Searching for Supersymmetry with the α_T variable in $p\bar{p}$ collisions with the CMS Detector at the Large Hadron Collider

Zoe Hatherell

A thesis submitted in fulfilment of the requirements
for the degree of Doctor of Philosophy
to Imperial College London
December 2011

Contents

Chapter 1

Searching for Supersymmetry with α_T in all-hadronic events

The analysis presented here represents a model-independent search for new physics in the all-hadronic channel, where the final state is defined by the presence of jets and missing energy. Designed to search for signs of supersymmetry whilst remaining sensitive to other new physics models, an inclusive strategy is used imposing restrictions only on the final state. Events are chosen based on their compatibility with a topology of heavy new particles pair-produced in p-p collisions, which decay through a chain with an end product which is stable and undetectable.

Isolating these new physics events from Standard Model background processes is essential in order to identify an excess. Controlling the dominant background from QCD Multijet processed is the central feature of the strategy, implementing use of the powerful discriminant, the α_T variable described in Chapter **??**. The remaining backgrounds from electroweak processes may then be accounted for using data-driven estimation techniques in muon and photon control samples.

1.1 Samples

This analysis uses datasets both from Monte Carlo simulation (MC) and of data recorded by the CMS detector in 2011.

1.1.1 Monte Carlo Simulation

Datasets of simulated events with calculated cross-sections are required for any analysis at the LHC. Due to the hadron-collider nature these rely on the PDF's

Standard Model Background

CMSSM SUSY Signal

For the purpose of understanding the possible yields from CMSSM SUSY, two mSUGRA parameter points are used. CMS has a dedicated set of 10 Low Mass(LM) points designed for initial data-taking from which we have chosen LM4 and LM6, the values of which are found in Table 1.1.

mSUGRA Point	m_0	$m_{1/2}$	A_0	$\tan \beta$	$\text{sign}(\mu)$
LM4	210 GeV	285 GeV	0	10	+
LM6	85 GeV	400 GeV	0	10	+

Table 1.1: The two CMSSM SUSY signal points used and their corresponding mSUGRA parameter values.

These points are chosen for their existence above the exclusion limit set previously, shown in Figure 1.1 on the exclusion plot from the 2010 iteration of this analysis[1].

The

1.1.2 Data Sample

This analysis considers data collected by CMS at $\sqrt{s} = 7$ TeV in 2011 between March and June, during the data-taking period known locally as Run2011A. Analyses use only the data taken whilst CMS was fully operational, and thus the data used were specified by the certified list of "good runs" that correspond to $1.1 fb^{-1}$ of integrated luminosity.

As described previously, loose requirements on the types of triggers passed allow the sorting of the data into each Primary Dataset (PD). For the hadronic analysis the HT dataset is used, with basic low-threshold H_T triggers required. Higher threshold triggers are applied subsequently as part of the analysis selection, detailed in Section 1.2.
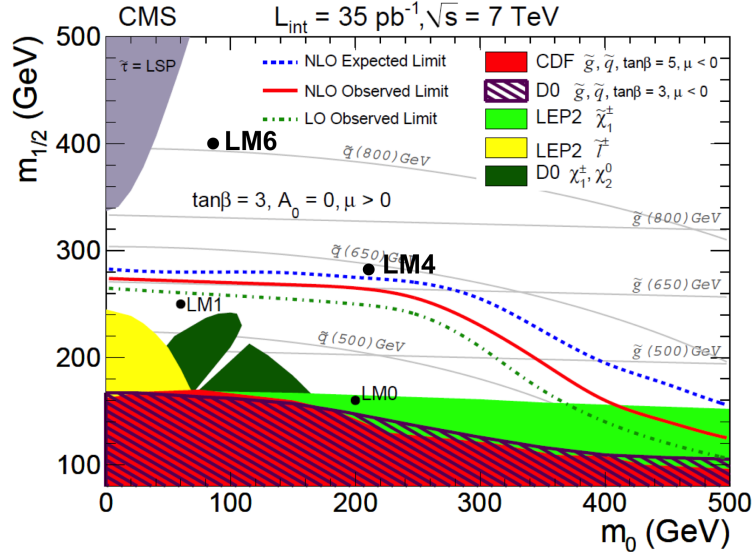
Figure 1.1: The exclusion limit set in the previous incarnation of this analysis with the $35pb^{-1}$ 2010 CMS dataset shown in the $m_0 - m_{1/2}$ plane with $\tan\beta = 10$, $A_0 = 0$ and $sign(\mu) = +$. Reference points LM4 and LM6 to be used in the 2011 analysis are illustrated, in the region yet to be excluded. Previously used reference points LM0 and LM1 are shown below the limit already excluded [1].

For the data-driven estimation techniques muon and photon control samples are defined. For the muon control sample, the HT PD is also used, as a low muon p_T requirement makes this better suited than the dedicated Muon PD. However the photon control sample uses the dedicated Photon PD that requires some basic threshold photon triggers to be passed.

1.2 Trigger

In order to select the signal events and minimise the contamination from backgrounds, a set of selection criteria is applied. As described previously in Section , data collected by the CMS detector is stored and organised according to the L1 and HLT trigger paths passed. Each given dataset then undergoes the offline event reconstruction described in Chapter RECON/IDEN, after which n-tuples are constructed from which we apply cuts on objects and analysis variables.

Previous incarnations of this analysis for the 2010 dataset REFF used a set of pure H_T triggers, however these are unsuitable for the 2011 analysis as these have too high thresholds for the analysis due to the increase in instantaneous

luminosity. The use of cross-object triggers is now employed, requiring events that pass thresholds in both H_T and $\displaystyle{\not}H_T$ for the signal region, and using the lowest un-prescaled thresholds available to ensure signal yields are accurate.

As this analysis makes use of those events which fail the selection criteria also, the hadronic control sample, the pre-scaled H_T triggers are still used taking into account the pre-scaled factors. In the muon control sample, due to the low p_T threshold we use the same triggers as for the hadronic signal sample, and the photon sample makes use of the single photon trigger paths.

Having passed one of the un-prescaled triggers the events are subjected to a series of cuts on objects and analysis variables in order to select the event topologies required and minimise the background contamination.

1.3 Object Definitions

1.3.1 Jets

The jets used in this analysis are Calo Jets, reconstructed as described in Section **??** using the anti-k_T jet clustering algorithm. In addition, a reconstructed jet must pass an additional selection in order to be considered for the analysis:

- Corrected jet transverse momentum requirement of $p_T > 50$ GeV

- Jet pseudo-rapidity $|\eta| < 3$ required to ensure within the fiducial range of the calorimeter systems.

- Passes "loose" jet identification criteria to reject jets resulting from unphysical energy using cuts in Table 1.2.

Any jet which pass the E_T and η requirements but fails the "loose" identification criteria is noted, and the event is marked as containing an "odd" jet, as the presence of such a particle reflects an event whose kinematics are poorly understood and may therefore lead to a misleading $\displaystyle{\not}H_T$.

1.3.2 H_T and $\displaystyle{\not}H_T$

The calculation of both H_T and $\displaystyle{\not}H_T$ is performed using only the jets selected by the selection above in Section 1.3.1.

Definition	Variable	Cut
Fraction of jet energy contributed by the "hottest" hybrid photo-diode	f_{HPD}	< 0.98
Minimum number of cells required to contribute 90% of the jet energy	N^{90}_{cells}	≤ 2
Fraction of jet energy contributed by deposits in ECAL	f_{EM}	> 0.01
Balance of the energy measured in the short(E_S) and long(E_L) HF fibers.	$R_{HF} = \frac{(E_S - E_L)}{(E_S + E_L)}$ (if $p_T^{jet} > 80$ GeV)	$R_{HF} > -0.9$ ($-0.9 < R_{HF} < 1$)

Table 1.2: Set of cuts applied in "loose" CaloJet ID used to reject jets resulting from fake calorimeter deposits representing unphysical energy. Devised using cosmic run data as a pure sample of non-collider "fake" jets, full details of which can be found in [2]

1.3.3 Muons

Although muons are not required by the analysis, a veto on them must be employed, based on muons that satisfy the following set of criteria:

- $p_T^\mu > 10$ GeV

- $|\eta| < 2.5$

- Relative Combined Isolation $= (Iso_{tracker} + Iso_{ECAL} + Iso_{HCAL})/p_T^\mu < 0.15$[1]

- Passes "tight" muon identification, using cuts shown in Table 1.3.

1.3.4 Electrons

Similarly electrons are also defined for veto purposes, with the definition of an electron in the analysis as that passing the following cuts:

- $E_T^e > 10 GeV$

- $|\eta| < 2.5$

---

[1]The components $Iso_{tracker}(Iso_{ECAL}, Iso_{HCAL})$ represent the sum of $p_T(E_T)$ in the relevant detector component, calculated in a cone of R = 0.3 in $\eta - \phi$ around there muon trajectory. The track hits used to reconstruct the muon are not used and any muon deposit in the calorimeters is removed via a smaller veto cone.

Definition	Variable	Cut
Reconstructed with outside-in algorithm	Global Muon	Required
Reconstructed with inside-out algortithm	Tracker Muon	Required
Global muon track fit quality	χ^2	< 10
Number of hits in the silicon tracker included in track	N_{trk}^{hits}	> 10
Number of pixel hits in N_{trk}^{hits}	N_{pixel}^{hits}	> 0
Number of hits in muon system included in Global Muon	N_{muon}^{hits}	≥ 1
Transverse impact parameter with respect to vertex	d_{xy}	< 2 mm

Table 1.3: *Set of cuts applied in "tight" Muon ID, taken from [3]*

- Combined Isolation = $(\sum_{tracker} p_T + \sum_{ECAL} E_T + \sum_{HCAL} E_T)/p_T^e < 0.15$

- Pass WP95 electron identification implemented using cuts in Table 1.4.

Definition	Variable	Barrel Cut	End-Cap Cut
RMS of the width in η of the crystals about the most energetic crystal in the seed	$\sigma_{i\eta i\eta}$	< 0.01	< 0.03
Difference in ϕ between track and supercluster	$\Delta\phi_{vtx}$	< 0.8	< 0.7
Difference in η between track and supercluster	$\Delta\eta_{vtx}$	< 0.007	< 0.01
Ratio of HCAL energy in $\Delta R = 0.15$ to ECAL seed energy	H / E	< 0.15	< 0.07

Table 1.4: *Set of cuts applied in "WP95" Electron ID, taken from [4] corresponding to an intended 95% efficiency for signal electrons in W events for electrons with $p_T > 20$ GeV.*

1.3.5 Photons

Photons in the analysis are defined by the following set of requirements:

---

[1]where electron $p_T > 20$ GeV, measured with a sample of $W \to e$ events.

- $p_T^\gamma > 25$ GeV

- $|\eta| < 2.5$

- Passes "tight" photon cut-based identification (including isolation) using cuts shown in Table 1.5.

Definition	Variable	Barrel Cut	End-Cap Cut
Tracker Isolation in a cone of R=0.4	Iso_{trk}	$< (2.0GeV + 0.001E_T^\gamma)$	
ECAL Isolation in an outer cone of R=0.4 (inner cone R=0.06 removed).	Iso_{ECAL}	$< (4.2GeV + 0.006E_T^\gamma)$	
HCAL Isolation in an outer cone of R=0.4 (inner cone R=0.15 removed).	Iso_{HCAL}	$< (2.2GeV + 0.0025E_T^\gamma)$	
RMS of the width in η of the crystals about the most energetic crystal in the seed	$\sigma_{i\eta i\eta}$	< 0.013	< 0.030
Ratio of HCAL energy in $\Delta R = 0.15$ to ECAL seed energy	H / E	< 0.05	< 0.05

Table 1.5: Set of cuts applied in "tight" Photon ID, taken from [5]

1.4 Pre-Selection

A basic selection of events used for comparison of distributions between data and monte-carlo events shall be known as "pre-selection", following the details set out in this section.

Vertex Requirement It is required that events have at least one non-fake good primary vertex with $N_{dof} > 4$. Constraints on the vertex position along the beam axis $|z_{vtx}| < 24$ cm and perpendicular to the axis of $\rho < 2$ cm must be satisfied. Events that have many fake tracks are identified as monster events and removed, by requiring that the ratio of High Purity tracks to the total number be greater than 25% in events with more than 9 tracks.

HCAL Barrel and End-cap (HBHE) Noise Fiter Events where noise has been identified in the HCAL are removed also, using an algorithm which checks for Photodetectors which have at east 17 out of 18 channels with an E > 1.5 GeV.

Following these conditions events are then selected by the following criteria, using the definitions of physics objects according to the criteria stated previously:

- Pass triggers as detailed in Section 1.2

- Require events with $N_{jet} \geq 2$

- $N_{muon} = N_{electron} = 0$ to reduce the effects of missing energy from neutrinos.

- $N_{photon} = 0$ to ensure a pure hadronic set of events.

- Events with > 0 "odd" jets are vetoed.

- Additional constraint on the transverse momentum of the two leading jets $p_T^{j1}, p_T^{j2} > 100$ GeV.

- Additional constraint on the leading jet pseudorapidity $|\eta^{j1}| < 2.5$.

- $H_T \geq 275$ GeV

- To protect the quantity α_T from the scenario where many jets exist below the momentum acceptance threshold, the missing energy variable $\not{H}_T$ estimated from jet measurment is compared to the quantity $\not{E}_T$ measured from the calorimters. If the ratio $R_{miss} = \not{H}_T / \not{E}_T > 1.25$, the event is rejected.

- $\alpha_T > 0.55$

1.5   An H_T Shape Analysis

Previous iterations of this analysis strategy with the 35pb^{-1} 2010 LHC dataset [1] used a cut-and-count strategy for all events passing the selection, defining the signal region by an $H_T > 375$ and using lower regions in H_T as control regions. The 2011 analysis follows the same selection but motivated by the increasing luminosity is undertaken as a Shape Analysis in bins of H_T, using the whole range $H_T > 275$ GeV as a signal region. This allows greater sensitivity to states of higher mass.

The set of lower bin edges are as follows: [275,325,375,475,574,675,775,875], where each bin is exclusive with an upper limit corresponding to the lower edge of

the next bin, except in the case of the final bin which is inclusive $H_T >875$. The background estimation techniques employed from data therefore are designed to identify the contribution in each distinct bin.

In order to include the two lowest bins in H_T is is necessary to scale the jet thresholds stated in Section **??** in order to maintain even event kinematics allowing a shape analysis approach. The background from $t\bar{t}$ + jets carries a bias to higher jet multiplicities compared to the other EWK components, and thus with identical jet definitions exhibits a turn on behaviour in H_T. In order to remedy this, the lowest two bins have the two jet energy requirements scaled by $^l/_{375.}$ where l represents the bin lower edge in question. This leads to the thresholds in Table 1.6.

H_T Region	Jet Definition	Second-Leading Jet Cut
$275 < H_T < 325$	$p_T > 36.7$ GeV	$p_T^{j1}, p_T^{j2} > 73.3$ GeV
$325 < H_T < 375$	$p_T > 43.3$ GeV	$p_T^{j1}, p_T^{j2} > 86.7$ GeV
$H_T > 375$	$p_T > 50$ GeV	$p_T^{j1}, p_T^{j2} > 100$ GeV

Table 1.6: The three different regions of jet scaling, with values indicated both for the basic definition of a jet used in the analysis, and the second-to-leading jet energy cut. The former is especially important as this alters the value of H_T as this is calculated using the jets in the event.

1.6 Data to Monte-Carlo Comparisons

1.6.1 Dependence of R_{α_T} on H_T

1.7 Data-Driven Background Estimation

1.7.1 Total background prediction

1.7.2 Estimating EWK background using high p_T using W+Jets events
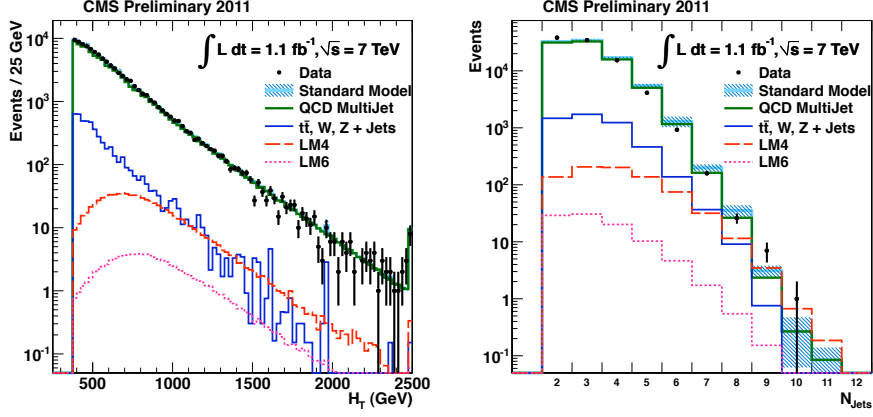
Types of decay contributing to Muon Control Sample

1.7.3 Estimation Z $\nu\bar{\nu}$ + jets background using photon + jets events
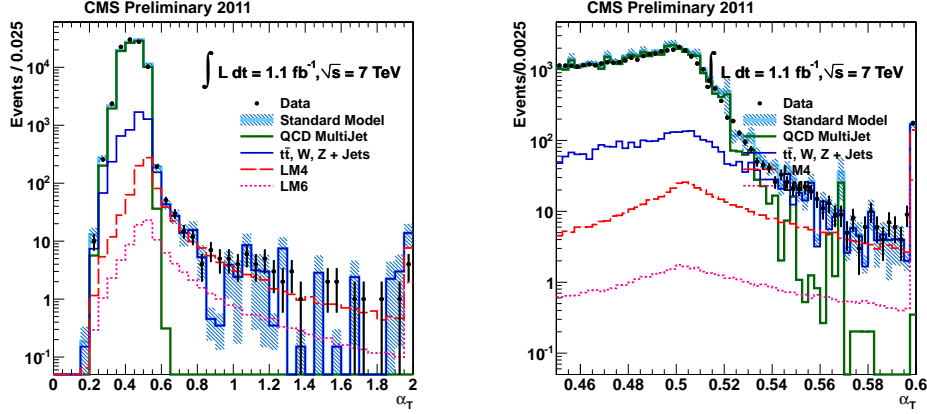
1.8 Systematic Uncertainties

1.9 Simultaneous Fit

1.10 Limits

1.11 Conclusion

(a) Comparison of H_T between data and MC for the hadronic selection for $H_T \geq 375$ GeV and MHT > 100 GeV.

(b) Comparison of the jet multiplicity between data and MC for the hadronic selection, for $H_T \geq 375$ GeV and MHT > 100 GeV.

(c) Comparison of the α_T distribution between data and MC for the hadronic selection, for $H_T \geq 375$ GeV and MHT > 100 GeV.

(d) Comparison of the α_T distribution highlighting the agreement on the sharply falling edge between Data and Monte Carlo for the hadronic selection, in the region $H_T \geq 375$ GeV.

Figure 1.2: Comparisons of 1.1 fb^{-1} 2011 7TeV CMS Data and equivalently weighted Monte-Carlo in basic kinematic quantities prior to the α_T selection cut.
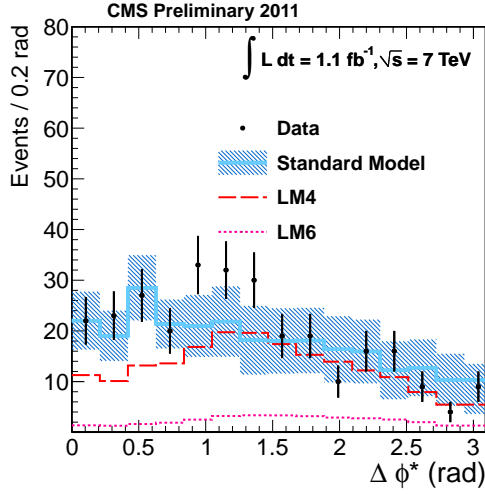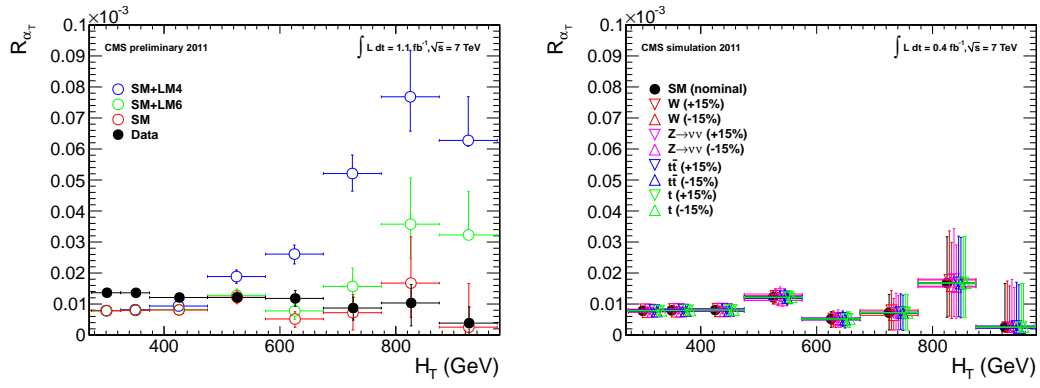
(a) $\Delta\Phi^*$ distribution after α_T selection.
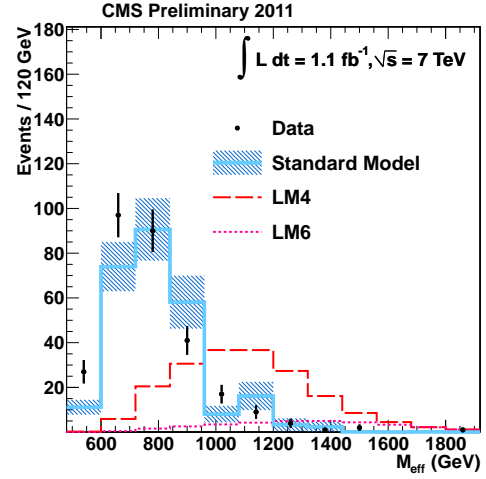
(b) The effective mass distribution, $M_{\text{eff}} = H_T + \text{MHT}$, of the events passing the α_T selection.
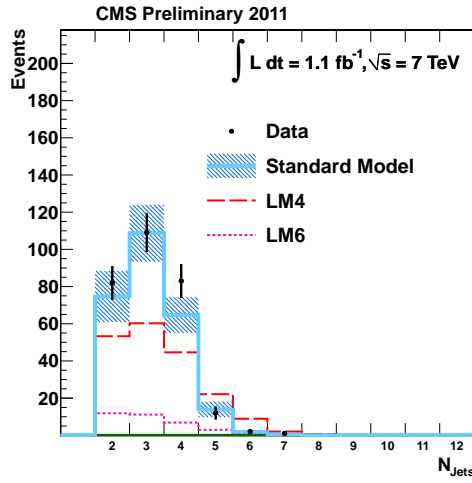
(c) Jet multiplicity after α_T selection.

Figure 1.3: Comparisons of 1.1 fb^{-1} 2011 7TeV CMS Data and equivalently weighted Monte-Carlo in basic kinematic quantities after the α_T selection cut.

Figure 1.4: (Left) The dependence of R_{α_T} on H_T for events with $N_{\text{jet}} \geq 2$. (Right) Dependence of rat on H_T when varying the effective cross-section of the four major EWK background components individually by $\pm 15\%$. (Markers are artificially offset for clarity.)
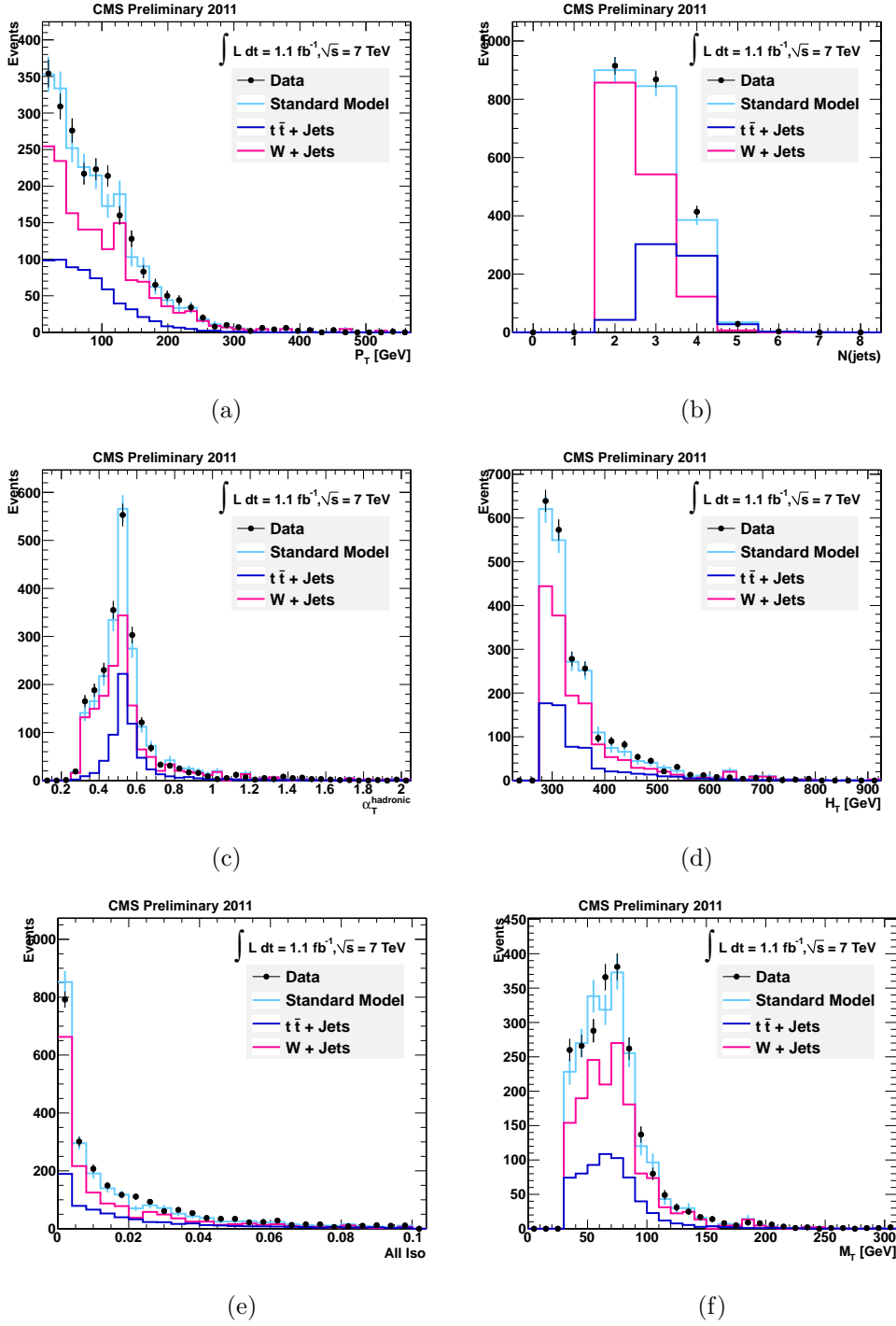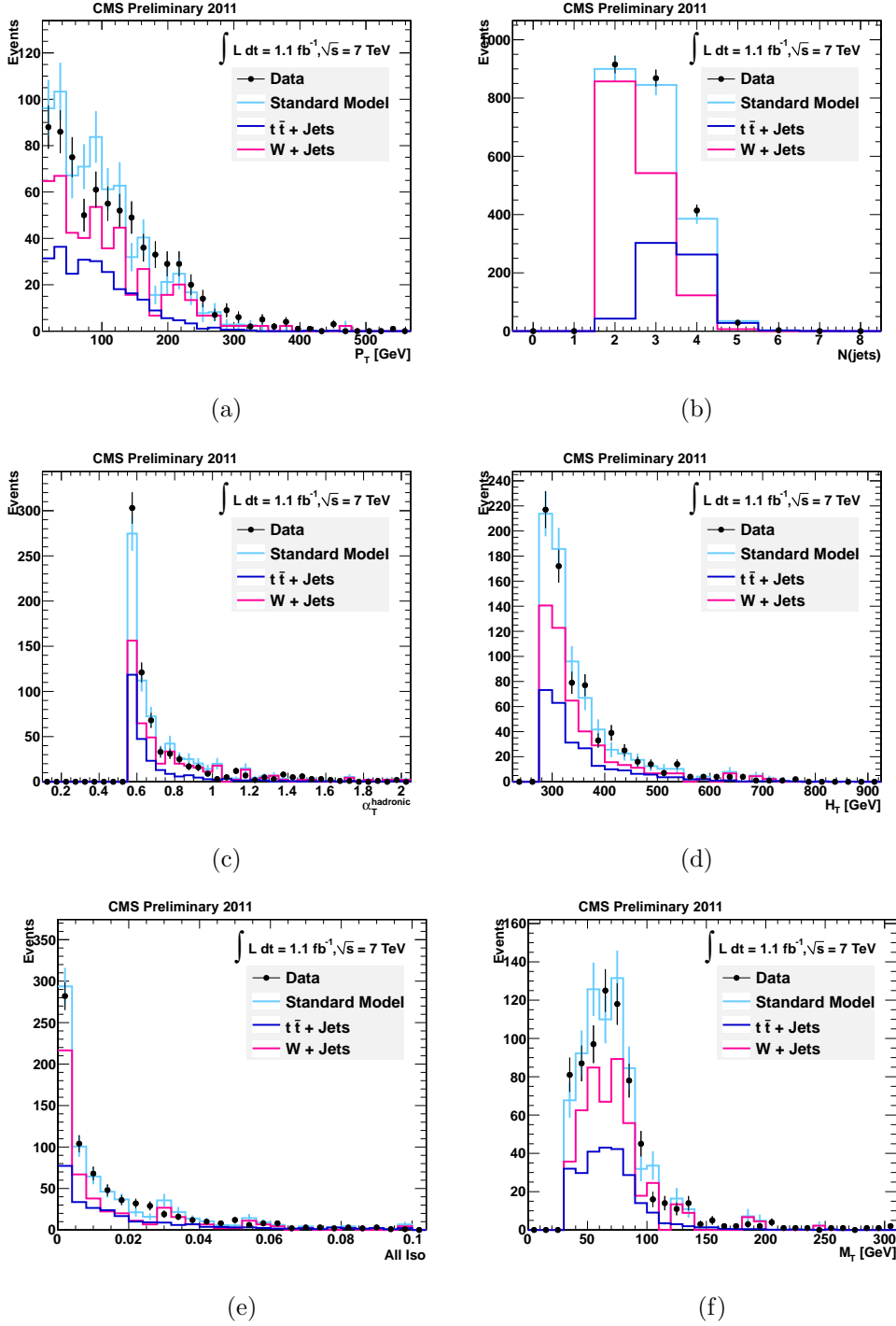
Figure 1.5: Data - Monte Carlo comparisons for the muon control selection before the $\alpha_T > 0.55$ cut is applied, shown for (a) α_T and (b) H_T, (c) Muon Combined Isolation and (d) M_T. A cut of HT >375 GeV has been applied, to select the region of fixed jet thresholds.
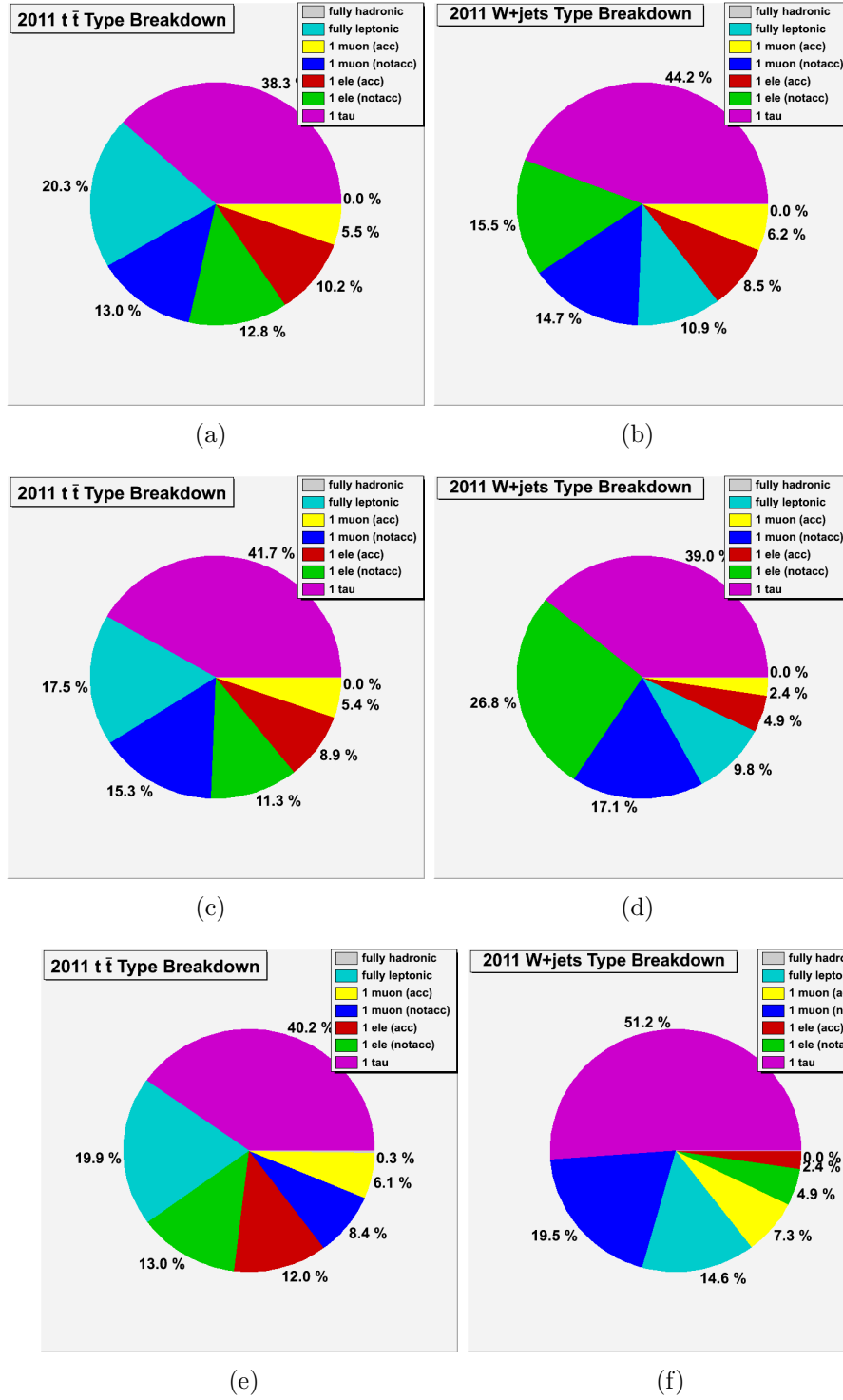
Figure 1.6: Data - Monte Carlo comparisons for the muon control selection after the $\alpha_T > 0.55$ cut is applied, shown for (a) H_T and (b) M_T, (c) Muon Combined Isolation and (d) M_T. A cut of HT >375 GeV has been applied, to select the region of fixed jet thresholds.

16

Figure 1.7: Type breakdown of decays resulting in an event selected by the Muon Control selection, shown using Monte-Carlo truth information separately for $t\bar{t}+$ jets (left) and $W+$ jets (right) events. The breakdown is made separately for each jet-scale case: $275<H_T<325$ GeV (top), $325<H_T<375$ GeV (middle), and $H_T>375$ GeV (bottom).
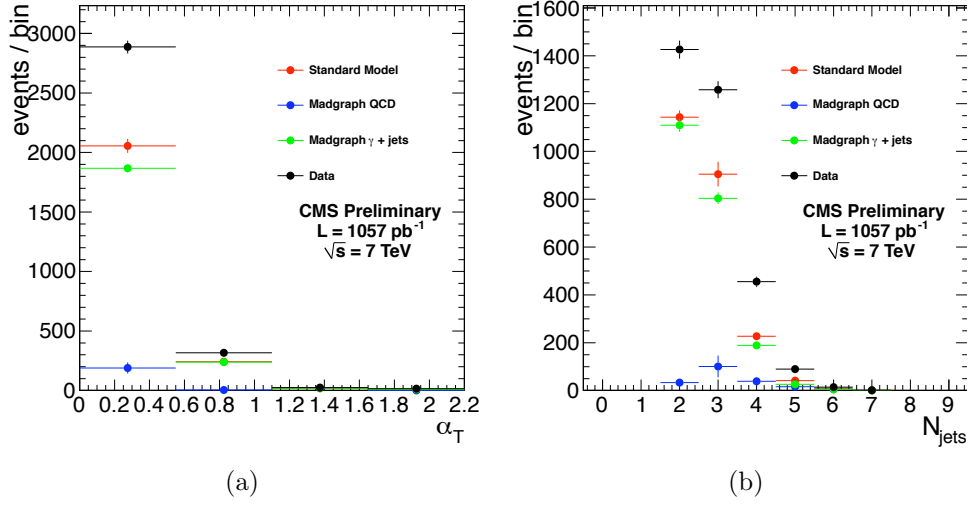
(a)   (b)

Figure 1.8: Data-MC comparisons for the photon control sample. $H_T > 375$ GeV and $\cancel{H}_T/H_T > 0.4$ are required. Left: the distribution of α_T. Right: the distribution of the number of jets.
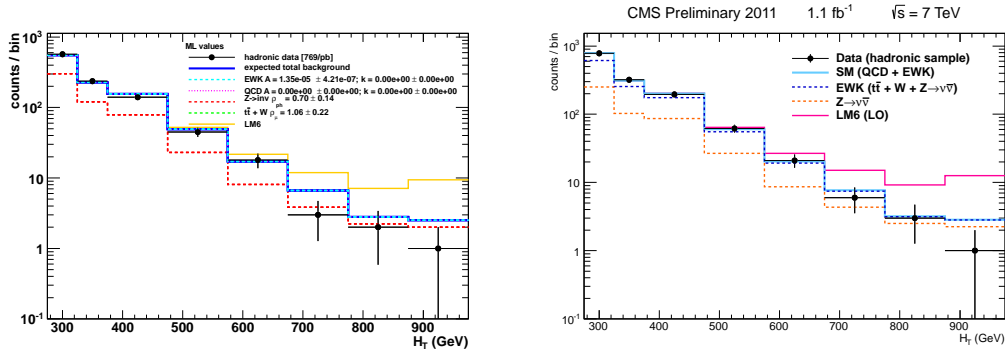

Figure 1.9: H_T distribution for events in the hadronic signal sample for scenario a) (left) and scenario b) (right). Shown are the events observed in data (black points), the outcome of the fit (blue line) and a breakdown of the individual background contributions as predicted by the control samples. A possible signal contribution from benchmark point LM6 is indicated as well (yellow line).
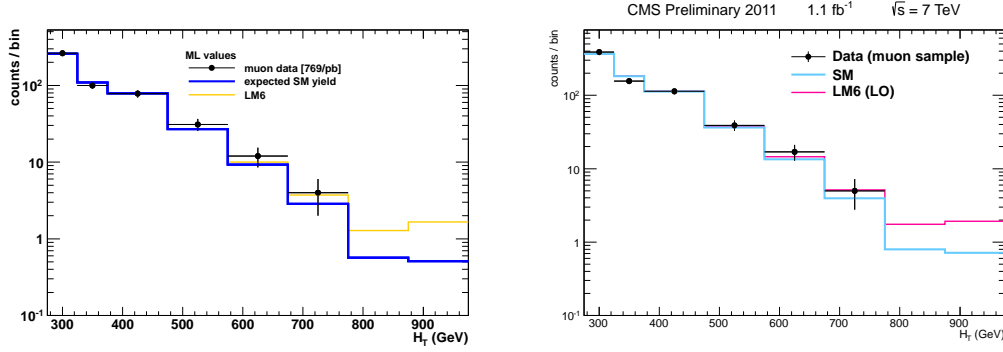
Figure 1.10: H_T distribution for events selected in the muon control sample for scenario a) (left) and scenario b) (right). Shown are the events observed in data (black points), the outcome of the fit (blue line) and the MC expectation (dashed line). A possible signal contribution from benchmark point LM6 is indicated as well (yellow line).
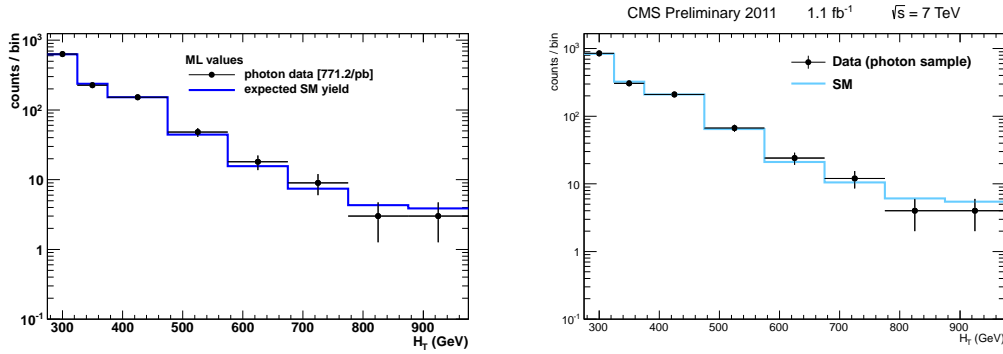

Figure 1.11: H_T distribution for events selected in the photon control sample for scenario a) (left) and scenario b) (right). Shown are the events observed in data (black points), the outcome of the fit (blue line) and the MC expectation (dashed line).
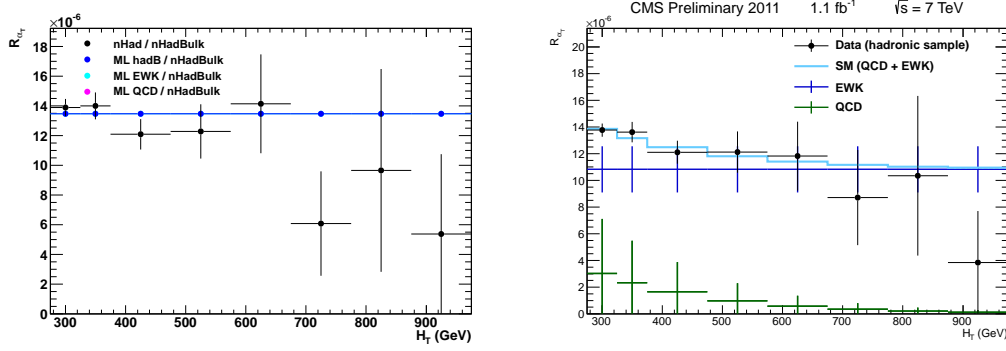
Figure 1.12: R_{α_T} as a function of H_T as observed in data (black points) and the results of the fit assuming different scenarios: a) (left) and b) (right) .
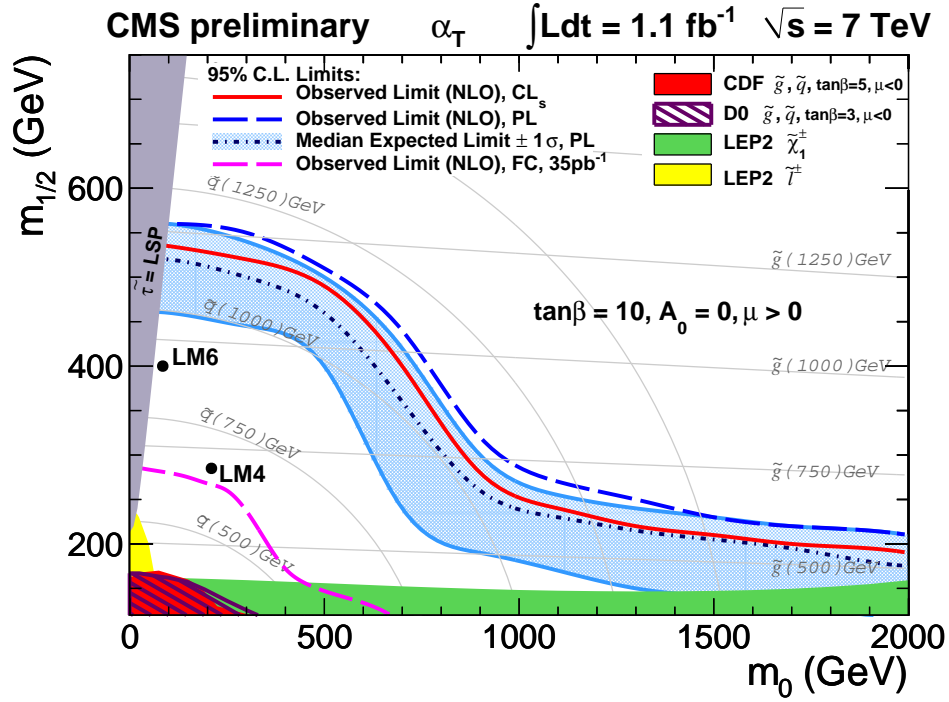

Figure 1.13: Observed and expected 95% CL exclusion contours in the CMSSM $(m_0, m_{1/2})$ plane $(\tan\beta = 10, A_0 = 0, \mu > 0)$ using NLO signal cross sections using the Profile Likelihood (PL) method. The expected limit is shown with its 68% CL range. The observed limit using the CL_s method is shown as well.

Bibliography

[1] Vardan Khachatryan et al. "Search for Supersymmetry in pp Collisions at 7 TeV in Events with Jets and Missing Transverse Energy". *Phys.Lett.*, **B698** (2011) 196–218, 1101.1628.

[2] The CMS Collaboration. "Calorimeter Jet Quality Criteria for the First CMS Collision Data". **CMS-PAS-JME-09-008**.

[3] The CMS Collaboration. "Performance of muon identification in pp collisions at $\sqrt{s} = 7$ TeV". **CMS-PAS-MUO-10-002**.

[4] N. Adam et al. "Measurements of Inclusive W and Z Cross Sections in pp Collisions at $\sqrt{s} = 7$ TeV". **CMS AN 2010/116**.

[5] The CMS Collaboration. "Isolated Photon Reconstruction and Identification at at sqrt(s) = 7 TeV". **CMS-PAS-EGM-10-006**.