\langle H_T \rangle^{N-1} - \langle H_T \rangle^0} \left(f_{Zinv}^{N-1} - f_{Zinv}^0 \right) , \quad (27)$$

1024 where f_{Zinv}^0 and f_{Zinv}^{N-1} are both floating parameters which are limited between zero and one.

1025 8.3 Electroweak control samples

1026 Let:

$$Z_{inv}^i \equiv f_{Zinv}^i \times EWK^i \quad (28)$$

$$ttW^i \equiv (1 - f_{Zinv}^i) \times EWK^i \quad (29)$$

1027 The variable Z_{inv}^i thus represents the expected number of $Z \rightarrow \nu\bar{\nu}$ events in H_T bin i of the
 1028 hadronically selected sample, and the variable ttW^i represents the expected number of events
 1029 from SM W -boson production (including top quark decays) in H_T bin i of the hadronically
 1030 selected sample.

1031 In each bin i of H_T , there are three measurements: n_{ph}^i , n_μ^i , and $n_{\mu\mu}^i$, representing the event
 1032 counts in the photon, single-muon, and double-muon control samples. Each of these measure-
 1033 ments has a corresponding yield in simulated data: MC_{ph}^i , MC_μ^i , and $MC_{\mu\mu}^i$. The simulation

Table 13: The systematic parameters used in H_T bins.

| H_T bin (i) | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
|-------------------------|---|---|---|---|---|---|---|---|
| syst. parameter (j) | 0 | 0 | 0 | 0 | 1 | 1 | 2 | 2 |

1034 also gives expected amounts of Z_{inv} and $t\bar{t} + W$ in the hadronically-selected sample: $MC_{Z_{\text{inv}}}^i$
 1035 and $MC_{t\bar{t}+W}^i$. After defining

$$r_{ph}^i = \frac{MC_{ph}^i}{MC_{Z_{\text{inv}}}^i}; r_{\mu\mu}^i = \frac{MC_{\mu\mu}^i}{MC_{Z_{\text{inv}}}^i}; r_\mu^i = \frac{MC_\mu^i}{MC_{t\bar{t}+W}^i} \quad , \quad (30)$$

1036 these likelihood functions are used:

$$L_{ph} = \prod_i \text{Pois}(n_{ph}^i | \rho_{phZ}^j \cdot r_{ph}^i \cdot Z_{\text{inv}}^i) \quad (31)$$

$$L_{\mu\mu} = \prod_i \text{Pois}(n_{\mu\mu}^i | \rho_{\mu\mu Z}^j \cdot r_{\mu\mu}^i \cdot Z_{\text{inv}}^i) \quad (32)$$

$$L_\mu = \prod_i \text{Pois}(n_\mu^i | \rho_{\mu W}^j \cdot r_\mu^i \cdot ttW^i + s_\mu^i) \quad . \quad (33)$$

1037 Equation 31 can be used to estimate the maximum likelihood value for Z_{inv}^i (the expectation
 1038 for the $Z \rightarrow \nu\bar{\nu} + \text{jets}$ background in the hadronic signal region) given the observations n_{ph}^i in
 1039 the photon control sample and the ratios r_{ph}^i . A similar construction is used when estimating
 1040 Z_{inv}^i from the di-muon control sample (Equ. 32) and ttW^i from the single muon control sample
 1041 (Equ. 33). The measurements in each of the control samples and the hadronic signal region,
 1042 along with the ratios r_{ph}^i , $r_{\mu\mu}^i$, and r_μ^i , are all considered simultaneously through the relation-
 1043 ships defined in Equs. 21, 28, and 29. The ratios r_{ph}^i , $r_{\mu\mu}^i$, and r_μ^i are simply the inverse of the
 1044 translation factor (1/TF) defined in Equ. 5 (Sec. 5.1). More specifically, MC_{ph}^i , MC_μ^i , and $MC_{\mu\mu}^i$
 1045 are the yields obtained from MC after applying the selection criteria for the photon, single
 1046 muon and di-muon samples, as defined by Equ. 6 (Sec. 5.1). The variables $MC_{t\bar{t}+W}^i$ and $MC_{Z_{\text{inv}}}^i$
 1047 are defined by Equs. 7 and 8 (Sec. 5.1), respectively.

1048 The parameters ρ_{phZ}^j , $\rho_{\mu\mu Z}^j$, and $\rho_{\mu W}^j$ represent “correction factors” that accommodate the sys-
 1049 tematic uncertainties associated with the control-sample-based background constraints. The
 1050 quantities σ_{phZ}^j , $\sigma_{\mu\mu Z}^j$, and $\sigma_{\mu W}^j$ represent the relative systematic uncertainties for the control
 1051 sample constraints, taken into account with the following terms:

$$L_{\text{EWK syst.}} = \prod_j \text{Gaus}(1.0 | \rho_{\mu W}^j, \sigma_{\mu W}^j) \times \text{Gaus}(1.0 | \rho_{\mu\mu Z}^j, \sigma_{\mu\mu Z}^j) \times \text{Gaus}(1.0 | \rho_{phZ}^j, \sigma_{phZ}^j) \quad . \quad (34)$$

1052 Three parameters per control sample are used to span the eight H_T bins, as shown in Table 13.
 1053 Alternatively, the single muon sample can be used to constrain the total EWK background thus:

$$r'_\mu^i \equiv \frac{MC_\mu^i}{MC_{t\bar{t}+W+Z_{\text{inv}}}^i} ; \quad (35)$$

$$L_\mu = \prod_i \text{Pois}(n_\mu^i | \rho_{\mu W}^j \cdot r'_\mu^i \cdot \text{EWK}^i + s_\mu^i) . \quad (36)$$

1054 The photon and di-muon likelihoods are dropped, as are the parameters f_{Zinv} . The variable
 1055 MC_μ^i again represents the yields obtained from MC, as defined by Equ. 6 (Sec. 5.1), after apply-
 1056 ing the selection criteria for the single muon sample. The variable $MC_{t\bar{t}+W+Z_{\text{inv}}}^i$ is defined by
 1057 Equ. 9 (Sec. 5.1).

1058 8.4 Contributions from signal

1059 Let x represent the cross section for a particular signal model, and let l represent the recorded
 1060 luminosity. Let ϵ_{had}^i (resp. ϵ_μ^i) be the analysis efficiency as simulated for the model in H_T bin
 1061 i of the hadronic (resp. single muon control) sample. Let δ represent the relative uncertainty
 1062 on the signal yield, assumed to be fully correlated among the bins, and let ρ_{sig} represent the
 1063 “correction factor” to the signal yield which accommodates this uncertainty. Let f represent an
 1064 unknown multiplicative factor on the signal cross section, for which an allowed interval shall
 1065 be determined.

1066 Then the expected hadronic signal yield s^i from Equation 20 is written as $s^i \equiv f \rho_{\text{sig}} x l \epsilon_{\text{had}}^i$,
 1067 and the “signal contamination” in the muon control sample s_μ^i from Equation 33 is treated
 1068 analogously: $s_\mu^i \equiv f \rho_{\text{sig}} x l \epsilon_\mu^i$. The systematic uncertainty on the signal efficiency is included via
 1069 an additional term in the likelihood:

$$L_{\text{sig}} = \text{Gaus}(1.0 | \rho_{\text{sig}}, \delta) . \quad (37)$$

1070 8.5 Total likelihood

1071 The likelihood function for a given selection k is the product of the terms described in the
 1072 previous sections:

$$L^k = L_{\text{hadronic}}^k \times L_\mu^k \times L_{ph}^k \times L_{\mu\mu}^k . \quad (38)$$

1073 There are $3 + N$ nuisance parameters associated with each selection: A_{QCD} , f_{Zinv}^0 , f_{Zinv}^{N-1} , $\{\text{EWK}^i\}_{i=0}^{N-1}$.
 1074 The 11 parameters k_{QCD} , ρ_{sig} , and ρ_{phZ}^j , $\rho_{\mu\mu Z}^j$, $\rho_{\mu W}^j$, with $j = 0, 1, 2$ are shared among selections.
 1075 The total likelihood is thus

$$L = L_{\text{sig}} \times L_{\text{EWK syst.}} \times \prod_k L_{\text{hadronic}}^k \times L_\mu^k \times L_{ph}^k \times L_{\mu\mu}^k . \quad (39)$$

1076 9 Final Results

1077 9.1 Standard Model

1078 Contributions to the likelihood from signal are dropped, and the likelihood function is maxi-
 1079 mized over all parameters using RooFit [24] and MINUIT [25]. The resulting parameter values

are listed in Tables 40 and 41 (Appendix G). The fit yields, as well as their uncertainties determined with an ensemble of pseudo-experiments, are compared with the observations in Tables 42, 43, 44, and 45.

Goodness-of-fit is determined using a method from [26]. The value L_{\max}^{data} of the likelihood function at its maximum is noted. The likelihood function (using the ML parameter values) is then used as a p.d.f. for the observations to generate many pseudo-experiments. For each pseudo-experiment, the likelihood function is again maximized over all parameters, and the value L_{\max} entered in a histogram. The quantile of L_{\max}^{data} in this histogram is interpreted as a p -value.

Figure 11 shows a comparison of the observed yields and SM expectations given by the simultaneous fit in bins of H_T with exactly zero b-tagged jets per event, in the hadronic signal and three control samples for the region $\alpha_T > 0.55$. Similarly, Figs. 12, 13 and 14 show the comparison between observed and expected yields when requiring exactly one, two and at least three b-tagged jets per event, respectively. For all four b-tag categories, good agreement is observed between the observed yields and the SM expectation from the simultaneous fit in all bins of each data sample. No significant excess above the SM expectation is observed. For illustrative purposes, App. K shows similar plots for the hadronic signal region with contributions from various example signal models superimposed on the SM-only expectation as given by the fit.

Table 14: Comparison of the measured yields in the different H_T and b jet multiplicity bins for the hadronic sample with the SM expectations and combined statistical and systematic uncertainties given by the simultaneous fit.

| H_T (GeV) | 275–325 | 325–375 | 375–475 | 475–575 | 575–675 | 675–775 | 775–875 | 875– ∞ |
|----------------------|----------------------|----------------------|----------------------|----------------------|----------------------|----------------------|----------------------|---------------------|
| 0 b jets SM | 2933^{+56}_{-52} | 1139^{+17}_{-40} | 783^{+17}_{-27} | 261^{+14}_{-8} | $81.5^{+6.5}_{-6.5}$ | $34.2^{+4.0}_{-3.8}$ | $10.4^{+2.8}_{-1.8}$ | $5.3^{+1.7}_{-1.1}$ |
| 0 b jets Data | 2919 | 1166 | 769 | 255 | 91 | 31 | 10 | 4 |
| 1 b jet SM | 630^{+26}_{-25} | 271^{+10}_{-16} | 202^{+10}_{-6} | $78.0^{+6.9}_{-1.9}$ | $24.2^{+2.9}_{-2.0}$ | $10.6^{+1.7}_{-1.3}$ | $2.9^{+0.9}_{-0.5}$ | $2.2^{+0.7}_{-0.4}$ |
| 1 b jet Data | 614 | 294 | 214 | 71 | 20 | 6 | 4 | 0 |
| 2 b jets SM | 162^{+13}_{-12} | $61.8^{+4.8}_{-6.3}$ | $58.8^{+4.8}_{-2.6}$ | $28.0^{+3.5}_{-1.1}$ | $9.0^{+1.4}_{-1.0}$ | $7.1^{+1.4}_{-1.0}$ | $0.6^{+0.3}_{-0.2}$ | $0.9^{+0.4}_{-0.2}$ |
| 2 b jets Data | 160 | 68 | 52 | 19 | 11 | 7 | 0 | 2 |
| ≥ 3 b jets SM | $10.5^{+3.5}_{-2.2}$ | $7.1^{+2.2}_{-1.8}$ | $5.8^{+1.4}_{-0.9}$ | $3.1^{+1.0}_{-0.7}$ | $1.7^{+0.5}_{-0.4}$ | $0.7^{+0.5}_{-0.4}$ | $0.1^{+0.1}_{-0.1}$ | $0.2^{+0.1}_{-0.1}$ |
| ≥ 3 b jets Data | 10 | 8 | 8 | 1 | 0 | 0 | 0 | 0 |

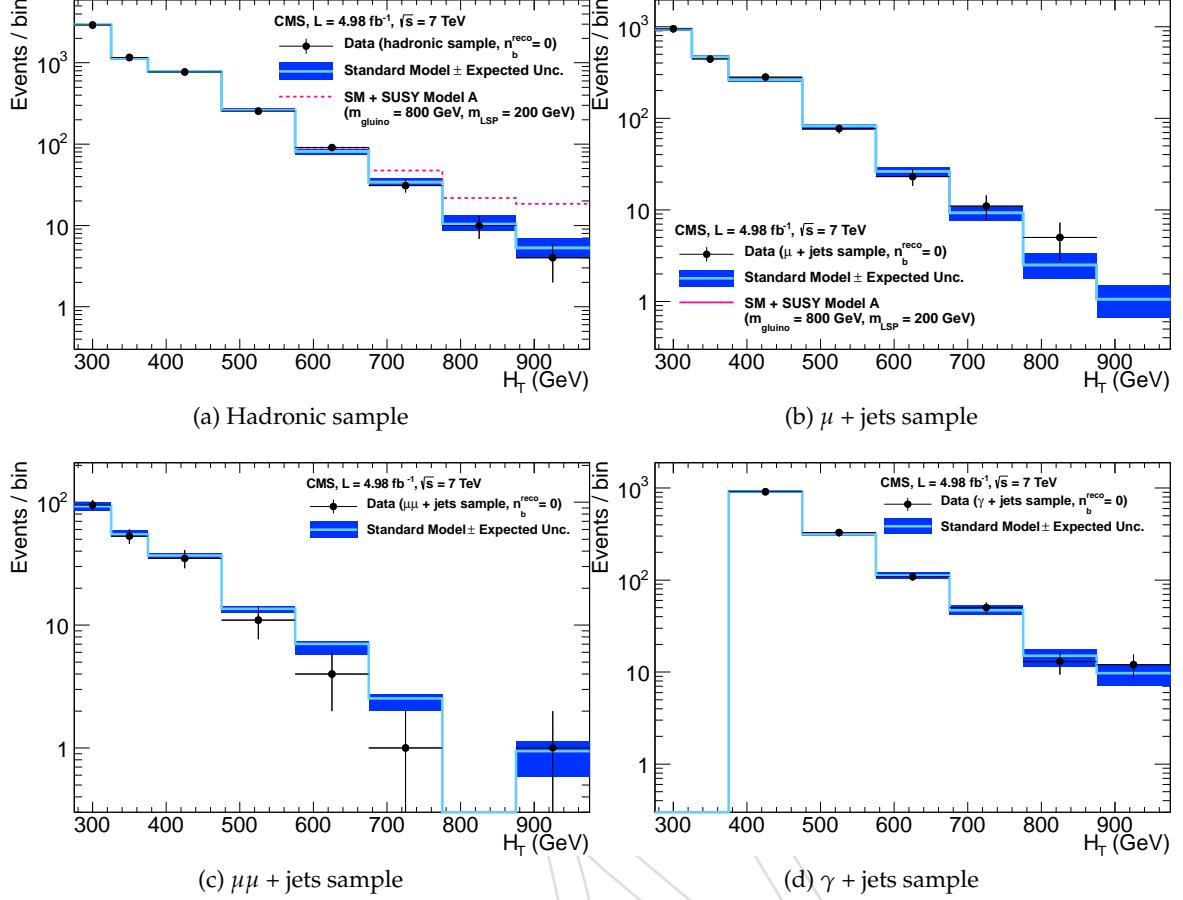


Figure 11: Comparison of the observed yields and SM expectations given by the simultaneous fit in bins of H_T for the (a) hadronic, (b) $\mu + \text{jets}$, (c) $\mu\mu + \text{jets}$ and (d) $\gamma + \text{jets}$ samples when requiring exactly zero reconstructed b-jets. The observed event yields in data (black dots) and the expectations and their uncertainties, as determined by the simultaneous fit, for all SM processes (light blue solid line with dark blue bands) are shown. For illustrative purposes only, an example signal model is superimposed on the SM expectation (magenta solid line). The expected signal contamination in the control samples is negligible.

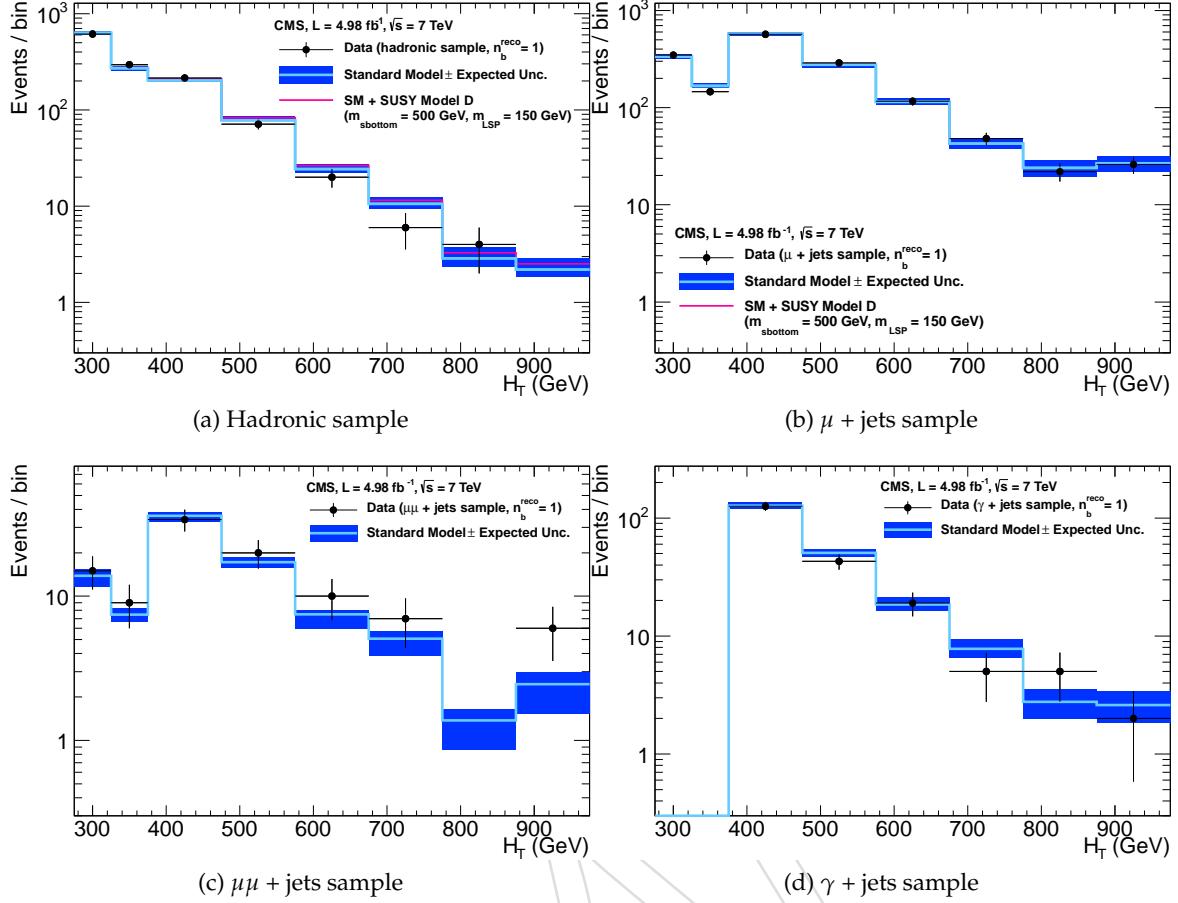


Figure 12: Comparison of the observed yields and SM expectations given by the simultaneous fit in bins of H_T for the (a) hadronic, (b) $\mu + \text{jets}$, (c) $\mu\mu + \text{jets}$ and (d) $\gamma + \text{jets}$ samples when requiring exactly one reconstructed b-jet. The observed event yields in data (black dots) and the expectations and their uncertainties, as determined by the simultaneous fit, for all SM processes (light blue solid line with dark blue bands) are shown. For illustrative purposes only, an example signal model is superimposed on the SM expectation (magenta solid line). The expected signal contamination in the control samples is negligible.

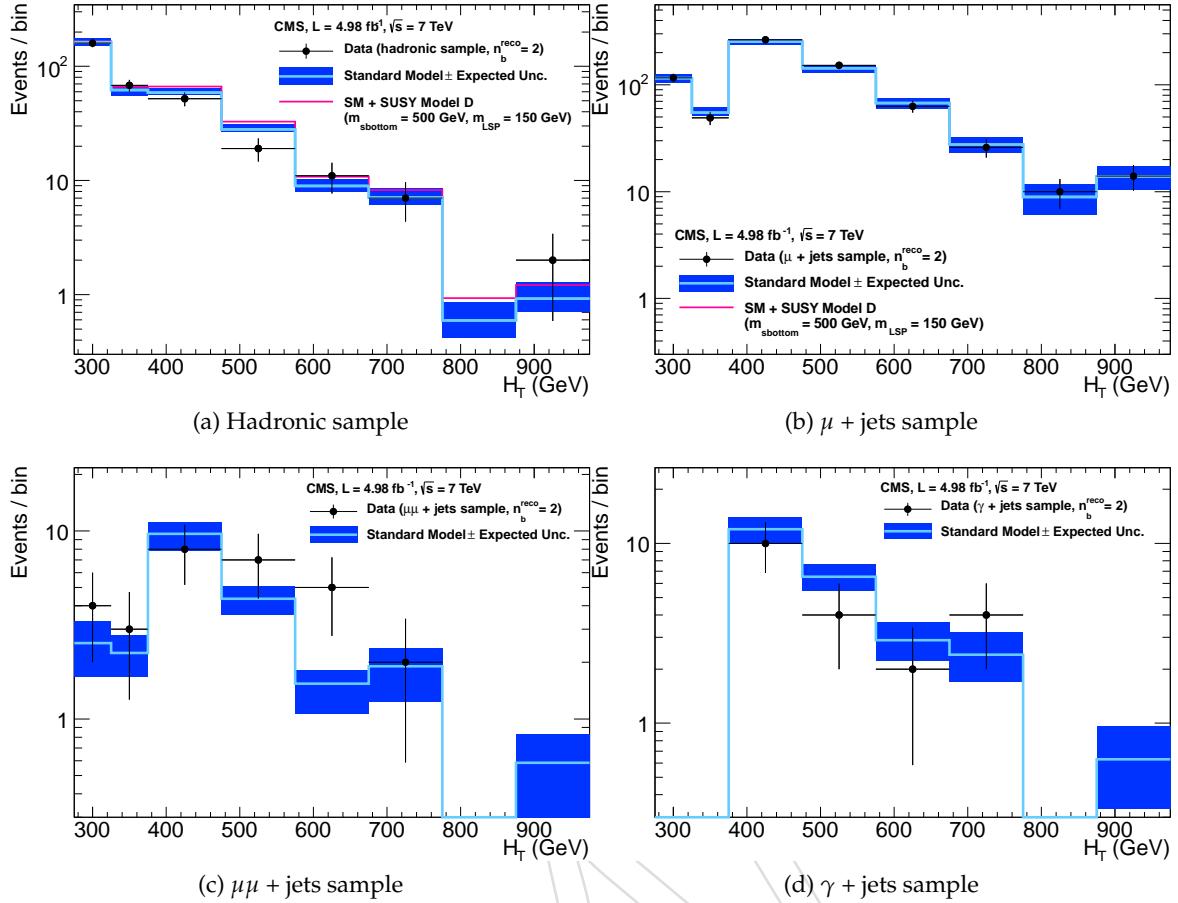


Figure 13: Comparison of the observed yields and SM expectations given by the simultaneous fit in bins of H_T for the (a) hadronic, (b) $\mu + \text{jets}$, (c) $\mu\mu + \text{jets}$ and (d) $\gamma + \text{jets}$ samples when requiring exactly two reconstructed b-jets. The observed event yields in data (black dots) and the expectations and their uncertainties, as determined by the simultaneous fit, for all SM processes (light blue solid line with dark blue bands) are shown. For illustrative purposes only, an example signal model is superimposed on the SM expectation (magenta solid line). The expected signal contamination in the control samples is negligible.

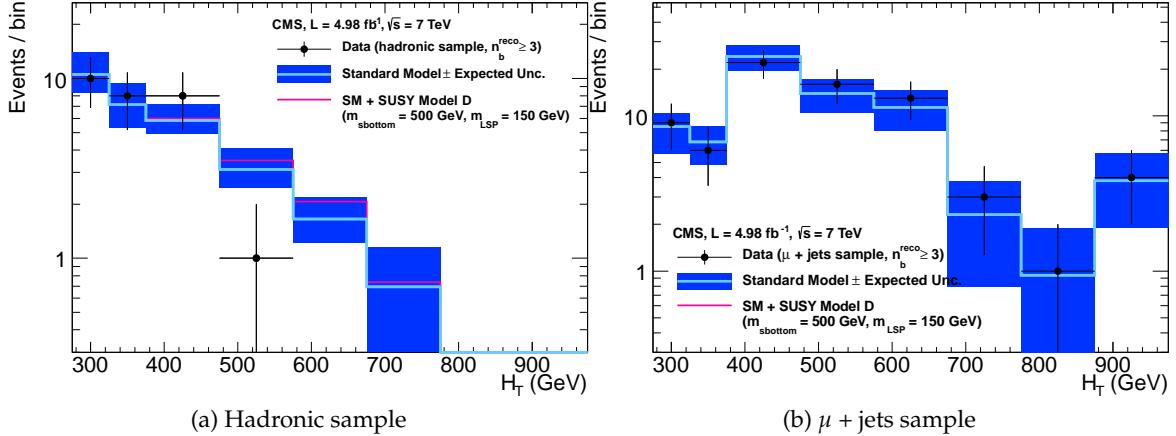


Figure 14: Comparison of the observed yields and SM expectations given by the simultaneous fit in bins of H_T for the (a) hadronic and (b) $\mu + \text{jets}$ samples when requiring at least three reconstructed b-jets. The observed event yields in data (black dots) and the expectations and their uncertainties, as determined by the simultaneous fit, for all SM processes (light blue solid line with dark blue bands) are shown. For illustrative purposes only, an example signal model is superimposed on the SM expectation (magenta solid line). The expected signal contamination in the $\mu + \text{jets}$ control sample is negligible.

1098 9.2 SUSY

1099 Limits are set in the parameter space of the CMSSM. At each point in the parameter space of the
1100 CMSSM, the SUSY particle spectrum is calculated with SOFTSUSY [27], and signal events are
1101 generated at leading order with PYTHIA 6.4 [28]. Inclusive, process-dependent, next-to-leading
1102 order calculations with next-to-leading logarithmic corrections [29] (NLO+NLL) of SUSY pro-
1103 duction cross sections are obtained with the program PROSPINO [30] and CTEQ6 [31] parton
1104 distribution functions. The simulated signal events are reweighted so that the distribution of
1105 the number of reconstructed vertices per beam crossing from the simulation matches that ob-
1106 served in data. Experimental uncertainties on the SM background prediction (10 – 40%), the
1107 luminosity measurement, and the total selection efficiency times acceptance for the considered
1108 signal model (16%, Table 10) are included in the calculation of the limit. The dominant sources
1109 of uncertainty on the signal efficiency times acceptance are derived from systematic variations
1110 of parton distribution functions, and corrections applied to jet energies and b jet efficiency and
1111 mistag rates. Although signal contributions to the total yield in each of the four considered data
1112 samples are allowed, the only significant signal contribution originates from the hadronic data
1113 sample in the case of the CMSSM. Although signal contributions to the total yield in each of
1114 the three considered data samples are allowed, the only relevant signal contribution originates
1115 from the hadronic data sample in the case of the CMSSM.

1116 Figure 15 shows the observed and expected exclusion limits at 95% confidence level (CL) in the
1117 $(m_0, m_{1/2})$ plane for $\tan \beta = 10$ and $A_0 = 0 \text{ GeV}$, calculated with NLO+NLL SUSY production
1118 cross sections and the CL_s method [32]. For this choice of parameter values, squark masses
1119 below 1250 GeV are excluded at 95% CL, as are gluino masses below the same value for the
1120 region $m_0 < 600 \text{ GeV}$. In the region $600 < m_0 < 3000 \text{ GeV}$, gluino masses below 700 GeV are
1121 excluded, while the squark mass in the excluded models varies in the range 1250 – 2500 GeV,
1122 depending on the value of m_0 . The mass limits are determined with the observed limit for the
1123 nominal production cross section less 1σ theoretical uncertainty.

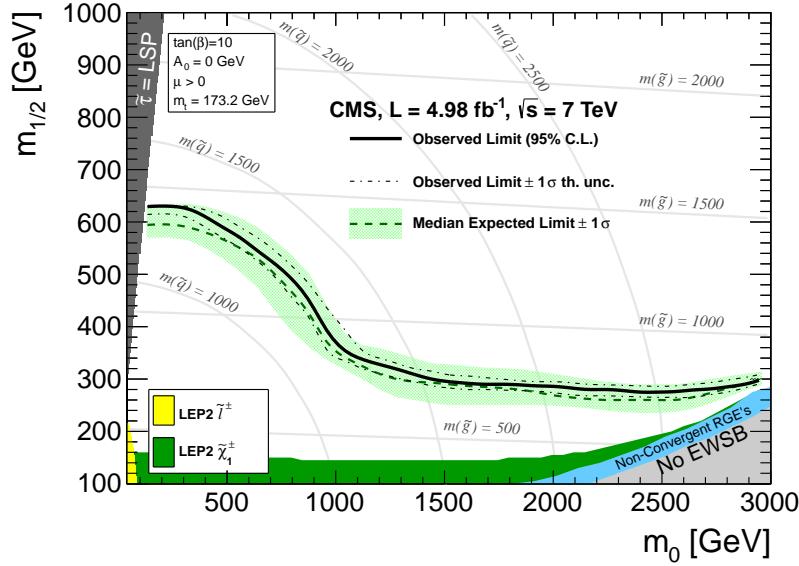


Figure 15: Exclusion contours at 95% CL in the CMSSM ($m_0, m_{1/2}$) plane ($\tan \beta = 10, A_0 = 0, \mu > 0$) calculated with NLO+NLL SUSY production cross sections and the CL_s method. The solid black line indicates the observed exclusion region. The dotted-dashed black lines represent the observed excluded region when varying the cross section by its theoretical uncertainty. The expected median exclusion region (green dashed line) $\pm 1\sigma$ (green band) are also shown.

The data observations are also interpreted using simplified models that characterise third generation squark production and compressed spectra scenarios, where the mass difference between the primary produced sparticle (e.g. a squark or a gluino) and the LSP is rather small. The production and decay modes of the models under consideration are summarised in Table 15. The simplified models T1 and T2 are used to characterise the pair production of gluinos and first or second generation squarks, respectively, depending on their mass as well as on the LSP mass. The latter four simplified models describe various production and decay mechanisms in the context of third generation squarks.

Experimental uncertainties on the SM background predictions (10 – 40%), the luminosity measurement, and the total selection efficiency times acceptance for the considered signal model (12%–18%, Tables 11 and 12) are included in the calculation of the limit. Again, the presence of signal events in the control samples has been accounted for. Signal efficiency in the kinematic region defined by $0 < m_{\tilde{g}(\tilde{q})} - m_{\text{LSP}} < 200 \text{ GeV}$ or $m_{\tilde{g}(\tilde{q})} < 350 \text{ GeV}$ is due in part to the presence of initial state radiation. Given the large associated uncertainties, no interpretation is provided for this kinematic region. In the case of model T1tttt, for which pair-produced gluinos decay to top-antitop quark pairs and the LSP, the region is enlarged to cover $0 < m_{\tilde{g}} - m_{\text{LSP}} < 400 \text{ GeV}$.

Figure 16 shows the upper limit on the cross section at 95% CL as a function of $m_{\tilde{q}}$ or $m_{\tilde{g}}$ and m_{LSP} for various simplified models. The solid thick black line indicates the observed exclusion region assuming NLO+NLL [29, 30] SUSY cross section for squark pair production in the limit of decoupled gluinos (or vice versa). The thin black lines represent the observed excluded region when varying the cross section by its theoretical uncertainty. The dashed purple lines indicate the median (thick line) $\pm 1\sigma$ (thin lines) expected exclusion regions.

The most stringent mass limits on the pair-produced sparticles are obtained at low LSP masses, while the limits typically weaken for compressed spectra, i.e. points close to the diagonal. In

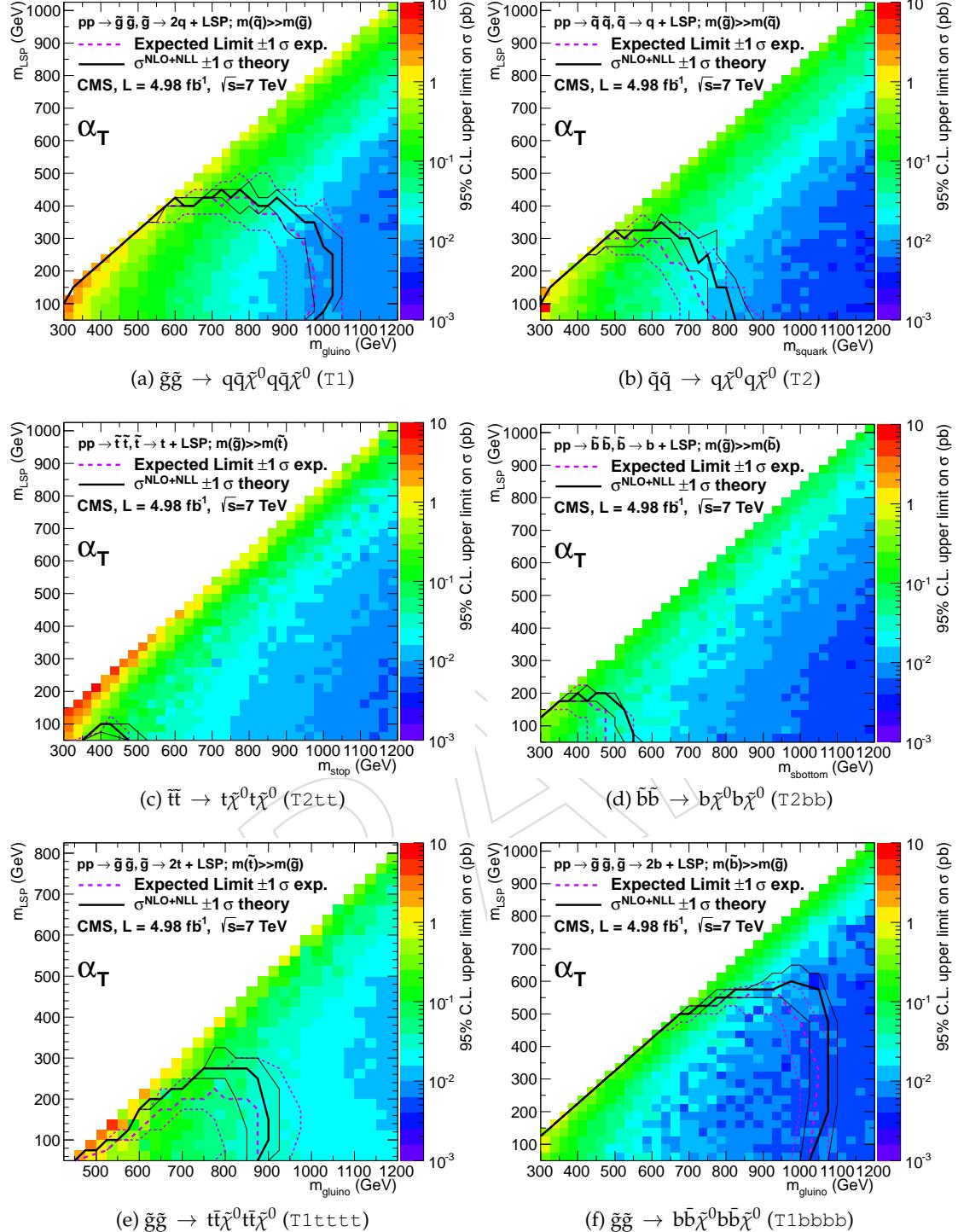


Figure 16: Upper limit on cross section at 95% CL as a function of $m_{\tilde{q}}$ or $m_{\tilde{g}}$ and m_{LSP} for various simplified models. The solid thick black line indicates the observed exclusion region assuming NLO+NLL SUSY production cross section. The thin black lines represent the observed excluded region when varying the cross section by its theoretical uncertainty. The dashed purple lines indicate the median (thick line) $\pm 1\sigma$ (thin lines) expected exclusion regions.

particular, for all of the considered simplified models, there is an LSP mass beyond which

Table 15: Production and decay modes for various simplified models.

| Model | Production and decay modes | Figure showing limit |
|--------|--------------------------------------------------------------------------------|----------------------|
| T1 | $\tilde{g}\tilde{g} \rightarrow q\bar{q}\tilde{\chi}^0 q\bar{q}\tilde{\chi}^0$ | 16a |
| T2 | $\tilde{q}\bar{q} \rightarrow q\tilde{\chi}^0 q\tilde{\chi}^0$ | 16b |
| T2tt | $\tilde{t}\bar{t} \rightarrow t\tilde{\chi}^0 t\tilde{\chi}^0$ | 16c,17 |
| T2bb | $\tilde{b}\bar{b} \rightarrow b\tilde{\chi}^0 b\tilde{\chi}^0$ | 16d |
| T1tttt | $\tilde{g}\tilde{g} \rightarrow t\bar{t}\tilde{\chi}^0 t\bar{t}\tilde{\chi}^0$ | 16e |
| T1bbbb | $\tilde{g}\tilde{g} \rightarrow b\bar{b}\tilde{\chi}^0 b\bar{b}\tilde{\chi}^0$ | 16f |

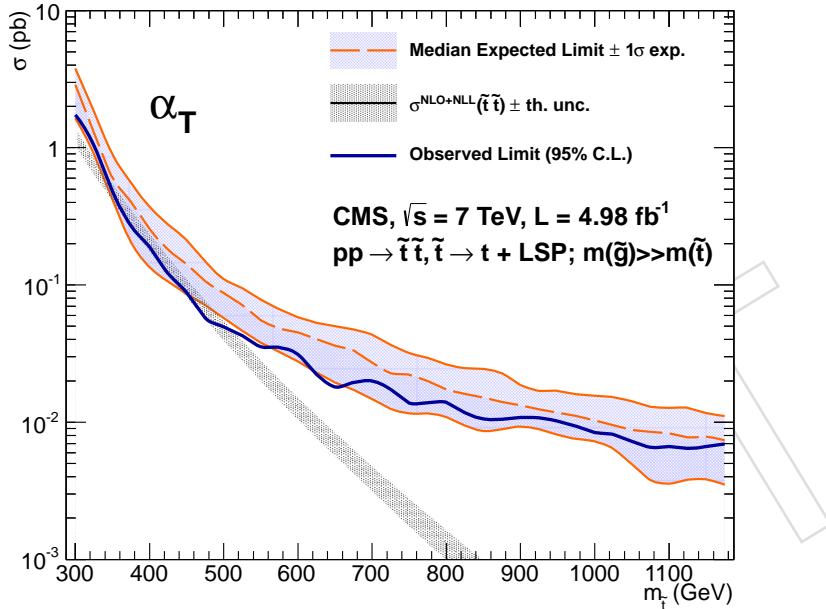


Figure 17: Excluded cross section versus top squark mass for a model in which pair-produced top squarks decay to two top quarks and a two neutralinos. The solid blue line indicates the observed cross section upper limit (95% CL) as a function of the top squark mass, $m_{\tilde{t}}$. The dashed orange line and blue band indicate the median expected excluded cross section with experimental uncertainties. The solid black line with grey band indicates the NLO+NLL SUSY top squark pair-production cross section and theoretical uncertainties.

no limit can be set. This is illustrated in Figure 16a, where the most stringent limit on the gluino mass is obtained at around 950 TeV for low LSP masses, while this limit weakens to below 900 GeV when the LSP mass reaches 350 GeV. For LSP masses above 400 GeV, no gluino masses can be excluded. Table 15 summarises these two extreme cases for models A to F. The estimates on the mass limits are determined with the observed limit for the nominal production cross section less 1σ theoretical uncertainty.

No exclusion of direct top squark pair production (model T2tt) assuming NLO+NLL production cross section is expected with the analysed dataset and for LSP masses greater than 50 GeV. Figure 17 shows the observed upper limit at 95% CL on the cross section as a function of the top squark mass ($m_{\tilde{t}}$) only, for a fixed LSP mass of $m_{\text{LSP}} = 50$ GeV.

1160 **10 Summary**

1161 A search for supersymmetry is reported, based on a data sample of pp collisions collected at
 1162 $\sqrt{s} = 7\text{ TeV}$, corresponding to an integrated luminosity of 4.98 fb^{-1} . Final states with two or
 1163 more jets and significant \cancel{E}_T , as expected from high-mass squark and gluino production and
 1164 decays, have been analysed. An exclusive search has been performed in a binned signal region
 1165 defined by the scalar sum of the transverse energy of jets, H_T , and the number of jets identified
 1166 to originate from a bottom quark. The sum of standard model backgrounds per bin has been
 1167 estimated from a simultaneous binned likelihood fit to hadronic, $\mu + \text{jets}$, $\mu\mu + \text{jets}$, and $\gamma +$
 1168 jets samples. The observed yields are found to be in agreement with the expected contribu-
 1169 tions from standard model processes. Limits in the CMSSM ($m_0, m_{1/2}$) plane for $\tan\beta = 10$,
 1170 $A_0 = 0\text{ GeV}$, and $\mu > 0$ have been derived. For this choice of parameter values, gluino masses
 1171 below 700 GeV are excluded at 95% CL. The exclusion increases to 1250 GeV for squarks and
 1172 gluinos of comparable mass. Furthermore, exclusion limits are also set in simplified models,
 1173 with a special emphasis on third generation squarks and compressed spectra scenarios. In the
 1174 considered models with gluino pair production and for small LSP masses, typical exclusion
 1175 limits of the gluino mass are around 1 TeV . For simplified models with squark pair production,
 1176 first or second generation squarks are excluded up to around 750 GeV and bottom squarks are
 1177 excluded up to around 500 GeV , again for small LSP masses. However, for the simplified mod-
 1178 els under consideration, the most constraining limit on the LSP mass is significantly below
 1179 1 TeV , indicating that a large range of SUSY parameter space is yet to be probed by the LHC.

1180 **References**

- 1181 [1] L. Randall and D. Tucker-Smith, “Dijet Searches for Supersymmetry at the Large Hadron
 1182 Collider”, *Phys. Rev. Lett.* **101** (2008) 221803.
 1183 doi:[10.1103/PhysRevLett.101.221803](https://doi.org/10.1103/PhysRevLett.101.221803).
- 1184 [2] CMS Collaboration, CMS Collaboration, “SUSY searches with dijet events”, CMS Physics
 1185 Analysis Summary SUS-08-005, CMS, 2008.
- 1186 [3] CMS Collaboration, CMS Collaboration, “Search strategy for exclusive multi-jet events
 1187 from supersymmetry at CMS”, CMS Physics Analysis Summary SUS-09-001, CMS, 2009.
- 1188 [4] CMS Collaboration Collaboration, “Search for Supersymmetry at the LHC in Events with
 1189 Jets and Missing Transverse Energy”, *Phys. Rev. Lett.* **107** (Nov, 2011) 221804.
 1190 doi:[10.1103/PhysRevLett.107.221804](https://doi.org/10.1103/PhysRevLett.107.221804).
- 1191 [5] CMS Collaboration, “Search for Supersymmetry in pp Collisions at 7 TeV in Events with
 1192 Jets and Missing Transverse Energy”, *Phys. Lett.* **B698** (2011) 196–218,
 1193 arXiv:1101.1628. doi:[10.1016/j.physletb.2011.03.021](https://doi.org/10.1016/j.physletb.2011.03.021).
- 1194 [6] G. L. Kane et al., “Study of constrained minimal supersymmetry”, *Phys. Rev. D* **49** (1994)
 1195 6173. doi:[10.1103/PhysRevD.49.6173](https://doi.org/10.1103/PhysRevD.49.6173).
- 1196 [7] A. H. Chamseddine, R. Arnowitt, and P. Nath, “Locally Supersymmetric Grand
 1197 Unification”, *Phys. Rev. Lett.* **49** (1982) 970. doi:[10.1103/PhysRevLett.49.970](https://doi.org/10.1103/PhysRevLett.49.970).
- 1198 [8] J. Alwall, P. Schuster, and N. Toro, “Simplified Models for a First Characterization of New
 1199 Physics at the LHC”, *Phys. Rev. D* **79** (2009) 075020, arXiv:0810.3921.
 1200 doi:[10.1103/PhysRevD.79.075020](https://doi.org/10.1103/PhysRevD.79.075020).

- [9] J. Alwall, M.-P. Le, M. Lisanti et al., "Model-Independent Jets plus Missing Energy Searches", *Phys.Rev.* **D79** (2009) 015005, arXiv:0809.3264.
doi:10.1103/PhysRevD.79.015005.
- [10] D. Alves, N. Arkani-Hamed, S. Arora et al., "Simplified Models for LHC New Physics Searches", arXiv:1105.2838. Official summary of results from the 'Topologies for Early LHC Searches' workshop, SLAC, September 2010.
- [11] M. Cacciari, G. P. Salam, and G. Soyez, "The anti- k_T jet clustering algorithm", *JHEP* **04** (2008) 063. doi:10.1088/1126-6708/2008/04/063.
- [12] CMS Collaboration, "Determination of Jet Energy Calibration and Transverse Momentum Resolution in CMS", *JINST* **6** (2011) P11002, arXiv:1107.4277. doi:10.1088/1748-0221/6/11/P11002.
- [13] CMS Collaboration, "Identification and filtering of uncharacteristic noise in the CMS hadron calorimeter", *JINST* **5** (2010) T03014. doi:10.1088/1748-0221/5/03/T03014.
- [14] CMS Collaboration, CMS Collaboration, "Electromagnetic calorimeter commissioning and first results with 7 TeV data", CMS Note CMS-NOTE-2010-012, CMS, 2010.
- [15] CMS Collaboration, CMS Collaboration, "Electron reconstruction and identification at $\sqrt{s} = 7$ TeV", CMS Physics Analysis Summary EGM-10-004, CMS, 2010.
- [16] CMS Collaboration, CMS Collaboration, "Performance of muon identification in pp collisions at $\sqrt{s} = 7$ TeV", CMS Physics Analysis Summary MUO-10-002, CMS, 2010.
- [17] CMS Collaboration, CMS Collaboration, "Isolated Photon Reconstruction and Identification at $\sqrt{s} = 7$ TeV", CMS Physics Analysis Summary EGM-10-006, CMS, 2010.
- [18] CMS PAS BTV 09 001 "Algorithms for b-Jet identification in CMS".
- [19] CMS PAS BTV 11 001 "Performance of b-Jet identification in CMS".
- [20] CMS Collaboration, CMS Collaboration, "Data-Driven Estimation of the Invisible Z Background to the SUSY MET Plus Jets Search", CMS Physics Analysis Summary SUS-08-002, CMS, 2008.
- [21] Z. Bern, G. Diana, L. Dixon et al., "Driving Missing Data at Next-to-Leading Order", *Phys.Rev.* **D84** (2011) 114002, arXiv:1106.1423.
doi:10.1103/PhysRevD.84.114002.
- [22] <https://twiki.cern.ch/twiki/bin/viewauth/CMS/BtagPOG>.
- [23] CMS Collaboration, CMS Collaboration, "Absolute Calibration of the Luminosity Measurement at CMS: Winter 2012 Update", CMS Physics Analysis Summary SMP-12-008, CMS, 2012.
- [24] W. Verkerke and D. Kirkby, "RooFit".
<http://root.cern.ch/drupal/content/roofit>.
- [25] F. James and M. Roos, "Minuit: A System for Function Minimization and Analysis of the Parameter Errors and Correlations", *Comput.Phys.Commun.* **10** (1975) 343–367.
doi:10.1016/0010-4655(75)90039-9.

- [26] G. D. Cowan, “Statistical data analysis”. Oxford Univ. Press, Oxford, 1998.
- [27] B. C. Allanach, “SOFTSUSY: a program for calculating supersymmetric spectra”, *Comput. Phys. Commun.* **143** (2002) 305. doi:10.1016/S0010-4655(01)00460-X.
- [28] T. Sjöstrand, S. Mrenna and P. Z. Skands, “PYTHIA 6.4 Physics and Manual”, *JHEP* **05** (2006) 026. doi:10.1088/1126-6708/2006/05/026.
- [29] M. Krämer, A. Kulesza, R. van der Leeuw et al., “Supersymmetry production cross sections in pp collisions at $\sqrt{s} = 7$ TeV”, arXiv:1206.2892.
- [30] W. Beenakker et al., “Squark and gluino production at hadron colliders”, *Nucl. Phys. B* **492** (1997) 51. doi:10.1016/S0550-3213(97)00084-9.
- [31] J. Pumplin et al., “New generation of parton distributions with uncertainties from global QCD analysis”, *JHEP* **07** (2002) 012, arXiv:hep-ph/0201195.
- [32] Particle Data Group Collaboration, “Review of particle physics”, *J. Phys. G* **37** (2010) 075021. doi:10.1088/0954-3899/37/7A/075021.

DRAFT

1253 A Distributions for the hadronic signal region

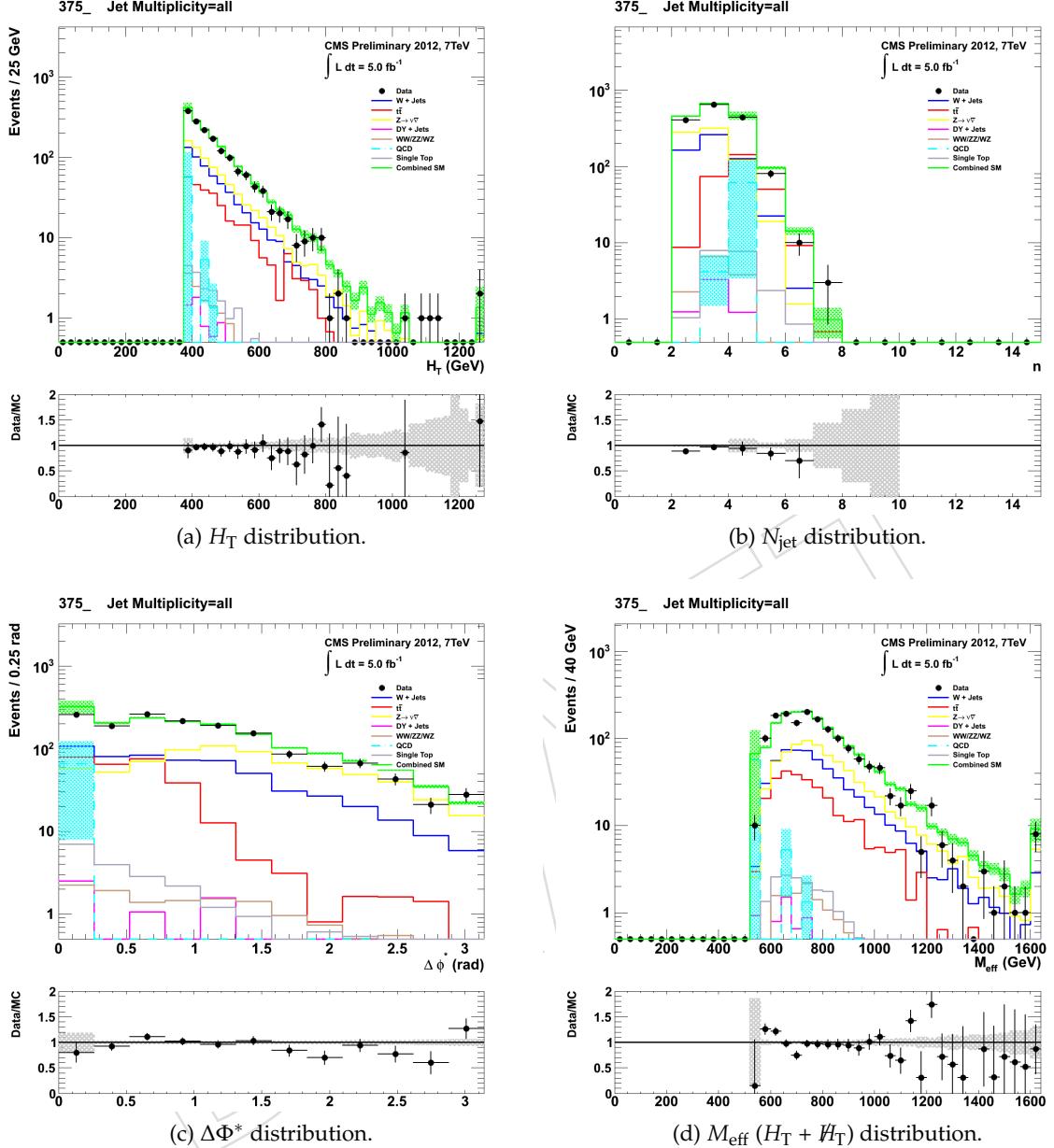


Figure 18: Data–MC comparisons of key variables for the hadronic signal region, following the full-selection, which includes $\alpha_T > 0.55$. In addition, the requirement $H_T > 375$ GeV is imposed. Bands represent the uncertainties due to the limited size of MC samples. No requirement is made on the number of b-tagged jets in the event.

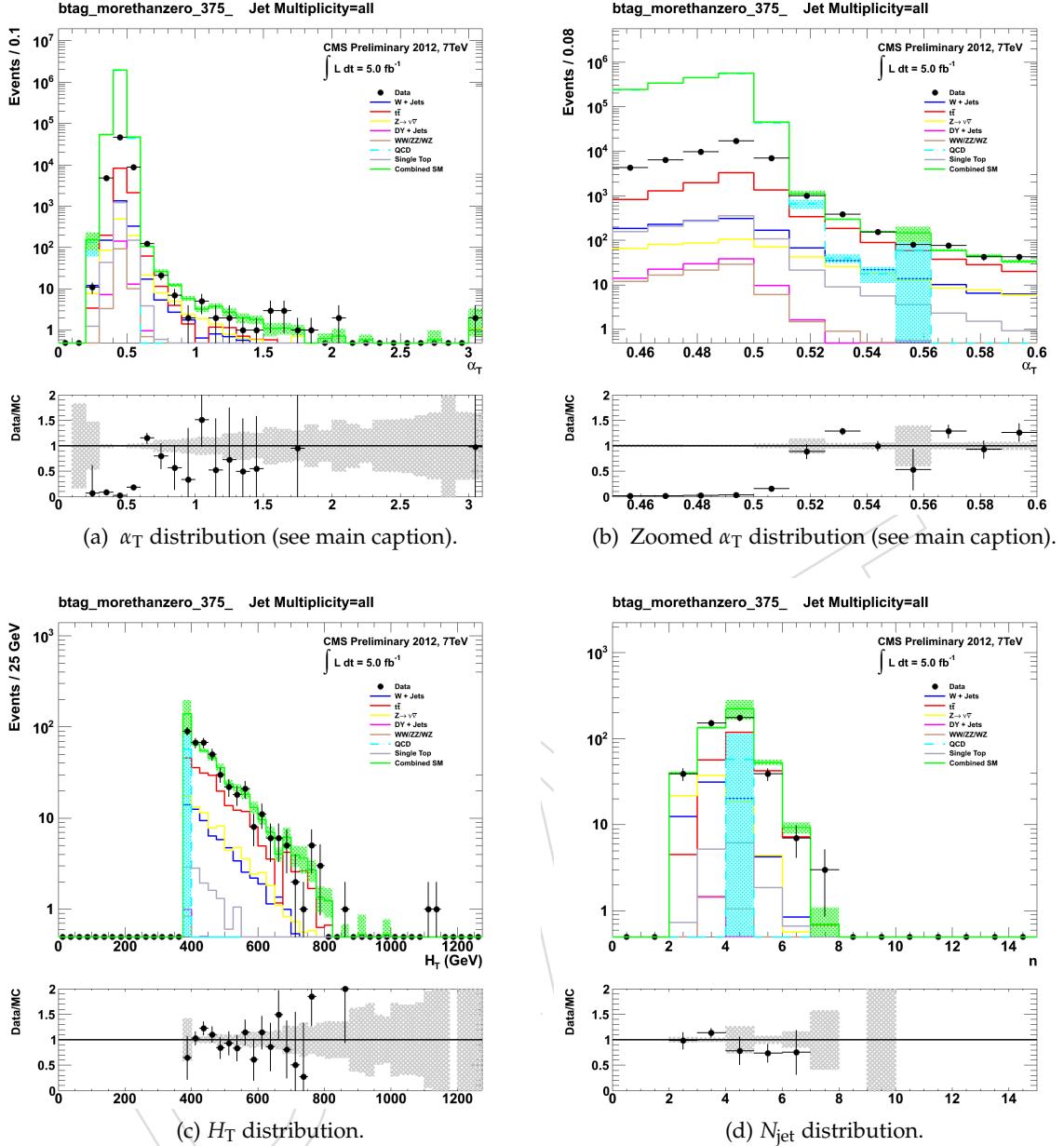


Figure 19: Data–MC comparisons of key variables for the b-tagged hadronic signal region, defined by $H_T > 375 \text{ GeV}$ and $\alpha_T > 0.55$. At least one b-tagged jet per event is required. Bands represent the uncertainties due to the limited size of MC samples. *The discrepancy in the α_T distributions for values $\alpha_T < 0.55$ is due to the trigger not being simulated in the MC simulation.*

1254 B Additional information on triggers

Table 16: List of HT triggers used.

| H_T bin (GeV) | Trigger |
|-------------------|--------------|
| $275 < H_T < 325$ | HLT-HT250_v* |
| $325 < H_T < 375$ | HLT-HT300_v* |
| $375 < H_T < 475$ | HLT-HT350_v* |
| $H_T > 475$ | HLT-HT400_v* |

Table 17: List of HT_AlphaT triggers used.

| $275 < H_T < 325$ GeV | $325 < H_T < 375$ GeV |
|-------------------------|-------------------------|
| HLT-HT250_AlphaT0p53_v* | HLT-HT300_AlphaT0p52_v* |
| HLT-HT250_AlphaT0p55_v* | HLT-HT300_AlphaT0p53_v* |
| HLT-HT250_AlphaT0p58_v* | HLT-HT300_AlphaT0p54_v* |
| HLT-HT250_AlphaT0p60_v* | HLT-HT300_AlphaT0p55_v* |

| $375 < H_T < 475$ GeV | $H_T > 475$ GeV |
|-------------------------|-------------------------|
| HLT-HT350_AlphaT0p51_v* | HLT-HT400_AlphaT0p51_v* |
| HLT-HT350_AlphaT0p52_v* | HLT-HT400_AlphaT0p52_v* |
| HLT-HT350_AlphaT0p52_v* | |
| HLT-HT350_AlphaT0p53_v* | |

Table 18: List of Mu-HT triggers used.

| $H_T > 275$ GeV | $H_T > 375$ GeV |
|-------------------|-------------------|
| HLT_Mu5-HT200_v* | HLT_Mu5-HT200_v* |
| HLT_Mu8-HT200_v* | HLT_Mu8-HT200_v* |
| HLT_Mu15-HT200_v* | HLT_Mu15-HT200_v* |
| HLT_Mu30-HT200_v* | HLT_Mu30-HT200_v* |
| HLT_Mu40-HT200_v* | HLT_Mu40-HT200_v* |
| | HLT_Mu40-HT300_v* |

Table 19: List of Photon triggers used.

| $H_T > 375$ GeV |
|-------------------------------|
| HLT_Photon75_CaloIdVL_v* |
| HLT_Photon75_CaloIdVL_IsoL_v* |
| HLT_Photon90_CaloIdVL_v* |
| HLT_Photon90_CaloIdVL_IsoL_v* |
| HLT_Photon125_v* |
| HLT_Photon135_v* |

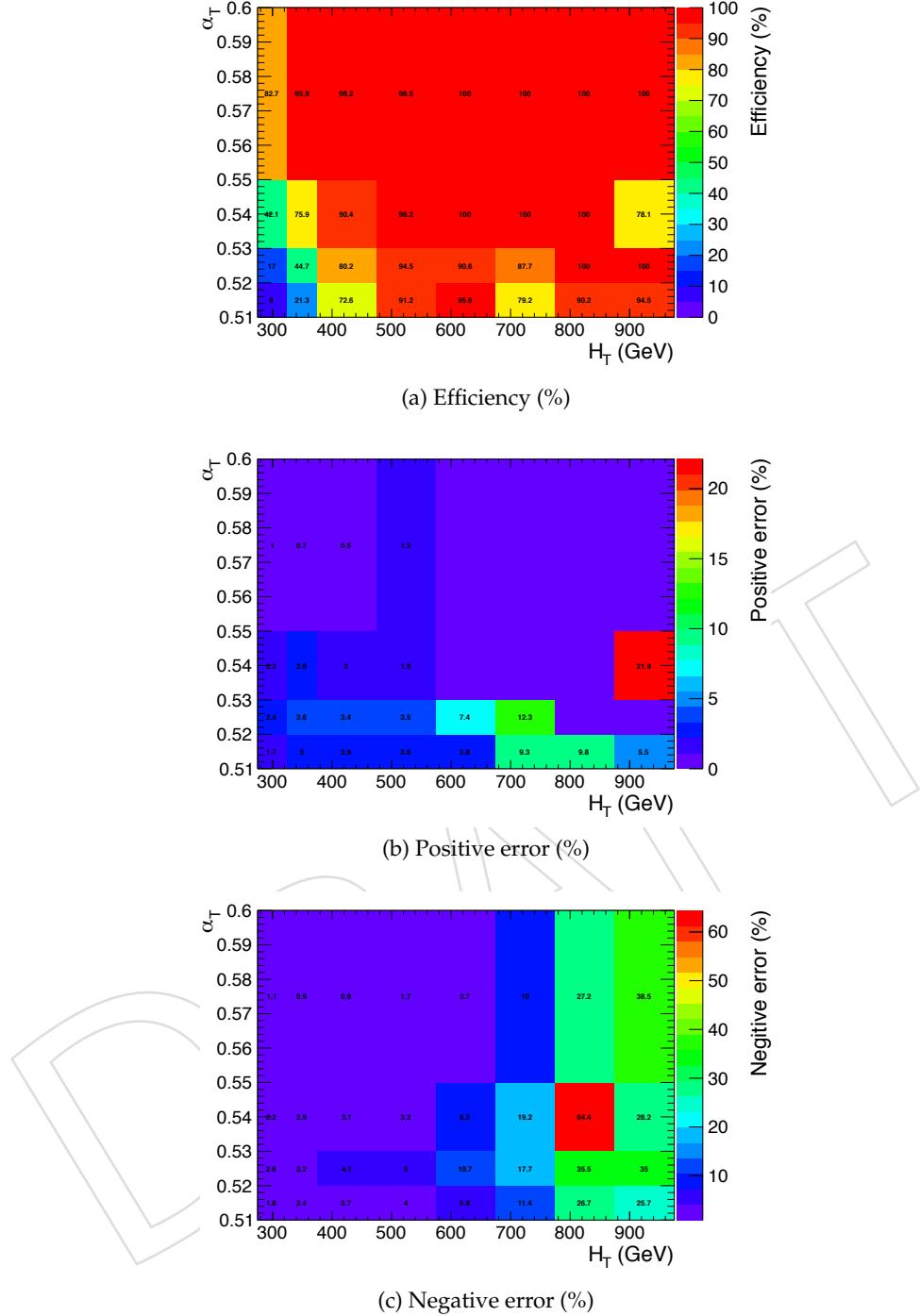


Figure 20: Efficiency and associated errors of the H_T _AlphaT trigger in offline bins of H_T and α_T .

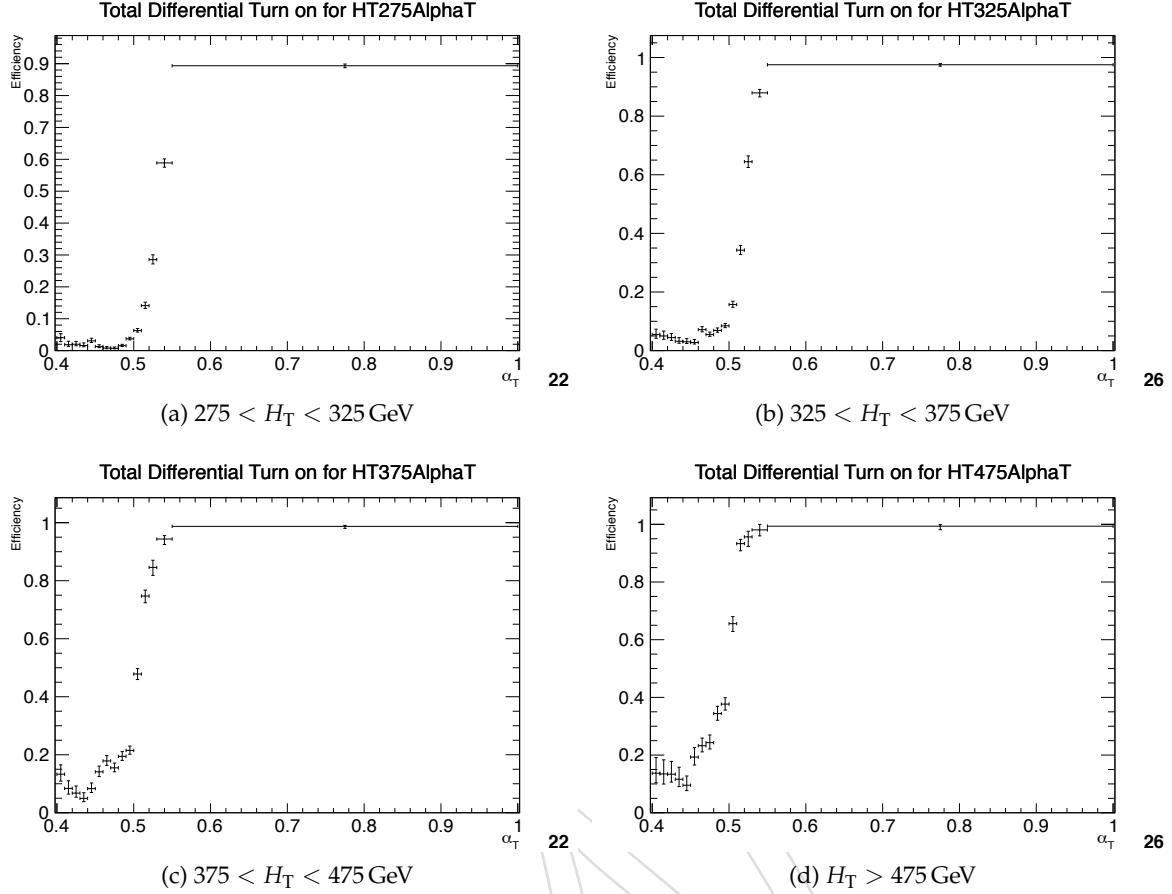


Figure 21: Efficiency turn-on curves for the α_T triggers used to collect events for four different HT regions.

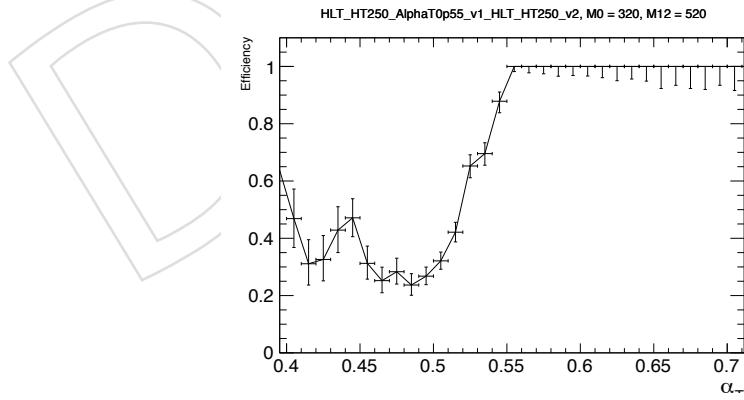


Figure 22: Efficiency turn-on curve for the representative model RM1, with $m_0 = 320 \text{ GeV}$ and $m_{1/2} = 520 \text{ GeV}$, using the α_T trigger with thresholds $H_T > 250 \text{ GeV}$ and $\alpha_T > 0.55$ and an offline signal region defined by $H_T > 275 \text{ GeV}$ and $\alpha_T > 0.55$.

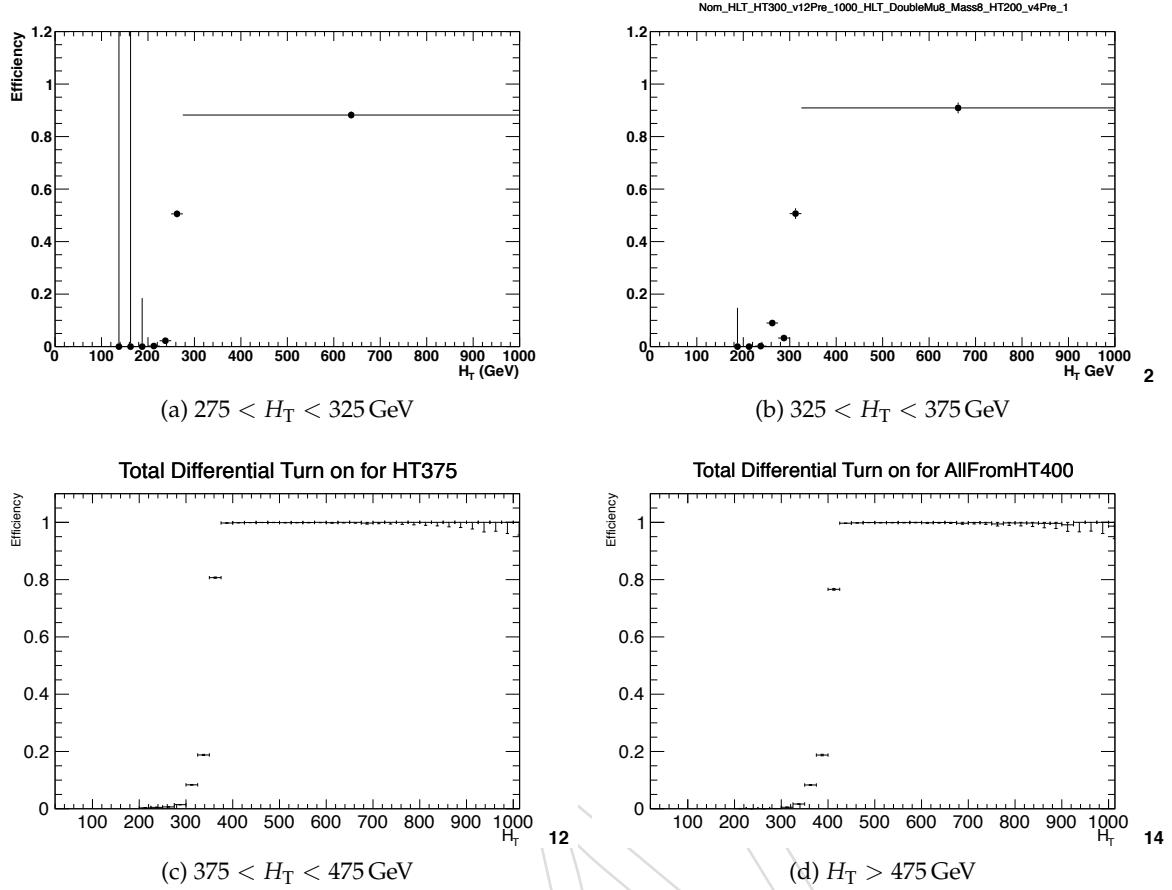


Figure 23: Efficiency turn-on curves for the H_T triggers used to collect events for four different H_T regions.

1255 **C Additional information for the control samples**

1256 **C.1 Distributions of b-tag multiplicity**

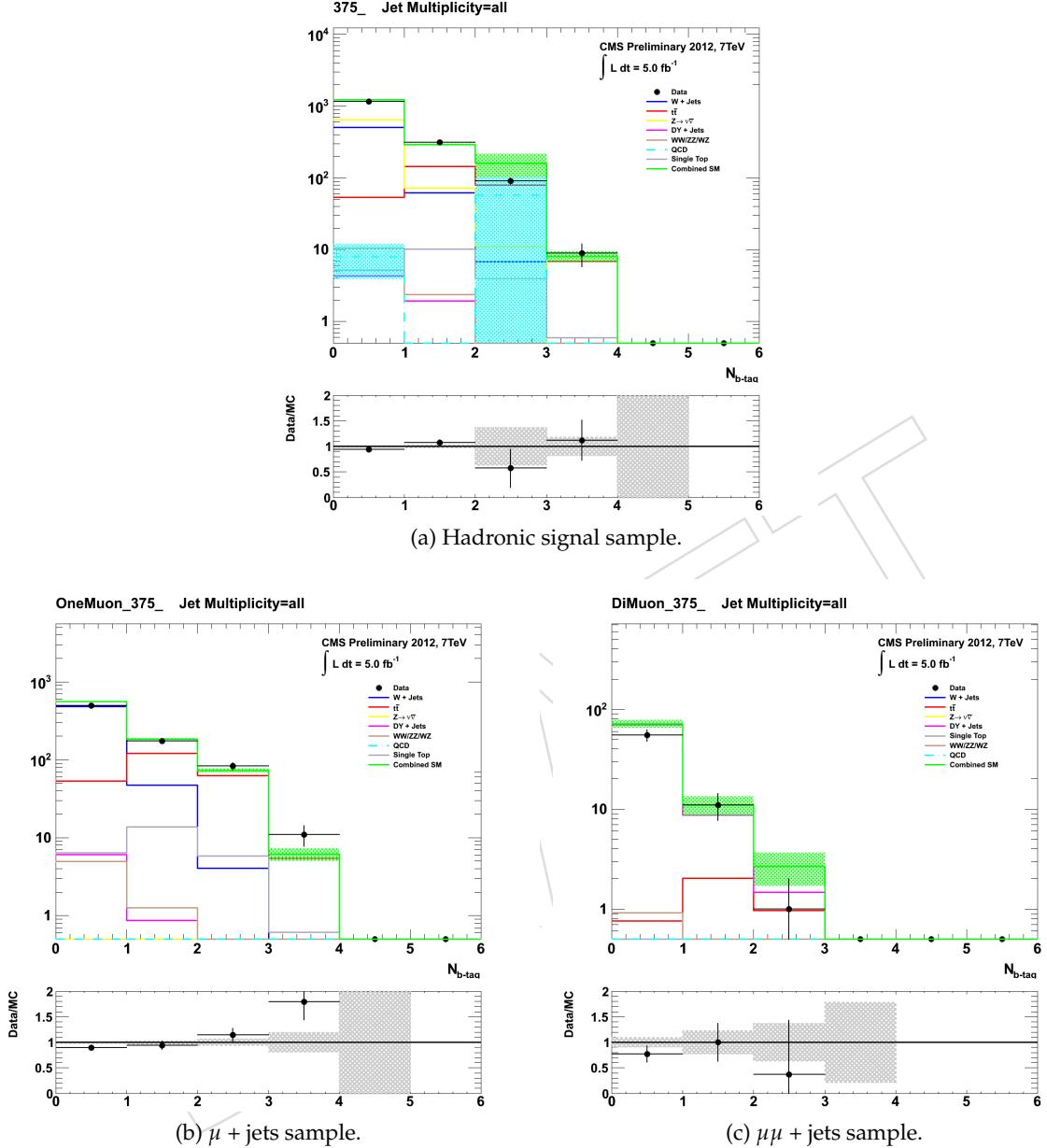


Figure 24: Data-MC comparison of the number of b-tagged jets per event in the (a) hadronic signal sample, (b) $\mu + \text{jets}$ sample and (c) $\mu\mu + \text{jets}$ control sample. Bands represent the uncertainties due to the limited size of MC samples.

1257

C.2 Distributions and predictions from the $\mu + \text{jets}$ sample

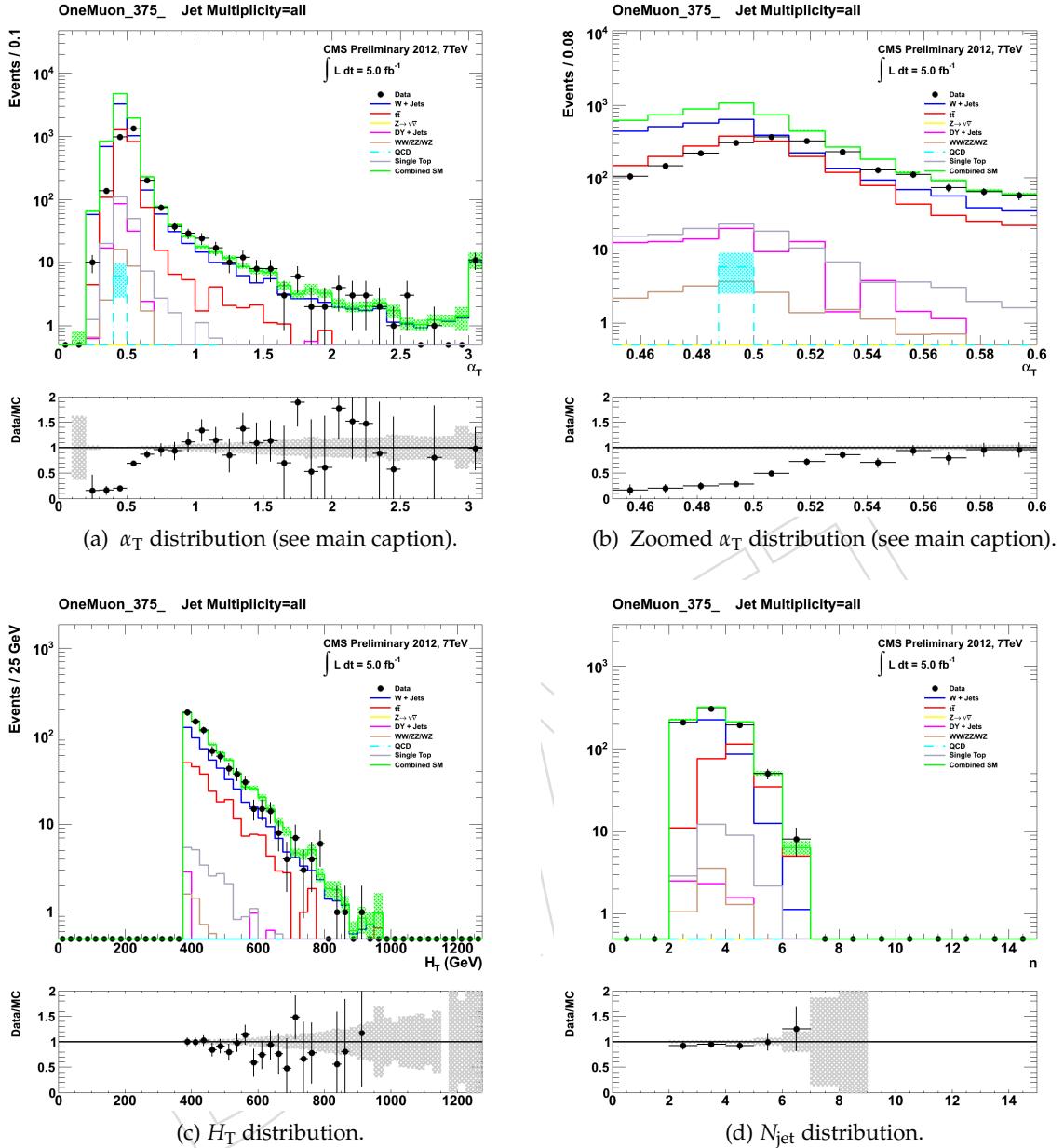


Figure 25: Data–MC comparisons of key variables for the muon control sample, for the region $H_T > 375 \text{ GeV}$ and $\alpha_T > 0.55$. Bands represent the uncertainties due to the limited size of MC samples. No requirement is made on the number of b-tagged jets in an event. *The discrepancy in the α_T distributions for values $\alpha_T < 0.55$ is due to the trigger not being simulated in the MC simulation.*

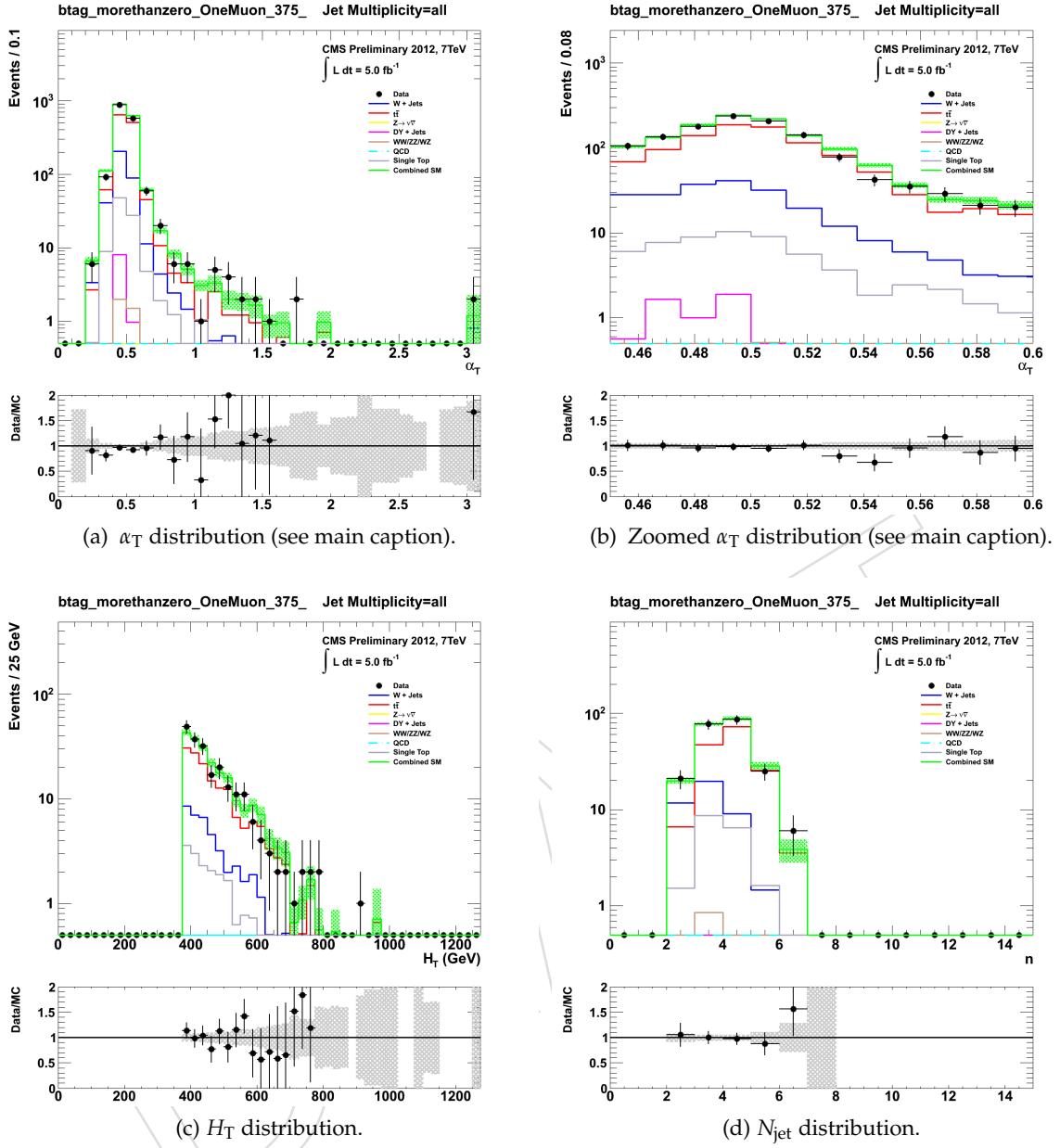


Figure 26: Data-MC comparisons of key variables for the b-tagged muon control sample, for the region $H_T > 375 \text{ GeV}$ and $\alpha_T > 0.55$. Bands represent the uncertainties due to the limited size of MC samples. At least one b-tagged jet per event is required. *The discrepancy in the α_T distributions for values $\alpha_T < 0.55$ is due to the trigger not being simulated in the MC simulation.*

Table 20: Predicted event yields in the hadronic signal region from $W + \text{jets}$, $t\bar{t}$ and single top processes, using the $\mu + \text{jets}$ control sample and requiring exactly zero b-tags per event.

| α_T bin | 0.55–∞ | 0.55–∞ | 0.55–∞ | 0.55–∞ |
|------------------------------------------|----------------------|--------------------|---------------------|--------------------|
| H_T bin (GeV) | 275–325 | 325–375 | 375–475 | 475–575 |
| $t\bar{t} + W$ Hadronic selection MC | 1371.69 ± 65.45 | 609.21 ± 53.68 | 392.15 ± 6.42 | 135.29 ± 7.23 |
| $\mu + \text{jets}$ selection MC | 994.22 ± 54.36 | 541.33 ± 42.49 | 1805.57 ± 19.72 | 823.49 ± 10.58 |
| Translation factor | 1.38 ± 0.10 | 1.13 ± 0.13 | 0.22 ± 0.00 | 0.16 ± 0.01 |
| $\mu + \text{jets}$ selection yield data | 949 | 444 | 1707 | 748 |
| $t\bar{t} + W$ prediction | 1309.31 ± 104.09 | 499.67 ± 63.55 | 370.74 ± 11.57 | 122.89 ± 8.12 |

| α_T bin | 0.55–∞ | 0.55–∞ | 0.55–∞ | 0.55–∞ |
|------------------------------------------|-------------------|-------------------|------------------|------------------|
| H_T bin (GeV) | 575–675 | 675–775 | 775–875 | 875–∞ |
| $t\bar{t} + W$ Hadronic selection MC | 46.53 ± 3.00 | 16.70 ± 1.39 | 6.07 ± 0.63 | 3.88 ± 0.46 |
| $\mu + \text{jets}$ selection MC | 359.51 ± 5.81 | 172.02 ± 3.76 | 82.76 ± 2.29 | 99.68 ± 2.61 |
| Translation factor | 0.13 ± 0.01 | 0.10 ± 0.01 | 0.07 ± 0.01 | 0.04 ± 0.00 |
| $\mu + \text{jets}$ selection yield data | 305 | 148 | 81 | 87 |
| $t\bar{t} + W$ prediction | 39.48 ± 3.46 | 14.36 ± 1.71 | 5.94 ± 0.92 | 3.39 ± 0.55 |

Table 21: Predicted event yields in the hadronic signal region from $W + \text{jets}$, $t\bar{t}$ and single top processes, using the $\mu + \text{jets}$ control sample and requiring exactly one b-tag per event.

| α_T bin | 0.55–∞ | 0.55–∞ | 0.55–∞ | 0.55–∞ |
|------------------------------------------|--------------------|--------------------|--------------------|--------------------|
| H_T bin (GeV) | 275–325 | 325–375 | 375–475 | 475–575 |
| $t\bar{t} + W$ Hadronic selection MC | 440.72 ± 15.82 | 209.67 ± 9.60 | 158.78 ± 9.89 | 64.47 ± 9.27 |
| $\mu + \text{jets}$ selection MC | 386.73 ± 15.05 | 195.07 ± 9.97 | 637.30 ± 15.68 | 321.54 ± 12.33 |
| Translation factor | 1.14 ± 0.06 | 1.07 ± 0.07 | 0.25 ± 0.02 | 0.20 ± 0.03 |
| $\mu + \text{jets}$ selection yield data | 347 | 146 | 568 | 288 |
| $t\bar{t} + W$ prediction | 395.43 ± 29.81 | 156.92 ± 16.87 | 141.51 ± 11.19 | 57.75 ± 9.24 |

| α_T bin | 0.55–∞ | 0.55–∞ | 0.55–∞ | 0.55–∞ |
|------------------------------------------|--------------------|------------------|------------------|------------------|
| H_T bin (GeV) | 575–675 | 675–775 | 775–875 | 875–∞ |
| $t\bar{t} + W$ Hadronic selection MC | 20.73 ± 3.86 | 10.34 ± 2.80 | 2.20 ± 0.74 | 1.67 ± 0.45 |
| $\mu + \text{jets}$ selection MC | 146.40 ± 10.00 | 64.44 ± 8.78 | 30.25 ± 4.16 | 35.21 ± 3.90 |
| Translation factor | 0.14 ± 0.03 | 0.16 ± 0.05 | 0.07 ± 0.03 | 0.05 ± 0.01 |
| $\mu + \text{jets}$ selection yield data | 116 | 48 | 22 | 26 |
| $t\bar{t} + W$ prediction | 16.43 ± 3.59 | 7.70 ± 2.59 | 1.60 ± 0.67 | 1.24 ± 0.43 |

Table 22: Predicted event yields in the hadronic signal region from $W +$ jets, $t\bar{t}$ and single top processes, using the $\mu +$ jets control sample and requiring exactly two b-tags per event.

| α_T bin | 0.55–∞ | 0.55–∞ | 0.55–∞ | 0.55–∞ |
|--------------------------------------|--------------------|------------------|-------------------|-------------------|
| H_T bin (GeV) | 275–325 | 325–375 | 375–475 | 475–575 |
| $t\bar{t} + W$ Hadronic selection MC | 128.63 ± 5.56 | 58.67 ± 3.99 | 52.86 ± 2.98 | 24.68 ± 2.40 |
| $\mu +$ jets selection MC | 122.36 ± 6.78 | 64.33 ± 4.87 | 255.01 ± 7.16 | 138.20 ± 5.36 |
| Translation factor | 1.05 ± 0.07 | 0.91 ± 0.09 | 0.21 ± 0.01 | 0.18 ± 0.02 |
| $\mu +$ jets selection yield data | 116 | 49 | 264 | 152 |
| $t\bar{t} + W$ prediction | 121.94 ± 14.20 | 44.69 ± 7.84 | 54.73 ± 4.82 | 27.15 ± 3.59 |

| α_T bin | 0.55–∞ | 0.55–∞ | 0.55–∞ | 0.55–∞ |
|--------------------------------------|------------------|------------------|------------------|------------------|
| H_T bin (GeV) | 575–675 | 675–775 | 775–875 | 875–∞ |
| $t\bar{t} + W$ Hadronic selection MC | 8.05 ± 1.28 | 5.92 ± 1.55 | 0.76 ± 0.40 | 0.72 ± 0.35 |
| $\mu +$ jets selection MC | 68.90 ± 4.29 | 26.68 ± 2.24 | 13.74 ± 1.59 | 13.27 ± 1.59 |
| Translation factor | 0.12 ± 0.02 | 0.22 ± 0.06 | 0.06 ± 0.03 | 0.05 ± 0.03 |
| $\mu +$ jets selection yield data | 63 | 26 | 10 | 14 |
| $t\bar{t} + W$ prediction | 7.36 ± 1.56 | 5.77 ± 1.95 | 0.55 ± 0.35 | 0.76 ± 0.43 |

Table 23: Predicted event yields in the hadronic signal region from $W +$ jets, $t\bar{t}$ and single top processes, using the $\mu +$ jets control sample and requiring at least three b-tags per event.

| α_T bin | 0.55–∞ | 0.55–∞ | 0.55–∞ | 0.55–∞ |
|--------------------------------------|------------------|-----------------|------------------|------------------|
| H_T bin (GeV) | 275–325 | 325–375 | 375–475 | 475–575 |
| $t\bar{t} + W$ Hadronic selection MC | 10.12 ± 0.32 | 4.74 ± 0.23 | 5.02 ± 0.21 | 3.39 ± 0.25 |
| $\mu +$ jets selection MC | 9.56 ± 0.37 | 4.70 ± 0.24 | 20.54 ± 0.48 | 14.50 ± 0.41 |
| Translation factor | 1.06 ± 0.05 | 1.01 ± 0.07 | 0.24 ± 0.01 | 0.23 ± 0.02 |
| $\mu +$ jets selection yield data | 9 | 6 | 22 | 16 |
| $t\bar{t} + W$ prediction | 9.52 ± 3.21 | 6.05 ± 2.48 | 5.38 ± 1.18 | 3.74 ± 0.98 |

| α_T bin | 0.55–∞ | 0.55–∞ | 0.55–∞ | 0.55–∞ |
|--------------------------------------|-----------------|-----------------|-----------------|-----------------|
| H_T bin (GeV) | 575–675 | 675–775 | 775–875 | 875–∞ |
| $t\bar{t} + W$ Hadronic selection MC | 1.28 ± 0.15 | 1.07 ± 0.16 | 0.13 ± 0.05 | 0.13 ± 0.04 |
| $\mu +$ jets selection MC | 8.73 ± 0.41 | 3.46 ± 0.18 | 2.07 ± 0.15 | 2.65 ± 0.19 |
| Translation factor | 0.15 ± 0.02 | 0.31 ± 0.05 | 0.06 ± 0.02 | 0.05 ± 0.01 |
| $\mu +$ jets selection yield data | 13 | 3 | 1 | 4 |
| $t\bar{t} + W$ prediction | 1.91 ± 0.58 | 0.93 ± 0.55 | 0.06 ± 0.09 | 0.19 ± 0.11 |

1258

C.3 Distributions and predictions from the $\mu\mu + \text{jets}$ sample

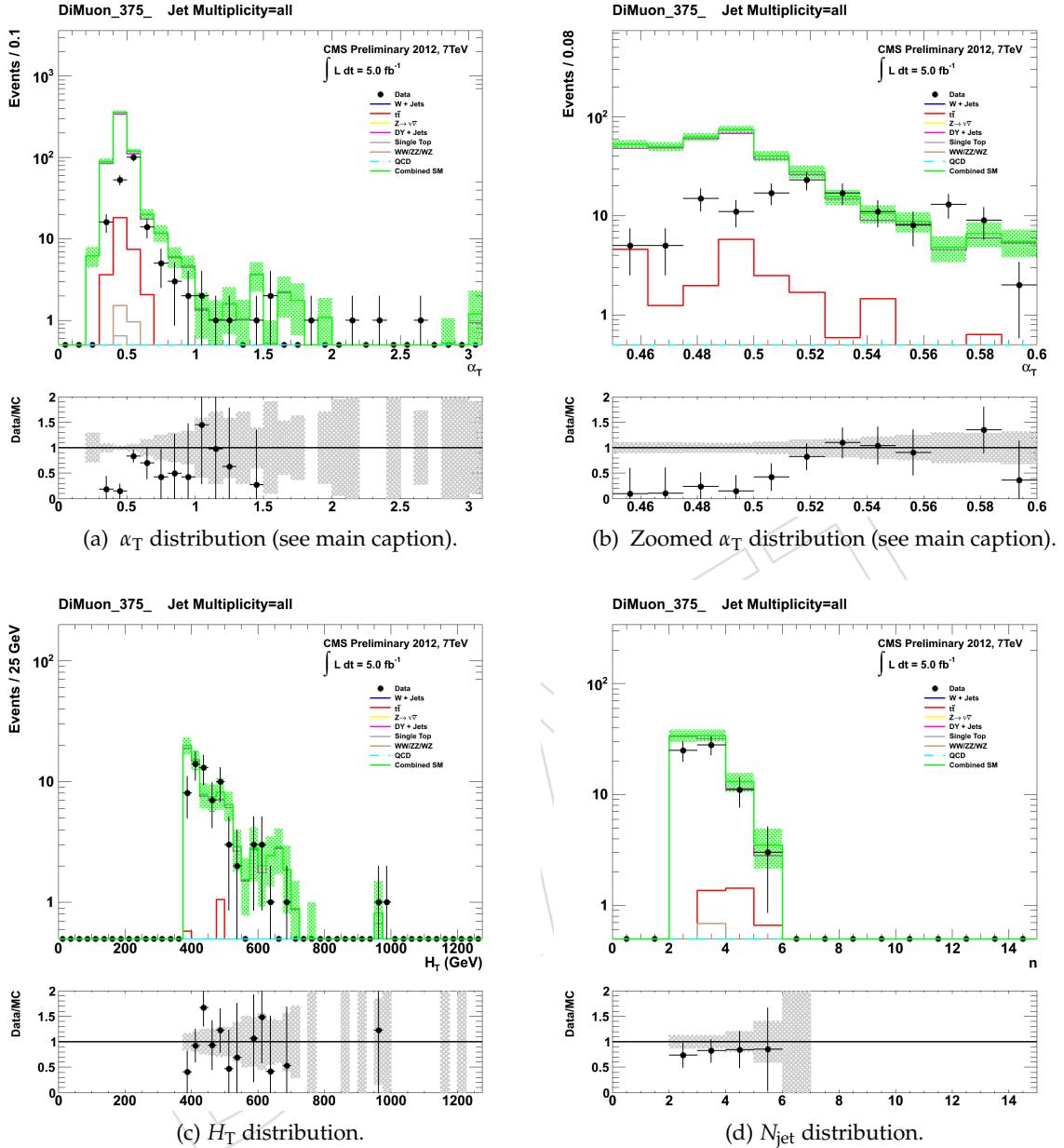


Figure 27: Data–MC comparisons of key variables for the muon control sample, for the region $H_T > 375 \text{ GeV}$ and $\alpha_T > 0.55$. Bands represent the uncertainties due to the limited size of MC samples. No requirement is made on the number of b-tagged jets in an event. *The discrepancy in the α_T distributions for values $\alpha_T < 0.55$ is due to the trigger not being simulated in the MC simulation.*

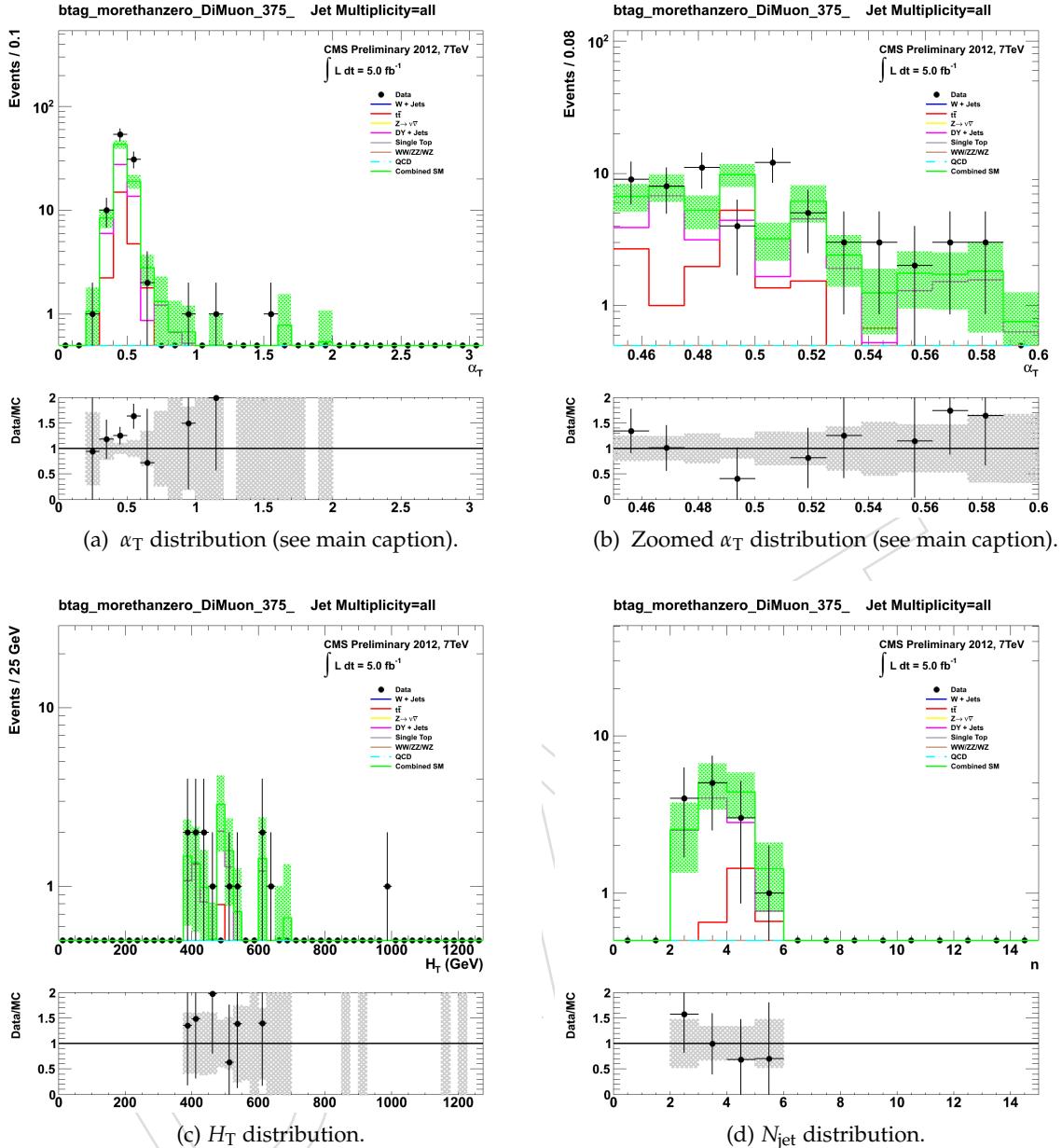


Figure 28: Data–MC comparisons of key variables for the b-tagged di-muon control sample, for the region $H_T > 375 \text{ GeV}$ and $\alpha_T > 0.55$. Bands represent the uncertainties due to the limited size of MC samples. At least one b-tagged jet per event is required. *The discrepancy in the α_T distributions for values $\alpha_T < 0.55$ is due to the trigger not being simulated in the MC simulation.*

Table 24: Predicted event yields in the hadronic signal region from $Z \rightarrow \nu\bar{\nu} + \text{jets}$ process, using the $\mu\mu + \text{jets}$ control sample and requiring exactly zero b-tag per event.

| α_T bin | 0.55–∞ | 0.55–∞ | 0.55–∞ | 0.55–∞ |
|----------------------------------------------------|----------------------|--------------------|--------------------|--------------------|
| H_T bin (GeV) | 275–325 | 325–375 | 375–475 | 475–575 |
| $Z \rightarrow \nu\bar{\nu}$ hadronic selection MC | 1271.56 ± 21.47 | 638.88 ± 9.49 | 495.89 ± 8.43 | 176.90 ± 9.45 |
| $\mu\bar{\mu} + \text{jets}$ selection MC | 99.11 ± 6.24 | 67.16 ± 5.63 | 251.54 ± 10.85 | 117.57 ± 7.44 |
| Translation factor | 12.83 ± 0.84 | 9.51 ± 0.81 | 1.97 ± 0.09 | 1.50 ± 0.12 |
| $\mu\bar{\mu} + \text{jets}$ selection yield data | 95 | 53 | 216 | 86 |
| $Z \rightarrow \nu\bar{\nu}$ prediction | 1218.86 ± 148.15 | 504.16 ± 81.49 | 425.83 ± 35.07 | 129.40 ± 17.59 |

| α_T bin | 0.55–∞ | 0.55–∞ | 0.55–∞ | 0.55–∞ |
|----------------------------------------------------|-------------------|------------------|------------------|------------------|
| H_T bin (GeV) | 575–675 | 675–775 | 775–875 | 875–∞ |
| $Z \rightarrow \nu\bar{\nu}$ hadronic selection MC | 67.75 ± 4.11 | 22.44 ± 1.77 | 9.45 ± 0.97 | 6.43 ± 0.73 |
| $\mu\bar{\mu} + \text{jets}$ selection MC | 51.70 ± 4.71 | 23.31 ± 3.09 | 13.44 ± 3.28 | 10.62 ± 1.99 |
| Translation factor | 1.31 ± 0.14 | 0.96 ± 0.15 | 0.70 ± 0.19 | 0.60 ± 0.13 |
| $\mu\bar{\mu} + \text{jets}$ selection yield data | 48 | 23 | 5 | 11 |
| $Z \rightarrow \nu\bar{\nu}$ prediction | 62.90 ± 11.39 | 22.14 ± 5.74 | 3.51 ± 1.81 | 6.65 ± 2.48 |

Table 25: Predicted event yields in the hadronic signal region from $Z \rightarrow \nu\bar{\nu} + \text{jets}$ process, using the $\mu\mu + \text{jets}$ control sample and requiring exactly one b-tag per event.

| α_T bin | 0.55–∞ | 0.55–∞ | 0.55–∞ | 0.55–∞ |
|----------------------------------------------------|--------------------|-------------------|-------------------|------------------|
| H_T bin (GeV) | 275–325 | 325–375 | 375–475 | 475–575 |
| $Z \rightarrow \nu\bar{\nu}$ hadronic selection MC | 156.21 ± 3.34 | 80.79 ± 1.93 | 61.36 ± 1.41 | 24.19 ± 1.49 |
| $\mu\bar{\mu} + \text{jets}$ selection MC | 16.30 ± 2.24 | 8.91 ± 1.93 | 31.90 ± 2.87 | 15.97 ± 1.71 |
| Translation factor | 9.58 ± 1.33 | 9.07 ± 1.98 | 1.92 ± 0.18 | 1.51 ± 0.19 |
| $\mu\bar{\mu} + \text{jets}$ selection yield data | 15 | 9 | 34 | 20 |
| $Z \rightarrow \nu\bar{\nu}$ prediction | 143.72 ± 42.15 | 81.64 ± 32.53 | 65.41 ± 12.75 | 30.30 ± 7.74 |

| α_T bin | 0.55–∞ | 0.55–∞ | 0.55–∞ | 0.55–∞ |
|----------------------------------------------------|------------------|-----------------|------------------------|-----------------|
| H_T bin (GeV) | 575–675 | 675–775 | 775–875 | 875–∞ |
| $Z \rightarrow \nu\bar{\nu}$ hadronic selection MC | 9.67 ± 0.69 | 3.58 ± 0.35 | 1.41 ± 0.12 | 1.04 ± 0.13 |
| $\mu\bar{\mu} + \text{jets}$ selection MC | 8.73 ± 1.01 | 3.81 ± 1.24 | 1.09 ± 0.11 | 1.91 ± 1.69 |
| Translation factor | 1.11 ± 0.15 | 0.94 ± 0.32 | 1.29 ± 0.17 | 0.54 ± 0.49 |
| $\mu\bar{\mu} + \text{jets}$ selection yield data | 10 | 7 | 0 | 6 |
| $Z \rightarrow \nu\bar{\nu}$ prediction | 11.08 ± 3.81 | 6.58 ± 3.32 | $0.00^{+1.15}_{-0.00}$ | 3.27 ± 3.20 |

Table 26: Predicted event yields in the hadronic signal region from $Z \rightarrow \nu\bar{\nu} + \text{jets}$ process, using the $\mu\mu + \text{jets}$ control sample and requiring exactly two b-tags per event.

| α_T bin | 0.55–∞ | 0.55–∞ | 0.55–∞ | 0.55–∞ |
|----------------------------------------------------|-------------------|------------------|------------------------|------------------------|
| H_T bin (GeV) | 275–325 | 325–375 | 375–475 | 475–575 |
| $Z \rightarrow \nu\bar{\nu}$ hadronic selection MC | 16.82 ± 0.87 | 8.29 ± 0.62 | 6.58 ± 0.57 | 2.28 ± 0.32 |
| $\mu\bar{\mu} + \text{jets}$ selection MC | 3.03 ± 0.70 | 2.93 ± 3.19 | 8.07 ± 1.32 | 2.44 ± 0.61 |
| Translation factor | 5.55 ± 1.31 | 2.83 ± 3.08 | 0.81 ± 0.15 | 0.93 ± 0.27 |
| $\mu\bar{\mu} + \text{jets}$ selection yield data | 4 | 3 | 8 | 7 |
| $Z \rightarrow \nu\bar{\nu}$ prediction | 22.19 ± 12.17 | 8.48 ± 10.47 | 6.52 ± 2.58 | 6.54 ± 3.07 |
| α_T bin | 0.55–∞ | 0.55–∞ | 0.55–∞ | 0.55–∞ |
| H_T bin (GeV) | 575–675 | 675–775 | 775–875 | 875–∞ |
| $Z \rightarrow \nu\bar{\nu}$ hadronic selection MC | 0.99 ± 0.22 | 0.33 ± 0.10 | 0.10 ± 0.03 | 0.07 ± 0.02 |
| $\mu\bar{\mu} + \text{jets}$ selection MC | 1.12 ± 0.35 | 0.44 ± 0.31 | 0.15 ± 0.18 | 0.13 ± 0.04 |
| Translation factor | 0.89 ± 0.34 | 0.74 ± 0.56 | 0.69 ± 0.89 | 0.55 ± 0.22 |
| $\mu\bar{\mu} + \text{jets}$ selection yield data | 5 | 2 | 0 | 0 |
| $Z \rightarrow \nu\bar{\nu}$ prediction | 4.43 ± 2.60 | 1.48 ± 1.60 | $0.00^{+1.15}_{-0.00}$ | $0.00^{+1.15}_{-0.00}$ |

DRAFT

¹²⁵⁹ **C.4 Distributions and predictions from for the $\gamma + \text{jets}$ sample**

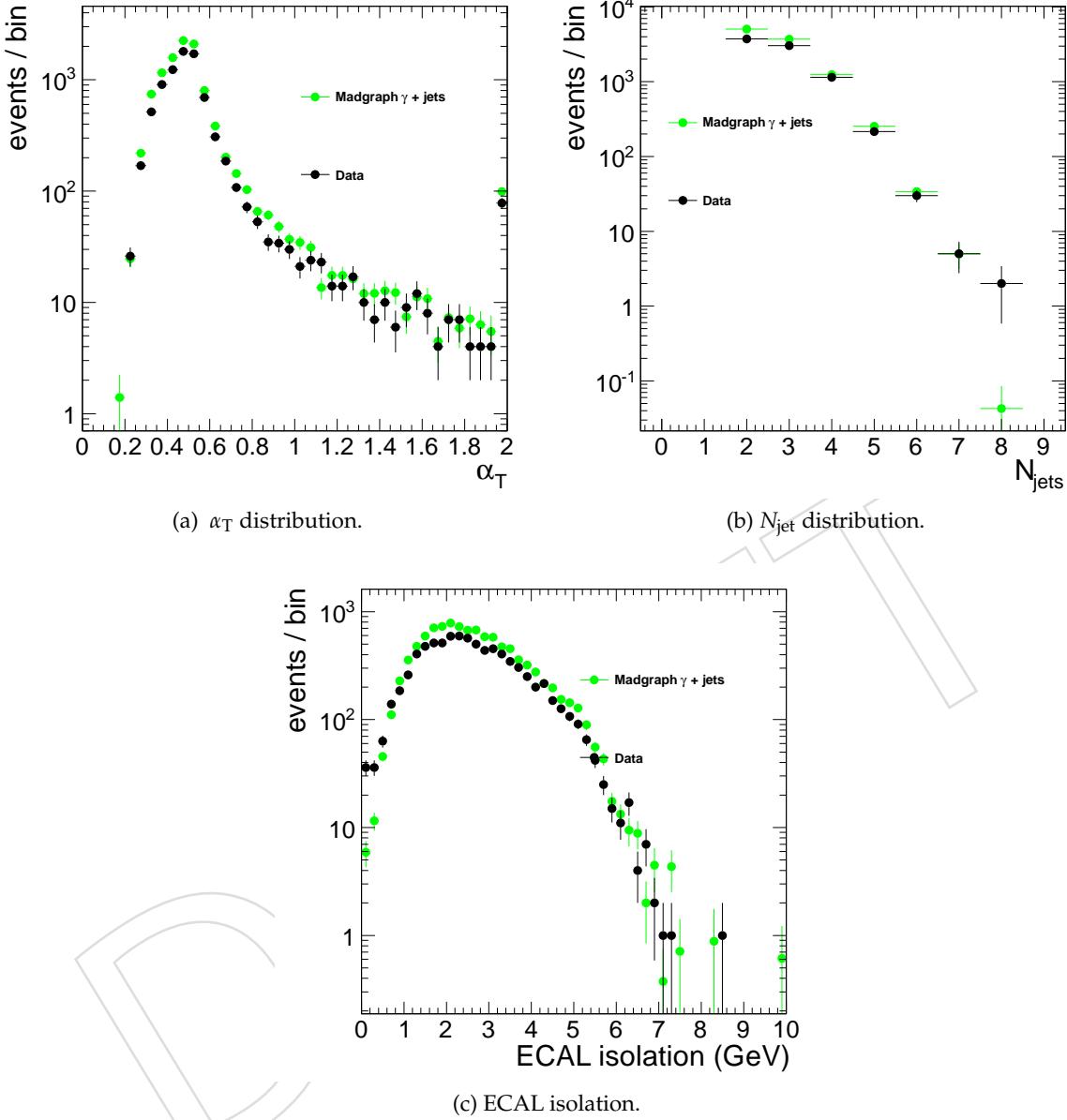


Figure 29: Data–MC comparisons of key variables for the photon control sample, for the region $H_T > 375 \text{ GeV}$. Errors on the MC expectations represent the uncertainties due to the limited size of MC samples. No requirement is made on the number of b-tagged jets in an event.

Table 27: Predicted event yields in the hadronic signal region from $Z \rightarrow \nu\bar{\nu} +$ jets process, using the $\gamma +$ jets control sample and requiring exactly zero b-tags per event.

| α_T bin | 0.55– ∞ | 0.55– ∞ | 0.55– ∞ | 0.55– ∞ |
|----------------------------------------------------|---------------------|-------------------|---------------------|--------------------|
| H_T bin (GeV) | 275–325 | 325–375 | 375–475 | 475–575 |
| $Z \rightarrow \nu\bar{\nu}$ hadronic selection MC | 1271.56 ± 21.47 | 638.88 ± 9.49 | 495.89 ± 8.43 | 176.90 ± 9.45 |
| $\gamma +$ jets selection MC | - | - | 1185.00 ± 28.41 | 389.90 ± 15.62 |
| Translation factor | - | - | 0.42 ± 0.01 | 0.45 ± 0.03 |
| $\gamma +$ jet selection yield data | - | - | 909 | 328 |
| $Z \rightarrow \nu\bar{\nu}$ prediction | - | - | 380.39 ± 16.86 | 148.82 ± 12.89 |

| α_T bin | 0.55– ∞ | 0.55– ∞ | 0.55– ∞ | 0.55– ∞ |
|----------------------------------------------------|-------------------|------------------|------------------|------------------|
| H_T bin (GeV) | 575–675 | 675–775 | 775–875 | 875– ∞ |
| $Z \rightarrow \nu\bar{\nu}$ hadronic selection MC | 67.75 ± 4.11 | 22.44 ± 1.77 | 9.45 ± 0.97 | 6.43 ± 0.73 |
| $\gamma +$ jets selection MC | 163.90 ± 9.98 | 50.86 ± 5.53 | 17.09 ± 3.16 | 13.90 ± 2.82 |
| Translation factor | 0.41 ± 0.04 | 0.44 ± 0.06 | 0.55 ± 0.12 | 0.46 ± 0.11 |
| $\gamma +$ jet selection yield data | 109 | 50 | 13 | 12 |
| $Z \rightarrow \nu\bar{\nu}$ prediction | 45.06 ± 5.80 | 22.06 ± 4.30 | 7.18 ± 2.51 | 5.55 ± 2.06 |

Table 28: Predicted event yields in the hadronic signal region from $Z \rightarrow \nu\bar{\nu} +$ jets process, using the $\gamma +$ jets control sample and requiring exactly one b-tag per event.

| α_T bin | 0.55– ∞ | 0.55– ∞ | 0.55– ∞ | 0.55– ∞ |
|----------------------------------------------------|-------------------|------------------|-------------------|------------------|
| H_T bin (GeV) | 275–325 | 325–375 | 375–475 | 475–575 |
| $Z \rightarrow \nu\bar{\nu}$ hadronic selection MC | 156.21 ± 3.34 | 80.79 ± 1.93 | 61.36 ± 1.41 | 24.19 ± 1.49 |
| $\gamma +$ jets selection MC | - | - | 139.50 ± 4.32 | 50.32 ± 3.32 |
| Translation factor | - | - | 0.44 ± 0.02 | 0.48 ± 0.04 |
| $\gamma +$ jet selection yield data | - | - | 126 | 43 |
| $Z \rightarrow \nu\bar{\nu}$ prediction | - | - | 55.42 ± 5.38 | 20.67 ± 3.66 |

| α_T bin | 0.55– ∞ | 0.55– ∞ | 0.55– ∞ | 0.55– ∞ |
|----------------------------------------------------|------------------|-----------------|-----------------|-----------------|
| H_T bin (GeV) | 575–675 | 675–775 | 775–875 | 875– ∞ |
| $Z \rightarrow \nu\bar{\nu}$ hadronic selection MC | 9.67 ± 0.69 | 3.58 ± 0.35 | 1.41 ± 0.12 | 1.04 ± 0.13 |
| $\gamma +$ jets selection MC | 22.63 ± 1.49 | 7.50 ± 0.15 | 2.92 ± 1.79 | 2.52 ± 1.64 |
| Translation factor | 0.43 ± 0.04 | 0.48 ± 0.05 | 0.48 ± 0.30 | 0.41 ± 0.28 |
| $\gamma +$ jet selection yield data | 19 | 5 | 5 | 2 |
| $Z \rightarrow \nu\bar{\nu}$ prediction | 8.12 ± 2.02 | 2.38 ± 1.08 | 2.41 ± 1.84 | 0.83 ± 0.84 |

Table 29: Predicted event yields in the hadronic signal region from $Z \rightarrow \nu\bar{\nu} + \text{jets}$ process, using the $\gamma + \text{jets}$ control sample and requiring exactly two b-tags per event.

| α_T bin | 0.55–∞ | 0.55–∞ | 0.55–∞ | 0.55–∞ |
|----------------------------------------------------|------------------|-----------------|------------------|-----------------|
| H_T bin (GeV) | 275–325 | 325–375 | 375–475 | 475–575 |
| $Z \rightarrow \nu\bar{\nu}$ hadronic selection MC | 16.82 ± 0.87 | 8.29 ± 0.62 | 6.58 ± 0.57 | 2.28 ± 0.32 |
| $\gamma + \text{jets}$ selection MC | - | - | 11.54 ± 2.11 | 3.96 ± 0.94 |
| Translation factor | - | - | 0.57 ± 0.12 | 0.57 ± 0.16 |
| $\gamma + \text{jet}$ selection yield data | - | - | 10 | 4 |
| $Z \rightarrow \nu\bar{\nu}$ prediction | - | - | 5.70 ± 2.14 | 2.30 ± 1.30 |

| α_T bin | 0.55–∞ | 0.55–∞ | 0.55–∞ | 0.55–∞ |
|----------------------------------------------------|-----------------|-----------------|------------------------|------------------------|
| H_T bin (GeV) | 575–675 | 675–775 | 775–875 | 875–∞ |
| $Z \rightarrow \nu\bar{\nu}$ hadronic selection MC | 0.99 ± 0.22 | 0.33 ± 0.10 | 0.10 ± 0.03 | 0.07 ± 0.02 |
| $\gamma + \text{jets}$ selection MC | 2.24 ± 0.88 | 0.72 ± 0.37 | 0.21 ± 0.07 | 0.21 ± 0.06 |
| Translation factor | 0.44 ± 0.20 | 0.46 ± 0.27 | 0.49 ± 0.22 | 0.33 ± 0.12 |
| $\gamma + \text{jet}$ selection yield data | 2 | 4 | 0 | 0 |
| $Z \rightarrow \nu\bar{\nu}$ prediction | 0.88 ± 0.78 | 1.83 ± 1.42 | $0.00^{+1.15}_{-0.00}$ | $0.00^{+1.15}_{-0.00}$ |

DRAFT

C.5 Illustrative total predictions for SM backgrounds
C.5.1 Predictions using the $\mu + \text{jets}$, $\mu\mu + \text{jets}$ and $\gamma + \text{jets}$ samples

Table 30: Total SM predictions for the eight H_T bins of the hadronic signal region and requiring exactly zero b-tags. The predictions are based on the $\mu + \text{jets}$ and $\mu\mu + \text{jets}$ samples, and are illustrative only, as the SM expectation for the final result is provided by the simultaneous fit.

| α_T bin | 0.55– ∞ | 0.55– ∞ | 0.55– ∞ | 0.55– ∞ |
|-----------------------------------------------------------------|----------------------|----------------------|--------------------|--------------------|
| H_T bin (GeV) | 275–325 | 325–375 | 375–475 | 475–575 |
| t \bar{t} + W prediction from $\mu + \text{jets}$ | 1309.31 ± 104.09 | 499.67 ± 63.55 | 370.74 ± 11.57 | 122.89 ± 8.12 |
| Z $\rightarrow v\bar{v}$ prediction from $\mu\mu + \text{jets}$ | 1218.86 ± 148.15 | 504.16 ± 81.49 | 425.83 ± 35.07 | 129.40 ± 17.59 |
| Z $\rightarrow v\bar{v}$ prediction from $\gamma + \text{jets}$ | - | - | 354.46 ± 18.12 | 136.77 ± 12.63 |
| Total SM prediction (μ and $\mu\mu$ samples) | 2528.17 ± 181.06 | 1003.83 ± 103.34 | 796.57 ± 36.92 | 252.28 ± 19.38 |
| Total SM prediction (μ and γ samples) | - | - | 691.68 ± 21.08 | 254.94 ± 14.93 |
| Hadronic yield from data | 2919 | 1166 | 769 | 255 |

| α_T bin | 0.55– ∞ | 0.55– ∞ | 0.55– ∞ | 0.55– ∞ |
|-----------------------------------------------------------------|--------------------|------------------|------------------|------------------|
| H_T bin (GeV) | 575–675 | 675–775 | 775–875 | 875– ∞ |
| t \bar{t} + W prediction from $\mu + \text{jets}$ | 39.48 ± 3.46 | 14.36 ± 1.71 | 5.94 ± 0.92 | 3.39 ± 0.55 |
| Z $\rightarrow v\bar{v}$ prediction from $\mu\mu + \text{jets}$ | 62.90 ± 11.39 | 22.14 ± 5.74 | 3.51 ± 1.81 | 6.65 ± 2.48 |
| Z $\rightarrow v\bar{v}$ prediction from $\gamma + \text{jets}$ | 45.06 ± 5.80 | 22.06 ± 4.30 | 7.18 ± 2.51 | 5.55 ± 2.06 |
| Total SM prediction (μ and $\mu\mu$ samples) | 102.37 ± 11.91 | 36.51 ± 5.99 | 9.45 ± 2.03 | 10.04 ± 2.54 |
| Total SM prediction (μ and γ samples) | 84.54 ± 6.75 | 36.42 ± 4.63 | 13.12 ± 2.67 | 8.93 ± 2.13 |
| Hadronic yield from data | 91 | 31 | 10 | 4 |

Table 31: Total SM predictions for the eight H_T bins of the hadronic signal region and requiring exactly one b-tag. The predictions are based on the $\mu + \text{jets}$ and $\mu\mu + \text{jets}$ samples, and are illustrative only, as the SM expectation for the final result is provided by the simultaneous fit.

| α_T bin | 0.55– ∞ | 0.55– ∞ | 0.55– ∞ | 0.55– ∞ |
|-----------------------------------------------------------------|--------------------|--------------------|--------------------|-------------------|
| H_T bin (GeV) | 275–325 | 325–375 | 375–475 | 475–575 |
| t \bar{t} + W prediction from $\mu + \text{jets}$ | 395.43 ± 29.81 | 156.92 ± 16.87 | 141.51 ± 11.19 | 57.75 ± 9.24 |
| Z $\rightarrow v\bar{v}$ prediction from $\mu\mu + \text{jets}$ | 143.72 ± 42.15 | 81.64 ± 32.53 | 65.41 ± 12.75 | 30.30 ± 7.74 |
| Z $\rightarrow v\bar{v}$ prediction from $\gamma + \text{jets}$ | - | - | 46.99 ± 5.91 | 22.95 ± 5.72 |
| Total SM prediction (μ and $\mu\mu$ samples) | 539.16 ± 51.63 | 238.56 ± 36.65 | 206.92 ± 16.96 | 88.05 ± 12.05 |
| Total SM prediction (μ and γ samples) | - | - | 184.58 ± 10.20 | 75.08 ± 7.94 |
| Hadronic yield from data | 614 | 294 | 214 | 71 |

| α_T bin | 0.55– ∞ | 0.55– ∞ | 0.55– ∞ | 0.55– ∞ |
|-----------------------------------------------------------------|------------------|------------------|-----------------|-----------------|
| H_T bin (GeV) | 575–675 | 675–775 | 775–875 | 875– ∞ |
| t \bar{t} + W prediction from $\mu + \text{jets}$ | 16.43 ± 3.59 | 7.70 ± 2.59 | 1.60 ± 0.67 | 1.24 ± 0.43 |
| Z $\rightarrow v\bar{v}$ prediction from $\mu\mu + \text{jets}$ | 11.08 ± 3.81 | 6.58 ± 3.32 | - | 3.27 ± 3.20 |
| Z $\rightarrow v\bar{v}$ prediction from $\gamma + \text{jets}$ | 8.12 ± 2.02 | 2.38 ± 1.08 | 2.41 ± 1.84 | 0.83 ± 0.84 |
| Total SM prediction (μ and $\mu\mu$ samples) | 27.51 ± 5.24 | 14.28 ± 4.20 | 1.60 ± 0.67 | 4.50 ± 3.23 |
| Total SM prediction (μ and γ samples) | 24.55 ± 4.12 | 10.09 ± 2.80 | 4.02 ± 1.96 | 2.06 ± 0.94 |
| Hadronic yield from data | 20 | 6 | 4 | 0 |

Table 32: Total SM predictions for the eight H_T bins of the hadronic signal region and requiring exactly two b-tags. The predictions are based on the $\mu + \text{jets}$ and $\mu\mu + \text{jets}$ samples, and are illustrative only, as the SM expectation for the final result is provided by the simultaneous fit.

| α_T bin | 0.55– ∞ | 0.55– ∞ | 0.55– ∞ | 0.55– ∞ |
|---------------------------------------------------------------------|--------------------|-------------------|------------------|------------------|
| H_T bin (GeV) | 275–325 | 325–375 | 375–475 | 475–575 |
| $t\bar{t} + W$ prediction from $\mu + \text{jets}$ | 121.94 ± 14.20 | 44.69 ± 7.84 | 54.73 ± 4.82 | 27.15 ± 3.59 |
| $Z \rightarrow \nu\bar{\nu}$ prediction from $\mu\mu + \text{jets}$ | 22.19 ± 12.17 | 8.48 ± 10.47 | 6.52 ± 2.58 | 6.54 ± 3.07 |
| $Z \rightarrow \nu\bar{\nu}$ prediction from $\gamma + \text{jets}$ | - | - | 4.88 ± 2.04 | 1.96 ± 1.22 |
| Total SM prediction (μ and $\mu\mu$ samples) | 144.14 ± 18.70 | 53.17 ± 13.08 | 61.25 ± 5.47 | 33.68 ± 4.73 |
| Total SM prediction (μ and γ samples) | - | - | 51.49 ± 5.41 | 29.36 ± 4.71 |
| Hadronic yield from data | 160 | 68 | 52 | 19 |

| α_T bin | 0.55– ∞ | 0.55– ∞ | 0.55– ∞ | 0.55– ∞ |
|---------------------------------------------------------------------|------------------|-----------------|-----------------|-----------------|
| H_T bin (GeV) | 575–675 | 675–775 | 775–875 | 875– ∞ |
| $t\bar{t} + W$ prediction from $\mu + \text{jets}$ | 7.36 ± 1.56 | 5.77 ± 1.95 | 0.55 ± 0.35 | 0.76 ± 0.43 |
| $Z \rightarrow \nu\bar{\nu}$ prediction from $\mu\mu + \text{jets}$ | 4.43 ± 2.60 | 1.48 ± 1.60 | 0.00 ± 0.00 | 0.00 ± 0.00 |
| $Z \rightarrow \nu\bar{\nu}$ prediction from $\gamma + \text{jets}$ | 0.88 ± 0.78 | 1.83 ± 1.42 | 0.00 ± 0.00 | 0.00 ± 0.00 |
| Total SM prediction (μ and $\mu\mu$ samples) | 11.78 ± 3.03 | 7.25 ± 2.52 | 0.55 ± 0.35 | 0.76 ± 0.43 |
| Total SM prediction (μ and γ samples) | 8.24 ± 1.75 | 7.60 ± 2.41 | 0.55 ± 0.35 | 0.76 ± 0.43 |
| Hadronic yield from data | 11 | 7 | 0 | 2 |

Table 33: Total SM predictions for the eight H_T bins of the hadronic signal region and requiring at least three b-tags. The predictions are based on the $\mu + \text{jets}$ sample only, and are illustrative only, as the SM expectation for the final result is provided by the simultaneous fit.

| α_T bin | 0.55– ∞ | 0.55– ∞ | 0.55– ∞ | 0.55– ∞ |
|------------------------------------------|------------------|-----------------|------------------|------------------|
| H_T bin (GeV) | 275–325 | 325–375 | 375–475 | 475–575 |
| Hadronic selection MC | 10.61 ± 0.32 | 4.98 ± 0.23 | 5.20 ± 0.21 | 3.46 ± 0.25 |
| $\mu + \text{jets}$ selection MC | 9.56 ± 0.37 | 4.70 ± 0.24 | 20.54 ± 0.48 | 14.50 ± 0.41 |
| Translation factor | 1.11 ± 0.05 | 1.06 ± 0.07 | 0.25 ± 0.01 | 0.24 ± 0.02 |
| $\mu + \text{jets}$ selection yield data | 9 | 6 | 22 | 16 |
| Total SM prediction | 9.99 ± 3.37 | 6.35 ± 2.60 | 5.57 ± 1.22 | 3.82 ± 1.00 |
| Hadronic yield data | 10 | 8 | 8 | 1 |

| α_T bin | 0.55– ∞ | 0.55– ∞ | 0.55– ∞ | 0.55– ∞ |
|------------------------------------------|-----------------|-----------------|-----------------|-----------------|
| H_T bin (GeV) | 575–675 | 675–775 | 775–875 | 875– ∞ |
| Hadronic selection MC | 1.33 ± 0.15 | 1.09 ± 0.16 | 0.13 ± 0.05 | 0.13 ± 0.04 |
| $\mu + \text{jets}$ selection MC | 8.73 ± 0.41 | 3.46 ± 0.18 | 2.07 ± 0.15 | 2.65 ± 0.19 |
| Translation factor | 0.15 ± 0.02 | 0.31 ± 0.05 | 0.06 ± 0.02 | 0.05 ± 0.01 |
| $\mu + \text{jets}$ selection yield data | 13 | 3 | 1 | 4 |
| Total SM prediction | 1.98 ± 0.60 | 0.94 ± 0.56 | 0.06 ± 0.09 | 0.19 ± 0.11 |
| Hadronic yield data | 0 | 0 | 0 | 0 |

¹²⁶² **C.5.2 Predictions using the $\mu + \text{jets}$ sample only**

Table 34: Total SM predictions using the $\mu + \text{jets}$ sample only, while requiring exactly zero b-tags per event. These predictions are illustrative only, as the SM expectation for the final result is provided by the simultaneous fit.

| α_T bin | 0.55– ∞ | 0.55– ∞ | 0.55– ∞ | 0.55– ∞ |
|------------------------------------------|----------------------|----------------------|---------------------|--------------------|
| H_T bin (GeV) | 275–325 | 325–375 | 375–475 | 475–575 |
| Hadronic selection MC | 2643.25 ± 74.32 | 1248.09 ± 55.16 | 888.04 ± 12.27 | 312.19 ± 15.87 |
| $\mu + \text{jets}$ selection MC | 994.22 ± 54.36 | 541.33 ± 42.49 | 1805.57 ± 19.72 | 823.49 ± 10.58 |
| Translation factor | 2.66 ± 0.16 | 2.31 ± 0.21 | 0.49 ± 0.01 | 0.38 ± 0.02 |
| $\mu + \text{jets}$ selection yield data | 949 | 444 | 1707 | 748 |
| Total SM prediction | 2523.04 ± 175.41 | 1023.68 ± 104.22 | 839.56 ± 25.13 | 283.57 ± 18.13 |
| Hadronic yield data | 2919 | 1166 | 769 | 255 |

| α_T bin | 0.55– ∞ | 0.55– ∞ | 0.55– ∞ | 0.55– ∞ |
|------------------------------------------|-------------------|-------------------|------------------|------------------|
| H_T bin (GeV) | 575–675 | 675–775 | 775–875 | 875– ∞ |
| Hadronic selection MC | 114.28 ± 6.36 | 39.14 ± 2.61 | 15.51 ± 1.27 | 10.30 ± 0.93 |
| $\mu + \text{jets}$ selection MC | 359.51 ± 5.81 | 172.02 ± 3.76 | 82.76 ± 2.29 | 99.68 ± 2.61 |
| Translation factor | 0.32 ± 0.02 | 0.23 ± 0.02 | 0.19 ± 0.02 | 0.10 ± 0.01 |
| $\mu + \text{jets}$ selection yield data | 305 | 148 | 81 | 87 |
| Total SM prediction | 96.96 ± 7.90 | 33.67 ± 3.64 | 15.18 ± 2.13 | 8.99 ± 1.28 |
| Hadronic yield data | 91 | 31 | 10 | 4 |

Table 35: Total SM predictions using the $\mu + \text{jets}$ sample only, while requiring exactly one b-tag per event. These predictions are illustrative only, as the SM expectation for the final result is provided by the simultaneous fit.

| α_T bin | 0.55– ∞ | 0.55– ∞ | 0.55– ∞ | 0.55– ∞ |
|------------------------------------------|--------------------|--------------------|--------------------|--------------------|
| H_T bin (GeV) | 275–325 | 325–375 | 375–475 | 475–575 |
| Hadronic selection MC | 596.92 ± 17.09 | 290.46 ± 9.95 | 220.14 ± 10.09 | 88.66 ± 9.76 |
| $\mu + \text{jets}$ selection MC | 386.73 ± 15.05 | 195.07 ± 9.97 | 637.30 ± 15.68 | 321.54 ± 12.33 |
| Translation factor | 1.54 ± 0.07 | 1.49 ± 0.09 | 0.35 ± 0.02 | 0.28 ± 0.03 |
| $\mu + \text{jets}$ selection yield data | 347 | 146 | 568 | 288 |
| Total SM prediction | 535.59 ± 38.68 | 217.39 ± 22.42 | 196.20 ± 13.11 | 79.42 ± 10.37 |
| Hadronic yield data | 614 | 294 | 214 | 71 |

| α_T bin | 0.55– ∞ | 0.55– ∞ | 0.55– ∞ | 0.55– ∞ |
|------------------------------------------|--------------------|------------------|------------------|------------------|
| H_T bin (GeV) | 575–675 | 675–775 | 775–875 | 875– ∞ |
| Hadronic selection MC | 30.40 ± 4.03 | 13.92 ± 2.85 | 3.61 ± 0.76 | 2.72 ± 0.48 |
| $\mu + \text{jets}$ selection MC | 146.40 ± 10.00 | 64.44 ± 8.78 | 30.25 ± 4.16 | 35.21 ± 3.90 |
| Translation factor | 0.21 ± 0.03 | 0.22 ± 0.05 | 0.12 ± 0.03 | 0.08 ± 0.02 |
| $\mu + \text{jets}$ selection yield data | 116 | 48 | 22 | 26 |
| Total SM prediction | 24.09 ± 4.23 | 10.37 ± 2.96 | 2.63 ± 0.86 | 2.01 ± 0.57 |
| Hadronic yield data | 20 | 6 | 4 | 0 |

Table 36: Total SM predictions using the $\mu + \text{jets}$ sample only, while requiring exactly two b-tags per event. These predictions are illustrative only, as the SM expectation for the final result is provided by the simultaneous fit.

| α_T bin | 0.55– ∞ | 0.55– ∞ | 0.55– ∞ | 0.55– ∞ |
|------------------------------------------|--------------------|------------------|-------------------|-------------------|
| H_T bin (GeV) | 275–325 | 325–375 | 375–475 | 475–575 |
| Hadronic selection MC | 145.45 ± 5.71 | 66.96 ± 4.05 | 59.44 ± 3.04 | 26.96 ± 2.47 |
| $\mu + \text{jets}$ selection MC | 122.36 ± 6.78 | 64.33 ± 4.87 | 255.01 ± 7.16 | 138.20 ± 5.36 |
| Translation factor | 1.19 ± 0.08 | 1.04 ± 0.10 | 0.23 ± 0.01 | 0.20 ± 0.02 |
| $\mu + \text{jets}$ selection yield data | 116 | 49 | 264 | 152 |
| Total SM prediction | 137.89 ± 15.86 | 51.01 ± 8.80 | 61.54 ± 5.22 | 29.65 ± 3.81 |
| Hadronic yield data | 160 | 68 | 52 | 19 |

| α_T bin | 0.55– ∞ | 0.55– ∞ | 0.55– ∞ | 0.55– ∞ |
|------------------------------------------|------------------|------------------|------------------|------------------|
| H_T bin (GeV) | 575–675 | 675–775 | 775–875 | 875– ∞ |
| Hadronic selection MC | 9.04 ± 1.32 | 6.25 ± 1.55 | 0.86 ± 0.40 | 0.79 ± 0.35 |
| $\mu + \text{jets}$ selection MC | 68.90 ± 4.29 | 26.68 ± 2.24 | 13.74 ± 1.59 | 13.27 ± 1.59 |
| Translation factor | 0.13 ± 0.02 | 0.23 ± 0.06 | 0.06 ± 0.03 | 0.06 ± 0.03 |
| $\mu + \text{jets}$ selection yield data | 63 | 26 | 10 | 14 |
| Total SM prediction | 8.26 ± 1.67 | 6.09 ± 2.00 | 0.63 ± 0.36 | 0.84 ± 0.44 |
| Hadronic yield data | 11 | 7 | 0 | 2 |

Table 37: Total SM predictions using the $\mu + \text{jets}$ sample only, while requiring at least three b-tags per event. These predictions are illustrative only, as the SM expectation for the final result is provided by the simultaneous fit.

| α_T bin | 0.55– ∞ | 0.55– ∞ | 0.55– ∞ | 0.55– ∞ |
|------------------------------------------|------------------|-----------------|------------------|------------------|
| H_T bin (GeV) | 275–325 | 325–375 | 375–475 | 475–575 |
| Hadronic selection MC | 10.61 ± 0.32 | 4.98 ± 0.23 | 5.20 ± 0.21 | 3.46 ± 0.25 |
| $\mu + \text{jets}$ selection MC | 9.56 ± 0.37 | 4.70 ± 0.24 | 20.54 ± 0.48 | 14.50 ± 0.41 |
| Translation factor | 1.11 ± 0.05 | 1.06 ± 0.07 | 0.25 ± 0.01 | 0.24 ± 0.02 |
| $\mu + \text{jets}$ selection yield data | 9 | 6 | 22 | 16 |
| Total SM prediction | 9.99 ± 3.37 | 6.35 ± 2.60 | 5.57 ± 1.22 | 3.82 ± 1.00 |
| Hadronic yield data | 10 | 8 | 8 | 1 |

| α_T bin | 0.55– ∞ | 0.55– ∞ | 0.55– ∞ | 0.55– ∞ |
|------------------------------------------|-----------------|-----------------|-----------------|-----------------|
| H_T bin (GeV) | 575–675 | 675–775 | 775–875 | 875– ∞ |
| Hadronic selection MC | 1.33 ± 0.15 | 1.09 ± 0.16 | 0.13 ± 0.05 | 0.13 ± 0.04 |
| $\mu + \text{jets}$ selection MC | 8.73 ± 0.41 | 3.46 ± 0.18 | 2.07 ± 0.15 | 2.65 ± 0.19 |
| Translation factor | 0.15 ± 0.02 | 0.31 ± 0.05 | 0.06 ± 0.02 | 0.05 ± 0.01 |
| $\mu + \text{jets}$ selection yield data | 13 | 3 | 1 | 4 |
| Total SM prediction | 1.98 ± 0.60 | 0.94 ± 0.56 | 0.06 ± 0.09 | 0.19 ± 0.11 |
| Hadronic yield data | 0 | 0 | 0 | 0 |

1263 **D Additional information on QCD multi-jet background estimation**

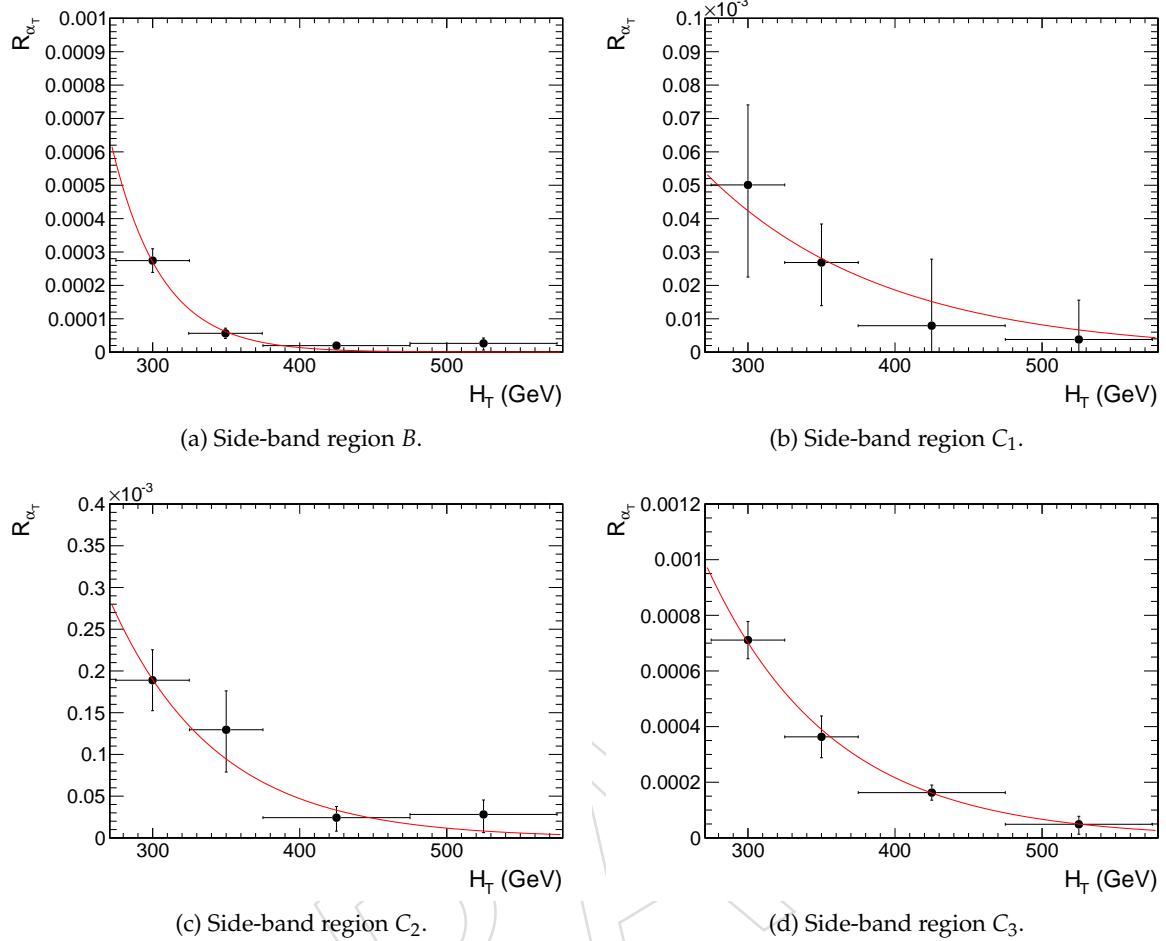


Figure 30: $R_{\alpha_T}(H_T)$ and exponential fit for various data side-bands. Linear y-axis scale.

1264 **E Closure tests and systematic uncertainties**

1265 **E.1 Defining muon samples without an α_T requirement**

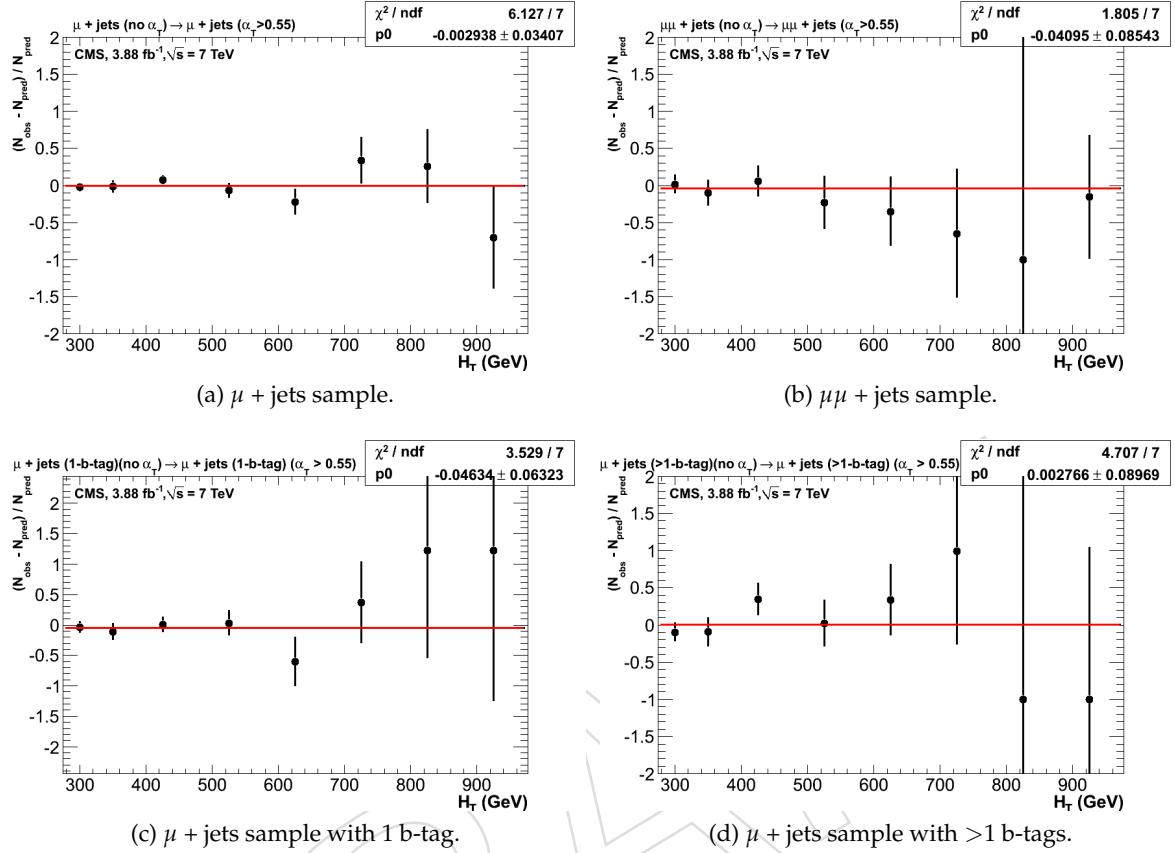


Figure 31: Closure tests that demonstrate the MC modelling of the α_T acceptance. The closure tests are performed for both the inclusive analysis with (a) the $\mu + \text{jets}$ sample and (b) the $\mu\mu + \text{jets}$ control sample. Similar tests are performed for the b-tag analysis using (a) the $\mu + \text{jets}$ sample and a requirement of exactly one b-tag, and (b) the $\mu + \text{jets}$ sample and a requirement of at least two b-tags. The red lines indicate the constant best fit value across all H_T bins.

1266 **E.2 Closure tests for inclusive analysis**

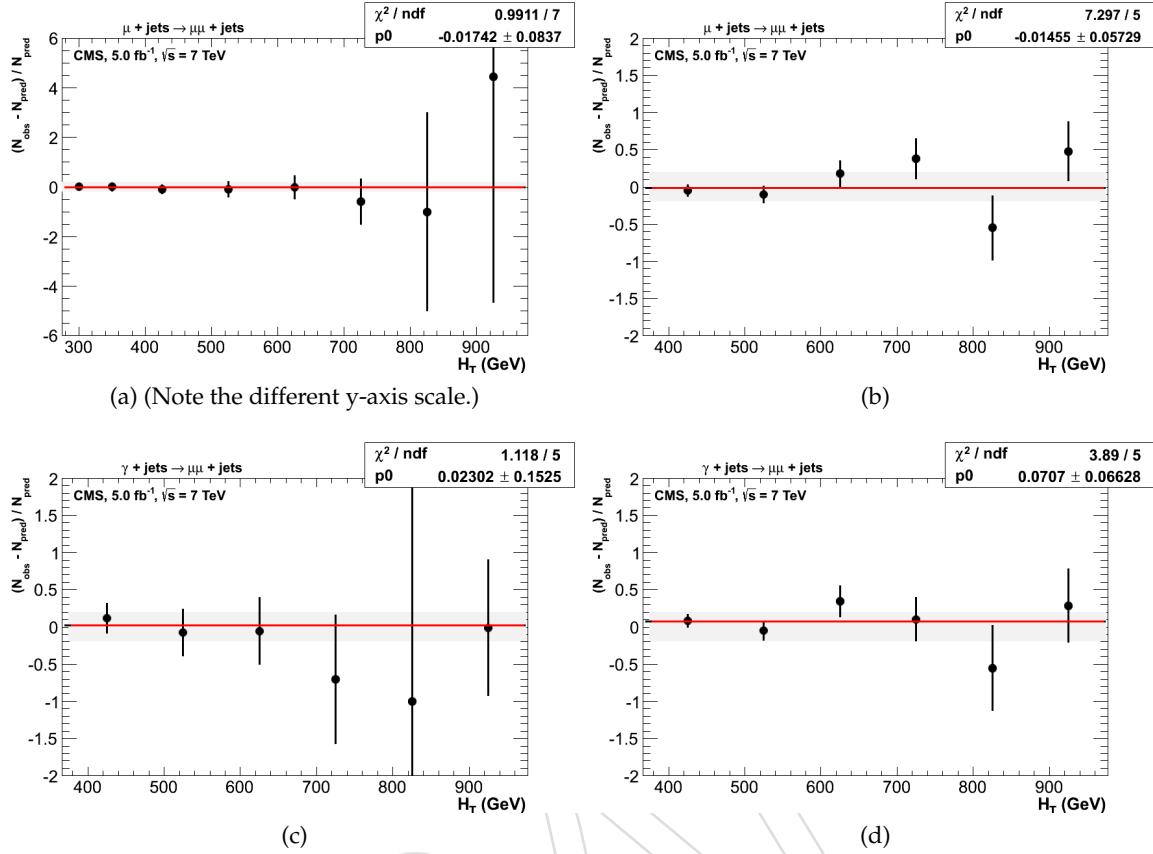


Figure 32: Closure tests using yields from one control to predict yields in another sample. The two plots on the left show closure tests which use “low stats” muon samples recorded with the HT_AlphaT triggers and defined by offline selection criteria that include an α_T requirement: (a) $\mu + \text{jets}$ sample $\rightarrow \mu\mu + \text{jets}$ sample and (c) $\gamma + \text{jets}$ sample $\rightarrow \mu\mu + \text{jets}$ sample. Similarly, the plots on the right show the same closure tests but using “high-stats” muon samples recorded with Mu_HT triggers and defined with no offline α_T requirement. The same tests are performed: (b) $\mu + \text{jets}$ sample $\rightarrow \mu\mu + \text{jets}$ sample and (d) $\gamma + \text{jets}$ sample $\rightarrow \mu\mu + \text{jets}$ sample. These closure tests are only possible for the six highest H_T bins due to the trigger conditions. The red lines indicate the constant best fit value across all H_T bins.

1267 **E.3 Closure tests for b-tag analysis**

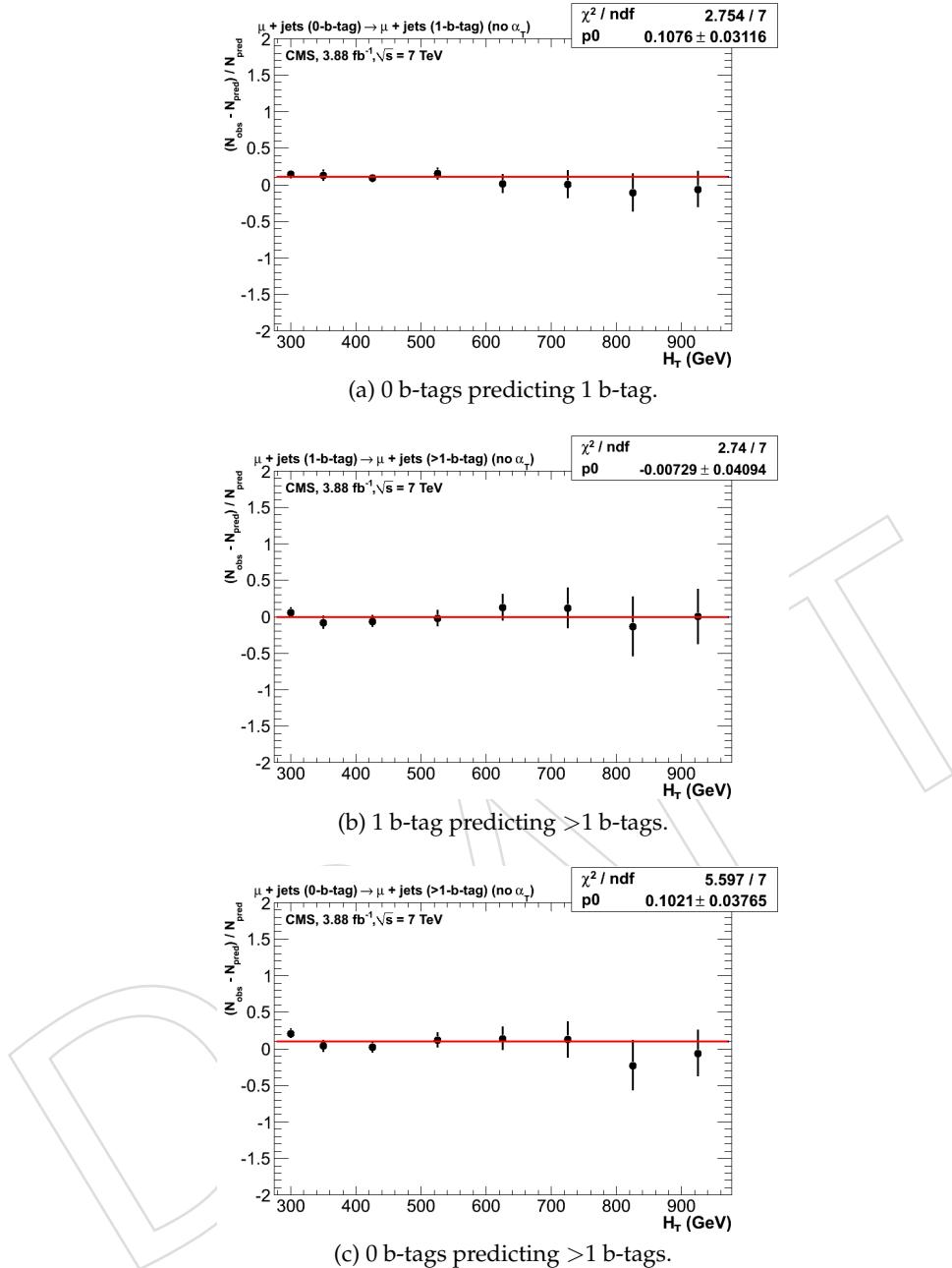


Figure 33: Closure tests with the $\mu + \text{jets}$ sample that demonstrate the MC modelling of the b-tagging algorithm and of different sample compositions by for different b-tag multiplicities: (a) 0 b-tags \rightarrow 1 b-tag, (b) 1 b-tags $\rightarrow \geq 2$ b-tags, (c) 0 b-tags $\rightarrow \geq 2$ b-tags.

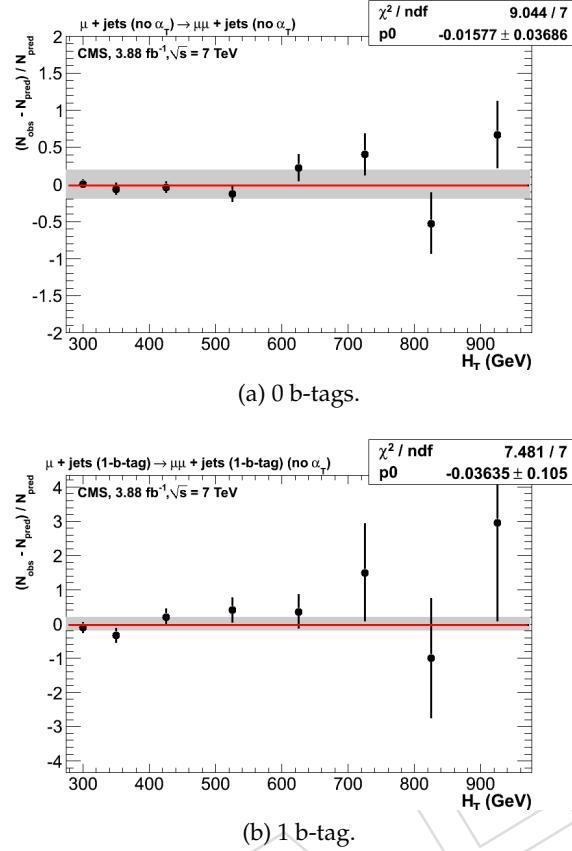


Figure 34: Closure tests using the $\mu + \text{jets}$ sample to predict the yields in a $\mu\mu + \text{jets}$ sample, for events with (a) exactly 0-b-tags and (b) exactly 1-b-tags.

1268 **E.4 Closure tests concerning pile-up**

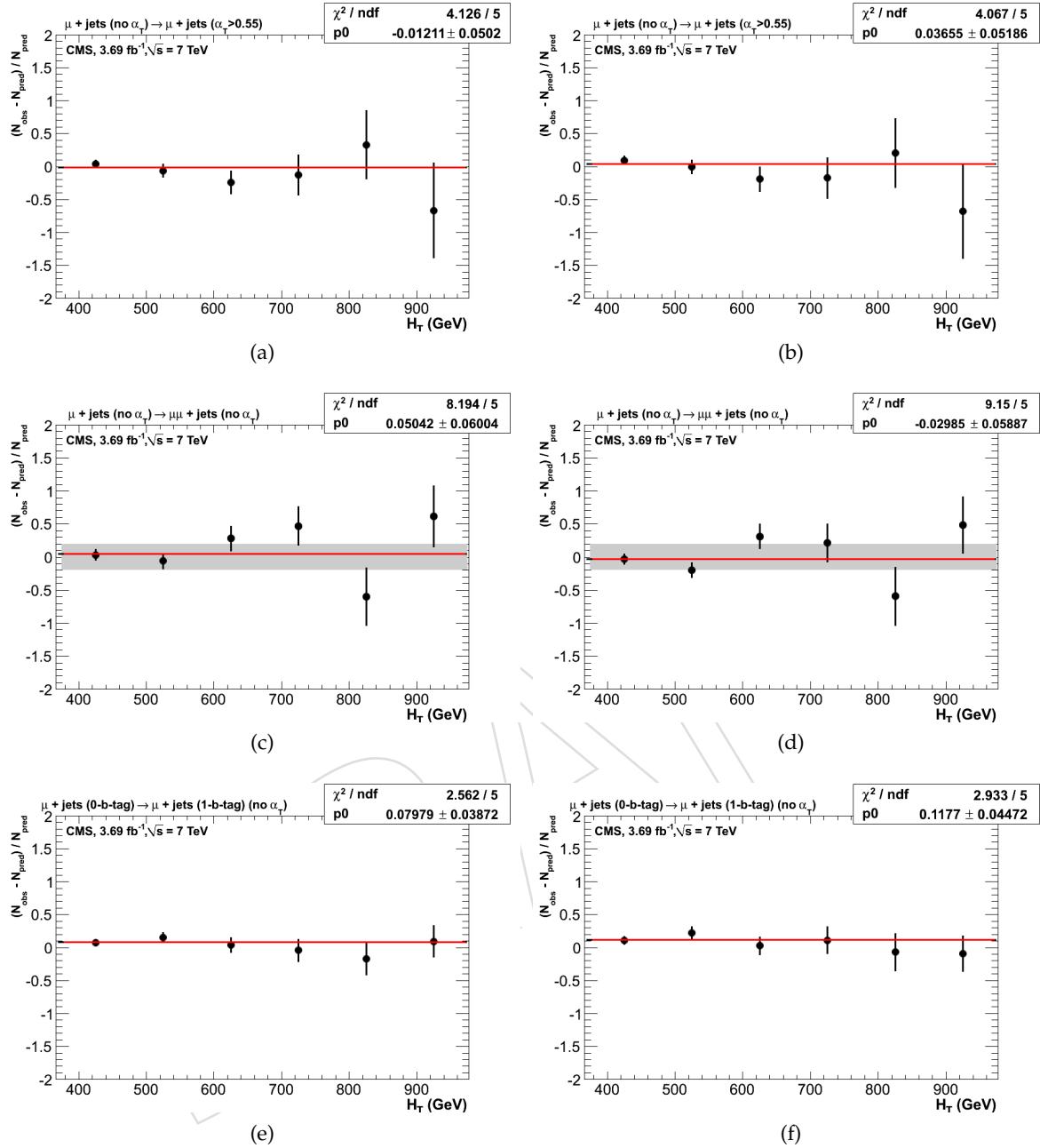


Figure 35: Closure tests using yields from one control to predict yields in another sample. The three plots on the left show closure tests from the inclusive analysis, which uses jets that are not corrected for the effects of pile-up. On the right, the jets in the analysis *are* corrected for pile-up effects by applying the `L1Offset` jet energy correction. The three closure tests are: probing the MC modelling of the α_{T} acceptance with the $\mu + \text{jets}$ sample (a) without and (b) with `L1Offset` jet energy corrections; using the $\mu + \text{jets}$ sample to predict yields in the $\mu\mu + \text{jets}$ sample (a) without and (b) with `L1Offset` jet energy corrections; and using a 0 b-tagged $\mu + \text{jets}$ sample (a) without and (b) with `L1Offset` jet energy corrections. The red lines indicate the constant best fit value across all H_{T} bins.

1269 **F Experimental systematic uncertainties on signal efficiency times
1270 acceptance**

1271 **F.1 Systematics due to jet energy scale uncertainties**

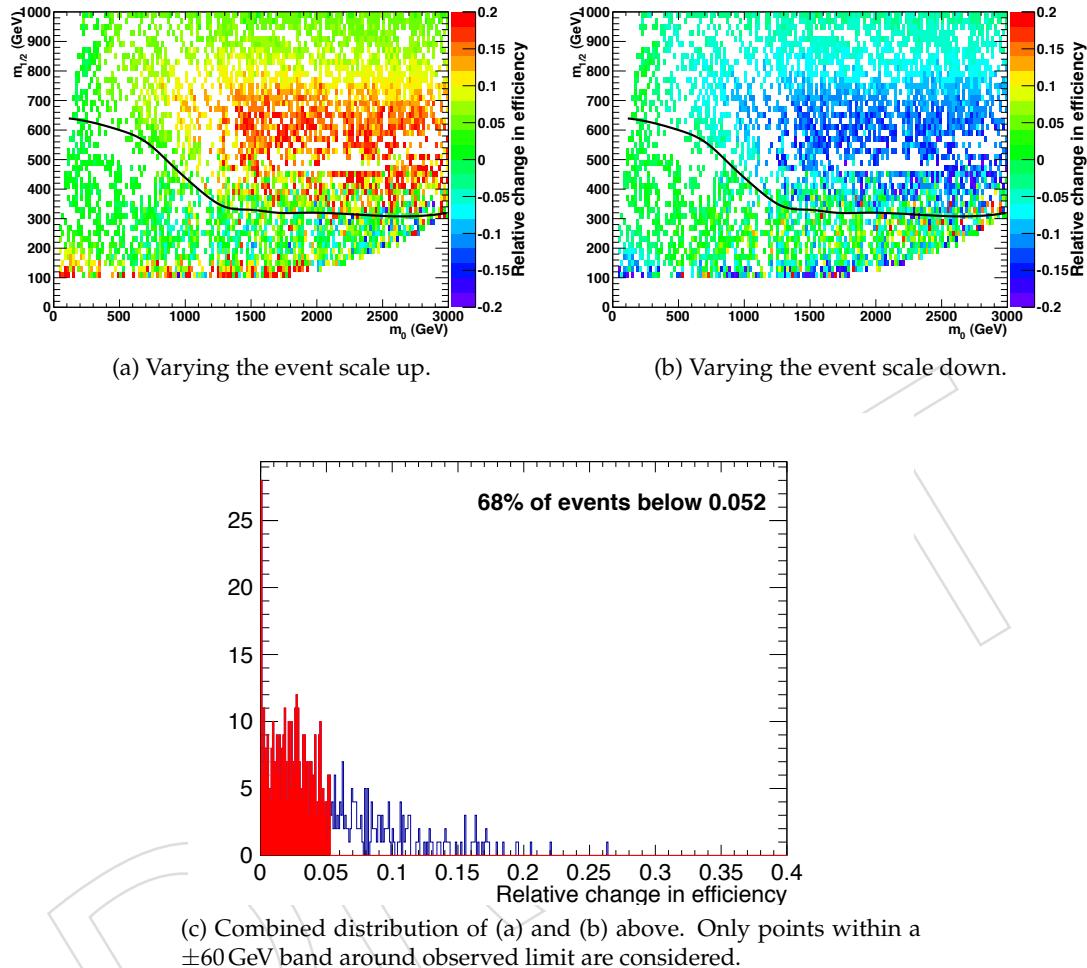


Figure 36: The effect of jet energy scale variations on signal efficiency in the CMSSM plane. All plots show the relative change in efficiency. No requirement is made on the number of reconstructed b jets. The red shaded area is bounded by the 68th percentile.

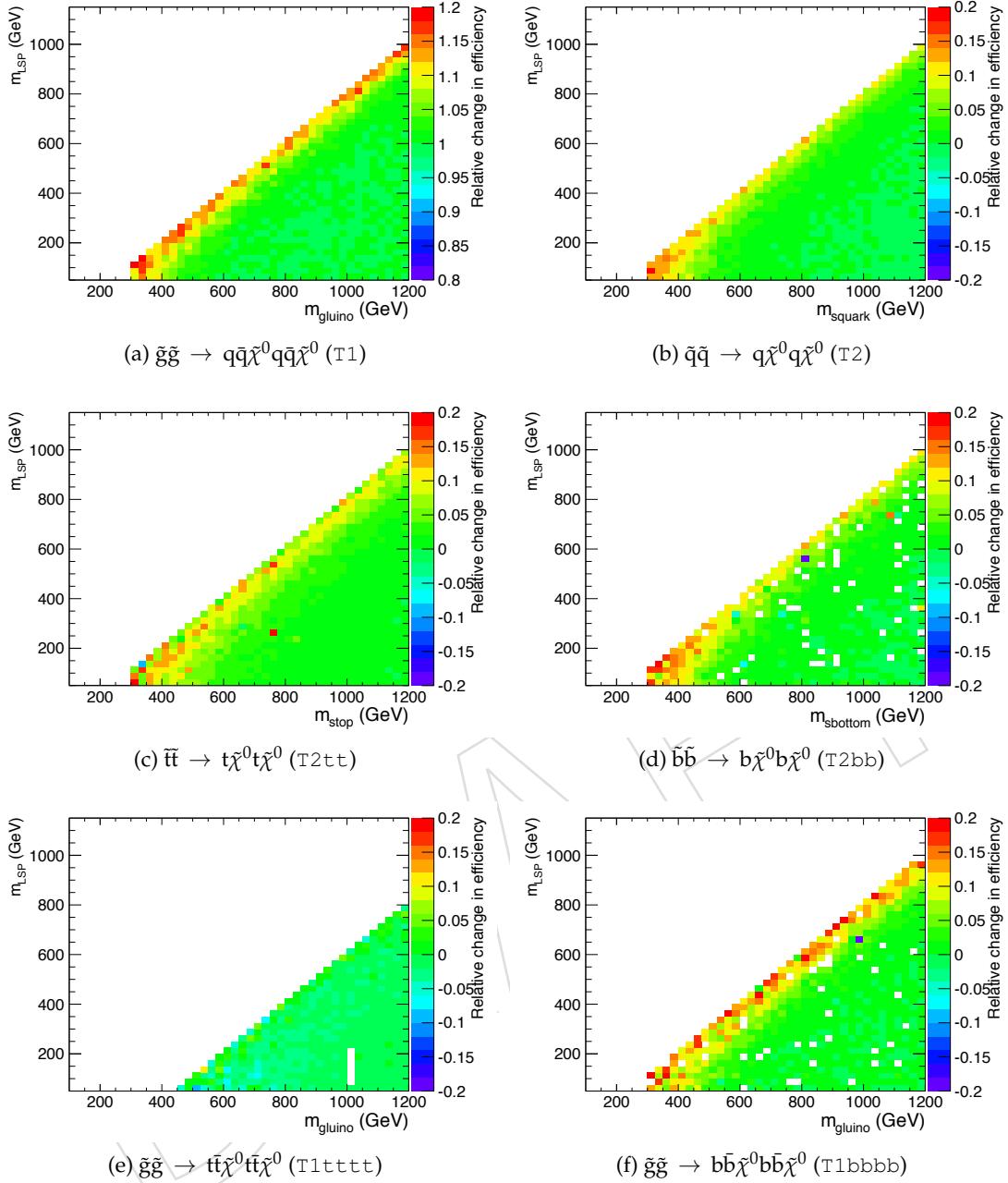


Figure 37: The fractional change in signal efficiency due to systematically increasing all jet energies, for various topologies. No requirement is made on the number of reconstructed b jets.

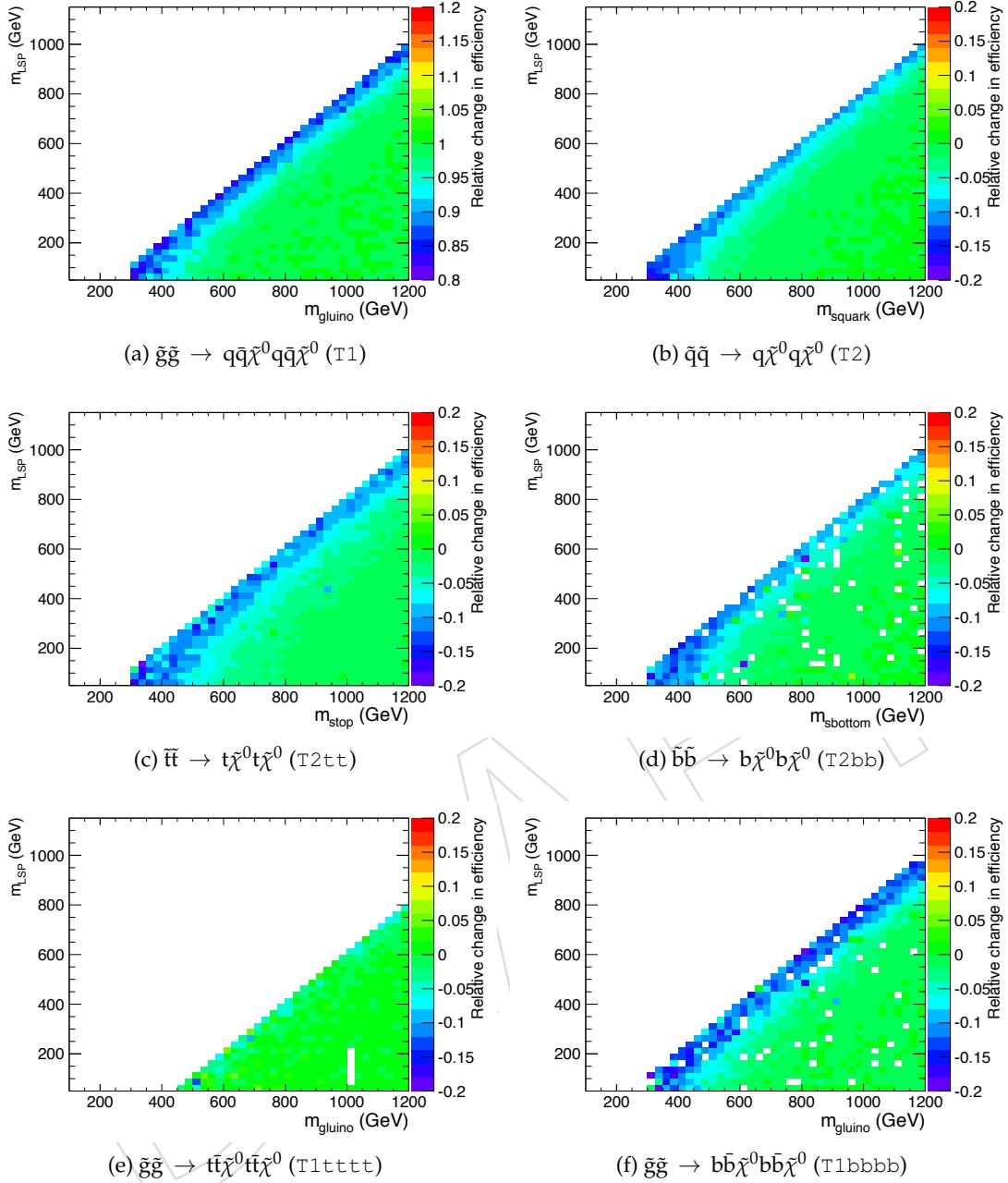


Figure 38: The fractional change in signal efficiency due to systematically decreasing all jet energies, for various topologies. No requirement is made on the number of reconstructed b jets.

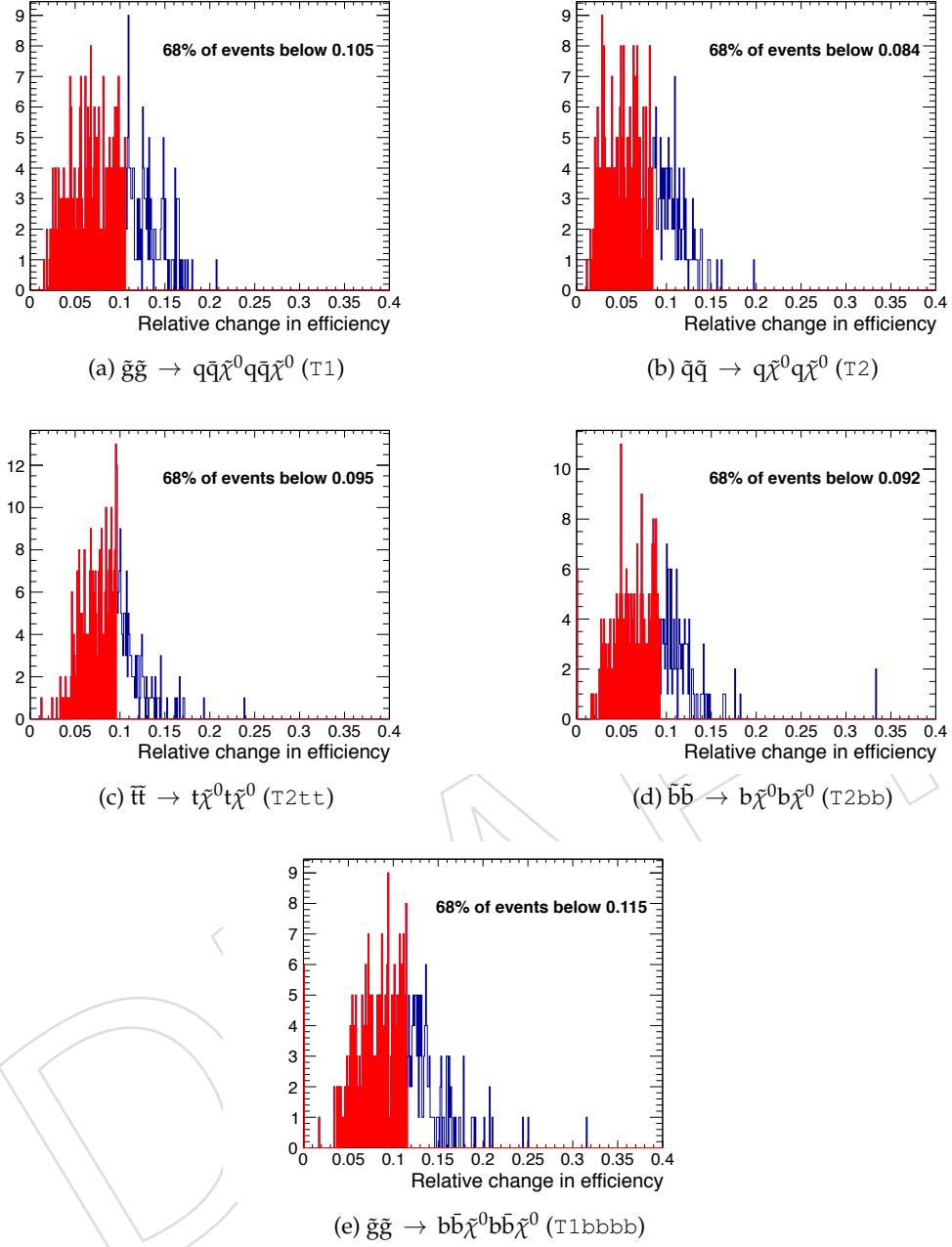


Figure 39: The fractional change in signal efficiency near to the diagonal due to systematically increasing or decreasing all jet energies, for various topologies. No requirement is made on the number of reconstructed b jets.

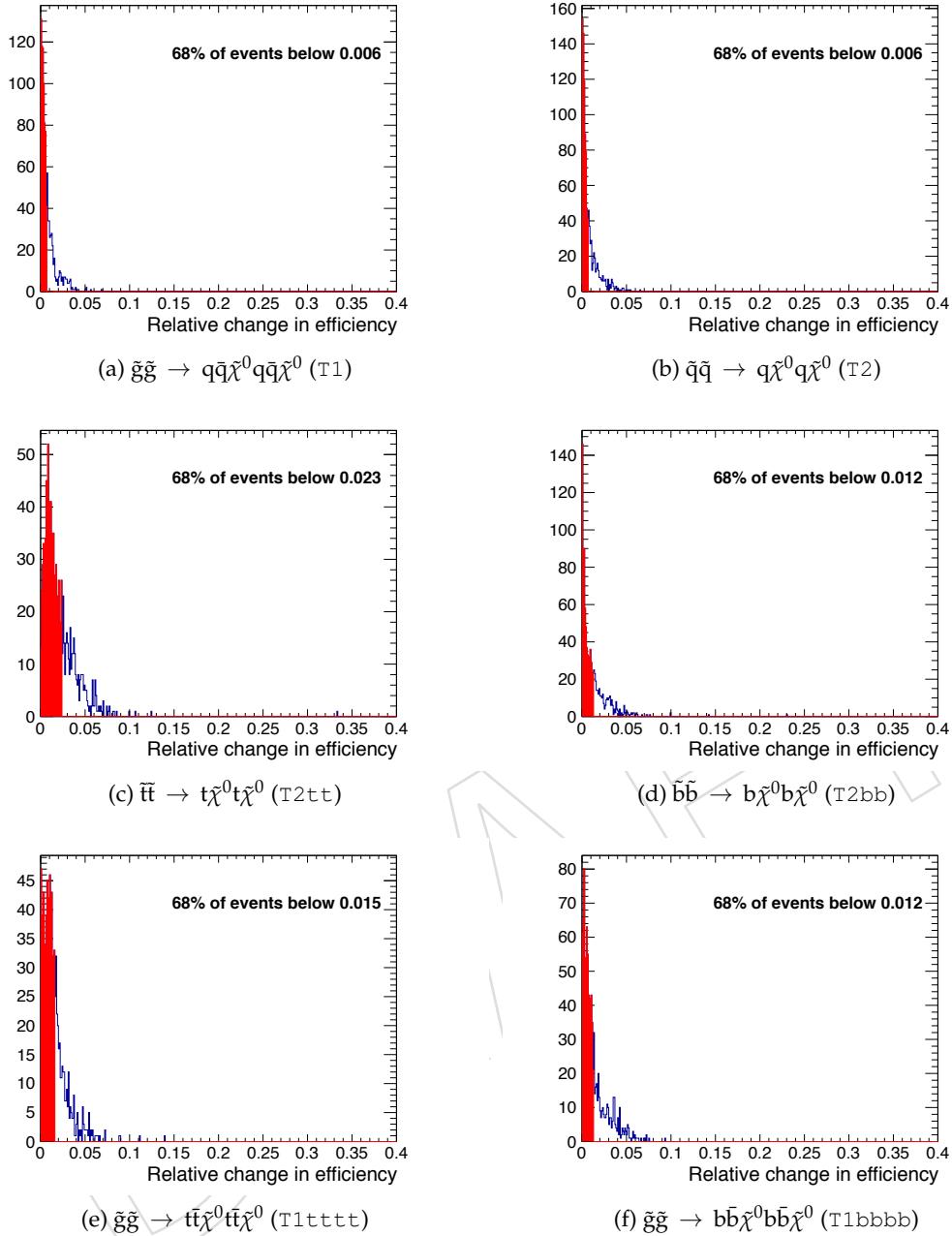


Figure 40: The fractional change in signal efficiency far from the diagonal due to systematically increasing or decreasing all jet energies, for various topologies. No requirement is made on the number of reconstructed b jets.

1272 **F.2 Systematics due to the MHT/MET cut**

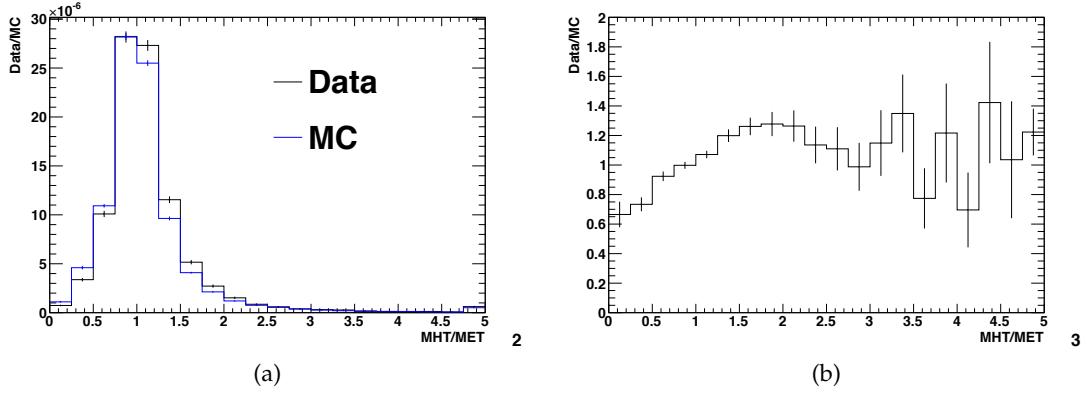


Figure 41: (a) Comparison of the MHT/MET distributions from data and MC, and (b) the ratio of the distributions, data/MC, as a function of the MHT/MET cut value.

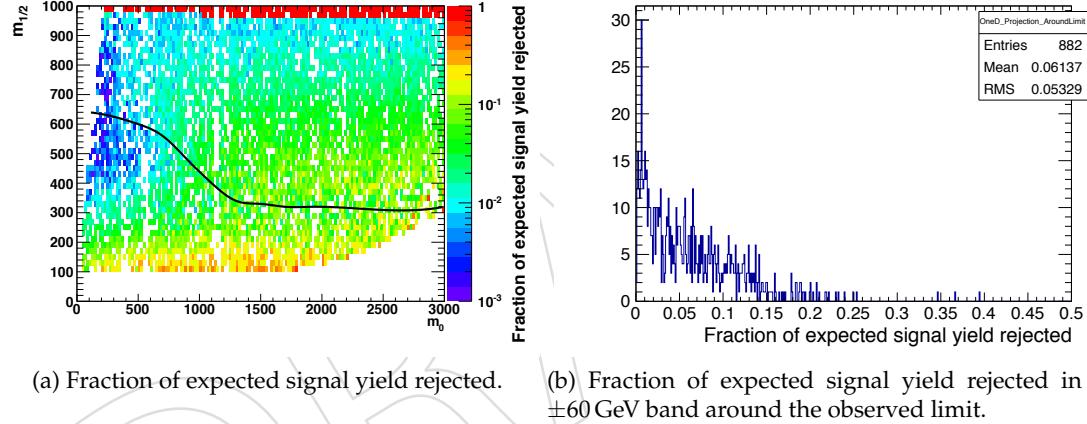


Figure 42: The fraction of expected signal yield that is rejected by the MHT/MET cut in the CMSSM plane. No requirement is made on the number of reconstructed b jets.

Table 38: Conservative estimates of inefficiency (%) for the MHT/MET cut when considering model points in the region near to the diagonal (i.e. small mass splitting and compressed spectra) for various simplified models.

| | T1 | T2 | T2tt | T2bb | T1tttt | T1bbbb |
|------|------|-----|------|------|--------|--------|
| Near | 10.9 | 3.5 | 20.4 | 3.9 | - | 10.8 |
| Far | 3.2 | 0.9 | 3.1 | 1.2 | 32.4 | 3.7 |

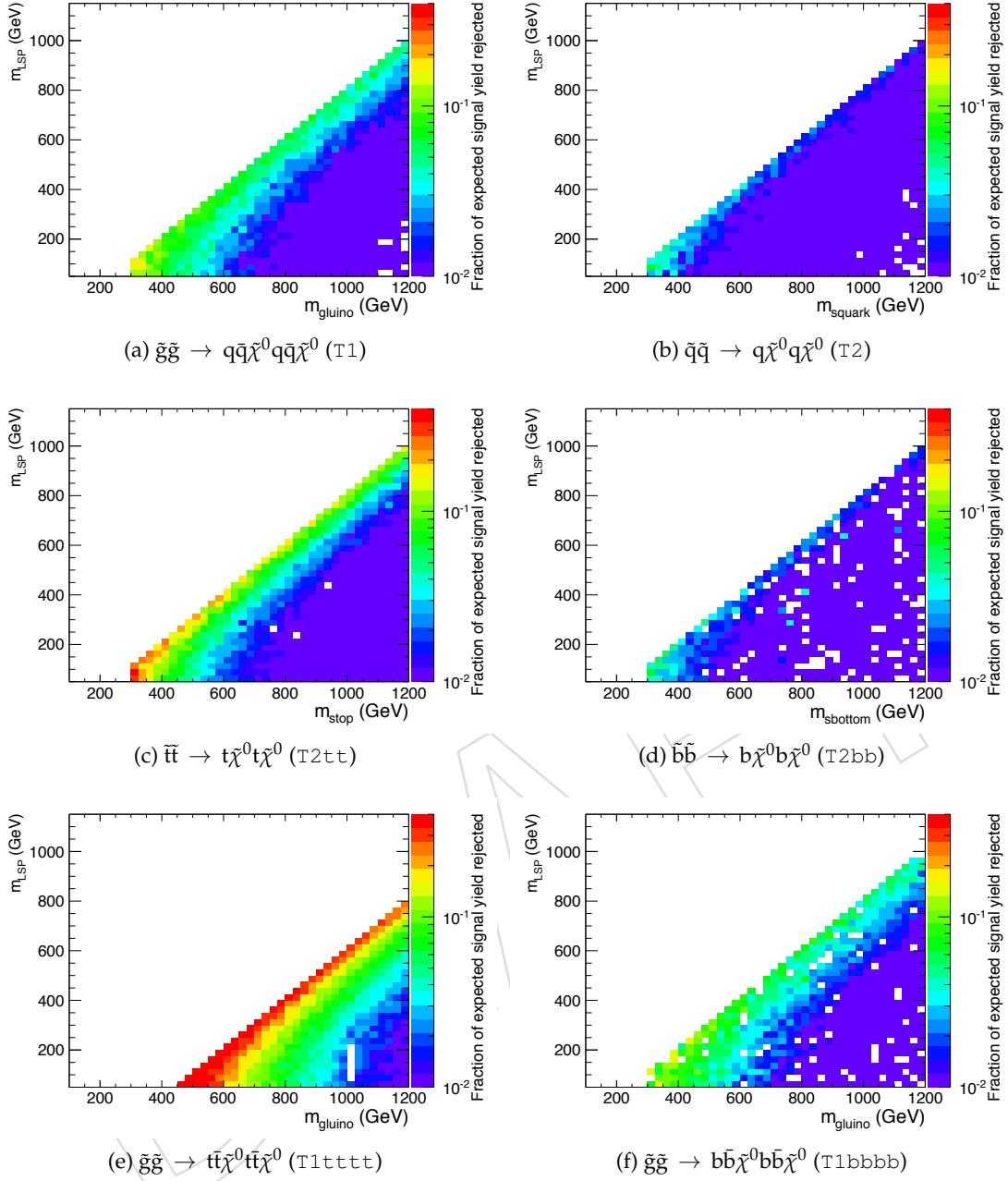


Figure 43: The fraction of expected signal yield that is rejected by the MHT/MET cleaning cut, for various topologies. No requirement is made on the number of reconstructed b jets.

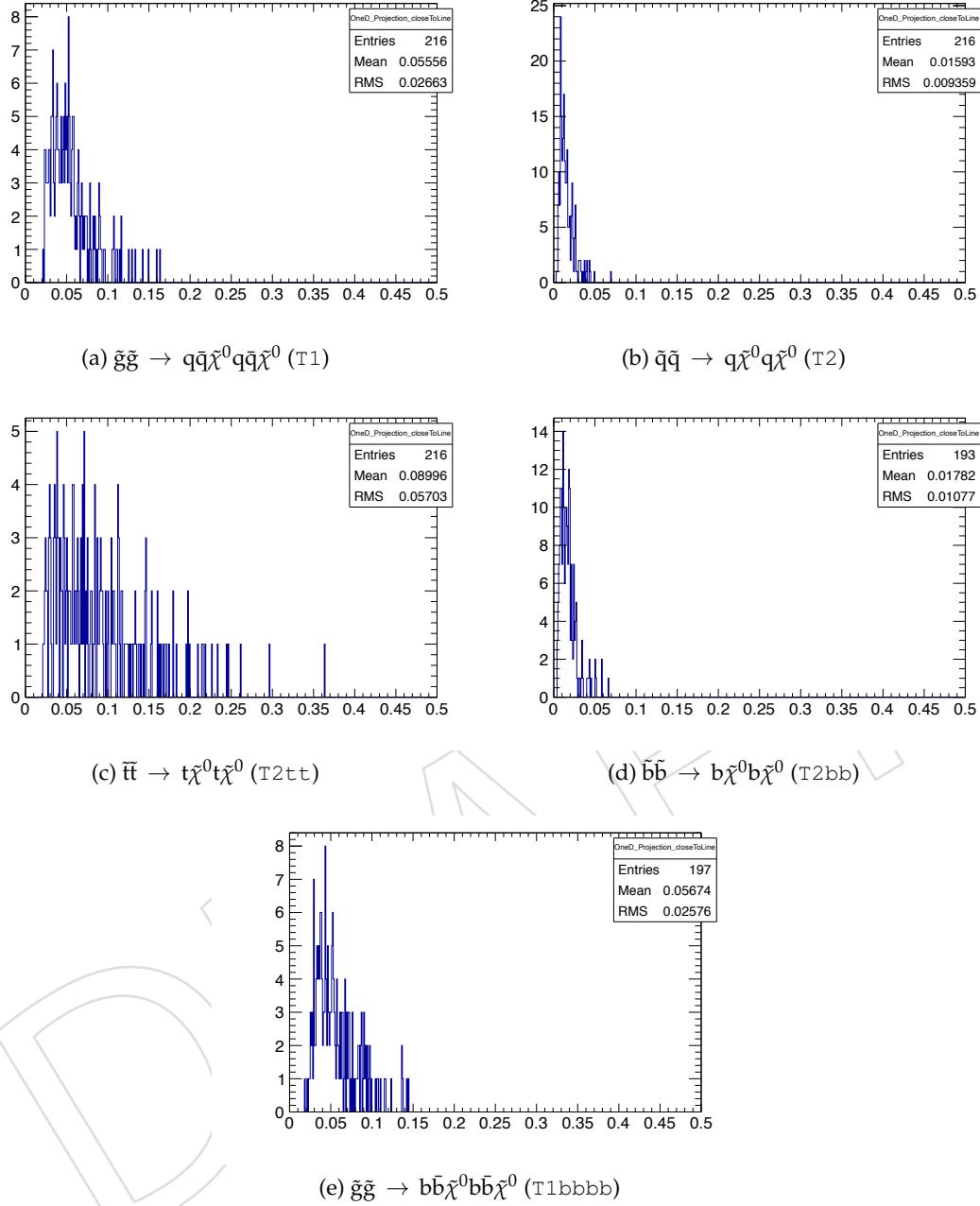


Figure 44: The fraction of expected signal yield that is rejected by the MHT/MET cleaning cut, near to the diagonal, for various topologies. No requirement is made on the number of reconstructed b jets.

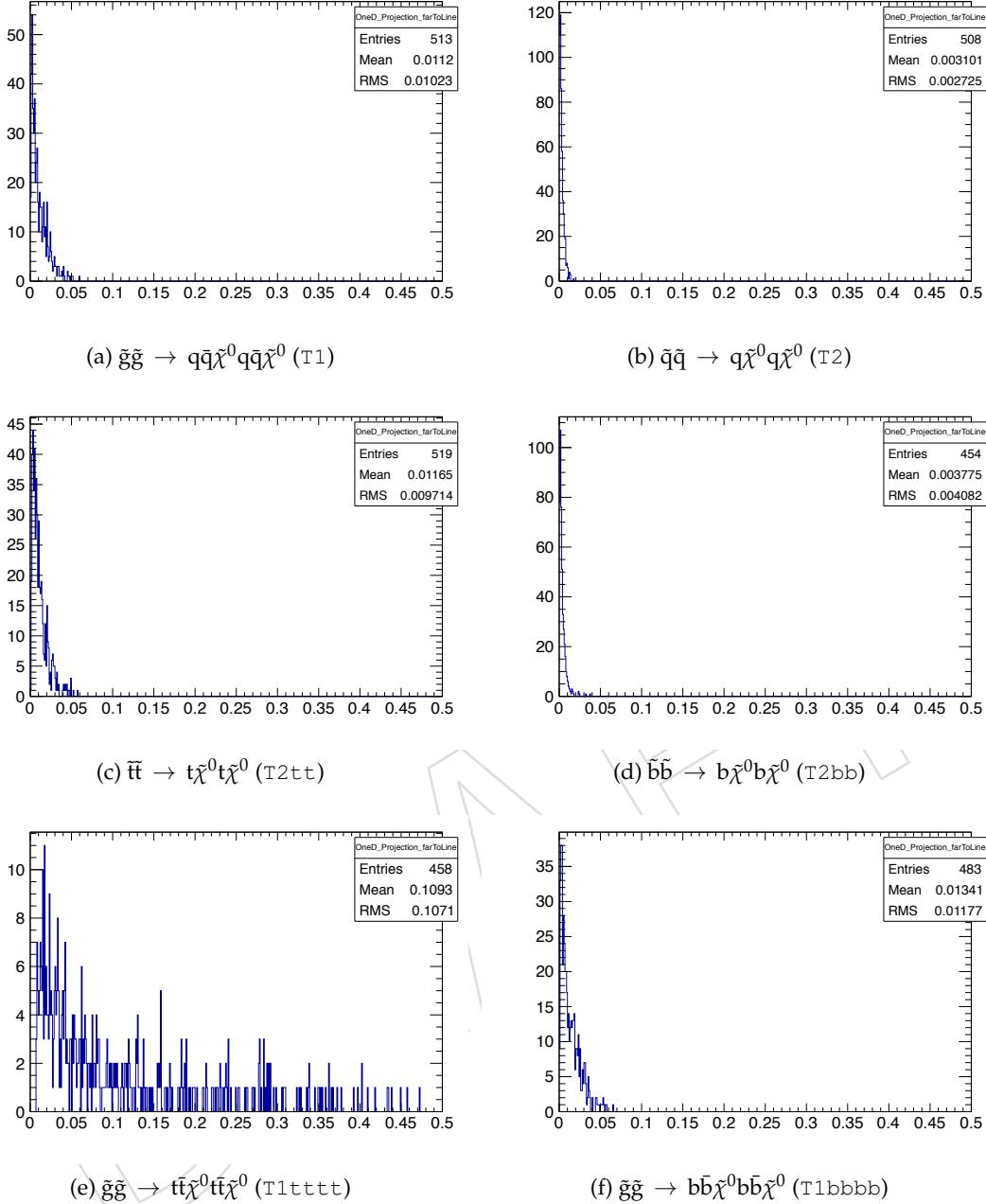
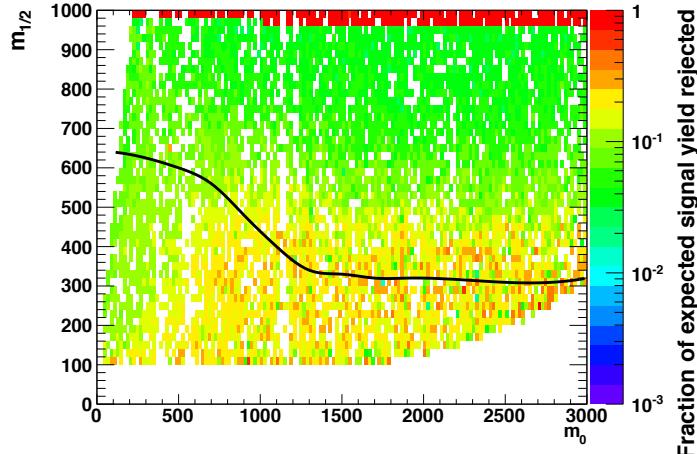
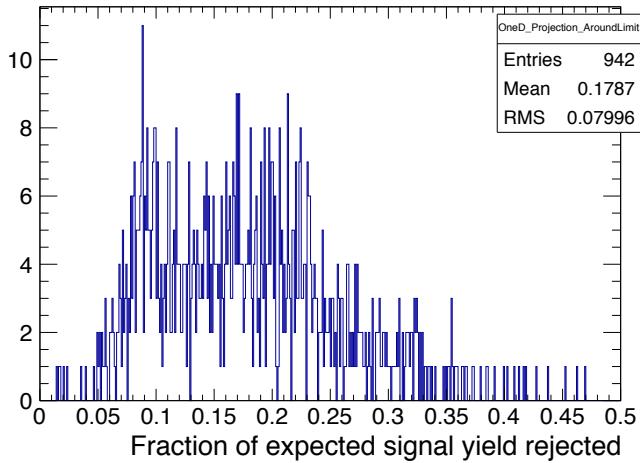


Figure 45: The fraction of expected signal yield that is rejected by the MHT/MET cleaning cut, far from the diagonal, for various topologies. No requirement is made on the number of reconstructed b jets.

¹²⁷³ F.3 Systematics due to the dead ECAL cut



(a) Fraction of expected signal yield rejected.



(b) Fraction of expected signal yield rejected in a ± 60 GeV band around the observed limit.

Figure 46: The fraction of expected signal yield that is rejected by the dead ECAL cut in the CMSSM plane. No requirement is made on the number of reconstructed b jets.

Table 39: Conservative estimates of inefficiency (%) for the dead ECAL cut when considering model points in the region near to the diagonal (i.e. small mass splitting and compressed spectra) for various simplified models.

| | T1 | T2 | T2tt | T2bb | T1tttt | T1bbbb |
|------|------|-----|------|------|--------|--------|
| Near | 13.3 | 6.4 | 23.7 | 7.5 | - | 16.4 |
| Far | 13.6 | 5.8 | 9.4 | 6.2 | 27.6 | 13.9 |

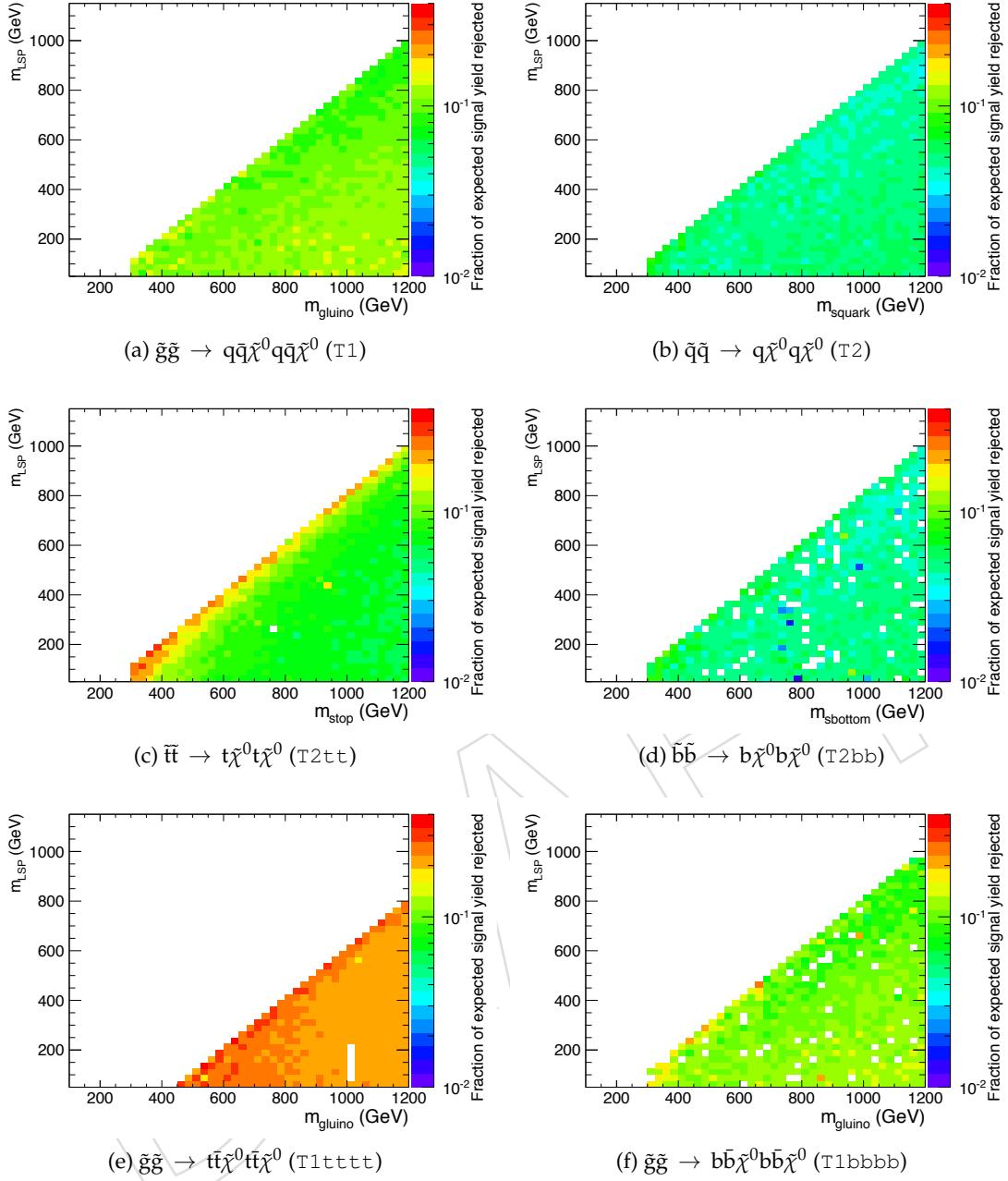


Figure 47: The fraction of expected signal yield that is rejected by the dead ECAL cleaning cut, for various topologies. No requirement is made on the number of reconstructed b jets.

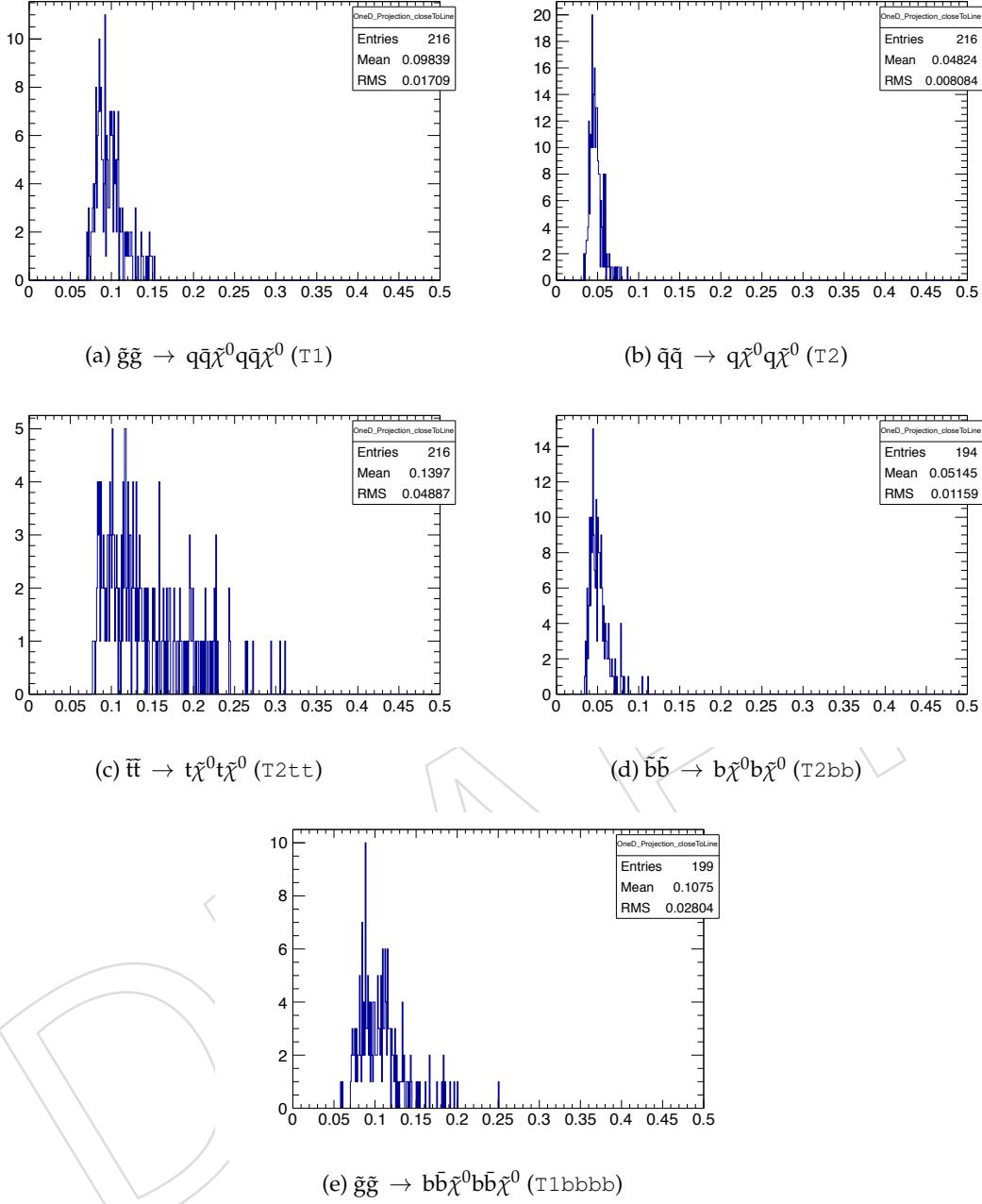


Figure 48: The fraction of expected signal yield that is rejected by the dead ECAL cleaning cut, near to the diagonal, for various topologies. No requirement is made on the number of reconstructed b jets.

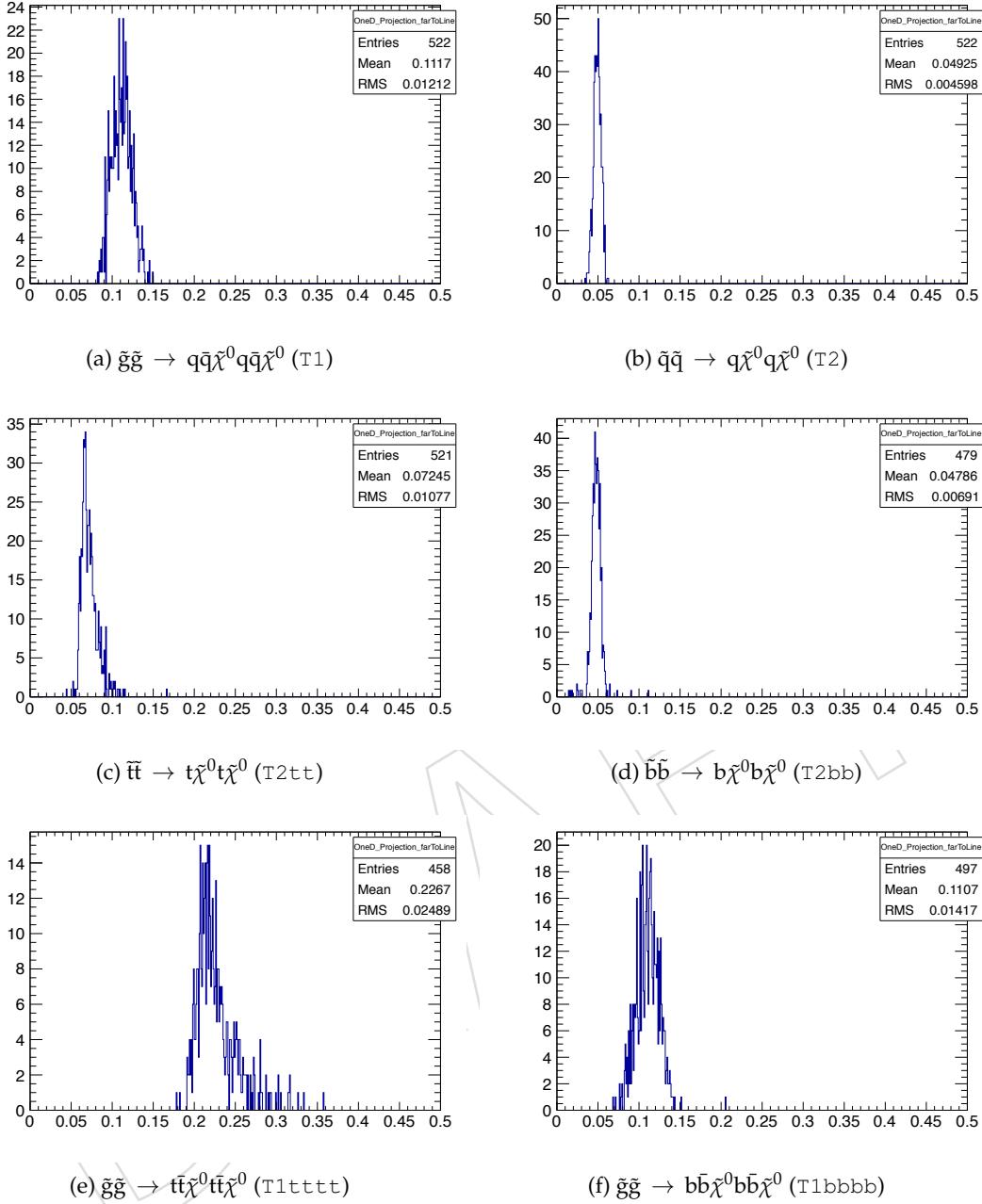
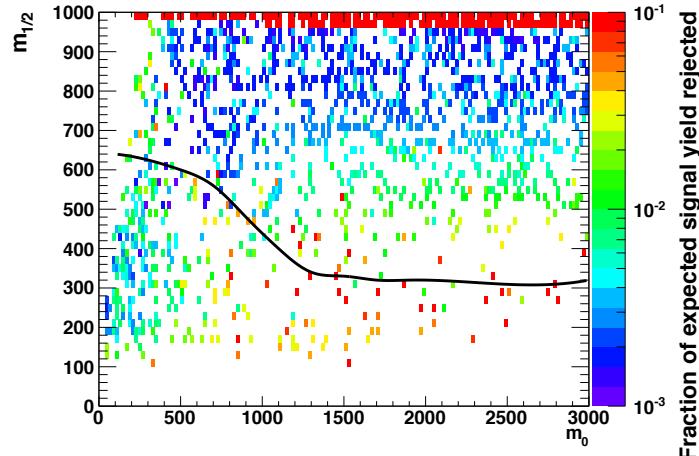
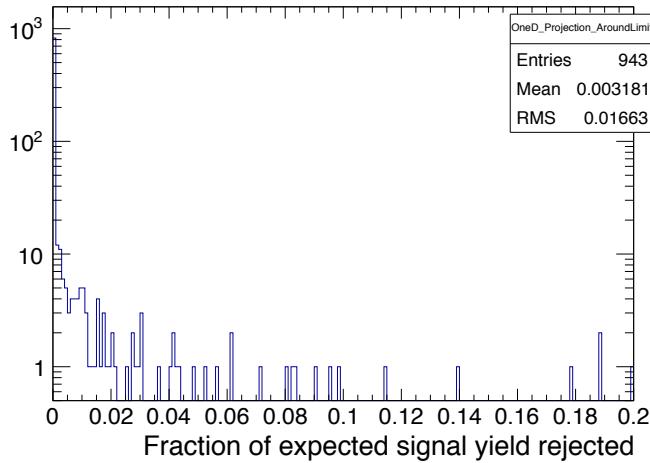


Figure 49: The fraction of expected signal yield that is rejected by the dead ECAL cleaning cut, far from the diagonal, for various topologies. No requirement is made on the number of reconstructed b jets.

¹²⁷⁴ **F.4 Systematics due to the lepton and photon vetoes**



(a) Fraction of expected signal yield rejected.



(b) Fraction of expected signal yield rejected in a ± 60 GeV band around the observed limit.

Figure 50: The fraction of expected signal yield that is rejected by the dead ECAL cut in the CMSSM plane. No requirement is made on the number of reconstructed b jets.

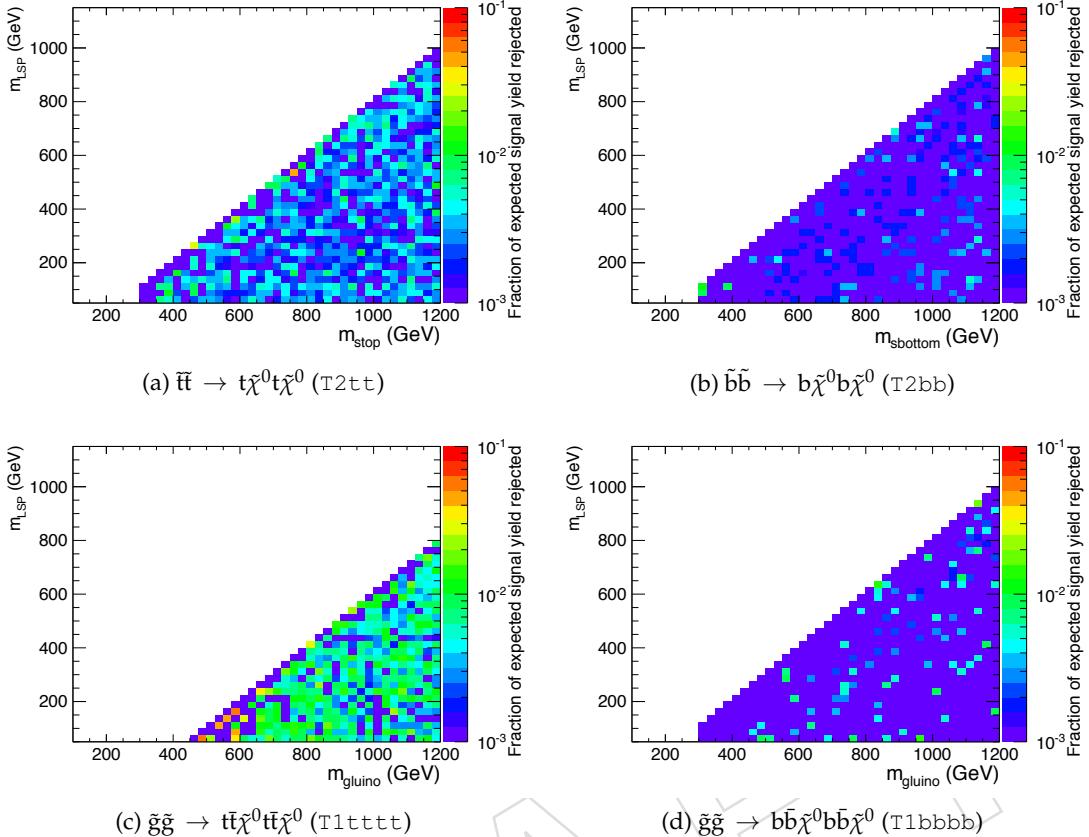


Figure 51: The fraction of expected signal yield that is rejected by the lepton and photon vetoes, for various topologies. No requirement is made on the number of reconstructed b jets.

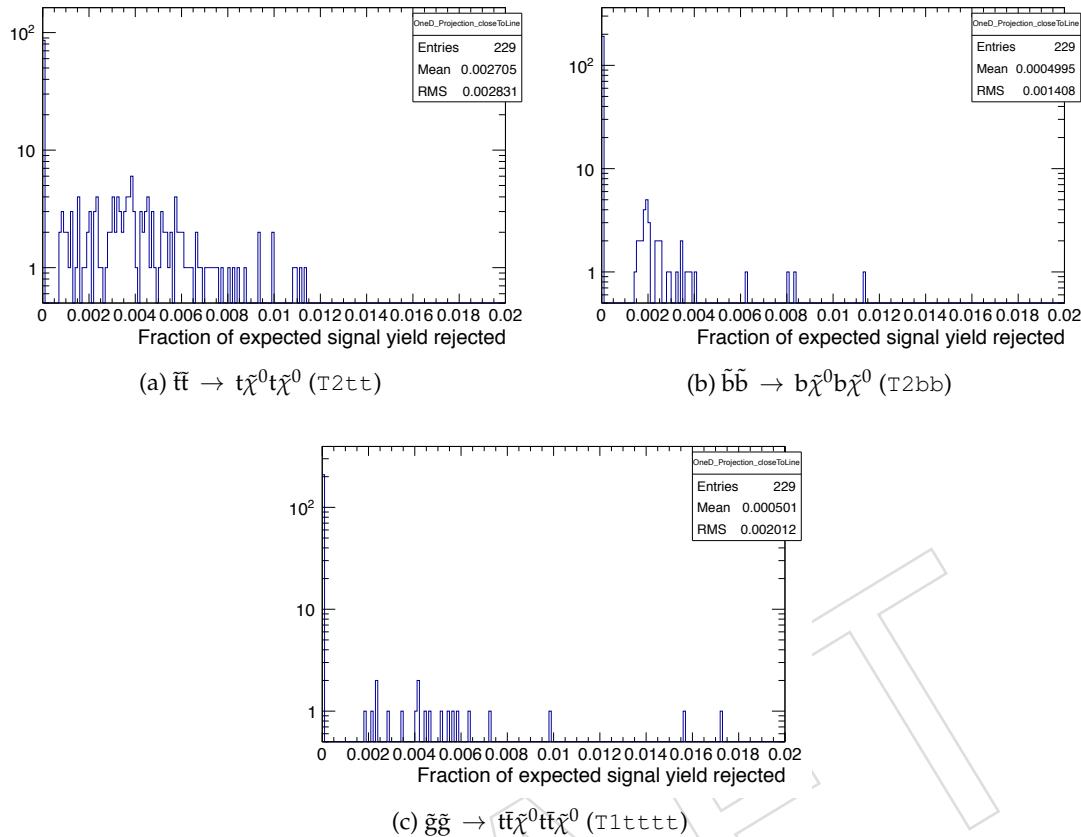


Figure 52: The fraction of expected signal yield that is rejected by the lepton and photon vetoes, near to the diagonal, for various topologies. No requirement is made on the number of reconstructed b jets.

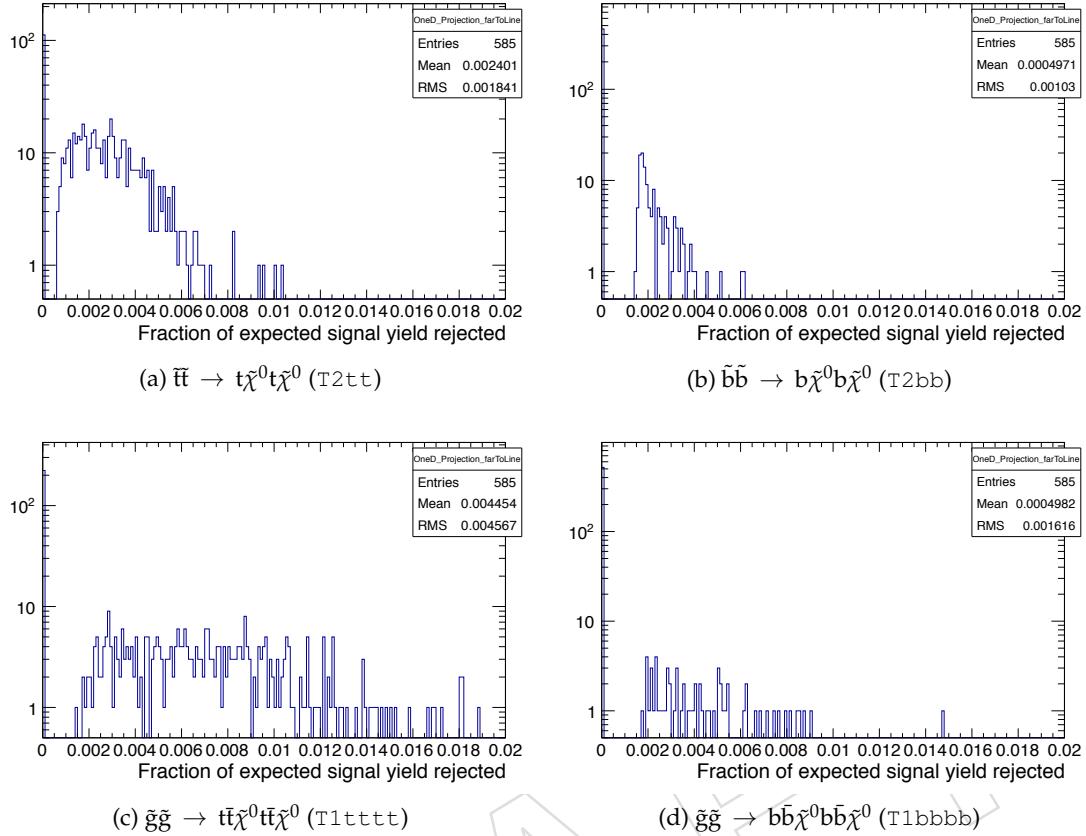


Figure 53: The fraction of expected signal yield that is rejected by the lepton and photon vetoes, far from the diagonal, for various topologies. No requirement is made on the number of reconstructed b jets.

1275 **G Nuisance parameters in the likelihood model**

Table 40: SM-only maximum-likelihood parameter values (common to all b-tag multiplicities).

| name | value | error |
|---------------------|----------|---------|
| k_{QCD} | 1.33e-02 | 9.7e-03 |
| $\rho_{\mu\mu Z}^0$ | 1.02 | 0.08 |
| $\rho_{\mu\mu Z}^1$ | 1.12 | 0.15 |
| $\rho_{\mu\mu Z}^2$ | 1.18 | 0.31 |
| $\rho_{\mu W}^0$ | 1.07 | 0.10 |
| $\rho_{\mu W}^1$ | 1.04 | 0.11 |
| $\rho_{\mu W}^2$ | 1.02 | 0.25 |
| $\rho_{\gamma Z}^0$ | 0.96 | 0.07 |
| $\rho_{\gamma Z}^1$ | 0.92 | 0.10 |
| $\rho_{\gamma Z}^2$ | 1.15 | 0.24 |

DRAFT

Table 41: SM-only maximum-likelihood parameter values.

| $N_{\text{b-tag}} = 0$ | | | $N_{\text{b-tag}} = 1$ | | |
|------------------------|----------|---------|------------------------|----------|---------|
| name | value | error | name | value | error |
| A_{QCD} | 1.07e-04 | 2.9e-04 | A_{QCD} | 2.61e-05 | 6.7e-05 |
| EWK^0 | 2.37e+03 | 1.9e+02 | EWK^0 | 4.94e+02 | 5.3e+01 |
| EWK^1 | 9.99e+02 | 9.6e+01 | EWK^1 | 2.37e+02 | 2.9e+01 |
| EWK^2 | 7.43e+02 | 5.2e+01 | EWK^2 | 1.93e+02 | 1.6e+01 |
| EWK^3 | 2.58e+02 | 1.6e+01 | EWK^3 | 7.72e+01 | 6.3e+00 |
| EWK^4 | 8.12e+01 | 6.9e+00 | EWK^4 | 2.41e+01 | 2.6e+00 |
| EWK^5 | 3.41e+01 | 4.0e+00 | EWK^5 | 1.06e+01 | 1.5e+00 |
| EWK^6 | 1.04e+01 | 2.3e+00 | EWK^6 | 2.88e+00 | 7.3e-01 |
| EWK^7 | 5.29e+00 | 1.4e+00 | EWK^7 | 2.19e+00 | 5.5e-01 |
| f_{Zinv}^0 | 0.49 | 0.03 | f_{Zinv}^0 | 0.27 | 0.04 |
| f_{Zinv}^7 | 0.74 | 0.07 | f_{Zinv}^7 | 0.43 | 0.07 |

| $N_{\text{b-tag}} = 2$ | | | $N_{\text{b-tag}} \geq 3$ | | |
|------------------------|----------|---------|---------------------------|----------|---------|
| name | value | error | name | value | error |
| A_{QCD} | 6.64e-06 | 1.8e-05 | A_{QCD} | 3.18e-07 | 1.1e-06 |
| EWK^0 | 1.28e+02 | 1.6e+01 | EWK^0 | 8.88e+00 | 3.1e+00 |
| EWK^1 | 5.32e+01 | 7.9e+00 | EWK^1 | 6.74e+00 | 2.0e+00 |
| EWK^2 | 5.64e+01 | 5.4e+00 | EWK^2 | 5.72e+00 | 1.2e+00 |
| EWK^3 | 2.78e+01 | 2.9e+00 | EWK^3 | 3.11e+00 | 7.9e-01 |
| EWK^4 | 8.95e+00 | 1.3e+00 | EWK^4 | 1.66e+00 | 4.8e-01 |
| EWK^5 | 7.07e+00 | 1.3e+00 | EWK^5 | 6.94e-01 | 4.0e-01 |
| EWK^6 | 5.93e-01 | 2.3e-01 | EWK^6 | 5.96e-02 | 6.1e-02 |
| EWK^7 | 9.20e-01 | 3.0e-01 | EWK^7 | 1.79e-01 | 9.9e-02 |
| f_{Zinv}^0 | 0.11 | 0.03 | | | |
| f_{Zinv}^7 | 0.20 | 0.07 | | | |

¹²⁷⁶ **H Maximum likelihood yields**

Table 42: Comparison of the measured yields in the different H_T bins of the hadronic and control samples with the SM expectations and combined statistical and systematic uncertainties given by the simultaneous fit. Exactly zero b-tags per event are required.

| H_T Bin (GeV) | 275–325 | 325–375 | 375–475 | 475–575 | 575–675 | 675–775 | 775–875 | 875– ∞ |
|---------------------|----------------------|----------------------|----------------------|----------------------|----------------------|----------------------|----------------------|---------------------|
| SM hadronic | 2933^{+56}_{-52} | 1139^{+17}_{-40} | 783^{+17}_{-27} | 261^{+14}_{-8} | $81.5^{+6.5}_{-6.5}$ | $34.2^{+4.0}_{-3.8}$ | $10.4^{+2.8}_{-1.8}$ | $5.3^{+1.7}_{-1.1}$ |
| Data hadronic | 2919 | 1166 | 769 | 255 | 91 | 31 | 10 | 4 |
| SM μ +jets | 940^{+26}_{-34} | 466^{+25}_{-6} | 262^{+9}_{-12} | $82.2^{+4.1}_{-6.6}$ | $26.3^{+3.0}_{-3.2}$ | $9.2^{+1.5}_{-1.6}$ | $2.5^{+0.8}_{-0.7}$ | $1.1^{+0.5}_{-0.4}$ |
| Data μ +jets | 949 | 444 | 281 | 77 | 23 | 11 | 5 | 0 |
| SM $\mu\mu$ +jets | $93.1^{+6.6}_{-7.4}$ | $54.9^{+4.4}_{-2.1}$ | $36.8^{+1.8}_{-2.1}$ | $13.6^{+0.7}_{-1.1}$ | $7.0^{+0.3}_{-1.3}$ | $2.5^{+0.2}_{-0.5}$ | $0.1^{+0.0}_{-0.0}$ | $0.9^{+0.2}_{-0.4}$ |
| Data $\mu\mu$ +jets | 95 | 53 | 35 | 11 | 4 | 1 | 0 | 1 |
| SM γ +jets | – | – | 913^{+33}_{-22} | 314^{+13}_{-15} | 112^{+9}_{-9} | $47.1^{+5.7}_{-5.4}$ | $15.0^{+2.7}_{-3.5}$ | $9.7^{+2.1}_{-2.6}$ |
| Data γ +jets | – | – | 909 | 328 | 109 | 50 | 13 | 12 |

Table 43: Comparison of the measured yields in the different H_T bins of the hadronic and control samples with the SM expectations and combined statistical and systematic uncertainties given by the simultaneous fit. Exactly one b-tag per event is required.

| H_T Bin (GeV) | 275–325 | 325–375 | 375–475 | 475–575 | 575–675 | 675–775 | 775–875 | 875– ∞ |
|---------------------|----------------------|---------------------|----------------------|----------------------|----------------------|----------------------|----------------------|----------------------|
| SM hadronic | 630^{+26}_{-25} | 271^{+10}_{-16} | 202^{+10}_{-6} | $78.0^{+6.9}_{-1.9}$ | $24.2^{+2.9}_{-2.0}$ | $10.6^{+1.7}_{-1.3}$ | $2.9^{+0.9}_{-0.5}$ | $2.2^{+0.7}_{-0.4}$ |
| Data hadronic | 614 | 294 | 214 | 71 | 20 | 6 | 4 | 0 |
| SM μ +jets | 336^{+17}_{-19} | 168^{+12}_{-7} | 575^{+19}_{-22} | 276^{+12}_{-17} | 115^{+10}_{-9} | $42.6^{+5.5}_{-5.4}$ | $24.0^{+4.6}_{-4.6}$ | $26.8^{+4.7}_{-4.8}$ |
| Data μ +jets | 347 | 146 | 568 | 288 | 116 | 48 | 22 | 26 |
| SM $\mu\mu$ +jets | $13.9^{+1.7}_{-2.2}$ | $7.4^{+0.9}_{-0.8}$ | $36.1^{+2.4}_{-3.0}$ | $17.2^{+1.3}_{-1.6}$ | $7.5^{+0.6}_{-1.5}$ | $5.1^{+0.6}_{-1.2}$ | $1.4^{+0.3}_{-0.5}$ | $2.4^{+0.5}_{-0.9}$ |
| Data $\mu\mu$ +jets | 15 | 9 | 34 | 20 | 10 | 7 | 0 | 6 |
| SM γ +jets | – | – | 128^{+9}_{-8} | $50.7^{+4.1}_{-3.7}$ | $18.5^{+2.8}_{-2.2}$ | $7.8^{+1.6}_{-1.3}$ | $2.8^{+0.8}_{-0.8}$ | $2.6^{+0.8}_{-0.8}$ |
| Data γ +jets | – | – | 126 | 43 | 19 | 5 | 5 | 2 |

Table 44: Comparison of the measured yields in the different H_T bins of the hadronic and control samples with the SM expectations and combined statistical and systematic uncertainties given by the simultaneous fit. Exactly two b-tags per event are required.

| H_T Bin (GeV) | 275–325 | 325–375 | 375–475 | 475–575 | 575–675 | 675–775 | 775–875 | 875– ∞ |
|---------------------|---------------------|----------------------|----------------------|----------------------|----------------------|----------------------|---------------------|----------------------|
| SM hadronic | 162^{+13}_{-12} | $61.8^{+4.8}_{-6.3}$ | $58.8^{+4.8}_{-2.6}$ | $28.0^{+3.5}_{-1.1}$ | $9.0^{+1.4}_{-1.0}$ | $7.1^{+1.4}_{-1.0}$ | $0.6^{+0.3}_{-0.2}$ | $0.9^{+0.4}_{-0.2}$ |
| Data hadronic | 160 | 68 | 52 | 19 | 11 | 7 | 0 | 2 |
| SM μ +jets | 116^{+10}_{-11} | $55.1^{+6.4}_{-4.5}$ | 254^{+13}_{-15} | 143^{+10}_{-12} | $67.6^{+7.4}_{-7.6}$ | $27.6^{+4.6}_{-4.6}$ | $8.9^{+2.8}_{-2.8}$ | $13.9^{+3.5}_{-3.5}$ |
| Data μ +jets | 116 | 49 | 264 | 152 | 63 | 26 | 10 | 14 |
| SM $\mu\mu$ +jets | $2.5^{+0.8}_{-0.9}$ | $2.2^{+0.6}_{-0.6}$ | $9.7^{+1.5}_{-1.8}$ | $4.4^{+0.7}_{-0.8}$ | $1.5^{+0.3}_{-0.5}$ | $1.9^{+0.5}_{-0.7}$ | $0.2^{+0.1}_{-0.1}$ | $0.6^{+0.2}_{-0.3}$ |
| Data $\mu\mu$ +jets | 4 | 3 | 8 | 7 | 5 | 2 | 0 | 0 |
| SM γ +jets | – | – | $12.0^{+2.0}_{-2.1}$ | $6.5^{+1.1}_{-1.1}$ | $2.9^{+0.8}_{-0.7}$ | $2.4^{+0.8}_{-0.7}$ | $0.3^{+0.1}_{-0.1}$ | $0.6^{+0.3}_{-0.3}$ |
| Data γ +jets | – | – | 10 | 4 | 2 | 4 | 0 | 0 |

Table 45: Comparison of the measured yields in the different H_T bins of the hadronic and control samples with the SM expectations and combined statistical and systematic uncertainties given by the simultaneous fit. At least three b-tags per event are required.

| H_T Bin (GeV) | 275–325 | 325–375 | 375–475 | 475–575 | 575–675 | 675–775 | 775–875 | 875– ∞ |
|------------------|----------------------|---------------------|----------------------|----------------------|----------------------|---------------------|---------------------|---------------------|
| SM hadronic | $10.5^{+3.5}_{-2.2}$ | $7.1^{+2.2}_{-1.8}$ | $5.8^{+1.4}_{-0.9}$ | $3.1^{+1.0}_{-0.7}$ | $1.7^{+0.5}_{-0.4}$ | $0.7^{+0.5}_{-0.4}$ | $0.1^{+0.1}_{-0.1}$ | $0.2^{+0.1}_{-0.1}$ |
| Data hadronic | 10 | 8 | 8 | 1 | 0 | 0 | 0 | 0 |
| SM μ +jets | $8.5^{+1.8}_{-2.9}$ | $6.8^{+1.7}_{-2.0}$ | $24.1^{+4.1}_{-4.7}$ | $13.9^{+3.1}_{-3.4}$ | $11.3^{+3.2}_{-3.4}$ | $2.3^{+1.5}_{-1.5}$ | $0.9^{+0.9}_{-0.9}$ | $3.8^{+1.9}_{-1.9}$ |
| Data μ +jets | 9 | 6 | 22 | 16 | 13 | 3 | 1 | 4 |

1277 I Validating the asymptotic CL_s method

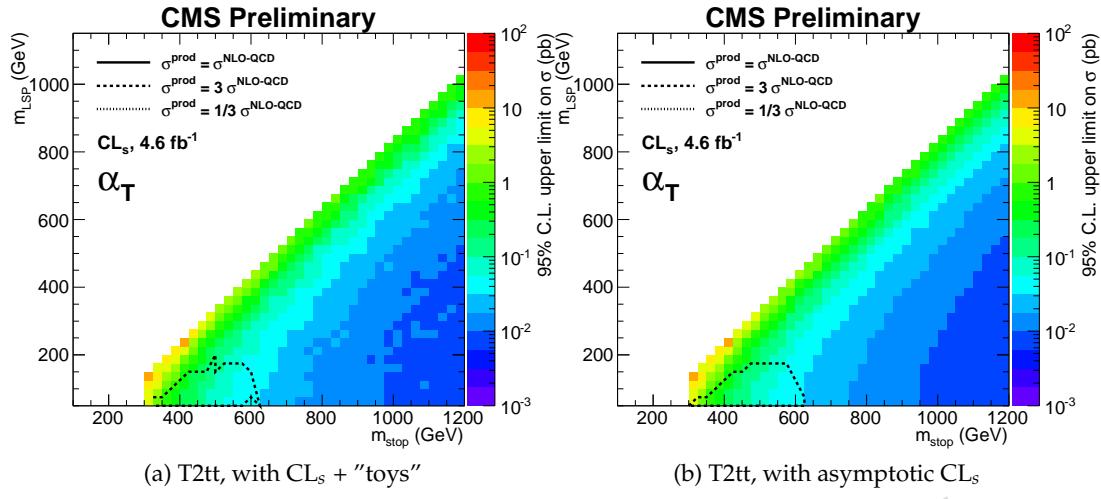


Figure 54: Validating the asymptotic CL_s approach.

DRAFT

1278 J Two-dimensional hadronic signal region

1279 This appendix describes a change to the analysis reported in Ref. [4] that was pre-approved
 1280 by the SUSY group on 1st March 2012, but is not currently used by the analysis presented in
 1281 this note. It concerns enlarging the hadronic signal region to cover more signal phase space, by
 1282 extending the region to lower values of the α_T variable.

1283 In order to suppress QCD multi-jet events to a negligible level, a rather tight cut of $\alpha_T > 0.55$
 1284 is required at low values of H_T . However, at higher values of H_T , such a tight α_T cut is not nec-
 1285 essary. This is due in part to the increase in the average jet energy, hence improving jet energy
 1286 resolutions. Furthermore, the relative impact of any mis-measurement or a jet falling below
 1287 threshold is reduced as the scale of the event (i.e. H_T) increases. Thus, lower α_T thresholds
 1288 can be used in conjunction with larger values of H_T to achieve the same background rejection
 1289 efficiency while increasing the signal acceptance, as described below.

The α_T variable is a relative quantity that is closely related to the ratio \cancel{H}_T/H_T . Assuming $\Delta H_T \approx 0$ (valid for well-measured QCD multi-jet events), a given α_T value cut corresponds to an effective cut on the ratio \cancel{H}_T/H_T , as shown in Equ. 40:

$$\frac{\cancel{H}_T}{H_T} = \sqrt{1 - \frac{1}{4\alpha_T^2}} \quad (40)$$

1290 Thus, by adding the bins at lower α_T values, the implicit threshold on \cancel{H}_T/H_T is also reduced,
 1291 which opens additional phase space and increases sensitivity to signal topologies with either
 1292 long decay cascades or compressed spectra that result in small values of \cancel{H}_T/H_T . By lowering
 1293 the α_T threshold from 0.55 to 0.52, one can reduce the \cancel{H}_T/H_T requirement by as much as $\sim 35\%$,
 1294 as shown in Table 46.

1295 Therefore, it is desirable to extend the one-dimensional H_T shape analysis to two dimensions, in
 1296 which α_T represents the second dimension. Preliminary studies show that indeed a 2D analysis
 1297 will increase the sensitivity for low- \cancel{H}_T signatures such as models present at high m_0 in the
 1298 CMSSM model-space or for compressed spectra in general.⁶ The gains in sensitivity are shown
 1299 in Fig. 55.

1300 Indeed, this 2D approach has already been foreseen in the α_T trigger strategy, described be-
 1301 low, which provides adequate coverage in the two dimensions to allow a binning in H_T and
 1302 α_T as summarized by Table 47. This choice of H_T and α_T binning is driven mainly by trig-
 1303 ger constraints (efficiencies, rates) and the requirement of negligible contamination from QCD
 1304 multi-jet events (motivated through MC studies).

Table 46: Effective \cancel{H}_T thresholds for different α_T and H_T thresholds.

| α_T | \cancel{H}_T/H_T | H_T (GeV) | \cancel{H}_T (GeV) |
|------------|--------------------|-------------|----------------------|
| 0.55 | 0.42 | 275 | 115 |
| 0.55 | 0.42 | 575 | 240 |
| 0.55 | 0.42 | 775 | 323 |
| 0.53 | 0.33 | 575 | 191 |
| 0.52 | 0.27 | 775 | 213 |

1305 In the analysis described in Ref. [4], the hadronic signal region is defined by $H_T > 275$ GeV
 1306 and $\alpha_T > 0.55$, which is divided into eight bins in H_T : two bins of width 50 GeV in the range

⁶These studies are currently being redone using CMS MC simulations for the CMSSM and SMS

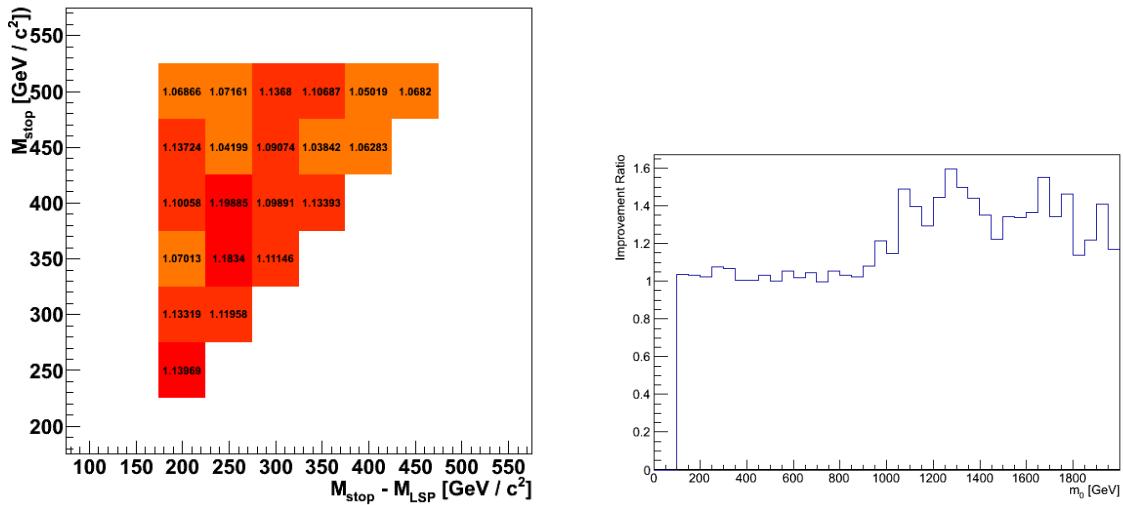


Figure 55: Gains in the “natural SUSY” (left) and CMSSM (right) model spaces using the 2D approach. The factors represent the relative increase in integrated luminosity required to exclude at the same confidence level a given point in the two different SUSY model spaces if the original 1D approach is used, relative to the 2D approach. Therefore, a value above unity indicates an improvement in sensitivity when using the 2D approach.

1307 $275 < H_T < 375 \text{ GeV}$, five bins of width 100 GeV in the range $375 < H_T < 875 \text{ GeV}$, and a
 1308 final open bin, $H_T > 875 \text{ GeV}$, as shown in Table 47. In the analysis presented in this note,
 1309 the hadronic signal region is extended by adding a second dimension, represented by the α_T
 1310 variable. Four additional bins, with the same boundaries in H_T as described above, are added
 1311 in the region $H_T > 575 \text{ GeV}$ and $0.53 < \alpha_T < 0.55$; and a further two bins are added, again with
 1312 the same H_T boundaries, in the region $H_T > 775 \text{ GeV}$ and $0.52 < \alpha_T < 0.53$. Again, this binning
 1313 is summarised in Table 47.

Table 47: Choice of H_T and α_T bins used in the analysis. Bins with jet p_T thresholds scaled down are marked with a dagger symbol (\dagger).

| α_T | | H_T (GeV) | | | | | | | |
|----------------|----------------------|----------------------|---------|---------|---------|---------|---------|---------------|--|
| 0.55– ∞ | 275–325 [†] | 325–375 [†] | 375–475 | 475–575 | 575–675 | 675–775 | 775–875 | 875– ∞ | |
| 0.53–0.55 | - | - | - | - | 575–675 | 675–775 | 775–875 | 875– ∞ | |
| 0.52–0.55 | - | - | - | - | - | - | 775–875 | 875– ∞ | |

1314 K HT distributions for example signal models

1315 K.1 T1 topology: $m_{\tilde{g}} = 800 \text{ GeV}$, $m_{\text{LSP}} = 200 \text{ GeV}$

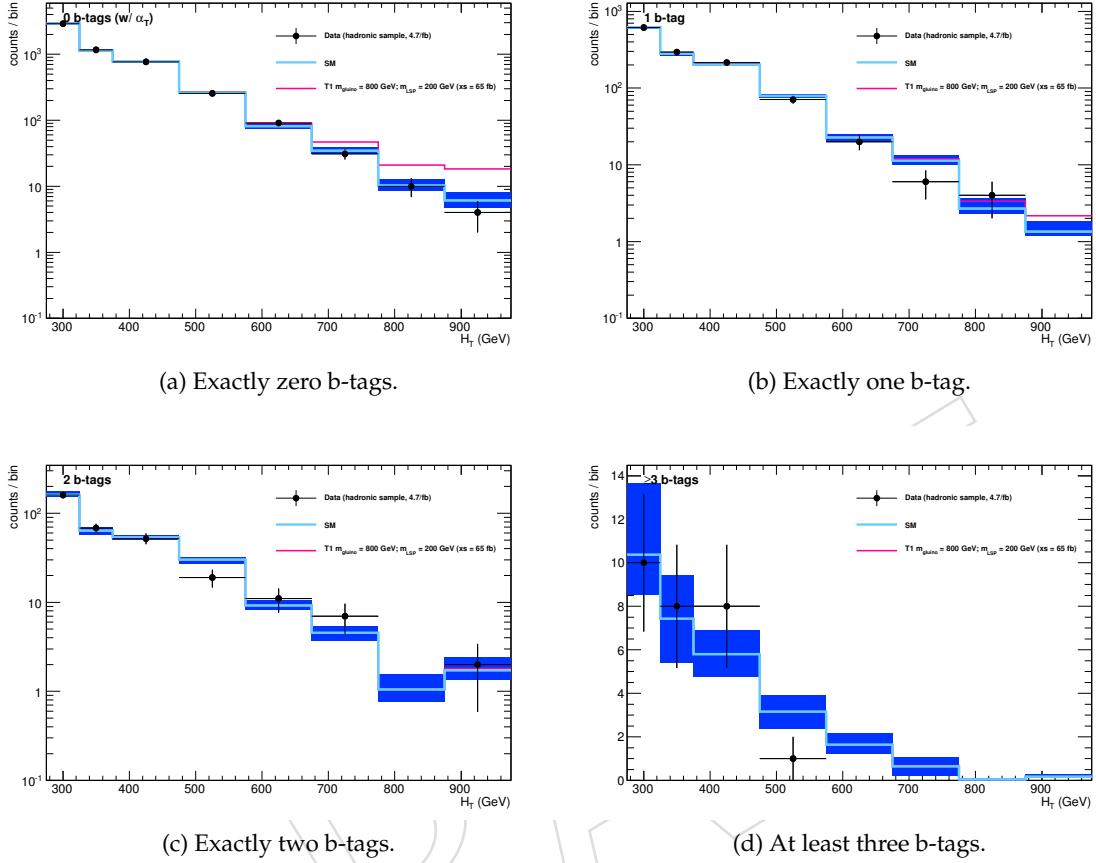


Figure 56: Comparison of the observed yields and SM expectations given by the simultaneous fit in bins of H_T for events in the hadronic signal sample with (a) exactly zero, (b) exactly one, (c) exactly two and (d) at least three b-tags. Shown are the observed event yields in data (black dots with error bars representing the statistical uncertainties) and the expectations and uncertainties, as determined by the simultaneous fit, for all SM processes (light blue solid line). For illustrative purposes, an example model, defined by the topology T1 and the masses $m_{\tilde{g}} = 800 \text{ GeV}$ and $m_{\text{LSP}} = 200 \text{ GeV}$ (i.e. near the observed limit line), is superimposed on the SM expectation (magenta solid line).

1316 **K.2 T2tt topology: $m_{\tilde{q}} = 400 \text{ GeV}$, $m_{\text{LSP}} = 50 \text{ GeV}$**

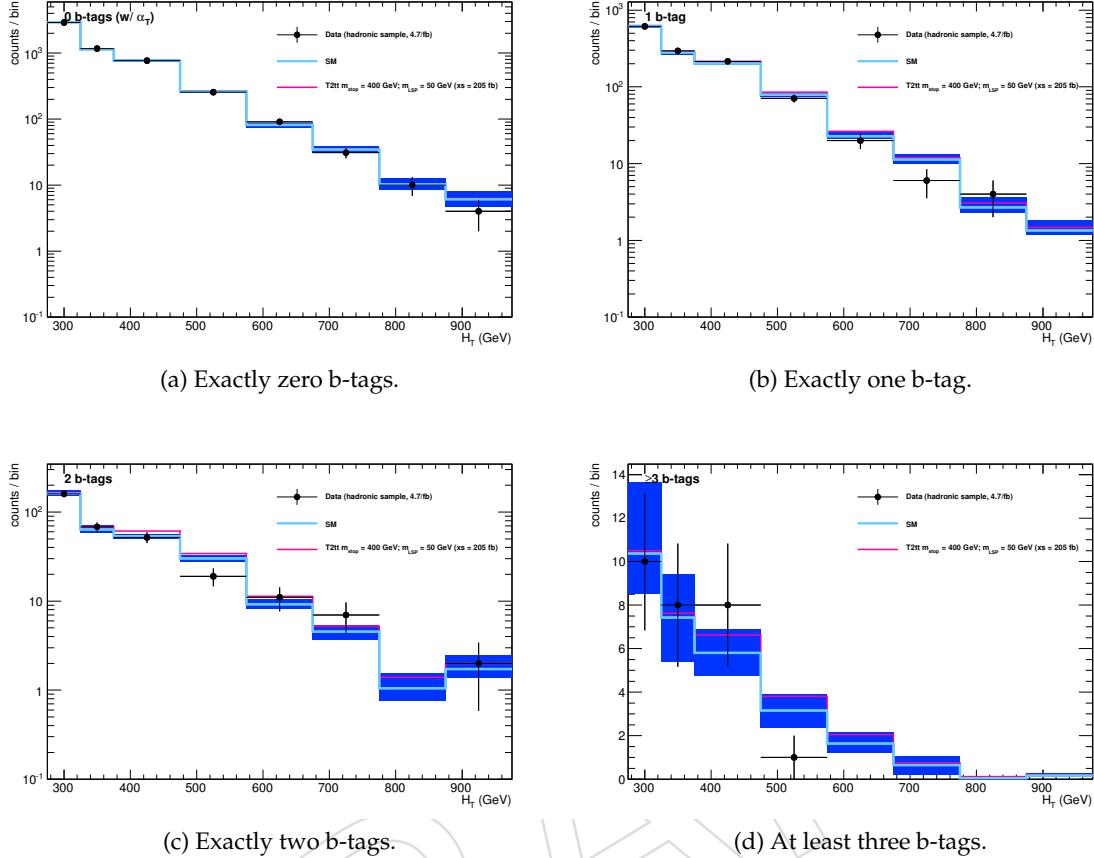


Figure 57: Comparison of the observed yields and SM expectations given by the simultaneous fit in bins of H_T for events in the hadronic signal sample with (a) exactly zero, (b) exactly one, (c) exactly two and (d) at least three b-tags. Shown are the observed event yields in data (black dots with error bars representing the statistical uncertainties) and the expectations and uncertainties, as determined by the simultaneous fit, for all SM processes (light blue solid line). For illustrative purposes, an example model, defined by the topology T2tt and the masses $m_{\tilde{q}} = 400 \text{ GeV}$ and $m_{\text{LSP}} = 50 \text{ GeV}$ (i.e. near the observed limit line), is superimposed on the SM expectation (magenta solid line).

1317 **K.3 T2bb topology: $m_{\tilde{q}} = 500$ GeV, $m_{\text{LSP}} = 150$ GeV**

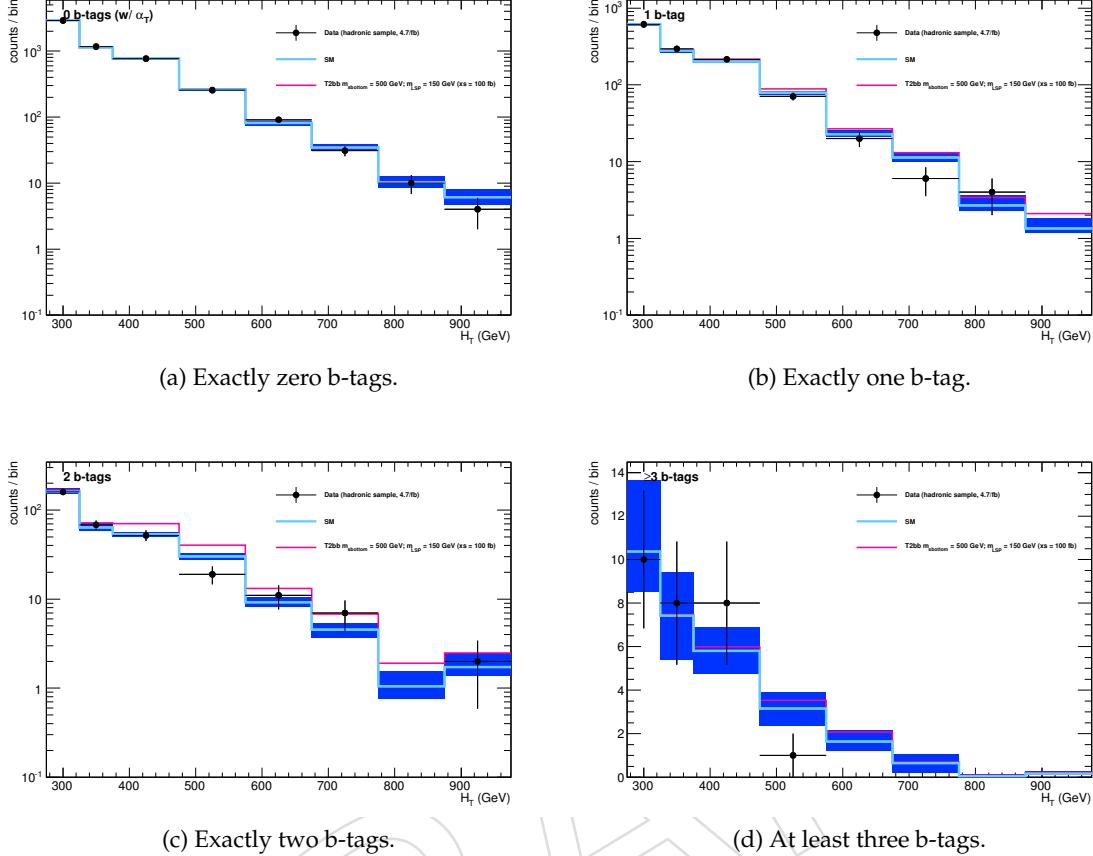


Figure 58: Comparison of the observed yields and SM expectations given by the simultaneous fit in bins of H_T for events in the hadronic signal sample with (a) exactly zero, (b) exactly one, (c) exactly two and (d) at least three b-tags. Shown are the observed event yields in data (black dots with error bars representing the statistical uncertainties) and the expectations and uncertainties, as determined by the simultaneous fit, for all SM processes (light blue solid line). For illustrative purposes, an example model, defined by the topology T2bb and the masses $m_{\tilde{q}} = 500$ GeV and $m_{\text{LSP}} = 150$ GeV (i.e. near the observed limit line), is superimposed on the SM expectation (magenta solid line).

1318 **K.4 T1bbbb topology: $m_{\tilde{g}} = 1000 \text{ GeV}$, $m_{\text{LSP}} = 500 \text{ GeV}$**

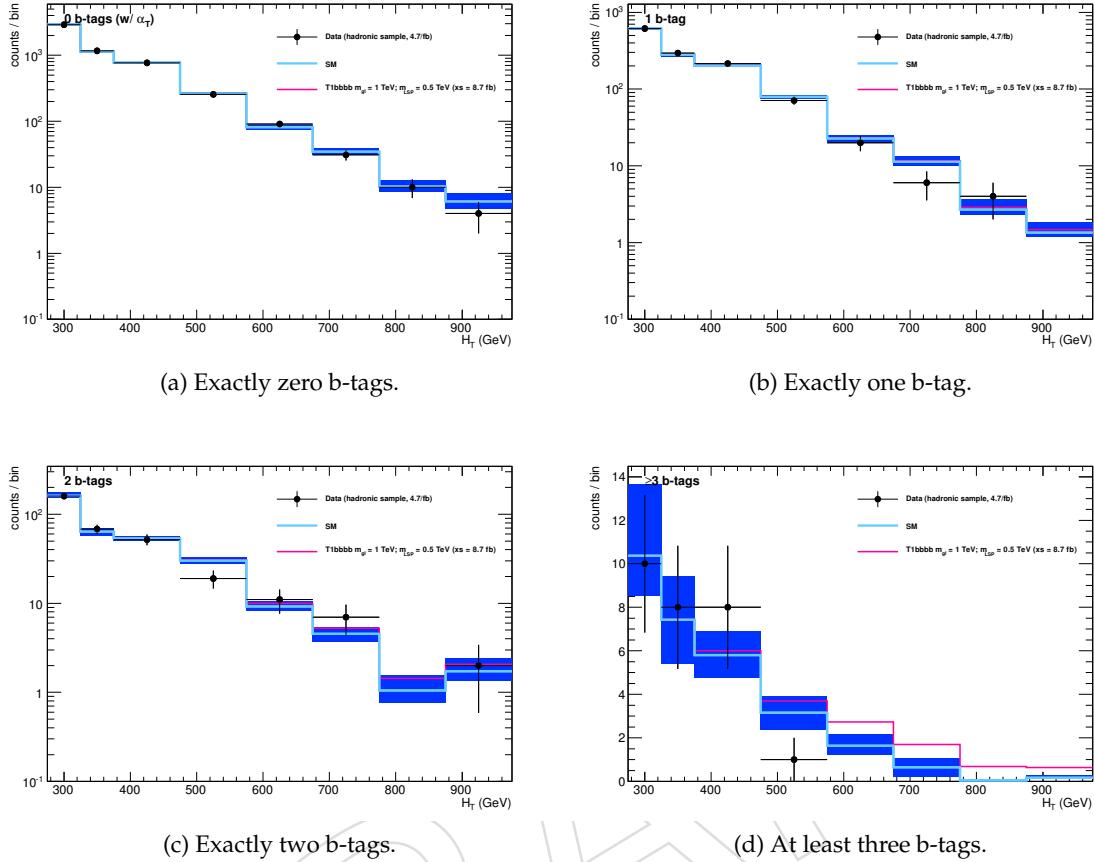


Figure 59: Comparison of the observed yields and SM expectations given by the simultaneous fit in bins of H_T for events in the hadronic signal sample with (a) exactly zero, (b) exactly one, (c) exactly two and (d) at least three b-tags. Shown are the observed event yields in data (black dots with error bars representing the statistical uncertainties) and the expectations and uncertainties, as determined by the simultaneous fit, for all SM processes (light blue solid line). For illustrative purposes, an example model, defined by the topology T1bbbb and the masses $m_{\tilde{g}} = 1000 \text{ GeV}$ and $m_{\text{LSP}} = 500 \text{ GeV}$ (i.e. near the observed limit line), is superimposed on the SM expectation (magenta solid line).

1319 L Comparison of asymptotic CLs vs CLs+toys for T1bbbb

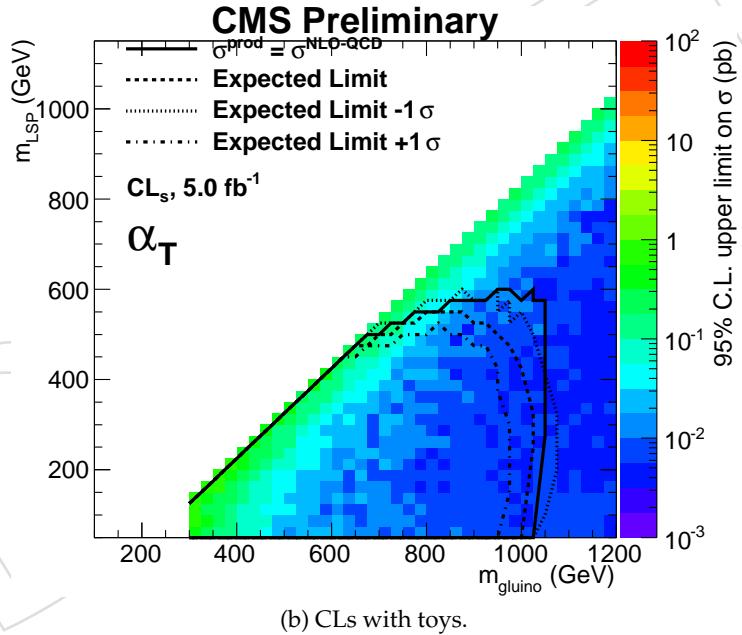
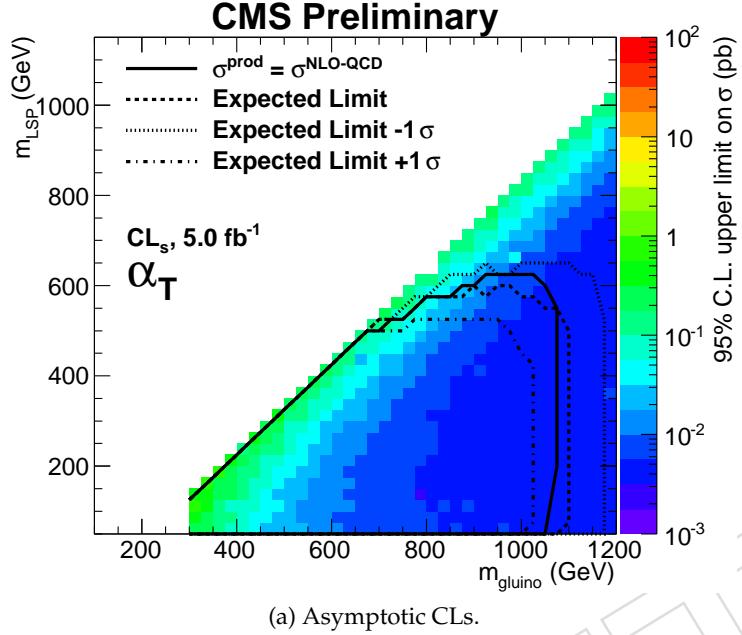


Figure 60: Comparison of limits in the T1bbbb plane using (a) asymptotic CLs and (b) CLs with toys. The observed limit (solid line) is shown on each plot, along with the expected limit (dashed line) and the -1σ (dotted line) and $+1\sigma$ (dot-dashed line) uncertainties. Due to small event counts at mid to high HT in the $n_b \geq 3$ slice, the asymptotic and toy based calculations do not seem to agree well. The limits given by the CLs with toys approach have been validated for a few model points and seem appropriate.

1320 **M Signal efficiency**

1321 **M.1 CMSSM**

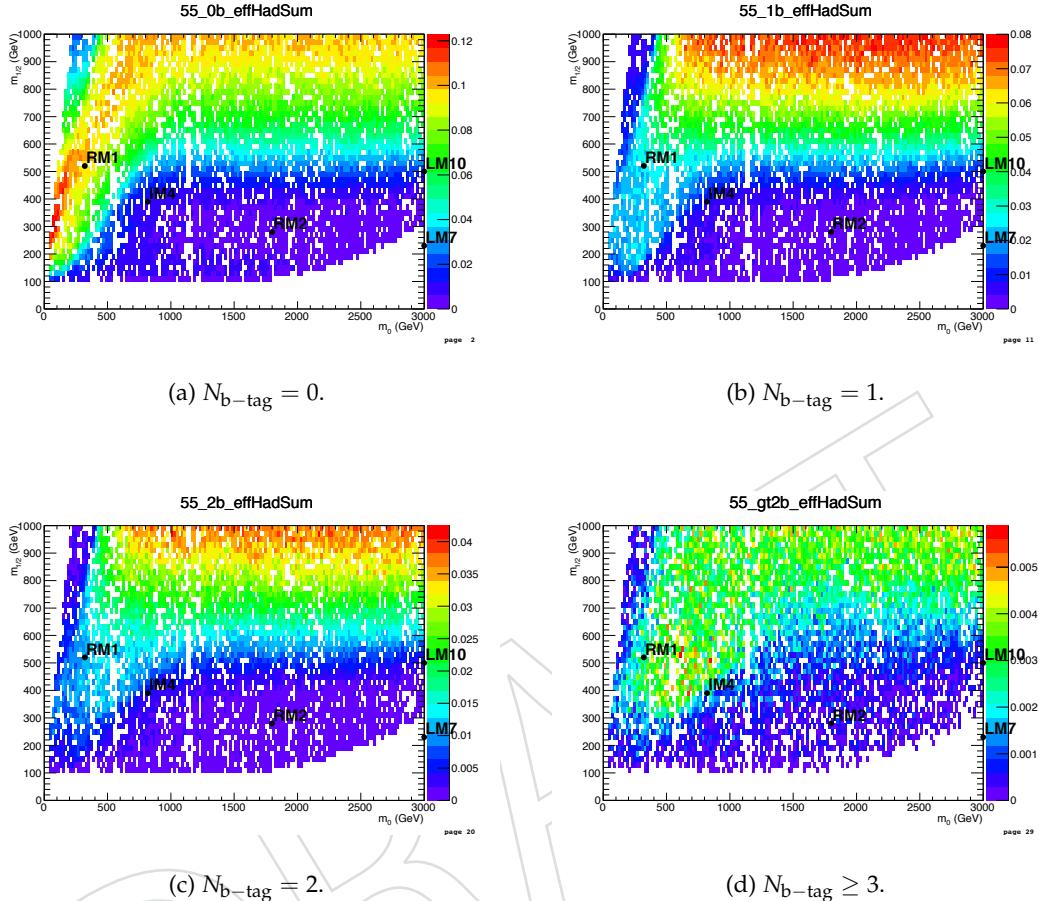


Figure 61: Signal efficiency in the $(m_0, m_{1/2})$ plane of the CMSSM, of the full hadronic signal selection, integrating over all eight H_T bins and requiring (a) exactly zero, (b) exactly one, (c) exactly two, and (d) at least three b-tags per event.

1322 M.2 T1

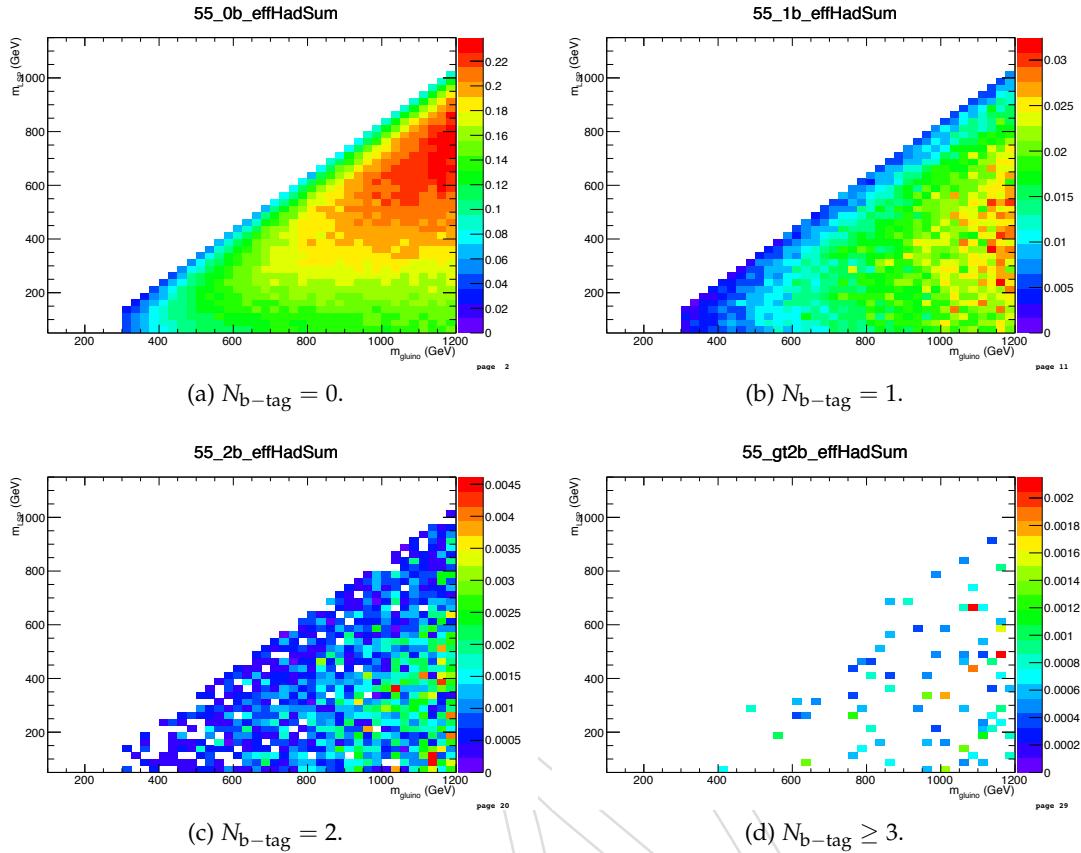


Figure 62: Signal efficiency in the $(m_{\text{gl}}, m_{\text{LSP}})$ plane of the T1 simplified model, of the full hadronic signal selection, integrating over all eight H_T bins and requiring (a) exactly zero, (b) exactly one, (c) exactly two, and (d) at least three b-tags per event.

1323 M.3 T2

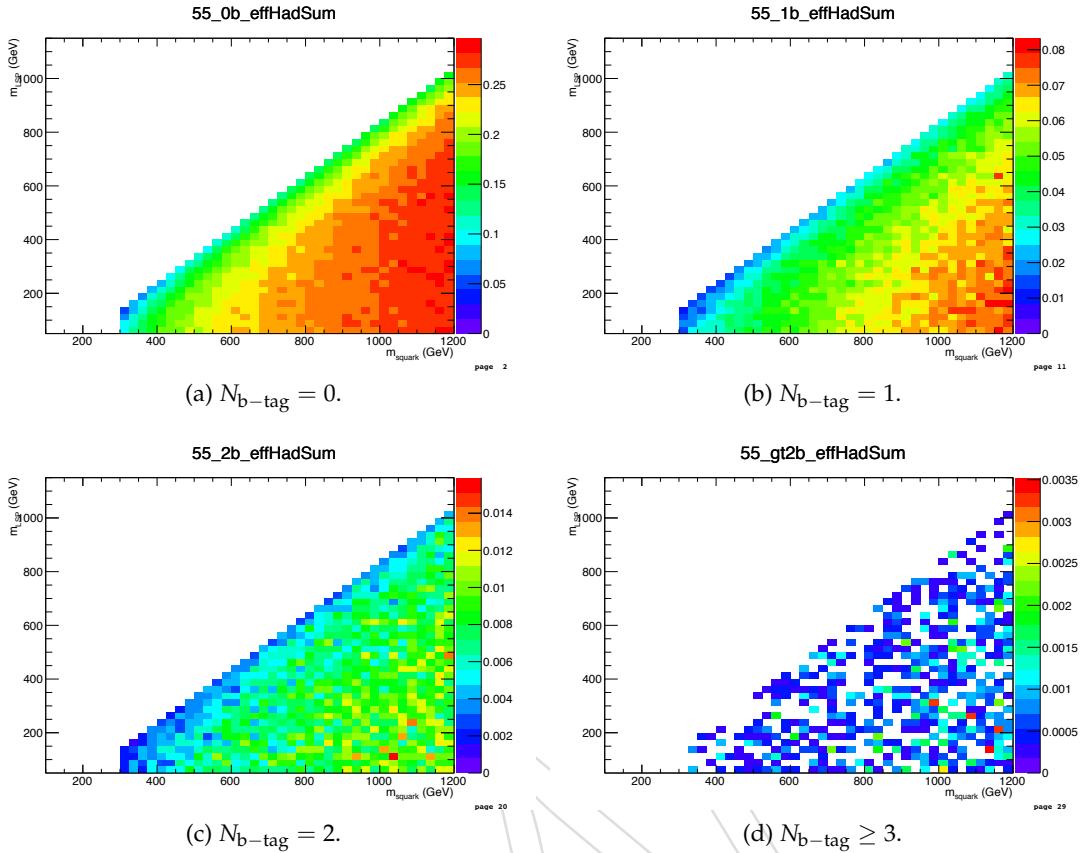


Figure 63: Signal efficiency in the $(m_{\text{sq}}, m_{\text{LSP}})$ plane of the T2 simplified model, of the full hadronic signal selection, integrating over all eight H_T bins and requiring (a) exactly zero, (b) exactly one, (c) exactly two, and (d) at least three b-tags per event.

1324 M.4 T2tt

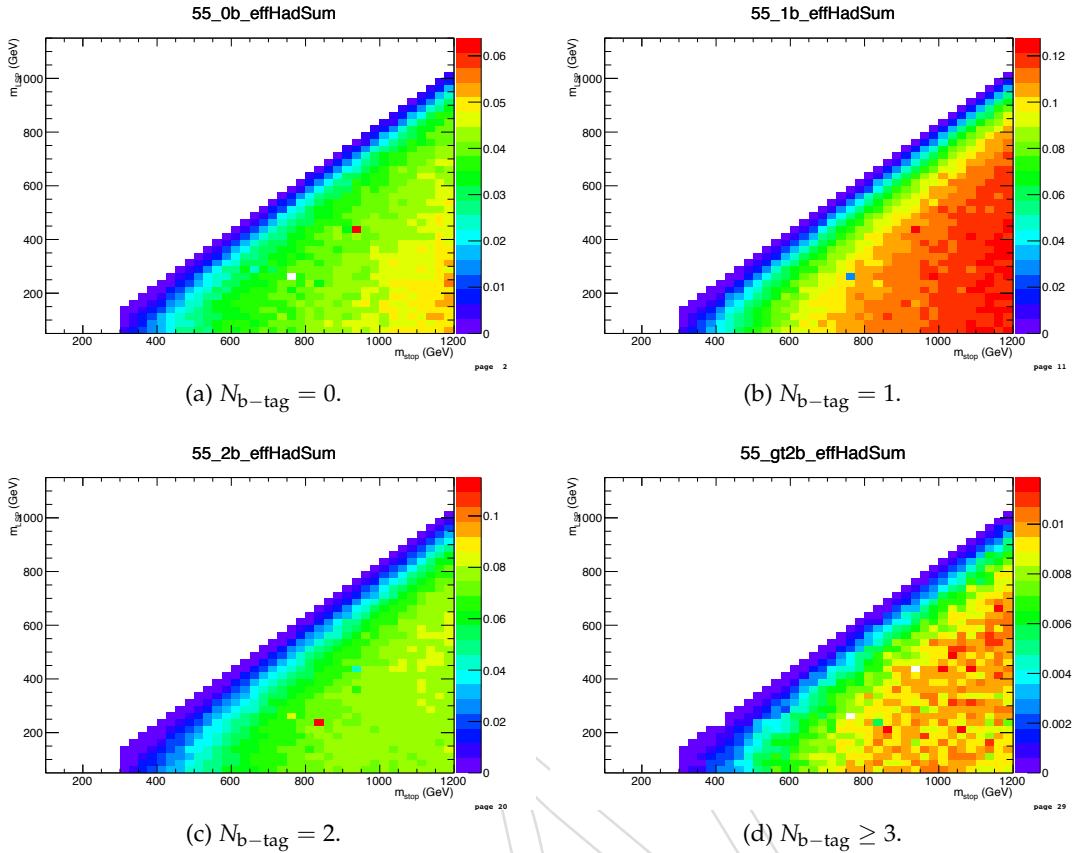


Figure 64: Signal efficiency in the (m_{sq}, m_{LSP}) plane of the $T2tt$ simplified model, of the full hadronic signal selection, integrating over all eight H_T bins and requiring (a) exactly zero, (b) exactly one, (c) exactly two, and (d) at least three b-tags per event.

1325 M.5 T2bb

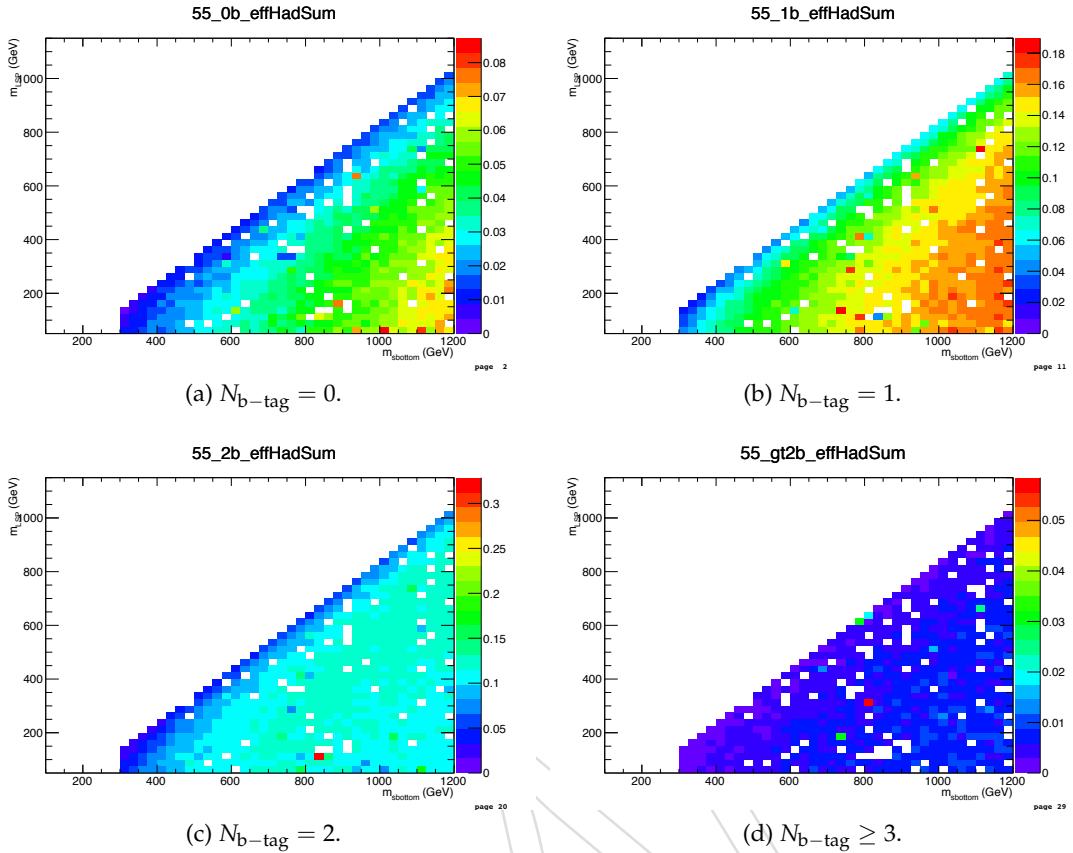


Figure 65: Signal efficiency in the (m_{sq}, m_{LSP}) plane of the T2bb simplified model, of the full hadronic signal selection, integrating over all eight H_T bins and requiring (a) exactly zero, (b) exactly one, (c) exactly two, and (d) at least three b-tags per event.

1326 M.6 T1tttt

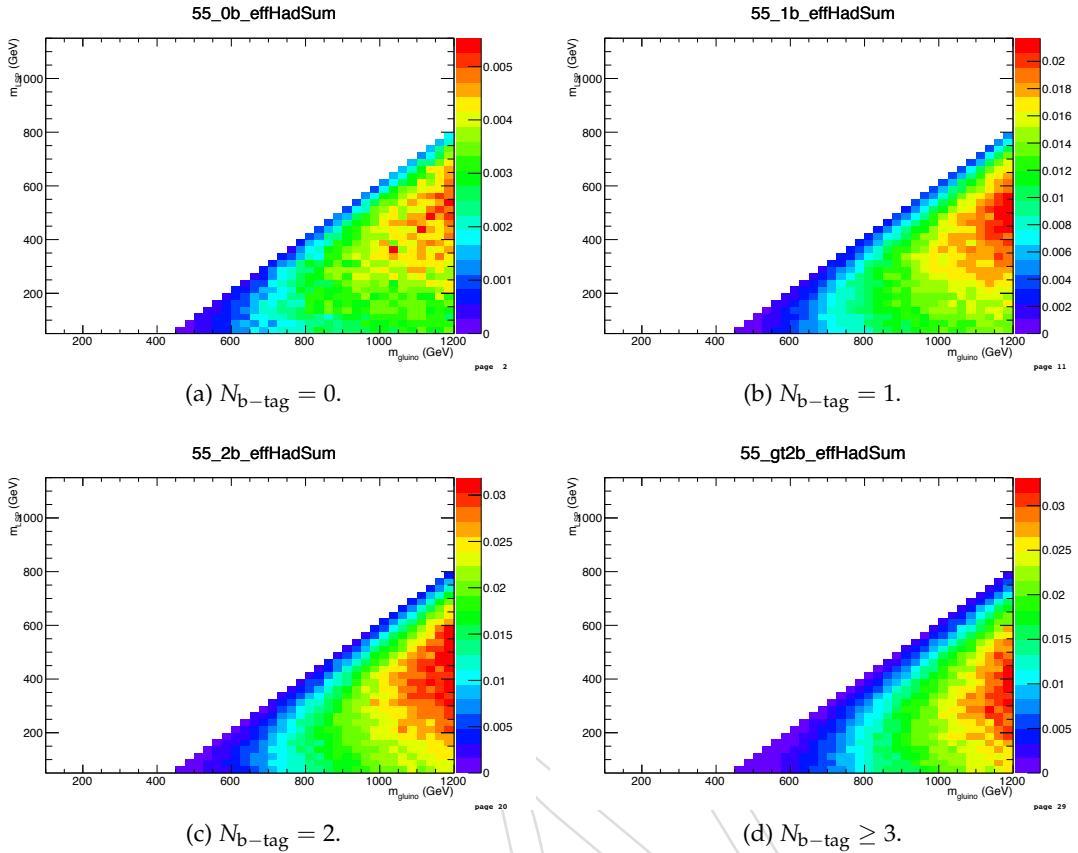


Figure 66: Signal efficiency in the $(m_{\text{glino}}, m_{\text{LSP}})$ plane of the T1tttt simplified model, of the full hadronic signal selection, integrating over all eight H_T bins and requiring (a) exactly zero, (b) exactly one, (c) exactly two, and (d) at least three b-tags per event.

1327 **M.7 T1bbbb**

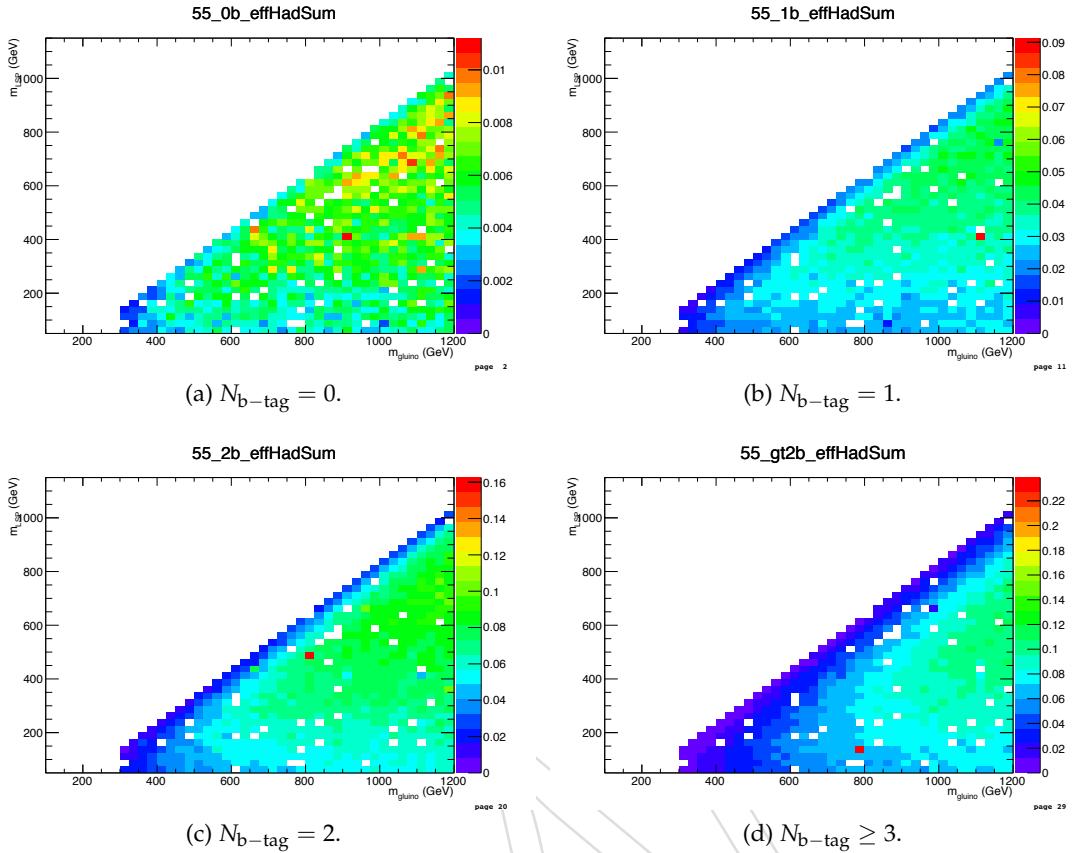


Figure 67: Signal efficiency in the $(m_{\text{glino}}, m_{\text{LSP}})$ plane of the T1bbbb simplified model, of the full hadronic signal selection, integrating over all eight H_T bins and requiring (a) exactly zero, (b) exactly one, (c) exactly two, and (d) at least three b-tags per event.

1328 M.8 Signal contamination for T1tttt

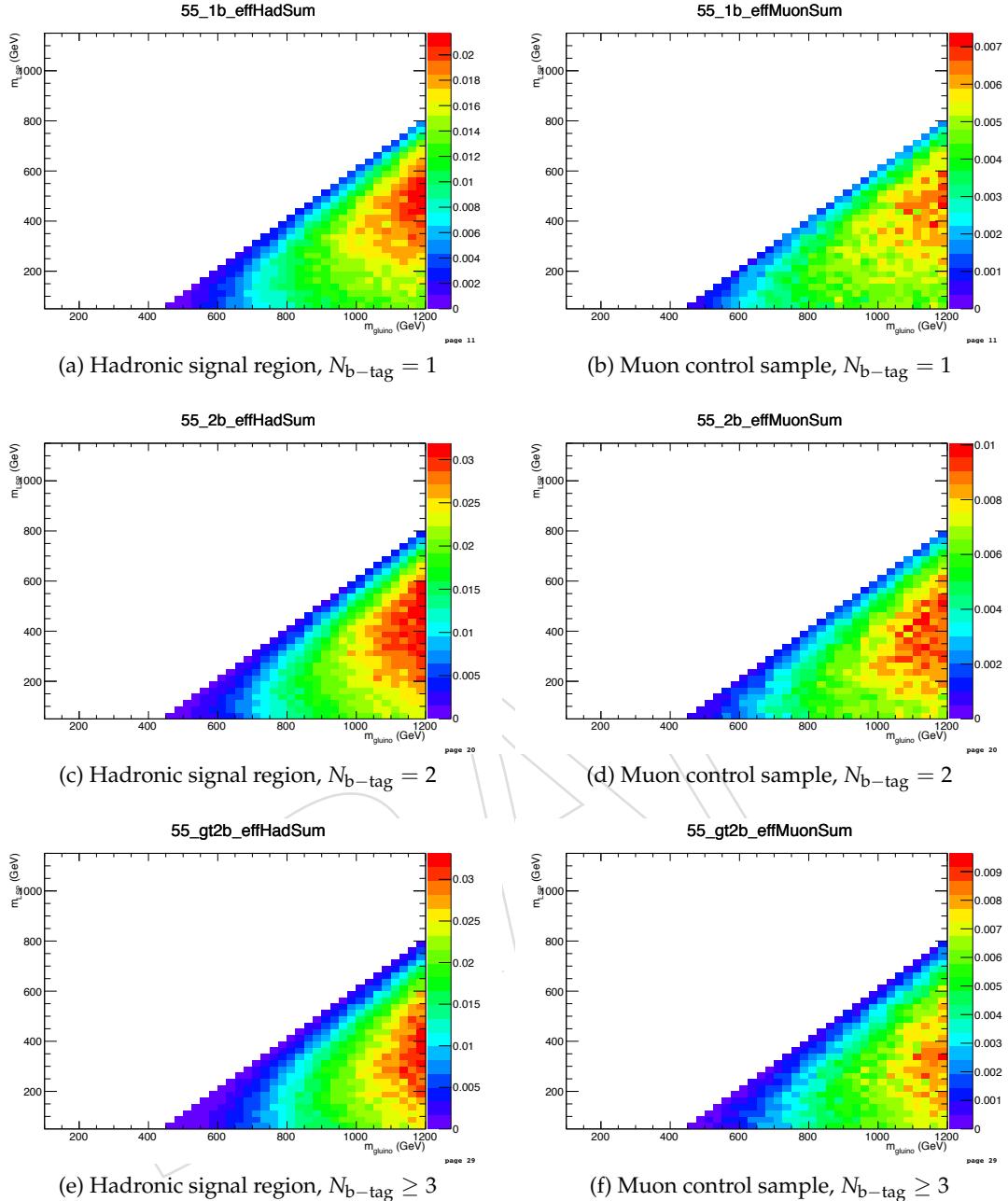


Figure 68: Signal efficiency in the planes of simplified model T1tttt, of the (left) hadronic signal sample selection or (right) single muon control sample selection, integrating over all eight H_T bins and requiring (top) exactly one, (middle) exactly two, or (bottom) at least three b-tags per event.