Searches for New Symmetries in pp Collisions with the Razor Kinematic Variables at $\sqrt{s} = 7$ TeV
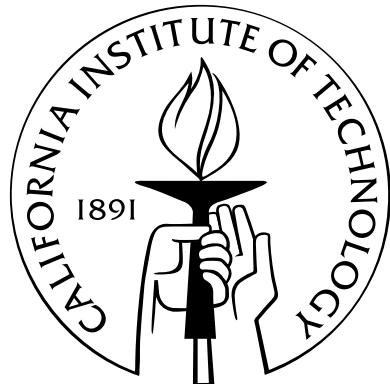
Thesis by

Christopher S. Rogan

In Partial Fulfillment of the Requirements

for the Degree of

Doctor of Philosophy

California Institute of Technology

Pasadena, California

2013

(Defended February 12, 2013)

Acknowledgements

Firstly, I would like to thank my Ph.D. advisor, Maria Spiropulu, whose passion and knowledge for experimental physics is truly contagious. It has been a singular pleasure and privilege to work with her these past years. I would also like to thank Harvey Newman, my original advisor for the first two years of my doctoral studies. I have benefitted enormously from his thoughtful comments and criticisms. It has been my great fortune to have received guidance from these two quite incredible mentors.

Special thanks is reserved for Maurizio Pierini, Joseph Lykken and Alvaro De Rújula. A large portion of this thesis describes work which followed from close collaboration with these remarkable individuals while much of my knowledge of particle physics, in both theory and practice, is the result of their generosity and patience in our interactions. I would also like to thank all the members of the Caltech CMS group, and I hope that I have been able to reciprocate even a fraction of the expertise, enthusiasm, and affection that I have received from this exceptional collection of individuals.

Finally, I would like to thank my parents, Gill and Sean, and Sami. Without their unwavering love and support I would never have been able to complete this thesis. It is dedicated to them.

Abstract

The construction and LHC phenomenology of the *razor* variables M_R, an event-by-event indicator of the heavy particle mass scale, and R, a dimensionless variable related to the transverse momentum imbalance of events and missing transverse energy, are presented. The variables are used in the analysis of the first proton-proton collisions dataset at CMS (35 pb^{-1}) in a search for superpartners of the quarks and gluons, targeting indirect hints of dark matter candidates in the context of supersymmetric theoretical frameworks. The analysis produced the highest sensitivity results for SUSY to date and extended the LHC reach far beyond the previous Tevatron results. A generalized inclusive search is subsequently presented for new heavy particle pairs produced in $\sqrt{s} = 7$ TeV proton-proton collisions at the LHC using $4.7 \pm 0.1 \text{ fb}^{-1}$ of integrated luminosity from the second LHC run of 2011. The selected events are analyzed in the 2D *razor*-space of M_R and R and the analysis is performed in 12 tiers of all-hadronic, single and double leptonsfinal states in the presence and absence of b-quarks, probing the third generation sector using the event heavy-flavor content. The search is sensitive to generic supersymmetry models with minimal assumptions about the superpartner decay chains. No excess is observed in the number or shape of event yields relative to Standard Model predictions. Exclusion limits are derived in the CMSSM framework with gluino masses up to 800 GeV and squark masses up to 1.35 TeV excluded at 95% confidence level, depending on the model parameters. The results are also interpreted for a collection of simplified models, in which gluinos are excluded with masses as large as 1.1 TeV, for small neutralino masses, and the first-two generation squarks, stops and sbottoms are excluded for masses up to about 800, 425 and 400 GeV, respectively.

With the discovery of a new boson by the CMS and ATLAS experiments in the $\gamma\gamma$ and 4ℓ final states, the identity of the putative Higgs candidate must be established through the measurements of its properties. The spin and quantum numbers are of particular importance, and we describe a method for measuring the J^{PC} of this particle using the observed signal events in the $H \to ZZ^* \to 4\ell$ channel developed before the discovery. Adaptations of the razor kinematic variables are introduced for the $H \to WW^* \to 2\ell2\nu$ channel, improving the resonance mass resolution and increasing the discovery significance. The prospects for incorporating this channel in an examination of the new boson J^{PC} is discussed, with indications that this it could provide complementary information to the $H \to ZZ^* \to 4\ell$ final state, particularly for measuring CP-violation in these decays.

Contents

II   Symmetries Beyond the Standard Model   139

8   Beyond the Standard Model Physics: Motivations and Possibilities 140

9   Razor Searches for Supersymmetry   151

Chapter 1

Introduction

It is often said that physics is the study of the world around us. Newtonian dynamics describes the observed motion of the earth around the sun, thermodynamics how a car engine operates. These are the physical laws that hold at the energy and length scales we experience in everyday life. Their predictive power has allowed for the development of contemporary civilization over centuries and eventually flight and space travel. They also include symmetries that have shaped our aesthetic tastes for physical theories. These laws are invariant over time and space, there is not preferred direction to the universe, and classical fields communicate the forces between matter. But there is physics that describe worlds very much unlike the one we experience, which exist at a different scale.

Just as Galilean invariance has given way to Lorentz invariance and finally Einstein's general relativity, we have seen that the laws of physics change at extreme speeds and energy densities. Matter can approach the speed of light but never reach it, a contemporary analogue of Zeno's paradox. The molecule, atom, and nucleus have all been split open to reveal a quantum world with its own laws and forces. This is the world that elementary particle physics describes. The primary difficulty in studying it is its removal from our own.

The particle/wave duality of light and matter indicates that size and energy go inversely; in order to probe increasing smaller length scales, one must use ever-increasing energies. In some sense, it is as simple as $E = mc^2$. In order to produce new, massive particles that interact at these small length scale we must produce interactions at a

commensurate energy scale. With the exceptions of electrons and protons, the first elementary particles discovered at the beginning of the 20th century came from cosmic rays. This source of high-energy particles has since been superseded by terrestrial particle accelerators in laboratories of increasing size and energy. The first of these involved shooting high-energy particles at fixed targets, in the model of the earliest experiments probing atomic structure through scattering. In the Newtonian world this would be sufficient, but special relativity indicates this is not the optimal way for converting the energy of relativistic particles into new matter through interactions. The advent of colliding beams particles, first electrons/positrons followed by proton/proton and proton/antiproton, has improved this efficiency substantially.

Similarly, the technology with which physicists study these high-energy particles has evolved with their energies. The photographic emulsion plates that captured the interactions of the cosmic rays have given way to cloud chambers, then bubble chambers, and now whole detectors made of millions of silicon microstrips, taking our ability to visualize the trajectories of charged particles from analogue to digital. Fixed target detectors have evolved into instruments with nearly 4π geometrical coverage around the interaction point of colliding beams. Never before have we been able to so efficiently control so much energy in so little space and so completely reconstruct the interactions of the quantum world.
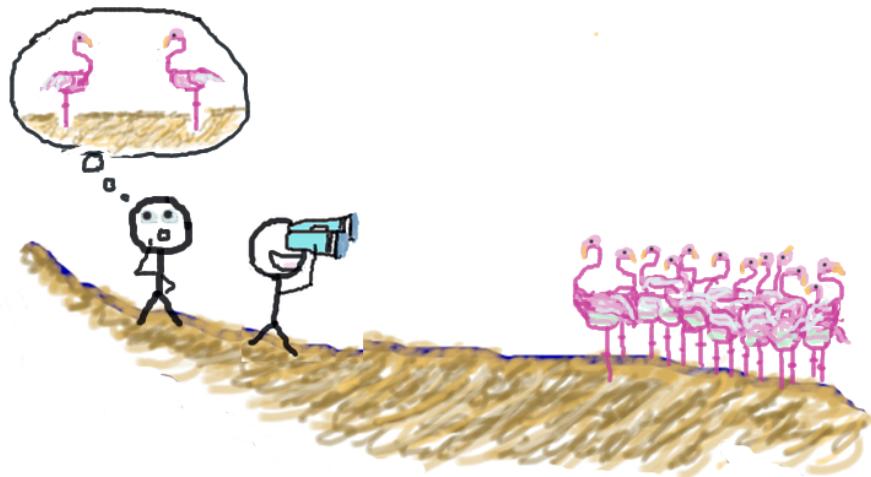
Of course, our understanding of elementary particles has grown with our access to them. With increasing numbers of particles discovered, we have been able to piece together the structure of the quantum world. The discovery of the anti-electron established that the symmetry of Dirac's equations had physical significance [1]. Particle-puzzle-pieces of mesons and baryons appearing from the sky were assembled into the Eightfold Way [2] and finally QCD [3, 4]. The quark model was confirmed once energies were achieved that could resolve the lumpiness of the proton, with quark and gluons jets following from energies large enough to overcome the strong force. And with this structure has come new symmetries, absolute, approximate and broken. The Standard Model (SM) [5, 6] of particle physics developed to explain these high-energy interactions has proved to be extremely successful, and for the last 50

years has survived myriad experimental tests, predicting the existence of the W, Z, and top before their discoveries. But this is not the end of the story. Just as the laws of physics have evolved from one energy scale to another, they could change again as we enter a new world at the TeV frontier.

This thesis concerns this frontier of elementary particle physics, opened through the highest energy particle collider ever realized, the Large Hadron Collider (LHC) and the Compact Muon Solenoid (CMS) experiment which is build to study these proton-proton collisions. It is composed of two parts: The first describes the Standard Model of particle physics and the CMS detector, and studies of the known elementary particles at previously unreachable energies. Chapter 2 gives an introduction to the particle content of the Standard Model, along with overviews of CMS and the LHC machine. This is followed by chapters explaining each of the subcomponents of the CMS detector, and how they are used to measure the properties of elementary particles. Chapter 3 describes the CMS tracking and muon detectors and how, combined with the CMS magnetic field, their measurements are used to reconstruct the trajectories and momenta of charged particles. The design and operation of the electromagnetic and hadronic calorimeters are explained in chapters 4 and 5, respectively, along with the reconstruction of electrons, photons and QCD jets. Chapter 6 talks about the physics of W and Z bosons, and how they can be used to both calibrate the detector and study the SM. Finally, part one concludes with a discussion of the Higgs boson. The CMS and ATLAS experiments recently announced the discovery of a new boson which could be the Higgs [7,8]. Chapter 7 explains how we can prove that it is *the* Higgs through measurements of its quantum numbers, with estimates of how long remains before we can conclude that the SM is complete.

Part II of this thesis is about what might be waiting beyond the SM. In chapter 8, the shortcomings of the SM are explained, along with motivations for expecting evidence of new, heavy degrees of freedom not included in the SM to appear at the LHC. The phenomenology of the theories describing this beyond the SM (BSM) physics is explained and used to motivate the development of the razor kinematic variables which can be used to search for this new physics. The derivation of these variables

is explained in chapters 9 and 10, along with the details of searches for BSM physics using 35 pb^{-1} and 4.7 fb^{-1} of $\sqrt{s} = 7$ TeV CMS data, respectively. Null results are interpreted in the context of models of supersymmetry, with constraints placed on its hypothetical parameters. The phenomenology of the 2D razor-space is described in detail, for both hypothetical signal events and SM backgrounds. Finally, chapter 11 describes new kinematic variables designed for future searches at the LHC which can improve our resolution of the scale of new physics and perhaps uncover the symmetries that govern BSM phenomena.

Part I

CMS and the Standard Model

Chapter 2

CMS and Fundamental Particles

2.1 The Standard Model of Particle Physics

The Standard Model (SM) of particle physics describes the interactions of all of the known elementary particles. The are the fermions, the spin 1/2 matter that makes up our universe, and the particle mediators of the forces which act on these fermions, the spin 1 bosons. The fermions are composed of two groups, leptons and quarks, each with three families of increasing mass and each particle with a corresponding antiparticle. The force carriers (photons, W/Z bosons, gluons) transmit the electromagnetic, weak and strong forces, respectively.

Each of the fermions carries an electric charge. The leptons (electrons, muons, taus) all have charge 1, which is defined by convention to be -1 for leptons and 1 for anti-leptons. The quarks have fractional charges, with quarks coming in an up and down type for each family, with charges 2/3 and $-1/3$, respectively. For each lepton family there is also a neutrino which is neutral and approximately massless. The charged Dirac fermions and their electromagnetic interactions can be described by quantum electrodynamics (QED) [9]. This theory, and the full SM, are gauge theories based on symmetry groups and the requirement of local gauge invariance. Electromagnetism is based on the $U(1)_{em}$ symmetry group, meaning that its laws are invariant under complex phases applied to particle fields. Local gauge invariance further requires that the theory be invariant under *space-time-dependent* phase changes, requiring that the theory is *continuously* invariant under gauge transforma-

tions. In going from a global symmetry to a local one, the charged particles require a force to communicate between different points in space-time in order to ensure that the gauge symmetry is not broken. Hence "gauge-ing" $U(1)_{em}$, or requiring that the Lagrangian is invariant under these continuous transformations, introduces a massless spin 1 gauge boson, the photon, which communicates the electromagnetic force between charged particles. QED, like all the gauge symmetries of the SM, is a renormalizable theory meaning that it is free from divergences resulting from loop Feynman diagrams contributing to particles' self-energies. The physical consequence is that the QED interaction-strength coupling, α_{em}, effectively changes as a function energy, *running* with the scale of the interaction.

In addition to electromagnetism, the quarks also interact through the strong force, which is described by a gauge theory based on the $SU(3)_C$ group, called quantum chromodynamics (QCD) [3, 4]. In addition to electric charge, the quarks also carry a color charge, corresponding to one of three colors. The requirement of local gauge invariance introduces eight massless gauge bosons, the gluons, which communicate the strong force. Unlike the photon, these gluons also carry color charge, with two separate color indices which allow the quarks of different colors to interact through its exchange. The running of the strong force coupling, α_s, through renormalization of QCD is quite different from the QED coupling due to the different group structure and corresponding form factors. This has profound consequences on the phenomenology of colored particles. QCD has the property of *asymptotic freedom*, implying that the strong force actual weakens between particles at higher energies large. Unlike the electromagnetic force which diminishes with distance, the strong force increases on short distance scales. Colored particle also exhibit a phenomena called *confinement*, whereby they can not be isolated singularly or detected directly. The result is that as two quarks pull apart from each other, the gluon fields form narrow tubes of color charge, pulling them together like a rubber band. If there is enough energy to pull the quarks apart, like in LHC collisions of protons, at some point it becomes energetically favorable for a new quark/anti-quark pair to appear from the vacuum along the color tube. For very energetic colored particles, this process will repeat recursively, leading

to a of jet color-neutral baryons and meson flying in the directions of the initial colored particle. Asymptotic freedom also implies that at small scales the size of baryons and mesons the individual quarks behave as free particles. At LHC collision energies, the quarks and gluons of colliding protons interact directly, as if they were free partons.

In the SM, the interactions between the particles are described by a non-Abelian Yang-Mills type gauge theory based on the group $SU(3)_C \times SU(2)_L \times U(1)_Y$, where $SU(3)_C$ is QCD. The SM particles can be organized according to their representations in the $SU(2)_L$ subgroup. The chiral-left component of the lepton and neutrino (which are all left-handed, as far as we know) form an $SU(2)_L$ doublet, while the chiral-right leptons are each in their own singlet. Similarly, the chiral-left components of each family's up and down type quarks are a $SU(2)_L$ doublet, while the chiral-right quarks are each singlets. Gauge-ing $SU(2)_L \times U(1)_Y$ results in a collection of massless spin 1 gauge bosons. The force carriers of the weak force, the $W^{\pm}$ and Z bosons, do have mass, which they acquire through spontaneous electroweak symmetry breaking (EWSB) [10–15]. This is accomplished through the addition of a complex $SU(2)_L$ doublet of spin zero fields, the Higgs field, to the theory. The Higgs acquires a nonzero vacuum expectation value (VEV) from its quadratic potential which breaks the $SU(2)_L \times U(1)_Y$, or electroweak symmetry. It is denoted such because it is a unified description of the weak and electromagnetic forces, which are only bifurcated by the breaking of a global symmetry from EWSB. The VEV is left invariant by once combination of $SU(2)_L \times U(1)_Y$ generators, which gives $U(1)_{em}$ electromagnetism. On the other hand, the other gauge bosons associated with $SU(2)_L \times U(1)_Y$ eat the degrees of freedom associated with the Higgs doublet, giving masses to the spin 1 weak bosons through the addition of a longitudinal degree of freedom. These are, in turn, linear combinations of the broken $SU(2)_L \times U(1)_Y$ generators' gauge bosons. Only one neutral scalar of the Higgs doublet is left, which is associated with the Higgs boson. This same boson gives the fermions their masses through Yukawa couplings which realize fermion mass terms in the Lagrangian when the Higgs acquires a VEV.

2.2　The CMS Experiment

The Compact Muon Solenoid (CMS) [16] is a multipurpose detector which surrounds one of the LHC interaction points. In these high energy proton-proton collisions, sprays of particles will fly out of the point of interaction. With a collection of dedicated subdetectors, CMS is able to measure the properties of these particles, reconstruct their identities and trajectories, and interpret the particles and interactions contained in the event. CMS is shown in figure 2.1.
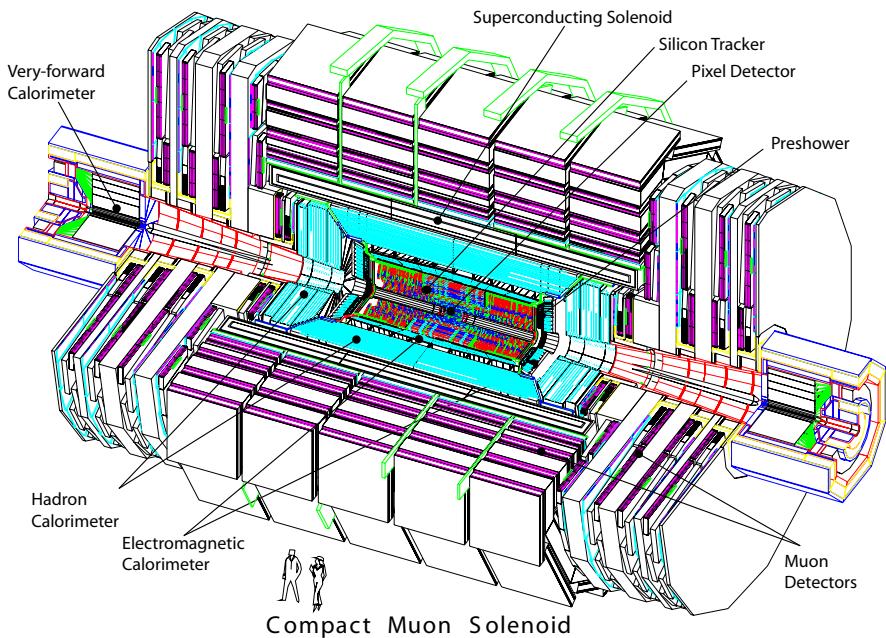
Figure 2.1: Illustration of the CMS detector

The design of CMS is based around the superconducting 4 T solenoidal magnet, which is centered on the beam pipe symmetrically around the interaction point. Charged particles traversing this field will bend as they travel out from the interaction point, and the large field is required to measure their momentum precisely. Inside and around the magnet are layers of subdetectors, each designed to detect and measure specific types of particles. The cylindrical shape of the magnet informs the geometries of these subdetectors; each includes a barrel component arranged in layers of fixed radius from the beam pipe while endcap components are placed in

layers perpendicular to the beam line, covering each end of the experiment. This enclosed geometry yields almost 4π angular coverage for observing all of the particles produced in interactions, with small holes through which the beams pass. The CMS coordinate system has the origin centered at the nominal interaction point, with the y-axis pointing vertically upward, the x-axis pointing radially inward toward to the LHC center, and the z axis points along the beam line. ϕ is measured from the x-axis in the x-y plane, transverse to the beam line, while the polar angle θ is measured from the z-axis. For convenience, pseudorapidity is defined as $\eta = \log\left[\tan\theta/2\right]$ and is equivalent to the rapidity of a massless particle traveling from the origin of CMS.

Each layer of the CMS detector is built to measure a different type of particle. At the innermost part of CMS, inside the solenoidal magnet barrel, lies the silicon pixel tracker, which in turn is surrounded by a silicon strip tracker. These subdetectors are used to reconstruct the trajectories of charged particles traveling through the CMS tracker volume. The large magnetic field combined with the fine spatial resolution of the tracker results in excellent track reconstruction performance. The inner tracker is surrounded by an electromagnetic calorimeter, composed of scintillating lead tungstate crystals. This detector is used to precisely reconstruct the energies of photons and electrons and to identify them through their electromagnetic showers. Outside of the electromagnetic calorimeter is a hadron calorimeter, made of interleaved layers of brass absorber and plastic scintillator. This dense sampling calorimeter is needed to stop the jets of hadrons which are produced in each collision from flying through the detector, measuring their energy in the process. Finally, outside of the barrel magnet sit muon detectors, which can measure the trajectories of muons as they pass through gaseous ionization chambers. With the combined information of each subdetector layer, events can be precisely reconstructed. In total, CMS has hadron calorimeters covering the region $|\eta| < 5$, electromagnet calorimeter crystals to $|\eta| < 3$ and muon and tracking coverage extending to $|\eta| < 2.4$. The partons which interact in LHC collisions only carry a fraction of the protons momentum, samples from its parton density function (PDF). As a result, LHC collisions can have large longitudinal momentum imbalances along the beam line. In the transverse

plane, momentum is approximately conserved, meaning that the large calorimeter pseudorapidity coverage is essential in measuring the energy of all of the high energy particles in the collision in order to infer the transverse momenta of any weakly interacting particles.

While designed with the optimized reconstruction of particles in mind, each subdetector must also satisfy the extreme running conditions of the LHC environment. Each detector is made of radiation-hard components that can withstand the enormous particle flux from collisions. The crystals of the electromagnetic calorimeter must be monitored in real time for radiation-induced transparency changes while the photodetector technology used in each of the calorimeters changes with increasing pseudorapidity in order to withstand more radiation. Similarly, the muon chamber technology is chosen according to the expected muon flux and local magnetic field. In the following chapters each of these subdetectors is described in detail.

2.3   The Large Hadron Collider

The Large Hadron Collider (LHC) [17] is a two-ring superconducting hadron collider. It is installed in the 27 km LEP tunnel at CERN, in Geneva, Switzerland, and at its design performance will provide 14 TeV collisions between 7+7 TeV protons with an instantaneous luminosity of $L = 10^{34}$ cm^{-1}s^{-1}. At present, the LHC machine has achieved 8 TeV collisions and a luminosity exceeding 5×10^{33} cm^{-1}s^{-1}. The protons are supplied to the LHC through an injector chain of smaller accelerators. Linear accelerators (LINACS) feed protons into the Proton Synchrotron Booster (PSB), followed by the Proton Synchrotron (PS) which accelerates them to 25 GeV. This is followed by the Super Proton Synchrotron (SPS), bringing protons to 450 GeV, and finally to the LHC, as illustrated in figure 2.2.

The design of the LHC reflects a balance between the desired performance and restrictions from both its location and cost. With the LEP tunnel as a location, the maximum beam energy is limited by the superconducting magnet strength. Furthermore, the desired collision energy precludes electron beams because of large syn-
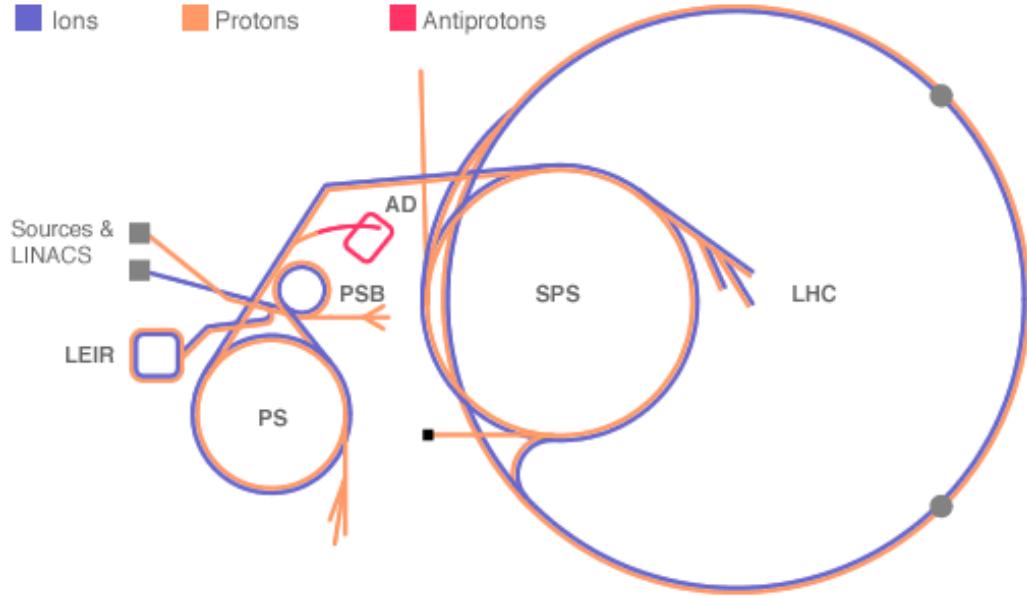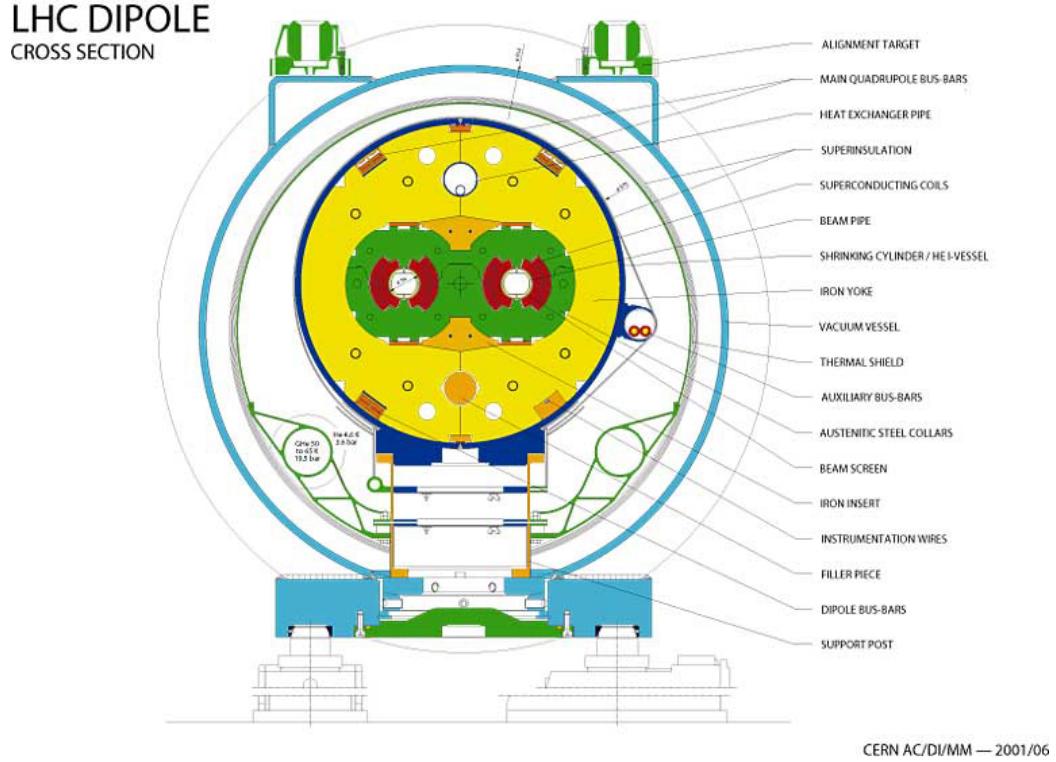
Figure 2.2: Diagram of LHC proton injection chain.

chrotron radiation. The machine luminosity can be written as

$$L = \frac{N_b^2 n_b f_{rev} \gamma_r}{4\pi \epsilon_n \beta_*} F \ , \tag{2.1}$$

where N_b is the number of particles per bunch, n_b the number of bunches per beam, f_{rev} the revolution frequency, γ_r the relativistic gamma factor, ϵ_n the normalized transverse beam emittance, β_* the beta function at the collision point, and F a geometric factor to a small nonzero crossing angle of the beams at the interaction point. The strong luminosity dependence on N_b implies that a proton/antiproton collider concept cannot be used, as at the Tevatron, because of the difficulty in achieving the necessary antiproton beam intensity. Hence, the LHC is chosen to be a proton/proton colliding machine. Each beam then requires a magnetic field pointing in an opposite direction in order to push it around the LHC ring. Due to cost restrictions, and the size of the LEP tunnel a "two-in-one" design was chosen for the LHC, where both proton beams are contained in the same beam pipe, cryostat and field. The primary magnets are then dipoles, with the twin-bore design illustrated in figure 2.3.

Figure 2.3: Illustration of twin-bore LHC dipole magnet. Each of the counterrotating proton beams are contained within the same cryostat.

The full LHC ring is composed of 1,232 dipole magnets, each with NbTi supercon-ductors cooled by superfluid helium to a temperature below 2 K and able to operate at fields above 8 T. They must all have practically identical characteristics in order to ensure successful operation, with variations in the field shape and strength not exceeding approximately 10^{-4} during operation.

Chapter 3

Charged Particles and the CMS Detector

3.1   The CMS Solenoidal Magnet

The excellent momentum resolution when measuring charged particles with the silicon tracker, described in section 3.2, is possible because of the enormous field strength of the CMS superconducting magnet [18–21]. It is designed to reach a 4 T field in the cylindrical free bore of 6 m diameter and 12.5 m length, where the silicon tracker and barrel calorimeters are placed, and has achieved 3.8 T for *in situ* operation. An illustration of the CMS magnetic is shown in figure 3.1, along with a picture of the steel support yoke.
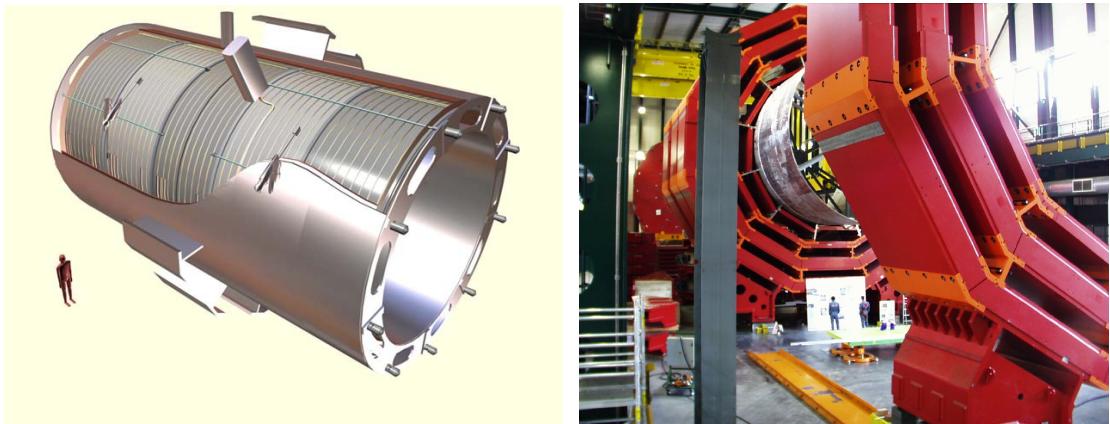
Figure 3.1: (Left) Artist's rendition of magnet cryostat, with view of five models composing the cold mass. (Right) Steel magnet support yokes.

A number of new features, relative to previous magnets for particle physics applications, are introduced in order to achieve the strength and size of the CMS magnet. Due to the number of ampere-turns required to generate the field (4.2×10^7 amp/turn) the winding is composed of four layers of NdTi conductors, as opposed to the usual one. The flux is returned through the 10K t steel yoke, which consists of 5 barrel wheels and two endcaps. Despite the conductor being mechanically reinforced with an aluminum alloy the large ratio between stored energy (2.6 GJ) and cold mass (220 t) causes large mechanical deformations during the energizing of the magnet, of order 0.15%. This stored-energy-to-mass ratio, $E/M = 11.6$, distinguishes the CMS magnet from other detector magnets, as shown in figure 3.2.
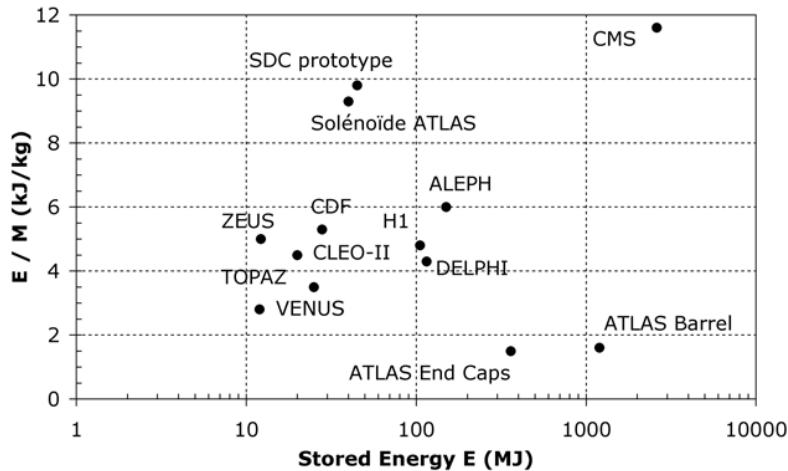
Figure 3.2: (Left) Energy-over-mass ratio E/M for a collection of particle-physics detector magnets. (Right) Steel yoke during early stage of assembly. The 5 barrel wheels support the vacuum chamber of the superconducting coil while one of two endcaps is visible at the back.

The magnitude and field direction of the CMS magnet are illustrated in figure 3.3. The return field is large enough to saturate 1.5 m of iron, meaning that dedicated muon detectors can be placed outside the iron yoke and calorimeters, giving full geometric coverage. The strength and uniformity of the magnetic field in the regions of these muon detectors inform the choice of technology used, as described in section 3.4.
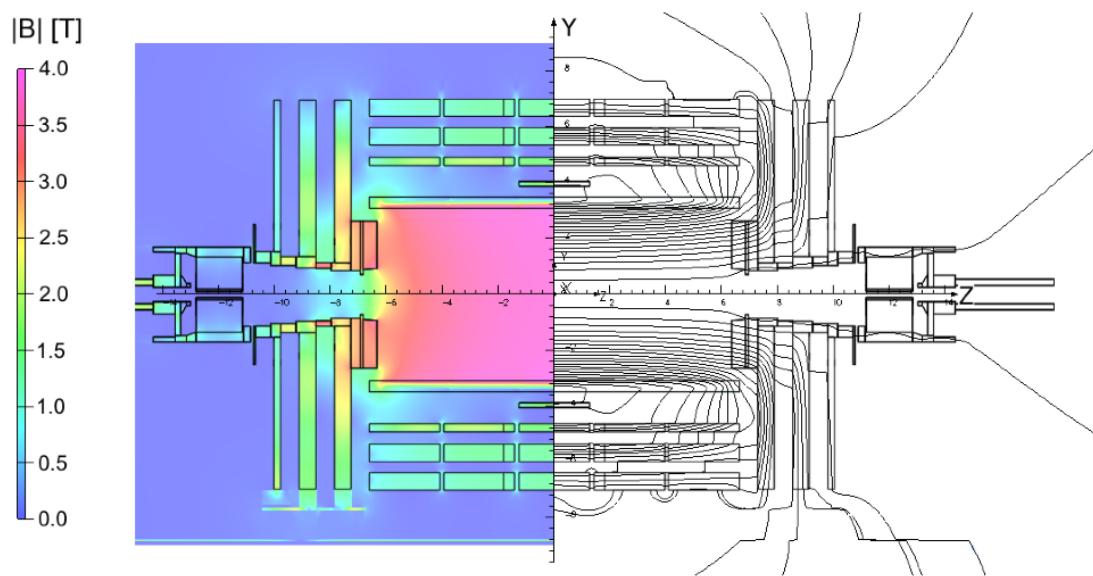
Figure 3.3: Diagram of the magnetic field throughout a longitudinal section of the CMS experiment. (Left) Color z-scale indicates the value of $|B|$. (Right) Field lines in and outside of the iron yoke, with each representing a magnetic flux increment of 6 Wb. [27].

3.2　The CMS Tracker

The design of the CMS tracker is motivated by both the required physics performance and the instantaneous LHC running conditions. Charged particles emerging from collisions must be measured efficiently and precisely as they move through the 4 T magnetic field which covers the full tracker volume. At LHC design luminosity [22] there are expected to be about 1000 particles coming from upwards of 20 overlapping pp interactions, occurring every 25 ns, with the extreme particle flux resulting in a high radiation environment. The tracker must be able to distinguish each of these tracks, measure their trajectory and the interaction primary vertices they come from, and do so quickly in order to correctly identify the bunch crossing while maintaining radiation resistant.

The CMS tracker [16], shown in figure 3.6, represents a balance of these considerations with the corresponding material budget from on-detector electronics and cooling. It consists of two main detectors: a silicon pixel detector, covering the region from 4 to 15 cm in radius (which will have the highest hit-density rate), and 49 cm on either side of the collision point along the LHC beam axis, and a silicon strip detector, covering the region from 25 to 110 cm in radius and within 280 cm on either side of the collision point along the beam axis. With a total of about 200 m^2 of active silicon area the CMS tracker is the largest ever built [23, 24].

3.2.1　Silicon Pixel Detector

The CMS silicon pixel detector includes about 66 million active elements which instrument a surface area of approximately 1 m^2. It is designed to provide at least three high-precision hits for each track. This accomplished through three concentric cylindrical barrel layers at average radii 4.3, 7.3 and 10.2 cm, respectively, and four fan-blade covers for the endcaps, 35.5 and 48.5 cm from the interaction point. The geometry of the pixel detector is illustrated in figure 3.5. The barrel layers have an active length of 53 cm which, along with the endcaps, provides three-hit coverage up to $|\eta| < 2.2$, with two-hit coverage to $|\eta| < 2.5$.
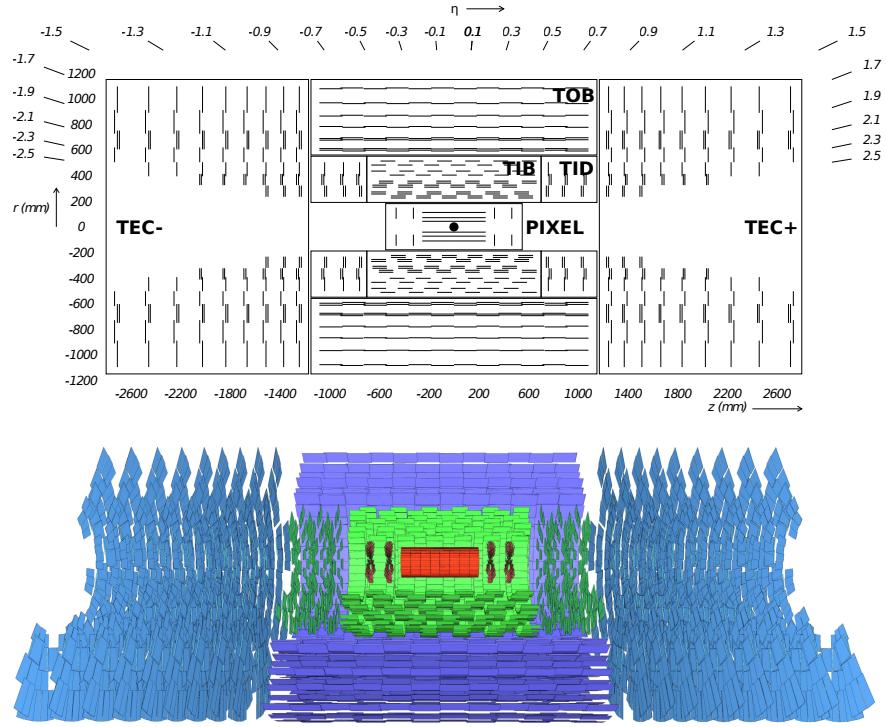
Figure 3.4: *r-z* slice of the CMS Tracker. (Top) Diagram of tracker with strip layers and sub-detectors. Double lines indicate overlapping strips for stereo readout. (Bottom) Three-dimensional visualization of the tracker.

The active elements are n-in-n $100\,\mu\text{m}\times150\,\mu\text{m}$ pixels [16], which achieve a spatial resolution between 15 and 20 μm and with occupancy below 1% with expected particle fluxes. These pixels are oriented with the smaller pitch in the azimuthal direction in the barrel and the radial direction in the disks. The resolution in the azimuthal direction is enhanced by significant Lorentz drift of the collected electrons resulting from the 3.8 T magnetic field, which leads to charge sharing in that direction and therefore improves the resolution. The endcaps benefit from both azimuthal and radial charge sharing through a 20 degree rotation of the disks about their radial axes with respect to the disk planes.
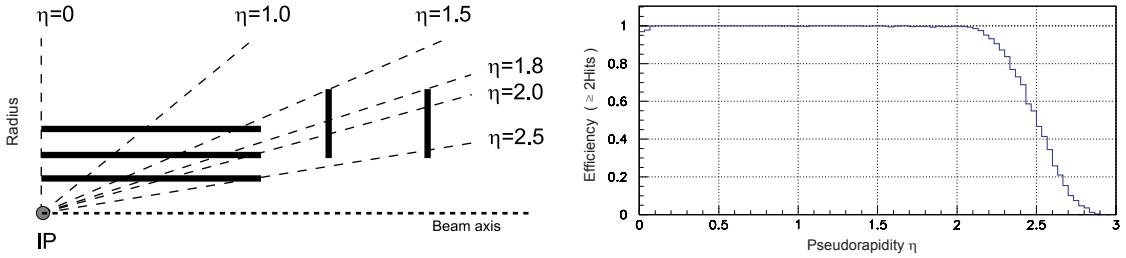
Figure 3.5: (Left) Geometrical layout of the pixel detector. (Right) Pixel hit coverage as a function of pseudorapidity.

3.2.2 Silicon Strip Detector

The CMS silicon strip detector has 9.3 million active elements covering an active surface area of $198\,\mathrm{m}^2$. The detector is composed of three distinct subsystems: The Tracker Inner Barrel and Disks (TIB/TID), the Tracker Outer Barrel (TOB) and the Tracker EndCaps (TEC).

The pitches of each of the tracker layers follows from the expected particle flux, with 10 cm x $80\,\mu$m cells in the TIB layers (20 cm $< r < 55$ cm) to a pitch as large as $183\,\mu$m in the TOB, with an occupancy of around 2%/3% per strip. At larger radii, the strip length must be increased in order to accommodate additional channels. The strip capacitance scales with its length, resulting in the magnitude of electronics noise scaling linearly with increased size. In order to maintain a signal-to-noise ratio well above 10, two different thicknesses of silicon micro-strip sensors are used with widths 320 and $500\,\mu$m, respectively. The thicker strips have correspondingly larger signal and are used at larger radii.

The TIB and TID extend in radius to 55 cm and are composed of four barrel layers, supplemented by three disks at each end, of $320\,\mu$m thick silicon microstrip sensors. This subdetector provides up to four r-ϕ measurements on a track, with the strips oriented parallel to the beam axis in the barrel and radially in the disks. The strip pitch is $80\,\mu$m in the inner pair of TIB layers and $120\,\mu$m in the outer pair of TIB layers, while in the TID, the mean pitch varies between 100 and $141\,\mu$m. This results in single point resolutions ranging between 23 and 35 μm.

Outside the TIB/TID is the TOB, with an outer radius of 116 cm. It consists of six barrel layers of 500 μm thick microstrip sensors with strip pitches of 183 μm in the first four layers and 122 μm in the last pair of layers, extending to ±118 cm in z, which give 53 and 35 μm single point resolutions.

On either side of the beam line from the inner and outer barrel trackers are the two TEC trackers, which cover $124 < |z| < 280$ cm and $22.0 < r < 113.5$ cm. Each one is comprised of nine disks, which are in turn made of up to seven rings of radial-strip silicon detectors. The sensor thicknesses are 320 μm in the inner four rings, increasing to 500 μm in the outer three. The average radial strip pitch varies from 97 to 184 μm.

For the inner two layers of the TIB and TOB, the inner two rings of the TID and TEC and the fifth ring of the TEC a second microstrip detector module is included which is mounted flat to the first with a stereo angle of 100 mrad. These second strips enables a measurement of the orthogonal coordinate (z in the barrel and r on the disks) with a single point resolution of 230 and 530 μm in the TIB and TOB, respectively.

With all the planes running efficiently the silicon tracker provides between 8 to 14 high precision measurements of track impact points up to $|\eta| < 2.4$, not counting stereo modules. The expected number of tracker hits and material budget, as a function of pseudorapidity, are shown in figure 3.6.

3.3   Track and Vertex Reconstruction

3.3.1   Track Reconstruction

Track reconstruction in CMS consists of combining the hits in the various tracking layers, three-dimensional points which indicate the path the took through the detector, into a reconstructed trajectory of the particle. This is accomplished through an iterative procedure in which track seeds are used to grow trajectories by searching for compatible hits, repeating to construct additional tracks.

The track reconstruction algorithm first requires and estimate of the proton-proton
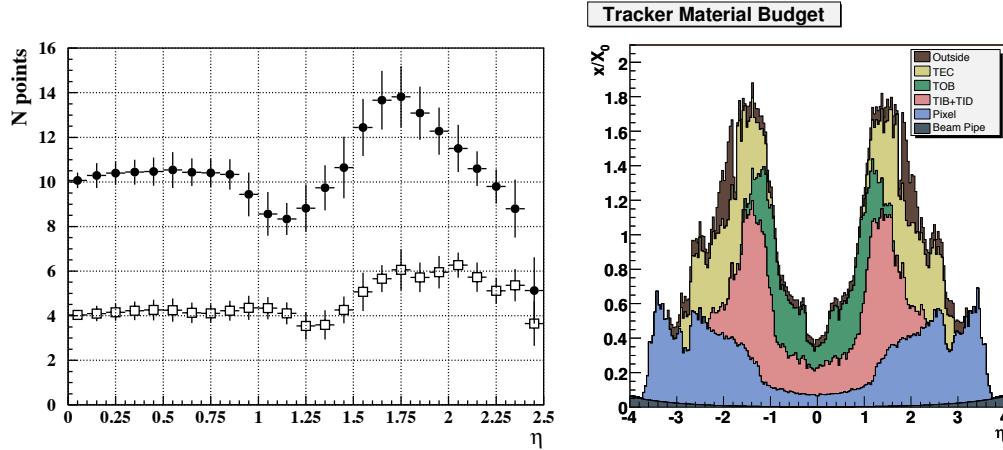
Figure 3.6: (Left) Number of measurement points in the strip tracker as a function of pseudorapidity. Open squares indicate the number of stereo layers while filled circles correspond to all layers. (Right) Material budget of different tracker subdetectors in units of radiation lengths.

interaction region, or beam spot. In the reconstruction algorithm, the transverse location of the beam spot is used as an initial estimate for the primary interaction point. The beam spot is measured over many tracks and events through an iterative χ^2 fit which exploits the correlation between the transverse impact parameter (d_{xy}) and the angle of the track at the point of closest approach (ϕ_0). Fill-to-fill variations of the beam spot are found to be at the level of $\sim$0.5 mm in x and y, and $\sim$2 cm in z [25].

CMS track reconstruction proceeds according to the combinatorial track-finder (CFT) algorithm. In the reconstruction of the tracks of a collision event, an initial round of track and vertex reconstruction is performed using only pixel hits around the beam spot position. The pixel vertices found at this stage are then included among the hits from the strip layers. Next, tracks are seeded from either triplets of hits in the tracker or pairs of hits with an additional constraint from the beam spot or a pixel vertex. These seeds provide an initial estimate of the track's trajectory, with corresponding uncertainty. Each seed is then extrapolated to the other layers of the tracker searching for compatible hits according to the equations of motion of a charged particle in a constant magnetic field, accounting for multiple scattering

and energy loss in the traversed material. As hits are found they are added to the trajectory, the track is refit, and the track parameters and uncertainties are updated. This procedure repeats iteratively until either the boundary of the tracker is reached or no more compatible hits can be found. At this point, an additional search for hits is performed starting from the outermost tracker layer hits and proceeding inwards. Finally, the collection of hits associated with a trajectory is fit to obtain the best estimate of the track parameters. This procedure constitutes one iteration of the CTF algorithm.

In total, six CTF iterations are performed for each event. At the end of each iteration, the reconstructed tracks are filtered to remove likely fakes and to provide a means of quantifying the quality of the remaining tracks. This is accomplished by appealing to the number of hits, the normalized χ^2 of the track, and the compatibility of the track originating from a pixel vertex. Tracks that pass the tightest selection are labelled *High Purity*. Between each iteration, the hits that are unambiguously assigned to the tracks reconstructed and accepted in the previous iteration are removed from the collection of tracker hits, leaving the remaining hits to be used in building additional tracks.

The first two CTF iterations use pixel triplets and pixel pairs as seeds to find prompt (consistent with the beam spot) tracks with $p_T > 0.9$ GeV/c. This is followed by an iteration using only pixel triplet as seeds for low-momentum prompt tracks. In order to identify tracks displaced from the beam spot, the next iteration uses combinations of pixel and strip layers as seeds. Finally, tracks lacking pixel hits are seeded by strip pairs in the final two iterations.

3.3.2 Primary Vertex Reconstruction

The reconstruction of interaction vertices in events begins with the collection of reconstructed tracks. These tracks are grouped according to similar z-coordinate at the point of closest approach to the beam line. The assignment of these groups follows from an adaptive vertex fit, where each of the tracks associated to a vertex

are assigned a weight between 0 and 1 based on their proximity to the commonly determined vertex.

The resolution of the measured primary vertex strongly depends on the number of tracks used in the fit and on the transverse momentum of those tracks. Early LHC collisions at 0.9 and 2.36 TeV were used to measure this resolution [25], with the results shown in figure 3.7. Here, the tracks in each event with only one reconstructed vertex are randomly partitioned into two different set. The difference in the positions of the reconstructed vertex from each set is then interpreted as the resolution multiplied by $\sqrt{2}$ to account for the independent fluctuations of each collection. This resolution is studied as a function of the number and average p_T of the tracks used in the vertex fit, in each CMS coordinate direction.
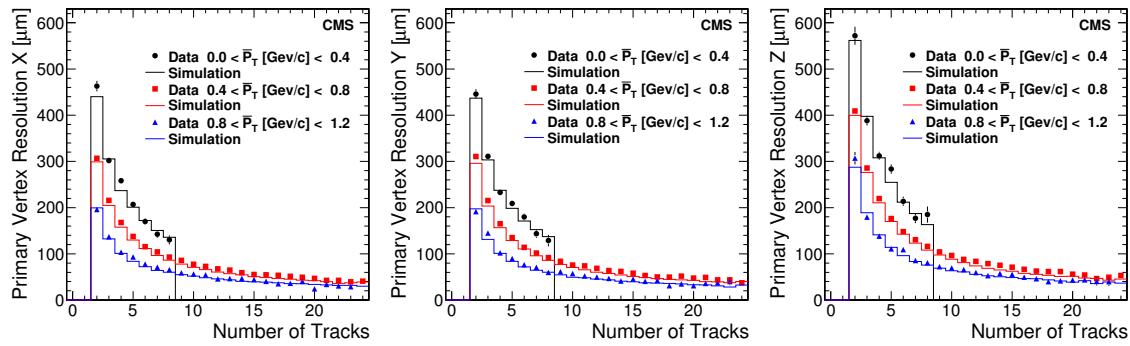
Figure 3.7: Primary vertex resolution distributions in (left) x, (center) y, and (right) z as a function of the number of tracks used in the vertex fit. Data and simulation are compared for different average track-p_T ranges [25].

3.4 The CMS Muon Detectors

As indicated in the acronym CMS, muon detection is of primary importance to the CMS physics program. Since, unlike other charged particles, muons only leave MIP deposits in the calorimeters and travel through the magnetic yoke, they provide a striking signature of interesting processes, particularly in high pile-up conditions at high instantaneous luminosity [17, 26]. The muon detector system is designed to identify muons and measure their trajectories with high precision, over the entire kinematic range of LHC collisions. The performance requirements for the muon system with the solenoidal field at 4 T are listed below [27]. The *in situ* field strength of 3.8 T results in an approximately 5% degradation to these values.

Momentum Resolution

- Standalone muon: 8-15% at 10 GeV/c, 20-40% at 1 TeV

- Global muon (+tracker): 1-1.5% at 10 GeV/c, 6-17% at 1 TeV

Charge Assignment

- Correct to 99% confidence level up to 7 TeV/c

These performance requirements are achieved through three different types of gaseous particle detectors distributed over a cylindrical barrel region and two planar endcaps. The chosen detector solution consists of approximately 25,000 m^2 of reliable, robust and inexpensive muon detector planes.

Each of the muon subdetectors utilizes gas ionization, with chambers of either drift tubes, cathode strip proportional plates or resistive plates. Each chamber is run independently such that they can be used, along with the silicon tracker, as a spectrometer with the CMS solenoidal magnetic field and the flux return providing charged particle bending over the detector volume. The geometry of the muon system follows from the detector and magnetic field-line shapes, with chambers arranged into barrel and endcap components. In the barrel, chambers are arranged into stations at

a fixed radial distance r from the beam line. Similarly, endcap stations are put at fixed distances along the beam direction z from the interaction point. An illustration of the muon system geometry for a longitudinal slice of CMS is shown in figure 3.8. There are four stations in the barrel and in each endcap, labeled MB1–MB4 and ME1–ME4±, respectively. Along the beam line, the barrel stations are divided into 5 wheels while the endcap stations are divided into rings, ME1/n–ME4/n, where n increases with radial distance from the beam axis.
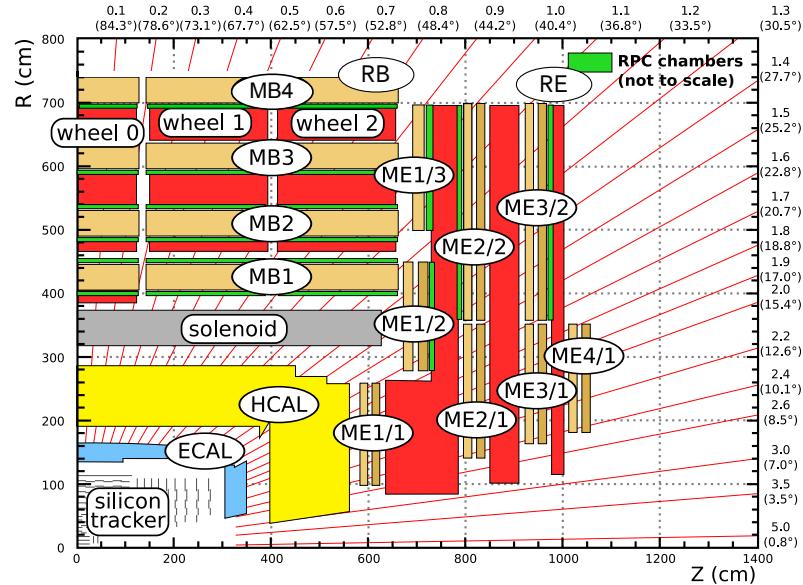
Figure 3.8: r–z cross section of a quadrant of the CMS detector with beam axis (z) running horizontally and radius (r) increasing upward, with the interaction point in the lower left corner. The various muon stations and the steel disks are shown in red.

3.4.1 Muon Drift Tubes

In the barrel region where the magnetic field is mostly uniform with a small strength (≤ 0.4 T) the muon system is composed of drift tube (DT) chambers with rectangular cells and sophisticated electrical field shaping. These DT chambers cover the pseudorapidity region $|\eta| < 1.2$. The four stations and five wheels of DTs are further divided into 12 ϕ-segments per wheel. Each segment contains eight layers of tubes measuring the position in the bending plane (r–ϕ) and four layers for the longitudinal plane (z).

The basic element of the DT system is the drift cell, illustrated in figure 3.9, each with a transverse size of $42 \times 13\,\mathrm{mm}^2$. The cells are filled with a noble gas mixture (85%/15% of Ar/CO2) which has a saturated drift velocity of about 55.5 μm/ns, with a maximum drift time of almost 400 ns. A 50 μm diameter gold-plated stainless-steel anode wire runs through the center of each cell. The wire operates at a voltage of +3600 V, creating an electric field between the wire and the cathode strips at the sides of the cell. Four electrodes are used to shape the effective drift field, operating at −1800 and +1800 V, respectively, on each of the cell sides.
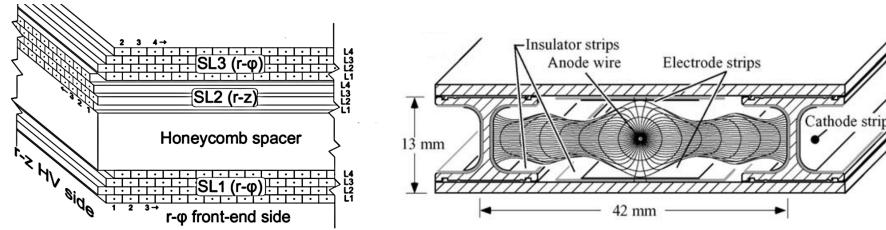
Figure 3.9: (Left) Sliced view of a DT chamber. (Right) Cross-sectional view of a DT cell with drift lines and isochrones. Cathode and anode strips run perpendicular to the viewing direction.

Four layers of parallel cells, staggered with respect to each other to maximize position resolution, form a superlayer (SL). Each chamber consists of two SLs that measure the r–ϕ coordinates using wires parallel to the beam axis, and one orthogonal SL that measures the r–z coordinate (except for the outermost barrel station). Each chamber is about 2.5 m long, with transverse lengths ranging from 1.9 to 4.1 m moving out radially from the beam line.

3.4.2   Muon Cathode Strip Chambers

In the endcap regions of CMS the performance requirements for the muon system are different with respect to the barrel. The muon flux, along with background rates, are high and the magnetic field is strong and nonuniform. CMS uses cathode strip chambers (CSC) in this region. With a short drift path the CSC chambers have a fast response time, reducing the sensitivity to the nonuniform magnetic field. These

CSCs cover the $|\eta|$ region from 0.9 to 2.4 with four stations of chambers at different distances along the beam line from the interaction point, with faces perpendicular to the beam.

Each CSC is made up of six layers, each of which provides a 2D measurement of the muon trajectory, as illustrated in figure 3.10 (left). Cathode strips run radially outward through the CSCs and provide a measurement in the r–ϕ bending plane while perpendicular wires provide a coarse measurement of the radial distance. Each of the CSCs operates as a standard multiwire proportional counter (MWPC), with the additional feature of a cathode-strip readout which can precisely measures the position at which a muon or other charged particle crosses the gas volume [28], as indicated by the illustration of the gas ionization avalanche profile shown in figure 3.10 (right).
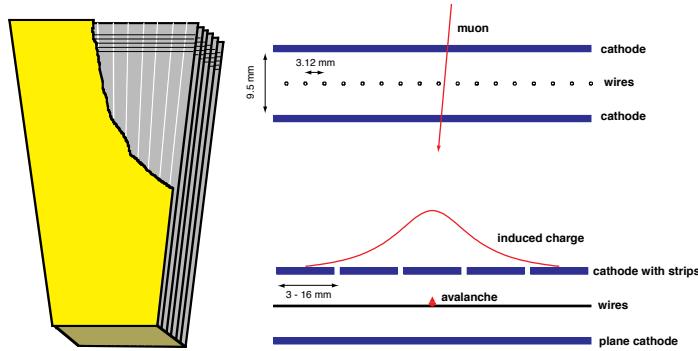
Figure 3.10: (Left) Diagram of a CSC. Each is made of six layers with the orientations of the wires and strips illustrated by a few examples. (Right) Cross-sectional views of the gas gap in a CSC with the anode wires and cathode planes running parallel. The gas ionization avalanche and resulting induced charge distribution on the cathode strips is illustrated.

Several different sizes of CSCs are used, ranging in length from about 1.7 to 3.4 m in the radial dimension. All chambers are filled with a gas mixture of 50% CO_2, 40% Ar, and 10% CF_4. Through the gas runs 80 cathode strips projected towards the beam line and anode wires with a diameter of 50 μm. These anode wires are grouped according to 5 to 16 wires, with widths from 16 to 51 mm, which limits the position resolution in the wire coordinate direction. The ME1/1 chambers are operated at an anode voltage of 2.9 kV with the others at 3.6 kV.

3.4.3 Muon Resistive Plate Chambers

.

In addition to the DT and CSC muon detectors, there is also a dedicated triggering detector system with excellent time resolution made of resistive plate chambers (RPC). These detectors are located in both the barrel and endcap regions where they can provide a fast, independent trigger with a looser p_T threshold (relative to the other detectors) over a large pseudorapidity range ($|\eta| < 1.6$).

Each of the RPCs is a double-gap chambers which are operated in avalanche mode to allow for high rates. An illustration of the RPC geometry is shown in figure 3.11. Each of the 2 mm thick gas gaps is filled with a mixture consisting of 95.2% freon, 4.5% isobutane and 0.3% sulphur hexafluoride, surrounded by two 2 mm thick resistive bakelite plates. The plates are coated with a thin conductive graphite layer, with a voltage of about 9.6 kV applied. The readout strips are aligned along fixed η in between the 2 gas gaps. When a charged particle crosses an RPC the gas will become ionized in both gap volumes and the avalanches generated by the large electric field over the gaps induce an image charge which is detected by the readout strips.

The RPCs are grouped in stations like the DTs and CSCs, with four in the barrel and three in the endcap. The innermost barrel stations have two RPC layers along the outside of the DT chambers, with each layer divided into 2 or 3 η partitions called rolls (figure 3.11). The RPC endcaps stations are divided into three rings with increasing radial distance from the beam line, with 36 chambers in each ring covering the full azimuthal range.

3.5 Muon Reconstruction

Muons are identified and their momenta measured in CMS using the combination of the muon and inner tracker detectors. The muon reconstruction in a collision event begins by first identifying hits in the detection layers of the muon DT and CSC sys-
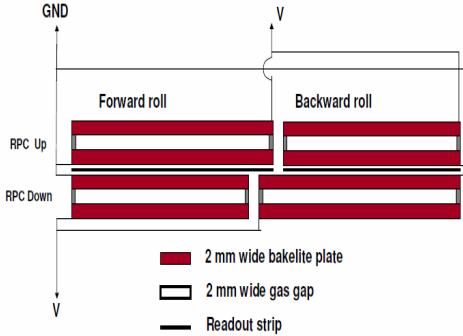
Figure 3.11: Illustration of a generic barrel RPC with two roll partitions.

tems. During this "local" reconstruction phase straight-line track segments from these hits seed muon candidates. The reconstructed muon tracks from the hits associated with these seeds are created in the subsequent "global" reconstruction, where information from the inner tracker can also be used. Muon tracks reconstructed using hits from only the muon detectors alone are called "standalone muons," while those which combine information from the central tracker and muon chambers are called "global muons." The muon system can also be used simply to tag extrapolated tracks coming from the central tracker measurements; these tracks are denoted "tracker muons." For muons with momenta below $\sim$200 GeV/c, tracker muons have better resolution than global muons, with the contribution from the later contributing at higher transverse momentum.

The direction of the magnetic field changes as the muons pass from the solenoidal barrel to the return yoke, causing a reversal of the curvature in the muon's trajectory. This means that the measurements from the innermost muon stations in the barrel and endcap are crucial for muons with transverse momenta up to a few hundred GeV since they provide the largest sagitta. For higher momenta muons the importance of the outer stations increases as multiple scattering effects become less important. The combined DT and CSC muon detector elements cover the full pseudorapidity interval $|\eta| < 2.4$ with no acceptance gaps, ensuring good muon identification over a the entire range. Offline reconstruction efficiency for the muons is typically 96-99% except in
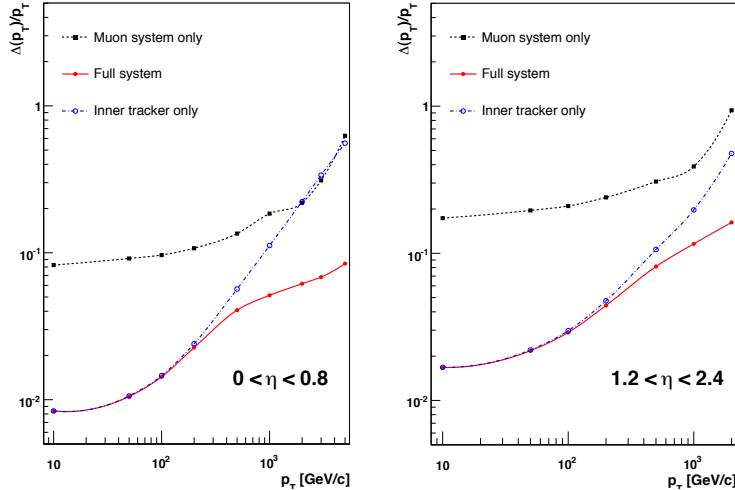
Figure 3.12: Reconstructed muon transverse momentum resolution as a function of transverse momentum using the muon detectors only, the inner tracking detectors only or both. Results are provided for muons with $|\eta| < 0.8$ (Left) and $1.2 < |\eta| < 2.4$ (Right) [16].

gaps between the 5 wheels of the yoke (at $|\eta| = 0.25$ and 0.8) and the transition region between the barrel outer wheels and the endcap disks [29]. The amount of absorbing material between the interaction point and the first muon station reduces the contribution of punch-through to about 5% of all muons reaching the first station, and to about 0.2% of all muons reaching further muon stations.

The combination of the muon and tracker measurements yields excellent muon resolution, as shown in figure 3.12, with complementary features that ensure continued performance for the entire range of muon momenta in LHC collisions.

3.6 Identification of b-quark jets

Jets that arise from bottom-quark hadronization and decay (*b*-quark jets) are characteristic of the final states for a wide range of interesting physics processes, like the decay of top quarks and various SUSY particles. These events can be selected from among otherwise large backgrounds with jets from gluons and light quarks by tagging, or identifying, these jets based on their distinguishing properties. Bottom quarks have hard fragmentation functions and the relatively large mass, and a long lifetime of the heavy flavor hadrons. The CMS tracking system is well suited to identify secondary decay vertices coming from long-lived hadrons and use their properties to identify *b*-tagged jets.

The *b*-tagging algorithm used in the search for evidence of SUSY in chapter 10 begins with the collection of reconstructed jets in the event. The creation and identification of these jets is described in detail in section 5.2. For each jet, the collection of reconstructed tracks is queried for tracks falling in a cone $\Delta R < 0.5$ around the jet axis, with a maximal distance to the axis of 0.2 cm. These tracks must then satisfy several additional requirements in order to be considered for the *b*-tagging algorithm: each must be a high purity track (see section 3.3.1, have a p_T of at least 1 GeV/c, a fit χ^2/ndof < 5 and have transverse and longitudinal impact parameters (IP), d_{xy} and d_z, smaller than 0.2 and 17 cm, respectively.

It is these track impact parameters with respect to the primary vertex which is used to distinguish decay products of a *b*-hadron from prompt tracks. The IP is calculated in three dimensions, relying particularly on the excellent resolution of the pixel detector along the *z*-axis. Each tracks' impact parameters are given a sign according to the scalar product of the vector pointing from the primary vertex to the point of closest approach with the jet direction. This means that tracks originating from the decay of particles traveling along the jet axis will tend to have positive IP values while the impact parameters of other prompt tracks will have both positive and negative signs. The resolution on the impact parameter depends strongly on p_T and η of the track. In order to account for this dependency, the impact parameter

significance S_{IP}, defined as the ratio between the IP and its estimated uncertainty, is used as a discriminating observable. The distribution of IP values and significances for selected tracks associated to jets in 2011 CMS running are shown in figure 3.13, indicating good agreement with expectations from the CMS full simulation. The IP significance has discriminating power between the decay products of b and non-b jets, following from displaced decay products of b-mesons.
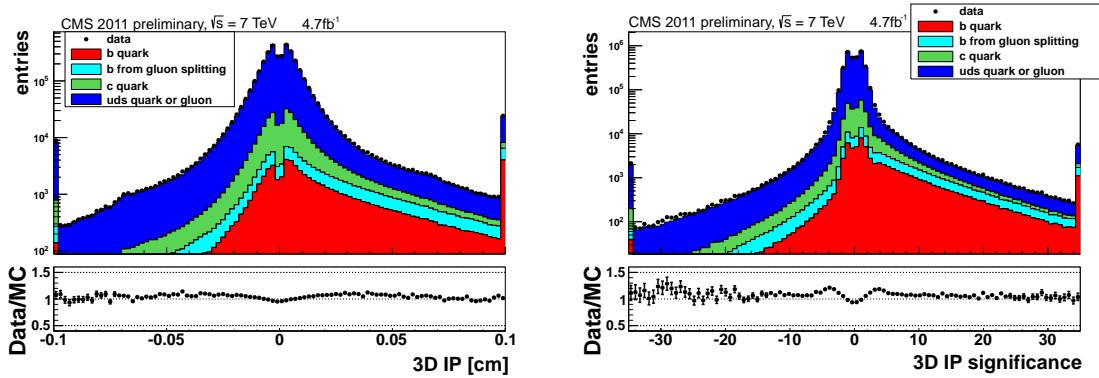
Figure 3.13: The impact parameter (Left) and corresponding significance (Right) for selected tracks associated with jets [30].

The IP significances of tracks associated to a single jet are combined to form a b-tagging discriminant using the *Track Counting* (TC) algorithm. Here, the tracks associated to a jet are sorted according to decreasing values of the IP significance. This ranking requirement biases the IP significance for the first track to higher values, while values of the following jets provide a largely unbiased indicator of the displacement of the tracks' vertices since probability to have several tracks with high positive values is low for light flavor jets. The TC algorithm uses the IP significance of the second and third ranked tracks to calculate a discriminator value, with two different versions tuned to yield either high efficiency (TCHE) or high purity (TCHP). The analyses described in this thesis use the TCHE algorithm with medium (TCHEM) working point, corresponding to a cut of 3.3 on the discriminator whose distribution is shown in figure 3.14 (Reft).

Even this simple tagging algorithms depends on high tracking efficiency and a reliable estimation of track parameters and their uncertainties, which makes it poten-
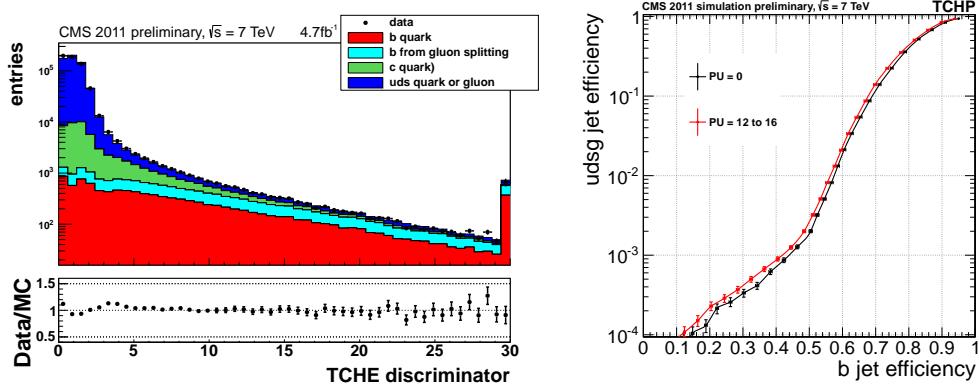
Figure 3.14: (Left) Discriminator values for the TCHE algorithm for data and simulated events. (Right) Light-flavor mistagging rate vs. b-tagging efficiency for different pile-up scenarios [30].

tially sensitive to changes in the running conditions of the experiment. The impact of high pile-up on the b-tagging performance is evaluated in figure 3.14 (right), which shows the rate for mistagging light-flavor jets as a function of b-tagging efficiency for the TCHE discriminant evaluated on jets reconstructed in collision events. We observe that the b-tagging performance is largely insensitive to the number of interaction vertices in these events.

Chapter 4

Electrons, Photons, and the CMS ECAL

4.1 The CMS Electromagnetic Calorimeter

The CMS electromagnetic calorimeter (ECAL) is designed to provide a fast response with excellent energy resolution to electrons and photons incident on its face. Furthermore, the device must be radiation tolerant, maintaing performance in the high particle-flux LHC environment. The ECAL is a homogeneous crystal calorimeter made of 75,848 lead tungstate ($PbWO_4$) crystals. The detector consists of a barrel region (EB), covering up to pseudorapidity $|\eta| = 1.48$, and two endcaps (EE), that extend the coverage up to $|\eta| = 3.0$. A silicon/lead pre-shower detector (ES) is installed in front of the crystal calorimeter in the endcaps in order to improve the γ/π^0 discrimination and the vertex reconstruction for photons, covering a pseudorapidity region $1.65 < \eta < 2.6$. An illustration of the CMS ECAL is shown in figure 4.1 [31,32].

When electrons and photons pass through the ECAL crystals they lose energy through interaction with the Coulomb fields of the crystal matter constituents, bringing the crystal medium into an excited state. In the quick return to the ground state, blue scintillation light is released in the resulting electromagnetic shower. This light is detected by avalanche photodiodes (APDs) in the barrel region [33, 34] and by vacuum phototriodes (VPTs) in the endcaps [35].
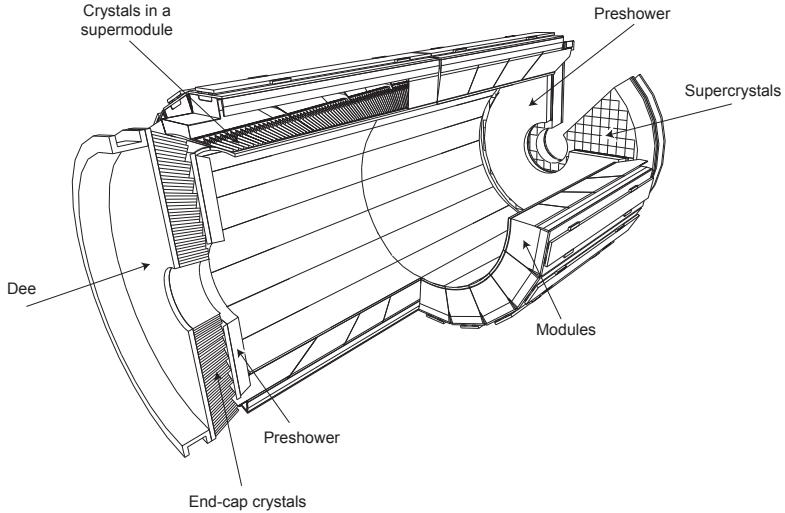
Figure 4.1: Illustration of the CMS ECAL showing the arrangement of crystal modules, supermodules and endcaps, with preshowers in front.

4.1.1 ECAL PbWO$_4$ Crystals

The properties of the ECAL PbWO$_4$ make them ideal for the CMS detector [36]. Their high density (8.28 g/cm^3) and short radiation length (0.89 cm), with correspondingly small Moliere radius (2.2 cm) mean that the calorimeter can be compact, and with fine granularity the direction and shower shape of incident particles can be measured accurately [37,38]. Example EB and EE crystals are shown in figure 4.2.

The scintillation mechanism in the crystals is fast, such that the decay time is on the order of the LHC designed inter-bunch-crossing time, with 80% of light emitted over 25 ns. This light is blue-green with a broad maximum at approximately 425 nm [38,39]. From this scintillation light about 4.5 photoelectron per MeV are collected in the APDs and VPTs. The crystals are polished after machining in order to maximize internal reflection and hence light collection.

4.1.2 ECAL Crystal Geometry

The EB is composed of 36 supermodules which, in turn, consist of 1,700 tapered crystals with a frontal area of approximately 2.2×2.2 cm^2 and a length of 23 cm, corresponding to 25.8 radiation lengths. This granularity is approximately the same
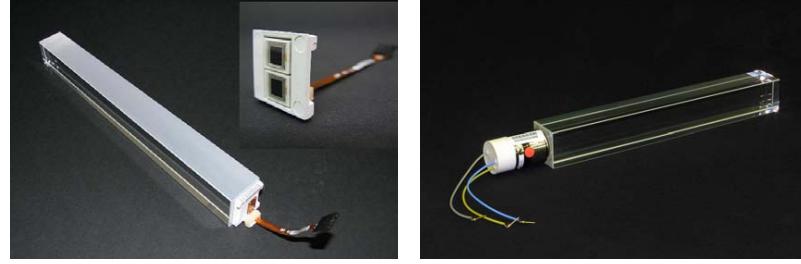
Figure 4.2: ECAL PbWO$_4$ crystals with attached photodetectors. (Left) EB crystal with APD. Two APDs in insert. (Right) EE crystal with VPT.

as the Moliere radius of the ECAL crystals. The crystals are mounted in a quasi-projective geometry, relative to the interaction point, to avoid gaps in the geometric coverage from intercrystal cracks. Each crystal covers $\Delta\eta \times \Delta\phi = 0.174 \times 0.174$ in projective-space.

The crystals individually wrapped in a 0.1 mm thick alveolar structure made with an aluminum layer facing the crystal and two layers of glass fiber-epoxy resin. The crystals are arranged into modules, each containing 400 or 500 crystals depending on η, such that the nominal crystal-to-crystal distance is 0.35 mm, with 0.5 mm between the crystals of each module. A supermodule is made of four modules separated by 4 mm thick aluminum webs, illustrated in figure 4.3.
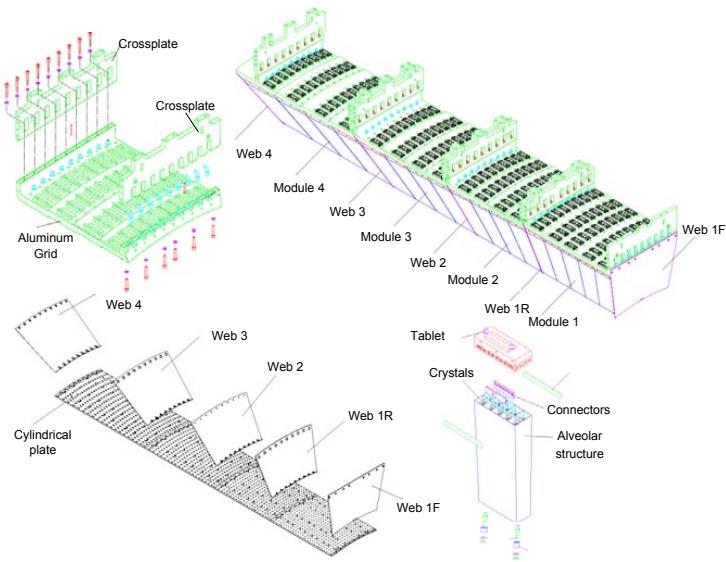
Figure 4.3: Layout of ECAL barrels crystals and modules into one supermodule.

The two ECAL EEs are constructed from four half-disk "dees," each composed of 3662 tapered crystals. Each EE crystal has a frontal area of 2.86×2.86 cm^2 and a length of 22 cm, corresponding to 24.7 radiation lengths. As for the barrel, EE crystals are arranged in a quasi-projective geometry, focussed at a point 1.3 m farther than the nominal interaction point along the beam line, with off-pointing angles between 2° and 8°. The crystals are grouped into 5×5 supercrystals, with 138 in each dee. They are arranged in a rectangular x-y grid, with 18 partial supercrystals on the inner (around the beam line) and outer circumference.

In front of each of the ECAL endcaps is the preshower ES, consisting of two orthogonal planes of silicon strip sensors interleaved with two planes of lead absorbers (2 and 1 X_0 respectively). The sensors have an active area of 61×61 mm^2, divided into 32 strips. They are grouped into "ladders" of 7, 8, or 10 sensors.

4.1.3 ECAL Energy Resolution

The ECAL barrel energy resolution for electrons is measured in test-beams (see section 4.2 for a description) to be [40]:

$$\frac{\sigma_E}{E} = \frac{2.8\%}{\sqrt{E(\text{GeV})}} \oplus \frac{12\%}{E(\text{GeV})} \oplus 0.3\%, \tag{4.1}$$

where the three contributions correspond to the stochastic, noise and constant terms, respectively. This is measured through the reconstruction of electrons from a 4×4 mm^2 collimated beam incident in the center a a single crystal, to minimize shower leakage effects. The energy is reconstructed from the 3×3 surrounding crystals. For the *in situ* environment, crystal intercalibration [41] and transparency monitoring, the CMS magnetic field and pile-up energy contributions must all be controlled to maintain a resolution of 0.5% for 100 GeV particles. In practice, the ultimate ECAL energy resolution in the *in situ* running environment depends on the material budget in front of the ECAL detectors and calibration of electrons and photons to account for inter-crystal energy leakage, radiation-induced crystal transparency changes (section 4.1.4) and the recovery of Bremsstrahlung radiation. The energy resolution of

electrons from Z decays is measured to be better than 2% in the central region of the ECAL barrel and 3% to 4% elsewhere. Similarly, the energy resolution of reconstructed photons following from 125 GeV Higgs boson decays varies between 1.1% and 2.5% in the ECAL barrel and between 2.2% and 5% in the endcaps [42].

4.1.4 ECAL Crystal Laser Monitoring System

.

The ECAL crystals are radiation resistant in that their scintillation mechanism is not altered through electromagnetic irradiation [43]. On the other hand, the crystals show a rapid loss of optical transmission under irradiation due to the production of color centers. While this damage will self-anneal, the transient color centers reduce the transparency of the crystals and absorb a fraction of the transmitted light, effectively reducing the response of the crystals to electromagnetic showers. This effect results in a dose-rate dependent oscillation of the crystals' transparency and energy response which follows the LHC collision fill-cycle, as shown in figure 4.4. In order to maintain the ECAL resolution these transparency fluctuations are monitored and corrected for using the ECAL laser monitoring system [44, 45].

Crystals' transparencies are monitored by injecting laser pulses through optical fibers directly into the crystals. Their crystals' APD response is normalized by the signal from silicon PN photodiodes, which receives the same laser light as the crystals. The ratio of an APD response to that of the PN, $R(t) = APD(t)/PN(t)$, indicates the relative response as a function of time. A blue laser ($\lambda = 440$ nm) is used to monitor crystal transparency, chosen to correspond to the PbWO$_4$ scintillation peak. Despite this attempt to match the light spectra, the different optical paths of scintillation and laser light through the crystal, respectively, means that the relationship between the diminished energy response of the crystals to the laser light and to incident electromagnetic particles is not linear. For small attenuations ($R(t) < 10\%$) the
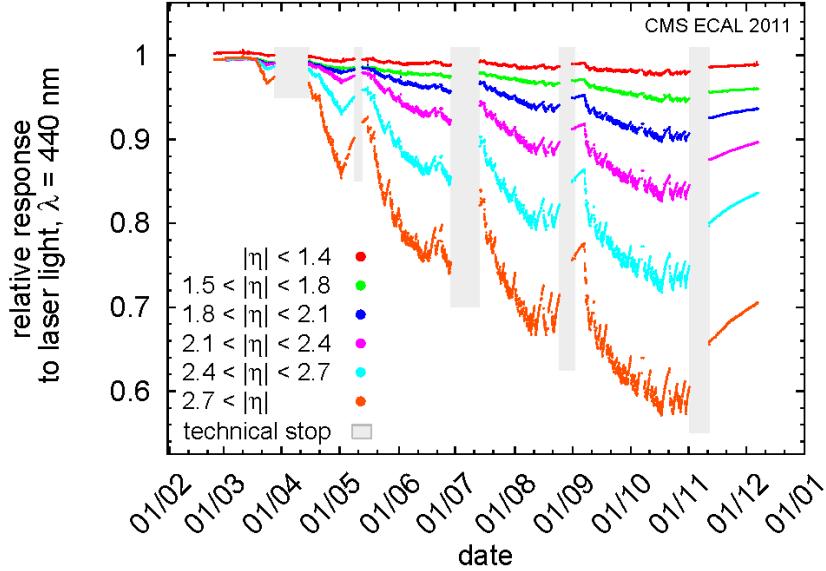
Figure 4.4: Relative crystal energy response as a function of time, as measured by the laser monitoring system. The response is averaged over the η ranges listed in the legend. Long periods without colliding beams are shaded.

relation between the response to scintillation and laser light can be modeled as

$$\frac{S(t)}{S(t_0)} = \left(\frac{R(t)}{R(t_0)} \right)^{\alpha} \, , \tag{4.2}$$

where $S(t)$ represents the scintillation light response and α is an effective parameter characteristic of each crystal which depends on the production method ($\alpha \sim 1.5$ for BCTP crystals and $\alpha \sim 1$ for SIC). Studies of the dynamics of these transparency changes and measurements of α were performed in dedicated test beams are described in section 4.2.

In addition to the blue laser a second wavelength in the infrared ($\lambda = 796$ nm), far from the scintillation emission peak, is also used to monitor the crystals. At this wavelength, the crystal response is little affected by the transparency changes and is used to monitor the stability of the system. In total there are three light sources, 2 blue and 1 near infrared, with duplication of the former to provide fault tolerance during *in situ* monitoring. Each source includes an Nd:YLF pump laser, a Ti:Sapphire laser and associated cooling an control electronics. The full pulse energy

is 1 mJ at the blue wavelength, corresponding to a 1.3 TeV particle in the ECAL, with a linear attenuator allowing for 1% steps down to 13 GeV. PN diode measurements monitor the intensity of the laser pulses to at a precision of 0.1%. The light pulses are distributed to the crystals and PN diodes via a system of optical fibers, illuminating one of 88 calorimeter regions at a time, with optical fiber fan-outs transmitting light to each crystal individually. Each crystals' response is measured once every 20 minutes during the LHC abort gap in between collision-filled bunch crossings.

4.2 Test Beam Studies of ECAL Crystal Transparency Changes

The lead tungstate crystals used in the CMS ECAL are radiation hard to high integrated doses but experience a dose-rate dependent transparency change, as described in section 4.1.4. During LHC accelerator operation this dose-rate for ECAL crystals will vary significantly, depending on the instantaneous luminosity and the location of the crystal in the ECAL. At design luminosity a typical dose-rate for the ECAL barrel crystals will be 15 rad/h and up to a factor of 100 higher for the crystals closest to the beam pipe. The transparency loss will largely recover when the irradiation stops, typically in a fast initial recovery, with a time constant on the order of tens of hours, followed by a slower recovery on the time-scale of hundreds to a few thousand hours. These transparency changes are monitored continuously by the laser monitoring system. The relationship between the crystals' time dependent response to laser light (R) and its response to scintillation light (S) is well modeled by [46]

$$\left(\frac{S}{S_0}\right) = \left(\frac{R}{R_0}\right)^{\alpha} \ , \tag{4.3}$$

in the regime where $\Delta S = S_0 - S$ is small ($\leq 10\%$). This relationship can be used to correct crystals' energy response for transparency changes, using a parameter α characteristic to each crystal.

In 2006 and 2007 test beams irradiations of a collection of ECAL crystals, installed in fully functioning ECAL supermodules, were performed in order to study the dynamics of crystal transparency change and understand the the intrinsic variations of crystals' α parameters. In these test beams, components of the CMS ECAL were placed in front of monoenergetic electron beams.

4.2.1 Test Beam Setup

During the 2006 ECAL test beam campaign nine ECAL barrel supermodules (SM) were intercalibrated and studied in the H4 beam line at the CERN SPS. The SMs were

mounted on a movable, computer operated table, allowing the electron beam to be directed at all crystals with the same quasi-projective geometry as the CMS detector in LHC running, relative to the nominal interaction vertex. As a result, the test beam setup differs from LHC operation only in the absence of the solenoidal magnetic field and material between the interaction point and ECAL barrel (beam line and tracker). The SMs were installed with final versions of readout electronics, high and low voltage systems, cooling system, temperature monitoring and laser monitoring.

Readout of the SM crystals was triggered using plastic scintillator tiles with a 20×20 mm^2 area, slightly smaller than the front faces of the crystals ($\approx 22 \times 22$ mm^2). A Time to Digital Converter (TDC) is used to measure the phase between the triggers given by these tiles and the ADC clock. The transverse position of the electron beam was measured using four layers of scintillating fibers hodoscopes. An impact point resolution of 250 μm is achieved [47] in both x and y (corresponding to η and ϕ, respectively, if the SM were installed in CMS). During the 2006 test beam, five different ECAL crystals were irradiated.

This same setup was used for the 2007 ECAL endcap test beam, with the exception of an improved movable table for the detector which allowed for more accurate positioning in front of the beam. In this test beam, 20 EE supercrystals were mounted in a 4×5 rectangular grid. The precision table allowed for shooting the beam at the corner between four crystals, irradiating all of them simultaneously and increasing the total number irradiated crystals by almost a factor of 10, relative to 2006.

4.2.2 Estimation of Irradiation Dose Rates

For 2006 test beam irradiations the beam was aligned such that its center was incident on the center of the face of the crystal being studied. Each crystal was irradiated for approximately ten hours with a continuous beam of either 120 GeV/c or 90 GeV/c momentum electrons, for crystals in supermodules 22 and 9, respectively. In order to permit comparison to earlier irradiation studies (and to better understand the radiation hardness of the crystals for *in situ* running) we estimate the dose on each

crystal during irradiation. This is done by modeling both the electron beam profile in the plane transverse to the crystal face and the electromagnetic shower profile within the crystal.

To model the shower profile we note that, on average, only 10% of the energy lies outside the cylinder with radius $R_M = 2.19$ cm, the Moliere radius characteristic of the crystals. These distributions are often described as the sum of two Gaussians, but for convenience we will estimate it as a single Gaussian. Assuming the transverse profile shape remains constant, the restriction that 90% of the energy lies within a radius R_M dictates that the transverse profile has $\sigma_T = 1.02$ cm.

In the longitudinal direction, we assume that the energy deposition is described by a gamma distribution [48]:

$$\frac{dE}{dt} = E_0 b \frac{(bt)^{a-1} e^{-bt}}{\Gamma(a)} \ , \tag{4.4}$$

where $t = x/X_0$ is a scale variable in units of radiation lengths and E_0 is the incident electron energy. The variables, a and b are set by noting that the longitudinal shower maximum, t_{max}, can be estimated as

$$t_{max} = (a-1)/b = \log(E_0/E_c) - 0.5 \ , \tag{4.5}$$

where E_c is the critical energy, defined as the energy at which the rates of loss from ionization and Bremsstrahlung are equal. For lead tungstate crystals, we estimate E_c = 10.8 MeV by noting that $R_M = X_0 E_s/E_c$ and assuming that $|dE/dt|_{brems} \approx E/X_0$. Here, E_s is the scale energy $\sqrt{4\pi/\alpha}\, m_e c^2 = 21.2$ MeV. For the parameter b we assume $b \approx 0.5$, a reasonable choice a range of materials [48].

With these assumptions, we estimate the dose at shower max (approximately 8.8 X_0 and 8.5 X_0 for 120 GeV/c and 90 GeV/c momentum electrons, respectively) as

$$Dose_{max}(e, 120\ GeV/c) = 3.7 \times 10^{-8} \times N_e (Gy) \ , \tag{4.6}$$

$$Dose_{max}(e, 90\ GeV/c) = 2.9 \times 10^{-8} \times N_e (Gy) \ , \tag{4.7}$$

where N_e is the number of incident electrons.

In order to model the the beam intensity, we use trigger counters placed in the beam line which record the number of particles crossing a 50×50 mm^2 area centered on the crystal under study. We approximate the beam profile as normally distributed in x and y and under that assumption measure for the 120 GeV/c momentum electron beam $(\sigma_x, \sigma_y) = (0.6$ cm, 1.1 cm$)$ and, for 90 GeV/c beam, $(\sigma_x, \sigma_y) = (1.0$ cm, 1.9 cm$)$.

With models for the beam and shower profiles we can calculate the average dose over the entire crystal by convoluting the two profiles, giving an average dose over the crystal volume as
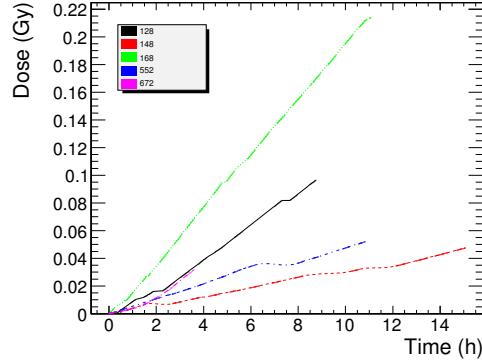
$$< Dose > (e, 120 \ GeV/c) = 7.0 \times 10^{-9} \times N (Gy) \qquad (4.8)$$

$$< Dose > (e, 90 \ GeV/c) = 4.2 \times 10^{-9} \times N (Gy) \ , \qquad (4.9)$$

where N is the number of hits recorded by the trigger counters. Using this estimate, we find that the average dose rates ranged from 0.005 Gy/h to 0.02 Gy/h, with the values for the five irradiated crystals from the 2006 test beam summarized in figure 4.5 (bottom). The integrated dose as a function of time for these irradiations is shown in figure 4.5 (top).

4.2.3   Treatment of Crystal Irradiation Data

During irradiation, electron-beam and laser data were recorded alternatively, with a period of roughly 10 minutes. An example irradiation, for crystal 168, supermodule 22, is shown in Figure 4.6. Here, each point shows the normalized electron and laser responses averaged over many events taken in the interval. The electron points correspond to a Gaussian fit to the 120 GeV electron energy distribution, with error bars shown. The laser points are calculated by fitting the distribution of 600 APD/PN values taken for each laser run with a Gaussian function. For each electron event, a corresponding laser response is calculated by linearly interpolating between the laser points that proceed and follow the electron run.

Crystal ID	e energy	Mean dose rate (Gy/h)	Int. mean dose (Gy)
128	90	0.0110	0.0964
148	90	0.0031	0.0475
168	120	0.0194	0.2141
552	90	0.0048	0.0511
672	90	0.0089	0.0317

Figure 4.5: (Top) Mean dose versus time for the various crystals under electron irradiation. (Bottom) Dose rates and integrated doses for the irradiated crystals.

The crystal energies are calculated from the 10 ADC samples read from the APDs of a crystal every 25 ns, as described in section 4.3, and used to calculate the pulse amplitude. This is converted into GeV, and an intercalibration constant is applied, derived using the S1 algorithm [49].

For the laser light energy reconstruction in the ECAL crystals, the difference in pulse shape between the laser and scintillation light means that the weights applied to the 10 ADC samples to calculate the pulse amplitude are no longer appropriate. Instead, the pedestals are defined for each crystal's channel on an event by event basis using the first 3 digitized samples, occurring before the laser signal in the 10 sample window. The maximum amplitude is obtained using a fit function [50]

$$A(t) = A_0 \left(\frac{t - t_0}{\beta} \right) e^{-\alpha \left(\frac{t - t_0 + \beta}{\beta} \right)} , \tag{4.10}$$

where β is the electronics decay time and the product $\alpha\beta$ is the electronics rise time. The response of the reference PN photodiodes to laser light is much slower, with a shaping time of about 750 ns, digitized in 50 samples at 40 MHz. Their maximum
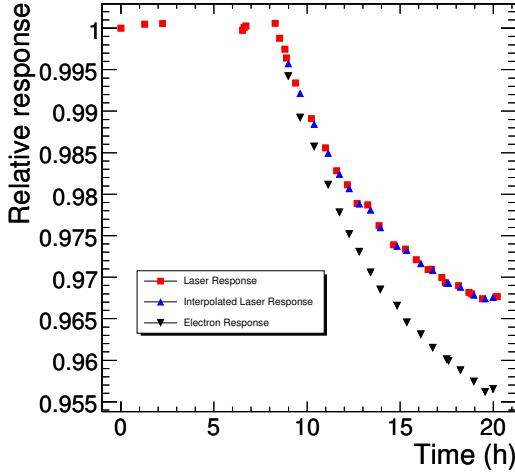
Figure 4.6: Normalized electron, laser and interpolated laser responses for a single crystal.

response is calculated with a 2nd-degree polynomial fit using the 16 samples around the maximum sample. The mean of the first five samples is subtracted as a pedestal.

The laser response for each crystal channel is given by the reconstructed APD amplitude divided by the PN amplitude corresponding to that channel, denoted APN/PN. Each laser run consists of 600 events for each channel, with the APD/PN value for the run calculated using an iterative Gaussian fit to these events' values. Subsequently, for each irradiated crystal, a reference channel was chosen in the same 5×5 trigger tower away from the electron beam. The irradiated crystal's APD/PN value for each laser run is then divided by the APD/PN value from the reference crystal in order to correct for variations in the laser pulse width.

4.2.3.1   Correction for Impact Point

The energy deposition of an electron in a single crystal depends, among other things, on its position of incidence on the face of the crystal. Electrons which hit the crystal near an intercrystal boundary will experience larger variations in response due to shower leakage, degrading the response resolution. A correction is developed by measuring the crystal response as a function of hodoscope measurement, effectively

parameterizing the response as a function of incidence position. The correction is calculated using the data itself, using the approach described in [49] with slight modifications briefly described here. The position dependence of the crystal response can be factorized as the product of two 4$^{\text{th}}$-order polynomials, $P_X(x)$ and $P_Y(y)$, each a function of one of the hodoscope coordinates

$$F(x, y) = P_X(x) \times P_Y(y) \ . \tag{4.11}$$

Using events selected from a 40 mm × 40 mm region in the (x,y) plane the mean response of the crystal vs. impact point is fit in the 20 mm × 20 mm region centered on the point of maximum response, determining the parameters of the polynomials. A unique pair of polynomials is calculated for each crystal studied. Subsequently, the corrected response for the ith event, S_i^{corr}, is calculated from the uncorrected response, S_i^{uncorr} by

$$S_i^{corr} = S_i^{uncorr} \frac{P_X^{MAX} \times P_Y^{MAX}}{P_X(x_i) \times P_Y(y_i)} = S_i^{uncorr} \times f(\vec{h}_i) \ , \tag{4.12}$$

where $(x_i, y_i) = \vec{h}_i$ are the impact coordinates of the ith event and P_X^{MAX} and P_Y^{MAX} are the maximum values of the polynomials.

4.2.4 Crystal α Parameter Measurement

The traditional approach for measuring α for a crystal is the *correlation plot* method, which involves fitting data points which each correspond to the average response of many laser or electron events for a single crystal, such as in Figure 4.6. A weakness of the approach is that the response changes over the time interval used to derive a single point are neglected, as are shower variations leading to sharing of energy with neighboring crystals. Finally, the method is performed on only one crystal at a time, rather than over multiple crystals in a way suited to *in situ* implementation, where there is no monoenergetic electron beam available for calibration. We propose a new approach which improves on the shortcomings of the correlation plot method, whereby we measure α by explicitly appealing to what the parameter is intended to

in the first place; the energy resolution is interpreted as a likelihood and α is chosen as the value which maximizes it. We describe this *energy resolution minimization* approach, along with the correlation plot method, in the context of these test beam irradiations.

4.2.4.1 Correlation Plot Approach

A typical example of the crystal response evolution is given in Figure 4.6, where the normalized electron and laser responses are plotted versus time. For each electron point, a corresponding laser point is calculated by linearly interpolating between the proceeding and following laser measurements. The alpha parameter from equation (4.3) is measured by comparing the electron responses to the corresponding interpolated laser responses. Specifically, equation (4.3) can be re-expressed using a logarithmic scale as

$$\log S = \alpha \log R + K \; , \tag{4.13}$$

where $K = \log S_0 - \alpha \log R_0$. Using this relation, the parameter α can be obtained from a linear fit of $\log S$ as a function of $\log R$. In this case, the normalization factors of the electron and laser responses contribute only to the constant K and do not directly affect the determination of α with their uncertainties. An example correlation plot and α fit are shown in Figure 4.7. In this case, the correlation plot is made using electron events incident in a 4 mm $\times$ 4 mm region centered on the crystal's point of maximum response, as measured by the hodoscope. This restriction is imposed in order to minimize variations from shower leakage. For each electron run, events are separated into bins by corresponding hodoscope values. The events in each bin are then grouped, according to time, into subruns such that there are at least 3000 events, giving a finer sampling of response loss evolution over the transverse geometry of the crystal face. The distribution of approximately 3,000 events is then fit using a simple Gaussian in a restricted range (one FWHM around the peak) in order to extract the mean value and error for that sample. This mean and error corresponds to one electron point in the correlation plot. The number 3000 is chosen so that,
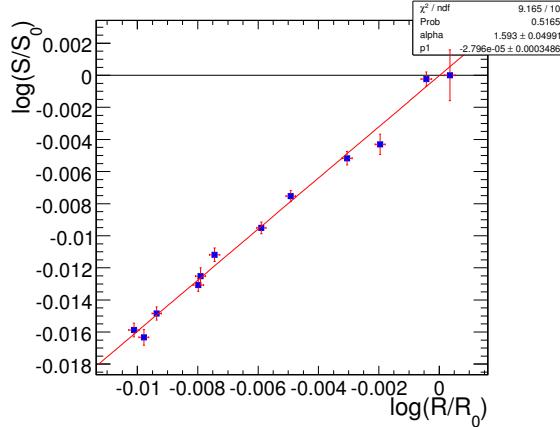
Figure 4.7: Correlation plot for crystal 128, supermodule 9, with the mean normalized electron energy loss plotted versus the mean normalized laser signal loss.

assuming roughly 0.5% energy resolution, the error on the mean of the distribution will be below 0.1%. The error in the measurement of α is then calculated by looking at the spread of α determined from independent correlation plots in 25 different 2 mm x 2 mm hodoscope bins in a 10 mm x 10 mm region on the face of the crystal, centered at the point of maximum response.

4.2.4.2   Energy Resolution Minimization Approach

In essence, the energy resolution minimization (ERM) approach to α parameter extraction is algorithmically similar to the correlation plot approach except, rather than grouping events into distributions according to time bins, events are considered individually. Each electron event, with a corresponding laser measurement, is used to construct an energy distribution whose standard deviation is explicitly minimized with respect to the alpha parameter.

During LHC running, the crystal response will be corrected as

$$S(t)_{corr} = S(t)_{raw} \left(\frac{R_0}{R(t)} \right)^{\alpha} . \tag{4.14}$$

Applying this correction to each individual electron event from test beam irradiation for a given crystal results in a monoenergetic energy distribution whose standard

deviation is a function of the crystal's α parameter. The ERM approach proceeds by minimizing the standard deviation divided by the mean of the electron response distribution with respect the α parameter. This ensures that the chosen α is the one which yields the best energy resolution, which should also correspond to the true value. In practice, it is more convenient to minimize the equivalent error function

$$Err(\alpha) = \frac{\sigma^2}{\bar{x}^2} + 1 = \frac{\sum_j^N (x_j)^2}{N} \frac{N^2}{(\sum_j^N x_j)^2} = \frac{\bar{x^2}}{\bar{x}^2} \; , \tag{4.15}$$

where x_j denotes the corrected energy of the jth event and is a function of α, σ is the standard deviation and N the number of events in the energy distribution. With the error function $Err(\alpha)$ written explicitly, the gradient with respect to α can be calculated as

$$\frac{\partial Err}{\partial \alpha} = \frac{2}{N\bar{x}^2} \sum_j^N \frac{\partial x_j}{\partial \alpha} \left(x_j - \frac{\bar{x^2}}{\bar{x}} \right) \; . \tag{4.16}$$

Equations (4.15) and (4.16) are general in that they can apply to different formulations of the corrected event energy x_j. For example, x_j can represent the response of the central crystal in an electron event, denoted E_1. Alternatively, x_j can represent the sum of the crystal responses from a matrix of crystals containing the incident electron, each response corrected with the corresponding crystal's α parameter. For this study, determination of α parameters will be done using both the E_1 formulation and E_{25} formulation, where E_{25} represents the sum of the crystal responses from a 5 x 5 matrix of crystals surrounding the central crystal upon which the electron is incident.

For the E_1 formulation, the energy response of the single crystal is especially sensitive to the impact position of the incident electron, as variations in shower containment result in variations of the deposition of energy in surrounding crystals. Subsequently, it is necessary to correct the E_1 response using the methods described in section 4.2.3.1. Denoting the raw central crystal response and corresponding normalized laser response for the jth event S_j and L_j, respectively, we write equations (4.15)

and (4.16) in the context of the E_1 formulation as

$$x_j := f(\vec{h}_j)S_jL_j^{-\alpha} \ , \tag{4.17}$$

$$Err_{E_1}(\alpha) = \frac{\sum_j^N(f(\vec{h}_j)S_jL_j^{-\alpha})^2}{N} \frac{N^2}{(\sum_j^N f(\vec{h}_j)S_jL_j^{-\alpha})^2} \ ,$$

$$\frac{\partial Err_{E_1}}{\partial \alpha} = \frac{2}{N(\sum_i^N f(\vec{h}_i)S_iL_i^{-\alpha})^2} \times$$

$$\sum_j^N \left[f(\vec{h}_j)S_jL_j^{-\alpha}\log(L_j)\left(\frac{\sum_i^N(f(\vec{h}_i)S_iL_i^{-\alpha})^2}{\sum_i^N f(\vec{h}_i)S_iL_i^{-\alpha}} - f(\vec{h}_j)S_jL_j^{-\alpha}\right)\right] \ ,$$

where $f(\vec{h}_j)$ is the containment correction factor from equation 4.12 and $\vec{h}_j = (x_j, y_j)$ is the hodoscope coordinates of the jth event.

For the E_{25} formulation we consider the α parameter of the central crystal of incidence, which we denote the kth crystal. Here, the containment correction used in the E_1 formulation is unnecessary since, by including the sum of the corrected responses from all crystals in a 5 x 5 matrix surrounding the central crystal, energy lost due to showering into neighboring crystals is already accounted for. Denoting, for the jth event, the raw crystal and normalized laser responses for the lth crystal as S_l^j and L_l^j, respectively, the error function and its gradient with respect to α_k can be expressed as

$$x_j := \sum_l^{25} c_l S_l^j (L_l^j)^{-\alpha_l} \ , \tag{4.18}$$

$$Err_{E_{25}}(\alpha_k) = \frac{\sum_j^N(\sum_l^{25} c_l S_l^j (L_l^j)^{-\alpha_l})^2}{N} \frac{N^2}{(\sum_j^N \sum_l^{25} c_l S_l^j (L_l^j)^{-\alpha_l})^2} \ ,$$

$$\frac{\partial Err_{E_{25}}}{\partial \alpha_k} = \frac{2}{N(\sum_i^N \sum_l^{25} c_l S_l^i (L_l^i)^{-\alpha_l})^2} \times$$

$$\sum_j^N \left[c_k S_k^j (L_k^j)^{-\alpha_k}\log(L_k^j)\left(\frac{\sum_i^N(\sum_l^{25} c_l S_l^i (L_l^i)^{-\alpha_l})^2}{\sum_i^N \sum_l^{25} c_l S_l^i (L_l^i)^{-\alpha_l}} - \sum_l^{25} c_l S_l^j (L_l^j)^{-\alpha_l}\right)\right] \ ,$$

where the sum over l is over the crystals in the 5 x 5 matrix and c_l is the inter-calibration constant of the lth crystal.

4.2.4.3 ERM Method Implementation

Due to the generality of the ERM formulation there is a great deal of freedom in specifying the specific details of its implementation, which should be chosen to fit the form of the data. For this study, electron data is grouped into runs such that, over the course of one run, the magnitude of crystals' transparency change is small compared to that over the course of the entire irradiation. As a result, it is possible to consider events from the same run together for preselection without *a priori* knowledge of the crystals' α parameters.

One issue is that, given the formulation of this approach, the error function and its gradient can be disproportionately dominated by outliers in the energy distributions. Additionally, distributions that are largely nonsymmetric can result in a systematic error in the determination of α, such as when there is a low energy tail in the distribution. For other studies, where the peak positions of quasi-Gaussian distributions are of interest, these problems are overcome by fitting the distribution in order to accurately measure the peak position. Here however, when it is necessary to use an explicit error function with a well-defined gradient, fitting the distributions in order to determine the standard deviation results in a non-smooth function of the α parameter, as shown in Figure 4.8 (left). A solution is to instead select events in the energy
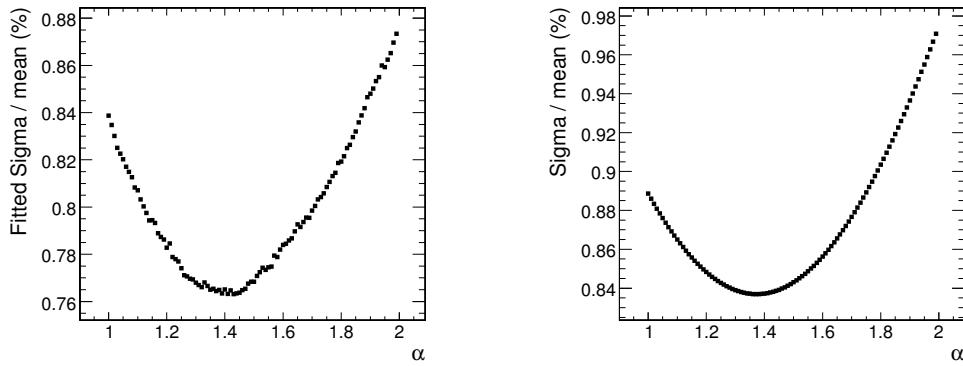
Figure 4.8: (Left) Fitted sigma/mean for crystal 168, supermodule 22. (Right) Sigma/mean for the same crystal, considering only events with energies near the peak response.

distribution in a symmetric way around the peak of the distribution and to exclude

53

outliers. This is accomplished by fitting the energy distributions for a particular ERM formulation (E_1 or E_{25}) using a Gaussian function within a restricted range around the peak position, and considering only events around the measured peak. The results of this procedure on the error function are shown in figure 4.8 (right). To ensure that the range of this event selection cut does not bias the measurement of the α parameter it is varied and α is recalculated, indicating no dependence on the cut range.

With an explicit error function and its gradient, the problem of α parameter measurement is reduced to that of a one-dimensional minimization. For this study, a parabolic interpolation was used that takes into account the gradient in order to speed convergence in the determination of α. The convergence qualities of the procedure for a sample crystal are demonstrated in figure 4.9.
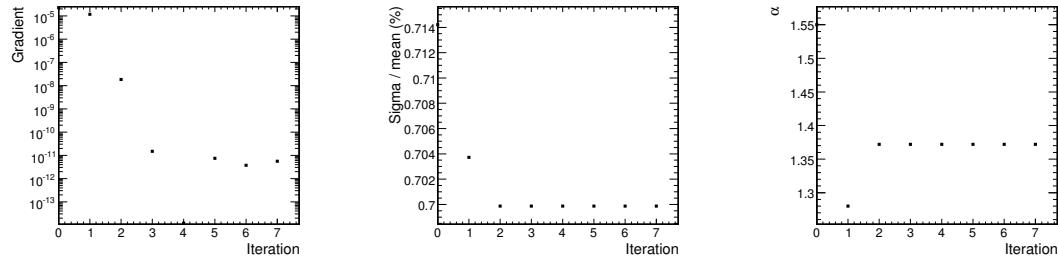
Figure 4.9: ERM parameters as a function of minimization iteration. (Left) Gradient of error function (equation (4.15)) with respect to α for the E_1 formulation. (Center) Standard deviation of the electron energy distribution. (Right) α.

4.2.5 Test Beam α Measurements from 2006

In addition to the requirements from each α measurement-method, events must satisfy a 10 mm $\times$ 10 mm hodoscope cut in the x and y hodoscope coordinates, centered at the position of maximum response for the crystal. For each crystal, the data used to measure α is split into 25 equally sized sets. From the resulting 25 independent measurements of α the error in the measurement is calculated from the spread in the distribution. The results of the measurements are summarized in Tab. 4.1. There

Crystal ID	Correlation plot α	E_1 α	E_{25} α	Change in response
128	1.52 ± 0.05	1.55 ± 0.02	1.52 ± 0.02	1.2 %
148	2.77 ± 0.51	1.58 ± 0.15	1.36 ± 0.24	0.5 %
168	1.34 ± 0.02	1.382 ± 0.004	1.364 ± 0.003	3.8 %
552	1.50 ± 0.11	1.55 ± 0.06	1.50 ± 0.03	0.9 %
672	0.76 ± 0.15	0.73 ± 0.15	0.65 ± 0.05	0.4 %

Table 4.1: Measured values for crystal α parameters using different methods.

is a strong correlation between the precision of the measurement of the α parameter for a crystal and the corresponding magnitude of transparency change during the irradiation period, as shown in Figure 4.10. This relationship is expected, since the the energy resolution of the crystal is more dependent on the α parameter when there is a greater change in response due to crystal transparency change, with a longer lever-arm for the correlation plot method and deeper error function minimum for the ERM approach. In general, we conclude that the minimization approaches are more precise than the correlation plot approach. Despite differences uncertainties, the values for α measured by each method are consistent, as illustrated in figure 4.11.
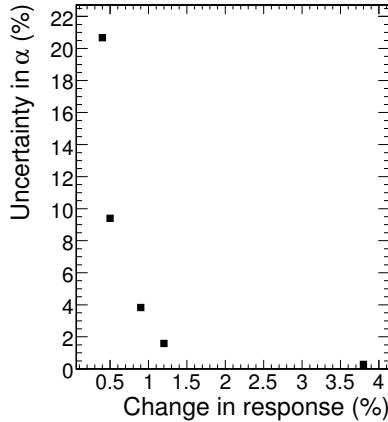
Figure 4.10: Magnitude of crystal response change during the irradiation run vs. uncertainty in the measured value of α (E_{25}) for all crystals
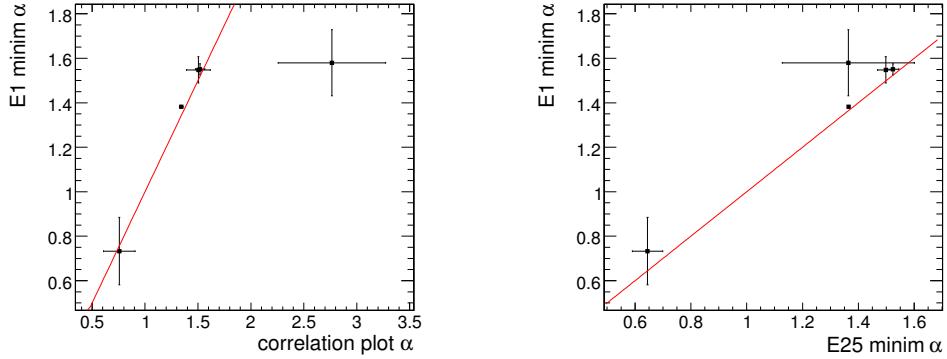
Figure 4.11: Correlation between α measured from the E_1 ERM approach and (left) correlation plot method, (right) E_{25} ERM.

4.2.5.1 Efficacy of Correction Procedure

Ultimately, the goal of the laser monitoring system is to correct for the energy loss due to radiation damage in order maintain the resolution of the calorimeter. Figure 4.12 demonstrates the difference in the resulting energy distributions for crystal 168 for the single crystal energy (E_1), the energy of the 3×3 matrix of crystals centered on 168 (E_9) and the energy of the surrounding 5 x 5 matrix of crystals (E_{25}) when using no laser correction and a correction with $\alpha = 1.36$ (the value measured for this crystal using the E_{25} ERM formulation). The distributions where made with a 10×10 mm^2 hodoscope cut and no containment correction was applied. The conversion of 40 MeV/ADC is used as the default of the reconstruction code, but should clearly be calibrated to a different value (the energy response peaks at 110 rather than 120 GeV) . When the laser monitoring correction is applied, the surrounding crystals are corrected with their respective normalized laser measurements using $\alpha = 1.6$, although since the magnitude of their transparency change is small the dependency of the E_9 and E_{25} values on these neighboring crystals' α values is negligible.

Although relatively small, a clear difference in energy resolution can be observed between the distributions corresponding to the transparency corrections using a default $\alpha = 1.6$ parameter and the measured α parameter for the irradiated crystal, as illustrated in Tab. 4.13. The contribution to energy resolution from uncertainty in
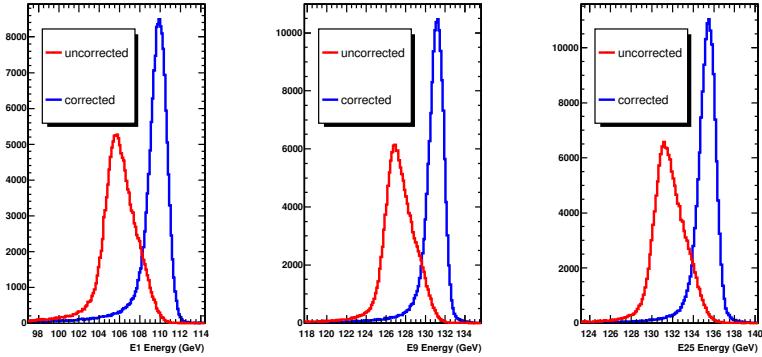
Figure 4.12: E_1 (left), E_9 (center) and E_{25} (right) energy distributions for crystal 168, with and without laser corrections.

	E_1 $\sigma(\text{E})/\text{E}$	E_9 $\sigma(\text{E})/\text{E}$	E_{25} $\sigma(\text{E})/\text{E}$
uncorrected	1.69%	1.41%	1.36%
$\alpha = 1.60$	0.722%	0.535%	0.539%
$\alpha = 1.36$	0.700%	0.528%	0.527%

Figure 4.13: Fitted energy resolutions for crystal 168 during irradiation with and without laser monitoring corrections. With corrections, the resolution is restored to values achieved before irradiation.

the crystal α parameter can be described by

$$\frac{\Delta E}{E} = \log\left(\frac{S}{S_0}\right) \frac{\Delta\alpha}{\alpha} \ , \tag{4.19}$$

where S/S_0 is the normalized crystal electron response. Observing the difference in peak energy between the energy distributions corrected with the measured α and default $\alpha = 1.6$ (110.5 and 109.9, respectively, for E_1), one can estimate the contribution to the energy resolution from the systematic uncertainty of α as $(110.5 - 109.9)/109.9 = 0.55\%$. Taking the value of $S/S_0 = 0.962$ for crystal 168 from Tab. 4.1 and an error in α of 16 %, equation (4.19) yields an estimation on the corresponding shift in the peak energy from using the incorrect α of 0.6 %, consistent with the value observed in data.

4.2.6 Test Beam α measurements from 2007

The EE Dee used in the 2007 test beam was equipped with 20 supercrystals, with individual crystals produced by either BTCP (russian, same as all the barrel) or SIC (chinese). The channel numbering and location of irradiated crystals is illustrated in figure 4.14. The 2007 test beam α measurements differed from 2006 in both the number of irradiations performed (12 unique positions, some multiple times in 2007) and the irradiation technique; Rather than irradiating the center of a single crystal, the electron beam was aimed at the corner between four crystals. This increased the number of irradiated crystals for which α could be measured by a factor of four.

Position	Irradiated Crystals	Manufacturer
1	131, 132, 151, 152	BTCP
2	136, 137, 156, 157	BTCP
3	236, 237, 256, 257	SIC
4	336, 337, 356, 357	BTCP
5	437, 438, 457, 458	BTCP
6	x431, 432, 451, 452	SIC
7	231, 232, 251, 252	BTCP
9	422, 423, 442, 443	BTCP
10	127, 128, 147, 148	SIC
11	123, 124, 143, 144	SIC
12	174, 175, 194, 195	BTCP

Figure 4.14: (Top) Channel numbering of the 20 super crystals used in the 2007 EE test beam. Positions in yellow correspond to crystals from BTCP, with purple indicating SIC. (Bottom) Irradiated crystals.

An additional advantage to the corner-shooting scheme is that the ERM method could be generalized to extract multiple values of α at once. Instead of the energy of a single crystal or 5×5 matrix the energy of the electron is estimated from a 6×6 matrix of crystals centered on the four irradiated ones. The total energy of these

crystals, E_{36}, is corrected from the laser measurement making its value dependent on the four central-crystal α values, and the energy resolution is minimized as a function of all their values simultaneously through a conjugate gradient-descent numerical minimization. The correlation between the α values measured with the correlation method and the four-crystal ERM approach are shown in figure 4.15 along with the distribution of α values for the BTCP crystals measured. The measurements from each method are in agreement, with more precise measurements following from the E_{36} four-α ERM approach. This is an important demonstration of the ERM approach being applied to multiple crystals simultaneously, indicating that the prospect for generalizing the method for an *in situ* α determination is good. The spread in α among the different BTCP crystals is as expected, $\sim$10%. The evaluation of the laser monitoring correction in section 4.2.5.1 indicates that not measuring each crystals α parameter independently will have an adverse effect on the ECAL energy resolution, with the 10% uncertainty in α when using one common value growing in significance with increasing transparency changes.
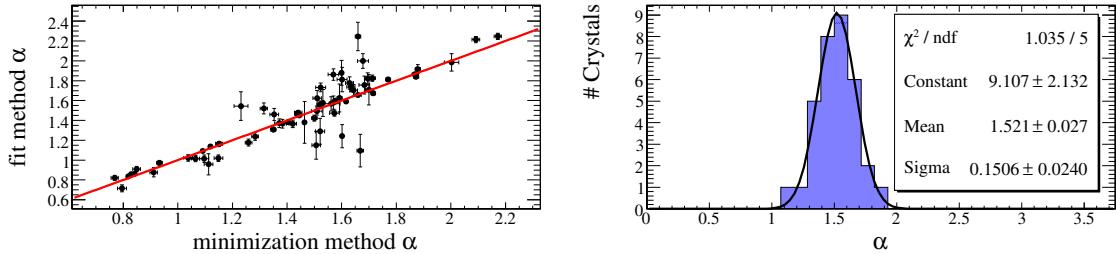
Figure 4.15: (Left) Correlation between values of α measured for endcap crystals using the fit method (correlation plot) and the minimization method (ERM) which measures α for the four central crystals simultaneously. (Right) Distribution of measured crystal α values for BTCP ECAL endcap crystals from test beam 2007 from ERM method. Values are corrected for the transverse orientation of light injection fibers to the dose profile (see section 4.2.6.1).

4.2.6.1 Effect of Fiber Geometry on the Measurement of α

In the 2007 test beam irradiations the beam center was aligned at the corner of four crystals. For endcap crystals, the laser light injection fibers are located at a specific corner, rather than in the center of the crystal face, as for the barrel. As a

result, the transverse dose profile has an orientation w.r.t. this fiber that is different for each crystal depending where the fiber is located. We identified four classes of orientations, corresponding to the fiber being located in the same corner of the crystal as the irradiation dose (denoted "1"), the fiber being located in the opposite corner (denoted "2") and the fiber being located in an adjacent corner ("0" and "3"). The adjacent corner orientation is further separated into two classes by the orientation of the crystal w.r.t. to the beam, which affects the dose profile w.r.t. the fiber. The raw measured α values for these different orientation classes are shown in figure 4.16.
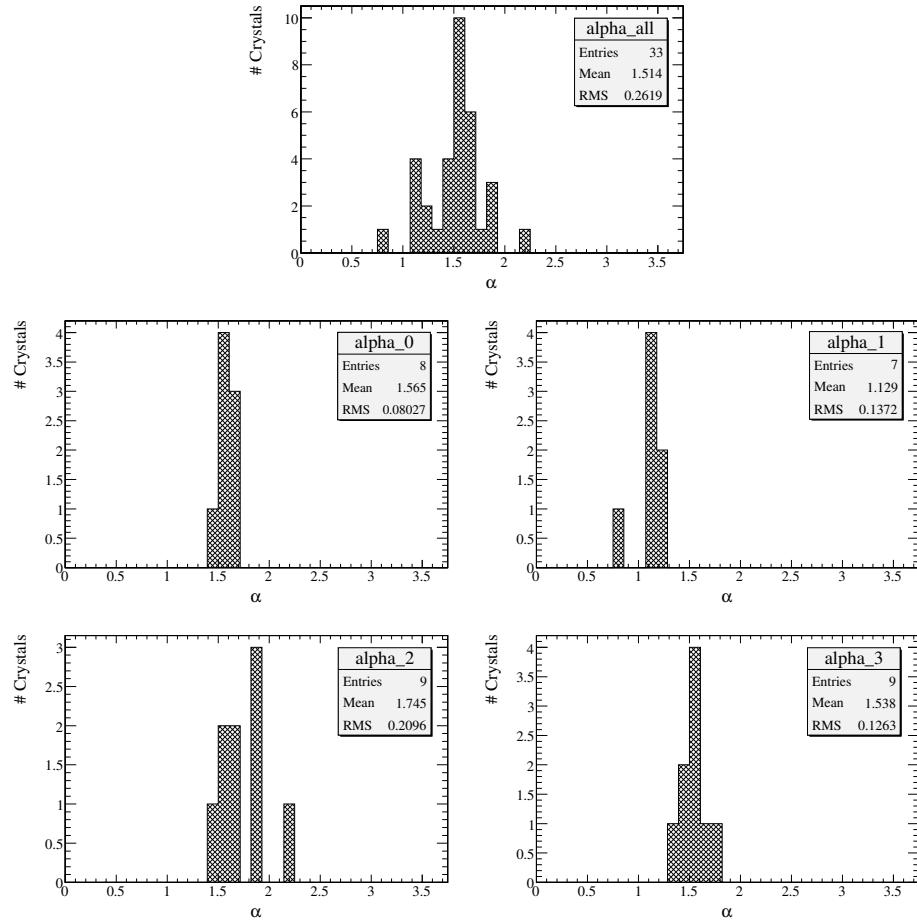
Figure 4.16: Measured α values for BTCP crystals in the 2007 test beam separated by fiber/dose orientation. (Top) All crystals. (Center left) Class 0. (Center right) Class 1. (Bottom left) Class 2. (Bottom right) Class 3.

We observe that the spread in α values within each orientation class is significantly

smaller than the spread over all the crystals, indicating that the different orientations result in a systematic change in the measured α value. This suggests that laser light sampling of the crystal is geometrically biased by where the light is injected, whereas the scintillation light from electromagnetic showers uniformly samples the whole crystal bulk, effectively experiencing different densities of color centers. This effect will not be present in *in situ* conditions, where the transverse gradient of the dose profile over a single crystal is negligible.

This effect is corrected for by normalizing the values from each separate class such that the mean of that individual distribution is equal to the mean from figure 4.16, yielding the distribution shown in figure 4.15. There are still residual effects from this dose/fiber orientation shift present in this corrected result, due to the fact that this should be a continuous correction, since the beam was asymmetric and not perfectly oriented at the corner of the four crystals. The larger spread in classes 1 and 2 (same corner/opposite corner) reflect this, since for these orientations the amount of transmission change (darkening) that the laser light samples is more sensitive to systematic misalignment of the beam in the corner of the four crystals, and any time dependent shift in the beam profile.

4.3 Electron Reconstruction

In a collision event, the signals from each ECAL crystals' photomultiplier are digitized by a sampling ADC, resulting in a series of 10 samples separated by 25 ns [51], each indicating the collected charge from the crystal's APD in that time slice. The amplitude of the signal is reconstructed from these samples [52] using a linear combination of their amplitudes: $A = \sum_j w_j \cdot S_j$, where S_j is the sample value in ADC counts and w_j is a weight, optimized using the measured average pulse shape. The pulse amplitude A_i, in ADC counts, of each crystal i is then multiplied by an ADC-to-GeV conversion factor G, which is measured separately for EB and EE crystals, and a crystal-by-crystal intercalibration constant C_i.

Since the lateral size of ECAL crystals are approximately one Moliere radius, the electromagnetic showers from incident particles generally spread over a few crystals in the lateral plane over the face of the ECAL. The ECAL clustering algorithm begins with the formation of "basic clusters" corresponding to local maxima of energy deposits. Due to the silicon tracker material in front of the ECAL (ranging from one to two radiation lengths depending on η) electrons and positrons will undergo Bremsstrahlung, with the magnetic field spreading this radiated energy in the ϕ direction. This energy is recovered through the formation of superclusters (SC), which are formed from groups of nearby basic clusters, extending further in ϕ in order to include clusters from radiated photons. In the barrel, this supercluster algorithm is called the "hybrid" algorithm, and is described in Ref. [53]. Due to differences in geometry, clustering in the EE and ES uses a slightly different algorithm, which merges together fixed-size 5×5 crystal basic clusters.

At this stage, the energies of superclusters are corrected to account for potential unclustered energy resulting from shower leakage Bremsstrahlung losses. These corrections are dependent on the type of the particle, its momentum, direction and impact point position. The supercluster energy can be expressed as

$$E_{e,\gamma} = F_{e,\gamma} \cdot \left[G \cdot \sum_i S_i(t) \cdot C_i \cdot A_i + E_{ES} \right] , \tag{4.20}$$

where the sum is over the crystals i belonging to the supercluster, $S_i(t)$ is the laser monitoring correction for crystal transparency changes and the factor $F_{e,\gamma}$ represents the supercluster energy correction. For endcap superclusters the preshower energy E_{ES} is also added.

Electron reconstruction proceeds by combining these energy-corrected superclusters with tracks reconstructed in the silicon tracker (section 3.3.1). This is performed by two complementary algorithms, where electrons are either seeded from ECAL superclusters or from tracks. In the former, the supercluster position is used to select pairs or triplets of hits in the innermost tracker layers in order to initiate the electron track reconstruction. The latter uses tracks as seeds and tries to match them to ECAL clusters by extrapolating the track measurement to the face of the ECAL. Regardless of the electron seed provenance, all the selected elements (track+SC) are used to reconstruct the electron tracks using including a modeling of the energy loss in the tracker material and a Gaussian Sum Filter (GSF) to fit the trajectories [53].

Electron candidates are built from the combination of ECAL superclusters and their associated GSF tracks and the properties of both, and their interconsistency, are used to identify electrons. The electron candidate's quality is based on the shower shape, its track/supercluster position and momentum agreement, and its isolation relative to signals in each of the subdetectors.

Shower shape

The shape of the energy shower in the ECAL and HCAL, reconstructed using the granularity of both detectors, is a good discriminator between energy deposition resulting from electromagnetic particles and those following from the hadronization of jets.

- **HCAL/ECAL energy ratio** The large number of radiation lengths covered by each ECAL crystal indicates that electromagnetic showers are unlikely to leak through the back of the crystal into the HCAL. As a result, the ratio of the electron's supercluster energy to that measured behind the electron in the HCAL, H/E, can be used to reject fake candidates.

- **Shower moment**: The shape of the ECAL supercluster can be used to identify candidates that are likely to result from electrons by looking at the distribution of energy through the crystals forming the cluster. The covariance $\sigma_{i\eta i\eta}$ is defined as

$$\sigma_{i\eta i\eta}^2 = \frac{\sum_i^{5\times 5} \left(\eta_i + \eta_{seed} - \bar{\eta}_{5\times 5}\right)^2 w_i}{\sum_i^{5\times 5} w_i} \ , \tag{4.21}$$

where the sum i is over the crystals in the cluster, E_i and η_i are the energy and pseudorapidity of the ith crystal and $w_i = \max(0, 4.7 + \log(E_i/E_{SC})$ is an energy-dependent weight (the same used in the determination of SC position). Small $\sigma_{i\eta i\eta}$ is indicative of the tightly clustered signature of an electron or photon while larger values show a more diffuse structure consistent with jet hadronization products. The η-direction covariant is used because it is not affected by shower-spreading in the azimuthal direction resulting from the magnetic field.

Track cluster comparison

Fake electron candidates can be rejected by requiring that the track and supercluster measurements of the electron momentum are consistent.

- **Position consistency**: Tracks are extrapolated to the face of the ECAL and the difference in angle, $\Delta\eta = \eta_{SC} - \eta_{\text{track}}$ and $\Delta\phi = \phi_{SC} - \phi_{\text{track}}$, is calculated.

- **E/p**: The ratio of ECAL cluster energy to track momentum should be in agreement for real electrons.

Isolation

The energy and momentum deposits measured from the subdetectors in the vicinity of electron candidates can be examined for traces of particles produced with the electrons, whose presence is indicative of particles coming from jets rather than promptly from, for example, W and Z bosons.

- **Tracker isolation**: Tracks in a cone around the electron candidate of $\Delta R = \sqrt{\Delta\eta^2 - \Delta\phi^2} = 0.3$ are selected, excluding those within $\Delta R < 0.04$ and also those falling in an $\eta - \phi = 0.015 \times 0.4$ strip to remove the electron track and deposits from photon conversions, respectively. The scalar sum of the p_T of these tracks corresponds to the tracker isolation.

- **ECAL isolation**: The sum of transverse energies of ECAL cells in a $\Delta R = 0.3$ cone around the candidate are summed, excluding the region $\Delta R < 0.06$ and an $\eta - \phi = 0.04 \times 0.4$ strip.

- **HCAL isolation**: The sum of transverse energies of HCAL cells in a $\Delta R = 0.3$ cone around the candidate are summed, excluding the region $\Delta R < 0.015$.

Chapter 5

Jets, Missing Transverse Energy, and the CMS HCAL

5.1 The CMS Hadron Calorimeter

The CMS detector must be able to reconstruct all of the particles produced in LHC pp collisions and, given the strongly interacting partons, many of these particles will be quarks and gluons which will hadronize into jets of particles. The CMS detector includes many interaction lengths of material in order to stop these particles and measure their energies. These measurements, along with the geometry and granularity of the HCAL, can be used to reconstruct the the momenta of jets appearing in collision events.

Hadrons are detected and reconstructed using the CMS hadron calorimeter (HCAL). The HCAL is composed of several subdetectors, covering different intervals of pseudorapidity and interaction depths. The HCAL barrel (HB) covers $|\eta| < 1.3$ and sits between the ECAL barrel and the solenoid's magnetic coil (1.77 m $< R <$ 2.95 m) with HCAL endcaps (HE) placed on either side behind the those of the ECAL, covering $|\eta| < 3$. Beyond that is the forward hadron calorimeter (HF), which completes the CMS calorimetric coverage to $|\eta| = 5.2$. In addition to geometric coverage, the HCAL system must also have enough stopping power to measure the full shower energies of incident hadrons. This is accomplished with the outer hadron calorimeter (HO) which is placed outside of the solenoid in front of the barrel muon systems. Between the full

HCAL detector, illustrated in figure 5.1, and the magnetic solenoid the material between the interaction point and the back of the HCAL detectors constitutes between 12 and 17 interaction lengths.
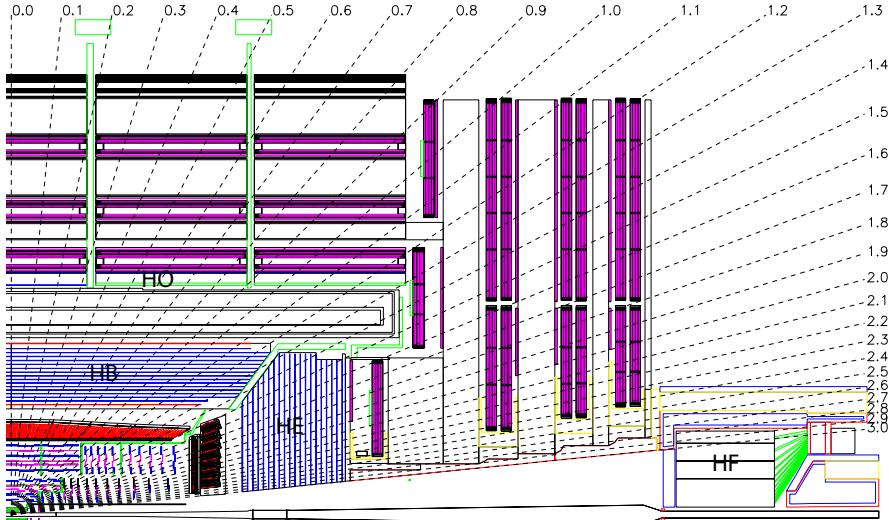
Figure 5.1: Longitudinal cross section of a quadrant of CMS, with the HCAL subdetectors HB, HE, HO and HF labelled. Interaction point is in the lower left corner. Dashed lines correspond to constant pseudorapidity.

5.1.1   Barrel Hadron Calorimeter

The HB is a sampling calorimeter composed of alternating layers of brass absorbers and plastic scintillating tiles. It consists of 36 azimuthal wedges which form two half barrels, one on either side of the interaction point, as shown in figure 5.2 (left). Each of these wedges is further divided into four azimuthal sectors, giving a granularity of $\Delta\phi = 0.087$. In the longitudinal direction, the plastic scintillators are divided in 16 intervals, constant in the interval of pseudorapidity they cover, which yields a granularity of $\Delta\eta = 0.087$, matching the azimuthal direction.

Each HB module has either 12 or 13 layers of 3.7 mm thick plastic scintillators, which are radiation hard. Between each of these layers are brass plates between 50.5 mm and 56.5 mm thick, increasing at larger radial distances from the beam line. The front and back plates are made of 40 and 75 mm thick steel, respectively. The layout
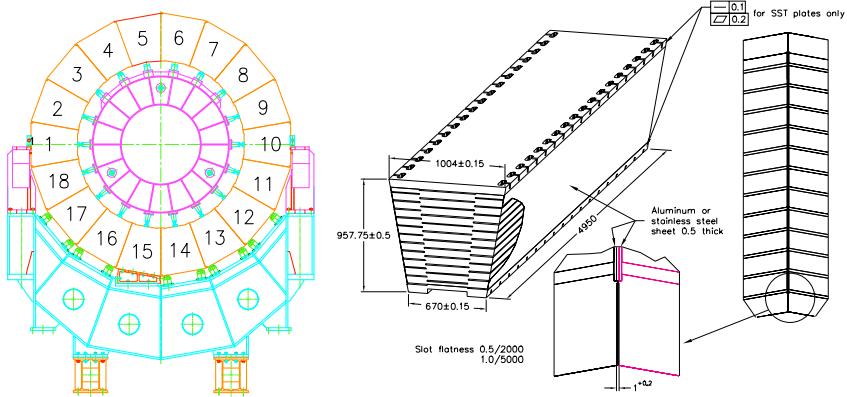
Figure 5.2: (Left) Illustration of transverse slice of HB with individual wedges numbered. (Right) Schematic view of an HB wedge, with design on scintillator sampling.

of an HB wedge is shown in figure 5.2 (right).

When particles are incident on the HCAL their shower develops in the absorber plates, while the energy of the particles produced in the shower are measured in the plastic scintillators. The scintillator light is collected with wavelength shifting fibers which brings the light to a hybrid photodiode [54], which has a gain of about 2000.

5.1.2   Endcap Hadron Calorimeter

The HE endcaps cover a large solid angle (13.2%) and as a result have a large particle flux, with approximately 34% of particles produced in LHC collisions falling in that interval. The same radiation-hard detector technology as the HB is used, with alternating layers of plastic scintillators connected to HPDs and brass absorber plates. Each HE layer is set at fixed distance from the interaction point along the beam line, meaning that the cells are arranged perpendicularly to the HB. As shown in figure 5.3 (left), the azimuthal granularity of the HE is the same as for the HB. The HE and HB sub-detectors are designed to overlap in pseudorapidity in the transition region between the two detectors in order to prevent gaps in geometrical coverage.

There are 17 layers of 9 mm thick scintillators in each HE unit, interleaved with 79 mm thick brass plates as shown in figure 5.3 (right). The HE constitutes about 10 interaction length for particles coming from the interaction vertex. Given the
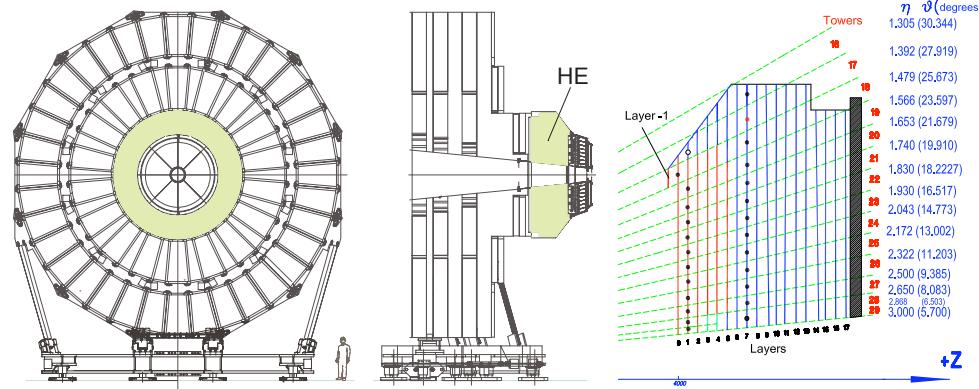
Figure 5.3: (Left) Illustration of HE endcap mounted on iron yoke, with transverse and longitudinal views. (Right) Diagram of longitudinal HE view, with layer segmentation shown as red and blue lines. Dotted lines point towards the interaction point and indicate the segmentation of separate HE towers.

geometry of the HE, the effective granularity of the individual cells changes from $\Delta\eta \times \Delta\phi = 0.087 \times 0.087$ in the region $|\eta| < 1.6$ to $\Delta\eta \times \Delta\phi \sim 0.17 \times 0.17$ at larger pseudorapidity.

5.1.3 Outer Hadron Calorimeter

In the barrel region the HB covers between 5.82 and 10.6 interaction lengths, depending on the polar angle relative to the interaction point. In order to complement this coverage, an outer hadron calorimeter (HO) is placed outside of the solenoidal magnet in order to catch the tails of hadronic showers that are not contained in the HB alone. With a contribution from the solenoidal coil of $1.4/\sin(\theta)$ interaction lengths, the HO gives the depth of CMS in the $|\eta| < 1.3$ barrel region to between 12 and 17 interaction lengths, depending on η. The HO is composed of one (two) cylindrical layers of scintillating fibers located outside (outside and inside) of a 19.5 cm thick piece of iron at a radial distance of four meters from the beam line. The HO geometry matches that of the barrel muon system, with five separate wheels at different fixed distances along the beam line. Only the HO for the central ring has two scintillator layers. Each of these rings has 12 segments in ϕ and six longitudinal slices, roughly matching the granularity of the HB. An illustration of the HO detector in CMS is

shown in figure 5.4.
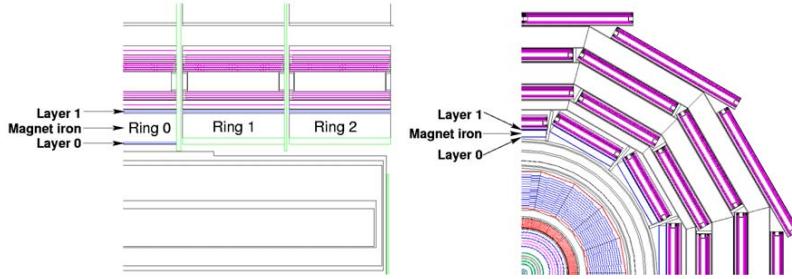
Figure 5.4: Longitudinal (left) and transverse (right) views of the HO detector, with different layers labeled.

The radial size of the HO is restricted by the muon system, with only 40 mm for the detector. In addition to an aluminum support structure, this leaves 16 mm for the detector layer, with 10 mm thick active scintillators. Like the HB and HE, scintillation light is collected by a wavelength shifting fiber and measured in HPDs.

5.1.4 Forward Hadron Calorimeter

In the far forward region, the LHC particle flux reaches unprecedented level at a particle detector, meaning that the components of the forward hadron calorimeter (HF) must be extremely radiation resistant. The scintillator-tile/wavelength-shifting fiber paradigm used in the other HCAL subdetectors would not withstand the expected LHC radiation rates, particularly up to $|\eta| = 5$ where 500 fb^{-1} of data would result in the HF experiencing $\sim$MGy of integrated dose [55]. For this reason, radiation-hard quartz fibers are used as the HF active material.

The HF consists of a steel absorber composed of 5 mm thick plates. Through the full depth of this absorber (165 cm $\sim$ 10 interaction lengths) run long quartz fibers, with smaller one starting at 22 cm from the front of the detector. Each of these fibers is 600 μm in diameter for the fused-silica core, extending to 800 μm with polymer hard-cladding and a protective acrylate buffer. When charged particles from showers in the steel absorbers pass through the fibers above the Cherenkov threshold ($E \geq 190$ KeV for electrons in this material) Cherenkov light is produced. This implies that

the HF is sensitive mostly to the electromagnetic component of showers [56], and the short and long quartz fibers can be used in conjunction to distinguish showers coming from electrons and photons from those of other particles.

Located 11.2 m from the interaction point, the HF is a cylindrical steel structure with an outer radius of 130 cm and extending within 12.5 cm of the beam line. It is subdivided azimuthally into 18 modular wedges, with a set of wedges on each side of the interaction region. The quartz fibers run parallel to the beam line and are grouped to form towers with granularity $\Delta\eta \times \Delta\phi = 0.175 \times 0.175$.

5.2   Jet Reconstruction

The energy depositions left by particles in the ECAL and HCAL are used to build representations of hadron jets by clustering them together. Each 5×5 cell of ECAL crystals is matched to single cells in either the HB or HE to form calorimeter towers (CaloTowers) which are the input to the clustering algorithm. Each CaloTower is interpreted as a massless particle with energy equal to the sum of measured energies in the constituent ECAL and HCAL cells. The direction of the CaloTower is assigned using the projective CMS geometry, assuming that the particles traveled from the interaction point. Section 5.4 discusses optimization schemes for this momentum assignment using reconstructed primary vertices in the event and the shape of showers within the CaloTowers.

The CaloTowers are clustered into jets using the anti-k_T clustering algorithm [57] with a size parameter $R = 0.5$ in the $\eta - \phi$ space, implemented in the `FastJet` package [58, 59]. The clustering is performed by four-momentum summation, such that each jet is the sum-total of all its constituents. The energies of these jets is then for corrected with jet-energy-scale (JES) factors derived from data. These are especially important in accounting for the energy lost in the noncompensating HCAL and the variable material budget between the interaction point and the calorimeters.

In 2011 running, the high instantaneous luminosity meant that single events can contain many interactions, leading to lower energy particles from softer interactions

biasing the reconstruction of jet energies coming from the primary interaction. This effect is corrected for by use of the *FastJet* pile-up subtraction algorithm. Here, the average energy density in $\eta - \phi$-space is calculated from the collection of CaloTowers used to cluster the jets. The area of each jet is then estimated, and its energy is corrected by an amount equal to the product of its area and the event energy density. Analyses using these jets (see sections B.7 and C.6) have demonstrated that this subtraction approach renders kinematic observables calculated from these jets momenta insensitive to pile-up conditions.

The performance the jet reconstruction algorithm is measured in samples of QCD multijet events using the dijet asymmetry method [60]. This approach exploits momentum conservation in the transverse plane for dijet events, using the imbalance of the jets as an estimator of the two jets' resolutions. The idealized topology of exactly two jets recoiling perfectly against each other in an event is violated by additional activity from the underlying event, soft radiation or lost energy from jet fragmentation effects. For dijet momentum balance, these effects are accounted for by measuring the dependence on extra event activity and extrapolating to zero contribution. The results of the asymmetry measurements from QCD multi-jet data are shown in figure 5.7.

5.3 Missing Transverse Energy Reconstruction

The presence of weakly interacting particles in collision events is inferred by appealing to conservation of momentum in the transverse plane of LHC collisions. By looking at the transverse balance of all of the reconstructed energy and momentum appearing in the detectors, the transverse momentum of any weakly interacting particles is inferred. The estimate of this quantity is denoted missing transverse energy (MET),[1] and it is defined as the negative vectorial sum of the momenta of all measured constituents in

---

[1]MET is a misnomer, since it is a vectorial quantity representing momenta. Historically, it has been calculated from the measured energies in calorimeters.
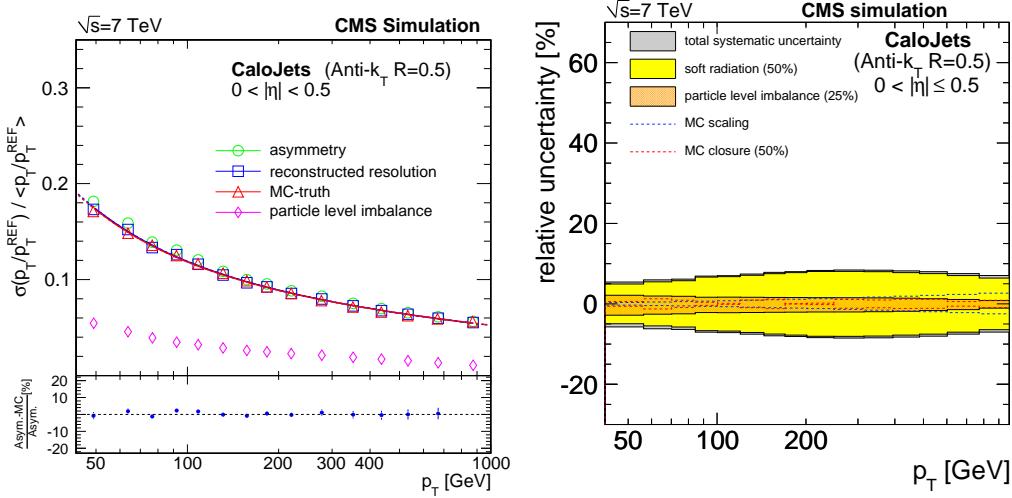
Figure 5.5: (Left) Jet resolution measured using the asymmetry method for jets in $|\eta| < 0.5$. The reconstruction-level (green circle) and particle-level (magenta diamond) results are shown together with the final measurement (blue square), compared to the generator-level MC (denoted as MC-truth) derived resolution (red triangle). (Right) Relative systematic uncertainty of the asymmetry method.

the final state

$$\vec{\mathrm{MET}} = -\sum_i \vec{p}_T^{\,i} \,, \tag{5.1}$$

where the sum is over all of the momentum reconstructed in the detector. Since a large portion of final state particles are neutral, and cannot be measured in the tracker, the ECAL and HCAL are of primary importance in MET reconstruction. A measurement of MET can proceed from the collection of the CaloTower constituents that are used to cluster jets, or by considering a different representation of the energy measured in the detector. In the studies described in this thesis, MET is reconstructed using the particle-flow algorithm [61,62] which attempts to construct every particle present in the collision event individually. Charged hadrons, electrons and muons are reconstructed from tracks in the tracker matched to the relevant calorimeters while photons and neutral hadrons are reconstructed from energy clusters separated from the extrapolated positions of tracks in ECAL and HCAL, respectively. A neutral particle overlapping with charged particles in the calorimeters is identified as a calorimeter

energy excess with respect to the sum of the associated track momenta. The energy of charged hadrons is determined from a combination of the track momentum and the corresponding ECAL and HCAL energy, corrected for zero-suppression effects, and calibrated for the nonlinear response of the calorimeters. Finally, the energy of neutral hadrons is obtained from the corresponding calibrated ECAL and HCAL energy.

By identifying each particle independently, calibrations are applied according to the particle type which allows the effective resolution of the HCAL measurements to be improved through use of the more-precise silicon tracker measurements. Similarly, the identification of photons, electrons and muons allows for particle-specific energy calibrations to be applied. This results in an improved MET resolution, in both scale and direction, relative to the calorimeter-only based analogue.

The performance of the PF MET measurement was evaluated in the earliest data from 2010, corresponding to about $10 \, \text{nb}^{-1}$ of 7 TeV collision data. During this early run period the low instantaneous luminosity allowed for low-prescale minimum bias triggers. In this dataset, the vast majority of events feature low $\sqrt{\hat{s}}$ QCD, where there are no intrinsic sources of hard, weakly interacting particles. This implies that the MET measured in these events is reflective of the MET resolution, which is larger than smaller effects that would cause the true transverse momentum imbalance to deviate from exactly zero in these minimum bias events. The PF MET distribution from this early data is shown in figure 5.6. The MET resolution depends strongly on the scalar sum of the measured energy of the particles used to calculate it. This is due to the fact that this resolution depends on the individual resolutions of the detectors used to calculated MET, which in turn depend on the total energy rather than the magnitude of its imbalance. Since MET is a vectorial quantity, it is usually decomposed into two scalar quantities by projecting it along two perpendicular axes in the transverse plane. The MET projected along the CMS x and y coordinate axes is shown in figure 5.6. The width of the distribution for this scalar quantity is indicative of the MET resolution. A more useful decomposition can be performed, event by event, along the axes parallel and perpendicular to the transverse thrust
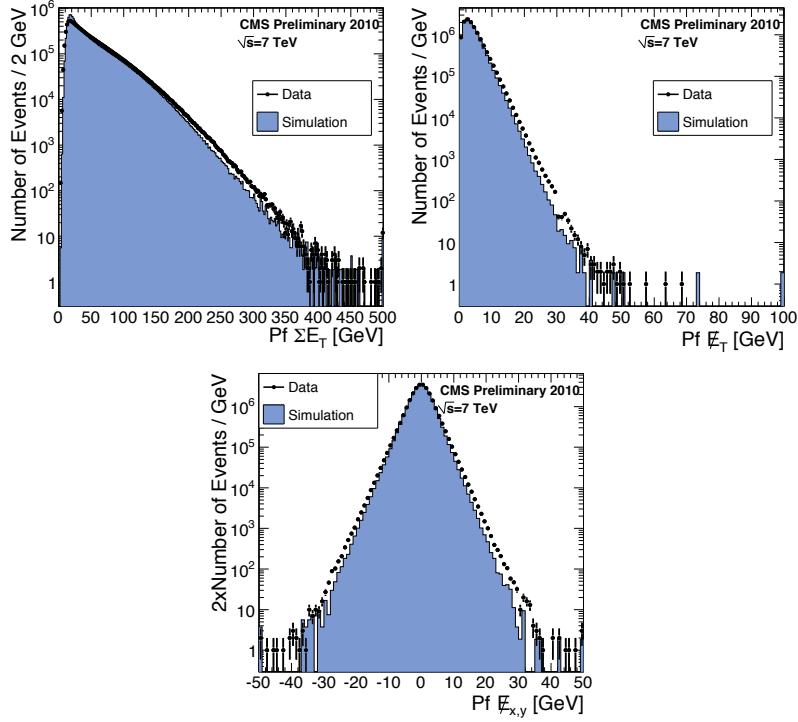
Figure 5.6: PF MET distributions measured from early 2010 data. (Left) Magnitude of the scalar sum of PF particle transverse momenta. (Right) Magnitude of the PF MET. (Bottom) PF MET magnitude projected onto the x and y directions, respectively.

axis. At larger scalar energy sums, many events will show characteristics of QCD dijet events, where the energy in the event is collimated along one axis in the transverse plane. This transverse thrust axis can be calculated using the precision silicon tracker by maximizing the quantity

$$
\text{TT} = \max_{\phi_{\text{TT}}} \frac{\sum_i |p_T^i \cos(\phi_{\text{TT}} - \phi_i)|}{\sum_i p_T^i} \; , \tag{5.2}
$$

where the sum i is over the tracks in the event with transverse momentum p_T^i and azimuthal angle ϕ_i. TT is the event thrust and ϕ_{TT} represents the transverse thrust axis. The MET can be projected along these two axes into a TT parallel ($\text{MET}_\parallel$) and perpendicular ($\text{MET}_\perp$) component

$$
\text{MET}_\parallel = \vec{\text{MET}} \cdot \hat{n}_{\text{TT}} \quad , \quad \text{MET}_\perp = \sqrt{|\vec{\text{MET}}|^2 - \text{MET}_\parallel^2} \; , \tag{5.3}
$$

where $\hat{n}_{\mathrm{TT}}$ gives the direction of the transverse thrust axis. The parallel MET component's resolution, with increasing scalar energy, is most sensitive to the calorimeter noncompensating response to jets and its resolution. Conversely, the perpendicular component depends strongly on the noise in the calorimeters. The resolution of PF and calorimeter-only MET, as a function of the calorimeter scalar energy sum, are shown in figure 5.7 for the first CMS 7 TeV data.
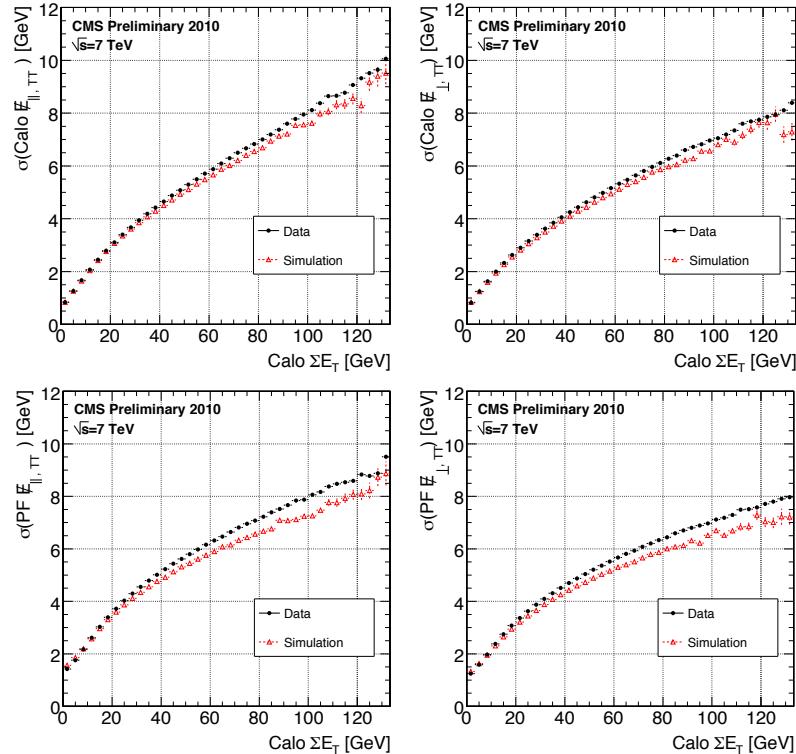
Figure 5.7: MET resolution as a function of the scalar sum of calorimeter energy in the event. (Top) MET from calorimeter measurements only. (Bottom) PF MET.

The PF improved over the calorimeter-only approach in resolution, although not to the magnitude predicted in simulated events. The calorimeter-only MET performance was in reasonable agreement with expectations. This was attributed to the early calibrations used in the PF reconstruction algorithm for charged and neutral hadrons. With subsequent calibrations the MET performance was improved, and it performance measured with the full 2010 CMS dataset [63].

5.4 Optimization of Momentum Assignment for Calorimeter Energy Deposits

The projective geometry of CMS associates each calorimeter cell with a position (η,ϕ) in pseudorapidity and azimuthal angle, respectively. When jets are clustered from the energy depositions in these cells, the momentum of these energy depositions are assigned a direction which implicitly assumes that the jet resulted from an interaction at the center of CMS, (0,0,0) in (x, y, z) coordinates. This assumption is justified by the fact that the size of the beam spot is negligible in the transverse plane and small on the z axis, when compared to the distance of the inner surface of ECAL from the center of the detector. It is know from previous work at the Tevatron [64] that an accurate projection of the energy deposits with respect to the true position of the event vertex is preferable, especially for search analyses based on multiple jets and/or MET. In particular, by removing the bias induced by incorrectly projecting the energy deposits to an incorrect interaction vertex, event migration from the core to the tails of the jet p_T and MET distributions can be reduced.

We explore two complementary approaches to improving the direction assignment to CaloTowers and jets. The first is to correct their direction for the true position of the primary vertex. As described in section 3.3.2, the position of every interaction vertex in LHC collisions is reconstructed precisely by the silicon pixel detector and strip tracker. Effectively, this measurement can be used to improve the position resolution of the calorimeters. We also consider a momentum assignment scheme that exploits the added granularity of the ECAL relative to the HCAL cells and CaloTowers. This *variable positioning* approach uses the shower shape in the 5×5 ECAL grid assigned to each CaloTower in order to improve the position assignment. We evaluate the effect of applying these improved position-assignment strategies to the CaloTowers before jet clustering (*a priori* corrections) or to the clustered jets (*a posteriori* corrections). Both correction schemes are found to improve jet position and energy resolution and reduce biases in their determination.

5.4.1   Correcting for the Primary Vertex Position

By default, each CaloTower is assigned an η_d and ϕ_d coordinate based on its position with respect to the nominal $(0, 0, 0)$ interaction point. Assuming that the interaction occurs precisely at this point, this assignment is sufficient, as it uniquely describes the CaloTowers' positions on the 2-sphere perpendicular to the nominal interaction point. When the primary event interaction occurs in a position displaced from the nominal interaction point these two coordinates are no longer describing the rapidity of the particles that impact a given CaloTower. We specify the CaloTower's position in three dimensions in order to assign its "physics" position on the 2-sphere perpendicular to the actual event interaction point. More specifically, the position of CaloTowers (and subsequently the jets clustered from them) can be described as a function of the η_d and ϕ_d coordinate (relative to $(0, 0, 0)$) and a "depth," or reference length, specifying the distance of the CaloTower/jet from the beam line, or the CMS z-axis. The geometry of a displaced primary vertex is illustrated in figure 5.8.
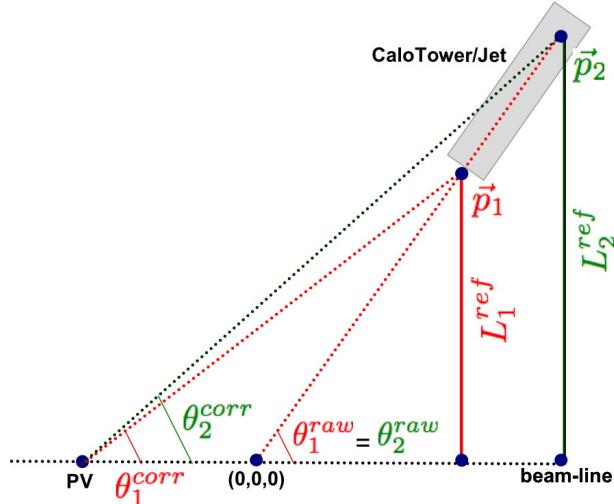
Figure 5.8: The geometry of a primary event vertex displacement along the CMS z-axis (beam line). In addition to a value of η_d and ϕ_d relative to the nominal interaction point, an additional parameter, L^{ref}, is necessary to uniquely identify the position of a CaloTower/Jet.

When the interaction occurs at the point $(0, 0, 0)$ the points $\vec{p}_1$ and $\vec{p}_2$ have an equivalent η_d and ϕ_d, regardless of the reference length, L^{ref}, chosen. This is a

direct consequence of the projective geometry of the CMS detector. On the other hand, when the event primary vertex (denoted as **PV** in figure 5.8) is displaced from $(0, 0, 0)$ the resulting value for the corrected η (or θ) coordinate for the CaloTower/jet depends on the chosen reference length, L^{ref}. This length corresponds to the distance between the energy depositions and the beam line and essentially requires the entire CaloTower/jet be described as a single point. Strategies for assigning this reference length are discussed in section 5.4.2.

Given the dimensions of the beam-spot (large spread on the order of 10 cm in z, 10's of μm in x/y), we only consider displacements of the primary event vertex in the CMS z coordinate, which implies that these corrections improve only the η and p_T/E_T measurements of the CaloTowers/jets. Significant displacements in the x/y (transverse) coordinate of the interaction point from $(0, 0, 0)$ should also be corrected for should they occur. The geometry of this correction implies that central CaloTowers and jets are more sensitive to the z-coordinate of the primary event vertex position relative to more forward objects. This effect is demonstrated in figure 5.9. Here, a toy Monte Carlo was implemented assuming a calorimetry object at each value of η_d shown. The primary interaction vertex position was varied in z, taken from a Gaussian distribution with mean 0 and $\sigma = 5.3$ cm. Additionally, a reference length of $L^{ref} = 159$ cm was used for all values of η_d (roughly the back of the ECAL). Given the relative magnitudes as a function of η, one finds that barrel calorimetric objects' reconstructed positions are more sensitive to the choice of L^{ref}.

5.4.2 CaloTower and Jet Depth Assignment for PV Corrections

As illustrated in figure 5.8 the correction to CaloTowers/jets is sensitive to the choice of depth, especially in the barrel region. In principle, the correct choice for this value would correspond to the distance from the beam line of the intersection between a line projected from $(0, 0, 0)$ through the detector η_d and ϕ_d and the actual trajectory of the physics object (jet or single particle) from the actual interaction point when the
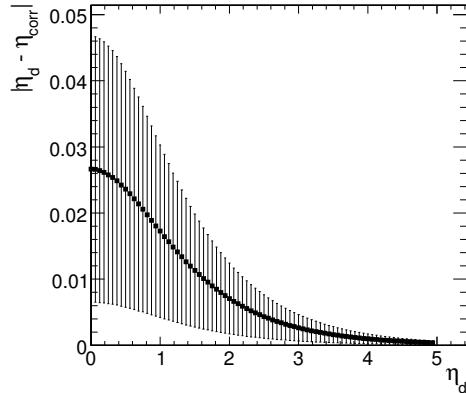
Figure 5.9: Toy Monte Carlo study of the event PV smearing around $(0, 0, 0)$. The z coordinate of the primary vertex is assumed to be normal distributed with mean 0 and $\sigma = 5.3$ cm. A reference length, L^{ref}, of 159 cm is used. Points correspond to the mean of the $|\eta_d - \eta_{corr}|$ distribution while the error bars correspond to the RMS.

detector has infinite granularity. In practice, it is not possible to determine this point given the finite granularity of the calorimeters, bending of the charged components of jets in the magnetic field, and the fact that the calorimeters are designed to be projective relative to the nominal interaction point. A well-motivated approach would be to assign L^{ref} according to the location of the longitudinal shower maximum of the jet/single particle. Unfortunately, for hadronic showers this point varies from event to event, even at a fixed incident energy for a single hadron. This is due to, among other things, fluctuations in the starting point of the shower, varying π^0 content in the shower and differences in noncompensation between the ECAL and HCAL components of the CaloTower. One could assign the depth as a function of the energy and electromagnetic fraction of the CaloTower or jet on an event-by-event basis. This ratio however strongly depends on the hadronization model chosen when simulating jets and is sensitive to the relative calibrations of the different components of the CaloTower (ECAL and HCAL) and their relative energy resolutions. To avoid introducing sensitivity to these effects we instead identify a "global optimal depth" which optimizes the results of these corrections. The jet η resolution as a function of η is shown in figure 5.10 for different choices of this depth. Here, we define "resolution"
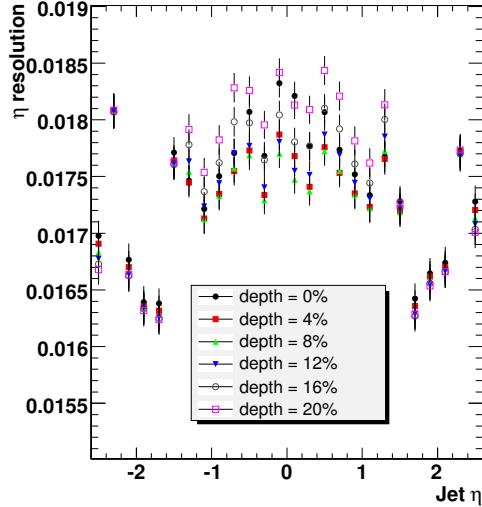
Figure 5.10: Jet η resolution as a function of η for jets with $p_T > 50$ GeV/c and *a priori* primary event vertex correction.

as the *mean absolute residual*. For a Gaussian distribution with mean zero this is equivalent to $\sigma \times \sqrt{2/\pi} \sim 0.8 \times \sigma$. It is calculated from a simulated sample of QCD multijet events by comparing the reconstructed jet momentum to that of generator-level jets clustered from the simulated particles in the event. The depth is quoted as a percentage of the distance between the front of the ECAL ($\sim$140 cm) and back of the HO ($\sim$410 cm) for the barrel region. Correspondingly, for the endcap region L^{ref} is defined as a percentage of the distance along the z axis between the front of the EE and back of the HE (320 cm and 570 cm from (0, 0, 0), respectively). The optimal choice for this depth is revisited in the context of other position corrections.

5.4.3 Variable Positioning for CaloTowers

In the absence of a displaced event primary vertex, each CaloTower is assigned a default η_d and ϕ_d coordinate. This assignment reflects the granularity of the HCAL cells, which have roughly a one-to-one correspondence with the CaloTowers. Fortunately, there is additional position information available from the increased granularity in η_d and ϕ_d of the individual ECAL crystals associated with each CaloTower in the

barrel and endcap regions. In general, the strategy is to use the granularity of the ECAL cells in the transverse plane w.r.t. the nominal interaction point in order to derive an event-by-event variable position for each CaloTower based on ECAL cell energy depositions. There are a number of ways to implement this strategy, including not only variations of different parameters but entire formulations. For example, one could treat each of the CaloTower's calorimetric cells, ECAL and HCAL, as a separate four-vectors and combine them in some prescribed way. Depending on the approach to this four-vector combination, this can lead to massive CaloTowers. Alternatively, one could use just ECAL cells to determine an η_d and ϕ_d for the entire CaloTower, and set the depth according to some fixed parameter (as described in section 5.4.2) or as a function of CaloTower electromagnetic fraction, E_T, η_d, etc. Such an approach would, as a result, be sensitive to these calorimeters' relative calibrations, resolutions and variations in hadronization and hadronic shower development in the calorimeters. For the sake of robustness we will consider here only those approaches that use the ECAL granularity to calculate an η_d and ϕ_d coordinate (relative to $(0,0,0)$) and will assign a value of L^{ref} according to an approximate "global optimization." With this formulation, the calculation of the CaloTower positions can be expressed as

$$\eta_{CT} = \frac{\sum_i^{cells} \eta_i w_i}{\sum_i^{cells} w_i} \quad , \quad \phi_{CT} = \frac{\sum_i^{cells} \phi_i w_i}{\sum_i^{cells} w_i} \ ,\tag{5.4}$$

where η_{CT} and ϕ_{CT} are the positions of the CaloTower in the transverse plane relative to $(0,0,0)$, the sum is over the ECAL cells assigned to the tower, and w_i is a cell weighting factor. With this parameterization, variations include different choices for the calculation of the weights (energy dependence, η dependence) and restrictions on which cells are included in the sum (for example, an absolute minimum cell energy or a cut on the cell energy relative to the total electromagnetic component of the CaloTower). We consider two different weighting schemes. The first, denoted the "log-method," uses logarithmic energy weighting

$$w_i = w_0 + \log\left(\frac{E_i}{\sum_j E_j}\right) \ ,\tag{5.5}$$

where the sum is over the same ECAL cells summed in equation (5.4) and E_i is the measured energy of the ith ECAL cell. An additional requirement is that only positive weights w_i are considered, so that the parameter w_0 acts as a relative energy cutoff. Without optimization in the context of the CaloTowers, we use the value $w_0 = 4.2$, the same as in the assignment of ECAL cluster positions [65], implicitly including only cells with more than 1.5% of the total electromagnetic energy in the sum. Furthermore, the "Scheme B" threshold requirements are maintained for all the CaloTower calorimetric cells.

Thresholds (GeV)				
HB	HO	HE	$\sum$EB	$\sum$EE
0.9	1.1	1.4	0.2	0.45

Table 5.1: Scheme B minimum calorimeter energy thresholds for CaloTowers.

The effect on the jet η resolution of varying the value of L^{ref}, for both *a priori* and *a posteriori* primary vertex corrections, is illustrated in figure 5.11 for this variable positioning scheme. The two types of primary vertex correction result in degenerate performance. Interestingly, the depth hierarchy is enhanced when using the variable positioning scheme relative to the fixed positioning case, demonstrating that a value of depth = 6% is roughly optimal, which is consistent with the "optimal" range for fixed CaloTower positioning discussed in section 5.4.2. It is clear that this is not precisely the optimal value for all values of η but also that, in a range of several percent around this value, η resolution performance is effectively degenerate. Hence, we choose depth = 6% as a benchmark value. The η distribution for *a priori* primary vertex corrected leading jets ($p_T > 50$ GeV/c) is shown in figure 5.12 for fixed and variable CaloTower positioning, respectively. For fixed positioning, the spikes in the η distribution are clearly visible, with 54 of them corresponding precisely to the 54 positions of CaloTower η segmentation in this η range. For variable positioning, this feature is less pronounced, resulting in a much smoother η distribution.

We consider a second formulation for CaloTower position calculation called the
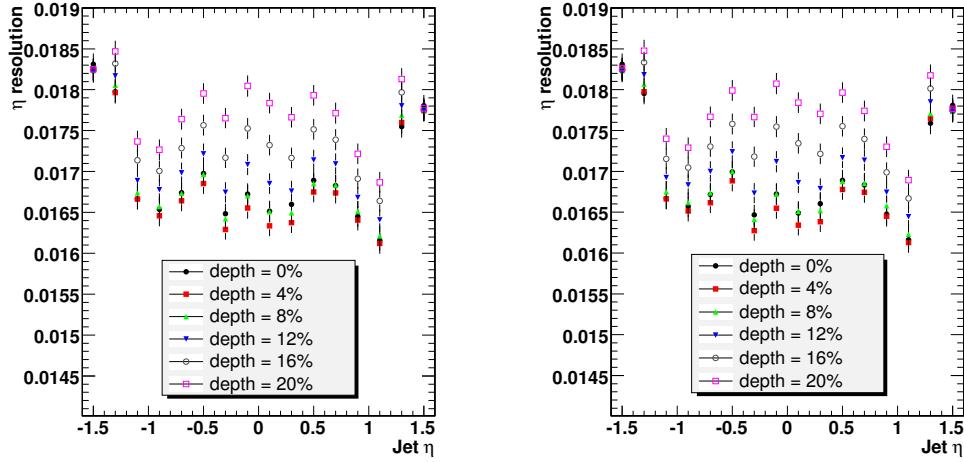
Figure 5.11: Jet η resolution as a function of η for jets with $p_T > 50$ GeV/c for (Left) *a priori* primary event vertex correction and (Right) *a priori* corrections. A logarithmic-weighted variable positioning scheme is applied.
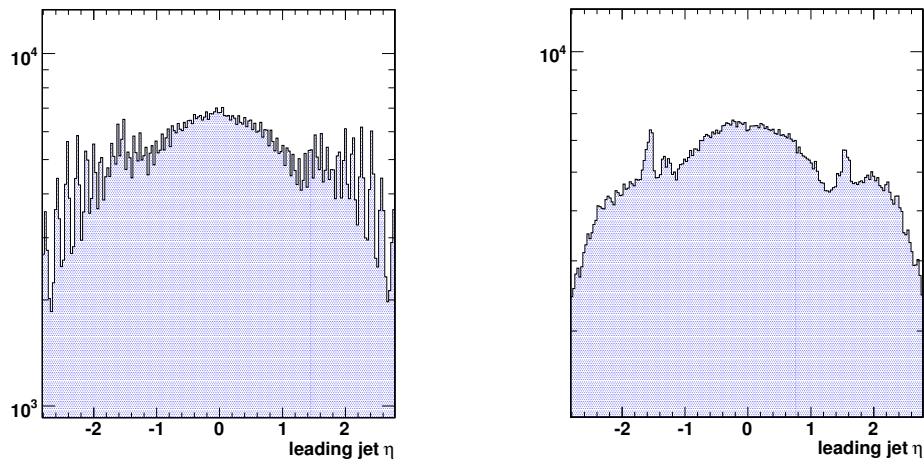
Figure 5.12: Distribution of leading jet η for jets with $p_T > 50$ GeV/c. (Left) Fixed CaloTower positioning. (Right) variable positioning.

"linear-method," where the ECAL cell weights are chosen as

$$w_i = E_i \; . \tag{5.6}$$

With no explicit energy requirements applied to the ECAL cells, this approach is roughly equivalent to combining the four-vectors of each of the ECAL cells when deriving the CaloTower position, with either an energy or E_T recombination scheme.[2] The jet ϕ and η resolutions for the log-method and linear-method position calculation approaches are compared in figure 5.13. The two methods perform almost identically w.r.t. to this metric (and also in position bias). The linear-method performs slightly better in ϕ resolution than the log-method while the opposite appears true for η resolution. One possible explanation for this effect is the difference in the physics that the two weighting approaches address. The log-method is motivated by the fact that hadronic and electromagnetic showers can be parameterized with exponential-like functions for the transverse shower development [66]. Hence the better the transverse profile of a shower can be described by a single exponential, the better the log-method will perform. Conversely, the more a shower drops off linearly with radius, the better the linear-method description will perform. This consideration affects both the η and ϕ resolutions. For ϕ resolution, there is an additional effect due to the bending of charged particles in the magnetic field which degrades the position resolution and introduces additional bias. This effect is directly related to the transverse momenta of the particles and, as a result, is better captured by the linear-method. We proceed by choosing the log-method as default for the variable positioning scheme, keeping in mind that marginal improvements from other ECAL weighting approaches could be achieved.

---

[2]Due to the small η range of ECAL cells within each CaloTower the factor $1/\cosh(\eta)$, which is applied to translate measured energy to E_T, is approximately the same for each ECAL cell.
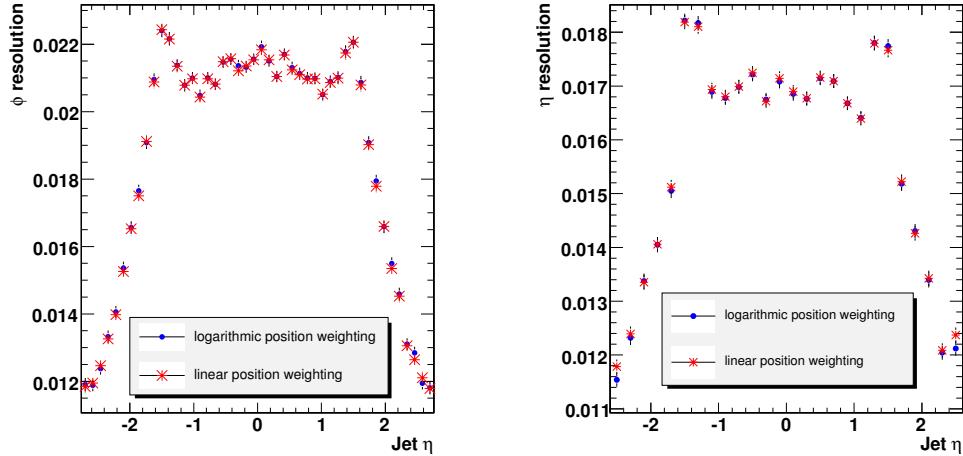
Figure 5.13: Jet position resolution as a function of η for jets with $p_T > 50 \,\mathrm{GeV}/c$ with two different variations on the weighting of ECAL cells for the position calculation. (Left) ϕ resolution. (Right) η resolution.

5.4.4   Corrected Jet Momentum Resolution and Bias

In figure 5.14 we compare the jet resolution achieved with the different correction schemes considered: primary vertex corrections, variable CaloTower positioning using the granularity of the ECAL and schemes where these corrections are performed before and after jet clustering, respectively. We observe that not correcting for the primary vertex position results in an appreciable degradation of the jet η resolution, especially in the central region where the jets' positions are most sensitive to the geometry of the correction. The ϕ resolution is unchanged by this correction. The variable positioning systematically improves jet position resolution relative to fixed positioning, except in the region of nontrivial overlap between the HB and HE ($|\eta| \sim 1.3$). With the exception of this region, which can benefit from a better description of the complicated ECAL, HB and HE overlap geometry, the improvements from the variable positioning scheme over fixed positioning are as large as 50% for jets in the calorimeter endcaps.

While jet position resolution is an important metric, it is also essential to ensure that low position bias is achieved. Not correcting for the primary vertex position results in a significant position bias as the primary vertex z coordinate deviates from 0, as demonstrated in figure 5.15. The variable positioning yields lower η bias, although
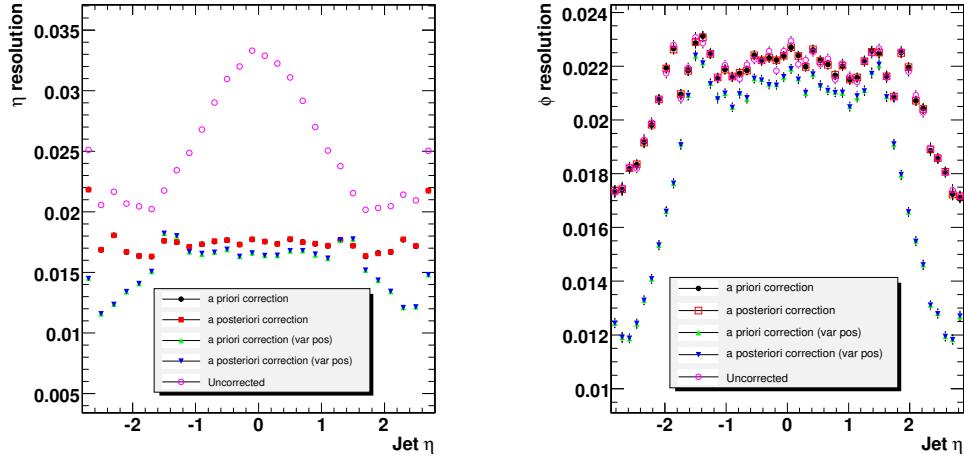
Figure 5.14: Jet position resolution as a function of η for jets with $p_T > 50$ GeV/c. (Left) η resolution. (Right) ϕ resolution.

this result should be put in the context of figure 5.15 in that negligible bias (relative to the η resolution) results from correcting for the primary vertex position, regardless of CaloTower position calculation.
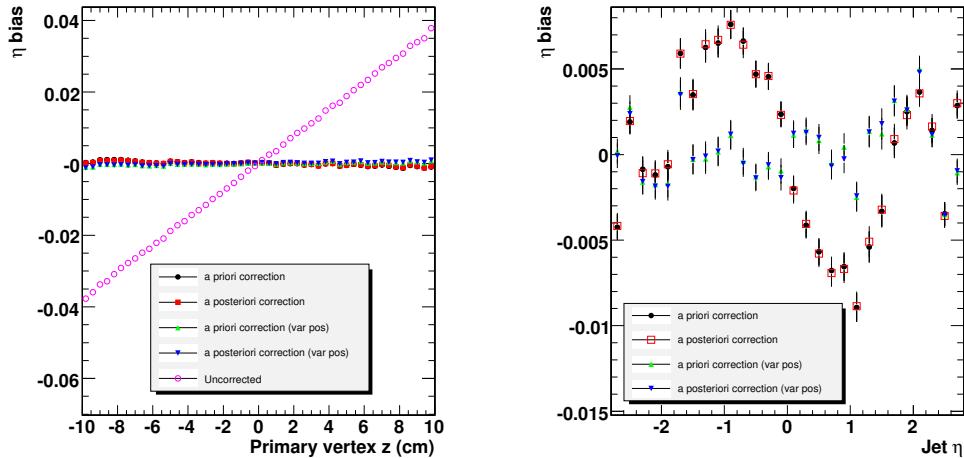
Figure 5.15: Jet η bias as a function of the primary event vertex z position. (Left) Jets with $p_T > 50$ GeV/c. (Right) soft jets with $p_T > 10$ GeV/c.

Chapter 6

The W and Z Boson Standard Candles

The fact that the W and Z bosons are massive is of tremendous theoretical and practical importance. These masses regulate the strength of the weak force, allowing for the hydrogen fusion reaction in the sun and for life to exist. They also indicate that the gauge symmetries of the SM are broken, through the predicted Higgs mechanism. For studying LHC collision events these masses are relevant not only because of the physics these bosons are associated with, but also because of their utility as *standard candles* which can be used to calibrate the detectors and measure their performance.

The relatively large masses of the W and Z bosons means that they are an abundant source of isolated, high transverse momentum leptons and neutrinos. $Z(\ell\ell)$ events can be used with a tag-and-probe technique, described in section 6.1, to measure the efficiency of lepton reconstruction algorithms in data, exploiting the Z mass in selecting events without biasing the measurement. Heavy bosons can also be used to tag particular event topologies in order to study their other properties. $Z(\mu\mu)$+jets events can be used to measure the jet energy scale by using the well-measured Z kinematics to calibrate the recoiling jets' [60]. $W(\ell\nu)$ events provide a control sample of events with weakly interacting particles which allows for measurements of the MET performance which and can be selected by identifying the decay lepton.

W and Z bosons, the force carriers of the electroweak part of the SM, can also be used to study strong interactions by identifying the bosons in events with associated

jet production. Studies of the scaling behavior of W/Z+jets yields with increasing number of jets is described in section 6.2, and produces a measurement which directly test perturbative QCD.

6.1   Lepton ID Efficiency Measurements Using $\mathbf{Z}(\ell\ell)$

A challenge for validating reconstruction and identification algorithms is doing so for leptons in events collected from LHC collisions. Simulated events give the benefit of knowing the true magnitude or identity of every quantity and particle in the, which provides a simple metric for comparison with reconstructed particles. Such information is not available in data, where the actual performance must be measured in order to understand the errors associated with physics analyses. In the case of lepton identification, the Z boson and its decays can be used as a standard candle to tune and measure lepton reconstruction performance. As an example, we consider the tag-and-probe measurement of the lepton identification efficiencies for the SUSY search analysis described in chapter 10.

From 2011 $Z(\ell\ell)$ events are selected by requiring that events contain two lepton candidates, either two electrons or two muons. Furthermore, one of these candidate (the tag) must satisfy a tight lepton identification requirement and the invariant mass of the two candidates must be close to the Z pole. This yields a relatively pure sample of $Z(\ell\ell)$, which can be used to evaluate the efficiency of identifying the second lepton candidate (the probe) as good, already knowing that the candidate is likely an actual lepton from independent information. The efficiencies of the offline lepton selection are measured using this tag-and-probe approach in 1.55 fb^{-1} of 2011 7 TeV data in different kinematic regions (bins in p_T and η) which match the granularity of ECAL for electron and of the muon chambers for the muons.

For $Z(ee)$ events, the electron selection efficiency can be factorized into the two contributions: the selection requirements involving the invariant mass and for an elec-

tron to be identified from the probe candidate. After the event selection, the only requirement on the probe electron is that it has a supercluster with loose properties. A fit is performed to determine the electron identification efficiency for supercluster candidates by considering the $m(e^{\text{tag}}e^{\text{probe}})$ invariant mass distribution. $Z(ee)$ events are modeled as a signal peak while potential non-$Z(ee)$ backgrounds are represented as a falling continuum. We assume that peaking events are real electrons, while continuum events are fake candidates since the contribution from continuum Drell-Yan di-lepton production is small at these masses relative to the fake contributions from QCD multi-jet events. These shapes are used to constrain the two contributions in a maximum likelihood fit, from which the electron identification efficiency is extracted by simultaneously considering a sample with probes failing identification and those satisfying it. Two example fits for two η bins of samples with electron probes $20 < p_T < 25$ GeV/c are shown in figure 6.1.
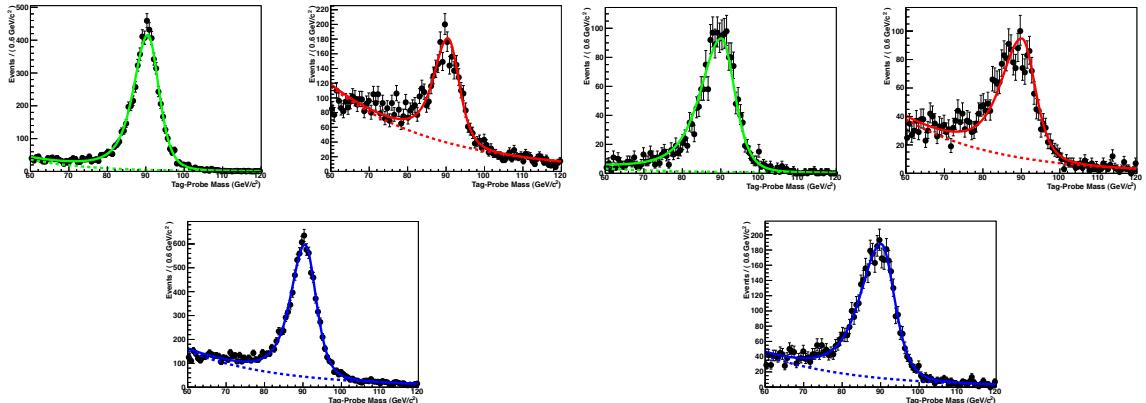
Figure 6.1: Distribution of e^+e^- invariant mass in the electron tag-and-probe sample. Probes are selected to have $20 < p_T < 25$ GeV/c and (Left) $|\eta| < 1$ and (Right)$1.566 < |\eta| < 2.0$. Super-imposed are the likelihood functions from the fit for samples of passing probes (green), failing probes (red), all probes (blue). Selection applied to probe corresponding to WP80 electron identification described in section 10.4.

This procedure is repeated for samples from each p_T and η bin, and performed similarly for a sample of $Z(\mu\mu)$ events to measure muon identification efficiencies. The measured efficiencies are then compared to values extracted from simulated $Z(\ell\ell)$ events, with the same procedure applied. Each pair of these efficiencies are used to

form data/MC ratios which can be used to correct other simulated events samples for residual data/simulation differences. The data/MC efficiency ratios measured for electrons and muons (using the identification criteria from section 10.4) are summarized in figure 6.2.
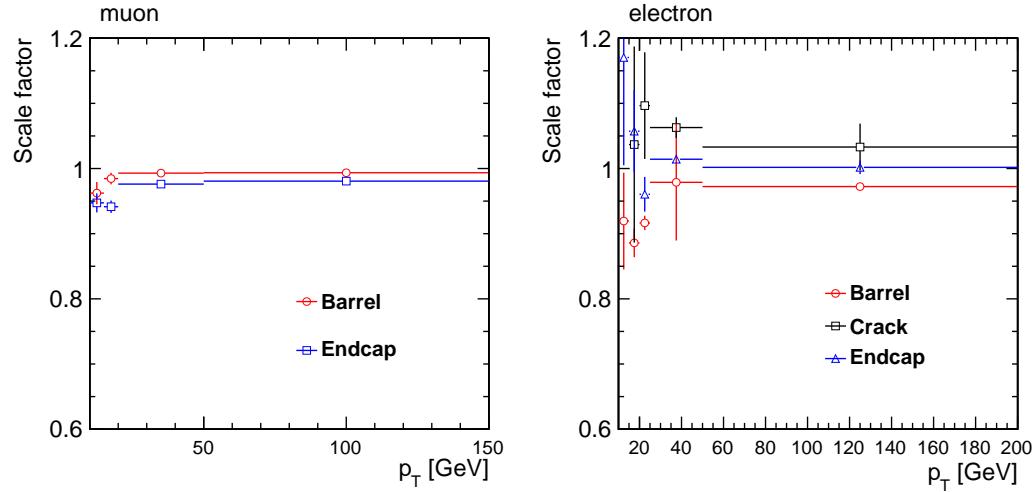
Figure 6.2: Offline lepton selection efficiency data/MC scale factors measured in different p_T and η bins for (Left) muons and (Right) electrons in the full 1.55 fb^{-1} dataset.

6.2 W/Z Production with Associated Jets

With the combined performance of all the subdetectors, CMS has the ability to efficiently select on-shell W and Z bosons in collision events through their decays to leptons and neutrinos. This well-controlled signature can be used to tag events with these bosons and measure other properties of the events in a largely unbiased way, such as the study of jet production in association with W or Z bosons, providing stringent tests of perturbative QCD calculations. Such a test can be performed through the measurement of the $W(Z) + n$ jets cross section, for different n. At present, next-to-leading-order (NLO) predictions are available for n up to four [67–70] and indicate that that these cross section exhibit Berends-Giele scaling with increasing n [71–73].

This hypothetical scaling follows

$$C_n \equiv \frac{\sigma_n}{\sigma_{n+1}} = \alpha \,, \tag{6.1}$$

where $\sigma_n = \sigma(W(Z)+ \geq n$ jets$)$, and α is a constant. At tree level, this behavior can be explained by noting that each extra strong emission comes with an additional factor of α_s, the QCD coupling. Hence the ratio of cross sections for different numbers of emissions should follow a power law related scaling as α^n. Recent works show that this scaling is not altered by fixed order QCD corrections [67, 68] and previous experimental measurements [74–78] have shown no significant deviation from scaling. The $W(Z) + n$ jets cross section is not only of theoretical importance, but also of practical importance for new physics searches where this process constitutes a large background, such as those described in chapters 9 and 10.

6.2.1 $\sigma(\mathbf{V}+ \geq \mathbf{n}$ jets$)$ **Cross Section Measurement Strategy**

The CMS measurement of the W and $Z + n$ jet cross sections [79] is performed on 36 pb^{-1} of 2010 $\sqrt{s} = 7$ TeV collision data. In this analysis, the high-p_T electrons and muons from W and Z decays are used to trigger and select events. The lepton selection is identical to that described in section 9.4 for the 2010 search for supersymmetry, for this measurement requiring that each event contains a lepton has a $p_T > 20$ GeV/c. The efficiency and uncertainties for triggering, reconstructed and identifying leptons are evaluated using an inclusive (with respect to jet activity) sample of $Z(\ell\ell)$ events with a tag-and-probe technique like that described in section 6.1.

The identification requirements on leptons significantly reduce backgrounds from non-W/Z contributions, particularly isolation requirements which reject events where leptons are located near other particle activity in the detectors, indicative of leptons coming from decays in jets. Despite a large rejection factor, a portion of these events survive the selection. These other backgrounds have a different rate dependence with requirements on the number of reconstructed jet and, as a result, would bias the W and $Z + n$ jet cross section measurements if not properly accounted for. To
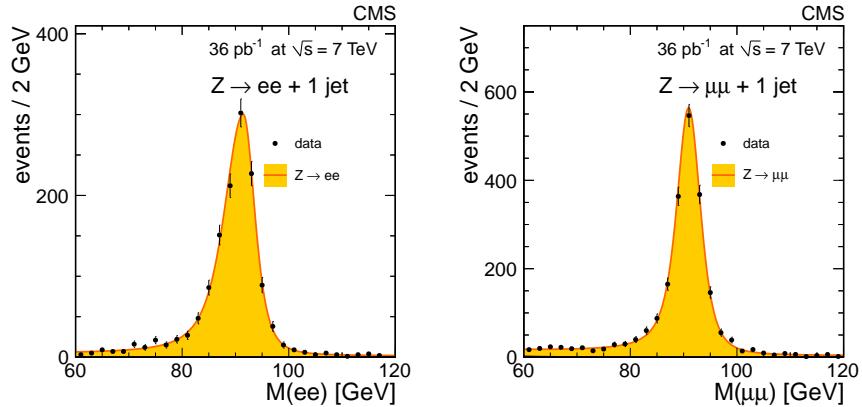
Figure 6.3: Di-lepton invariant mass for the $Z + 1$ jet samples in the electron channel (Left) and the muon channel (Right). The fit result for the signal is indicated by the yellow-filled area while the backgrounds are too small to be seen for this jet multiplicity.

discriminate between W/Z+jets events and these backgrounds we use the known masses of the W and Z bosons.

$Z(\ell\ell)$+jets event cross sections are measured using an extended maximum-likelihood fit to the di-lepton invariant mass ($M(\ell\ell)$) distribution. As for the tag-and-probe analysis, the invariant mass provides powerful discrimination between $Z(\ell\ell)$ events, which peak with a narrow width at the Z-pole, and other backgrounds which, without on-shell Zs, populate a falling continuum. Example fits for $Z + 1$ jet samples are shown in figure 6.3. The contamination from background processes with hard, prompt leptons are dominated by $t\bar{t}$ and W+jets and comprises a relatively small yield, without a peaking structure. The shape of backgrounds without heavy bosons is determined from a control sample selected with inverted lepton identification and isolation criteria. These fits are performed independently for exclusive samples with different numbers of reconstructed jets, where jets are counted according to how many have $|\eta| < 2.4$ and $E_T > 30$ GeV, and are a distance greater than $\Delta R = 0.3$ from the reconstructed leptons in the event. The η acceptance is limited to the region covered by the silicon tracker, restricting the analysis to jets that can benefit from tracker information ffor more precise measurements of jet energy.

Measuring the cross sections of $W(\ell\nu)$+jets events is more difficult than for Z

bosons for several reasons. With only one lepton in the final state, backgrounds are naturally larger without a second lepton requirement to suppress them. Furthermore, the decay neutrino is weakly interacting and escapes the detector unseen. This means that the invariant mass of the W cannot be measured directly. There are also other backgrounds containing W bosons, namely events with top quarks which decay to Wb. These are predominantly $t\bar{t}$+jets events, and having at least two jets in the event means this background grows, relative to $W(\ell\nu)$+jets, with increasing jet multiplicity.

The first two challenges can be overcome by using an analogue to the invariant mass suitable for open final states. The measured MET in each event can be interpreted as the transverse momentum of the escaping neutrino and, with the measured lepton, is used to calculate the transverse mass, $M_T(\ell\nu)$:

$$M_T(\ell\nu) = \sqrt{2(|\vec{p}_T^{\,\ell}||\vec{\mathrm{MET}}| - \vec{p}_T^{\,\ell} \cdot \vec{\mathrm{MET}})} \, , \tag{6.2}$$

where $\vec{p}_T^{\,\ell}$ is the lepton transverse momentum and $\vec{\mathrm{MET}}$ is the measured missing transverse momentum. For $W(\ell\nu)$ events, this variable will have a Jacobian-peaking structure with a kinematic edge at the W mass, while other backgrounds fall as a continuum. The transverse mass is a precursor to the razor variables, described in chapters 9 and 10, which are mass-sensitive variables for studying events with multiple weakly interacting particles.

Of course, events with top quarks will also have a peaking M_T distribution. This background is controlled by counting the number of b-tagged jets, $n_{\mathrm{jet}}^{\mathrm{b-tagged}}$, appearing in the event. $W(\ell\nu)$+jets events contain predominantly light-flavored jets while events with top quarks will naturally have an enhancement in b-quarks. Hence, the total selected single lepton sample can be organized into two components, one which exhibits a peaking structure in $M_T(\ell\nu)$, dominated by W+jets and $t\bar{t}$, and another which does not, dominated by QCD multijet events. Similarly, events with top quarks naturally contain an enhancement of b-tagged jets while W+jets and QCD events do not. The yield of each of these three contributions is measured through a two-dimensional fit to the $M_T(\ell\nu)$ distribution and the number of b-tagged jets. Each dimension is used
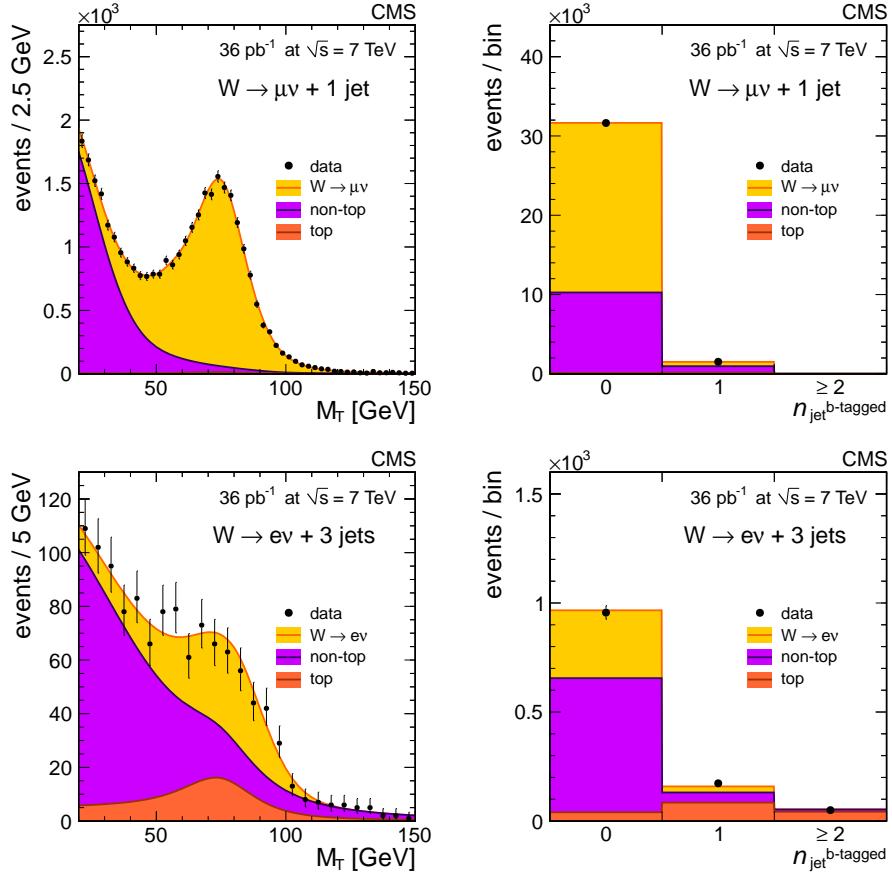
Figure 6.4: Fit results for the $W(\ell\nu) + n$ jets sample with $n = 1$ in the muon final state (Top) and $n = 3$ for the electron final state (Bottom). (Left) $M_T(\ell\nu)$ projection. (Right) Number of b-tagged jets projection. Fit results for each signal and background species in the fit are identified according to color.

to discriminate against each of the background types. The likelihood function is built under the assumption that there are few b-quark jets in the signal events, meaning that W events produced in association with heavy-flavor jets are attributed to background yields. Given the statistical precision of the measurement, this assumption has a negligible effect on the W+jets cross section result. This fits are performed using exclusive jet multiplicity bins for $n \leq 3$, with the last bin selected inclusively with $n \geq 4$. Examples of these fits are shown in figure 6.4.

6.2.2 Results

The yields measured for $Z(\ell\ell)$+jets and $W(\ell\nu)$+jets events, in different jet multiplicity bins, are corrected for lepton identification efficiencies which can vary with the number of jets due to isolation requirements. Many of the uncertainties from these measurements are reduced by considering the $V+n$ jets cross sections relative to the inclusive W and Z cross sections, $\sigma(V+ \geq n \text{ jets})/\sigma(V)$, and to the $(n\text{-}1)$ jets cross sections, $\sigma(V+ \geq n \text{ jets})/\sigma(V+ \geq (n-1) \text{ jets})$. An advantage of using ratios is that common uncertainties, like the absolute normalization from integrated luminosity measurements and the lepton selection and identification efficiencies, cancel. The measured $V+n$ jets cross sections are shown in figures 6.5 and 6.6 for W and Z bosons, respectively. These yields are not the absolute values measured in the $M(\ell\ell)$ and M_T distributions, but rather are *unfolded* quantities which account for the matrix which describes the probability of observing a certain number of reconstructed jets as a function of the true number of particle-jets in the event, derived from simulated events. This unfolding contributes to the systematic uncertainties of the measurements, along with jet-energy-scale uncertainties and associated pile-up effects.

Finally, these measurements can be used to test the Berends-Giele scaling hypothesis and measure the scaling parameter C_n from equation (6.1). To allow for deviation from constant scaling due to for example, phase-space effects, a second parameter, β, is introduced:

$$C_n = \alpha + \beta\, n \; . \tag{6.3}$$

The V+jets cross section yields are fit using this parameterization with the measured values of α and β shown in figure 6.7 for each of the different final stats (W/Z, electron/muon). The scaling expressed in equation (6.3) is not expected to hold for $n = 0$ due to the different production kinematics where no jets recoil against the vector boson, so it is not included the in the fit. The Berends-Giele scaling hypothesis is confirmed to work well up to the production of four jets, with the β parameter lying within one standard deviation from zero for both the W+jets and Z+jets cases.
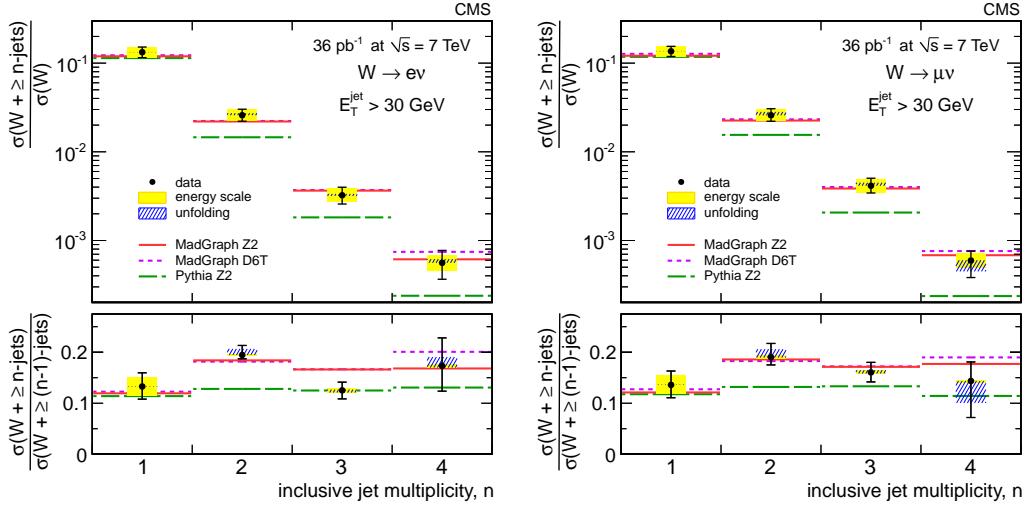
Figure 6.5: $W(\ell)+n$ jet cross sections. (Top) The ratios $\sigma(W+ \geq n \text{ jets})/\sigma(W)$ and (Bottom) $\sigma(W+ \geq n \text{ jets})/\sigma(W+ \geq (n-1) \text{ jets})$ in the (Left) electron and (Right) muon channels. Measured yields are compared with expectations from simulated events generated with MADGRAPH [80] and PYTHIA [213]. The uncertainties due to the energy scale and unfolding procedure are shown as yellow and hatched bands, respectively, while the error bars represent the total uncertainty.
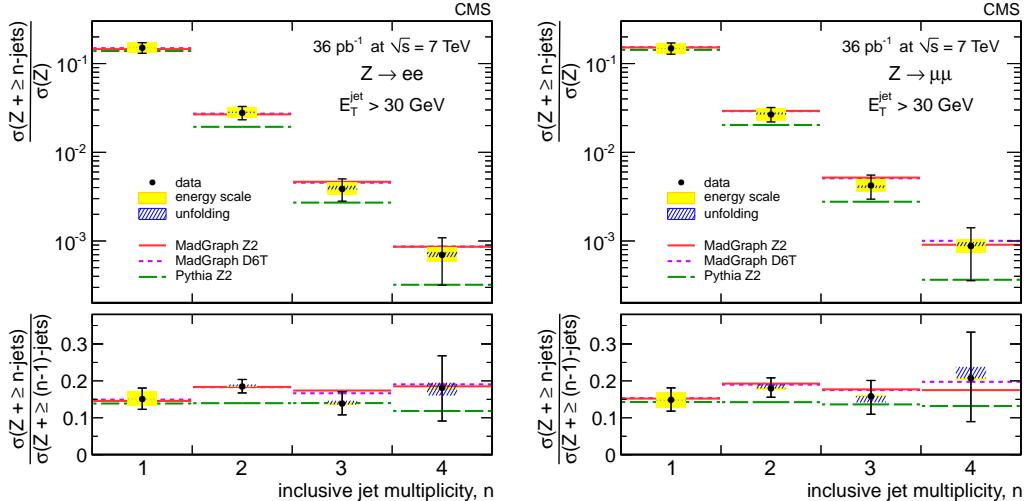
Figure 6.6: $Z(\ell\ell)+n$ jet cross sections. (Top) The ratios $\sigma(Z+ \geq n \text{ jets})/\sigma(Z)$ and (Bottom) $\sigma(Z+ \geq n \text{ jets})/\sigma(Z+ \geq (n-1) \text{ jets})$ (bottom) in the (Left) electron and (Right) muon channels. Measured yields are compared with expectations from simulated events generated with MADGRAPH [80] and PYTHIA [213]. The uncertainties due to the energy scale and unfolding procedure are shown as yellow and hatched bands, respectively, while the error bars represent the total uncertainty.
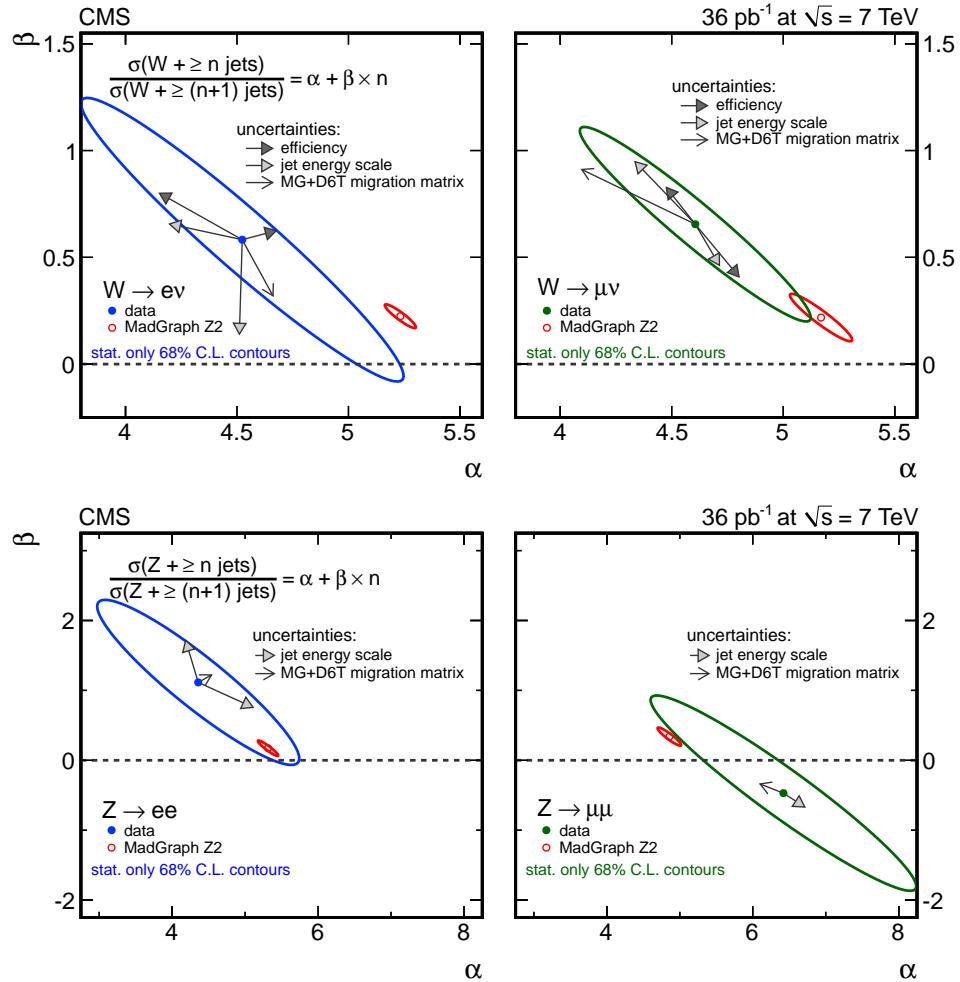
Figure 6.7: Fit results for the scaling parameters α and β after pileup subtraction, efficiency corrections, and unfolding of detector resolution effects. (Top) W+jets, (Bottom) Z+jets, (Left) electrons and (Right) muons. The ellipses correspond to 68% confidence level contours considering the statistical uncertainty only, for both data and simulation. The arrows show the displacement of the central value when varying each indicated parameter by its estimated uncertainty. The arrows labelled "MG+D6T migration matrix" correspond to the uncertainty from the unfolding procedure, which uses simulated events generated with MADGRAPH [80].

Chapter 7

Spinning the Higgs

On July 4th, 2012 the CMS and ATLAS experiments at the CERN LHC announced the discovery of a new particle. Both experiments observed a boson in the $\gamma\gamma$ and $Z(\ell\ell)Z^*(\ell\ell)$ final states, with each experiment indicating a discovery with signifiance exceeding 5 σ. CMS measures the mass of this new boson to be 125.3 ± 0.4 (stat) $\pm$ 0.5 (syst) while ATLAS finds 126.0 ± 0.4 (stat) ± 0.4 (syst), indicating a consistent observation between the experiments. This discovery comes in the context of a search for the SM Higgs boson and the cross section, decay channels and mass are in agreement, so far, with expectations from the Higgs. It may be that the final component of the SM has finally revealed itself.

Because the idea is so venerable, one may have grown insensitive to how special a Higgs boson would be. Its quantum numbers must be those of the vacuum, which its field permeates. Its couplings to the electroweak gauge bosons $W^\pm$ and Z are proportional to their masses, as are its couplings to quarks and leptons. Any deviation from the predicted quantum numbers or couplings of a putative Higgs boson would have deep ramifications for particle physics. After this discovery, the experimental program for Higgs physics must be focused on the rigorous determination of these fundamental quantities to confirm whether or not it is *the* Higgs.

We discuss here how the quantum numbers of a Higgs look-alikes (HLLs) can be measured. The study focuses on the so-called "golden channel" for Higgs physics, namely the Higgs decay $H \to ZZ^* \to 4\ell$, where Z^* denotes that one of the Zs may be strongly off-shell, as is the cases for the events attributed to the new boson in CMS

and ATLAS. This channel has the advantage that the kinematics of the Higgs and its decay products are fully reconstructible from a completely leptonic final state. The information in the corresponding decay angles can be used to infer the quantum numbers of the decaying H, and distinguish between SM Higgs bosons and an imposter. This study, performed in 2010 with with Alvaro De Rújula, Joseph Lykken, Maurizio Pierini and Maria Spiropulu [81] did not have the benefit of knowing the mass of a putative Higgs and so considers three test masses, 145, 200 and 350 GeV/c. Similarly, the study examined the possibility that an HLL could have spin 1. If the observed resonances in the $\gamma\gamma$ and 4ℓ final states correspond to the same bosons this possibility is excluded by the Landau-Yang theorem [82,83], and hence we will not discuss it in this context. Regardless, theories of physics beyond the SM predict that there could be other new bosons which decay through the golden channel, such that these test mass results and spin 1 cases would still apply if there are additional discoveries. We will see that early CMS studies of the new bosons quantum numbers, applying the techniques from this study, are in relatively good agreement with the expectations from the $m_H = 145$ GeV/c^2 test case described here.

In our analysis we compare a SM Higgs signal to a variety of Higgs look-alikes. We consider the most general Lorentz invariant couplings of a massive, spinless boson to ZZ or ZZ^*; this corresponds to gauge-invariant couplings up to dimension six. Some of the corresponding HLLs can be considered as modifications of the SM Higgs properties via P or CP violation or Higgs compositeness. Another spin 0 HLL corresponds to a new massive pseudoscalar, a particle occurring in models with extended Higgs sectors such as supersymmetry. We also discuss as one of our HLLs a massive spin 2 resonance coupling to the ZZ energy-momentum tensor, not necessarily with the universality of a graviton-like coupling. Although universally-coupled massive gravitons are already experimentally excluded in the relevant mass range [84], general spin 2 HLLs are a natural example of our study of spin discriminations.

7.1 The Golden Channel: $\mathbf{H \to ZZ^* \to 4\ell}$

The purpose of this study is to quantify the degree to which one can discriminate a SM Higgs boson from HLLs at, or close to, the moment of discovery at the LHC. There is a vast literature about determining Higgs properties from signals in a variety of final states (for a review, see [85]), but this research mostly addresses only the related question of whether it is possible *at all* to determine Higgs quantum numbers and couplings at a hadron collider. The current situation in this respect is similar to the LHC experimental program for supersymmetry, where only recently are there quantitative studies of the potential to discriminate supersymmetry look-alikes at the moment of discovery [86–90].

Our study focuses on the so-called "golden channel" for Higgs physics, namely the Higgs decay $H \to ZZ^* \to \ell_1^+ \ell_1^- \ell_2^+ \ell_2^-$, where $\ell_{1,2}^\pm$ denotes an electron or a muon, and Z^* denotes that one of the Zs may be strongly off-shell. Approximately half of the events will be $\mu^+ \mu^- e^+ e^-$, where all four leptons are easily distinguishable, and even in the 4μ and $4e$ final states all four leptons can be distinguished by the requirement that one or both Z bosons are reconstructed within an on-shell mass window. A well-measured, four-body, closed kinematic final state provides many independent observables for determining properties of the observed resonance; thus this channel provides more information than, e.g., the Higgs decay into two photons, where the photon polarizations are not measured.

We factorize the HLL problem into observables related to production and observables related to decay. In this paper we perform a systematic analysis including all of the information from the putative Higgs decays, leaving the analysis of Higgs versus HLL production to later work. While this factorization of production and decay is not completely clean, we find that the resulting model-dependent uncertainty introduced into the decay analysis is small. A full analysis will include production information and could produce stronger results than those presented here, since large cross section differences are expected between SM Higgs production and the production of many Higgs look-alikes. However, including Higgs and HLL production also introduces

new theoretical and measurement uncertainties and assumptions involving associated hadronic jets and the parton distribution functions that describe the initial state.

The analysis depends on five distinct angles that describe the $H \to ZZ^* \to 4\ell$ decay process. In the case where one of the Z bosons is strongly off-shell, the SM Higgs versus HLL decays also differ in their dependence on the reconstructed Z^* invariant mass. Because we are interested in HLL discrimination with small data samples, at or near the moment of discovery, we need to use all of the decay information in the events, including not just the distributions but also the correlations between all five (or six) of the relevant observables. Previous analyses of the Higgs golden mode decay properties have examined the dependence on some of the relevant angular distributions [91–96] and have shown the potential for LHC measurements to discriminate a SM Higgs from look-alikes with different spin and parity assignments or CP properties [85, 94–110]. However, none of these studies utilized all of the decay information in the events, and all of them have ignored the effects of detector phase space sculpting of the angular distributions, which are accounted for here.

We will denote the putative Higgs and its mass by H and m_H, regardless of whether it is a SM Higgs or a look-alike. This notation is also used to describe background events, where the four-lepton object is treated as a Higgs or HLL in the sense that m_H stands for $m_{4\ell}$. Since the events are fully reconstructible the lab frame kinematics of the candidate H particles are known: their transverse momentum p_T, pseudorapidity η, and azimuthal angle. These three variables define the direction and boost from the lab frame to the H rest frame. All other observables can then be defined with respect to the H rest frame, as illustrated in figure 7.1.

The H azimuthal angle plays no physical role, while the p_T and η distributions influence the way the detector selects events, sculpting the distributions of the final-state lepton's directions and energies. Once an event is boosted back to the 4ℓ rest-system (the rest system of the two initial-state fusing partons), the memory of p_T and η is lost, modulo these phase space acceptance effects.

In the approximation that the final-state leptons are massless, 12 observables are measured per event. Since all 12 are well-measured there is no experimental reason

not to reexpress these in terms of whatever combinations most naturally capture the underlying physics. Thus we choose four observables to be m_H and the three production observables just described that define the H rest frame. The remaining eight observables are taken to be the two reconstructed masses of the Z bosons together with six decay angles defined with respect to the H rest frame.

In the H rest frame the reconstructed Z bosons are back-to-back. We label these bosons as Z_1, Z_2 and take the direction of Z_2 as defining the positive z-axis. Because of Bose symmetry, the labeling is arbitrary; in the case of an $e^+e^-\mu^+\mu^-$ final state we will follow the literature [106] and choose Z_2 to be the Z boson that decayed to muons. We then adopt the additional convention that the transverse direction of the μ^- lies along the positive y-axis; thus the Z_2 decay leptons lie in the y-z plane.

With the above choices, the reconstructed Z boson masses m_1 and m_2 also define the longitudinal boosts from the H rest frame to the rest frames of the decaying Z_1 and Z_2 bosons. The boost parameters are given by

$$\gamma_1 = \frac{m_H}{2m_1}\left(1 + \frac{m_1^2 - m_2^2}{m_H^2}\right) , \qquad (7.1)$$

$$\gamma_2 = \frac{m_H}{2m_2}\left(1 - \frac{m_1^2 - m_2^2}{m_H^2}\right) . \qquad (7.2)$$

We let θ_1, φ_1 denote the ℓ_1^- decay angles in the Z_1 rest frame, while θ_2, φ_2 denote the ℓ_2^- decay angles in the Z_2 rest frame.

There are two additional angles Θ, Φ defining the direction of the initial state partons as reconstructed in the H rest frame. For a gluon-gluon initial state these angles measure a rotation from the z-axis defined above to the direction of the initial state gluon with positive z-component of momentum. For quark-antiquark ($q\bar{q}$) initiated production of an HLL we have the problem that we do not know event-by-event which proton contributed the antiquark; this is resolved by symmetrizing the expected angular distributions under the replacement $\cos\Theta \to -\cos\Theta$.

As expected, one combination of the three azimuthal angles Φ, φ_1 and φ_2 is physically redundant. We take advantage of this fact to make the replacements $\varphi_1 \to \Phi + \phi$, $\varphi_2 \to \Phi$. Thus ϕ then represents the azimuthal rotation between the Z_2 and

Z_1 decay planes.

In summary, the 4-momenta of the process $gg \to H \to Z_1 Z_2 \to \ell_1^- \ell_1^+ \ell_2^- \ell_2^+$ are explicitly parametrized in the H rest frame as

$$
\begin{aligned}
p_{g_2} &= \frac{m_H}{2}\left(1, S\cos\Phi, S\sin\Phi, C\right), \\
p_{g_1} &= \frac{m_H}{2}\left(1, -S\cos\Phi, -S\sin\Phi, -C\right), \\
k &= m_H\left(1, 0, 0, 0\right), \\
p_2 &= m_2\left(\gamma_2, 0, 0, \beta_2\gamma_2\right), \\
p_1 &= m_1\left(\gamma_1, 0, 0, -\beta_1\gamma_1\right), \\
p_{\ell_2^-} &= \frac{m_2}{2}\left(\gamma_2(1+\beta_2 c_2), 0, s_2, \gamma_2(\beta_2+c_2)\right), \\
p_{\ell_2^+} &= \frac{m_2}{2}\left(\gamma_2(1-\beta_2 c_2), 0, -s_2, \gamma_2(\beta_2-c_2)\right), \\
p_{\ell_1^-} &= \frac{m_1}{2}\left(\gamma_1(1+\beta_1 c_1), -s\,s_1, -c\,s_1, -\gamma_1(\beta_1+c_1)\right), \\
p_{\ell_1^+} &= \frac{m_1}{2}\left(\gamma_1(1-\beta_1 c_1), s\,s_1, c\,s_1, -\gamma_1(\beta_1-c_1)\right).
\end{aligned}
\tag{7.3}
$$

Here k denotes the 4-momentum of H, while p_1, p_2 are the 4-momenta of Z_1, Z_2. We used the condensed notation $C, S = \cos\Theta, \sin\Theta$, $c, s = \cos\phi, \sin\phi$, $c_1, s_1 = \cos\theta_1, \sin\theta_1$, and $c_2, s_2 = \cos\theta_2, \sin\theta_2$.

Of the five relevant angles, Θ and Φ are Z-pair production angles, while the remaining three are 4ℓ production angles. We will use the notation

$$
\begin{aligned}
\vec{\Omega} &= \{\Phi, \cos\Theta\}, \\
\vec{\omega} &= \{\phi, \cos\theta_1, \cos\theta_2\}.
\end{aligned}
\tag{7.4}
$$

For a SM Higgs, the distributions in Θ and Φ are flat if we ignore the phase space acceptance effects inherent in any experimental analysis. In previous studies these two angles have typically been integrated over.

Although we have tried to conform to the literature in our parametrization of the decay angles, we note that the literature itself is divided over the choice of which decay plane orientation corresponds to $\phi=0$ rather than $\phi=\pi$. We conform to the

convention of Buszello et al. [106], which is opposite to that of Djouadi [85] and Bredenstein et al. [122].
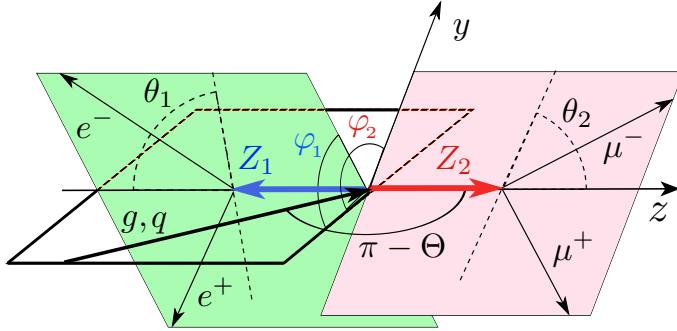
Figure 7.1: The Cabibbo-Maksymowicz angles [111] in the $H \to ZZ$ decays.

The decay amplitudes defined in the next section depend on two combinations of the boost parameters γ_1 and γ_2, defined by

$$\gamma_a \;=\; \gamma_1 \gamma_2 (1 + \beta_1 \beta_2) \,, \tag{7.5}$$

$$\gamma_b \;=\; \gamma_1 \gamma_2 (\beta_1 + \beta_2) \,, \tag{7.6}$$

which are in fact just the cosh and sinh of the rapidity difference of Z_2 and Z_1, such that

$$\gamma_a^2 - \gamma_b^2 = 1 \,. \tag{7.7}$$

More explicitly, we have

$$\gamma_a = \frac{1}{2 m_1 m_2} \left(m_H^2 - (m_1^2 + m_2^2) \right) \,. \tag{7.8}$$

7.2 The Quantum Numbers of the Higgs Boson and Other Possibilities

7.2.1 General Couplings to ZZ[*]

The vertex Feynman rules for the most general coupling of a spin-less particle to the polarization vectors ϵ_1^μ and ϵ_2^α of two Zs of four-momenta p_1 and p_2 are given by the expression

$$L_{\mu\alpha} = X\, g_{\mu\alpha} - (Y + i\,Z)\,\frac{k_\alpha k_\mu}{M_Z^2} + (P + i\,Q)\,\epsilon_{\mu\alpha}\frac{p_1 p_2}{M_Z^2}\,, \tag{7.9}$$

where we have suppressed repeated indices in the contraction of the four-index ϵ tensor, $k = p_1 + p_2$ and only Lorentz-invariance has been assumed. The dimensionless form factors X to Q are functions of k^2 and $p_1 \cdot p_2$ which, with no loss of generality, can be taken to be real (but for their absorptive parts, expected to be perturbatively small). The rescalings by $1/M_Z^2$ are just for definiteness, since the true mass scale of the underlying operators is as yet unspecified. In practice we also remove an overall factor of $ig M_Z / \cos\theta_W$, so that $X = 1$ corresponds to the tree level coupling of a SM Higgs boson.

Similarly, most general parity-conserving vertex describing the coupling of a $J = 2^+$ particle of polarization tensor $\epsilon^{\rho\sigma}$ to our two vector bosons is

$$\begin{aligned}
L^{\rho\sigma\mu\alpha} = \quad & X_0\, m_H^2\, g^{\mu\rho}\, g^{\alpha\sigma} \\
& + (X_1 + i\,Y_1)\,(p_1^\alpha\, p_2^\rho\, g^{\sigma\mu} + p_1^\rho\, p_2^\mu\, g^{\sigma\alpha}) \\
& + (X_2 + i\,Y_2)\, p_1^\rho\, p_2^\sigma\, g^{\mu\alpha},
\end{aligned} \tag{7.10}$$

where we have dropped contributions that have more than two derivatives or are odd under parity, and again with all coefficients real. The special case of tree level graviton-like couplings corresponds to

$$X_0 = -\frac{1}{2}\kappa\,, \quad X_1 = \kappa\,, \quad X_2 = -\kappa\,, \tag{7.11}$$

with all other coefficients vanishing and κ an overall coupling strength.

These general couplings, with naive mass dimensions $d = 3$, 4, and 5, can arise from $SU(2)_L \times U(1)_Y$ invariant operators of dimension 5, 6, or higher. Since, for HLLs with nonvanishing weak charges this parentage introduces model dependence, we relegate it to a brief discussion in Appendix $A.2$.

7.2.2 "Pure" Cases of Specified $\mathbf{J^{PC}}$

We specify in this section the results for four cases (scalar, pseudoscalar, vector and axial vector) that would be "pure" in the sense of having a single dominant term in their HZZ couplings, which we use to define their spin and parity. This allows one to illustrate the mass and angular dependences of the predictions, setting the stage for the later discussion of the impure cases for which P and/or CP are not symmetries of the theory, and to establish comparisons with the existing literature.

The general expressions for the angular correlations in the ZZ^* case (which includes ZZ when the two Z masses are fixed at M_Z) are given in section A.4, where

$$\eta \equiv \frac{2\, c_v\, v_a}{(c_v^2 + c_a^2)} \simeq 0.15, \tag{7.12}$$

denotes the quantity arising from the SM couplings of the Z bosons to the final state leptons.

7.2.2.1 The Standard Higgs, $\mathbf{J^{PC} = 0^{++}}$

The tree level SM coupling of the Higgs to two Zs of polarization ϵ_1 and ϵ_2 is $\propto \epsilon_1 \cdot \epsilon_2$ (see equation (7.9)). The angular distribution of the leptons in $H \to ZZ \to 4\,l$ decay, for on or off-shell Zs of mass m_1 and m_2, is

$$\frac{d\Gamma[0^+]}{dc_1\, dc_2\, d\phi} \propto m_1^2\, m_2^2\, m_H^4 \big[1 + c_1^2 c_2^2 + (\gamma_b^2 + c^2)s_1^2 s_2^2$$
$$+ 2\gamma_a\, c\, s_1 s_2\, c_1 c_2 + 2\eta^2(c_1 c_2 + \gamma_a\, c\, s_1 s_2)\big]. \tag{7.13}$$

7.2.2.2 A Pure Pseudoscalar, $\mathbf{J^{PC} = 0^{-+}}$

The coupling of a $J^{PC}=0^{-+}$ pseudoscalar to two Zs of polarization ϵ_1 and ϵ_2 and four-momenta p_1 and p_2 is proportional to $\epsilon[\epsilon_1, \epsilon_2, p_1, p_2]$, (see equation (7.9)). The angular distribution of the leptons in its $ZZ \to 4l$ decay is

$$\frac{d\Gamma[0^-]}{dc_1\, dc_2\, d\phi} \propto m_1^4\, m_2^4\, \gamma_b^2$$
$$\left(1 + c_1^2 c_2^2 - c^2 s_1^2 s_2^2 + 2\,\eta^2\, c_1 c_2\right) . \tag{7.14}$$

7.2.2.3 A Pure Massive Graviton, $\mathbf{J^{PC} = 2^{++}}$

Since the general analysis of spin 2 coupling to off-shell Zs is quite cumbersome, we will only quote results for the example of a positive parity spin 2 with graviton-like couplings produced by gluon fusion and decaying to two on-shell Z's. Defining the on-shell ratio $x \equiv m_H/M_Z$ and using the massive graviton formalism of [112], we obtain the tree level angular distribution

$$\frac{d\Gamma[gg \to \text{graviton} \to ZZ]}{dC\, dc_1\, dc_2\, d\Phi\, d\phi} \propto 16x^4C^2 + 2(x^4 + 16)S^4 + s_1^2 s_2^2[(x^4 + 16)S^4 - 4x^2(x^2 + 4)S^2 + 4x^4]$$
$$+8x^2 S^2\Big[[2 + S^2 + (2 - 3S^2)c_2^2]s_1^2\cos(\Phi + \phi)^2 + [2 + S^2 + (2 - 3S^2)c_1^2]s_2^2\cos^2\Phi\Big]$$
$$+S^4 s_1^2 s_2^2[x^4\cos(2\Phi + \phi)^2 + 16\,c^2] - (s_1^2 + s_2^2)[(x^2 + 4)^2C^4 + 2(3x^4 - 16)C^2 + (x^2 - 4)^2]$$
$$+2S^2 c_1\, c_2\, s_1\, s_2\, \Big[x^2\,[2(x^2 + 4) - (x^2 + 12)S^2]\cos(2\Phi + \phi) + 4\,[4x^2 - (3x^2 + 4)S^2]c\Big] . \tag{7.15}$$

We note the $\cos^4\Theta$ dependence characteristic of a spin 2 resonance.

7.2.3 Tests of Symmetries

Now we discuss the behavior of the HZZ couplings under various symmetries, including CP and Bose-Einstein statistics. The discussion attempts to clarify the literature on these issues.

Consider the $J=0$ case. The most general coupling of a spin-less particle to the polarization vectors ϵ_1 and ϵ_2 of two Zs is that of equation (7.9). In computing the ensuing $H \to ZZ^* \to 4\ell$ process one finds that the XP interference term is of the

form:

$$\frac{d\Gamma[0, \text{Todd}]}{dc_1 \, dc_2 \, d\phi} \propto$$
$$2 \, m_1^3 \, m_2^3 \, m_H^2 \, \gamma_b \, s_1 \, s_2 \, s \left[s_1 \, s_2 \, c + \gamma_a \left(c_1 c_2 + \eta^2 \right) \right] , \qquad (7.16)$$

where the term $\sin \theta_1 \sin \theta_2 \sin \phi \propto \vec{p}_{e^+} \cdot \vec{p}_{\mu^-} \times \vec{p}_{\mu^+}$. By definition, this observable is $\tilde{T}$-odd: it changes sign as all three-momenta are reversed (the tilde in "$\tilde{T}$-odd" emphasizes that past and future are not being interchanged).

The Born approximation is, by definition, the result of squaring the amplitude dictated by the Lagrangian to lowest order in its couplings: a quadratic result, in our case, in any pair of the quantities X to Q in equation (7.9). To this order, a $\tilde{T}$-odd observable must vanish if CP is a symmetry, as shown in [113]. Thus, a nonvanishing $\tilde{T}$-odd observable such as that of equation (7.16) can only arise if CP-invariance is violated.

The XQ interference term resulting from equation (7.9) is

$$\frac{d\Gamma[0, \text{Codd}]}{dc_1 \, dc_2 \, d\phi} \propto$$
$$-2 \, \eta \, m_1^3 \, m_2^3 \, m_H^2 \, \gamma_b \left[c_1 + c_2 \right] \left(1 + c_1 c_2 + \gamma_a \, s_1 s_2 \, c \right) . \qquad (7.17)$$

This term is CP odd and $\tilde{T}$-even, a combination not addressed by the theorem quoted above. It is a C-odd observable, in that it changes sign under the interchange of $p_{e^+} \leftrightarrow p_{e^-}$ and $p_{\mu^+} \leftrightarrow p_{\mu^-}$, tantamount to $\cos \theta_i \leftrightarrow -\cos \theta_i$ in our chosen notation.

7.2.4 Tests of Compositeness

If the couplings of an HLL conserve P and CP, but the object is not point-like, there will be deviations from the standard $g_{\mu\nu}$ coupling to Zs. To lowest order in the dimensions of the corresponding effective operators, these will be of two types. The first is a nonvanishing Y in equation (7.9), and the second is a nontrivial form for X. Barring large effects (quite conceivable in a model with multiple SM Higgs-like fields) deviations in X are much harder to limit or measure than a nonzero

Y/X which is governed by the shapes of angular distributions. Contributions to Y can arise from gauge invariant operators of dimension 5 containing a non SM-like spin 0 HLL (Appendix A.2) or from higher dimension operators containing the SM Higgs [114–116].

It is useful to introduce the notation $\tan \xi \equiv Y/X$. In this notation, the "composite" HLL angular distribution is of the form:

$$d\Gamma_C = \cos^2\xi \, d\Gamma_{XX} + \cos\xi \sin\xi \, d\Gamma_{XY} + \sin^2\xi \, d\Gamma_{YY} \,, \tag{7.18}$$

where $d\Gamma_{XX}$ is the standard result of equation (7.13). The interference term is

$$\frac{d\Gamma_{XY}}{dc_1 \, dc_2 \, d\phi} \propto \\ -2 \, m_1^3 \, m_2^3 \, m_H^2 \, \gamma_b^2 \, s_1 \, s_2 \, (c_1 c_2 c + \gamma_a \, s_1 s_2 + \eta^2 \, c) \,, \tag{7.19}$$

and the last term is

$$\frac{d\Gamma_{YY}}{dc_1 \, dc_2 \, d\phi} \propto m_1^4 \, m_2^4 \, \gamma_b^4 \, s_1^2 s_2^2 \,. \tag{7.20}$$

Contrary to all of the other cases we study, the interference term in this instance is between two operators whose P and C are identical: the HLL is not point-like, but it is 'pure' 0^{++}. As a consequence, the angular distribution of the interference term is not very different from that of the XX and YY terms and the interference can, for certain values of Y/X, be very destructive. This can be seen even at the level of the $H \to ZZ$ branching fraction, the integral of equation (7.18) over $\cos\theta_1$, $\cos\theta_2$, and ϕ,

$$\Gamma_C \propto m_1^2 m_2^2 \left[2\cos^2\xi + (\gamma_a \cos\xi - m_1 m_2 \gamma_b^2 \sin\xi)^2 \right]. \tag{7.21}$$

If ξ has a value close to the (mass-dependent) point of maximal interference, the golden mode channel can be suppressed by a large factor. For this to happen X and Y ought to be of the same order of magnitude, signifying a low dynamical scale for a composite Higgs.

7.3 Measuring Higgs Quantum Numbers

7.3.1 Statistical Approach

In this section we discuss the statistical formulation we use to address comparisons between different hypotheses as well as relevant measurements for the characterization of an HLL resonance. We focus on four statistical approaches:

- (1) Comparisons between two "pure" spin-parity hypotheses (such as 0^+ vs. 0^-).

- (2) Comparisons between two spin-parity hypotheses, with at least one of the two being an "impure" admixture of two pure HLL states (e.g., 0^+ vs. a combination of 0^+ and 0^-). This case is similar to (1), except for the presence of one or more nuisance parameters.

- (3) The measurement of mixing parameters in the case of impure Higgs look-alikes.

Each of these cases involves attempting to establish the nature of a newly discovered particle. As described in section A.1, statistical subtraction techniques based on a fit to the four-lepton invariant mass distribution can be used to effectively remove background events from the same. In this study, measurements of the Higgs quantum numbers are performed on samples composed exclusively of signal events, with background assumed to have been removed. This is judged to give a good approximation to an actual experimental analysis.

The cases (1) and (2) involve tests between different J^P interpretations for signal events appearing in the four-lepton resonance. In the (1) scenario the two hypotheses under consideration are simple, i.e., the corresponding likelihoods are fully specified once the values $\vec{X}$ are fixed. In the (2) case the unknown mixing angles for the impure hypothesis, referred to as $\vec{\xi}$ (and including, e.g., various mixing angles), are treated as nuisance parameters. The analysis in case (3) is a traditional parameter estimate, based on the ML fit, for which we obtain a confidence interval by using the Feldman-Cousins approach [117]. We discuss the three cases starting from the last.

7.3.1.1 Coupling Admixtures

Consider the example of a one-parameter mixture of two types of HZZ coupling, such as the composite case discussed in section 7.2.4. For a fixed value of the resonance mass m_H and the mixing angle ξ, equation (7.18) is the theoretical probability distribution of the events as a function of the variables $\vec{X}$ for ZZ and ZZ^* final states. The experimental *pdf* is a numerical representation of the result of sieving (with a specific detector and its resolution, trigger and analysis requirements) a very large number of events, generated with the theoretical *pdf* of equation (7.18). This experimental *pdf*, referred to as P, is a function $P=P_{m_H}(\xi, \vec{X})$ of m_H, (which is kept fixed through this exercise), ξ, and $\vec{X}$. The dependence on $\vec{\Omega} \equiv \{\cos\Theta, \Phi\}$ is, in this example, exclusively a phase space acceptance effect.

Many experiments with a fixed number of events N_S are simulated, assuming the same detector response. The probability of each event, evaluated with the experimental *pdf*, is P_i. The likelihood of a given experiment is $\mathcal{L}(\xi) = \prod_{i=1}^{N_S} P_i$. The experimentally measured value of the ξ parameter, $\hat{\xi}$ corresponds to the value that maximizes $\mathcal{L}(\xi)$. The simulation is repeated many times, as a function of the true value of the mixing angle ξ. Running many experiments one can derive the confidence interval, i.e. the range covering the true value of ξ for some confidence level and some measured value $\hat{\xi}$ [117].

It is customary to estimate the error (or the number n of standard deviations σ) in the measured ξ from the expression $\mathcal{L}(\xi_{max} \pm n\,\sigma) = \mathcal{L}(\xi_{max}) - n^2/2$. While this method is accurate for large samples with Gaussian errors, it is not the one used to draw the σ contours in figure 7.2 (where $\xi=\xi_{XQ}$ as given in equation (7.25) and in the similar figures of section 7.3.2). Instead, the confidence level (CL) is evaluated measuring the frequency of a given result in the set of generated pseudoexperiments.
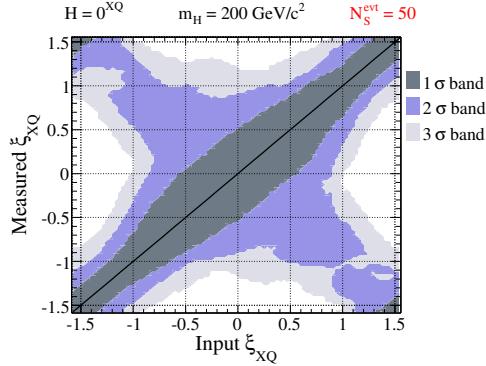
Figure 7.2: Confidence intervals for measured values of ξ_{XQ} for a C-violating $J{=}0$ resonance with a mass 200 GeV/c^2.

7.3.1.2 Confronting J^P hypotheses

Consider two hypotheses, $\mathbb{H}_{0,1}$, for the spin-parity assignment of a signal candidate sample, detected via its ZZ mass peak and background-subtracted using the $_sPlot$ method. Large numbers of events are generated assuming each hypothesis and used to construct two unbinned experimental *pdfs*: $P_{\mathbb{H}_{0,1}} \equiv P_{m_H}(\vec{X} \mid \mathbb{H}_{0,1})$. For our pure spin-parity cases, the simple nature of the hypotheses considered guarantees through the Neyman-Pearson (NePe) lemma [118] that the hypothesis test is *universally most powerful*. Next, we explicitly identify one hypothesis as $\mathbb{H}_0$ and the other as $\mathbb{H}_1$. Additionally, we specify the test *statistic* Λ which we define as the log-likelihood ratio $\log[\mathcal{L}(\mathbb{H}_1)/\mathcal{L}(\mathbb{H}_0)]$. Finally, we must a priori choose the acceptable probability level α of rejecting $\mathbb{H}_0$ in favor of $\mathbb{H}_1$, even though $\mathbb{H}_0$ is true (Type I error). We generate a series of pseudoexperiments with a fixed number of events N_S to construct the *pdf* of Λ for the two hypotheses. A typical result is illustrated in figure 7.3. We first generate pseudoexperiments considering $\mathbb{H}_0$ as true. For each experiment we construct two likelihoods $\mathcal{L}(\mathbb{H}_0) \equiv \prod_{i=1}^{N_S} P_{\mathbb{H}_0}(\vec{X}_i)$ for the correct interpretation of the true theory, and $\mathcal{L}(\mathbb{H}_1) \equiv \prod_{i=1}^{N_S} P_{\mathbb{H}_1}(\vec{X}_i)$ for its incorrect interpretation. With the ensemble of experiments one constructs the distribution $P(\Lambda \mid \mathbb{H}_0)$ with $\Lambda \equiv \log[\mathcal{L}(\mathbb{H}_1)/\mathcal{L}(\mathbb{H}_0)]$. The result is the leftmost (red) curve in figure 7.3. The exercise is repeated with the pseudoexperiments generated considering $\mathbb{H}_1$ as true and the result is the rightmost

(blue) curve in the figure. An a priori chosen value of α implicitly defines a value $\hat{\Lambda}(\alpha)$ via

$$\alpha = \int_{\hat{\Lambda}(\alpha)}^{\infty} P(\Lambda \,|\, \mathbb{H}_0) \, d\Lambda \; . \tag{7.22}$$

This fixed value $\hat{\Lambda}(\alpha)$ implies that

$$\beta(\alpha) = \int_{-\infty}^{\hat{\Lambda}(\alpha)} P(\Lambda \,|\, \mathbb{H}_1) \, d\Lambda \tag{7.23}$$

is the probability of accepting $\mathbb{H}_0$ even though $\mathbb{H}_1$ is correct (Type II error). The value $1 - \beta$ is called the *power* of the test. When the real experiment is performed, a specific value Λ_{exp}, is obtained for Λ. The associated *p-value* $= \int_{\Lambda_{\text{exp}}}^{\infty} P(\Lambda \,|\, \mathbb{H}_0) \, d\Lambda$, is compared to α to determine if the measurement favors one hypothesis versus the other.

Instead of the α and β values, the significance σ is commonly used. To convert to an equivalent number of σ's using figure 7.3 we calculate the same α-area in a Gaussian distribution centered at 0 with $\sigma=1$. The number n of α-equivalent standard deviations is obtained by inverting

$$\alpha = \frac{1}{\sqrt{2\pi}} \int_n^{\infty} dx \, e^{-x^2/2} \; . \tag{7.24}$$
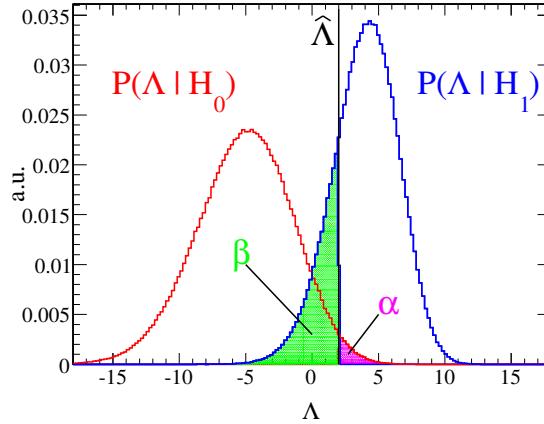
Figure 7.3: Distribution of Λ for $m_H=200$ GeV/c^2 and $N_S=23$, constructed with $\sim 10^9$ pseudoexperiments. The hypotheses being confronted are $\mathbb{H}_0=0^+$ and $\mathbb{H}_1=0^-$.

The a priori (subjective) choice of α (and subsequently β and corresponding significances) is heavily discussed in the literature. *The Physical Review*, for example, requires a 5σ (3σ) significance to claim discovery (evidence). The caveat is, of course, that when one minimizes as much as possible the probability of an error of Type I (wrongly claiming a discovery) one risks making an error of Type II (and, e.g., delaying the claim of a discovery to the next luminosity upgrade).

A pure vs. impure HLL hypothesis test has an additional complication due to the dependence of the likelihood function on the mixing angles $\vec{\xi}$ in at least one of the two hypotheses. In this case, we are testing the simple (i.e. mixing angle independent) hypothesis against a class of alternative hypotheses, connected by the variation of a continuous unknown parameter(s). The test is performed by comparing the simple hypothesis to the impure hypothesis with values of $\vec{\xi}$ that best fit the data.

The impure vs. impure Higgs look-alike test is technically identical to the pure vs. impure. Here, we try to exclude some value of the mixing angle parameter for one of the two composite hypotheses in favor of the alternative impure hypothesis, where the mixing angles are treated as nuisance parameters. With fixed mixing angles, one impure look-alike becomes a simple hypothesis (like a pure one) tested against an impure hypothesis.

7.3.2   Results

We present results for three HLL masses: m_H=145, 200, and 350 GeV/c^2, using pseudoexperiments built with the full $\vec{X}$ *pdf*.

7.3.2.1 0^+ vs. 0^-

We consider here two different "pure" scalar hypotheses: 0^+, corresponding to a SM Higgs, and 0^-, a pseudoscalar. Neither of these possibilities has an explicit dependence on the angles $\vec{\Omega}$ in their differential cross section, meaning that only the variables $\vec{\omega}$ (and the off-shell Z mass, $m_2=M_{Z^*}$, for $m_H < 2M_Z$) are used to discriminate between the two hypotheses.
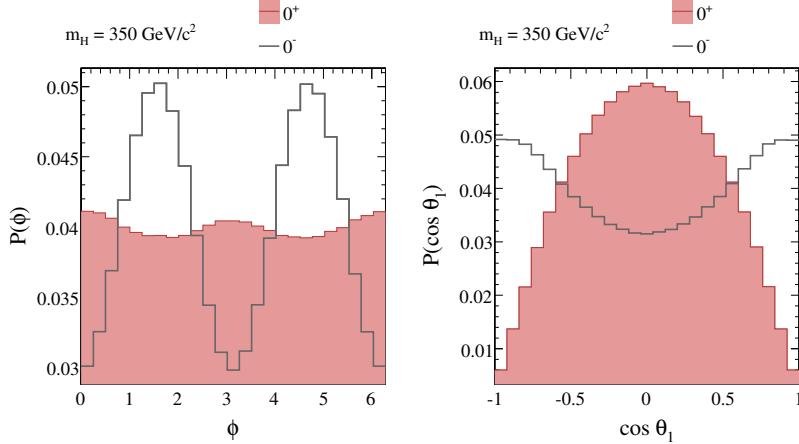
Figure 7.4: Distributions of the variables ϕ (left) and $\cos\theta_1$ (right) for 0^+ and 0^- resonances with m_H=350 GeV/c^2. All distributions are normalized to a unit integral.

In figure 7.4 we show the distributions in ϕ and $\cos\theta_1$ at m_H=350 GeV/c^2 for J^P=0^+ and 0^-. These angular variables (along with $\cos\theta_2$, whose distribution is identical to that of $\cos\theta_1$ except when Z_2 is off-shell) provide the discrimination between these two hypotheses at all masses m_H. For masses m_H below the $2\,M_Z$ threshold, the kinematic factors in equations (7.13) and (7.14) result in the differential cross section dependences on the off-shell Z mass M_{Z^*} that differ for the 0^+ and 0^- cases. This is illustrated in figure 7.5 (left) for m_H=145 GeV/c^2. For all the discriminating variables we consider, the ability to distinguish between two hypotheses is degraded when their correlations are neglected. This is shown in figure 7.5 (right) where we present the results of the NePe hypothesis test between 0^+ and 0^- for likelihoods built using different subsets of variables and correlations thereof. Specifically $P(M_{Z^*}, \vec{\omega})$ denotes the use of the full set of variables while in $P(\vec{\omega})$ the probability distribution of M_{Z^*} is ignored. The product of all one-dimensional probabilities, ignoring correlations, is $\prod_i P(X_i)$. As expected, the likelihood including all discriminating variables and their correlations is optimal. The other two definitions give similar results. We note that, regardless of the results, the use of $\prod_i P(X_i)$ is an improper approximation, since the X_i variables are far from being uncorrelated.

The significance for discriminating between the 0^+ and 0^- hypotheses (assuming
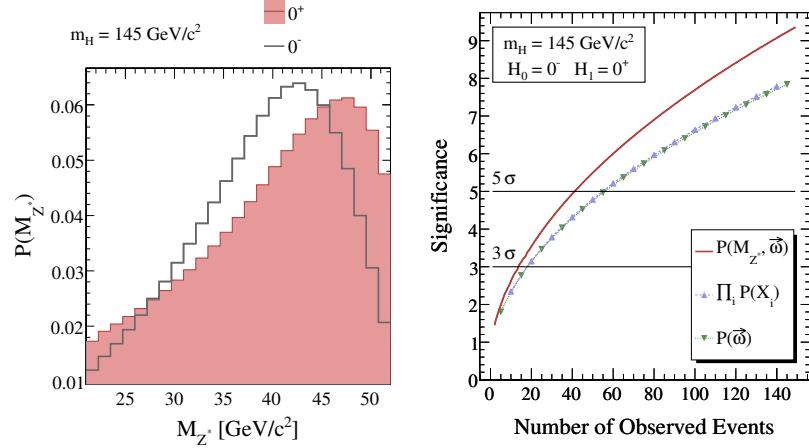
Figure 7.5: (Left) Distribution of M_{Z^*} for 0^+ and 0^- $H \to ZZ^*$ decays at m_H=145 GeV/c^2, normalized to a unit integral. (Right) Median significance for rejecting 0^- in favor of 0^+, assumed to be correct, as a function of N_S. The different likelihood constructions are specified in the text.

one or the other to be correct), as a function of N_S, where N_S is the number of observed signal events, is shown in figure 7.6 for m_H=145 GeV/c^2. In all cases, results correspond to the case where $\mathbb{H}_1$ is the true hypothesis (see section 7.3.1). The model discrimination is based on a NePe test between these simple hypotheses with test statistic $\log(\mathcal{L}[0^+]/\mathcal{L}[0^-])$. The variables $\vec{\omega}$ (and M_{Z^*}, when applicable), along with their correlations, are used in the likelihood construction. The significance for rejecting one hypothesis in favor of the other at around the time of $5\,\sigma$ excess in this single channel is better than $3\,\sigma$ for m_H=145, 200, and 350 GeV/c^2 while a $5\,\sigma$ discrimination can be achieved with twice the observed signal events (less than $\sim$40 events in both mass cases presented here).

7.3.3 0^+ vs. 2^+

We consider one "pure" spin 2 model: a J=2^+ heavy graviton-like resonance. A J=2 object has *pdfs* with nontrivial dependence on the angles $\vec{\Omega}$ up to quartic order in $\cos\Theta$. In figure 7.7 we show the corresponding distributions in the $\vec{\Omega}$ variables for m_H=200 and 350 GeV/c^2. The ability to discriminate between the 0^+ and J=2 hypotheses improves with increasing resonance mass. Despite the presence of quartic
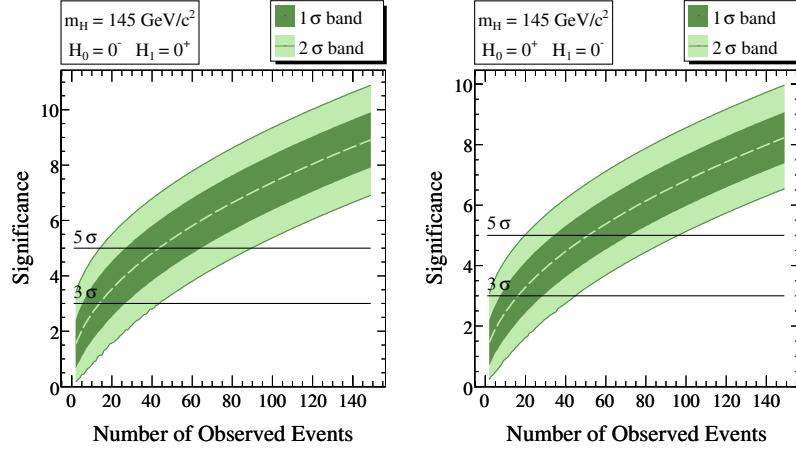
Figure 7.6: Significance for rejecting 0^- in favor of 0^+, assuming 0^+ is true (left), and vice-versa, $0^+ \leftrightarrow 0^-$ (right), for m_H=145. The dashed central line is the median significance. The 1 and $2\,\sigma$ bands correspond to 68% and 95% confidence intervals, centered on the median.
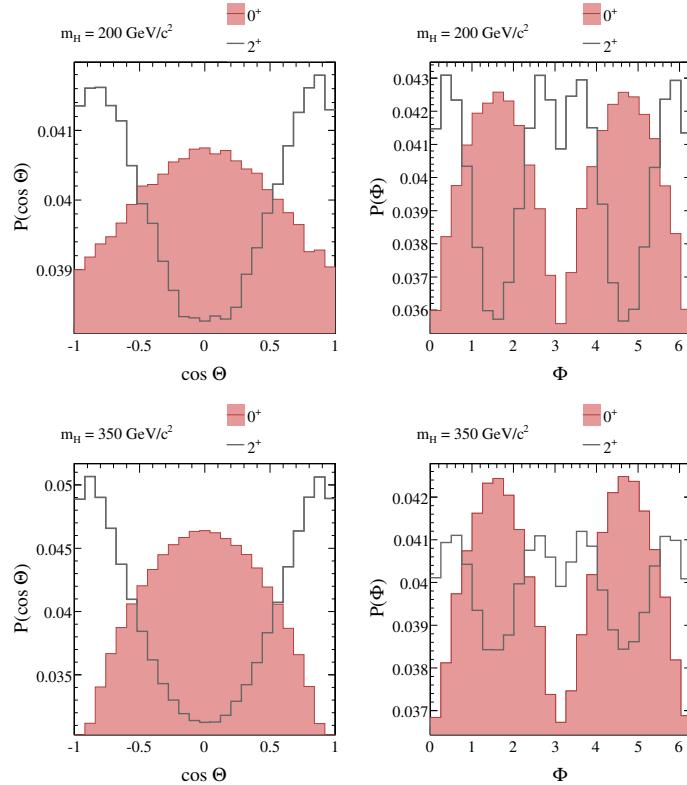
Figure 7.7: Distributions of the variables $\cos\Theta$ (left) and Φ (right) for 0^+, 2^+ resonances with masses of 200 and 350 GeV/c^2 (top, bottom). All distributions are normalized to a unit integral.

terms in $\cos\Theta$ in the 2^+ *pdf* and the absence of this variable in the 0^+ *pdf*, their corresponding one-dimensional *pdfs* are similar for the 0^+ and 2^+ resonances for values of m_H close to $2\,M_Z$, as shown in figure 7.7. Similar behavior is observed in the distributions of $\cos\theta_1$ and $\cos\theta_2$, as illustrated in figure 7.8. Nevertheless, the inclusion of all angular variables and their correlations improves the discrimination power between these hypotheses as shown in figure 7.9.
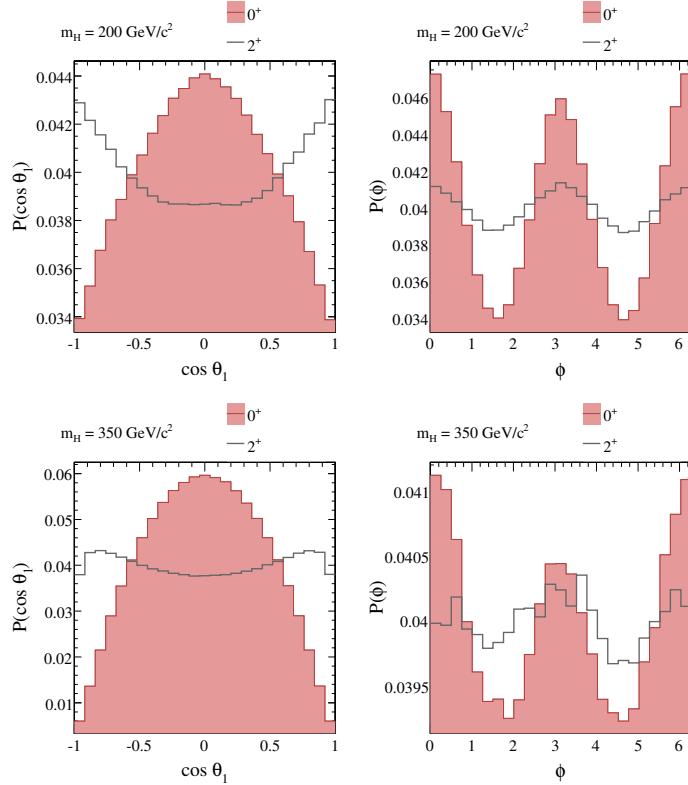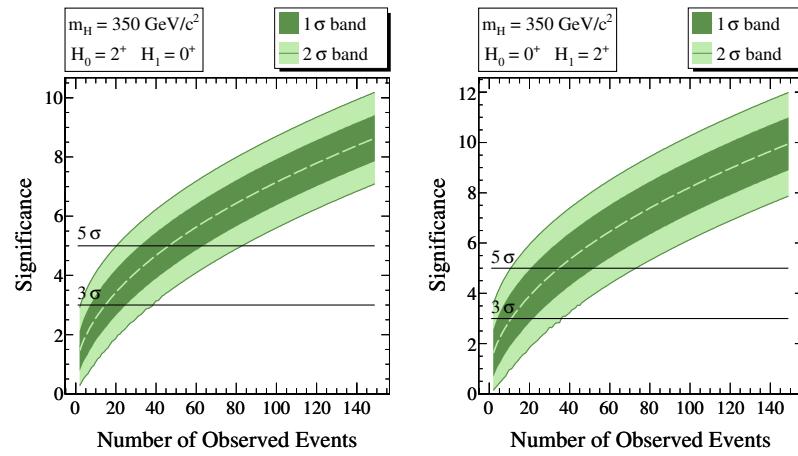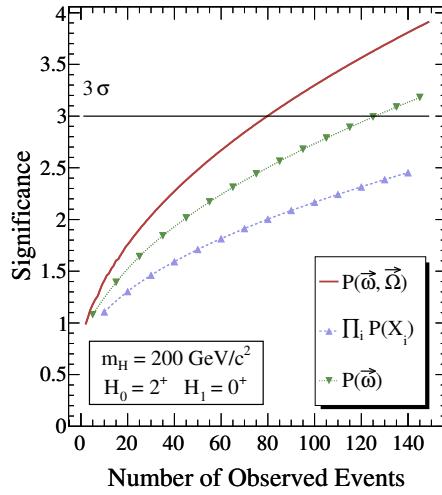
Figure 7.8: Distributions of the variables $\cos\theta_1$ (left) and ϕ (right) for 0^+, 2^+ resonances with masses of 200 and 350 GeV/c^2 (top, bottom). All distributions are normalized to a unit integral.

The significance for discriminating between 0^+ and 2^+ as a function of N_S, is summarized in figure 7.10 for m_H=350 GeV/c^2. For these tests the variables $\vec{\Omega}$ and $\vec{\omega}$ and their correlations were used in the likelihood. Model discrimination is based on the NePe test between simple hypotheses with test statistic $\log(\mathcal{L}[0^+]/\mathcal{L}[2^+])$ and $\log(\mathcal{L}[0^+]/\mathcal{L}[2^-])$.

Figure 7.9: Median significance for rejecting 2^+ in favor of 0^+, assuming 0^+ is true, for the different likelihood constructions discussed in the text.

Figure 7.10: Significance for rejecting 2^+ in favor of 0^+, assuming 0^+ is true (left) or vice-versa ($0^+ \leftrightarrow 2^+$, right), for $m_H = 350$ GeV/c^2.

7.3.4 0^+ vs. Mixed Scalar States

Consider the vertex Feynman rules of equation (7.9) for the most general Lorentz-covariant coupling $\mathcal{L}_{\mu\alpha}$ of a spinless object to a Z pair. Rather than studying the general case, for which any of the quantities X to Q can be nonzero, we investigate three cases, each with only two nonvanishing types of coupling, resulting in one free mixing "angle" and an overall normalization (which we ignore):

- $X \neq 0$, $P \neq 0$: A scalar whose ZZ coupling violates CP, described in terms of an angle ξ_{XP} as

$$\mathcal{L}_{\mu\alpha} \propto \cos(\xi_{XP})\, g_{\mu\alpha} + \sin(\xi_{XP})\, \epsilon_{\mu\alpha} p_1 p_2 / M_Z^2$$

- $X \neq 0$, $Q \neq 0$: A scalar whose ZZ coupling violates C, described in terms of an angle as

$$\mathcal{L}_{\mu\alpha} \propto \cos(\xi_{XQ})\, g_{\mu\alpha} + i \sin(\xi_{XQ})\, \epsilon_{\mu\alpha} p_1 p_2 / M_Z^2$$

- $X \neq 0$, $Y \neq 0$: A composite 0^+, parameterized in terms of an angle as

$$\mathcal{L}_{\mu\alpha} \propto \cos(\xi_{XY})\, g_{\mu\alpha} - \sin(\xi_{XY})\, k_\alpha k_\mu / M_Z^2$$

As a function of N_S we estimate the significance with which one can determine:

- (a) What range of values of the angles can be excluded in favor of a pure 0^+ for a SM-like resonance;

- (b) Whether a pure 0^+ can be excluded in favor of a nontrivial mixture when the resonance corresponds to one of the three mixed cases discussed above.

We consider first the example of a CP-violating HZZ coupling with m_H=350 GeV/c^2.

To address (a) we construct a series of simple hypothesis tests of the type we considered earlier for distinguishing between pure J^{PC} states. Specifically, for a given

number of observed signal events at a fixed value of m_H, we perform a NePe test
between two simple hypotheses: that the resonance is 0^+ (denoted hypothesis $\mathbb{H}_1$)
or that the resonance is $J{=}0$ with ξ_{XP} *fixed to a specific nonzero value* (denoted
hypothesis $\mathbb{H}_0$). The test statistic we use is $\log[\mathcal{L}^{XP}(\xi_{XP})/\mathcal{L}(0^+)]$, where $\mathcal{L}(0^+)$ and
$\mathcal{L}^{XP}(\xi_{XP})$ denote the likelihoods for a set of events agreeing with the hypotheses $\mathbb{H}_1$
and $\mathbb{H}_0$, respectively. The test cannot be performed for $\xi_{XP}{=}0$, since in this case the
$\mathbb{H}_0$ CP-violating hypothesis we want to test reduces to the alternative $\mathbb{H}_1$ hypothesis
(the CP-conserving SM Higgs).

The result of this test is the significance with which hypothesis $\mathbb{H}_0$ can be rejected
in favor of the hypothesis $\mathbb{H}_1$, or similarly, the significance with which a particular
value of ξ_{XP} can be excluded in favor of the 0^+ hypothesis. This test is then repeated
with different fixed values of ξ_{XP}, i.e. different NePe tests with different hypotheses
$\mathbb{H}_0$. The results for a large ensemble of such tests are shown in figure 7.11. Here,
$\mathbb{H}_0 = 0^{XP}$ denotes the simple $J{=}0$ CP-violating hypothesis with ξ_{XP} fixed at values
chosen on the x-axis.
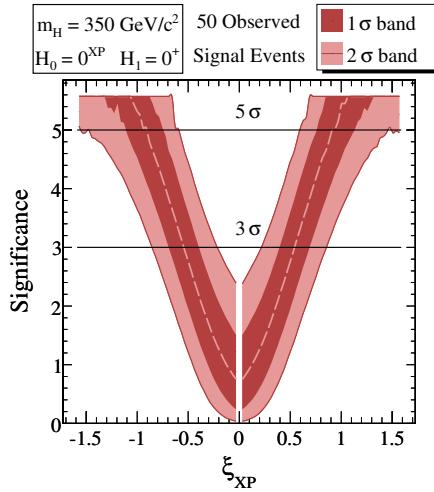
Figure 7.11: Significance for excluding values of ξ_{XP} in the CP-violating $J{=}0$ hypothesis in favor of the 0^+ one, assumed to be correct, for $m_H{=}350$ GeV/c^2 and $N_S{=}50$. The dashed line corresponds to the median of the significance. The 1 and $2\,\sigma$ bands correspond to 68% and 95% confidence intervals centered on the median value.

In this example we see that, for N_S=50, the significance for excluding a CP-violating coupling exceeds $3\,\sigma$ for $|\xi_{XP}| > 0.5$ and $5\,\sigma$ for $|\xi_{XP}| > 0.9$.

In addressing (b) we cannot construct a simple NePe test between 0^+ and a fixed-ξ_{XP} hypothesis. Instead, we treat ξ_{XP} as a nuisance parameter and choose a value, $\hat{\xi}_{XP}$, that maximizes the CP-violating likelihood for the given set of observed events. Specifically, we fix ξ_{XP} at a particular value (the "true" value) to generate events and perform NePe tests comparing ξ_{XP}=0 (denoted hypothesis $\mathbb{H}_0$) and $\xi_{XP} = \hat{\xi}_{XP}$ ($\mathbb{H}_1$). This test is repeated for many different values of the fixed "input" ξ_{XP}.

An example of results from an ensemble of these tests is shown in figure 7.12. Because of the addition of a nuisance parameter, the figure's interpretation is not simply related to the interpretation of figure 7.11, which answered question (a). What figure 7.12 shows is the expected significance with which one can exclude the SM hypothesis in favor of the CP-violating hypothesis with $\xi_{XP}=\hat{\xi}_{XP}$, as a function of the true value of ξ_{XP} (given on the x-axis). No a priori knowledge of the actual value of ξ_{XP} is required to perform this test. From figures 7.11 and 7.12 we observe
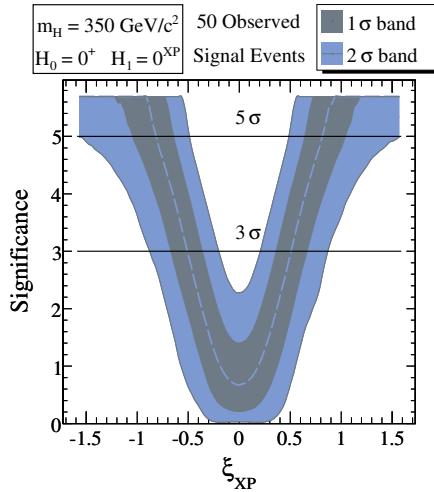
Figure 7.12: The significance for excluding a pure 0^+ in favor of a CP-violating HZZ coupling ($\xi_{XP} \neq 0$), assuming the latter to be correct, with ξ_{XP} given by its x-axis values. Example for N_S=50, m_H=350 GeV/c^2. Dashed line and bands as in figure 7.11.

that the expected significances are symmetric around ξ_{XP}=0. This is due to the *pdfs*

of the "pure 0^+" and "pure 0^-" terms being even under $\xi_{XP} \to -\xi_{XP}$, while the $\tilde{T}$-odd interference term vanishes under the integration of $\cos\theta_1$, $\cos\theta_2$ or ϕ. We shall see that there are exceptions to this trivial statement. Comparing these two figures we observe a remarkable similarity of the significances of the two tests. Since two different statistics are used, this is somewhat of a coincidence. To explain it, consider the example with $\xi_{XP}=\pi/5$, which corresponds to vertical slices of figures 7.11 and 7.12. We denote the two different test statistics $\Lambda^{\text{fix}}=\log[\mathcal{L}^{XP}(\xi_{XP})/\mathcal{L}(0^+)]$, with ξ_{XP} fixed at its true value, corresponding to a simple hypothesis test and $\Lambda^{\text{max}}=\log[\max\mathcal{L}^{XP}(\hat{\xi}_{XP})/\mathcal{L}(0^+)]$, *profiled* to the value $\hat{\xi}_{XP}$ at which it peaks. The distributions of Λ^{fix} and Λ^{max} are shown in figure 7.13.
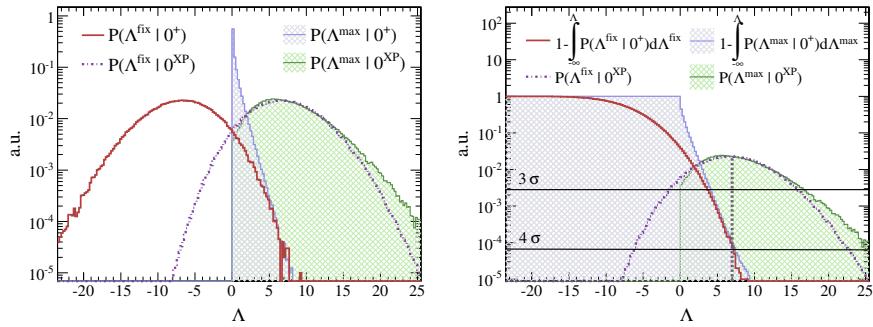
Figure 7.13: Distributions of the two statistics Λ, defined in the text, for $m_H=350$ GeV/c^2 and $N_S=50$. The hypotheses are $\mathbb{H}_0=0^+$, and $\mathbb{H}_1=0^{XP}$ with the CP-phase ξ_{XP} fixed at $\pi/5$. (Top) Probability distributions P($\Lambda|\mathbb{H}$). (Bottom) The same with the 0^+ results traded for 1 minus their cumulative values. The two nearly indistinguishable vertical dotted lines correspond to the median values of the P($\Lambda|\mathbb{H}_1$) distributions.

In the top figure the bell-shaped curves $P(\Lambda^{\text{fix}}|0^+)$ and $P(\Lambda^{\text{fix}}|0^{XP})$ are characteristic of a simple hypothesis test. The distributions of Λ^{max} have a sharp cutoff at $\Lambda^{\text{max}}=0$, since the 0^+ model is a member of the 0^{XP} family with $\xi_{XP}=0$, and $\max L^{XP}(\hat{\xi}_{XP})/L(0^+) \geq 1$, which are also features characteristic of this type of test.

The reason for two very different hypothesis tests to end up in the similar-looking results of figures 7.11 and 7.12 is that the statistically significant features of the different-looking $P(\Lambda)$ distributions shown in figure 7.13 are actually very similar. $P(\Lambda^{\text{fix}}|0^{XP})$ and $P(\Lambda^{\text{max}}|0^{XP})$ differ, but the distributions of ξ_{XP} close to the max-

ima are localized around the true input value, their median values and 68% and 95% confidence intervals are nearly identical (try to tell apart the two vertical dotted lines in the lower half of figure 7.13, at $\Lambda \sim 7$). Also, the tails of one-minus-cumulative distributions for $P(\Lambda^{\text{fix}}|0^+)$ and $P(\Lambda^{\text{max}}|0^+)$ coalesce for p-values exceeding $2\,\sigma$ significance, despite large differences in the distributions themselves.
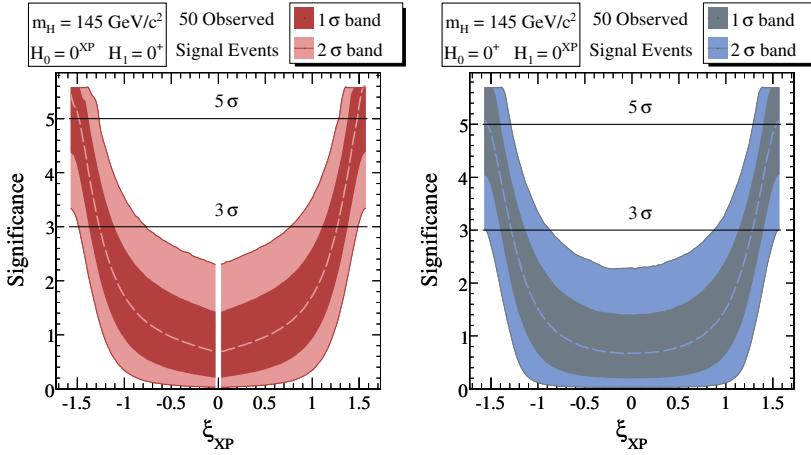
Figure 7.14: (Left) Significance for the exclusion of values of a CP-violating $\xi_{XP} \neq 0$ in favor of 0^+ (ξ_{XP}=0), assumed to be correct. (Right) Significance for excluding a pure 0^+ in favor of $\xi_{XP} \neq 0$, assumed correct with ξ_{XP} given by its x-axis values. Results for m_H= 145 and N_S=50.

In figure 7.14 we show the results for the distinction between pure 0^+ and CP-violating J=0 hypotheses for m_H=145 GeV/c^2. The "flat" behavior around ξ_{XP}=0 is due to the coupling strength of the 0^+ part relative to 0^-, an order of magnitude larger for m_H=145 GeV/c^2 and closer to unity for the higher m_H values. The corresponding results at m_H=350 GeV/c^2 are those of figures 7.11 and 7.12.

The next mixed J=0 case that we consider is that of a C-violating scalar, with mixing angle ξ_{XQ}. This scenario is very similar to that of the CP-violating scalar: only the interference term between the 0^+ and 0^- amplitudes is different (C-odd, instead of T-odd). The expected results of hypothesis tests distinguishing between a C-violating scalar and a 0^+ state are shown in figure 7.15. Comparing this figure with 7.14, we observe identical behavior in all the results. This shows that the relative strength between the 0^+ and 0^- parts of the matrix element squared, rather than the

nature of the interference term, is the most relevant factor in resolving the values of ξ_{XP} and ξ_{XQ}.
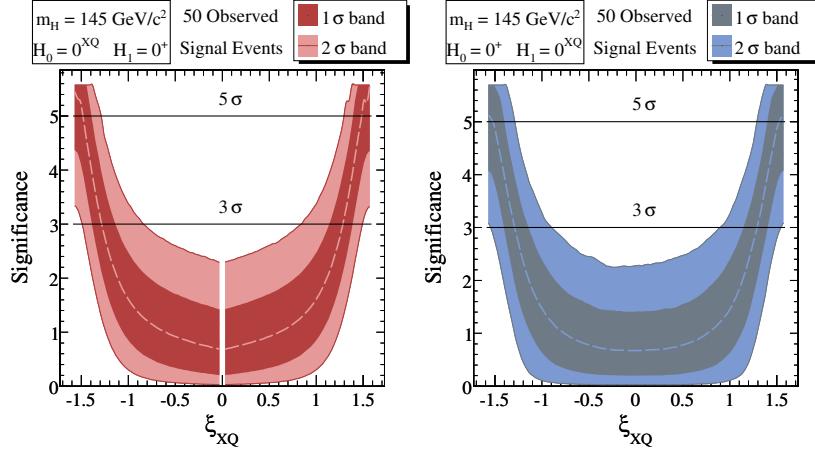
Figure 7.15: (Left) Significance for excluding values of a C-violating $\xi_{XQ} \neq 0$ in favor of 0^+ (ξ_{XQ}=0), assumed to be correct. (Right) Significance for excluding a pure 0^+ in favor of $\xi_{XQ} \neq 0$, assumed correct for the ξ_{XQ}-values on the x-axis. Hypothesis tests are for m_H=145 and N_S=50.

If a pure 0^+ hypothesis is rejected in favor of both $\xi_{XP} \neq 0$ and $\xi_{XQ} \neq 0$, the next question would be whether it is possible to distinguish between these two cases. To address this question, we perform a series of hypothesis tests similar to the one described to answer type (2) questions. Specifically, we first assume a given CP-violating $\xi_{XP} \neq 0$ as "true." We then assess the expected significance with which particular values of ξ_{XQ} can be excluded in favor of the true hypothesis. Hence, for each fixed value of ξ_{XP} we perform a test against the C-violating case using a fixed ξ_{XQ}. The test statistic is $\Lambda = \log[\max L^{XP}(\hat{\xi}_{XP})/L(\xi_{XQ})]$, where the 0^{XQ} hypothesis is simple (fixed ξ_{XQ}) and $L(\xi_{XP})$ is *profiled* "experiment by experiment." The test is repeated over a matrix of values for ξ_{XP} and ξ_{XQ}. Next, we switch the roles of the hypotheses to assess the significance for excluding given values of ξ_{XP} in favor of $\xi_{XQ} \neq 0$. The results are shown in figure 7.16. The color-coded z-axis is the median of the significance for ruling out the hypothesis $\mathbb{H}_0$ with the value of $\xi_{\mathbb{H}_0}$ given on the y-axis in favor of the $\mathbb{H}_1$ hypothesis with $\xi_{\mathbb{H}_1} \neq 0$, assumed to be correct for $\xi_{\mathbb{H}_1}$-values chosen on the x-axis.
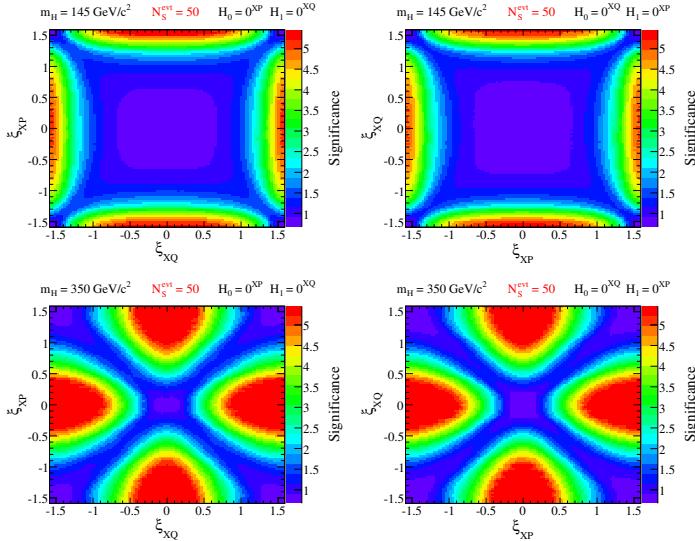
Figure 7.16: The median of the significance (colored z-"axis") for excluding values of $\xi_{\mathbb{H}_0}$ (y-axis) in favor of the $\xi_{\mathbb{H}_1} \neq 0$ hypothesis assuming as correct the values $\xi_{\mathbb{H}_1}$ of the x-axis. The tests are performed for $\mathbb{H}_1 = 0^{XP}$, $\mathbb{H}_0 = 0^{XQ}$ (Left) and $\mathbb{H}_1 = 0^{XQ}$, $\mathbb{H}_0 = 0^{XP}$ (Right); $m_H = 145$ and 350 GeV/c^2 (top and bottom), for $N_S = 50$.

The similarities between the C- and CP-mixed scalars are reflected in the $y \leftrightarrow x$ symmetries of figures 7.16. Moreover, switching the roles of the two hypotheses (comparing the figures on the left with those on the right) one only sees small changes. Still, the fact that the diagonals ($|\xi_{XP}| = |\xi_{XQ}|$) are not all at the same significance shows that the tests are sensitive to the differences between the $\tilde{T}$- and C-odd interference terms, but it would require an order of magnitude larger N_S to draw $5\,\sigma$-level conclusions over most of the (ξ_{XP}, ξ_{XQ}) plane. For example, we show in figure 7.17 the significance with which one can distinguish between the two cases, as a function of the number of observed events, for $\xi_{XY,XQ} = \pi/4$ and $m_H = 200$ GeV/c^2. The ambiguity between ξ_{XP}^{meas}, $-\xi_{XP}^{meas}$, $\xi_{XQ} = \xi_{XP}^{meas}$ and $\xi_{XQ} = -\xi_{XP}^{meas}$ would be very hard to lift.

The last $J = 0$ mixed case that we consider has unique features; this is the "composite Higgs" in which a term $\propto k_\mu k_\nu$ is present in the HZZ coupling. This case is different from the previous ones in that a composite scalar has well defined $J^{PC} = 0^{++}$, regardless of the value of the angle ξ_{XY} characterizing the mixing between its point-
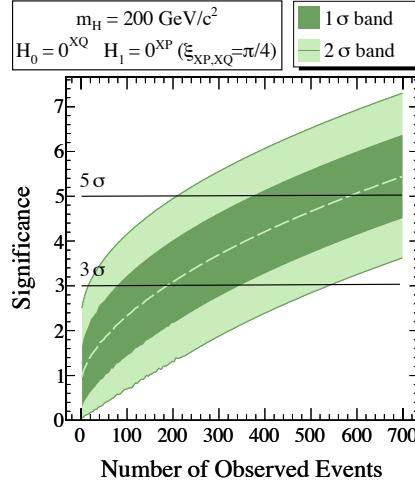
Figure 7.17: The significance for excluding the C-violating $J=0$ hypothesis in favor of a CP-violating case, assuming the latter to be correct, with $\xi_{XP,XQ}=\pi/4$. Example for $m_H=200$ GeV/c^2.

like and derivative couplings. As a consequence, the angular integrals of their interference term do not vanish, and there is no symmetry around $\xi_{XY}=0$. All the terms in the *pdf* having the same discrete symmetries and similar angular dependences; there happen to be large cancellations in the *pdf* for a "critical" m_H-dependent value of ξ_{XY}, as in the example shown in figure 7.18 for the fully angular-integrated result.
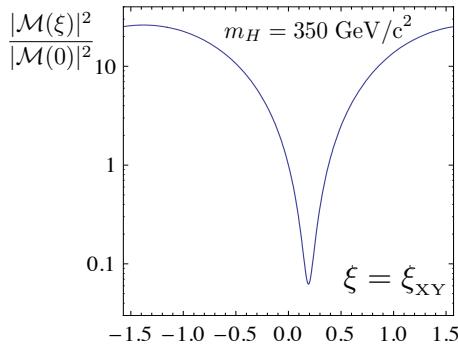
Figure 7.18: The fully angularly-integrated matrix element squared for a "composite" 0^+, showing a strong destructive interference at a given ξ_{XY}. The result, shown here for $m_H=350$ GeV/c^2, is normalized to $\xi_{XY}=0$.

The appearance of an order of magnitude enhancement of the squared matrix element in figure 7.18 for $O(1)$ values of ξ_{XY} can be regarded as an artifact of our choosing

a rather low mass scale (M_Z) in the definition of the dimensionless coupling Y in equation (7.9); if, e.g., we instead chose the compositeness scale at m_H=350 GeV/c^2, this enhancement would be much smaller. Nevertheless the possible enhancement from a nonzero Y coupling, and the possible suppression from XY interference, signifies an interesting scenario: it is possible to discover an HLL that is in fact a 0^{++} resonance, and is produced by exactly the same pp production processes as a SM Higgs, but for which the cross section times branching fraction to ZZ is several times higher *or* several times lower than Standard Model expectation.

We evaluate the significance with which one can distinguish between a point-like and a composite 0^+ using the same hypothesis-test approach described earlier for the CP-violating scalar case. The results are shown in figure 7.19. We observe a nontrivial behavior of the significance values at and around the critical ξ_{XY}. Interestingly, the qualitative nature of these cancellations also changes with mass. For m_H=145 GeV/c^2 the composite scalar with ξ_{XY} near the critical point is 0^+-like, relative to nearby values of ξ_{XY}. For m_H=350 GeV/c^2, it is very difficult to distinguish between the composite and elementary hypotheses, except if ξ_{XY} is close to critical. Near this critical value the significance is greatly improved, because after the large cancellations the angular distributions of the pure 0^+ and the mixed case no longer resemble each other.

As we discussed for the C- and CP-violating cases, an additional question is whether one can distinguish a composite scalar from other mixed scalars. We find that, compared to the composite case, the two other mixed cases are nearly identical. The results for the distinction between the CP-violating and composite cases are shown in figure 7.20. For large values of ξ_{XY} and ξ_{XP}, it is possible to distinguish between the two hypotheses at a large significance with a mere N_S=50. For m_H=350 GeV/c^2, the composite scalar is very similar to the point-like 0^+ and cannot be distinguished from it except if ξ_{XY} is near its critical point. Replacing the CP-violating scalar with the C-violating one yields results nearly identical to the ones in figure 7.20.
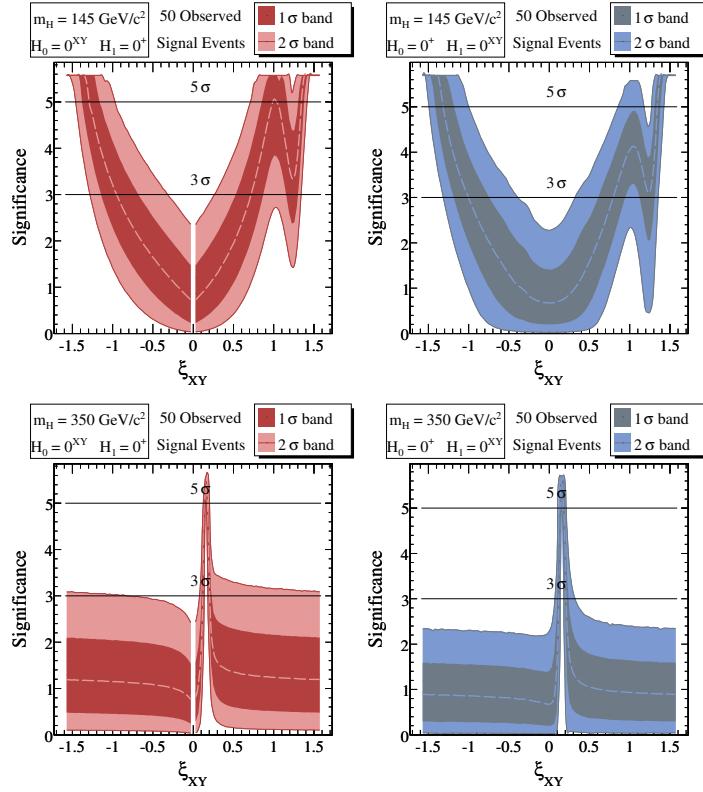
Figure 7.19: (Left) significance for excluding values of ξ_{XY} in favor of a point-like 0^+ (ξ_{XY}=0), assumed to be correct. (Right) significance for excluding a point-like 0^+ in favor of a "composite" one ($\xi_{XY} \neq 0$), assumed correct for the ξ_{XY} values on the x-axis, for m_H=145 and 350 GeV/c^2 (top and bottom) and N_S=50.
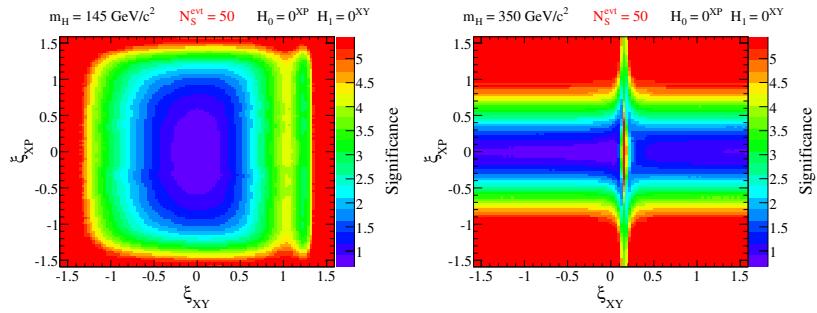
Figure 7.20: The median of the significance (colored-labeled z-"axis") for excluding values of ξ_{XP} (y-axis) in favor of the composite scalar assuming it to be correct with the ξ_{XY} values of the x-axis, for m_H=145 and 350 GeV/c^2 (left and right) and N_S=50.

7.3.5 Lagrangian Parameter Estimation in Mixed J = 0 Cases

Were one to find out from real data and the hypothesis tests discussed in the previous section that a mixed $J=0$ state is the preferred description, the next item in the context of this analysis would be the measurement of its mixing parameters (in a larger context one would include at this stage the measurement of decay branching ratios).

We have seen in section 7.3.4 that our hypothesis tests can demonstrate, if correct, that a standard 0^+ particle is disfavored relative to a mixed scalar with unspecified HZZ coupling ratios (or mixing angles). In these tests, the angles were treated as nuisance parameters. Their measurement proceeds along the same line; the preferred value is simply that which maximizes the likelihood, but the treatment of confidence intervals need be different.

More specifically, each mixed hypothesis family is characterized by mixing angles $\vec{\xi}$. For each "experiment," N events are simulated, each one characterized by a vector $\vec{x}_e = \{\vec{\omega}, \vec{\Omega}, M_{Z^*}\}|_e$. The likelihood for a particular family of hypotheses is $\mathcal{L}(\vec{\xi}) = \prod_{e=1}^{N} P_e(\vec{x}_e, \vec{\xi})$. The measured values of the mixing angles, $\vec{\xi}_{\mathrm{meas}}$, are chosen to be those that maximize the likelihood. To assign confidence intervals to these measurements we use a fully frequentist approach. An ensemble of "experiments" is performed with fixed input values $\vec{\xi} = \vec{\xi}_{\mathrm{input}}$. For each experiment, the measured values of $\vec{\xi}$ are taken from the maximization of the likelihood. This procedure is repeated for a fine matrix of input values, covering the allowed parameter space. From the probability distribution functions $P(\vec{\xi}_{\mathrm{meas}}|\vec{\xi}_{\mathrm{input}})$, estimated using this ensemble of experiments, the Feldman-Cousins unified approach [117] is used to choose which elements of probability are included in confidence intervals.

As an example, consider the CP-violating scalar case, discussed in section 7.3.4. The confidence intervals for measured values of ξ_{XP} (the mixing parameter that characterizes this hypothesis) are shown in figure 7.21 for different values of m_H. The way to interpret these figures is as follows: For a particular set of data (one experiment,
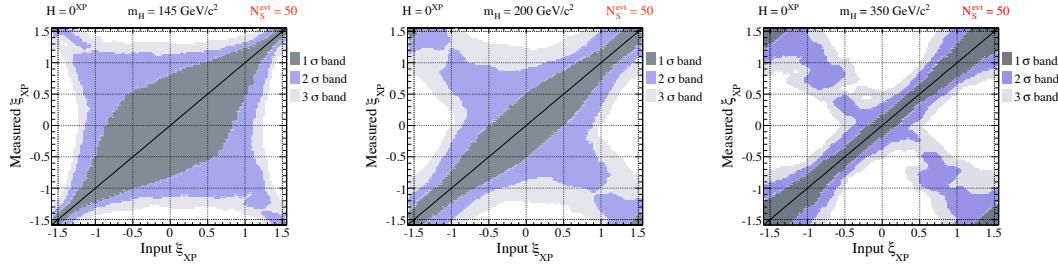
Figure 7.21: Confidence intervals for measured values of ξ_{XP} for a CP-violating $J=0$ resonance, for $m_H=145$, 200 and 350 GeV/c^2 (left, center and right), all for $N_S=50$. For measured values of ξ_{XP} on the y-axis, confidence intervals should be read horizontally, see text.

which in this case includes $N_S=50$ observed events) an input value of ξ_{XP} (to be read on the x-axis) results in a measured value to be read (with its error bands) on the y axis. The confidence intervals are obtained by drawing a horizontal line passing through the measured ξ_{XP}. The overlap of this line with the $n\,\sigma$ bands dictates which values of "input ξ_{XP}" should be included in the $n\,\sigma$ confidence intervals. For example, for $m_H=200$ GeV/c^2 (middle of figure 7.21) we see that, if $\xi_{XP}^{\rm meas}=0$, the $3\,\sigma$ confidence interval is approximately $\xi_{XP} \in [-1, 1]$.

The $1\,\sigma$ bands in figure 7.21 are centered on the diagonal $\xi_{XP}^{\rm meas}=\xi_{XP}^{\rm input}$, implying that there is no significant bias in the measurement. In addition to this, the $2\,\sigma$ and $3\,\sigma$ bands also cover most of the diagonal $\xi_{XP}^{\rm meas} = -\xi_{XP}^{\rm input}$. This confirms our observation from section 7.3.4 that our ability to pin down this parameter comes predominantly from measuring the relative strengths of the 0^+ and 0^- parts of the *pdf* rather than the nature ($\tilde{T}$-odd) of its interference term. An increased number of observed events is needed to fully resolve this sign ambiguity.

In figure 7.21 we see that for $m_H=145$ GeV/c^2 (but not for $m_H=200$ GeV/c^2) the size of the confidence intervals for ξ_{XP} decreases with increasing $|\xi_{XP}|$. This is due to the effective coupling strengths of the 0^+ and 0^- parts of the *pdf* differing by a factor of ~ 10 at $m_H=145$ GeV/c^2 but not at the other masses. Hence, at the lowest mass, only at $\tan^2(\xi_{XP}) \sim 10$ does the *pdf* exhibit 0^+- and 0^--like behaviors of similar magnitude.
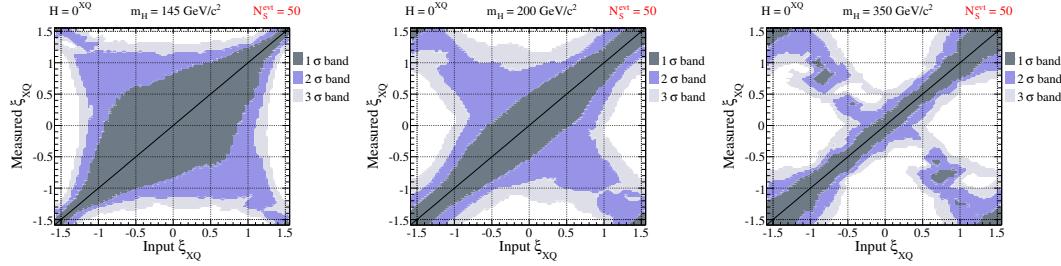
Figure 7.22: Confidence intervals for measured values of ξ_{XQ} for a C-violating $J=0$ resonance for m_H=145, 200 and 350 GeV/c^2 (left, center and right), all for N_S=50. For measured values of ξ_{XQ} on the y-axis, confidence intervals should be read horizontally, see text.
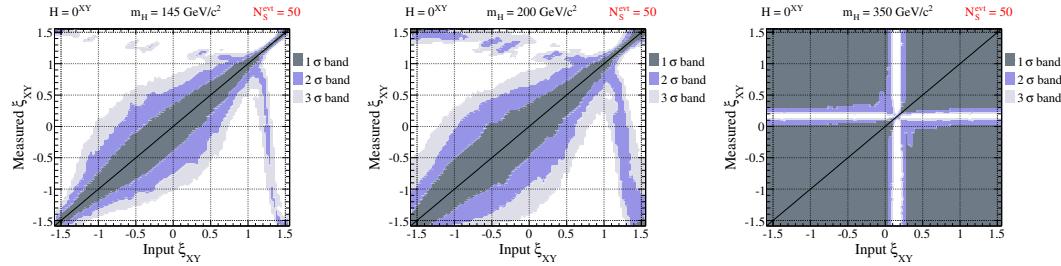
Figure 7.23: Confidence intervals for measured values of ξ_{XY} for a "composite" $J=0$ resonance, for m_H=145, 200 and 350 GeV/c^2 (left, center and right), all for N_S=50. For measured values of ξ_{XY} on the y-axis, confidence intervals should be read horizontally.

Confidence intervals for measurements of the parameter ξ_{XQ} for a scalar with C-violating HLL couplings are shown in figure 7.22; These are nearly identical to those in figure 7.21, reflecting the difficulty of discriminating the $\xi_{XP} \neq 0$ and $\xi_{XQ} \neq 0$ hypotheses, as discussed in section 7.3.4. For the C-odd case, the sign ambiguity of ξ_{XQ}^{meas} is slightly worse than for the $\tilde{T}$-odd one as demonstrated by the $1\,\sigma$ confidence bands appearing on the $\xi_{XQ}^{\mathrm{meas}} = -\xi_{XQ}^{\mathrm{input}}$ diagonal for m_H=350 GeV/c^2. This is also expected, since the C-odd interference term is proportional to the relatively small number $\eta \approx 0.15$, see equation (7.17). One's ability to distinguish between $J=0$ C- and $\tilde{T}$-odd admixtures relies on the resolution of the interference terms. With a factor of 10 more statistics ($N_S \sim 500$), one would be able to resolve the sign ambiguity in ξ_{XP} and ξ_{XQ} and to distinguish between the two cases.

The confidence intervals associated with measurements of ξ_{XY} for a composite scalar are shown in figure 7.23. We observe that, for m_H=145 and 200 GeV/c^2, the $1\,\sigma$ intervals are centered on the diagonal $\xi_{XY}^{meas} = \xi_{XY}^{input}$. There are no bands along $\xi_{XY}^{meas} = -\xi_{XY}^{input}$, since the interference term is of a different nature than that of the discrete-symmetry violating cases. The extensions of the 2 and 3 σ bands along almost horizontal and vertical lines around $\xi_{XY} \sim 1.3$ result from large cancellations in the *pdf*, discussed in section 7.3.4.

The figure for m_H=350 GeV/c^2 is hard to decipher. With a magnifier one sees that at the critical value of ξ_{XY} the confidence intervals are tiny. Everywhere else, the intervals essentially include all possible values *except* the critical one. This is tantamount to saying that at this mass we cannot tell, on the basis of our analysis, a composite from a point-like scalar *unless* is has a particular value of ξ_{XY}, a fact made clearer by figure 7.19.

7.4 Conclusions and Outlook

These studies demonstrate that small signal samples in the $ZZ \rightarrow 4\ell$ or $ZZ^* \rightarrow 4\ell$ decay channels could be sufficient to characterize a putative Higgs particle. Below we summarize these each of the results.

7.4.1 Summary of Pure Case Discrimination

Amongst the many comparisons considered in our analysis, the ones between simple hypotheses are the most readily summarized. This we do in Tables 7.1 and 7.2 for m_H=145 GeV/c^2 for all pure-case comparisons between J=0, 1 parent particles, and in Tables 7.3, 7.4 (7.5, 7.6) for m_H=200 (350) GeV/c^2, for all pure-case comparisons between J=0, 1, 2 parent particles.

Overall, the discrimination power of the hypothesis tests is very impressive. The m_H=200 GeV/c^2 benchmark example is the one requiring the largest statistics to reach a given discrimination at a given level of confidence. Compared with the m_H=350 GeV/c^2 case, this is because various coefficients of the angular dependences

$\mathbb{H}_0 \Downarrow \mathbb{H}_1 \Rightarrow$	0^+	0^-	1^-	1^+
0^+	–	17	12	16
0^-	14	–	11	17
1^-	11	11	–	35
1^+	17	18	34	–

Table 7.1: Minimum number of observed events such that the median significance for rejecting $\mathbb{H}_0$ in favor of the hypothesis $\mathbb{H}_1$ (assuming $\mathbb{H}_1$ is right) exceeds $3\,\sigma$ with $m_H{=}145$ GeV/c^2.

$\mathbb{H}_0 \Downarrow \mathbb{H}_1 \Rightarrow$	0^+	0^-	1^-	1^+
0^+	–	52	37	50
0^-	44	–	34	54
1^-	33	32	–	112
1^+	54	55	109	–

Table 7.2: Same as Table 7.1, but requiring that the median significance exceeds $5\,\sigma$.

$\mathbb{H}_0 \Downarrow \mathbb{H}_1 \Rightarrow$	0^+	0^-	1^-	1^+	2^+
0^+	–	24	45	62	86
0^-	19	–	19	19	38
1^-	40	18	–	90	48
1^+	56	19	85	–	66
2^+	86	45	54	70	–

Table 7.3: Minimum number of observed events such that the median significance for rejecting $\mathbb{H}_0$ in favor of the hypothesis $\mathbb{H}_1$ (assuming $\mathbb{H}_1$ is right) exceeds $3\,\sigma$ with $m_H{=}200$ GeV/c^2.

vanish at the $m_H{=}2\,M_Z$ threshold. The $m_H{=}145$ GeV/c^2 example fares better than the 200 GeV/c^2 one for the same reason, amplified by the extra lever-arm supplied by a non-trivial M_{Z^*} distribution.

The tables also show that the discriminating power between two given hypotheses is approximately symmetric under the interchange of "right" and "wrong." Telling 1^+ from 1^- is always difficult but not impossible, a fact of relevance for a Z' look-alike analysis. The level of significance does not obey a naïve $N(\sigma) \propto \sqrt{N_S}$ law. However we find by inspection that an approximation of the form $N(\sigma) = a + b\sqrt{N_S}$ works well, allowing one to extrapolate to larger numbers of events than presented here.

Other lessons from the tables are case-by-case specific, reflecting the mass-dependent quantum-mechanical entanglement between the decay variables. Some examples are:

$\mathbb{H}_0 \Downarrow \mathbb{H}_1 \Rightarrow$	0^+	0^-	1^-	1^+	2^+
0^+	–	76	146	203	287
0^-	59	–	60	61	123
1^-	130	57	–	297	156
1^+	182	58	278	–	217
2^+	287	146	178	230	–

Table 7.4: Same as Table 7.3, but requiring that the median significance exceeds $5\,\sigma$.

$\mathbb{H}_0 \Downarrow \mathbb{H}_1 \Rightarrow$	0^+	0^-	1^-	1^+	2^+
0^+	–	8	21	24	11
0^-	9	–	22	22	36
1^-	24	22	–	81	46
1^+	26	22	80	–	56
2^+	15	39	55	73	–

Table 7.5: Minimum number of observed events such that the median significance for rejecting $\mathbb{H}_0$ in favor of the hypothesis $\mathbb{H}_1$ (assuming $\mathbb{H}_1$ is right) exceeds $3\,\sigma$ with $m_H{=}350$ GeV/c^2.

distinguishing the "natural-parity" $J{=}0^+$ and 1^- hypotheses for $m_H{=}145$ GeV/c^2 requires only a dozen signal events for $3\,\sigma$ discrimination. For 200 GeV/c^2, discriminating 0^+ from 0^- is relatively easy, but distinguishing 0^+ from 2^+ is difficult. For 350 GeV/c^2, contrariwise, 2^+ is relatively easy to disentangle from 0^+, but not from 0^-.

7.4.2 Summary of Mixed Cases, CP, and Compositeness Discrimination

We find that direct sensitivity to CP odd, parity odd XP interference effects, or to CP odd, parity even XQ interference effects, will require signal samples about

$\mathbb{H}_0 \Downarrow \mathbb{H}_1 \Rightarrow$	0^+	0^-	1^-	1^+	2^+
0^+	–	25	67	77	35
0^-	26	–	68	68	118
1^-	76	68	–	268	149
1^+	83	68	263	–	184
2^+	46	127	181	240	–

Table 7.6: Same as Table 7.5, but requiring that the median significance exceeds $5\,\sigma$.

an order of magnitude larger than considered here. We have also observed that with much smaller statistics it may be possible to conclude that a mix of X and P (or X and Q) couplings is favored over just the pure X (i.e., 0^+) or pure P (i.e., 0^-) couplings alone. Such a conclusion would be tantamount to demonstrating CP violation in the Higgs sector. However this scenario relies on large CP violation, and even in this favorable case one cannot tell an X and P mixture from an X and Q mixture without more data than what is required to establish discovery.

In the case of a composite Higgs, it may be conceivable that the Higgs is as "soft" as a pion, in the sense of having an inverse radius and a mass of comparable magnitude. In this scenario we have seen that the angular distributions associated to the X and Y couplings are similar after integrating over the decay angles. As a result there can be strong destructive interference between these contributions. For our lighter mass benchmarks we find good discrimination of pure 0^+ from the mixed composites. For the heavier m_H=350 GeV/c^2 example, discrimination based on decay angles is poor unless the strong interference effects are present; here we also observed that substantial enhancement or suppression of the HLL$\to ZZ$ branching fraction can provide another important discriminator.

For mixed cases, one could worry that certain combinations of exotic couplings might let an HLL successfully masquerade as a 0^+ Higgs, even when all the pure case exotics are excluded. For spin 1 HLLs we have shown that this does not happen. In fact we find that when we have an SM Higgs, the entire family of mixed coupling spin 1 HLLs can be excluded at approximately the same expected level of significance as for the pure 1^- or 1^+ cases. An even stronger result is that the general spin 0 hypothesis can be conclusively discriminated from the general spin 1 hypothesis, at or close to the moment of discovery.

7.4.3 Confronting CMS Data

In our analysis we focused on decay information, exploiting an approximate factorization between observables related to Higgs (or HLL) production and observables

related to decay. The factorization is only approximate because of phase space acceptance effects and, in the case of spin > 0 HLLs, correlations between the initial and final state particles. In a real data analysis one would want to include production information, which in turn would require a detailed knowledge of radiative corrections, PDFs, and full detector simulation for the HLLs. While beyond the scope of this study, such an analysis is currently being performed with CMS data.

The QCD corrections to the signal predictions for $d^2\sigma/dp_T\, d\eta$ are large, as is well-studied for the SM Higgs (see, e.g., [85, 119, 120] and references therein.) The impact on the total cross sections is not relevant to our analysis, but the corrections to the (p_T, η) distributions will modify the phase space acceptance effects on the distributions of the final-state leptons. For the SM Higgs these corrections are included at NLO in the CMS analysis, and a recent study shows that the effects of NNLO corrections on the final-state lepton distributions are not dramatic [121].

There are electroweak radiative corrections that directly involve the final-state leptons. For the SM Higgs these corrections have been computed and studied in detail [122]; the corrections are of the order of 5-10% and cause a mild distortion of the angular distributions. These effects are included in the CMS analysis, but they do not introduce anything conceptually new to the methodology proposed in this study, and their inclusion involves details of the experimental treatment of the vertex and subsequent radiations by electrons and muons.

Preliminary CMS studies [123] have tested the simple 0^+ and 0^{-1} cases, with the NePe hypothesis test-statistic distributions shown in figure 7.24. At this stage, the 0^- hypothesis is disfavored by more than 2 σ with respect to the SM Higgs. The sensitivity of the hypothesis test is roughly consistent with expectations from the $m_H = 145$ GeV/c^2 test case, summarized in Tab. 7.1.

Of course this is only the beginning in determining the identity of the putative Higgs candidate. Our treatment of couplings and HLLs was not exhaustive, since we have ignored gauge invariant operators with dimension greater than 6, have only examined one case of spin 2 HLL, and have not even mentioned the possibility of HLLs with spins higher than 2. In addition to the rich possibilities involving mixed scalars,
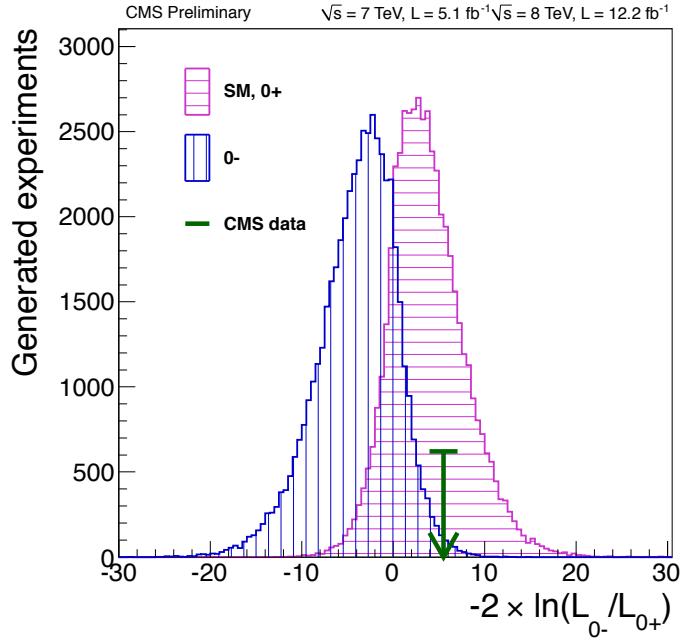
Figure 7.24: Expected distributions of the NePe test statistic comparing the the J^{PC} hypotheses 0^- and 0^+. The observed value, calculated using CMS data in the 4ℓ final state, is indicated by the arrow [123].

compositeness and gravitons there may be more exotic possibilities to consider.

Part II

Symmetries Beyond the Standard Model

Chapter 8

Beyond the Standard Model Physics: Motivations and Possibilities

The Standard Model of particle physics has succeeded in describing the physical world to a remarkable precision to energies up to the weak scale. To date there are no significant deviations between experimental observations and the predictions of the model. If the new boson observed by the CMS and ATLAS experiments proves to be consistent with the Higgs, then the Standard Model will be complete. As we enter an new energy regime with the LHC an important questions are what, if anything, lies beyond the Standard Model (BSM)? The answer is currently unknown, but in this chapter we discuss some reasons to expect that there is physics BSM and how it could manifest itself in LHC collisions. New physics models that mitigate the perceive shortcomings of the SM tend to involve new symmetries of nature; we discuss how these symmetries, and the resulting phenomenology of new physics, can inform our experimental searches for BSM possibilities.

8.1 The Aesthetics of the Standard Model

Despite its enormous theoretical and experimental success, the SM has several shortcomings. To begin with, there are a range of experimental observations that, while not in contradiction with the SM, are not explained by it. The universe we inhabit

is observed to be matter-antimatter asymmetric, and a dynamic explanation from the early universe would require particles and interactions not contained in the SM. Also within our universe, we have inferred the presence of dark matter which only interacts weakly with the SM particles our world is composed of, and so far only detectably through gravity. There is no particle candidate in the SM spectrum which can account for the observed abundance of dark matter, implying that it is composed of BSM particles. The observation of neutrino masses and mixings is, in some sense, evidence of new physics by construction since we have not observed right-handed neutrinos and they don't have a place in the SM. Each of these experimental conclusions strongly implies the presence of BSM physics, but its nature is unknown.

In addition to experimental challenges to the completeness of the SM there are aesthetic and theoretical aspects which suggest that there could be something beyond it. According to our understanding of the SM, all of the masses and mixings of fermions are free, unpredicted parameters. In the context of the SM, the number of fermion generations appears arbitrary (although it is interesting to note that three is the minimum for CP violation) and their masses span several orders of magnitude. The corresponding Yukawa couples, all playing a similar roll in the theory, range from order 1 to 10^{-6}, for no apparent reason. Similarly, the difference in size of the gauge couplings in unexplained. The unification of the weak and electromagnetic forces is an important component in the SM; perhaps the complete theory of the universe should include the unification of the electroweak and strong forces. Gravity is also not included in the SM and is 10^{32} times weaker than the weak force, with its unification an even more difficult prospect.

Perhaps the most instructive problem in the SM for divining the nature of new physics comes as a direct consequence of the masses of the SM particles and the difference in magnitude between the Planck and electroweak scales. If we consider a Dirac fermion (f) in the SM, its mass m_f come from its interaction with the Higgs field (H). When the Higgs field gains a VEV the chiral symmetry of the fermion is broken by a spontaneously generated mass term appearing in the Lagrangian. In addition, an interaction term between the Higgs and fermion appears ($-\lambda_f H \bar{f} f$). Just

as the Higgs gives the fermion mass, this interaction term contributes to quantum corrections to the Higgs mass through diagrams like the one shown in Fig. 8.1 (left). The shift to the Higgs mass from that diagram is

$$\Delta m_H^2 = \frac{|\lambda_f|^2}{16\pi^2} \left[-2\Lambda_{UV}^2 + 6m_f^2 \log(\Lambda_{UV}/m_f) + \cdots \right] , \qquad (8.1)$$

where Λ_{UV} is the cutoff scale for the fermion momentum running in the loop and the ellipses correspond to high-order terms in $1/\lambda^2$. The correction to the Higgs mass is quadratically divergent in Λ_{UV}, as are similar 1-loop diagrams coming from couplings to the heavy gauge bosons and the Higgs' self-interaction. The total effect of each of these contributions on the physical Higgs mass at 1-loop can be summarized as

$$m_H^2(\text{phys}) \approx m_H^2 + \frac{c}{16\pi^2}\Lambda_{UV}^2 , \qquad (8.2)$$

where m_H^2 is the parameter appearing in the Lagrangian, c depends on the various coupling constants of the SM and Λ_{UV} represents the ultraviolet completion of the SM, or the scale up to where the theory is valid. If the SM is to provide a description of nature all the way to the Planck scale, $M_P \sim 2.4^{18}$ GeV, then it naively seems that the physical Higgs mass should be of this same order. Perturbative unitarity arguments [124, 125] imply nearly the opposite, that the mass should be smaller than a few hundred GeV. The new boson discovered at 125 GeV by the CMS and ATLAS experiments [7, 8], if it is the Higgs, would confirm this fact. These considerations beg the question: how is the Higgs able to stay relatively light?

It is possible that the Lagrangian parameter m_H cancels the large Λ_{UV} term in the right-hand side of equation (8.2) but if Λ_{UV} is at the GUT scale $\sim 10^{16}$ this would require m_H to be fine-tuned to 1 part in 10^{26}. This is known as the *gauge hierarchy problem*. It is an issue of aesthetics and can be thought of as a question of symmetry. The masses of the fermions are protected by chiral symmetry, in that these masses break that symmetry, which in term protects them from quadratically divergent cutoff contributions. Similarly, quadratic divergences to the gauge boson
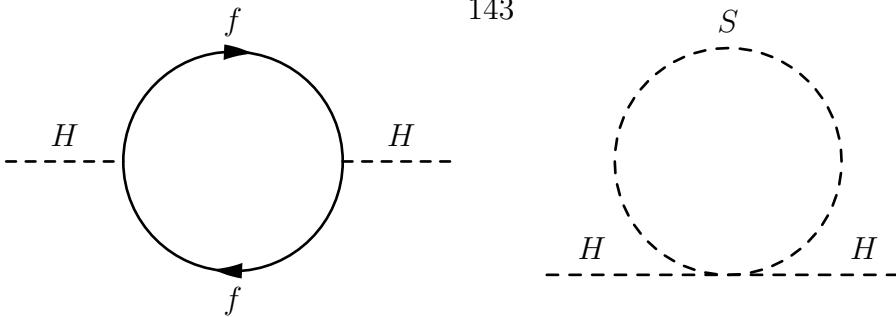
Figure 8.1: Examples of quadratically divergent Feynman diagrams contributing to the Higgs boson mass.

masses are removed through gauge-invariant dimensional regularization, such that they are prevented from being dragged to a high scale by the local gauge symmetries of the SM. On the other hand, there is no such symmetry in the SM which protects the Higgs mass. As a scalar, the number of degrees of freedom associated with a massive and massless particle are the same and hence, its mass will be sensitive to the UV completion of the theory in the absence of new BSM dynamics.

A new symmetry could keep the Higgs mass light by guaranteeing that the quadratically divergent 1-loop contribution is cancelled through the appearance of new particles in the theory with their own quantum corrections to the Higgs mass. If we consider a massive scalar (S) which also interacts with the Higgs, through a term $-\lambda_s |H|^2 |S|^2$ in the Lagrangian, then this particle will also result in 1-loop correction to the Higgs mass, illustrated in Fig. 8.1 (right), with the value calculated to be

$$\Delta m_H^2 = \frac{|\lambda_f|^2}{16\pi^2} \left[-2\Lambda_{UV}^2 + 6m_f^2 \log(\Lambda_{UV}/m_f) + \cdots \right] . \tag{8.3}$$

This contribution has an opposite sign relative to the fermion contribution; if there is a relation between the couplings λ_S and λ_f fixes their relative values of the right type the strongly divergent contributions from each particle will cancel. Such a symmetry is *supersymmetry* (SUSY) [126, 127]. For every SM particle there is an additional superpartner particle, with spin differing by 1/2, which is related by a new supersymmetry between bosons and fermions. In this case, the supersymmetry protects the scalar Higgs mass and ensures the cancellation of divergences.

There are other possible approaches to new symmetries which result in similar cancellations. Little Higgs [128, 129] models introduce new strong dynamics at scales above 10 TeV which include new same-statistic partners of the SM particles which cancel divergent effects. Some BSM possibilities do not even require these cancellations to mitigate the gauge hierarchy problem. Models of extra dimensions [130, 131] effectively reduce the the Planck scale by allowing gravity to propagate in a bulk other than the four dimensions we are most familiar with, such that its magnitude is geometrically reduced relative to the other forces. If this were the case, it could be that the Planck scale and the weak scale are actually the same, and the size of the extra dimensions puts Λ_{UV} at the same scale as the physical Higgs mass.

Each of these BSM theories suggest possible solutions to the gauge hierarchy problem through different approaches, but with strong similarities in their general implications. They all indicate that there is a strong reason to believe that *something* new happens at the TeV scale. Whether it is TeV^{-1} extra dimensions, or new symmetries manifested through new particles there should be new degrees of freedom which only appear at these higher energies related to the stabilization of the weak scale. As we shall see, the phenomenology of these BSM possibilities indicates that we may be able to infer their existence from new physics searches at the LHC.

8.2 Symmetric Possibilities

8.2.1 Supersymmetry

Supersymmetry (SUSY) is a hypothetical symmetry between fermions and bosons. In order to ensure cancellations to quadratically divergent contributions to scalar masses a new superpartner is added for each of the SM particles. The simplest SUSY model which reproduces the SM at the electroweak scale is the Minimal Supersymmetric Standard Model (MSSM) [132], and is a direct symmetrization of the SM fields, based on the $SU(3)_C \times SU(2)_L \times U(1)_Y$ gauge group.

Each of the SM particles are put into supermultiplets with their superpartners,

which have the same gauge quantum numbers. The fermion fields are promoted to chiral scalar superfields, with a $SU(2)$ superfield for the left-handed fermions and a singlet for the right-handed ones. Their scalar superpartners are sfermions (selectrons, smuons, staus), with one for each chirality. Similarly the gauge bosons are put into gauge superfields with their own fermionic superpartners, gauginos. The SM Higgs $SU(2)$ doublet becomes a doublet of left-chiral superfields. In order to give down-type fermions mass a second left-chiral superfield must be introduced, giving a more expansive Higgs sector than the SM. One of the SUSY Higgses is a light scalar, resembling the SM Higgs. The gauge-eigenstate fermionic superpartners of the SM analogues, winos and binos for the W and B fields, respectively, mix with the higgsinos to form mass eigenstates: Neutral gauginos and higgsinos mix to give four neutralinos, while the charged fields result in four charginos. An attractive feature of SUSY is that the three gauge couplings unify at the GUT scale, potentially pointing towards a unified theory.

In the SM, lepton and baryon number are conserved in all renormalizable interactions. For the MSSM, this is no longer the case since there are now scalars which carry these quantum numbers (the superpartners of the quarks and leptons) and B or L violating renormalizable interactions are possible which are also invariant under the SM gauge symmetries. The presence of such terms appearing in the Lagrangian is strongly constrained by experiment, particularly B-violation would result in decays of the proton, which are so-far unobserved. These terms can be forbidden by requiring that the theory is invariant under a new parity-like symmetry called R-parity, or matter parity. Under this new parity the the quark and lepton superfields are odd, while the Higgs and gauge superfields are even. The conserved R-parity can be expressed as

$$R = (-1)^{3(B-L)+2s} \ ,$$
(8.4)

where s is the spin of the field. This conserved $\mathbb{Z}_2$ parity has important phenomenological consequences for observing evidence of sparticles. It implies that there must

be an even number of superpartners in each interaction, implying that these sparticles are produced in pairs in LHC collisions. This same restriction means that the lightest supersymmetric particle (LSP) will be stable, unable to decay to SM particles. SUSY with weak-scale LSPs could potentially give a particle interpretation of dark matter. If produced in LHC collisions, the LSPs will escape the detector without interacting.

If SUSY exists, it must be a broken symmetry since sparticles with mass the same as their SM partners have been excluded experimentally. This implies that there are also supersymmetry-violating terms appearing in the Lagrangian. In order to maintain the cancellation of quadratic divergences these are restricted to *soft* breaking terms related to the scalar sparticle and gaugino masses. With these additional terms the MSSM has more than 100 new masses, phases and mixing angles relative to the SM.

Experimental search results are often interpreted in a simplified subspace of the MSSM called minimal supergravity (mSUGRA) [133]. Supergravity refers to the nature of SUSY breaking, which follows from a SUSY-violating hidden sector which communicates only through gravity with the SM. In the minimal SUGRA model, the first-two generation slepton and and gaugino mass matrices are assumed to be trivial, as are the Yukawa coupling matrices, with no complex phases in any soft terms, such that the full theory is described by only five parameters: Common soft mass parameters m_0 and $m_{1/2}$ for the sfermions and and gauginos, respectively, a universal trilinear coupling A_0 for the Yukawa interactions, the ratio of the Higgs' VEV's $\tan\beta$ and the sign of the Higgs mass parameter, μ. This tractable theory space is useful for presenting experimental results and is often referred to as the constrained MSSM or CMSSM.

There are a vast number of possible signatures for SUSY at the LHC. The hadron-hadron collider environment implies that the cross sections for strongly interacting sparticles will be larger than for the electroweak-inos. Heavy squarks and gluons could be produced, decaying through lighter sparticles to SM particles like heavy gauge bosons, leptons and jets. R-parity conservation leads to signatures of MET from escaping LSPs.

8.2.2 Little Higgs Models

An alternative approach to SUSY for explaining the lightness of the Higgs is a class of possibilities called Little Higgs Models [134–138]. Here, the Higgs is constructed as a pseudo-Goldstone boson of a nonlinearly realized approximate global symmetry, analogous to the pions and kaons in QCD. In these theories a larger symmetry group, like $SU(5)$, is broken to a smaller subgroup, like $SO(5)$. The broken symmetries result in a pseudo-Goldstone multiplet. The larger symmetry group must contain two copies of $SU(2) \times U(1)$ and two additional subgroups which contain generators transforming like $SU(2)$ doublets which commute with the former. It is then constructed that only the *combination* of both weak gauge interactions breaks all global symmetries acting on the Higgs, meaning the quadratically divergent contributions must involve both couplings, and can only appear at two loops. The Higgs is radiatively stable up to a cut-off of around 10 TeV while having gauge, Yukawa and self-interactions of order one. This is accomplished, as in SUSY, by adding new partners of the SM particles related by a new symmetry.

Like SUSY, some of the possible interactions associated with these Little Higgs models conflict with experimental constraints. These interactions can be removed by appealing to a new conserved parity, called T-parity. The new, heavy, particles associated with one copy of $SU(2) \times U(1)$ are T-parity even, while the SM fields from the other are T-parity odd. As for SUSY, this $\mathbb{Z}_2$ symmetry results in a potential dark matter candidate, as the lightest T-parity odd particle will be stable and weakly interacting. The dominant production mode in these models are new colored particles, analogous to the squarks and gluinos in SUSY. Similarly, there can appear heavy T-odd partners of the leptons and electroweak gauge bosons.

8.2.3 Extra Dimensions

Theories of extra dimensions feature a very different approach to explaining the hierarchy between the electroweak and Planck scales. If the SM is confined to a four dimensional space, but there are additional dimensions in which gravity can propagate,

the weakness of gravity can follow as a geometric consequence. These include models of compact extra dimensions [139–145] and also with warped geometries, where the hierarchy between scales is generated by large curvature of extra dimensions [146]. It is also possible to embed supersymmetry, or other BSM symmetry theories into these scenarios [147, 148].

In models of compact extra dimensions with small enough size, the SM fields can propagate in these compact dimensions resulting in a Kaluza-Klein (KK) tower of excitations from each SM particle. At the first level, the SM bosons have a new particle partner while each SM fermion has two. With one extra dimension, each of these new particles is odd under a KK-parity which is a remnant of the broken translational invariance in this fifth dimension. The lightest KK-odd particle is weakly interacting and due to this conserved $\mathbb{Z}_2$ symmetry would be a candidate for matter. At the LHC, a variety of KK odd partners could be pair produced if additional interactions are present in the model, with mass spectra potentially resembling SUSY.

8.3 BSM Phenomenology

While different in the physics they encompass, the phenomenological properties of models which attempt to mitigate the gauge hierarchy problem are very similar. Each introduces a spectrum of new particles which are partners with their SM counterparts under a new symmetry of nature. In the cases where experimentally disfavored interactions are removed by appealing to a further symmetry, each includes a new conserved quantum number or parity. This feature is particularly desirable because it means the lightest of these new particles is unable to decay to SM particles, so it is weakly interacting and could maybe explain the dark matter scattered throughout the universe. These BSM models also make predictions about the scale of new physics. In explaining why the electroweak and Planck scales are not the same, they predict the appearance of new, heavy degrees of freedom near the electroweak scale. In LHC collisions, this could be manifested through the production of new, heavy particles

which have exceeded the energy capabilities of previous accelerators. These particles can interact with the SM in a variety of ways, leading to a rich collection of possible BSM signatures.

In the searches for new physics described in chapters 9 and 10, two common characteristics of these hypothetical events are exploited. As is the case for W and Z particles, the decays of new massive particles can be identified by reconstructing the mass, or a mass-sensitive variable, from the decay products detected in these event. An excess of events at a particular mass not explicitly present in the SM would indicate the discovery of a new one. Hence the scale of new physics can be used to discover it. The conserved $\mathbb{Z}_2$ parities in these models imply that these new particles must be produced in pairs, and that at the end of each of their the decay chains at least one weakly interacting particle must appear (or disappear). This implies a very particular topology for LHC events: two new massive particles each decaying to a system of detectable SM particles and the appearance of missing transverse momentum. The razor variables used in these new physics searches were designed specifically to study this signature.

It can be argued that this $\mathbb{Z}_2$-parity-inspired topology is more general than the models discussed. The existence of dark matter, without an SM explanation, indicates that there is another type of matter that is yet to be identified. Its weakly interacting nature suggests that if this new particle(s) is heavier than some of the SM particles that there is a new, at least approximate, symmetry preventing its decay. The absence of a discovery at previous experiments, operating at lower energies, implies that any new particles which do interact with the SM have large masses. A light Higgs hints that the SM should interact with new degrees of freedom around the TeV. Occam's razor implies that each of these indications of BMS phenomena should be related, making searches at the LHC targeting this signature a well-motivated approach for observing evidence of BSM physics with new particle spectra satisfying these general characteristics.

Of course, the interactions of the SM are complicated, and the new physics which could appear at the LHC might be no different. The models discussed all contain

an additional copy of the SM, and have an enormous variety of possible interactions involving new particles. General consideration about the production and decay topologies of these models should be complemented by searching for anomalously large numbers of SM particles in these events, like leptons, jets and gauge bosons, which could come from the decays of new particles.

Chapter 9

Razor Searches for Supersymmetry

In chapter 8 several theories of new physics were described which mitigate perceived problems in the Standard Model, particularly related to the naturalness of the scales appearing and the amount of fine-tuning of theory parameters necessary to realize them. In each of these BSM theories, this is achieved by predicting new symmetries of nature which generally imply the existence of new, undiscovered fundamental particles and interactions which we would like to study at the LHC.

Of particular interest are theories that include a discrete $\mathbb{Z}_2$ symmetry, or a new type of parity quantum number like R-parity in SUSY, T-parity in Little Higgs models and KK-parity in models with extra dimensions. These theories are phenomenologically appealing since the $\mathbb{Z}_2$ symmetry often forbids interactions problematic to precision electroweak constraints from appearing in the Lagrangian. Additionally, these symmetries can prevent the lightest new particle from decaying, resulting in a possible particle explanation for the abundance of dark matter in the universe. In the following chapter, we describe a search for new particles associated with these BSM possibilities using the CMS detector with 35 pb^{-1} of $\sqrt{s} = 7$ TeV pp collision data. New event kinematic variables, denoted *razor* variables, are derived specifically for discovering and characterizing new BSM particles through interactions motivated by $\mathbb{Z}_2$ symmetries. We describe the phenomenology of the Standard Model in terms of razor variables and how this is used to infer the presence or absence of anomalous

events in our data sample in hadronic and leptonic final states. The results of this search are interpreted R-parity conserving SUSY scenarios, putting constraints on sparticle masses in hypothetical models.

9.1 Razor Kinematic Variables

In LHC pp collisions, a $\mathbb{Z}_2$ symmetry implies that new BSM particles can only be produced in pairs. Once produced, each of these new particles can decay through a cascade of SM particles, with an odd number of BSM particles appearing at each decay step. These decays can proceed down the new particle mass spectrum until reaching the lightest new particle which, since it is stable and will not decay to SM particles, is weakly interacting and will not be detected. This general BSM event signature is illustrated in figure 9.1.
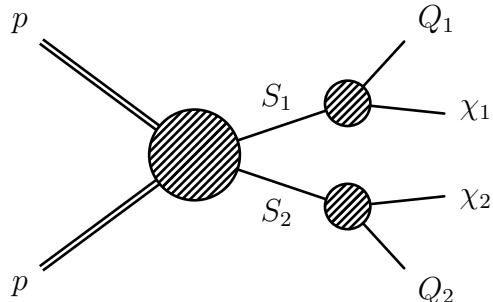
Figure 9.1: $\mathbb{Z}_2$ symmetry motivated BSM LHC event signature. Two massive particles, S_i, are produced in a pp collision and each decay to a system of detectable SM particles, Q_i, and a system of weakly interacting particles, χ_i.

In the past years the development of kinematical variables that assist the discovery of this type of event topology has been intense and rich [87,149–162]. In general, these methods try to exploit one or more characteristic features of these events which can distinguish them from similar ones with only SM particles. Coming from different decays, the systems of visible particles Q_i do not have to recoil against each other in momentum, as they would for SM processes without weakly interacting particles in the final state. Similarly, the presence of the particles represented by χ_i in these BSM events can be inferred by looking at the imbalance of transverse momentum

among the visible particles.[1] Finally, the fact that we haven't discovered these new particles in the past indicates that they are probably massive; the *raison d'être* for these new particles and symmetries, explaining the relation between the weak and Planck scales, implies they should generally have masses close to the weak scale. We derive two new, complementary kinematic variables which are independently sensitive to different distinguishing characteristics of these events, M_R and R. The variable M_R is sensitive to the *scale*, or masses, of these new physics particles. The dimensionless variable R, the razor, indicates the amount of transverse imbalance of momentum in collision events, a signature of independent decay chains initiated by pair-produced particles and decaying to weakly interacting particles.

9.1.1 The Scale of New Physics: $\mathbf{M_R}$

In order to derive the kinematic variable M_R, we consider the simplest topology corresponding to figure 9.1, where the interactions at the $S_i Q_i \chi_i$ vertices are direct two-body decays $S_i \to Q_i \chi_i$. This situation could represent, for example, the pair-production of right-handed squarks where each squark decays directly to a light quark and a weakly interacting, potentially massive, neutralino. If an event like this occurred in an LHC collision, the momentum of the decay products Q_i would be measured (if if they are within the angular acceptance of the detector) while the particles χ_i would escape undetected, taking their momentum with them. We assume here that $M_{S_1} = M_{S_2} \equiv M_S$ and $M_{\chi_1} = M_{\chi_2} \equiv M_\chi$. This means that, in each of these events, there are two new types of particles, S and χ, each with an unknown mass.

What we would like to have is a procedure, event by event, for accurately reconstructing the masses M_S and M_χ. If these masses are sufficiently different from the masses of the SM particles then we could use this information to distinguish events with these two new particles present from SM background events, looking for excesses at fixed values in the reconstructed M_S and M_χ distributions. Unfortunately this is

---

[1]Since the interaction which produces the particles S_i is initiated by the parton constituents of protons, which do not carry all of the protons' momentum, conservation of momentum will only approximately hold among the visible particles in the directions transverse to the beam-line in a collision event.

not possible. Since the particles χ_i escape undetected the kinematics of the particles M_S and M_χ cannot be reconstructed using only the visible Q_i decay products of the the system. The reason is that the event is kinematically underconstrained by the measurements we are able to make. As we shall see shortly, in this case the number of kinematic degrees of freedom (d.o.f.), N, and the number of kinematic constraints available in each reconstructed event, M, satisfy $N - M = 4$. This means that we would need four more kinematic constraints to fully solve the system (possibly only up to some number of discrete ambiguities); even if we knew the masses M_S and M_χ a priori we could not completely reconstruct these events.

With these considerations in mind, we can ask a slightly less ambitious question: Is there a *characteristic scale*, related to M_S and M_χ, that we can partially reconstruct event by event? The answer is yes, and to understand what this characteristic scale is we consider the different reference frames relevant to these types of events, illustrated in figure 9.2.
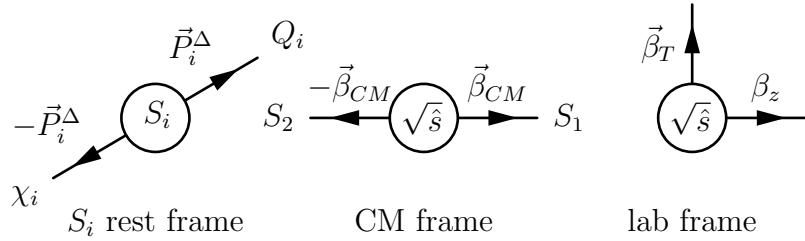
Figure 9.2: The four reference frames describing the pair production of particles S_i, each decaying $S_i \to Q_i \chi_i$. (Left) In each of the respective S_i rest frames the particles Q_i and χ_i are traveling with equal and opposite momentum, with the magnitude of their momentum set by the particle mass differences. (Center) In the CM frame, the two particles S_1 and S_2 are traveling with equal and opposite velocities β_{CM}, with $\sqrt{\hat{s}}$ representing the $S_1 + S_2$ CM object. (Right) In the laboratory frame, the CM system is traveling with a longitudinal velocity β_z due predominantly to differences in the longitudinal momentum of the interacting partons. In the transverse plane, the CM system can have a nonzero velocity $\vec{\beta}_T$ coming from other final state particles recoiling against the CM system.

Since each S_i is undergoing a two-body decay, the decay products are traveling with equal and opposite momentum in the S_i rest frame. If the masses M_S and M_χ are the same event by event, the magnitude of the decay products' momentum in the

S_i rest frames is also fixed. Assuming that the objects Q_i are approximately massless we observe that

$$|\vec{P}_1^{\Delta}| = |\vec{P}_2^{\Delta}| = \frac{M_{\Delta}}{2} \equiv \frac{1}{2} \frac{m_S^2 - m_{\chi}^2}{m_S} \ . \tag{9.1}$$

For this event topology, M_{Δ} is the characteristic scale. It is the same for each event of this type, and the momentum of the final state particles in the laboratory frame will be related to its value. To calculate M_{Δ} we must reconstruct an approximation of the S_i rest frames based only on our observations in the laboratory frame. This is accomplished through a series of physics-motivated approximations which effectively eliminate the extra unknown d.o.f. from the problem.

Let us denote the four-vectors of the particles S_i as s_1 and s_2. Similarly, we call the four-vectors of the Q_i and χ_i particles q_i and ν_i, respectively. Throughout this discussion, superscripts on these four-vectors (and corresponding three-momenta) indicate which reference frame they correspond to. In the $S_1 S_2$ rest frame (CM frame) the S_1 and S_2 four-vectors are given by

$$
\begin{aligned}
p[S_1] &\equiv s_1^{CM} = M_S \, \gamma_{\text{CM}} \{1, \vec{\beta}_{\text{CM}}\} = \{E_{S_1}^{CM}, \vec{s}_1^{\ CM}\} \ , \\
p[S_2] &\equiv s_2^{CM} = M_S \, \gamma_{\text{CM}} \{1, -\vec{\beta}_{\text{CM}}\} = \{E_{S_2}^{CM}, \vec{s}_2^{\ CM}\} \ ,
\end{aligned} \tag{9.2}
$$

such that $(s_1 + s_2)^2 = \hat{s} = 4\gamma_{CM}^2 M_S^2$, where $\hat{s}$ is the usual Mandelstam variable describing the hard partonic subprocess. The boost $\vec{\beta}_{CM}$, and corresponding γ_{CM} indicate how far above the $2M_S$ energy threshold the two S_i are produced. We need not consider off-shell production of the particles S_i here.

In their respective S_i rest frames (S-frames), the decay products of each S_i have four-momenta defined as

$$
\begin{aligned}
p[Q_i] &\equiv q_i^S = \frac{M_{\Delta}}{2} \{1, \hat{u}_i\} = \{E_{Q_i}^S, \vec{q}_i^{\ S}\} \ , \\
p[\chi_i] &\equiv \nu_i^S = \frac{M_{\Delta}}{2} \{R_{S\chi}, -\hat{u}_i\} = \{E_{\chi_i}^S, \vec{\nu}_i^{\ S}\} \ ,
\end{aligned} \tag{9.3}
$$

where $R_{S\chi} = \frac{M_S^2 + M_{\chi}^2}{M_S^2 - M_{\chi}^2}$ and each $\hat{u}_i$ is a unit vector. In the laboratory frame (l-frame),

the momentum of the particles $\vec{q_i}^{\,l}$ and $\vec{\nu_i}^{\,l}$ are related to their momenta in their respective S frames through a series of Lorentz transformations. To move from the rest frame of S_1 (S_2) to the CM frame, q_1^S and ν_1^S (q_2^S and ν_2^S) are boosted to a frame traveling at a velocity $\vec{\beta}_{\text{CM}}$ ($-\vec{\beta}_{\text{CM}}$) with respect to the S_1 (S_2) rest frame. Finally, to move from the CM frame to the lab frame, each of the final state particles is boosted to a frame traveling at a velocity $\vec{\beta}_L = (\vec{\beta}_T, \beta_z)$, where $\vec{\beta}_T$ and β_z are the transverse and longitudinal components of this boost, respectively. The transformations taking the final state particles from their respective S_i rest frames to the lab frame can be schematically described as

$$q_1^S, \nu_1^S \xrightarrow{\vec{\beta}_{\text{CM}}} q_1^{CM}, \nu_1^{CM} \xrightarrow{\vec{\beta}_L} q_1^l, \nu_1^l \ ,$$
$$q_2^S, \nu_2^S \xrightarrow{-\vec{\beta}_{\text{CM}}} q_2^{CM}, \nu_2^{CM} \xrightarrow{\vec{\beta}_L} q_2^l, \nu_2^l \ . \tag{9.4}$$

This series of Lorentz transformations is equivalent to moving through the reference frames of figure 9.2 from left to right. To calculate M_Δ in the respective S_i-frames we would need to perform the inverse series of transformations to the particles Q_i which we have measured in the lab frame. With this goal in mind, we observe that the under-constrained d.o.f. in this problem can be expressed in terms of these unknown boosts. In the final state, the two escaping χ particles represent $4 + 4 = 8$ unknown d.o.f., in that we don't know their three-momenta and masses. The constraints $M_{S_1} = M_{S_2}$ and $M_{\chi_1} = M_{\chi_2}$ yield two constraints, while assuming conservation of momentum in the direction transverse to the beam axis provides two additional constraints, leaving four under-constrained d.o.f.. Viewing the problem in terms of the unknown boosts $\vec{\beta}_{CM}$ and $\vec{\beta}_L$, we observe that these four d.o.f. are equivalent to the direction and magnitude of $\vec{\beta}_{CM}$ (three d.o.f.) and the longitudinal component of $\vec{\beta}_L$ (1 d.o.f.).[2] While these extra d.o.f. cannot be constrained by kinematic measurements, they can be eliminated through approximations motivated by the underlying physics of these events.

---

[2]If $\vec{\beta}_{CM}$ and the longitudinal component of $\vec{\beta}_L$ are known then conservation of transverse momenta can be used to calculate the transverse part of $\vec{\beta}_L$ up to a discrete ambiguity.

In practice, $|\vec{\beta}_T| \sim p_T^{CM}/\sqrt{\hat{s}} \leq p_T^{CM}/2M_S$, where p_T^{CM} is the transverse momenta of the CM system resulting from initial state radiation and underlying event particles. Even though p_T^{CM} scales with $\sqrt{\hat{s}}$, the spectrum of $|\vec{\beta}_T|$ will get softer for increasingly large values of M_S. Motivated by expectations of large masses M_S in the models we are searching for we approximate $\vec{\beta}_T \to 0$. Similarly, if the mass M_S is sufficiently large relative to the hadron-hadron collider energy $\sqrt{s}$, the particles S_1 and S_2 will be mostly produced near the $\sqrt{\hat{s}} \sim 2M_S$ threshold, such that $\gamma_{CM} \sim 1$. The approximation $\gamma_{CM} = 1$ implies that the particles S_i are produced exactly at threshold, with $\vec{\beta}_{CM} \to 0$.

With the $\gamma_{CM} = 1$ approximation the S_i rest frames and the CM frame are the same, and we denote this reference frame the *rough approximation frame*, or *R*-frame. In the *R*-frame the particles Q_i satisfy the constraint

$$|\vec{q}_1^{\,R}| = |\vec{q}_2^{\,R}| = \frac{M_\Delta}{2} \ . \tag{9.5}$$

Additionally, with the $\vec{\beta}_T \to 0$ approximation the *R*-frame is now related to the lab frame by a single longitudinal boost, which we will denote β_R. We can solve for β_R by using the constraint in equaton (9.5) and calculating the longitudinal boost that will move the visible particles Q_i to a reference frame where the magnitude of their momenta is equal. We find that

$$\beta_R = \frac{|\vec{q}_1^{\,l}| - |\vec{q}_2^{\,l}|}{q_{1z}^l - q_{2z}^l} \ . \tag{9.6}$$

Finally, we define the *R*-frame mass, M_R, as

$$M_R \equiv 2|\vec{q}_1^{\,R}| = 2|\vec{q}_2^{\,R}| = 2\sqrt{\frac{(|\vec{q}_1^{\,l}|q_{2z}^l - |\vec{q}_2^{\,l}|q_{1z}^l)^2}{(q_{1z}^l - q_{2z}^l)^2 - (|\vec{q}_1^{\,l}| - |\vec{q}_2^{\,l}|)^2}} \ . \tag{9.7}$$

If our approximations hold the *R*-frame will be equivalent to the two S_i rest frames and our variable M_R will be equal to M_Δ.

We have derived a variable, M_R, sensitive to the characteristic scale M_Δ which

we can calculate using only the visible particle momenta $\vec{q}_i^{\,l}$ measured in the lab frame. In the limit that our approximations hold, $\gamma_{CM} \to 0$ and $\vec{\beta}_T \to 0$ imply that $M_R \to M_\Delta$. We observe that M_R is invariant under longitudinal boosts such that even if $\gamma_{CM} \neq 1$, M_R is independent of the true value of β_z. In the following sections we will examine the efficacy of this approximate reconstruction approach, testing the other properties of M_R with a collection of toy models.

9.1.2 The $\gamma_{\mathrm{CM}} = 1$ Approximation

In order to understand how well our $\gamma_{CM} = 1$ approximation holds in practice we must consider how γ_{CM} is distributed in reality and what happens to M_R when $\gamma_{CM} \neq 1$. To address the former question, we consider the simple model with two scalar particles: Φ_0 with zero mass and Φ_1 with mass M_S, where these scalars are approximating the interacting partons and S_i particles from figure 9.1, respectively.

We consider contact interaction pair production of Φ_1 through a $\lambda|\Phi_0|^2|\Phi_1|^2$ vertex, where λ is the dimensionless $\Phi_0^2\Phi_1^2$ coupling, which we set to 1 without a loss of generality. The subprocess cross section is proportional to

$$\hat{\sigma}(\hat{s}) \propto \lambda^2 \frac{\sqrt{1 - 4M_S^2/\hat{s}}}{\hat{s}} \propto \lambda^2 \frac{\sqrt{1 - 1/\gamma_{CM}^2}}{\gamma_{CM}^2 M_S^2} \ . \tag{9.8}$$

From equation (9.8) we observe that $\gamma_{CM} = 1$ is kinematically forbidden, and that the cross section for the subprocess will decrease asymptotically as $1/\gamma_{CM}^2$.

Additional suppression of large values of γ_{CM} is caused by the parton distribution functions (PDFs) in hadron-hadron collisions. Assuming the two initial state Φ_0 particles are partons from colliding protons with momentum fractions x_a and x_b respectively, and PDFs $f_1(x)$ and $f_2(x)$ we can write the total cross section as

$$\frac{d\sigma}{dx_a dx_b} \propto [f_1(x_a)f_2(x_b) + a \leftrightarrow b]\hat{\sigma}(\hat{s} = sx_ax_b) \ , \tag{9.9}$$

where s is the proton-proton CM energy. Changing variables from x_b to γ_{CM} through the relation $sx_ax_b = 4\gamma_{CM}^2 M_S^2$ and integrating over x_a we find that the differential

cross section with respect to γ_{CM} is given by

$$\frac{d\sigma}{d\gamma_{CM}} \propto \frac{\sqrt{1-1/\gamma_{CM}^2}}{s\gamma_{CM}} \times$$

$$\int_{\frac{4\gamma_{CM}^2 M_S^2}{s}}^{1} [f_1(x_a)f_2(\frac{4\gamma_{CM}^2 M_S^2}{sx_a}) + a \leftrightarrow b]\frac{dx_a}{x_a}. \tag{9.10}$$

In figure 9.3 we show the probability distribution function for γ_{CM} for $\sqrt{s} = 7$ TeV pp collisions, where we have numerically integrated equation (9.10) for $q\bar{q}$-like (u and sea quark PDFs) and gg-like production. We use PDF parameterizations of the form $xf_i(x) = A_i x^{\delta_i}(1-x)^{\eta_i}(1+\epsilon_i\sqrt{x}+\gamma_i x) + A'_i x^{\delta'_i}(1-x)^{\eta'_i}$ with NNLO parameters determined from a global PDF fit at $Q^2 = 1$ GeV^2 [163]. Larger values of M_S result in lower values of γ_{CM}, with all distributions peaking at approximately $\gamma_{CM} \sim 1.1$ and falling quickly with increasing γ_{CM}. The PDFs are fast-falling functions of x, resulting in a steeply-falling γ_{CM} distribution. We conclude that, for nonresonant particle pair production, the $\gamma_{CM} = 1$ approximation is quite good.
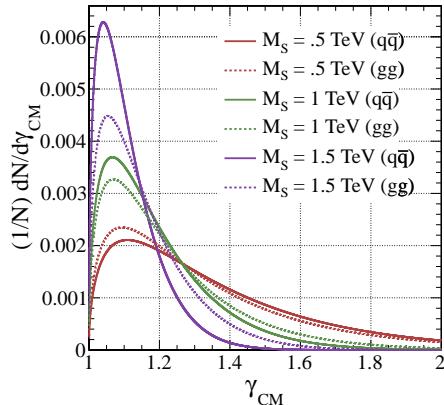
Figure 9.3: Distribution of γ_{CM} for $q\bar{q}$-like and gg-like production at $\sqrt{s} = 7$ TeV for different values of M_S.

The exact dependence of the subprocess cross section on γ_{CM} will vary depending on the nature of the interacting final and initial state particles in the $2 \to 2$ process, but the resulting distribution of γ_{CM} should be qualitatively similar to the result shown in figure 9.3: γ_{CM} exactly equal to 1 is kinematically forbidden, but values of

γ_{CM} near 1 are preferred to larger values due to the falling subprocess and total cross sections with increasing $\sqrt{\hat{s}} = \sqrt{s x_a x_b} \propto \gamma_{CM}$.

In order to understand the behavior of M_R when $\gamma_{CM} \neq 1$ we return to the example introduced in section 9.1.1. Using the same notation, we again consider the pair production of massive particles S_1 and S_2, and continue to use the approximation $\vec{\beta}_T \to 0$, this time with γ_{CM} not equal to 1. We consider a toy simulation of S pair production, with decays $S_i \to Q_i \chi_i$, where we have taken flat matrix elements for the angular distribution of the S_i decay products (decay axis randomly distributed as a sphere in the S_i rest frame). The resulting distributions of M_R, for different fixed values of γ_{CM}, are shown in figure 9.4. We observe that the peak value of M_R scales as $\gamma_{CM} M_\Delta$, with the width of the M_R distribution increasing with γ_{CM}. Hence, in practice, the distribution of M_R will peak near M_Δ, even when $\gamma_{CM} \neq 1$, with resolution degrading with increasing γ_{CM}.
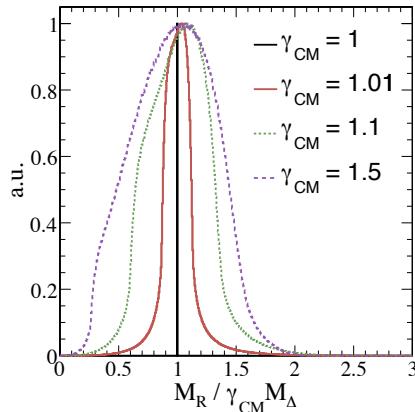
Figure 9.4: Distribution of M_R, in units of $\gamma_{CM} M_\Delta$, for different values of γ_{CM}. Distributions are normalized such that their maximum value is equal to one.

9.1.3 The Razor R

In section 9.1.1, we derived the kinematic variable M_R which is sensitive to the mass difference M_Δ for events of the type shown in figure 9.1, a property that can be used to distinguish these events from SM background processes. Unfortunately, for

searches in most final states this is not sufficient to suppress large backgrounds; the SM production cross section at $\sqrt{\hat{s}} \sim M_\Delta$ is generally much larger than the cross section of the models we are searching for. In order to be able to identify events with the production of new particles, we can exploit a characteristic of the class of $\mathbb{Z}_2$ symmetry model we are looking for: pair production of particles each independently decaying to a weakly interacting particle, leading to a transverse imbalance of momentum in the event.

In order to understand how this property can be used to distinguish these events from SM backgrounds, we consider a search for the pair production of squarks, each decaying to a quark jet and a weakly interacting particle. The largest background to this two jet and missing transverse momentum final state is QCD dijet production, where nonzero missing transverse energy can result from instrumental backgrounds, jet mis-measurements, finite detector acceptance and non-Gaussian tails in the detector response, in addition to the production of neutrinos. To understand how the variable M_R will behave for backgrounds of this type we consider the simple case of QCD dijet production in more detail. In the dijet rest frame, we express the two jets' four-vectors as

$$
\begin{aligned}
k_1 &= \frac{\sqrt{\hat{s}}}{2} \{1, \hat{v}\} \ , \\
k_2 &= \frac{\sqrt{\hat{s}}}{2} \{1, -\hat{v}\} \ ,
\end{aligned}
\tag{9.11}
$$

where $\sqrt{\hat{s}}$ is the dijet invariant mass and $\hat{v}$ is a unit vector. If we assume that the Lorentz transformation from the dijet rest frame to the laboratory frame is simply a longitudinal boost, β_z, (the CM system has no transverse momentum) we find that for this type of event $M_R = \sqrt{\hat{s}}$. Therefore, M_R will be distributed as $\sqrt{\hat{s}}$ for this background process, falling steeply, while the signal distribution will peak near M_Δ. The question of whether or not we can identify signal events in the presence of this background becomes a question of whether the effective dijet cross section is sufficiently small for $\sqrt{\hat{s}}$ in the range of the signal peak around M_Δ, which it is most likely not.

Examining the expression for M_R in equation (9.7), we see that there is additional kinematical information in our events that we have not yet exploited. For example, M_R is independent of the azimuthal angle, $\Delta\phi$, between the two final state jets. In QCD dijet events the jets should be largely back to back in the transverse plane, with $\Delta\phi$ peaking at π. On the other hand, the two jets in the SUSY signal events result from the decay of two separate squarks, implying that their direction in the transverse plane is largely independent of each other apart from spin-correlations and effects resulting from $\vec{\beta}_{CM} \neq 0$. Hence, the distribution of $\Delta\phi$ for signal events will be significantly flatter than for the background. Rather than simply cutting on the variable $\Delta\phi$, we incorporate this information into a new variable denoted M_T^R.

In this particular final state, we assume that in signal events there are two escaping weakly interacting particles χ_1 and χ_2 with four momenta ν_1^l and ν_2^l, with each particle "paired" with an observed jet with four-momenta q_1^l and q_2^l, respectively. From these four-vectors we define the variable $M_{2S} = \sqrt{(1/2)[(\nu_1^l + q_1^l)^2 + (\nu_2^l + q_2^l)^2]}$, which is equal to M_S for signal events. The only constraint we have on the four-vectors ν_i^l is that the vectorial sum of their transverse momenta should be equal to the observed missing transverse energy, $\vec{M}$. Setting $(\nu_i^l)^2 = 0$ (if only because we don't know it *a priori*) and minimizing M_{2S} over ν_{1z}^l and ν_{2z}^l yields

$$\min_{\nu_{iz}} M_{2S} = \sqrt{|\vec{q}_{1T}^{\,l}||\vec{\nu}_{1T}^{\,l}| - \vec{q}_{1T}^{\,l} \cdot \vec{\nu}_{1T}^{\,l} + |\vec{q}_{2T}^{\,l}||\vec{\nu}_{2T}^{\,l}| - \vec{q}_{2T}^{\,l} \cdot \vec{\nu}_{2T}^{\,l}} \ . \tag{9.12}$$

Motivated by the backgrounds we are considering, where missing transverse momentum often results from imperfect measurements of the jets' momenta, we assign half of the measured missing transverse momenta to each escaping particle such that $\vec{\nu}_{1T}^{\,l} = \vec{\nu}_{2T}^{\,l} = \vec{M}/2$ and define M_T^R as:

$$M_T^R \equiv \min_{\nu_{iz}} M_{2S}\Big|_{\vec{\nu}_{1T}^{\,l}=\vec{\nu}_{2T}^{\,l}} = \sqrt{\frac{|\vec{M}|}{2}(|\vec{q}_{1T}^{\,l}| + |\vec{q}_{2T}^{\,l}|) - \frac{1}{2}\vec{M} \cdot (\vec{q}_{1T}^{\,l} + \vec{q}_{2T}^{\,l})} \ . \tag{9.13}$$

Like M_R, the variable M_T^R also contains information about the scale of the process we are studying. If we assume that $\gamma_{CM} = 1$ then the M_T^R distribution has a kinematic

endpoint at M_Δ for signal events. We note that M_T^R is an additional measurement of the scale of the process that uses information independent of the M_R. Therefore, rather than cutting on M_T^R we form the dimensionless R-frame razor, R, as the ratio of M_T^R and M_R, such that $R \equiv M_T^R/M_R$. For the signal process, the distribution of R peaks near 0.5, since this is the ratio of two measurements of the same scale, M_Δ, with an additional geometrical factor due to the fact that M_T^R contains only transverse information. For the QCD dijet background, if $\vec{M} = 0$, then R is 0, for any value of $\sqrt{\hat{s}}$.

As was discussed previously, there are several mechanisms for the measurement of $\vec{M}$ to be nonzero in QCD dijet events. For example, one or both jets in the final state could be mis-measured due to calorimeter noncompensation, uninstrumented regions of the detector or weakly interacting particles, causing an imbalance in the event and resulting in nonzero missing transverse momentum. To evaluate how these possibilities affect the measured values for M_R and M_T^R in background events, we return to equation (9.11) which describes the kinematics of the dijet system in it's CM frame. We now realistically assume that the measured jet momenta, q_i^l, are scaled relative to their true values, so that $q_i^l = f_i k_i$. Here, we are assuming that the direction of the two jets is not changed, but rather that only a fraction f_i of the jets' momentum is observed, where $f_i > 0$, while $1 - f_i$ is incorrectly interpreted as missing transverse momentum. Additionally, without loss of generality we adopt the convention $f_1 \geq f_2$.

With these mis-measurements, we find that M_R takes a value:

$$M_R = \sqrt{\frac{4f_1^2 f_2^2 \hat{s}(\hat{v} \cdot \hat{z})^2}{(f_1 + f_2)^2(\hat{v} \cdot \hat{z})^2 - (f_1 - f_2)^2}} \ , \tag{9.14}$$

independent of the longitudinal boost, β_z, that takes the jets from their CM frame to the laboratory frame. The missing transverse energy can now be nonzero, with $\vec{M} = (f_2 - f_1)\vec{k}_{1T}$ and M_T^R can be expressed as

$$M_T^R = \sqrt{(f_1 - f_2)f_1 \frac{\hat{s}(1 - (\hat{v} \cdot \hat{z})^2)}{4}} \ . \tag{9.15}$$

From equation (9.14) we see that these mis-measurements *decrease* the value of M_R, assuming that $f_1 \lesssim 1$. Therefore the distribution of M_R for the background will not have events promoted to the tail of the distribution due to these types of mis-measurements; instead, these mis-measurements will suppress the background M_R distribution. Furthermore, if we require that $R > C$, where C is some cut value, this implies that $CM_R < M_T^R$. To understand the effect of this cut, we change variables $(\hat{v} \cdot \hat{z})^2 = \cos(\theta_1)^2$ and $f_1 = f_2 \cos(\theta_2)^2$. With these substitutions, the inequality $CM_R < M_T^R$ can be re-expressed as

$$16C^2 \cos(\theta_1)^2 \cos(\theta_2)^4 + \sin(\theta_1)^4 \sin(\theta_2)^6 < \tag{9.16}$$
$$4\sin(\theta_1)^2 \cos(\theta_1)^2 \sin(\theta_2)^2 \cos(\theta_2)^2 \ .$$

This inequality implies that if $C \geq 1/2$, no background events of this type will satisfy the requirement on R. If $C \sim 0.4$, some events can pass, but M_T^R will reach its allowed maximum, for fixed $\sqrt{\hat{s}}$, at $M_T^R \sim \sqrt{\hat{s}}/5$, with the razor inequality implying that $M_R < M_T^R/C \lesssim \sqrt{\hat{s}}/2$. Hence for this type of background event to result in $M_R \sim M_\Delta$, it must have $\sqrt{\hat{s}} > 2M_\Delta$. Therefore, we observe that adding a requirement on R to our event selection will remove most QCD dijet events with mis-measurements of the type described above.

Another possibility resulting in nonzero missing transverse momentum in these background events is that there are additional particles, whose vectorial sum of transverse momentum is nonzero, and that escape detection. For example, jets resulting from initial state radiation could remain unseen due to limited detector acceptance, causing a transverse imbalance in the visible momentum in the event. In order to understand the effect of this type of background on M_R and M_T^R, we consider a simple example. We denote the vectorial sum of the transverse momentum of particles escaping detection as $\vec{P}_T$. Returning again to the QCD dijet example described by equation (9.11), a nonzero value of $\vec{P}_T$ will result in two significant changes to the final state particle kinematics. Firstly, the missing transverse energy will be nonzero, with $\vec{M} = \vec{P}_T$. Secondly, this missing momentum will result in the dijet system un-

dergoing an additional transverse boost when moving from the dijet rest frame to the laboratory frame (any additional contribution to the longitudinal momentum imbalance in the event is absorbed into the longitudinal boost, β_z, which moves the dijets from their CM frame to the laboratory frame). Specifically, the dijets are moved to a frame traveling at a velocity $\vec{\beta} = \vec{M}/(\gamma\sqrt{\hat{s}})$, where $\gamma = (1 - |\vec{\beta}|^2)^{1/2}$ and $\sqrt{\hat{s}}$ is the dijet invariant mass. In this case, M_R is given by

$$M_R = \gamma\sqrt{\hat{s}} \left(1 - \frac{\gamma^2(\vec{\beta} \cdot \hat{v})^2}{(\hat{v} \cdot \hat{z})^2}\right)^{-1/2} , \qquad (9.17)$$

while M_T^R can be expressed as

$$M_T^R \sim \sqrt{\frac{\gamma\beta\hat{s}(\sqrt{(1 - (\hat{v} \cdot \hat{z})^2} + \gamma\beta)}{2}} , \qquad (9.18)$$

where the approximate equality holds up to order β^2. We observe that that for fixed $\sqrt{\hat{s}}$, after applying a requirement on R, remaining background events will have M_R with an upper bound that goes as $\sqrt{\gamma\beta\hat{s}}$ if the jets have a large transverse component in their rest frame, otherwise as $\gamma\beta\sqrt{\hat{s}}$. Recalling that $\gamma\beta = |\vec{P}_T|/\sqrt{\hat{s}}$, we observe that the asymptotic behavior of these upper bounds can be re-expressed as $|\vec{P}_T|$ and $(|\vec{P}_T|\sqrt{\hat{s}})^{1/2}$, respectively. Hence, we see that in order for these types of background events to populate the M_R distribution in the neighborhood of some value of M_Δ, the magnitude of the vectorial sum of the transverse momentum of any missing particles needs to be on the order of M_Δ, in addition to the invariant mass of the visible particles independently having the same scale.

In the case of the jets plus missing transverse momentum final state, this example is not only relevant for the QCD multijet background, but also for the so-called irreducible background $Z(\nu\nu)$+dijets. Here, $|\vec{P}_T| \sim p_T^Z$, and hence has an intrinsic scale on the order of M_Z. The distribution of M_R still falls off exponentially for this background when $M_Z \lesssim M_R$.

The variables M_R and R^2, used in conjunction, are powerful variables for distinguishing new physics events from the relevant backgrounds, exploiting both the scale

of new physics and transverse momentum balance of events to suppress backgrounds with spurious missing transverse momentum.

9.1.4 Generalizing to Inclusive Final States

To derive the variables M_R and R we have not only made assumptions about the kinematics of new physics events, but also about the underlying interactions within them. Specifically, we have assumed that the pair produced particles have the same mass and that they decay directly to a weakly interacting particle and visible SM particle. In reality, new particles of different types, with different masses, can be produced together and their decays can be complicated, and asymmetric, cascades with multiple steps. In this section we explore how the kinematic variables M_R and R behave in these more complicated production and decay scenarios and develop an approach for generalizing their application to events with more than two measured particles in the final state.

We first return to the example described shown in figure 9.1, except we now allow for the two massive particles, S_1 and S_2, to have different masses. Alternatively, we observe that allowing the masses of the weakly interacting particles resulting from the decays of S_1 and S_2 to be different will have a similar effect; the value of M_Δ for each decay chain is what dictates the kinematics of the event in the CM frame. Using the notation of section 9.1.1, we will assume that each of the two decay chains has a different value for $M_\Delta^i = \frac{M_{S_i}^2 - M_{\chi_i}^2}{M_{S_i}}$, such that $M_\Delta^2 = M_\Delta^1(1 + \delta) = M_\Delta(1 + \delta)$.

Assuming $\gamma_{CM} = 1$, we numerically integrate over flat matrix elements for the S_i decay angles to derive the distribution for M_R, for different values of δ, which is shown in figure 9.5. We find that M_R peaks precisely at the geometric mean of M_Δ^1 and M_Δ^2. Hence M_R, in some sense, is sensitive to the average characteristic scale of the two different decay chains.

If the particles S_i do not undergo direct two body decays, either by going through a multibody decay or through an additional decay step with an intermediate massive particle, then each of the two decay chains can produce more than one visible SM
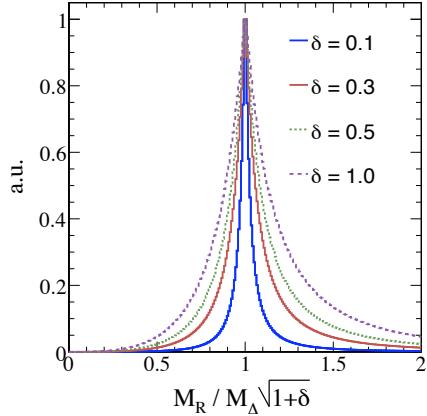
Figure 9.5: Distribution of M_R, in units of $M_\Delta\sqrt{1+\delta}$, for different values of δ. Distributions are normalized such that the maximum value is equal to 1.

particle in the final state. To generalize the kinematic variables to these cases, events with multiple particles in the final state are cast into a two object topology through the formation of two *mega-jets*, or two objects made from grouping all of the observed particles into two unique partitions. The two mega-jet four-momenta are defined as the sum of four-vectors for all the assigned objects, requiring that at least one object is assigned to each mega-jet. The combination of assignments that is chosen is the one which minimizes the invariant masses of the two mega-jets summed in quadrature. This choice of assignment combines particles together that are traveling in the same direction, attempting to group the common decay products of each particle S_i together. M_R and R are defined as before, except using the momenta of the two mega-jets to construct the variables.

As an example, we consider a more complicated decay topology, shown in figure 9.6. In this scenario, the particles S_1 and S_2 have the same mass (M_S), as do χ_1 and χ_2 (M_χ), except now one or both of the particles S_i undergoes a two-body decay to a visible particle, Q_{i+2}, and another particle, G_i, with mass $M_G = M_S(1-\delta)$. The particle G_i then decays to another visible particle, Q_i, and χ_i. Numerically integrating over all the decay angles in this scenario (using flat matrix elements) with $\gamma_{CM} = 1$, and requiring $R > 0.4$, we derive the distributions for M_R, shown in figure 9.7, where
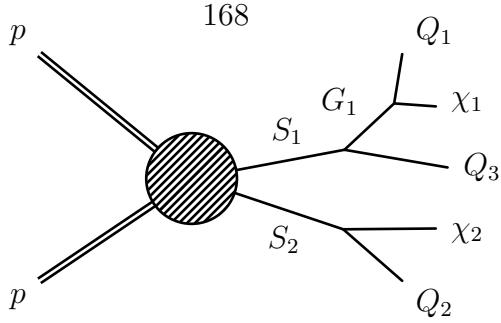
Figure 9.6: Two massive particles, S_i, are produced in a pp collision. S_i can decay directly to a visible SM particles, Q_i, and a weakly interacting particle, χ_i, as illustrated for the decay of S_2. Alternatively, S_i can decay directly to an intermediate heavy particle G_i and a visible SM particles, Q_{i+2}, with G_i subsequently decaying to a visible particle Q_i and a weakly interacting particle, χ_i, as illustrated for the decay of S_1.

we have assumed either one or both of the particles S_i decays through an intermediate G_i. We find that, in both of these cases, the resulting M_R distribution peaks at $M_\Delta = \frac{M_S^2 - M_\chi^2}{M_S}$, regardless of the value of δ (for the values considered here) and irrespective of whether all of the visible decay products resulting from a particular S_i are assigned to the same mega-jet. Even in more complicated cases, the variable M_R is able to resolve the characteristic scale of the parent particle and weakly interacting particle mass splitting.
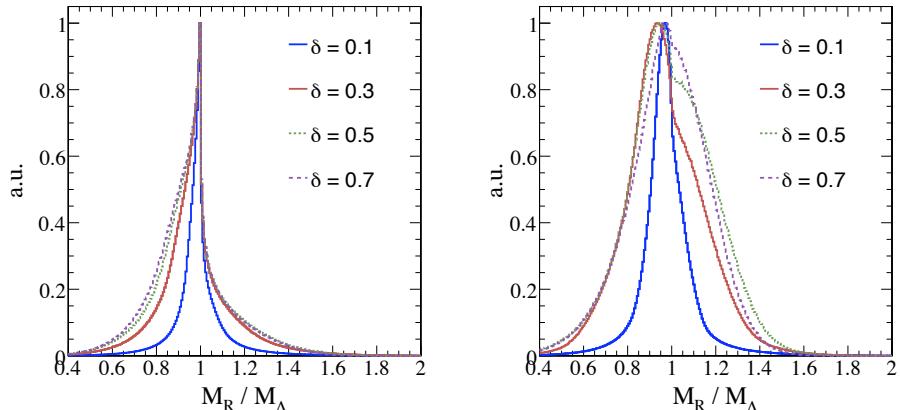
Figure 9.7: Distribution of M_R when one (left) or both (right) of the particles S_i decays to an intermediate particle G_i with mass $M_G = M_S(1 - \delta)$, for different values of δ. Distributions are normalized such that the maximum value is equal to 1

The variables M_R and R are well suited for searching for a general class of signals,

including variations that violate the symmetries used to derive them. M_R peaks at the characteristic scale of events, reflecting the mass splitting between the new massive particles initially produced and the weakly interacting particles they decay to. R can be used in conjunction with M_R, suppressing backgrounds at this characteristic scale using the transverse shape of events, even in the presence of detector mis-measurements and other experimentally difficult effects.

9.2 Razor Phenomenology and M$_R$ Scaling

In the canonical approach to searching for SUSY experimenters look for an excess of events with large missing transverse energy indicating the presence of escaping weakly interacting particles. The challenge is that the tails of the MET distribution are both difficult to model and difficult to clean a priori from spurious instrumental effects. In Section 9.1 we introduce a new approach to searching for SUSY phenomena using the razor kinematic variables, M_R and R, designed to compensate for these difficulties. Like MET, the variable M_R is sensitive to the characteristic mass scale of new physics events. When used in conjunction with R, background events events with M_R values which are not reflective of their true scale can be suppressed. In order to understand the phenomenology of the razor variables in signal and SM background events we consider a collection of data and simulated events corresponding to the search to the 2010 CMS search for SUSY described in section 9.3. The details of how these events are selected in data or simulated, how physics objects are reconstructed and the selection applied are explained in section 9.4.

We first consider simulated SM background events with jets and MET final states. One of the most The M_R distribution for some of the largest backgrounds to SUSY searches are shown in figure 9.8. For each of these backgrounds, the M_R distribution peaks at its respective scale. QCD multijet events, with an intrinsic scale of order Λ_{QCD}, peak at the minimum scale set by the jet p_T requirements applied to the objects used to calculate M_R. $t\bar{t}$+jets and $Z(\nu\nu)$+jets events, which can both have MET resulting from neutrinos in the event (from W decays for $t\bar{t}$) peak at scales set by

the top and Z masses, respectively. Without a requirement on R, the QCD multijet yield is larger than the other backgrounds (and potential signals) by several orders of magnitude, even at large values of M_R. As described qualitatively in section 9.1.3, an increasingly tight requirement on R suppresses backgrounds with M_R values larger than the *true* scale of the process. This effect is clear in figure 9.8; with tighter cuts on R the tail of the M_R distribution for these backgrounds is exponentially reduced, falling increasingly steeper. With a moderate cut on R the QCD multijet background is now the smallest background with M_R above a few hundred GeV by at least several orders of magnitude. The cuts on R reduce the tails of all the backgrounds in M_R, but significantly more so for backgrounds without neutrinos resulting from heavy boson decays. Backgrounds with high transverse momentum weakly interacting particles, like the signals we are searching for, can have large momentum imbalances in the transverse events plane, yielding naturally large values of R in events with M_R near the true scale of the process.
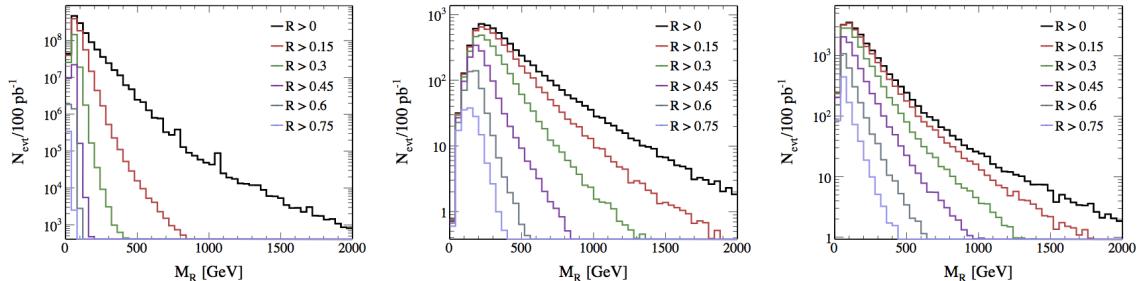
Figure 9.8: M_R distribution for simulated event samples, for different cuts on the razor, R. (Left) QCD multijet events. (Center) $t\bar{t}$+jets events. (Right) $Z(\nu\nu)$+jets events.

The signal events we are searching for are expected to exhibit the same behavior in the variables M_R and R as the SM backgrounds, but at a new, higher scale. Here, we consider two benchmark SUSY scenarios, LM1 and LM5, which are parameter points in the mSUGRA class of models and defined in section B.1. The M_R distribution, as a function of R requirement, for LM1 and LM5 events are shown in figure 9.9 and 9.10, respectively. Also included in those figures are the sparticle mass spectra for the two models. The signal models' events are composed primarily of strong

production of pairs of squarks or gluinos, which subsequently decay directly or through cascades down the sparticle mass spectrum to the LSP, which is weakly interacting and escapes detection. The correspondence between the behavior of M_R and these spectra is clear, with pronounced peaks corresponding to the squark/gluino and LSP mass splittings. With increasing cuts on R the mass peak remains largely unchanged while, like the backgrounds considered in figure 9.8, the right-hand tail of the distribution falls increasingly quickly.
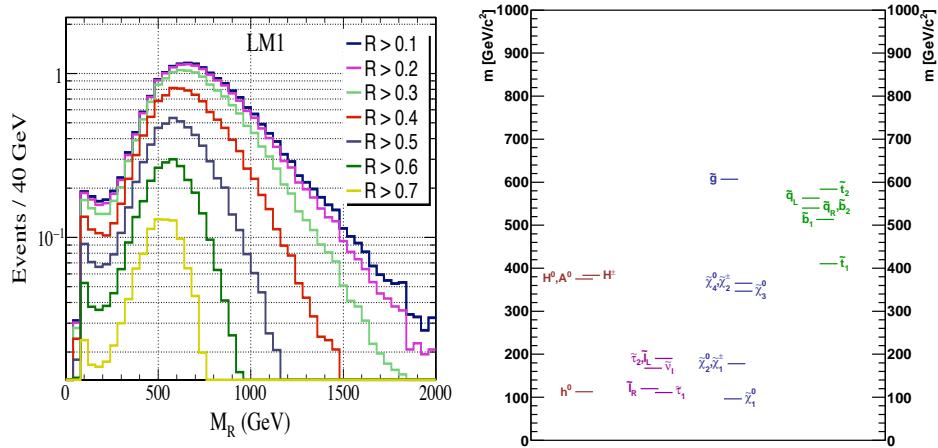
Figure 9.9: (Left) the M_R distribution after different R selections for the LM1 mSUGRA benchmark model. (Right) the superpartner spectra for the corresponding model. The prominent M_R peak corresponding to strong pair-production of squarks and gluinos with masses $\sim$600 GeV decaying to 100 GeV LSPs. The peak position indicates the characteristic scale of this mass splitting.

A search for evidence of new physics using the razor variables can be summarized most succinctly by considering the two-dimensional M_R vs. R razor plane and where the events from each process are expected to appear, shown in figure 9.11. SUSY events appear at larger values of M_R due to the large mass splittings in the particles produced while the SM backgrounds peak at their own, lower, mass scales. The pair of weakly interacting particles appearing in the R-parity conserving SUSY events leads to a transverse momentum imbalance and larger values of R, relative to backgrounds.

Identifying a region of phase space enriched in potential signal events and a small amount of background is straightforward in terms of the razor variables. The following
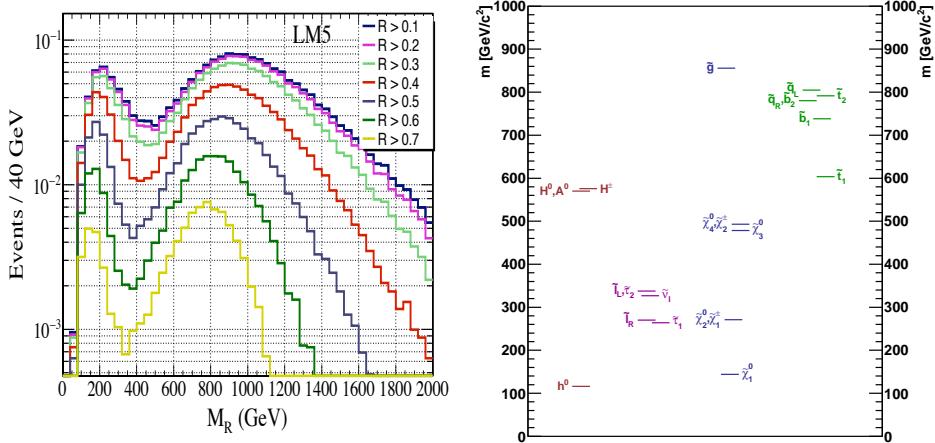
Figure 9.10: (Left) the M_R distribution after different R selections for the LM5 mSUGRA benchmark model. (Right) the superpartner spectra for the corresponding model. The prominent M_R peak corresponding to strong pair production of squarks and gluinos with masses ~ 850 GeV decaying to 150 GeV LSPs. The peak position indicates the characteristic scale of this mass splitting. A smaller M_R peak appears at a lower value from events with pair production of charginos and second mass eigenstate neutralinos.

section describes an approach to predicting SM backgrounds at large M_R/R in the razor plane using control measurements in the low M_R/R region.

9.2.1 M_R Scaling

The M_R distribution for simulated QCD multijet events, shown in figure 9.8 and 9.11 demonstrates a dramatic dependence on the value of the R requirement, with the slope of the M_R tail becoming steeper with increasing R cut. To confirm this predicted behavior in data, we select a QCD multijet control sample using prescaled jet triggers requiring at least two jets with an average uncorrected $p_T > 15$ GeV. Because of the low jet threshold, the QCD multijet background dominates this sample for low M_R, allowing us to study the M_R shape dependence on R thresholds.

The M_R distributions for events satisfying the QCD control selection, for different values of the R threshold, are shown in figure 9.24 (left). We find that the M_R distribution is exponentially falling, after a turn-on at low M_R resulting from the p_T threshold requirement on the jets entering the M_R calculation. This turn-on can be
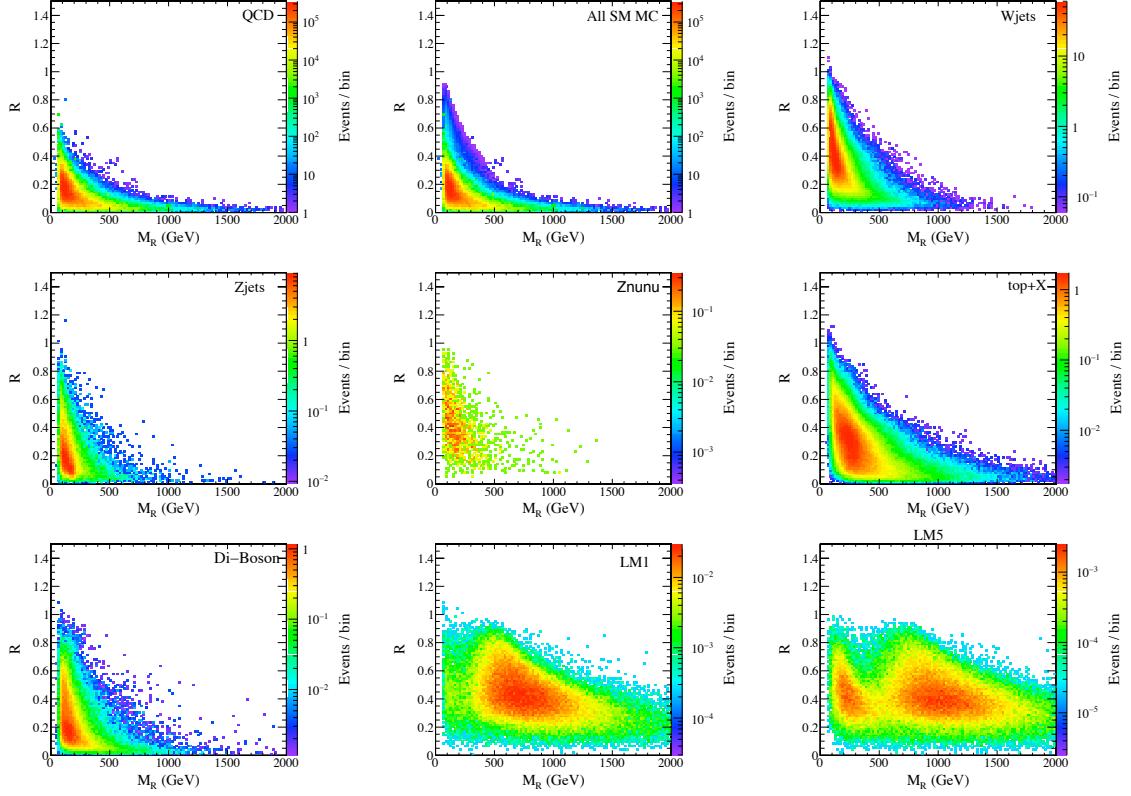
Figure 9.11: The razor plane: M_R versus R simulated event yields for L $= 10$ pb^{-1}. Shown are Monte Carlo simulated samples: QCD multijets, all simulated backgrounds combined, W+jets, Z+jets, $Z \to \nu\bar{\nu}$+jets, t+X, di-boson, CMS SUSY benchmark models LM1 and LM5, respectively and as labeled. Yields correspond to the baseline HAD box selection described in section 9.4.

modeled as an asymmetric Gaussian, while the tail is well described by an exponential function. We perform a likelihood based fit on the M_R distribution for different R thresholds, modeling it with the function

$$
f(x) = \begin{cases} f_1(x) = N_{\text{gaus}} e^{-\frac{(x-\mu)^2}{2\sigma_1^2}} & x \leq \mu \ , \\ f_2(x) = N_{\text{gaus}} e^{-\frac{(x-\mu)^2}{2\sigma_2^2}} & x \in (\mu, t) \ , \\ f_3(x) = N_{\text{exp}} e^{Sx} & x \geq t \ , \end{cases} \tag{9.19}
$$

where the parameters N_{gaus} and t follow from the others, which are floated in the fits,

and the continuity conditions

$$f_2(t) = f_3(t) , \quad f_2'(t) = f_3'(t). \tag{9.20}$$
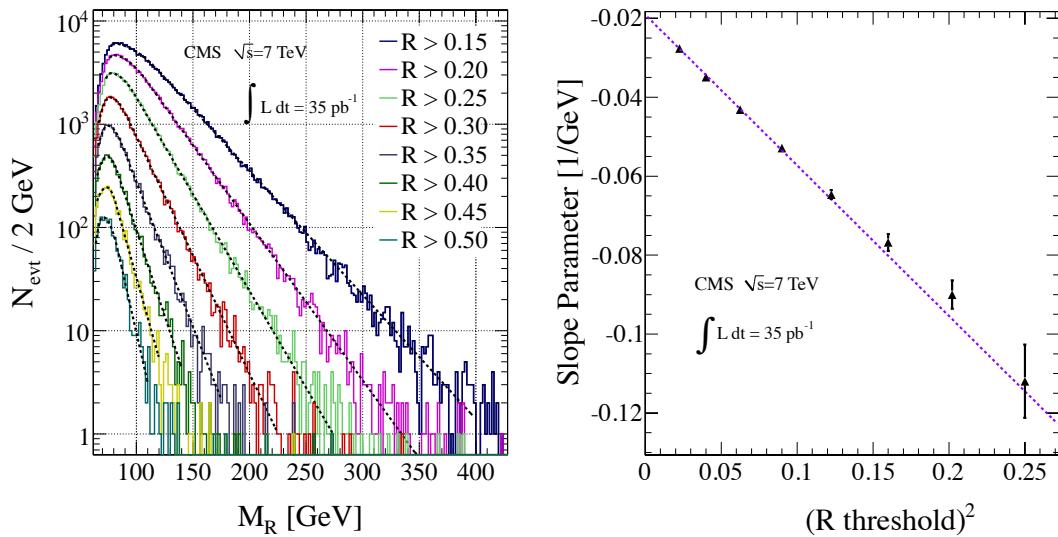
Figure 9.12: (Left) M_R distributions for different values of the R threshold for data events in the QCD control box. Fits of the M_R distribution to an exponential function and an asymmetric Gaussian at low M_R, are shown as dotted black curves . (Right) The exponential slope S from fits to the M_R distribution, as a function of the square of the R threshold for data events in the QCD control box.

The dependence of the exponential slope parameter, S, on the R threshold is shown in figure 9.24 (right). We observe that the slope parameter, indicating how quickly the tail of the M_R distribution falls, has a linear dependence with the square of the R threshold requirement, such that

$$S = a + b \cdot (R \text{ threshold})^2 . \tag{9.21}$$

The R threshold shapes the M_R distribution in a simple, and therefore predictable, way. Adherence to this M_R scaling implies that with knowledge of the parameters a and b from equation (9.21) we can predict the the shape of the M_R distribution at

large R thresholds.

The qualitative similarities between the shapes of the different SM background in the razor plane, illustrated in figure 9.11 is indicative of the fact that, at values of M_R past their intrinsic mass scale and at sufficiently high R threshold, each of the different SM backgrounds exhibit this same M_R scaling, albeit with different values of the parameters a and b. Apart from QCD multijet backgrounds, which we see populate the low M_R/R region of the razor plane, the other backgrounds are processes with energetic neutrinos and muons from massive vector boson decays (including Ws from top decays).

9.2.2   The Razor Box Concept

In some sense, the canonical R-parity conserving SUSY signature are events with jets and MET. Strong production of new sparticles, if kinematically accessible, will dominant over electroweak production at a hadron collider and the produced squarks and gluinos will decay into jets and escaping LSPs. A search for these events could proceed by vetoing the presence of reconstructed leptons but a peculiarity of the SM is that a large portion of the background events would still have leptons in them. The reason is that, at high R and M_R, most of the background made up of events with $W(\ell\nu)$ decays present, with high p_T neutrinos resulting in large values of R. The remaining background events also involve heavy vector bosons, through $Z(\nu\nu)$+jets production. Conversely, squark and gluon decays could also result in lepton final states, perhaps also through intermediate W and Z decays.

By classifying events according to their lepton content, in different *boxes*, we can isolate background processes like $W(\ell\nu)$+jets and use them to make inferences about other kinematically similar background processes, like $Z(\nu\nu)$+jets events. Additionally, we can search for signal events in these different final state boxes, benefitting from this classification for signals with an abundance of leptons. The largest background appearing in each of the final states considered in the razor SUSY search is events with on-shell W bosons which decay leptonically $W \rightarrow \ell\nu$. We recall that the

razor variables M_R and R are calculated using the momentum of two event mega-jets which are constructed from all of the objects recorded in the event. This means that, in order to result in at least two final state objects that can seed these mega-jets, $W(\ell\nu)$ decays must be accompanied by other final state objects, such as associated production jets or other top decay products in top+X events. In the former, the yields of $W+n$ jet events falls exponentially with increasing n (see Chapter 6 for more information about Berends-Giele scaling), with a factor of 5-10 less events for each additional reconstructed jet (depending on p_T threshold). Whether the lepton from the W decay is included among the objects making mega-jets dictates whether this background is composed of mostly $W(\ell\nu)+1$ jet or $W(\ell\nu)+2$ jet events. The value of R calculated in these events will depend largely on the kinematics of the neutrinos, and hence on the W themselves, as will M_R if decay leptons are included in its calculation.

In the SUSY search described in section 9.3 we consider three different final states, or boxes, based on the presence or absence of electrons and muons: the ELE, MU and HAD boxes, respectively. The SM backgrounds appearing in each of these boxes can generally be thought of as corresponding to the two different types of W decay events described above where, in the HAD box in particular, events of both qualitative types can appear. In order to understand how these two different background contributions behave in our M_R scaling model we consider $W(\ell\nu)$ events with two different kinematic views of the decay leptons:

- The lepton (muon or electron) participates with its energy and momentum in the mega-jet and M_R reconstruction. This is the default approach used in the SUSY search and corresponds to the MU and ELE boxes for single muon and single electron events, respectively.

- The lepton is treated as "invisible" both in the mega-jet, M_R and R reconstruction. This treatment corresponds to the MU* and ELE* boxes for single muon and single electron events, respectively. In this case, all of the W decay products are effectively weakly interacting, meaning these events will behave

kinematically like $Z(\nu\nu)$+jets events. This correspondence is exploited in the HAD box background prediction.

In the following, we describe the M_R scaling behavior for each of the SM backgrounds appearing in the ELE, MU, ELE* and MU* boxes, and how they are related to similar backgrounds in the HAD box, using simulated events. These scaling relations and correspondences between final states form the basis for the strategy of the CMS razor SUSY search.

9.2.3 MU* and ELE* Boxes: Leptons+Jets Processes with the Lepton Treated as Invisible

Some of the largest backgrounds in the HAD box are processes with a heavy vector boson's transverse momentum escaping the detector, such as $Z(\nu\nu)$+jets and $W(\mu\nu)$+jets. To derive control samples and study the behavior of kinematically similar processes with respect to the variables R and M_R, we implement a lepton box selection in which we treat the identified leptons as "invisible," kinematically mimicking the presence of an additional neutrino; we denote these selections the MU* and ELE* Boxes for muons and electrons, respectively.

By treating leptons as invisible in $W(\ell\nu)$+jets events we can create a source of kinematically similar events to $Z(\nu\nu)$+jets, one of the largest backgrounds in the hadronic final state. Technically, this is accomplished in reconstruction through two different approaches for electrons and muons. In the case a muon is identified satisfying the muon box criteria, it is assumed that the muon did not leave a significant calorimetric deposit, such that the reconstructed jet collection is not affected by its presence. In the calculation of the MET the momentum from the muon is ignored, effectively treating it as an escaping neutrino. For electrons, the treatment of the MET reconstruction is the same, ignoring the electron momentum in its calculation. If a reconstructed jet is found to match the direction of the electron three-momentum within 0.3 in $\sqrt{\Delta\eta^2 + \Delta\phi^2}$ it is removed from the collection of objects used in the cal-

culation of M_R and the mega-jets, effectively removing the electron energy footprint from among the reconstructed particles.

Figure 9.13: (Left) M_R distributions for different values of the cut on R for simulated $W(\mu\nu)$+jets events satisfying the MU* Box selection. Fits to the exponential part of the M_R distribution are shown as dotted black lines. (Right) Value of the exponential slope S from fits to the M_R distribution, as a function of R cut for simulated $W(\mu\nu)$+jets events satisfying the MU* Box selection.

The largest contributor to yields in the single lepton MU* and ELE* boxes is $W(\ell\nu)$+jets. Since the lepton does not participate in the construction of the hemispheres, at least two additional high-p_T objects must be present in the event, resulting in the selection of events with W bosons produced in association with two or more jets and MET from the recoiling W. The M_R distribution for simulated $W(\mu\nu)$+jets events, for different R cuts, is shown in figure 9.13 (left).

For M_R values above a characteristic scale (determined by the minimum jet p_T requirement and the minimum Q^2 for the partonic subprocess), the M_R distribution falls exponentially. The value of the exponential slope S that describes this exponential behavior exhibits M_R scaling, as was the case for QCD multijet events, illustrated in figure 9.13 (right).

The analogous figures for $W(e\nu)$+jets events are shown in figure 9.14. We observe both qualitatively and quantitatively the same behavior in the M_R distribution as a

function of R. With the lepton ignored in the calculation of the kinematic variables R and M_R, the $W(\mu\nu)$ and $W(e\nu)$ processes are kinematically identical as expected, and exhibit the same M_R scaling behavior.
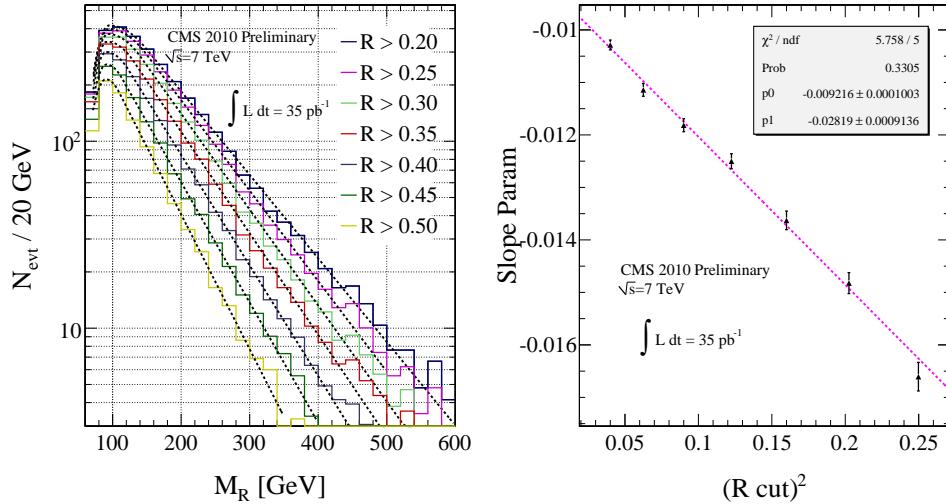
Figure 9.14: (Left) M_R distributions for different values of the cut on R for simulated $W(e\nu)$+jets events satisfying the ELE* Box selection. Fits to the exponential part of the M_R distribution are shown as dotted black lines. (Right) Value of the exponential slope S from fits to the M_R distribution, as a function of R cut for simulated $W(e\nu)$+jets events satisfying the ELE* Box selection.

The signature of a heavy vector boson decaying into two neutrino-like objects also reproduces the kinematic phase-space of $Z(\nu\nu)$+jets events contributing to the background of the HAD box. The M_R distribution for $Z(\nu\nu)$+jets events selected in the HAD box is shown in figure 9.15 (Left). As expected, these events behave in a nearly identical way to the $W(\ell\nu)$+jets events selected in the MU* and ELE* boxes. These similarities allow us to use the MU* and ELE* boxes to infer the shape of the M_R distribution and relative efficiency with respect to the R cuts for $Z(\nu\nu)$+jets events.

Apart from $W(\ell\nu)$+jets the other large background in the MU* and ELE* boxes is $t\bar{t}$+jets production with at least one W boson decaying leptonically. Kinematically, these events events are very similar to the previously discussed processes, in that they include a leptonically decaying W recoiling against jets. The M_R distributions for $t\bar{t}$
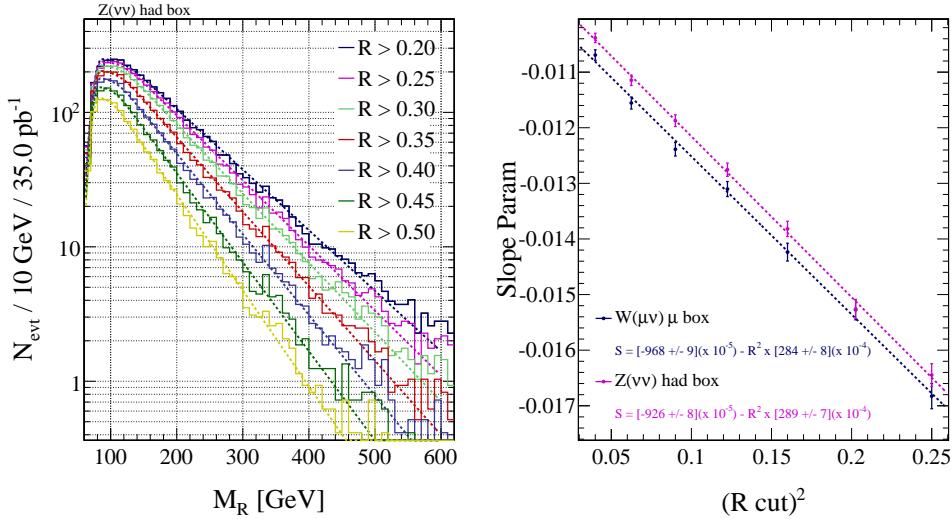
Figure 9.15: (Left) M_R distributions for different values of the cut on R for simulated $Z(\nu\nu)$+jets events satisfying the HAD Box selection. Fits to the exponential part of the M_R distribution are shown as dotted colored lines. (Right) Value of the exponential slope S from fits to the M_R distribution, as a function of R cut for simulated $W(\mu\nu)$+jets events satisfying the MU* Box selection and for $Z(\nu\nu)$+jets events that satisfy the HAD Box selection.

events with a lepton satisfying the MU* and ELE* Box requirements are shown in figure 9.16 and 9.17, respectively. As was the case for $W(\ell\nu)$+jets events, the $t\bar{t}$ M_R distribution falls exponentially once M_R exceeds the relevant scale for the process, with the slope of the exponential fall scaling with R^2. The primary difference between W and $t\bar{t}$ events is the difference in the process' scale. For example, if we consider $t\bar{t}$ events where both W's decay to leptons which are not observed in the detector then kinematically these events are identical to the SUSY-motivated events which we are searching for. In this case, the Ws play a role analogous to the escaping neutralinos in the canonical SUSY di-jet final state. Similarly, the two reconstructed b-jets are not constrained to be back to back in the transverse plane (as is the case for jets coming from squark decays for example). Hence, before falling off exponentially these $t\bar{t}$ events peak at a scale set by the top and W masses, analogous to the peaking behavior of signal events.

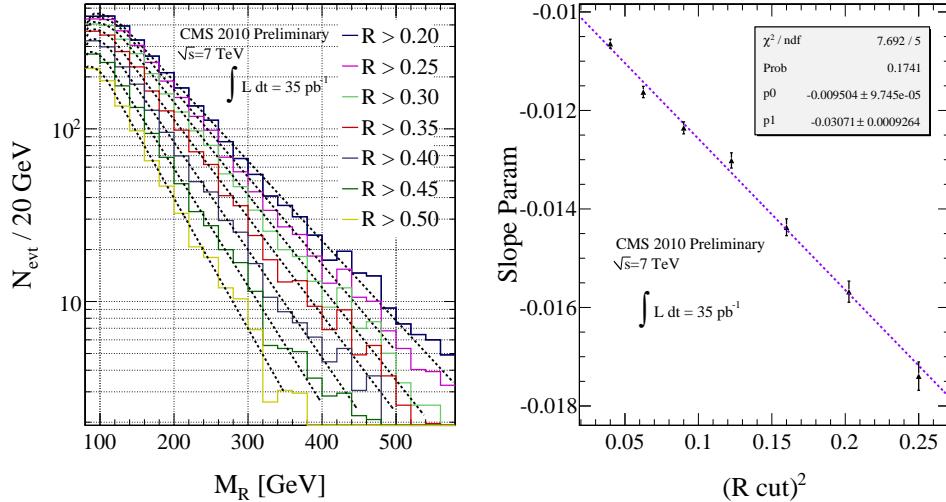Despite the differences in the intrinsic scale between $W(\ell\nu)$+jets and $t\bar{t}$+jets pro-
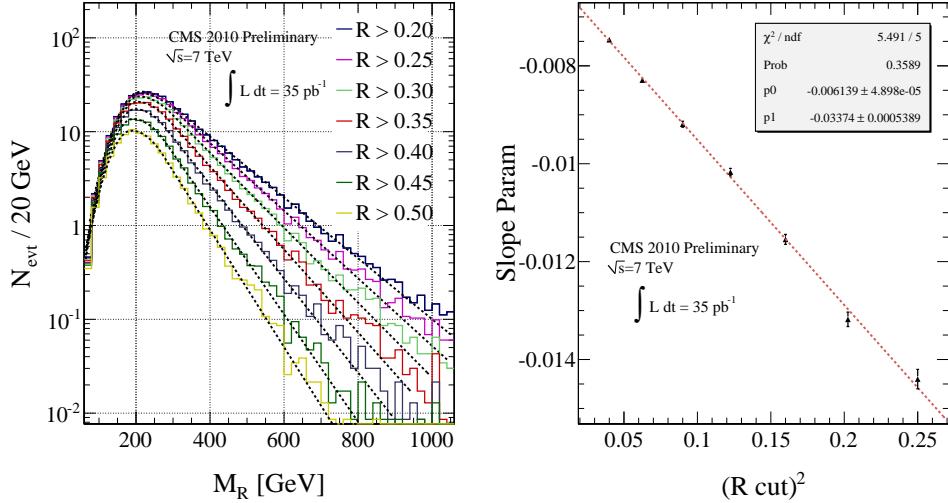
Figure 9.16: (Left) M_R distributions for different values of the cut on R for simulated $t\bar{t}$)+jets events satisfying the MU* Box selection. Fits to the exponential part of the M_R distribution are shown as dotted black lines. (Right) Value of the exponential slope S from fits to the M_R distribution, as a function of R cut for simulated $t\bar{t}$+jets events satisfying the MU* Box selection.
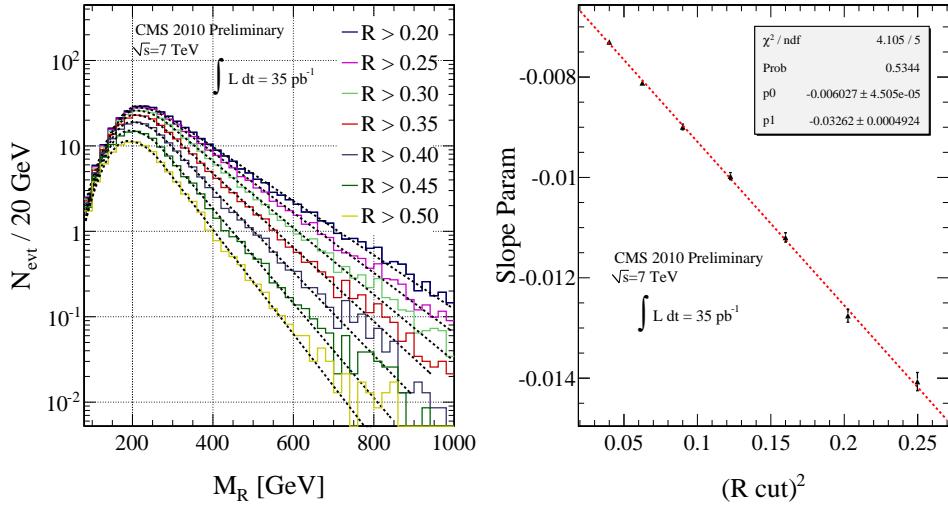
Figure 9.17: ELE* Box selection and treatment of leptons. (Left) M_R distributions for different values of the cut on R for simulated $t\bar{t}$+jets events satisfying the HAD Box selection. Fits to the exponential part of the M_R distribution are shown as dotted colored lines. (Right) Value of the exponential slope S from fits to the M_R distribution, as a function of R cut for simulated $t\bar{t}$+jets events satisfying the MU* Box selection.

cesses, we observe that their parameters b describing the R^2 scaling behavior are nearly identical. Comparing the M_R exponential slope parameters for $t\bar{t}$+jets events with those for $W(\ell\nu)$+jets in the MU* and ELE* Boxes, for example, we see that the largest difference in the R^2 scaling behavior of these slopes is in the parameter a. This is the result of the W's from top decays having some intrinsic transverse momenta on the order of $(M_{top}^2 - M_W^2)/2M_{top}$ compared to the "intrinsic" transverse momenta of the W in $W(\mu\nu)$ events being smaller and resulting largely from the minimum jet p_T requirements in the mega-jet calculation. This difference in intrinsic W p_T shifts the M_T^R distributions and R distributions relative to each other, acting as an effective "R-offset" which results in the $t\bar{t}$ slope being systematically steeper for $W(\mu\nu)$+jets events, as indicated by the systematically larger value of a. Despite this difference, the two processes have parameters b that are nearly identical.

The kinematic similarities between the previously discussed processes indicate that any differences between simulated events and data will affect each of these processes the same way in the context of the variables R and M_R. With insufficient integrated luminosity to resolve the R^2 scaling behavior of $t\bar{t}$ and Z events selected in di-lepton boxes we instead infer the shapes of these backgrounds by directly measuring the R^2 scaling parameters for $W(\ell\nu)$+jets events in data and deriving relative data/simulation slope correction factors to apply to the shapes of the other kinematically similar backgrounds.

The values of the R^2 scaling slope parameters measured in simulated events for processes contributing to the MU* and ELE* boxes are summarized in Table 9.1.

Process / Box	a	b
$W(\mu\nu)$+jets / μ^* Box	$(-95 \pm 1) \times 10^{-4}$	$(-307 \pm 9) \times 10^{-4}$
$W(e\nu)$+jets / e^* Box	$(-92 \pm 1) \times 10^{-4}$	$(-282 \pm 9) \times 10^{-4}$
$t\bar{t}$ / μ^* Box	$(-614 \pm 5) \times 10^{-5}$	$(-337 \pm 5) \times 10^{-4}$
$t\bar{t}$ / e^* Box	$(-603 \pm 5) \times 10^{-5}$	$(-326 \pm 5) \times 10^{-4}$
$Z(\nu\nu)$+jets / HAD Box	$(-926 \pm 8) \times 10^{-5}$	$(-289 \pm 7) \times 10^{-4}$

Table 9.1: Values of parameters describing the M_R evolution with R^2 for simulated events of different processes selected with the MU* Box, ELE* Box and HAD Box requirements. For the lepton boxes, leptons are treated as neutrinos, better reproducing the HAD box kinematic configuration of $Z(\nu\nu)$+jets.

9.2.4 MU and ELE Boxes: Leptons+Jets Processes with the Lepton Treated as Visible

As mentioned earlier, one of the largest backgrounds in the HAD box is $W(\ell\nu)$+jets events. These events, with respect to R and M_R, are of two qualitatively different types. One component is the one discussed in the previous section; namely, $W(\ell\nu)$ produced in association with two or more jets, with the charged lepton from the W decay contributing minimally or not at all to the calculation of R and M_R. Here, the missing transverse momentum is essentially the W transverse momentum.

The other $W(\ell\nu)$+jets background component to the HAD box results from the lepton giving a nontrivial contribution to the calculation of the mega-jets and R/M_R. Since the collection of objects used to calculate R/M_R is based on calorimeter deposi- tions, this occurs when the W decays to an electron, or a τ which subsequently decays either hadronically or to an electron. In each of these cases, part of the W decay is reconstructed as a jet and contributes to the mega-jets going into the calculation of M_R.

In order to isolate a control sample of events with phase-space kinematically sim- ilar to this "visible-lepton" background in the HAD box we implement a lepton box selection in which we treat the identified leptons as "visible," kinematically mimick- ing this background; we denote these selections the MU and ELE Boxes for muons and electrons, respectively. In this treatment, the identified lepton momentum con- tributes to the calculation of M_R or R and participates in the construction of the mega-jets. This implies two technically different procedures for reconstructed muons and electrons.

In the case a muon is identified satisfying the muon box criteria, it is assumed that the muon did not leave a significant calorimetric deposit, such that the reconstructed jet collection is not affected by its presence. Given this consideration, we include the muon in the list of objects used in the construction of the mega-jets without concern of redundancy in the jet collection. For electrons, if a reconstructed jet is found to match the direction of the electron three-momentum within 0.3 in $\sqrt{\Delta\eta^2 + \Delta\phi^2}$

it is removed from the collection of objects used in the calculation of M_R and the mega-jets and replaced by the reconstructed electron object. This is done in order to prevent mis-calibration of electrons when they are included in the jets collection and calibrated as such.

The backgrounds in the single lepton ELE and MU Boxes are essentially the same processes that appeared in the ELE* and MU* Boxes, with $W(\ell\nu)$+jets being the largest. Since the lepton is included in the mega-jet and M_R reconstruction, this background is predominantly composed of W bosons produced in association with at least one hadronic jet. This results in the $W(\ell\nu)$ yields in the ELE and MU Boxes being larger relative to those in the ELE* and MU* Boxes, respectively for smaller values of M_R. At the same time, this also implies that the observed MET in the event will result predominantly from the escaping neutrino only, rather than both of the W decay products. Since the neutrino is the primary source of MET, and the lepton comes from the same mother particle, we observe two different components to $W(\ell\nu)$+jets MU and ELE Boxes which can be understood as follows:

In one case, the decay lepton is sufficiently hard such that one of the mega-jets contains a large contribution from this lepton. Here, increased lepton momentum will roughly translate into increased M_R. Simultaneously, the neutrino is coming from the same W decay, so a harder lepton, on average, means a softer neutrino and a smaller value of R. This direct anti-correlation between the two particles results in component with a more steeply falling M_R distribution relative to the $W(\ell\nu)$+jets contribution to the MU* and ELE* boxes which is predominantly W+1 jet production. We denote this the 1st *component* of $W(\ell\nu)$+jets in the MU and ELE boxes.

In the second case, there are at least two jets recoiling against the W boson. Here, these two other jets are sufficient for forming two mega-jets, and the lepton needs only have a transverse momentum exceeding the lepton box threshold requirement $(p_T^\ell > 15 \text{ GeV}/c)$ and does not need to participate significantly in the construction of the mega-jets for the event to yield a large value of M_R. This can be achieved even if the lepton is anti-aligned with the direction of the W's transverse boost, assuming that this boost is sufficiently large (which is precisely the phase space of interest in the

high R/M_R plane). Hence, for this *2*nd *component* of $W(\ell\nu)$+jets there is no tug-of-war between the neutrino and lepton resulting from the W decay, since the neutrino can be directly aligned with the boosted W and the resulting MET will be roughly proportional to the W p_T, with a small offset correction due to the momentum of the lepton which will detract from the MET.

The M_R distribution for simulated $W(\mu\nu)$+jets and $W(e\nu)$+jets events selected in the MU and ELE boxes are shown in figure 9.18 and 9.19 (top left), respectively. We model the distributions with two different exponential components, such that the exponential part, $f_3(x)$, of the background shape function of equation (9.19) is replaced by

$$f_3(x) = N_{\exp} \left[e^{S_1 x} + f e^{S_2 x} \right].$$

<div align="right">(9.22)</div>

For simulated events we fit the M_R distributions using this function, floating each of the free parameters independently. The results of these fits for $W(\mu\nu)$+jets and $W(e\nu)$+jets events are shown in figure 9.18 and 9.19. We observe that the fits identify two distinct contributions to the M_R distribution, and that both contributions feature exponentially falling behavior with slopes that follow the R^2 scaling seen in the MU* and ELE* boxes. Comparing the MU and ELE Box results, we also see that this fit identifies the same two components in the muon and electron cases, in that the exponential slope parameters are inter-consistent.

This same two-component behavior is present in $W(\ell\nu)$+jets processes contributing to the hadronic box when the lepton contributes to the calculation of M_R and R, i.e., when the W decays to an electron which does not satisfy the electron ID criteria but makes a significant contribution to a reconstructed calorimeter jet or similarly if the W decays to a τ which hadronizes and appears as a jet. In figure 9.20 and 9.21 the results from 2-component exponential fits to these hadronic box samples are shown. We observe that these fits identify the same two components in M_R as in the MU and ELE boxes for $W(\ell\nu)$+jets, with inter-consistent values of the R^2 exponential slope scaling parameters.

A clear picture emerges as to the slopes parameters of the two components for
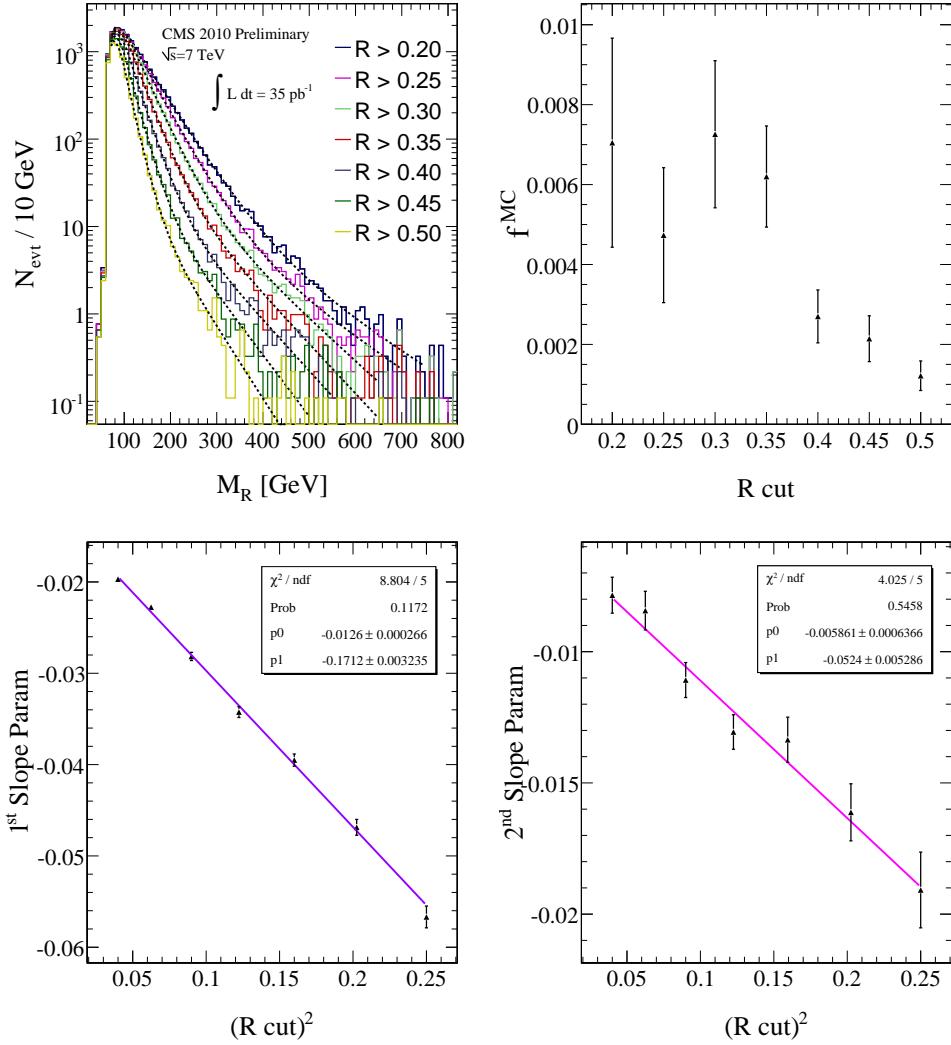
Figure 9.18: $W(\mu\nu)$+jets in the MU Box. (Top left) M_R distribution as a function of R. Two-component fits are shown as dotted black lines. (Top right) The parameter f^{MC} describing the relative amplitude of the second W+jets component to the first, as returned from the fit to simulated events. (Bottom left) Fitted values of the 1st component slope parameter as a function of R^2. (Bottom right) Fitted values of the 1st component slope parameter as a function of R^2.
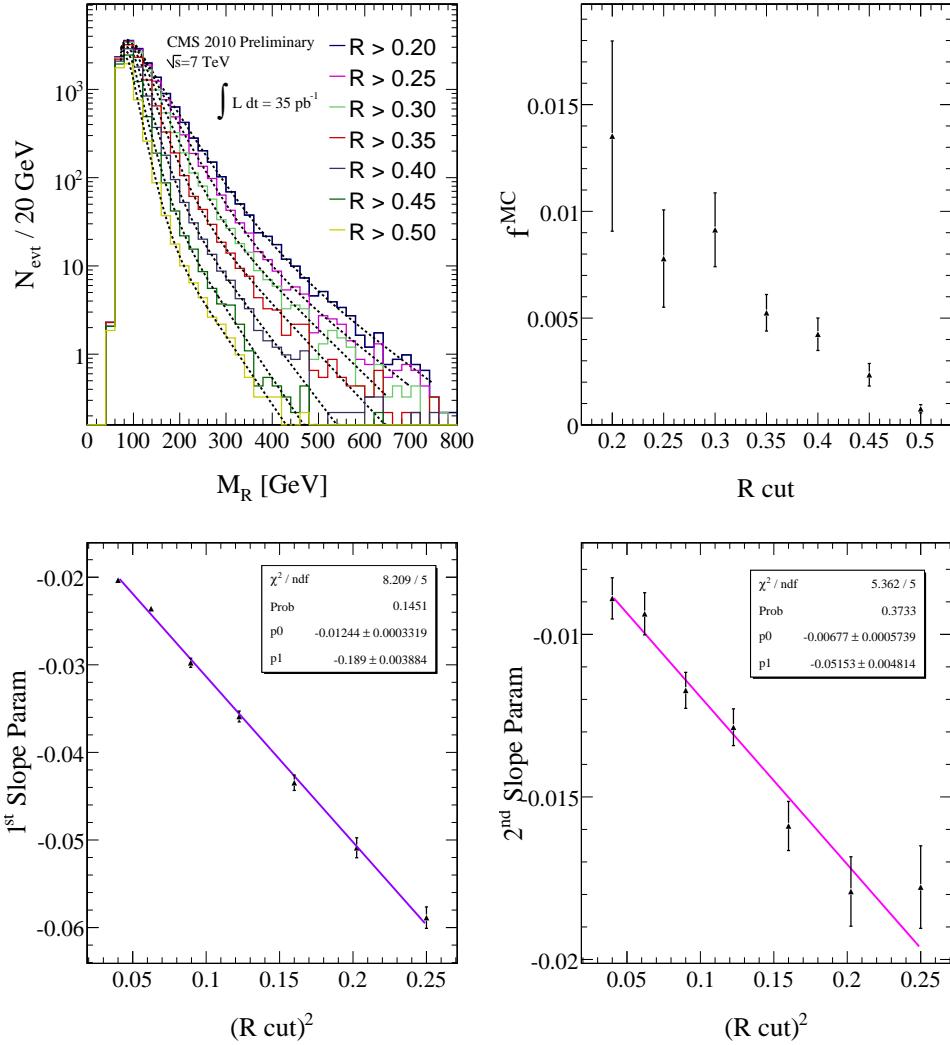
Figure 9.19: $W(e\nu)$+jets in the ELE Box. (Top left) M_R distribution as a function of R. (Top right) The parameter f^{MC} describing the relative amplitude of the second W+jets component to the first, as returned from the fit to simulated. (Bottom) Fits of the two slope parameters as a function of R^2
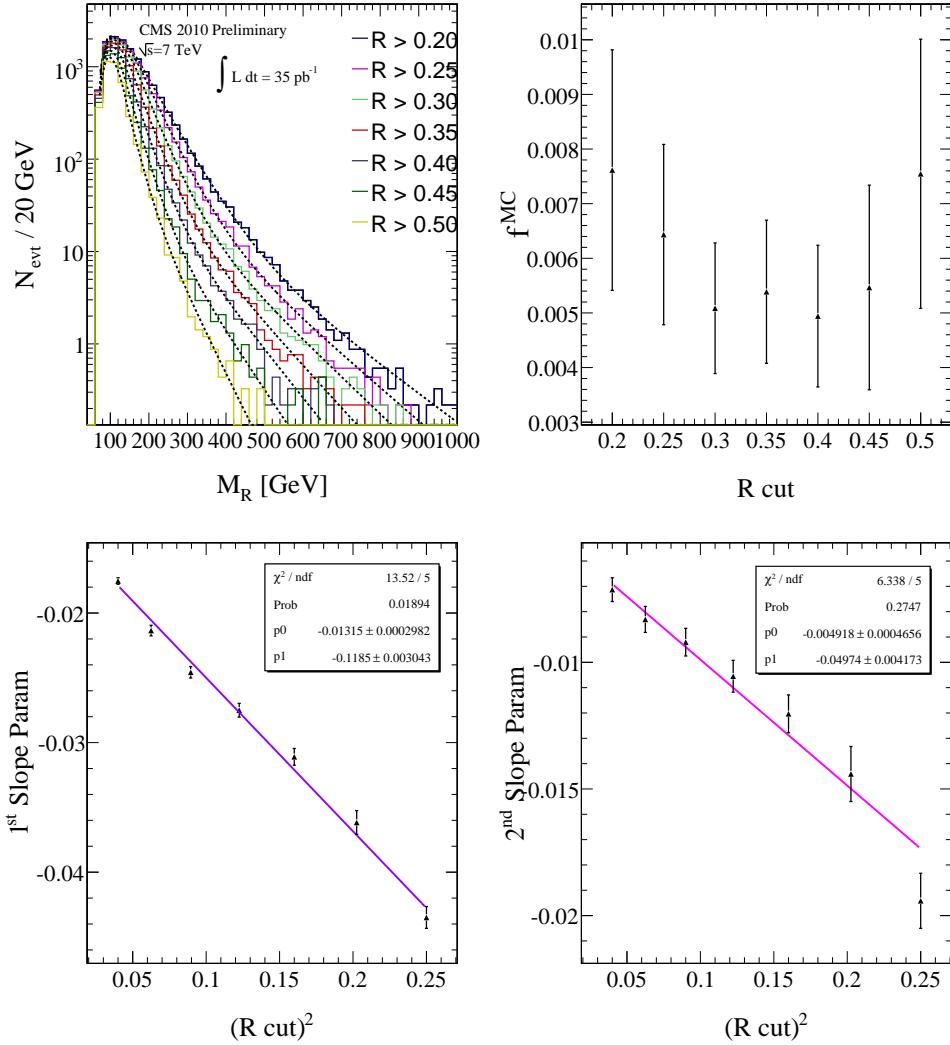
Figure 9.20: $W(e\nu)$+jets in the HAD Box. (Top left) M_R distribution as a function of R. (Top right) The parameter f^{MC} describing the relative amplitude of the second W+jets component to the first, as returned from the fit to simulated events. (Bottom) Fits of the two slope parameters as a function of R^2.
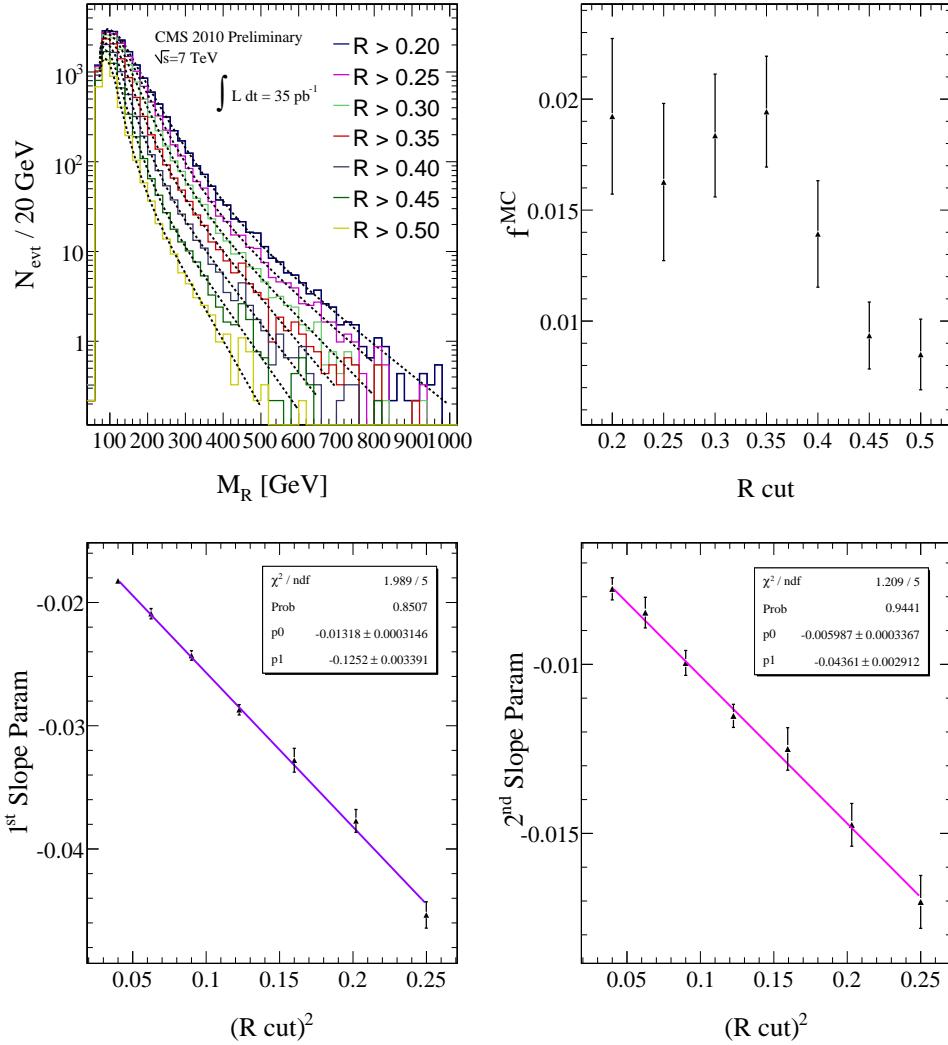
Figure 9.21: $W(\tau\nu)$+jets in the HAD Box. (Top left) M_R distribution as a function of R. (Top right) The parameter f^{MC} describing the relative amplitude of the second W+jets component to the first, as returned from the fit to simulated events. (Bottom) Fits of the two slope parameters as a function of R^2.

these $W(\ell\nu)$ processes. The f factor depends largely on the rate of $W+1$ jet to $W +$ ≥ 2 jet events in the high R/M_R region, and as a result is sensitive to the modeling of the jet energy scale near the jet p_T threshold of 30 GeV. Rather than assuming a value for this parameter from data, we float it in the lower M_R side-band of each of the signal boxes.

In the kinematic region probed by the SUSY search the two-component behavior in M_R does not appear in the other large background processes in the MU, ELE and HAD boxes. For each background process, we perform binned maximum likelihood fits to the M_R distributions, for different values of the R cut, fitting the distributions with a single exponential component and then independently with a double exponential function. We take the ratio of the likelihoods from the two fits and asses the significance of the second exponential component. We find that only the four processes discussed above have significant second exponential components (where we assume that the likelihood ratio of these fits is distributed as a $\chi^2(1)$ distribution). For example, The M_R distributions for simulated $t\bar{t}$+jets and $W(\mu\nu)$+jets events in the MU and HAD Boxes shown in figure 9.22 indicate that these distributions are well described by a single exponential function. The reason for this can be understood by looking at the slope parameters which describe the R^2 scaling of each of these processes. Comparing the values of the b slope parameters in these samples with those of the second component from $W(\ell\nu)$+jets events we observe good agreement. The reason is that these processes *only* have a significant 2nd $W(\ell\nu)$+jets component-like contribution. For example, in the case of $W(\mu\nu)$+jets events selected in the HAD box, the W decay muon cannot contribute significantly to the mega-jet construction since it does not leave a significant calorimetric deposit, and hence cannot result in a calorimetric jet. This implies that the W boson must be produced in association with at least two jets in order to form two mega-jets, resulting in a kinematic configuration equivalent to events constituting the 2nd $W(\ell\nu)$+jets component. Similarly, for semi- and fully-leptonic $t\bar{t}$+jets events, the lepton from the W decay does not need to contribute to the mega-jet construction since there are at least two other b-jets, decay products of the second W and any additional initial or final state radiation which

can form two mega-jets. Hence, these $t\bar{t}$ events appear kinematically, with respect to the variables R and M_R, as a W boson, whose decay products escape the detector, recoiling off of at least two jets.

The exponential slope parameters of the background processes to the MU, ELE and HAD Boxes measured from simulated events are summarized in Table 9.2. The kinematic similarities between these backgrounds indicate that any differences between simulated events and data will affect each of these processes the same way in the context of the variables R and M_R. With insufficient integrated luminosity to resolve the R^2 scaling behavior of $t\bar{t}$ and events selected in di-lepton boxes we choose instead to infer the shapes of these backgrounds by directly measuring the R^2 scaling parameters for $W(\ell\nu)$+jets events in data and deriving data/MC slope correction factors to apply to the shapes of the other backgrounds.

Process / Box	a	b
$W(\mu\nu)$+jets / μ Box (1st)	$(-126 \pm 3) \times 10^{-4}$	$(-171 \pm 3) \times 10^{-3}$
$W(\mu\nu)$+jets / μ Box (2nd)	$(-58 \pm 6) \times 10^{-4}$	$(-52 \pm 5) \times 10^{-3}$
$W(e\nu)$+jets / e Box (1st)	$(-124 \pm 3) \times 10^{-4}$	$(-189 \pm 4) \times 10^{-3}$
$W(e\nu)$+jets / e Box (2nd)	$(-67 \pm 6) \times 10^{-4}$	$(-52 \pm 5) \times 10^{-3}$
$W(e\nu)$+jets / HAD Box (1st)	$(-131 \pm 3) \times 10^{-4}$	$(-119 \pm 3) \times 10^{-3}$
$W(e\nu)$+jets / HAD Box (2nd)	$(-48 \pm 5) \times 10^{-4}$	$(-52 \pm 4) \times 10^{-3}$
$W(\tau\nu)$+jets / HAD Box (1st)	$(-132 \pm 3) \times 10^{-4}$	$(-125 \pm 3) \times 10^{-3}$
$W(\tau\nu)$+jets / HAD Box (2nd)	$(-60 \pm 3) \times 10^{-4}$	$(-43 \pm 3) \times 10^{-3}$
$W(\mu\nu)$+jets / HAD Box	$(-103 \pm 1) \times 10^{-4}$	$(-43 \pm 1) \times 10^{-3}$
$t\bar{t}$ / μ Box	$(-733 \pm 9) \times 10^{-5}$	$(-450 \pm 10) \times 10^{-4}$
Single top / μ Box	$(-97 \pm 2) \times 10^{-4}$	$(-670 \pm 20) \times 10^{-4}$
$t\bar{t}$ / e Box	$(-721 \pm 8) \times 10^{-5}$	$(-430 \pm 10) \times 10^{-4}$
Single top / e Box	$(-101 \pm 2) \times 10^{-4}$	$(-620 \pm 20) \times 10^{-4}$
$t\bar{t}(\mu + X)$ / HAD Box	$(-663 \pm 8) \times 10^{-5}$	$(-459 \pm 9) \times 10^{-4}$
$t\bar{t}$(no $W \to \mu\nu$) / HAD Box	$(-735 \pm 5) \times 10^{-5}$	$(-398 \pm 6) \times 10^{-4}$
Single top / HAD Box	$(-896 \pm 9) \times 10^{-5}$	$(-550 \pm 9) \times 10^{-4}$

Table 9.2: Values of parameters describing the M_R evolution with R^2 for simulated events of different processes selected with the MU Box, ELE Box and HAD BOX requirements. Two sets of slope parameters are listed for processes with a significant second component.
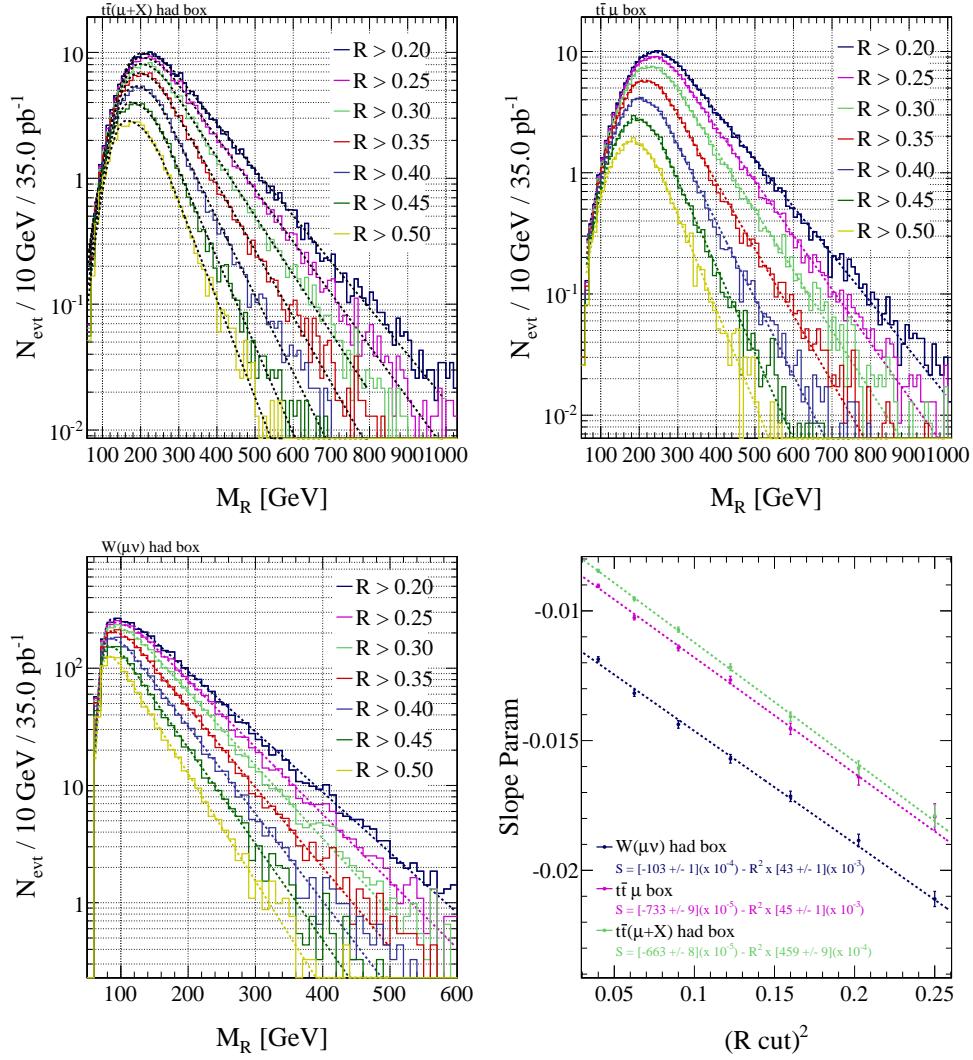
Figure 9.22: (Left) M_R distributions for different values of the cut on R for simulated $t\bar{t}(\mu + X)$ events satisfying the HAD Box selection. (Center) M_R distributions for different values of the cut on R for simulated $t\bar{t}$ events satisfying the MU Box selection. Fits to the exponential part of the M_R distribution are shown as dotted black lines. (Right) Value of the exponential slope S from fits to the M_R distribution, as a function of R cut for simulated $W(\mu\nu)$+jets and $t\bar{t}$ events satisfying different selections.

9.3 The 35 pb^{-1} CMS Razor Search for SUSY

Experimental limits from the Tevatron and LEP showed that superpartner particles, if they exist, are significantly heavier than their Standard Model counterparts. Thus proposed experimental searches for R-parity SUSY [164–168] at the LHC have focused on a combination of two SUSY signatures: multiple energetic jets and/or leptons from the decays of pair-produced superpartners, and large missing transverse energy from the two lightest superpartners produced in those same decay chains.
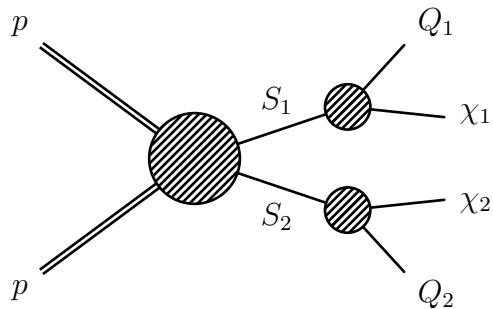
Figure 9.23: General R-parity conserving SUSY LHC event signature. Two massive sparticles, S_i, are produced in a pp collision and each decay to a system of detectable SM particles, Q_i, and a system of weakly interacting particles, χ_i.

Here, we present an inclusive search for SUSY based on the razor kinematic variables [169]. In this approach, all the reconstructed final state objects in each event a grouped into two *mega-jets*, mimicking the the expected R-parity conserving signal topology of two pair-produced sparticles each decaying to a system of visible SM particles and one or more stable, weakly interacting LSPs, illustrated in figure 9.23. From the three momenta of these mega-jets we calculate, event by event, the razor variables M_R and R, which are derived in section 9.1

$$
\begin{aligned}
M_R &\equiv 2\sqrt{\frac{(|\vec{p}^{\,j1}|p_z^{j2}-|\vec{p}^{\,j2}|p_z^{j1})^2}{(p_z^{j1}-p_z^{j2})^2-(|\vec{p}^{\,j1}|-|\vec{p}^{\,j2}|)^2}} \ , \\
M_T^R &\equiv \sqrt{\frac{|\vec{M}|(|\vec{p}_T^{\,j1}|+|\vec{p}_T^{\,j2}|)-\vec{M}\cdot(\vec{p}_T^{\,j1}+\vec{p}_T^{\,j2})}{2}} \ , \\
R &\equiv M_T^R/M_R \ ,
\end{aligned}
\tag{9.23}
$$

where $\vec{p}^{\,j1}$ and $\vec{p}^{\,j2}$ are the three momenta of the two mega-jets and $\vec{M}$ is the missing

transverse energy. As described in section 9.2, the phenomenological properties of these variables make them ideal for distinguishing between SM backgrounds and SUSY collision events.

The variable M_R is sensitive to the *scale* of events, such that for signal events of the type shown in figure 9.23 it will resolve the mass splittings between the parent sparticles S_i and the masses of the weakly interacting systems of particles, χ_i. The variable R is sensitive to the transverse imbalance of events, and will suppress topologies that have less than two weakly interacting particles in the final state. This is accomplished without making strong assumptions about the missing transverse energy spectrum or any details of the intermediate decay chains. If the difference between the masses of new sparticles and those of the SM are resolvable, SUSY events will appear at larger values of M_R and R than the vast majority of SM background events. Using the phenomenological M_R scaling of these backgrounds, as described in section 9.2.1, we can make measurements of these background shapes and yields at low M_R and R and use this information to predict these same backgrounds in the high M_R and R razor plane where we could observe excesses of signal events.

The strategy and execution of the 2010 CMS razor SUSY search, performed on 35 pb^{-1} of *pp* collision data at $\sqrt{s} = 7$ TeV, can be summarized as follows:

1. Events are selected by triggers identifying events by the presence of high transverse momentum calorimetric energy depositions or the presence of signals consistent with leptons. This is described in section 9.4.1.

2. Jets and leptons are reconstructed and identified in these events, as described in section 9.4.2.

3. The reconstructed objects in each event are combined into two mega-jets, which are used to calculate the variables M_R and R. Several baseline kinematic requirements are applied to clean the event samples of mis-reconstructed events. This procedure is described in section 9.4.3.

4. Events are assigned to *boxes* based on the presence or absence of a reconstructed

lepton. This box partitioning scheme allows us to isolate individual background processes based on final state particle content and kinematic phase space in order to measure their shape and yield. The details of how each box is defined and what measurements are made in each one are provided in section 9.5.

- **ELE and MU Boxes**: if there is a reconstructed electron or muon events are assigned to the ELE or MU boxes, respectively. These selections, along the razor variables, allow us to identify $W(\ell\nu)$+jets background control regions where we can measure the shape and normalization of this background and use it to predict W, Z and top quark backgrounds in the high M_R/R signal regions in each of the boxes

- **HAD Box**: Events without leptons are assigned here, and must also satisfy the H_T trigger signal selection trigger requirements. The backgrounds populating this box include QCD multi-jets, $Z(\nu\nu)$+jets, $W(\ell\nu)$+jets and top events, some with charged leptons that did not satisfy the standard CMS electron and muon selection.

- **QCD control Boxes**: For each of the ELE, MU and HAD boxes we also define a QCD multi-jet control sample analogue with additional requirements designed to enhance this contribution. For the lepton final states, we reverse the lepton identification requirements in order to get a multi-jet enriched, and EWK background suppressed, sample of events with leptons. In the hadronic final state, we use prescaled, low-threshold jet triggers to select events in low M_R/R kinematic phase-space, where the relative production cross sections ensure us a pure QCD multi-jet sample. The control samples are used to predict the QCD multi-jet background contributions to the ELE, MU and HAD boxes.

5. The R and M_R shape and normalization of various backgrounds are measured in different box control regions. These measurements are used to predict the SM backgrounds to the high R and M_R signal regions in the ELE, MU and HAD boxes, as described in section 9.6.

6. The observed yields in the ELE, MU and HAD box signal regions are compared with the predicted yields and used to make inferences about the presence or absence of SUSY events in these data samples. section 9.7 explains how these search results are used to constrain the parameter space of hypothetical SUSY models.

9.4 Event Selection and Reconstruction

Throughout this analysis, we use two different sources of collision event samples. The first are events triggered and reconstructed during 2010 pp collision running of the LHC using the CMS detector. The second are Monte Carlo (MC) events samples, where event generators are used to simulate the particles produced in proton collisions, propagate these particles through a representation of the CMS detector, digitize the hypothetical signals these particles would leave in the detector and reconstruct the event as if it were recorded in data taking. The different event samples used in this analysis are described in section B.1.

9.4.1 Trigger Selection

Events are recorded by the CMS detector if they satisfy one or more online trigger requirements. These triggers are based on fast, approximate reconstruction of the event that mimics later reconstruction requirements that will be applied in identifying jets and leptons. In the 2010 CMS razor search we employ triggers based on three different types of physics object based in data:

- **Muon triggers:** Events with high-p_T muons are selected and recorded online using the Level-1 muon trigger and the high-level trigger (HLT), which require information from the muon chambers (both Level-1 and HLT) and the inner tracker (just HLT). The HLT trigger paths used in this analysis consider HLT muons in the $|\eta| < 2.1$ region with p_T thresholds varying according to which instantaneous luminosity trigger menu was deployed during the run. The HLT

paths used, `HLT_Mu9`, `HLT_Mu11` and `HLT_Mu15`, with HLT p_T thresholds of 9, 11 and 15 GeV/c, respectively, are all found to be fully efficient for muons satisfying the offline identification and kinematic requirements.

- **Electron triggers:** High-p_T electrons are selected and recorded online using the Level-1 electron trigger and the HLT trigger which require information from the ECAL (with coarse granularity at Level-1 and the full granularity and precise energy calibration at the HLT). Depending on the HLT menu deployed during running, various electron trigger paths are used to efficiently select electron events. At HLT an ECAL cluster with $E_T > 15$ GeV is required for trigger paths, `HLT_Photon15_Cleaned_L1R`. In higher luminosity running, we use the path `HLT_Ele15_SW_CaloEleId_L1R`, which has additional shape requirements on the HLT ECAL cluster. Each of these triggers is determined to be fully efficient with respect to the offline electron reconstruction and identification requirements.

- **Jet triggers:** Events with jets are selected and recorded online using the Level-1 single jet trigger and the HLT di-jet trigger which require information from ECAL and HCAL. The trigger `HLT_DiJetAve15U` corresponds to a requirement of two jets with arithmetic E_T average above 15 GeV. These HLT jet energies are not corrected for non-unity calorimeter energy response. For high-luminosity running, these triggers were pre-scaled.

- **HT triggers:** The H_T trigger paths used to select signal events in the HAD box are recorded online using the Level-1 single jet trigger and additional HLT requirements based on calculations of H_T. HLT jets with uncorrected $E_T > 20$ GeV are included in the HLT H_T definition,

$$H_T^{HLT} = \sum_i^{jets} E_T^i \ . \tag{9.24}$$

Depending on the run period, three different H_T threshold triggers were the lowest threshold, un-pre-scaled trigger: `HLT_HT100U`, `HLT_HT140U` and `HLT_HT150U`.

section B.2 describes how the kinematic effects of these triggers are accounted for in the analysis

9.4.2 Physics Object Reconstruction and Identification

9.4.2.1 Primary Vertex Reconstruction

Selected events are required to have a least one Primary Vertex (PV) reconstructed in the event satisfying several conditions. The vertex must be constructed with at least 13 associated degrees of freedom (at least 14 tracks matched to this vertex) and must be within a distance $|\Delta z| < 25$ cm from the beam spot along the beam axis. When multiple PV are reconstructed in an event the one with highest associated $\sum_{track} p_T$ is used to project physics object from when calculating their momenta.

9.4.2.2 Muon Identification and Selection

CMS muon identification is based on a number of quality criteria resulting from muon candidate reconstruction and is designed to suppress the rates of non-muons resulting in mis-identification. More details about muon reconstruction can be found in section 3.5.

For the typical range of transverse momenta explored in this analysis, the muon momentum resolution is dominated by the inner tracker measurements. A good consistency between tracker and muon detector measurements is essential to reduce the contamination from muons produced in decays in flight of hadrons and from punch-through. For the muons in this search, each muon must be identified by two different algorithms, one that starts from the inner tracker information ("tracker muons"), and another one that starts from the segments in the muon chambers ("global muons"). A requirement of χ^2 per degree of freedom of < 10 is imposed on a global fit containing tracker and muon hits. The presence of at least two levels of muon stations in the measurement is required to ensures a sensible momentum estimate at the muon trigger level and to further suppress remaining punch-through and sail-through candidates, since these are unable to penetrate deeply in the iron yoke of CMS. For a

precise estimate of momentum and impact parameter, only tracks with more than 10 tracker hits and at least one hit in the pixel detector are used. Cosmic ray particles are rejected by requiring a transverse impact parameter distance to the beam spot position of less than 2 mm. Furthermore, the track must have at least one pixel detector hit. Dedicated studies of muons in cosmic runs show that the high-p_T cosmic contamination after these cuts is negligible in the present sample. More details and studies on muon identification in CMS at $\sqrt{s} = 7$ TeV can be found in reference [170].

For the definition of the MU box in this search we require that the sum of the transverse momentum of tracks in an isolation cone of 0.3 in $\Delta R = \sqrt{\Delta\eta^2 + \Delta\phi^2}$ around the muon candidate track are less than 15% of the candidate momentum. In addition to all of the above requirements, muons must have $|\eta| < 2.1$ and $p_T > 20$ GeV/c.

9.4.2.3 Electron Identification and Selection

Electrons are identified in the CMS detector as clusters of ECAL energy deposits matched to tracks from the silicon tracker. The reconstruction of electrons is described in detail in section 4.3. In order to qualify as a reconstructed electron candidates must satisfy a number of quality criteria based on the properties of its ECAL clusters, its matched track and the inter-consistency of the energy and momentum measurements of the two, respectively.

For electrons in this search, we require that candidates have an ECAL cluster with $E_T > 20$ GeV with $|\eta| < 1.4442$ for barrel (EB) clusters or $1.566 < |\eta| < 2.500$ for endcap (EE) clusters. ECAL clusters are required to match tracks using an algorithm [171] which accounts for possible energy loss due to Bremsstrahlung in the tracker layers. Particles misidentified as electrons are suppressed by requiring the track trajectory η and ϕ, extrapolated to the face of the ECAL from the interaction point, to matches the position of the ECAL cluster η and ϕ. Additional misidentification is reduced by limiting the amount of HCAL energy measured in a cone of $\Delta R < 0.15$ around the ECAL cluster direction and by requiring a narrow ECAL cluster width in η. Misidentified particles, as well as real electrons arising from jet

fragmentation, are suppressed by imposing isolation cuts: limits on the additional sums of HCAL E_T, ECAL E_T, and track p_T in a cone of $\Delta R < 0.3$ around the electron candidate direction. Tracks and ECAL energy associated with the electron are excluded from these sums. Limits range from 3% to 10% of the electron candidate E_T, depending on the subdetector and ECAL region.

9.4.2.4   Jet Identification and Selection and
Missing Transverse Energy Reconstruction

Jets are reconstructed from calorimeter towers which are composed of a 5 x 5 array of ECAL crystals and a HCAL module. The energy depositions in these towers are clustered into jets using the infrared and collinear-safe anti-kt jet algorithm [57] with a cone size $R_{\mathrm{cone}} = 0.5$ in the $(\eta \times \phi)$ space. More details about jet reconstruction in CMS events can be found in section 5.2. The jets are corrected for noncompensating calorimetric energy response using Monte Carlo derived corrections, and they are required to have a $p_T > 30$ GeV threshold and be within $|\eta| < 3.0$.

The missing transverse energy is calculated as the negative transverse vectorial sum of all of the reconstructed particle flow (PF) candidates in the event. The PF reconstruction algorithm is described in section 5.2. Effectively, this collection of PF candidates accounts for all of the reconstructed energy and momentum in each of the CMS subdetector systems.

In addition to jets clustered from calorimeter deposits, we also consider jets clustered from tracks (track-jets) [172] and PF candidates (PF-jets). Intermediate results using these alternative jet types provide useful cross-checks to the primary calorimeter jets in the analysis, in that they rely on measurements from different subdetectors and have different calibrations.

9.4.3 Kinematic Requirements and Mega-Jet Reconstruction

Each event is required to have at least two reconstructed objects (jets or leptons) in order to create two mega-jets. If there are more than two reconstructed objects in an event every partition of the objects into two non-empty groups is considered. A mega-jet is formed from the objects assigned in its partition by summing their four vectors,

$$p_{MJ}^n = \sum_{i \in n} p_c^i \, , \tag{9.25}$$

where the index n runs over the two mega-jet four vectors p_{MJ}^n and the index i indicates the objects' four vectors p_c^i which are assigned to mega-jet n. Out of all the possible partitions of the reconstructed objects into two mega-jets the one which minimizes the mega-jet masses, $(p_{MJ}^n)^2$, summed in quadrature is chosen, an algorithm adopted from [173].

This choice of mega-jet construction algorithm implicitly maximizes the momenta of the two mega-jets. On the other hand, the same choice leads to more *balanced* events than other potential algorithms, in that the total momentum of the two mega-jets $p_{MJ}^1 + p_{MJ}^2$ is also implicitly maximized. The net result is that alternative algorithms could yield larger values of M_R with smaller values of R, on average. In practice, we find the algorithm to be quite stable and effective at rejecting background with the kinematic variables of interest.

When constructing mega-jets, one can choose whether to include reconstructed leptons in the final state among the visible objects used to build the mega-jets; alternatively, the leptons can be treated as *invisible* and removed from consideration in the calculation of the kinematic variables (as if they were escaping weakly interacting particles). For backgrounds like $W(\ell\nu)$+jets, the former choice yields more transversely balanced mega-jets, and lower values of R, due to the fact that, since they come from the same decay, harder neutrinos (and hence larger MET) will also

produce leptons. For this background, when leptons are treated as invisible, the MET corresponds to the entire W transverse momentum, similar to the case of $Z(\nu\nu)$+jets events.

In this analysis, electrons are treated as visible, and included in the mega-jet calculation in the default ELE box treatment, while muons are treated as invisible for the MU box. This choice was made to maintain consistency with the reconstruction of the kinematic variables at the trigger level, which is based on calorimetric objects that are not sensitive to MIP muons. While the background in muon final states can be artificially less constrained by the kinematic variables with this choice, this allows us to use $W(\mu\nu)$+jets events to study and constrain the shape of $Z(\nu\nu)$+jets events in the HAD box.

The longitudinal boost velocity that relates the laboratory frame to the R frame described in section 9.1.1, β_R, is calculated from the mega-jets momentum, $\vec{p}_{MJ}^{\,i}$, as

$$\beta_R = \frac{|\vec{p}_{MJ}^{\,1}| - |\vec{p}_{MJ}^{\,1}|}{p_{MJ,z}^1 - p_{MJ,z}^2} \ . \tag{9.26}$$

If $\beta_R \leq 1$ the boost is well-defined, as is the variable M_R, and we can include the event in the selected event sample. A fraction of events have mega-jets that lead to ill-defined configurations with $\beta_R \geq 1$. In order to minimize the frequency of ill-defined events, we calculate β_R for each possible pair of mega-jets when deciding how to partition the objects in an event, considering only assignments that result in $\beta_R \leq 1$, if one exists. We further require that $\beta_R < 0.99$ in order to remove events for which the variables used in the analysis would be singular, in particular rare QCD multi-jet events that can have erroneous M_R values. The efficiency of this requirement for typical signal events is close to 100%.

This is the entirety of the baseline selection for the analysis. The razor variables M_R and R are calculated from the mega-jet's momentum, event by event, and the analysis consists of making measurements of and putting additional requirements on these kinematic variables.

9.5 Control Regions and Boxes

9.5.1 Hadronic QCD Control Box

In section 9.2 we describe the phenomenological properties of the SM background in terms of the razor variables, in particular M_R scaling. The tail of backgrounds' M_R distribution falls exponentially, with the slope of the decay set by the requirement on R. The shape of the M_R distribution for the QCD mutlijet background is particularly sensitive to this R requirement, changing more dramatically than other backgrounds. This is because these events have no source of high transverse momentum weakly-interacting particles that are isolated from the visible SM particles in the event, unlike backgrounds with on-shell W and Z bosons. This means that at larger values of M_R and R the rate of QCD mutlijet events is small relative to EWK and top backgrounds while at low M_R and R the mutlijet event yield dwarves the others by several orders of magnitude. In order to measure the shape of this background in the razor plane, which can be used to predict its contribution at large M_R and R, we must be able to select events at low M_R and R values with extremely high efficiency so that the shapes are not kinematically biased. This is not possible with the signal H_T triggers used to select events in the HAD box since the trigger requirement dramatically alters the M_R distribution in in the interval of interest for isolating QCD mutlijet events (see section B.2).

We define a hadronic QCD control box by applying the HAD box base-line selection (lepton veto) and requiring a low-threshold jet trigger, `HLT_DiJetAve15U`. The trigger requires at least two uncorrected jets with average $p_T > 15$ GeV, a threshold low enough such that it doesn't bias the offline M_R distribution in the range required for measuring the M_R scaling parameters for QCD mutlijet events. Given the high prescale of this trigger and the expected signal rates, there is no contamination from signal events in the QCD control box. In fact, in the M_R range considered for these mutlijet shape measurements there is negligible contamination from any SM background process (EWK, top) other than QCD mutlijets, an observation discussed in section B.8.

With a pure sample of QCD mutlijet events without kinematic bias from trigger requirements, we study the M_R shape of these events and measure its evolution as a function of R requirement. The M_R distributions for events selected in the hadronic QCD control box, for different values of the R cut, are shown in figure 9.24 (left). We observe that the M_R distribution, after an approximate Gaussian turn-on at low M_R resulting from the p_T threshold requirement on jets entering the mega-jet calculation, drops exponentially. The exponential slope, S, of the distribution scales linearly with the square of the R requirement, the characteristic feature of the M_R scaling described in section 9.2.1 and exhibited in figure 9.24 (right).
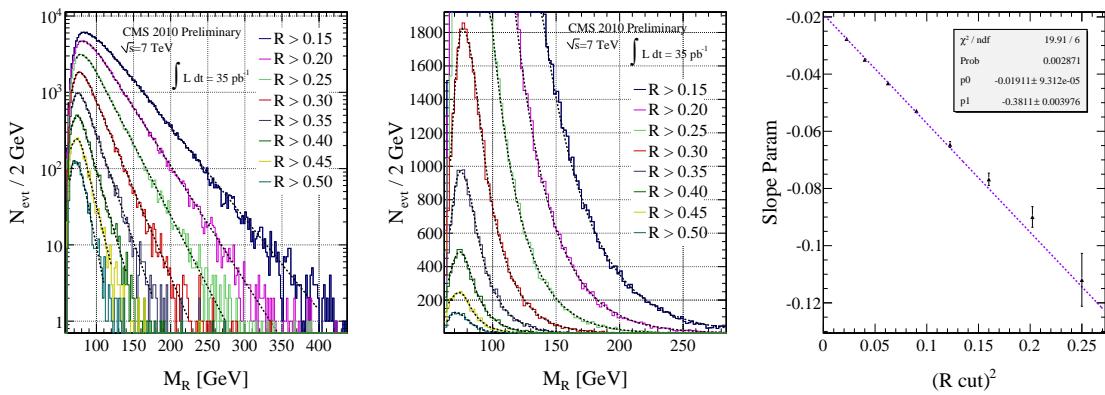
Figure 9.24: (Left and Center) M_R distributions for different values of the cut on R for events in data selected in the QCD control box. Fits to the exponential part of the M_R distribution are shown as dotted black lines. (Right) Value of the exponential slope S from fits to the M_R distribution, as a function of R cut.

In order better understand the M_R scaling behavior appearing in QCD mutlijet events we repeat the same set of measurements for events reconstructed with track-jets and PF-jets. The results for these alternative jet types are shown in figure 9.25 and 9.26. We observe the same qualitative M_R scaling of the slopes S.

If we parameterize the R requirement dependence on the M_R exponential slope as

$$S = a + b \cdot (R \text{ requirement})^2 \, , \tag{9.27}$$

we can then compare the quantitative features of the M_R scaling for QCD mutlijet
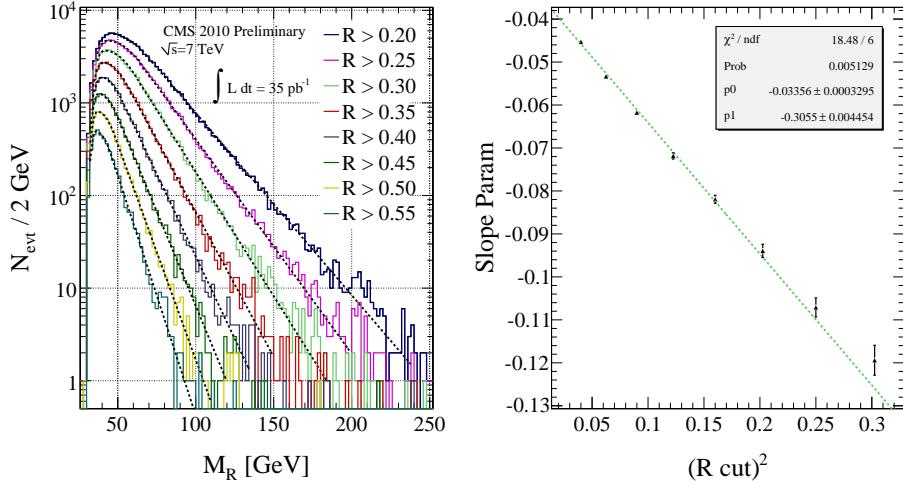
Figure 9.25: (Left) M_R distributions (using track-jets) for different values of the cut on R for events in data selected in the hadronic QCD control box. Fits to the exponential part of the M_R distribution are shown as dotted black lines. (Right) Value of the exponential slope S from fits to the track-jet M_R distribution, as a function of R cut.
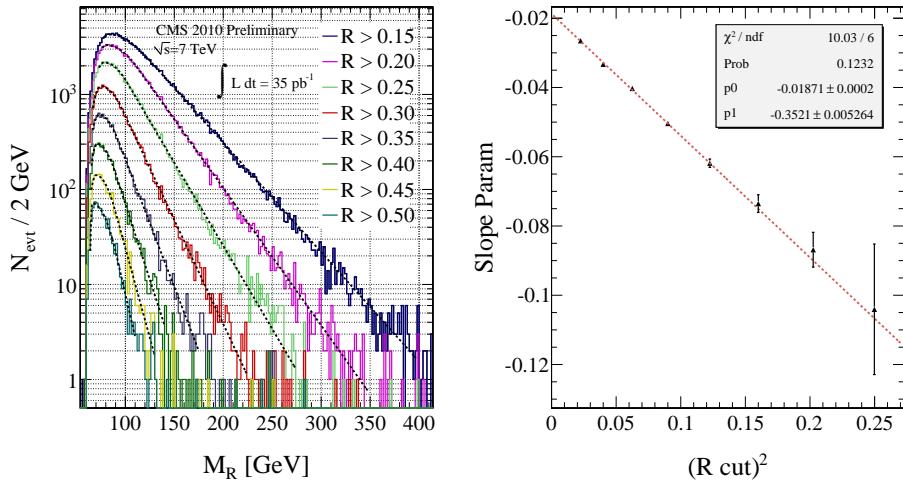
Figure 9.26: (Left) M_R distributions (using PF-jets) for different values of the cut on R for events in data selected in the hadronic QCD control box. Fits to the exponential part of the M_R distribution are shown as dotted black lines. (Right) Value of the exponential slope S from fits to the track-jet M_R distribution, as a function of R cut.

events using different jet types. The values of the parameters a and b describing the M_R scaling of the slopes S for all jet types are summarized in Table 9.3.

Comparing the M_R in these different jet types is interesting for a number of

Jet Type	a	b
Calo Jets	$(-1911 \pm 9) \times 10^{-5}$	$(-381 \pm 4) \times 10^{-3}$
Track Jets	$(-336 \pm 3) \times 10^{-4}$	$(-306 \pm 4) \times 10^{-3}$
PF Jets	$(-187 \pm 2) \times 10^{-4}$	$(-352 \pm 5) \times 10^{-3}$

Table 9.3: Measured values of the slope parameters a and b for different jet types in the QCD control box.

reasons. The calorimeter jets which the razor SUSY search is based on are reconstructed exclusively using the ECAL and HCAL while track-jets are clustered only from tracks reconstructed in the inner detector. This means that the majority of systematic detector-related effects which could alter the M_R and R distributions should be entirely different between these two jet types. Similarly, PF-jets provide a useful comparison to calorimeter jets in that they include information from other subdetectors apart from the calorimeters and are also calibrated independently to have unity energy response. This allows for meaningful quantitative comparisons between the scaling parameters of the different jet types.

The relative values of the parameter a between jet types is understood as the ratio of the relative physics object energy scales. Jet energy scale (JES) corrected calorimeter and PF-jets are calibrated so that the modal energy response is equal to the true jet energy in simulated events. Raw track-jets are made only of charged tracks, which only carry about 60% of a jet's momentum, on average, with large jet-by-jet fluctuations. This means that the momentum response of track-jets, relative to the true jet energy is only $\sim$60% of that for PF-jets or calorimeter jets. Comparing the a result of the PF and track-jets in the table above we measure

$$\frac{a(\text{PF Jet})}{a(\text{track Jet})} = \frac{187 \pm 2}{336 \pm 3} = (55 \pm 1) \times 10^{-2} \ , \tag{9.28}$$

in agreement with the expectations based on the relative energy response of the two jet types. Correspondingly, the value of the a parameters for calorimeter and PF-jets are indistinguishable. Energy scale dependencies are absorbed in the value of a. The differences in b among the types of jets result from variations in the implicit phase space requirements placed on jets in order to be reconstructed. The p_T thresholds

applied in order for each type of jet to enter their respective mega-jet calculations are chosen to correspond roughly to the same particle-level jets (30 GeV/c for calorimeter and PF jets, 15 GeV/c for track jets). This results in the similarities of measured values b for different jet types. Residual differences in b encapsulate the differences in geometrical acceptance and position/momentum resolution of the sub-detectors contributing to the different jet reconstruction schemes.

With the M_R scaling parameters a and b measured for QCD mutlijet background event we can predict their M_R slope for large R values, where we would like to search for SUSY events. When measuring the slope parameters a and b we fit the M_R exponential slopes, as a function of R requirement, as if they were statistically independent measurements. This approach neglects the fact that the dataset used in each fit is a small subset of the one with slightly looser R requirement. We study the effect of neglecting these statistical correlations in the fitting procedure in section B.4 and conclude that it is a small effect, with a correspondingly small systematic error added to background yield predictions which rely on the a and b parameters, described in section 9.6.

Finally, the HAD QCD control sample is useful for monitoring the stability of SM backgrounds' M_R shape as a function of time and instantaneous luminosity. Section B.7 discusses potential biases resulting from these effects and constrains them to be negligible based on measurements in with this control sample.

9.5.2   ELE and MU QCD Control Boxes

The lepton identification requirements which define the ELE and MU boxes are optimized to efficiently identify leptons coming from heavy boson decays while rejecting leptons resulting from the hadronization and decay products of quarks and gluons or mis-identified particles. These requirements heavily suppress the yields of QCD multijet events in the lepton final state boxes. However, we cannot a priori conclude that this multijet contribution to the SM background in our high M_R R signal regions is negligible. In order to infer the M_R shape and scaling behavior of these lepton fi-

nal state multijet events, we define lepton QCD control box requirements which can isolate a pure sample of QCD multijet events in selected data which have trigger-ing lepton candidates and kinematics similar to the backgrounds expected in the ELE and MU boxes.

The ELE and MU QCD control boxes are defined by the same baseline require-ments as the ELE and MU boxes, apart from the fact that the isolation requirements on the leptons are inverted, resulting in a sample of multijet events with non-isolated leptons. This is the same methodology used for the "anti-lepton" heavy flavor control samples described in [174, 175]).

The M_R distributions for data and simulated QCD multijet events satisfying the ELE box selection are shown in figure 9.27. A comparison between the M_R shape for simulated QCD multijet events in the ELE and ELE QCD control boxes demonstrates that they are indistinguishable with the available simulated event statistics. We see that the expected contribution of EWK and top backgrounds to the ELE QCD control box sample is negligible relative to the QCD multijet yield. The M_R distribution falls exponentially.
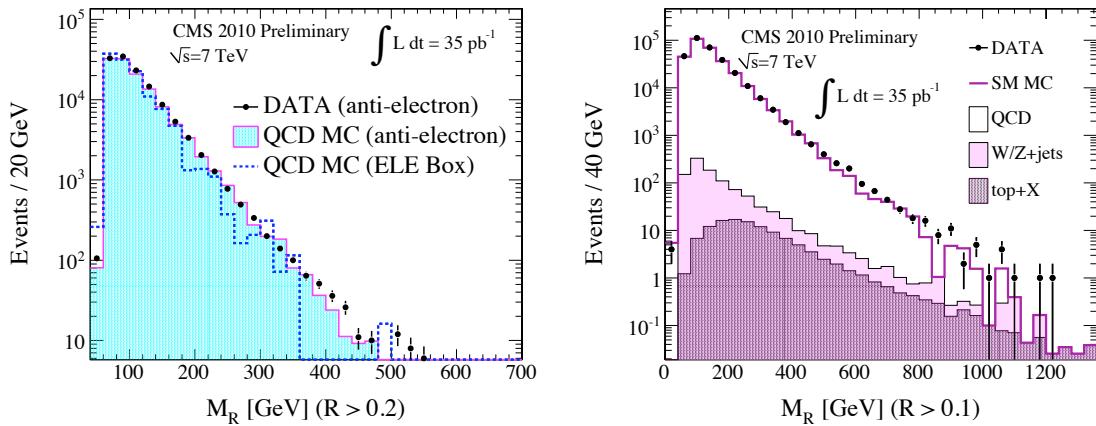
Figure 9.27: (Left) M_R distribution for events selected in the ELE box for data with QCD electron control selection and for simulated QCD events with default electron selection and with inverted electron isolation selection. A cut of $R > 0.1$ is ap-plied. (Right) M_R distribution for the ELE box QCD control selection for data and simulation.

We measure the M_R scaling parameters for events selected in the MU and ELE

QCD control boxes, with the M_R shape and slope fits shown in figure 9.28 and 9.29, respectively. The fit parameters a and b which describe the M_R exponential slopes in these lepton-enriched QCD multijet samples are summarized in Table 9.4 and are used to constrain the contribution to the ELE and MU box event yields from this background.

Figure 9.28: (Left) M_R distributions for different values of the cut on R for events in data satisfying the QCD muon control selection. Fits to the exponential part of the M_R distribution are shown as dotted colored lines. (Right) Value of the exponential slope S from fits to the M_R distribution, as a function of R cut for data events satisfying the QCD muon control selection.

QCD control selection	a	b
QCD HAD Box	(-1911 ± 9) x 10^{-5}	(-381 ± 4) x 10^{-3}
QCD MU Box	(-1576 ± 4) x 10^{-5}	(-1224 ± 8) x 10^{-4}
QCD ELE Box	(-1717 ± 2) x 10^{-5}	(-1902 ± 6) x 10^{-4}

Table 9.4: e Box Slope Measurements (MC and Data)

The values of a and b for QCD multijet events appearing in the ELE and MU QCD control boxes are not the same, due to differences in the kinematics of these final states and the experimental treatment of the relevant reconstructed physics object.

Figure 9.29: (Left) M_R distributions for different values of the cut on R for events in data satisfying the QCD electron control selection. Fits to the exponential part of the M_R distribution are shown as dotted colored lines. (Right) Value of the exponential slope S from fits to the M_R distribution, as a function of R cut for data events satisfying the QCD electron control selection.

9.5.3   ELE and MU Boxes

The high M_R and R part of the razor plane in the lepton final states is a region of phase space where we will look evidence of SUSY sparticle production. As we saw in section 9.5.1 and 9.5.2, differences in intrinsic scale and M_R scaling behavior between the SM backgrounds and hypothetical signal events mean that the relative yields from different contributions can vary dramatically across the razor plane. In this case, this allows us to identify a region in the razor plane in the ELE and MU boxes, not including the region where we will look for signal events, where the background is made up almost exclusively of $W(\ell\nu)$+jets events. Hence, while searching for new physics in one part of the razor plane we can measure the shape and yield of background events in another.

Using the MU Box selection, we identify M_R intervals for different R cuts where we expect the yield $W(\mu\nu)$+jets events to be significantly higher than other background

contributions. In particular, we are sensitive to the M_R scaling parameters of the 1st component of this $W(\ell\nu)$+jets background, with two distinct scaling components results from events with different numbers of associated jets produced with the W bosons. This two component phenomenology is discussed in section 9.2.4. The R^2 dependence of the M_R exponential slope is measured for these events and shown in figure 9.30.

Figure 9.30: (Left) M_R distributions for different values of the cut on R for data events satisfying the MU box selection. Fits to the exponential part of the M_R dist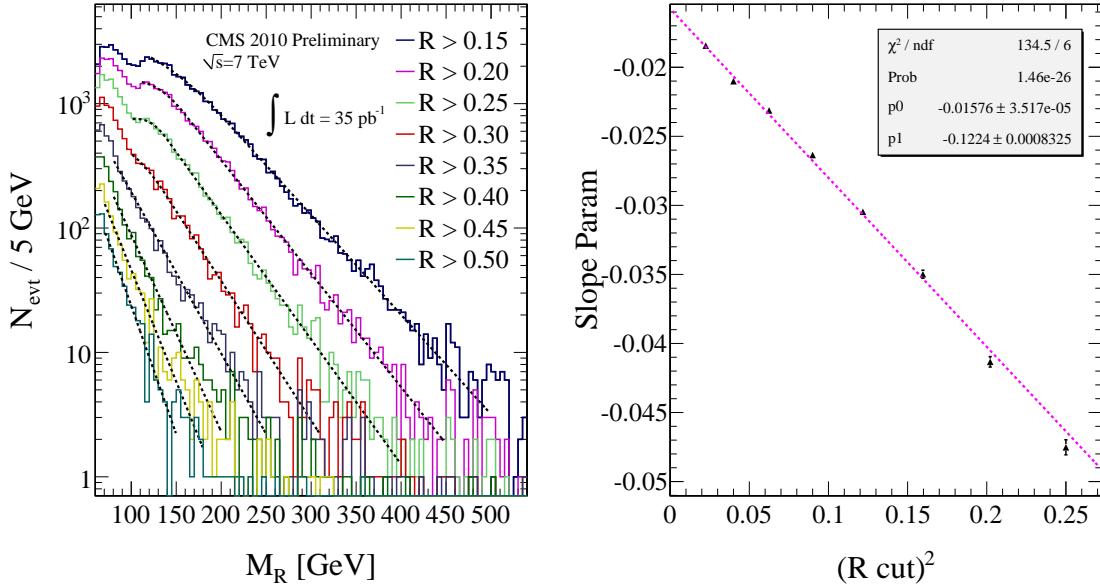ribution are shown as dotted colored lines. (Right) Value of the exponential slope S from fits to the M_R distribution, as a function of R cut.

Here, the M_R distribution is modeled with two independent exponential components, simultaneously floating both slopes along with their relative and absolute normalizations. From this fit we extract the slope parameters characterizing the exponential behavior of the 1st $W(\ell\nu)$+jets component. We observe that these exponential slopes exhibit M_R scaling behavior, in agreement with predictions, and find the values of the parameters a and b which describe the R^2 scaling to be in good agreement with the values extracted from simulated $W(\mu\nu)$+jets events.

The differences (and their uncertainties) between the values of these scaling parameters measured in data and in simulated events are used to construct data/simulation shape scale factors (SFs), $\rho^{\text{DATA/MC}}$, which are calculated as the ratio of the data and simulation measurements for the parameters a and b. From the MU Box we find

$$\rho(a)^{\text{DATA/MC}}_{MU} = \frac{117 \pm 3}{126 \pm 3} = 0.93 \pm 0.03 \ ,$$

$$\rho(b)^{\text{DATA/MC}}_{MU} = \frac{172 \pm 4}{171 \pm 4} = 1.00 \pm 0.03 \ .$$

This same procedure is repeated for $W(e\nu)$+jets events in the ELE Box. The results of the two component exponential fit, along with the extracted 1st component slope parameters are shown in figure 9.31. We observe R^2 scaling behavior for the slope of the extracted 1st $W(e\nu)$+jets component and find the values of the parameters describing this scaling in good agreement with the extracted values from simulated events. For the DATA/MC slope SFs we find from the ELE Box:

$$\rho(a)^{\text{DATA/MC}}_{ELE} = \frac{125 \pm 3}{124 \pm 3} = 1.00 \pm 0.03 \ ,$$

$$\rho(b)^{\text{DATA/MC}}_{ELE} = \frac{176 \pm 4}{189 \pm 4} = 0.93 \pm 0.03 \ .$$
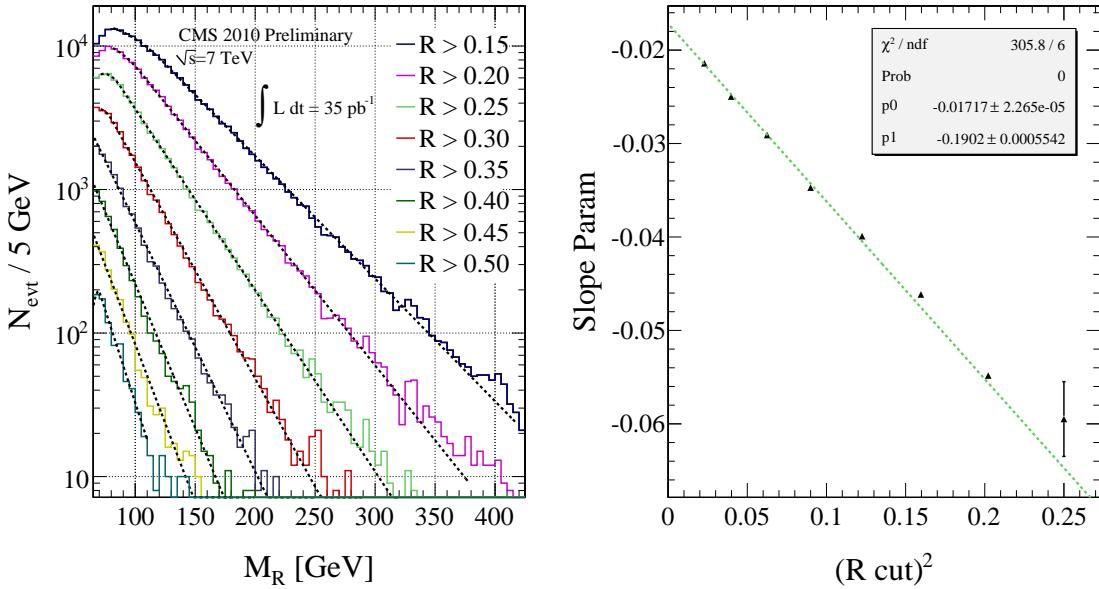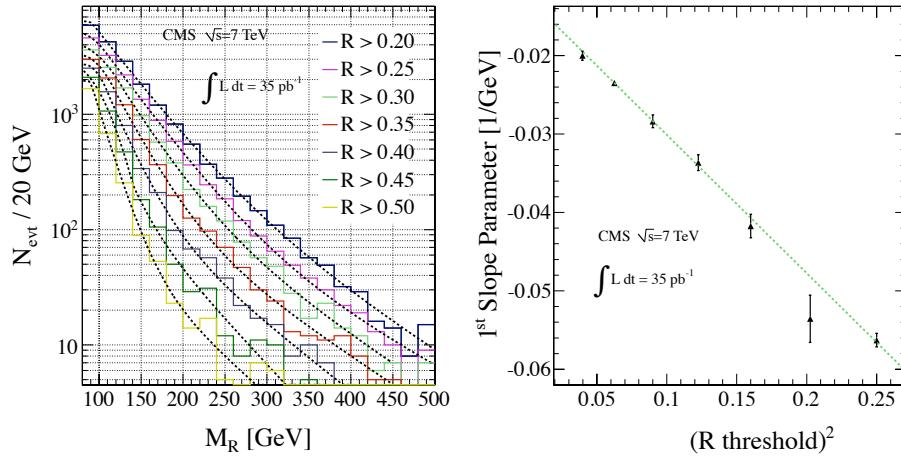

Figure 9.31: (Left) M_R distributions for different values of the cut on R for data events satisfying the ELE Box selection. Fits to the exponential part of the M_R distribution are shown as dotted colored lines. (Right) Value of the exponential slope S from fits to the M_R distribution, as a function of R cut.

The DATA/MC correction factors measured independently in the MU and ELE

boxes can be combined, yielding

$$\rho(a)_1^{\text{DATA/MC}} = 0.97 \pm 0.02 \ ,$$

$$\rho(b)_1^{\text{DATA/MC}} = 0.97 \pm 0.02 \ .$$

These shape SFs are used to predict the shapes of backgrounds to the ELE, MU and HAD boxes which are kinematically to the $W(\ell\nu)$+jets events studied here. These backgrounds are also predicted to exhibit M_R scaling, with shape parameters a and b taken from fits to simulated events (see section 9.2) and corrected with the shape SFs derived from the ELE and MU boxes.

These SFs are used to predict the shape parameters of all the backgrounds to the ELE, MU and HAD boxes, with the exception of the 2nd $W(\ell\nu)$+jets components in each box and $Z(\nu\nu)$+jets in the HAD Box. For these background processes, we can perform measurements using the ELE* and MU* selections described in section 9.2.3, where the reconstructed leptons are treated as if they were weakly interacting particles in the calculation of R and M_R. This means that the $W(\ell\nu)$+jets events that make up the majority of the background to the ELE* and MU* selections is kinematically almost identical to the $Z(\nu\nu)$+jets and 2nd $W(\ell\nu)$+jets component backgrounds.

We use the MU* box selection and lepton treatment in order to measure the M_R slopes of $W(\mu\nu)$+jets events in selected data. The fits of the MU* box M_R distribution, as a function of R cut, are shown in figure 9.32.

We find no significant second exponential component in the fit regions considered, and compare the M_R scaling parameters with expectations from simulated $W(\mu\nu)$+jets events satisfying the MU* box selection, observing that the SFs are indistinguishable from one. We use these values to derive DATA/MC shape correction SFs for the 2nd $W(\ell\nu)$+jets component and $Z(\nu\nu)$+jets events in the HAD Box:
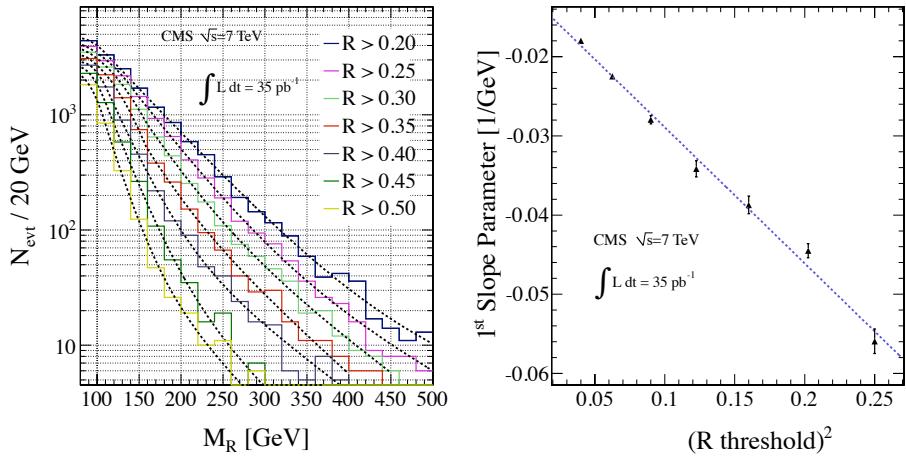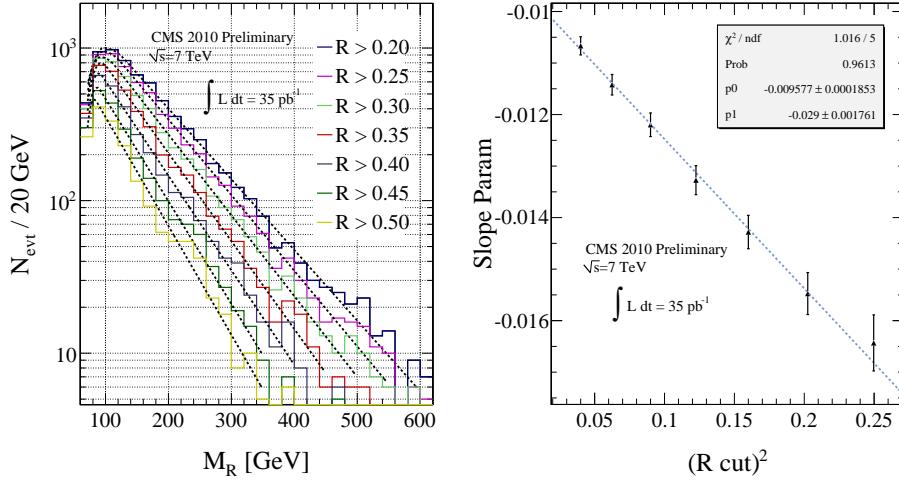
Figure 9.32: (Left) M_R distributions for different values of the cut on R for data events satisfying the MU* Box selection. Fits to the exponential part of the M_R distribution are shown as dotted colored lines. (Right) Value of the exponential slope S from fits to the M_R distribution, as a function of R cut.

$$\rho(a)_2^{\text{DATA/MC}} = \frac{96 \pm 2}{95 \pm 1} = 1.01 \pm 0.02 \ ,$$

$$\rho(b)_2^{\text{DATA/MC}} = \frac{29 \pm 2}{31 \pm 1} = 0.94 \pm 0.07 \ .$$

The results of these DATA/MC slope comparison measurements are summarized in Table 9.5.

Process / Box	a	b
DATA / MU Box ($W(\mu\nu)$+jets 1st)	(-117 $\pm$ 3) x 10^{-4}	(-172 $\pm$ 4) x 10^{-3}
$W(\mu\nu)$+jets / MU Box 1st	(-126 $\pm$ 3) x 10^{-4}	(-171 $\pm$ 3) x 10^{-3}
DATA / ELE Box ($W(e\nu)$+jets 1st)	(-125 $\pm$ 3) x 10^{-4}	(-176 $\pm$ 4) x 10^{-3}
$W(e\nu)$+jets / ELE Box 1st	(-124 $\pm$ 3) x 10^{-4}	(-189 $\pm$ 4) x 10^{-3}
DATA MU* Box ($W(\mu\nu)$+jets)	(-96 $\pm$ 2) x 10^{-4}	(-29 $\pm$ 2) x 10^{-3}
$W(\mu\nu)$+jets / μ^* Box	(-95 $\pm$ 1) x 10^{-4}	(-307 $\pm$ 9) x 10^{-4}

Table 9.5: Comparison of parameters describing the M_R evolution with R^2 for MU, MU* and ELE Box data and simulated events of different processes selected with the same box requirements.

9.6 Signal Region Yield Predictions and Observations

9.6.1 ELE and MU Box Background Predictions

In the ELE and MU final state boxes we define signal regions by applying additional requirements on the variables M_R and R, selecting events in the region of the razor plane where we expect potential signal events to appear with suppressed background yields. Without using observations of the event yields in these signal regions we predict the expected SM background contribution to these yields; the compatibility of these predictions to the actual, independent observations of events yields allows us to make statistical inferences about the contributions from SUSY signal events.

The procedure for assembling the total SM background predictions in the ELE and MU signal boxes can be summarized as follows:

- Calculate shapes in M_R for each SM process using M_R scaling parameters a and b and the signal region R requirement. If the parameters a and b were not measured directly in a data control box then we use the values measured from simulated events (section 9.2) corrected with shape scale factors ρ derived from kinematically similar control samples.

- Set the relative normalizations of the EWK and top backgrounds using process cross section measurements from CMS in different final states than the ones considered in this search.

- Set an overall normalization by measuring the event yield in a M_R side-band.

- With the EWK and top backgrounds fixed, determine the normalization of the residual QCD multijet background in the low M_R region using the shapes measured from the lepton QCD control boxes (section 9.5.2).

- Float the relative normalization of the first and second $W(\ell\nu)$+jets component in an intermediate M_R side-band, orthogonal to the signal region.

Here, we describe each of these steps in detail. In some cases, we briefly review measurements and observations that are described in greater detail earlier in this chapter.

The first step in the background prediction consists of setting the initial M_R *shape* of the EWK and top background components. The M_R distribution for each of these components is distributed (at sufficiently high M_R) as a falling exponential, $e^{S \cdot M_R}$, and the exponential slope S exhibits a dependence on the value of the R value that can be described as

$$S = a + b \cdot R^2 \ , \tag{9.29}$$

the characteristic feature of M_R scaling, described in section 9.2.1. Each process contributing as a background in the leptonic boxes is modeled as having different values of the parameters a and b, which are measured in simulated event samples (section 9.2). These MC measurements provide initial values for these parameters. In the case of the W+jets background, we identify regions in the razor plane of the different lepton boxes (ELE, MU, ELE*, MU*) where, because of selection and kinematics, we can select a pure sample of $W(\ell\nu)$+jets events and measure the M_R scaling slope parameters a and b from equation (9.29).

We compare the measured values of these parameters a and b between data and simulated events, calculating DATA/MC correction factors $\rho(a)$ and $\rho(b)$ defined as the ratio of the data and MC values of the parameters. This is done independently for $W(\ell\nu)$+jets events selected in each of the ELE, MU and MU* box selections. We observe that, in data and simulated events, the measurements of a and b agree between the electron and muon final states and between data and MC in each case. This implies that the DATA/MC correction factors ρ measured for each of the boxes are consistent with 1, with a precision set by measured errors of the data and MC values of a and b.

The values of the ρ parameters measured in the ELE and MU boxes correspond to regions of phase space where the background is composed almost entirely by the first W+jet component. The ρ parameters from these two boxes are combined to

give $\rho(a)_1$ and $\rho(b)_1$, DATA/MC scale factors that are used to correct the shape parameters a and b measured for other, kinematically similar SM backgrounds in simulated events. The scale factors $\rho(a)_1$ and $\rho(b)_1$ are used in the shape prediction for all of the backgrounds in the ELE, MU and HAD boxes, except for the W+jets second component and $Z(\nu\nu)$+jets in the HAD Box. For these two backgrounds, the ρ parameters measured from $W(\mu\nu)$+jets events in the MU* box are used. In the alternative MU* lepton treatment the muon is reconstructed as a "neutrino" in that it is excluded from the calculation of R and M_R as if it left no measurable momentum in the detector. This reproduces the kinematic configuration of the W+jets second component and $Z(\nu\nu)$+jets backgrounds, in that in each case there are at least two reconstructed jets recoiling against an invisible, heavy vector boson.

For each background i, we correct the values of a_i and b_i measured in simulated events by these parameters ρ, such that the values of these parameters used in the background prediction is given by

$$a_i = \rho(a)a_i^{MC} \quad , \quad b_i = \rho(b)b_i^{MC} \ . \tag{9.30}$$

These values are used to calculate, for a given R cut, the value of the M_R exponential slope parameter for each background process according to equation (9.29). The background shape prediction for each process i at high M_R is defined by this slope parameter and an additional normalization parameter B_i, such that the background shape of process i can be expressed as

$$f_i^{EXP}(M_R) = B_i e^{S_i M_R}. \tag{9.31}$$

For each process, before the distribution of M_R becomes exponential it has a turn-on region where it peaks at an M_R value, m_i, set by a characteristic scale for that process and the jet p_T and acceptance requirements. We find that this M_R region is well described by an asymmetric normal (Gaussian) distribution, with the widths $\sigma_i^R \neq \sigma_i^L$. We use values of m_i, σ_i^L and σ_i^R measured in simulated events, as a function

of the R cut, to model this part of the background function for the box prediction. The exponentially decaying portion of the background function is attached to the asymmetric Gaussian peak by requiring continuity of the background function and its first derivative in M_R. It is important to emphasize that these asymmetric Gaussian shape parameters derived from simulated events do not contribute to the background predictions in the signal region in that all measurements and fits that dictate this prediction are restricted to ranges where the M_R distribution is well described by an exponential function. Rather, this part of the prediction is included primarily to test for closure of this background prediction method in the low M_R region.

The shape of the M_R distribution for each background process i is fully determined by the parameters σ_i^R, σ_i^L, m_i and S_i (with corresponding systematic errors), except for the W+jets second component processes appearing in each box, whose exponential M_R distributions are given by

$$f_i^{EXP}(M_R) = B_i \left[e^{S^1 i M_R} + f_i e^{S_i^2 M_R} \right] , \tag{9.32}$$

with values f_i describing the relative normalization of the two components that are of the order 10^{-3} to 10^{-2}. We initialize each f_i to the value observed in simulation and ultimately float the values in fits to M_R side-bands in data.

With the shapes of the M_R background determined, the next step in the background prediction is to set the *normalization*, B_i, of each background. In general, the elements entering the normalization of a particular background i can be factorized as

- inclusive cross section, σ_i

- lepton trigger/reconstruction/ID efficiency, ϵ_i^ℓ

- Selection cuts efficiency (R cut), ϵ_i^R

- Integrated luminosity, L

For the EWK and top backgrounds in the lepton boxes, the W, Z and $t\bar{t}$ cross sections measured by CMS are used in normalizations, with corresponding errors.

The W and Z cross section measurements are performed in inclusive (with respect to jet multiplicity) electron and muon final states [176] while the $t\bar{t}$ cross section measurement is performed in the di-lepton channel. In each of these cases, the data-set overlap with the ones used in this analysis is negligible. The measured values of these cross sections are summarized below:

$$\sigma(pp \to WX) \times \mathrm{BF}(W \to \ell\nu) \;=\; 9.951 \pm 0.073 \text{ (stat)} \pm 0.280 \text{ (syst)} \pm 1.095 \text{ (lum) nb}$$

$$\sigma(pp \to ZX) \times \mathrm{BF}(Z \to \ell\ell) \;=\; 0.931 \pm 0.026 \text{ (stat)} \pm 0.023 \text{ (syst)} \pm 0.102 \text{ (lum) nb}$$

$$\sigma(pp \to t\bar{t}) \;=\; 194 \pm 72 \text{ (stat)} \pm 24 \text{ (syst)} \pm 21 \text{ (lum) pb .}$$

For the initial normalization, B_i^0, of a background process i we use values of ϵ_i^ℓ and ϵ_i^R calculated from simulated event samples. ϵ_i^ℓ is multiplied by a DATA/MC correction factor, ρ, measured using the tag-and-probe method with $Z(\ell\ell)$ events in data and simulated events [176]. The variables B_i^0 are expressed as a cross sections, in units of pb. We assign their values by calculating the integral, for a given R cut, of the M_R distribution from where it transitions to an exponential shape to infinity using simulated events, weighted to correspond to the expected yield for 1 pb^{-1} of data and using the above cross sections. This same integral is calculated analytically, as a function of B_i^0, using the background shape function of equation (9.31) or (9.32). We solve for the normalizations B_i^0 by requiring that the two integrals are equal. At this stage, all the parameters describing the shapes and normalizations of the different backgrounds contributing to the ELE and MU boxes predictions are specified.

The total background normalization at this stage is arbitrary, with the relative normalization of different background processes set by measured cross sections and parameters taken from simulation. The next step is to re-normalize the background predictions using an M_R control region in the data. We choose the region 125 GeV $< M_R <$ 175 GeV in the ELE and MU Boxes to measure the event yield, $N_{\mathrm{DATA}}^{\ell\ \mathrm{Box}}$, for different values of the R cut. In this interval the expected QCD multijet contribution is small and the dominant background process is $W(\ell\nu)$+jets. We propagate a

systematic error accounting for possible contamination from Z and $t\bar{t}$ events by evaluating the relative contributions of each process to these event yields, as predicted using the initial MC background normalizations. Errors on the MC efficiencies used in this initial normalization are included. A summary of these numbers is presented, for the MU and ELE Boxes, in Tables 9.6 and 9.7, respectively.

We compare $N_{\mathrm{DATA}}^{\ell\ \mathrm{Box}}$ with the sum of the integrals of each of the functional EWK and top background process predictions in the same interval, denoted $A_0^{\ell\ \mathrm{Box}}$. The ratio of these two numbers, $A_f^{\ell\ \mathrm{Box}}$ is calculated as

$$A_f^{\ell\ \mathrm{Box}} = \frac{N^{\ell\ \mathrm{Box}}}{A_0^{\ell\ \mathrm{Box}}} \pm \frac{\sqrt{N^{\ell\ \mathrm{Box}}}}{A_0^{\ell\ \mathrm{Box}}} \oplus \frac{\sigma(A_0^{\ell\ \mathrm{Box}})}{A_0^{\ell\ \mathrm{Box}}} \frac{N^{\ell\ \mathrm{Box}}}{A_0^{\ell\ \mathrm{Box}}} \ . \tag{9.33}$$

Process		uncertainty
$R > 0.4$	$N^{\mathrm{MU\ Box}} = 1237$	$A_{\mathrm{f}}^{\mathrm{MU\ Box}} = 38.1 \pm 2.4\ \mathrm{pb}^{-1}$
$W(\ell\nu)$+jets MU Box	93%	-
$Z(\ell\ell)$+jets MU Box	1.4%	19%
top+X MU Box	5.2%	40%
$R > 0.45$	$N^{\mathrm{MU\ Box}} = 743$	$A_{\mathrm{f}}^{\mathrm{MU\ Box}} = 41.2 \pm 3.1\ \mathrm{pb}^{-1}$
$W(\ell\nu)$+jets MU Box	92%	-
$Z(\ell\ell)$+jets MU Box	1.3%	19%
top+X MU Box	7.0%	40%
$R > 0.5$	$N^{\mathrm{MU\ Box}} = 389$	$A_{\mathrm{f}}^{\mathrm{MU\ Box}} = 42.0 \pm 3.8\ \mathrm{pb}^{-1}$
$W(\ell\nu)$+jets MU Box	91%	-
$Z(\ell\ell)$+jets MU Box	1.2 %	19%
top+X e Box	7.5%	40%

Table 9.6: Fraction of the contribution to the sum of integrals $A_0^{\mathrm{MU\ Box}}$ for each process, along with the fractional initial normalization. Measured values of $N^{\mathrm{MU\ Box}}$ and $A_{\mathrm{f}}^{\mathrm{MU\ Box}}$ are also listed.

The factor $A_f^{\ell\ \mathrm{Box}}$ has dimensions of integrated luminosity and, given the conventions of our initial normalizations, can be interpreted as measurement of the effective total integrated luminosity of the selected data sample,

$$A_f^{\ell\ \mathrm{Box}} = L \cdot \left(\frac{\epsilon_W^{R,DATA}}{\epsilon_W^{R,MC}} \right) \ , \tag{9.34}$$

Process	expected yields	uncertainty
$R > 0.4$	$N^{\text{ELE Box}} = 1085$	$A_{\text{f}}^{\text{ELE Box}} = 36.3 \pm 1.1 \text{ pb}^{-1}$
$W(\ell\nu)$+jets e Box	94%	-
$Z(\ell\ell)$+jets e Box	0.3%	19%
top+X e Box	5.5%	40%
$R > 0.45$	$N^{\text{ELE Box}} = 596$	$A_{\text{f}}^{\text{ELE Box}} = 36.5 \pm 1.5 \text{ pb}^{-1}$
$W(\ell\nu)$+jets e Box	93%	-
$Z(\ell\ell)$+jets e Box	0.2%	19%
top+X e Box	6.5%	40%
$R > 0.5$	$N^{\text{ELE Box}} = 288$	$A_{\text{f}}^{\text{ELE Box}} = 35.8 \pm 2.1 \text{ pb}^{-1}$
$W(\ell\nu)$+jets e Box	93%	-
$Z(\ell\ell)$+jets e Box	0 %	19%
top+X e Box	7%	40%

Table 9.7: Fraction of the contribution to the sum of integrals $A_0^{\text{ELE Box}}$ for each process, along with the fractional initial uncertainty of that processes' normalization. Measured values of $N^{\text{ELE Box}}$ and $A_{\text{f}}^{\text{ELE Box}}$ are also listed.

where $\epsilon_W^{R,DATA}$ and $\epsilon_W^{R,MC}$ are the R requirement efficiencies for W+jets measured in data and MC simulated events, respectively. The consistency of the measured factors with the CMS measured integrated luminosity of 36.1 pb^{-1}, listed in Tables 9.6 and 9.7, validates the two efficiencies used for the lepton boxes. The unfolding of these efficiencies from the integrated luminosity isn't required in the normalization procedure in that the parameter $A_f^{\ell \text{ Box}}$ is a measurement of the product of the integrated luminosity *and* of the DATA/MC correction factor which is applied to the other EWK and top background predictions.

The final normalization of each of the EWK and top backgrounds in the lepton boxes is calculated as the product of the initial normalization and the factors $A_f^{\ell \text{ Box}}$. At this stage, we have a prediction for the shape and normalization for the entire M_R distribution for each of these backgrounds.

The next step of the background prediction focuses on the low M_R region. The EWK and top background predictions are fixed, in both shape and normalization, with corresponding errors. Using the lepton box QCD multijet shapes measured from the data (section 9.5.2) we constrain the contribution of QCD multijet events to the ELE and MU box event yields by floating the normalization of this background in a

fit to the low M_R region 80 GeV $< M_R <$ 120 GeV. This M_R interval is independent of that used to normalize the EWK and top contributions. In these low M_R fits, the *pdf* is the sum of the QCD multijet and EWK/top predictions, where the shapes have all been fixed to measured values and only the normalization of the QCD multijet component is allowed to vary.

For $R > 0.4$ we find that the amplitude of the QCD component is consistent with zero in both the ELE and MU boxes, which is consistent with MC expectations. This same procedure is validated in closure tests covering the entire background estimation described in section B.9.

Finally, to complete the background prediction we return to the predominantly $W(\ell\nu)$+jets region of the M_R distribution just below the signal region. In a binned likelihood fit in the region $200 < M_R < 400$ GeV we float the parameter f describing the relative normalization between the two $W(\ell\nu)$+jets components while keeping all other background parameters fixed (shapes and normalizations). Because of the M_R interval considered, this fit is independent of the low M_R QCD multijet, and these two steps of the background prediction do not depend on each other.

Using the value of f that maximizes the fit likelihood in the M_R sideband, with corresponding systematic error as returned from the fit, the final background prediction in the lepton boxes is obtained. This is a prediction of the inclusive M_R background distribution, where each background is represented by a function whose shape and normalization have systematic errors corresponding to the various contributing factors that have been discussed. Every component that goes into the final background prediction is either measured directly from data or is inferred from values measured in simulated events multiplied by DATA/MC correction factors. Thus, each of the systematic uncertainties entering the background prediction is extracted using measurements from data. The final background prediction in the MU Box for different values of the R cut is shown in figure 9.33. Similarly, The final background prediction in the ELE Box for different values of the R cut is shown in figure 9.34.

In the lepton box signal regions, expected integrated yields are calculated as the integrals of the functional background predictions, with corresponding systematic

Figure 9.33: Final background predictions for the MU Box with (Left) $R > 0.4$ (Right) $R > 0.5$ (Bottom) $R > 0.45$. The size of the bands around the backgrounds indicates the systematic uncertainties.

uncertainties. The integrated yields for different R and M_R cuts from data, with the predicted yields, are summarized in Tables 9.8 and 9.9 for the ELE and MU Boxes, respectively. We observe agreement between the predicted and observed yields in all cases considered. A summary of the uncertainties entering these background measurements is presented in Table 9.10.
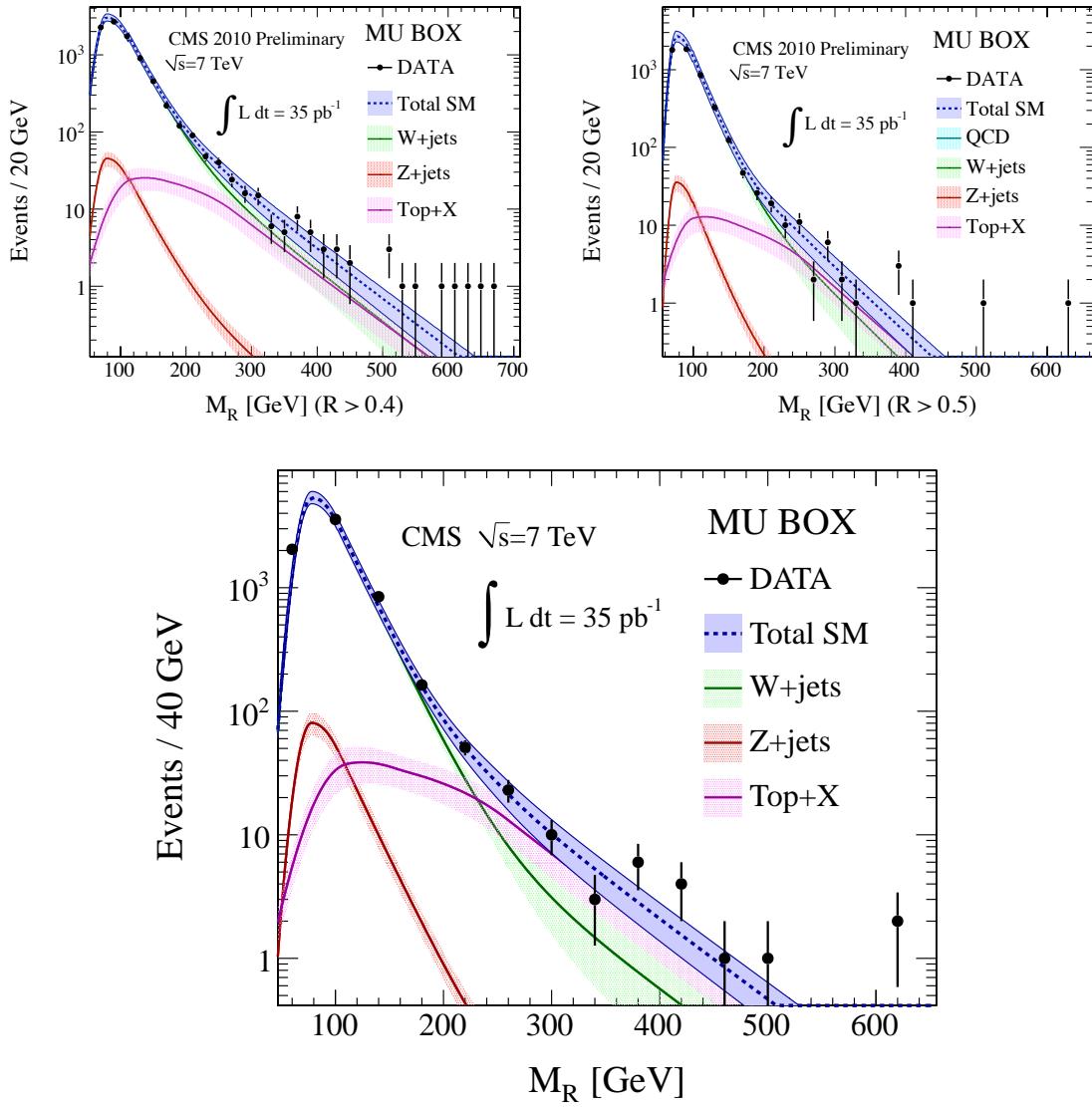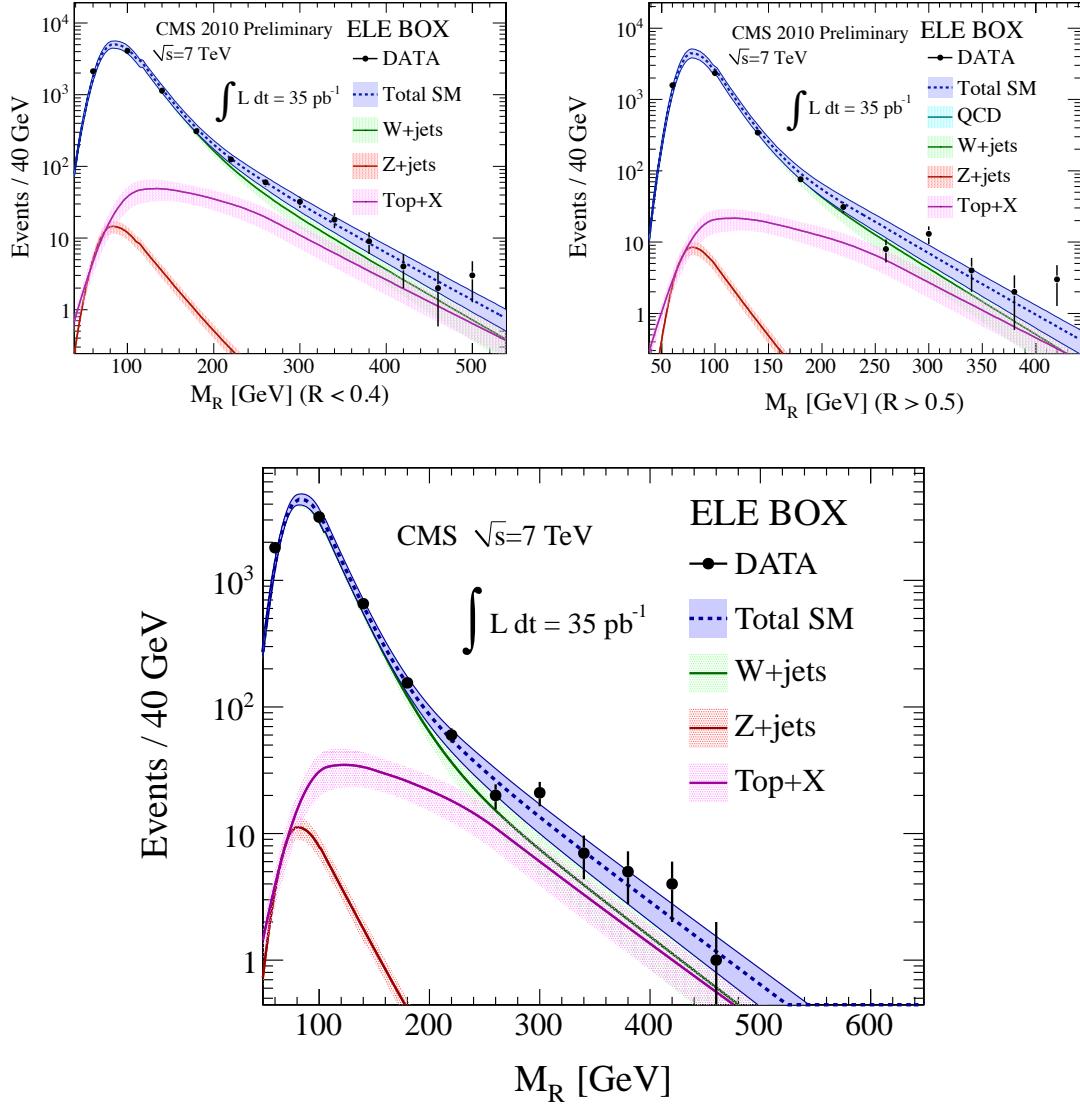

Figure 9.34: Final background predictions for the ELE Box with (Left) $R > 0.4$ (Right) $R > 0.5$ (Bottom) $R > 0.45$. The size of the bands around the backgrounds indicates the systematic uncertainties.

For the interpretation of the results in SUSY parameter space we choose the

R cut $/M_R$ cut	Predicted	Observed
$R > 0.40$ / $M_R > 400$ GeV	10.3 ± 3.1	9
$R > 0.40$ / $M_R > 500$ GeV	2.27 ± 0.78	0
$R > 0.40$ / $M_R > 600$ GeV	0.51 ± 0.19	0
$R > 0.45$ / $M_R > 400$ GeV	3.4 ± 1.1	5
$R > 0.45$ / $M_R > 500$ GeV	0.63 ± 0.23	0
$R > 0.45$ / $M_R > 600$ GeV	0.12 ± 0.05	0
$R > 0.50$ / $M_R > 400$ GeV	1.25 ± 0.46	3
$R > 0.50$ / $M_R > 500$ GeV	0.18 ± 0.07	0
$R > 0.50$ / $M_R > 600$ GeV	0.03 ± 0.01	0

Table 9.8: Predicted and observed yields for ELE Box with different R/M_R cuts.

R cut $/M_R$ cut	Predicted	Observed
$R > 0.40$ / $M_R > 400$ GeV	10.3 ± 3.6	18
$R > 0.40$ / $M_R > 500$ GeV	2.33 ± 0.91	10
$R > 0.40$ / $M_R > 600$ GeV	0.53 ± 0.23	4
$R > 0.45$ / $M_R > 400$ GeV	2.68 ± 0.99	8
$R > 0.45$ / $M_R > 500$ GeV	0.51 ± 0.20	3
$R > 0.45$ / $M_R > 600$ GeV	0.10 ± 0.04	2
$R > 0.50$ / $M_R > 400$ GeV	1.10 ± 0.45	3
$R > 0.50$ / $M_R > 500$ GeV	0.17 ± 0.07	2
$R > 0.50$ / $M_R > 600$ GeV	0.025 ± 0.010	1

Table 9.9: Predicted and observed yields for the MU Box with different R/M_R cuts.

Parameter	Description	Relative Magnitude
Slope parameter a	systematic bias from correlations in fits	5%
Slope parameter b	systematic bias from correlations in fits	10%
Slope parameter a	uncertainty from Monte Carlo	1%-10%
Slope parameter b	uncertainty from Monte Carlo	1%-10%
$\rho(a)^{\text{DATA/MC}}$	measured from DATA	3%
$\rho(b)^{\text{DATA/MC}}$	measured from DATA	3%
Normalization	systematic+statistical component	3%-8%
f	extracted in MLFit (W only)	15%-30%
$W/t\bar{t}$ cross section ratio	CMS measurements (top only)	40%
W/Z cross section ratio	CMS measurements (Z only)	19%

Table 9.10: Summary of non-negligible uncertainties entering the background predictions for the MU and ELE Boxes.

selections which result in optimal expected (without appealing to observed yields) reach in the CMSSM, described in section 9.7. For the ELE and MU boxes, we use

the selection $R > 0.45$ and $M_R > 500$. The SM process-by-process breakdown of the background predictions for these cuts are tabulated in section B.5.

To test the robustness of this background prediction approach against potential biases or oversights we have performed a number dedicated studies. Section B.3 explains studies of the potential for correlations between the lepton identification requirements used in this analysis and the variables M_R and R. We conclude that the lepton identification does not bias the shape of the kinematic variables. Section B.4 evaluates the effect of neglected statistical correlations between fit samples in our determinations of the shape parameters a and b. We assign systematic uncertainties, shown in Tab. 9.10, with magnitudes derived in this study.

In section B.6 we evaluate whether the background predictions shown in Tab. 9.8 and 9.9 are biased and whether the errors we quote actually cover a 68% probability interval. Our studies indicate that the central values we predict for backgrounds are accurate and that the errors have their intended meaning. In general, the predictions for background yields in the M_R/R based are in good agreement with observations. The largest discrepancies appear the the MU box. We find that, when taking into account systematic uncertainties and Poisson sampling statistics, the observed yields are consistent with statistical fluctuations around the predicted mean, within the quoted uncertainties as described in section B.6.1.

Finally, we repeat the full exercise of the background prediction, including measurements in control samples, on simulated event samples generated to have the same yields as what we have observed in data. In this controlled environment, we can test whether the method *closes*, in that it predicts the same background yields that are put in. This simulation closure test is described in section B.9, where we conclude that the method successfully closes.

9.6.2 HAD Box Background Prediction

The procedure for calculating the M_R background prediction in the HAD box is very similar to that for the predictions in the ELE and MU boxes, described in

section 9.6.1. The added complication in the HAD box is that events are selected with and H_T trigger, with nontrivial thresholds ranging from 100 to 150 GeV throughout LHC running and based on HLT jet energies that are not corrected for calorimeter noncompensation. This results in a nontrivial shaping of the M_R distribution for backgrounds in the HAD box since the efficiency for an event to pass these H_T triggers depends strongly on M_R, illustrated for simulated $Z(\nu\nu)$+jets events in the HAD box in figure 9.35. We observe that the effect of this trigger inefficiency only extends up to values of $M_R \sim 400$ GeV, such that it will efficiently select SUSY events, if they are present in the event sample, in the signal regions. On the other hand, we need unfold the effects of these nontrivial H_T trigger requirements in the M_R side-band (M_R values lower than the signal region requirement) if we would like to use this as an additional control region.
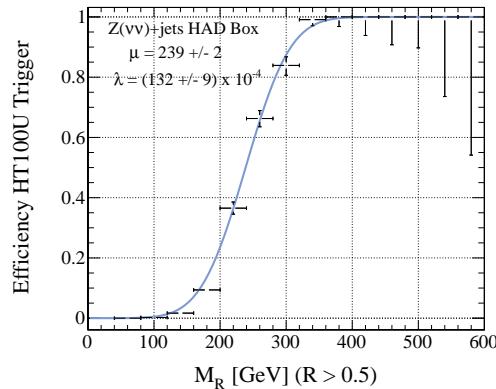
Figure 9.35: HLT_HT100U trigger efficiency as a function of M_R for simulated $Z(\nu\nu)$+jets events that satisfy the HAD box selection and $R > 0.5$.

The procedure for assembling the total SM background predictions in the ELE and MU signal boxes can be summarized as follows:

- Calculate *unbiased* (with respect to the H_T trigger) shapes in M_R for each SM process using M_R scaling parameters a and b and the signal region R requirement. If the parameters a and b were not measured directly in a data control box then we use the values measured from simulated events (section 9.2) corrected with shape scale factors ρ derived from kinematically similar control samples.

- Set the relative normalizations of the EWK and top backgrounds using process cross section measurements from CMS in different final states than the ones considered in this search.

- Set an overall normalization by measuring the event yields in ELE and MU box control regions.

- With the unbiased shapes and normalizations of the EWK and top backgrounds fixed (with the exception of the 2nd W component fraction, which is varied), we determine the normalization of the multijet background and parameters describing the H_T trigger efficiency as a function of M_R (see section B.2) from a maximum likelihood fit to the low M_R control region of the HAD box.

The first step is executed in the same manner as for the lepton boxes (section 9.6.1); each process contributing as background to the HAD box has its shape predicted predicted for the signal region R requirement using the M_R scaling relation described in equation (9.29). Each background is modeled as an exponential, with the slope calculated from the parameters a and b (unique to each background process) from equation (9.29). The parameters a and b are measured in simulated events for each background process contributing to the HAD box (see section 9.2), and these MC derived parameters are multiplied by correction factors ρ, according to equation (9.30) which are derived from kinematically similar event configurations in control regions (see section 9.5). The non-exponential part of the M_R distribution is described by an asymmetric normal distribution, with shape parameters measured from simulated events. The M_R interval used for the final fit in the HAD box is chosen as to minimize any dependence on these shape parameters in the background prediction (the predictions in the high M_R signal regions are not sensitive to these parameters). The initial normalizations of each background process, B_i^0, are calculated in the same way as the analogous factors in the lepton box background predictions.

Because of the nontrivial H_T trigger turn-on, we do not measure the overall background normalization in the HAD Box. We instead use the normalization factors $A_f^{\ell \text{ Box}}$ measured in the ELE and MU boxes. When using these measured normal-

izations in the lepton boxes, we don't need to account for lepton identification efficiencies, since any dependence on them cancels when making predictions within the same box and the M_R shape is insensitive to them (see section B.3). In order to apply these measured normalizations to predictions in the HAD box we must correct for discrepancies in lepton reconstruction and identification between simulation and data.

We measure the muon and electron reconstruction, ID and trigger efficiencies using $Z(\ell\ell)$ events and the tag-and-probe strategy, finding values [176]: $\epsilon^{\mu,meas} = 0.834 \pm 0.010$ and $\epsilon^{e,meas} = 0.753 \pm 0.023$. These measured values are compared with the analogous values measured in simulated $W(\ell\nu)$+jets events: $\epsilon^{\mu,MC} = 0.829$ and $\epsilon^{e,MC} = 0.776$. Using these comparisons, we construct DATA/MC correction factors $\rho(\epsilon^\ell)$ that we use to correct efficiencies measured in simulated events. These measurements describe these efficiencies for the *inclusive* VBTF selection with no requirements on R/M_R. With additional R/M_R requirements these efficiencies change by a few percent due to correlations between the lepton kinematics, the W boson kinematics and the reconstructed values of R/M_R. We use the lepton efficiencies measured in simulated events for the more restrictive regions of phase-space, corrected with $\rho(\epsilon^\ell)$. We also include the difference between the inclusive efficiencies and those from restricted phase space as a systematic error on the efficiency.

The efficiency corrected normalization factors, $A_{corr}^{\ell\ \text{Box}}$, measured in the lepton boxes are summarized in Table 9.11. We recall that these factors represent both a measurement of the integrated luminosity of the data sample, but also a DATA/MC correction factor for the efficiency of the (lepton unrelated) selection requirements. We observe agreement between these measured effective integrated luminosities and expectations, indicating that acceptance efficiencies related to R/M_R requirements are well modeled in the simulation.

Ultimately, the factors from the MU and ELE boxes are combined and used to normalize all of the EWK and top background contributions, with the exception of $Z(\nu\nu)$+jets. For this background, the corresponding normalization measured in the MU* Box is used.

Box / R cut	$A_{corr}^{\ell \text{ Box}}$
MU* Box $R > 0.40$	42.0 ± 4.7 pb^{-1}
MU* Box $R > 0.45$	39.8 ± 5.0 pb^{-1}
MU* Box $R > 0.50$	39.4 ± 5.7 pb^{-1}
MU Box $R > 0.40$	38.1 ± 2.4 pb^{-1}
MU Box $R > 0.45$	40.9 ± 3.3 pb^{-1}
MU Box $R > 0.50$	41.7 ± 4.5 pb^{-1}
ELE Box $R > 0.40$	37.4 ± 1.8 pb^{-1}
ELE Box $R > 0.45$	37.6 ± 2.1 pb^{-1}
ELE Box $R > 0.50$	36.9 ± 2.6 pb^{-1}

Table 9.11: Lepton efficiency-corrected normalizations, $A_{corr}^{\ell \text{ Box}}$, from measurements in the lepton boxes.

We are exploiting the kinematic similarities between backgrounds in different final state boxes, measuring normalizations in the lepton boxes in regions of the razor plane that are nearly identical for the different boxes. As a result, any systematic shortcomings in the Monte Carlo simulation description of R or M_R acceptance efficiencies or values of the exponential slope parameters (which should effect events in the different boxes the same way) are accounted for in this normalization procedure through DATA/MC correction factors.

At this stage, the unbiased M_R shape and normalizations of the EWK and top backgrounds in the HAD box are specified and fixed, with the exception of the parameter f relating the normalization of the W background 2nd component to the 1st. The unbiased multijet background M_R shape is calculated using the appropriate scaling parameters measured in the HAD QCD control box, as described in section 9.5.1. We multiply these unbiased predictions by H_T trigger efficiency curves, $f^{TRIG}(M_R \mid \mu, \lambda)$, with an example shown in figure 9.35. The parameters μ and λ describe the shape of the efficiency curve in our model, and are specific to each background process. A complete description of these trigger efficiency functions can be found in section B.2.

For Runs 2010A and 2010B the HAD box selects events using three different H_T threshold triggers: HLT_HT100U, HLT_HT140U and HLT_HT150U. This means that our trigger efficiency functions need to be integrated-luminosity-weighted averages of the respective curves for the different triggers. The shape parameters μ and λ

are measured in simulated events for each background process in the HAD box. In order to account for differences between data and simulation we introduce additional free parameters μ^{EWK} and λ^{EWK} which are common for all non-multijet background processes and are used to scale the MC derived values of μ and λ.

The total H_T trigger turn-on function which is multiplied with the unbiased M_R shape prediction can be expressed for the ith EWK or top background process as

$$
\begin{aligned}
F_i^{TRIG,EWK}(x \quad | \quad \mu^{EWK}, \lambda^{EWK}, \vec{\mu}_i, \vec{\lambda}_i) & \qquad\qquad (9.35) \\
= \sum_j^{N^{TRIG}} \epsilon^j f^{TRIG}(x \,| \mu^{EWK}, \lambda^{EWK}, \vec{\mu}_i, \vec{\lambda}_i) \ , &
\end{aligned}
$$

where $\vec{\mu}_i$ and $\vec{\lambda}_i$ are the trigger turn-on parameters measured from simulated events, independently for each process. The index j indicates which of the three H_T trigger thresholds is being referenced while the factors ϵ^j indicate what fraction of the HAD box data sample was selected with each trigger.

By introducing the additional parameters μ^{EWK} and λ^{EWK} and allowing them to float in a likelihood fit to the data, we permit our description of these turn-on curves to reflect possible deviations in the behavior of these shapes between data and simulation. For example, DATA/MC discrepancies in the jet energy scale which change the relative scales between uncorrected HLT level jets (used in the H_T triggers) and the reconstructed jets (used in the construction of M_R) would affect all the trigger turn-on functions in a uniform way. These discrepancies would be absorbed into the parameters μ^{EWK} and λ^{EWK}, allowing our modeling of these turn-ons to conform to the data. Similarly, any systematic discrepancy between the estimated integrated luminosity for which each trigger is used and the true value will be absorbed into μ^{EWK}, and will also be reflected in the error on this parameter.

We introduce similar flexibility into our modeling of the QCD trigger turn-on shapes by introducing additional parameters μ^{QCD} and λ^{QCD} and redefining the ensemble of trigger turn-ons for QCD multijet events as

$$F^{TRIG,QCD}(x \mid \mu^{QCD}, \lambda^{QCD}) = \sum_j^{N^{TRIG}} \epsilon^j f^{TRIG}(x \mid (\mathrm{HT}^j/\mathrm{HT}^1)\mu^{QCD}, \lambda^{QCD}) , \quad (9.36)$$

where HT^j denotes the H_T threshold (in units of energy) for the jth trigger. Hence, μ^{QCD} and λ^{QCD} are used to fully describe the shape of the QCD trigger turn-on curves, with the additional constraints that λ is the same for each of the different H_T threshold triggers, with respect to these events, and that the μ parameters for multijet event trigger turn-ons are related by the ratios of H_T thresholds, conclusions supported by observations in simulated QCD events (section B.2).

With all other elements (normalizations and shapes) fixed, the final HAD box background prediction follows from a binned likelihood fit of the total background shape to data in the interval $80 < M_R < 400$ GeV where the parameters μ^{EWK}, λ^{EWK}, μ^{QCD}, λ^{QCD} and a parameter A^{QCD}, which dictates the normalization of the QCD multijet background, are simultaneously floated in the fit. Additionally, the parameter, f, describing the relative normalization between the two $W(\ell\nu)$+jets components, is floated.

The values of these parameters that maximize the likelihood agree with expectations from simulation. In particular, we find that the QCD multijet trigger turn-on parameters, μ^{QCD} and λ^{QCD}, are in agreement with direct observations of the turn-on shape in the QCD control box.

The values of the parameters floated in the fit which correspond to the maximum of the likelihood, along with the errors on these parameters (as calculated from the fit) are used to calculate the final background prediction and its error (including both shape and normalization parameters). This allows for the uncertainty associated with these trigger turn-on curves to be treated in a rigorous and consistent, relative to manner with other uncertainties in the analysis.

The final HAD box background prediction for $R > 0.5$ is shown in figure 9.36. We find agreement between the predicted M_R distribution and the M_R yields observed in data, over the inclusive M_R distribution. In particular, predicted and observed

background yields in the high M_R signal regions are summarized in Table 9.12. A summary of the uncertainties going into these background predictions is listed in Table 9.13.

M_R cut cut	Predicted	Observed
$M_R > 400$ GeV	25.3 ± 5.5	29
$M_R > 500$ GeV	5.5 ± 1.4	7
$M_R > 600$ GeV	1.09 ± 0.32	3

Table 9.12: Predicted and observed yields for different M_R cuts with $R > 0.5$ in the HAD Box.

Parameter	Description	Relative Magnitude
Slope parameter a	systematic bias from correlations in fits	5%
Slope parameter b	systematic bias from correlations in fits	10%
Slope parameter a	uncertainty from Monte Carlo	1%-10%
Slope parameter b	uncertainty from Monte Carlo	1%-10%
$\rho(a)^{\text{DATA/MC}}$	measured from DATA	3%
$\rho(b)^{\text{DATA/MC}}$	measured from DATA	3%
Normalization	systematic + statistical component	8%
Trigger Parameters	systematic from fit toys	2%
f	extracted in MLFit (W only)	13%
$W/t\bar{t}$ cross section ratio	CMS measurements (top only)	40%
W/Z cross section ratio	CMS measurements (Z only)	19%

Table 9.13: Summary of non-negligible uncertainties entering the background predictions for the HAD Box.

We perform an ensemble dedicated studies to check that the HAD box background prediction is unbiased and that the errors are estimated correctly (section B.6), that the method closes for simulated event datasets (section B.9) and that lepton reconstruction and identification requirements do not bias shape and normalization measurements between final state boxes (section B.3).
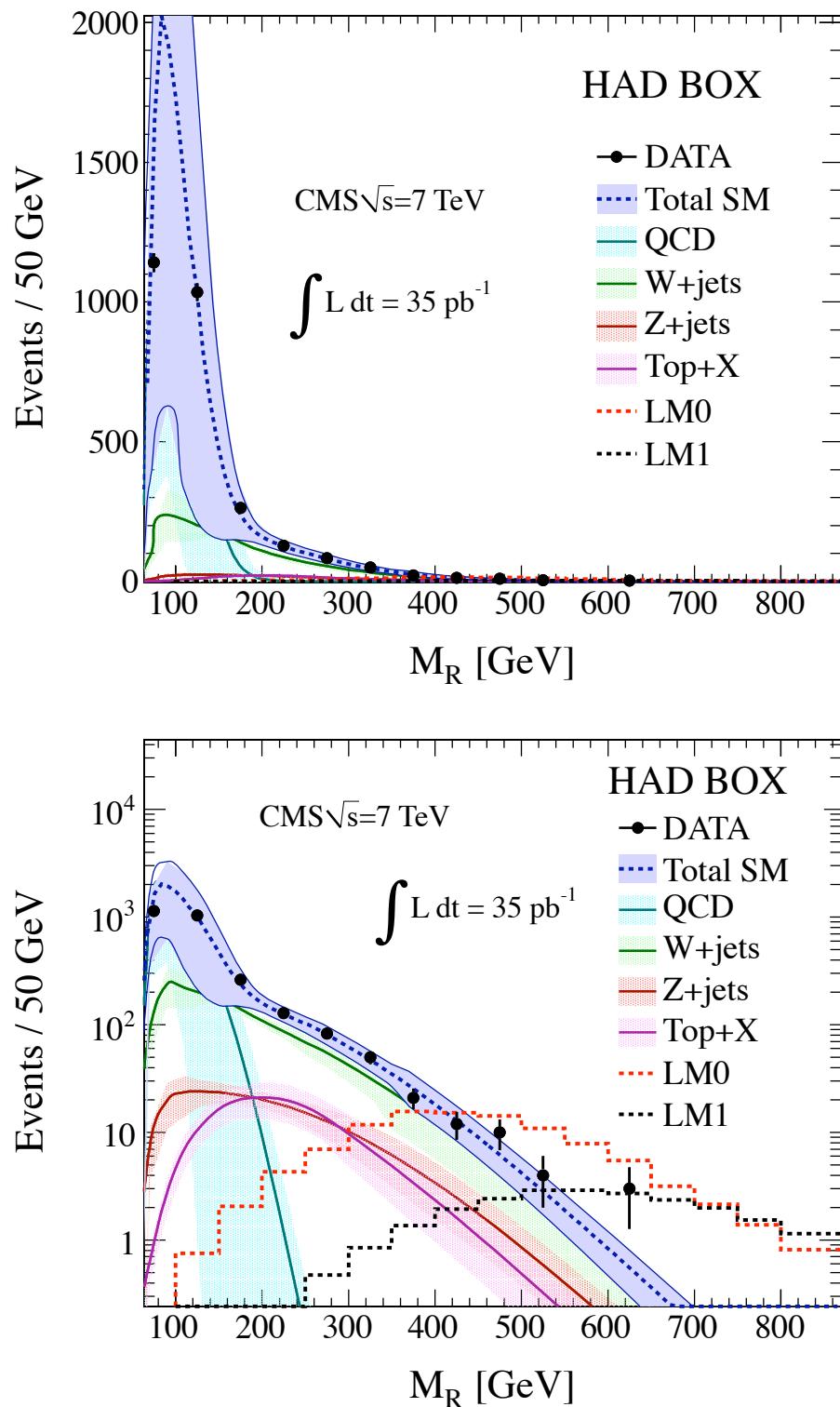
Figure 9.36: Final background prediction for the HAD Box with $R > 0.5$ linear scale (Left) and $R > 0.5$ log scale (Right).

9.7   Results Interpretation in SUSY Parameter Space

The predicted and observed yields from 35 pb^{-1} of data in the ELE, MU and HAD box are summarized in Table 9.14. Since we observe no significant excess of events beyond the SM expectations, we can make inferences about which SUSY models can be excluded by these results in the cases where we would have observed more events than we saw. These results are used to new physics possibilities through the calculation of a model-independent 95% confidence level (C.L.) limit on the number of signal events.

Final state box	M_R / R cut	predicted yield	observed yield
ELE box	500 GeV / 0.45	0.63 ± 0.23	0
MU box	500 GeV / 0.45	0.51 ± 0.20	3
HAD box	500 GeV / 0.45	5.5 ± 1.4	7

Table 9.14: Summary of predicted and observed yields used for SUSY interpretations

The likelihood for the number of observed events n in a particular box is modeled as a Poisson function, given the sum of the signal, s, and the background events, b. A posterior probability density function for the signal yield is derived using Bayes theorem, assuming a flat prior for the signal and a log-normal prior for the background shown in Fig 9.37 for each of the final state boxes. The different posteriors reflect the predictions and observations in each box. The excess of events observed in the MU box results in the posterior peaking at a small, nonzero value.

A 95% probability model independent upper limit is calculated by finding the signal yield s^* which satisfies

$$\int_0^{s^*} P(s)ds = 0.95 \int_0^{\infty} P(s)ds \ , \tag{9.37}$$

where $P(s)$ is the posterior pdf. These 95% probability intervals are indicated by the filled areas in figure 9.37. The observed limit in the HAD box is $s^* < 8.4$ (expected
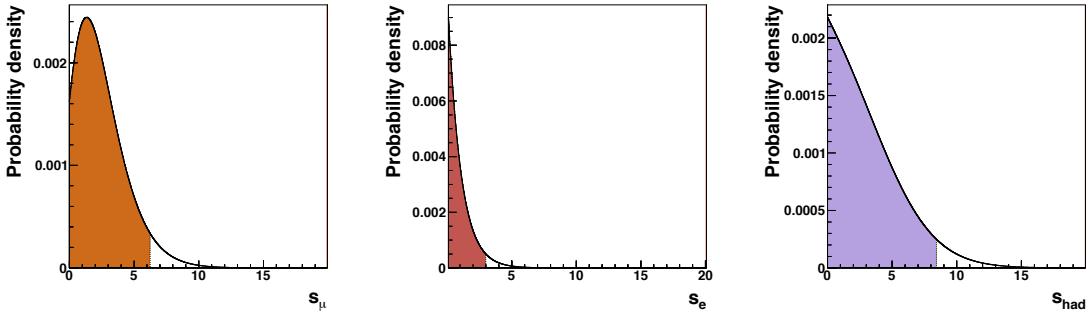
Figure 9.37: Posterior *pdf* for signal yield in the MU (Left), ELE (Center) and HAD (Right) boxes, obtained modeling the likelihood of the observed yield n as a Poisson function $P(n|s+b)$ and using a flat prior for the signal yield s and log-normal prior for the background yield b.

limit 7.2 ± 2.7); in the MU box $s^* < 6.3$ (expected 3.5 ± 1.1); and in the ELE box $s^* < 2.9$ (expected 3.6 ± 1.1). For 10% of the pseudoexperiments in the MU box the expected limit is worse than the observed. The stability of the result was studied against different choices of the prior. In particular, using the reference posterior derived with the methods described in Ref. [177] the observed limits in the HAD, MU and ELE boxes are 8.0, 5.3 and 2.9, respectively.

These results can be interpreted in the context of the CMSSM, which is a simplified subset of the full SUSY parameter space motivated by minimal supergravity scenarios for spontaneous soft supersymmetry breaking. Model points in the CMSSM are specified by five soft breaking parameters: three mass parameters m_0, $m_{1/2}$ and A_0 which are, respectively, universal scalar and gaugino masses and a universal trilinear scalar coupling, as well as $\tan\beta$, the ratio of the up-type and down-type Higgs vacuum expectation values, and the sign of the supersymmetric Higgs mass parameter μ. More details about the CMSSM can be found in Ch. 8. The models realized by scanning over these parameters are widely varied in their superpartner spectra, production channels and decay chains and can produce events in many different final states, although they aren't exhaustively representative of all SUSY possibilities. This means that by interpreting these results in this model framework, we can confront a significant collection of potential signal kinematics and final states. The efficiency for selecting

the events produced in these hypothetical models can vary from ∼1% to 15%, and are shown in section B.10 for each of the signal models considered and each final state selection.

We project the upper limits on s^* in each final state box on the m_0 and $m_{1/2}$ plane comparing by comparing them with the expected signal event yields and excluding any model if $s(m_0, m_{1/2}) > s*$. The systematic uncertainty on the signal yield (coming from the uncertainty on the luminosity, the selection efficiency, and the theoretical uncertainty associated to the cross section calculation) is modeled according to a log-normal prior. The uncertainty on the selection efficiency includes the effect of jet energy scale (JES) corrections (section B.12) corrections, parton distribution function (PDF) uncertainties [178]) (section B.11) and the description of initial-state radiation (ISR). All of these effects are summed in quadrature to calculate the the total systematic error on the signal yield and are summarized in Table 10.4.

box	MU	ELE	HAD
Experiment			
JES	1%	1%	1%
Data/MC ϵ	6%	6%	6%
$\mathcal{L}$ [179]	4%	4%	4%
Theory			
ISR	1%	1%	0.5%
PDF	3%-6%	3%-6%	3%-6%
Total	8%-9%	8%-9%	8%-9%
CMSSM			
NLO σ	16%-18%	16%-18%	16%-18%
Total	17%-19%	17%-19%	17%-19%

Table 9.15: Breakdown and total systematic uncertainties on the signal yield. For the CMSSM scan the NLO cross section uncertainty is included.

The observed limits from the HAD, ELE and MU boxes are shown in figure 9.38 in the CMSSM $(m_0, m_{1/2})$ plane for $\tan\beta = 3$ or $\tan\beta = 10$, $A_0 = 0$, $\text{sgn}(\mu) = +1$, along with the 68% probability band around the expected limits which is obtained by applying the same limit setting procedure described above to an ensemble of background-only pseudoexperiments. The band is calculated around the median of

the expected limit distribution. Observed limits are also interpreted for CMSSM models with $\tan\beta = 50$, $A_0 = 0$, $\text{sgn}(\mu) = +1$, shown in figure 9.39.

We also interpret these NULL search results in the context of simplified models [180–182]. These are SUSY-inspired signal event topologies with only two sparticles in the new physics spectrum: a strongly interacting squark or gluon which will be pair-produced in LHC collisions and a weakly interacting LSP, which the strongly interacting sparticles decay to. Cross section upper limits can be placed on these models directly as a function of the sparticle masses appearing in the spectra. The 95% C.L. cross sections upper limits as a function of the physical masses for two benchmark simplified models (four-flavor squark pair production and gluino pair production) are shown in figure 9.40. In the former, each squark decays to one quark and the LSP, resulting in final states with two jets and missing transverse energy. Similarly, the in the second model gluinos undergo three body decays to two light quarks and the LSP, yielding events with four jets and missing transverse energy.

The qualitative features of these simplified model results reinforce our understanding of the razor variables and the kinematics of these new physics events. The cross section upper limits in figure 9.40 and the selection efficiencies for these models (shown in section B.10) do not depend linearly on $m_{\text{squark/gluino}}$ or m_{LSP} but rather on

$$M_\Delta = \frac{m_{\text{squark/gluino}}^2 - m_{\text{LSP}}^2}{m_{\text{squark/gluino}}} \ , \tag{9.38}$$

due to the fact that the M_R peak position for these events scales with this value.

Figure 9.38: Observed (solid curve) and expected (dot-dashed curve) 95% CL limits in the $(m_0, m_{1/2})$ CMSSM plane with $\tan\beta = 3$ (Left), $\tan\beta = 10$ (Right) and $A_0 = 0$, $\text{sgn}(\mu) = +1$. Results are shown for the HAD box (Top), the MU box (Center) and the ELE box (Bottom). The $\pm$ one standard deviation equivalent variations in the uncertainties are shown as a band around the expected limits.
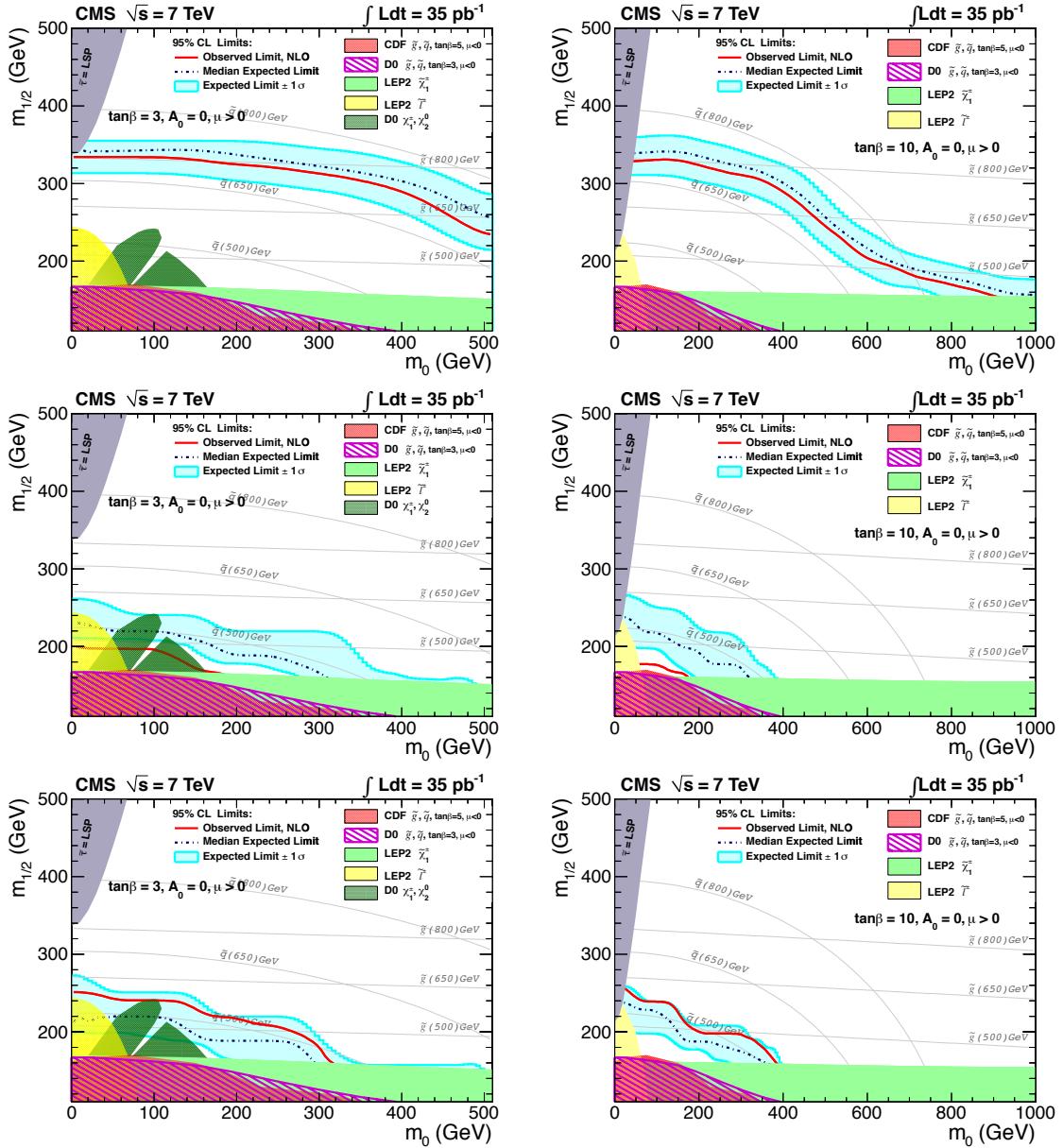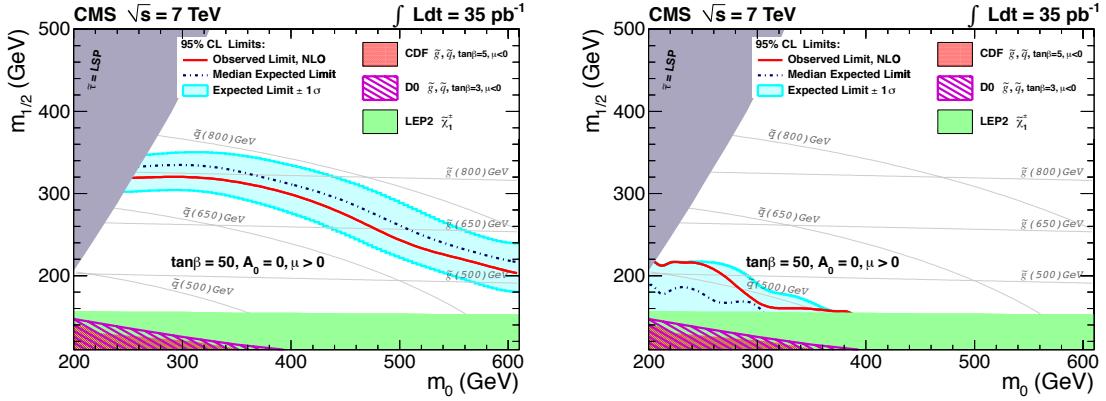
Figure 9.39: Observed (solid curve) and expected (dot-dashed curve) 95% CL limits in the $(m_0, m_{1/2})$ CMSSM plane with $\tan\beta = 50$ and $A_0 = 0$, $\mathrm{sgn}(\mu) = +1$. Results are shown for the HAD box (Left) and the ELE box (Right). The $\pm$ one standard deviation equivalent variations in the uncertainties are shown as a band around the expected limits.
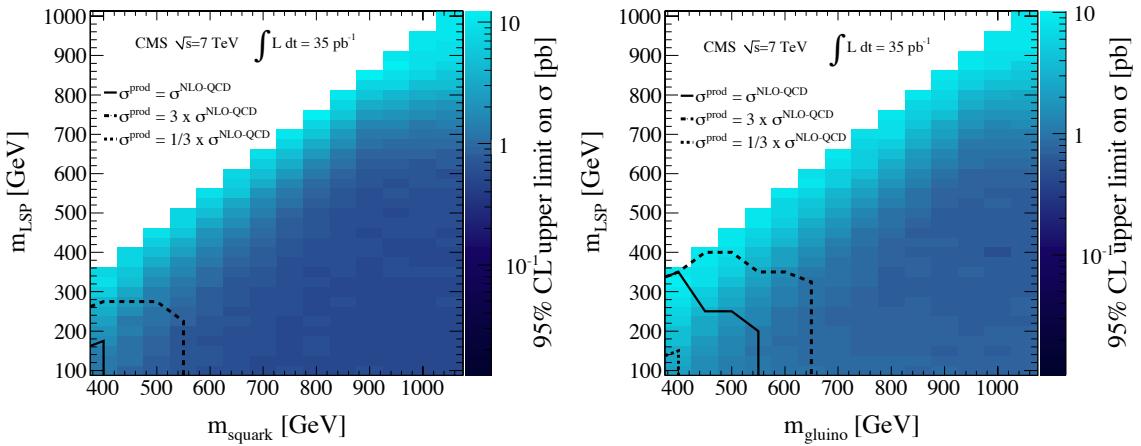
Figure 9.40: Upper limits on two simplified models: di-squark production (Left) resulting in a 2-jet + MET final state and di-gluino (Right) production resulting in a 4-jet + MET final state. The shade scale indicates the value of the cross section excluded at 95% C.L. for each value of m_{LSP} and m_{gluino} or m_{squark}. The solid and dashed contours indicate the 95% C.L. limits assuming the NLO cross section and its variations up and down by a factor of three.

9.8 Conclusion

We performed a search for squarks and gluinos using a data sample of 35 pb^{-1} integrated luminosity from pp collisions at $\sqrt{s} = 7$ TeV, recorded by the CMS detector at the LHC. The search was based on the razor variable, M_R and R, which are used to distinguish between events containing two or more weakly interacting particles resulting from the decays of new, heavy sparticles and the SM backgrounds in final states with jets, missing transverse energy and with and without leptons.

The search relied on predictions of the SM backgrounds determined from data samples dominated by SM processes. No significant excess over the background expectations was observed, and model-independent upper limits on the numbers of signal events were calculated. The results were presented in the $(m_0, m_{1/2})$ CMSSM parameter space. For simplified models the results were given as limits on the production cross sections as a function of the squark, gluino, and LSP masses.

The constraints placed by this analysis on SUSY parameter space demonstrate the strengths of the razor analysis approach; the simple exponential behavior of the various SM backgrounds when described in terms of the razor variables is useful in suppressing these backgrounds and in making reliable estimates from data of the background residuals in the signal regions. Hence, the razor method provides an additional powerful probe in searching for physics beyond the SM at the LHC.

Chapter 10

Searching Through Razor Space

In chapter 9 we described a search for new physics focusing on models with new, heavy particles and conserved $\mathbb{Z}_2$ symmetries, like R-parity and SUSY. Using the kinematic variables M_R and R, event yields in hadronic and single lepton final states were used to place constraints on new physics parameters and particle masses that were, in many cases, the most restrictive ever. But there are a number of ways that this search could be improved. In the following chapter, we describe a *new* search for new physics based on adaptations of the original razor variables using the CMS detector with 5 fb^{-1} of $\sqrt{s} = 7$ TeV pp collision data. With new variables that avoid ill-defined event configurations and an increase in kinematic phase-space and the number of final states the search considers, there is a large increase in sensitivity relative to the previous search. In chapter 10, we describe the search through the two dimensional M_R/R^2 plane in final states with zero, one or two leptons, with and without b-quark tagged jets. The results of this search are interpreted in R-parity conserving SUSY scenarios, along with a collection of $\mathbb{Z}_2$ symmetry-inspired simplified models, including cases featuring the production and decays of new-symmetry partners of SM tops, like stops in SUSY.

10.1 A Better $\mathrm{M_R}$

The kinematic variable M_R, derived in section 9.1.1, satisfies a number of useful properties. It is sensitive to the mass scale of new physics events and is invariant under

boosts along the beam axis, an essential feature at a hadron collider. Unfortunately it is not always well-defined. For some event configurations the longitudinal boost which takes particles from the lab frame to the R-frame, β_R, is greater than one, corresponding to an R-frame that is traveling faster than the speed of light. This can occur when the approximations used to derive M_R lead to unphysical cases. In this section we derive a new variable which shares all of M_R's useful properties, but that is always well-defined.

To see how this is accomplished, we return to the simple scenario of the production of two, identical particles of type S, each decaying to a visible SM particle Q and a weakly interacting particle χ, such that each decay chain has an identical particle content. This type of event is illustrated in figure 10.1.
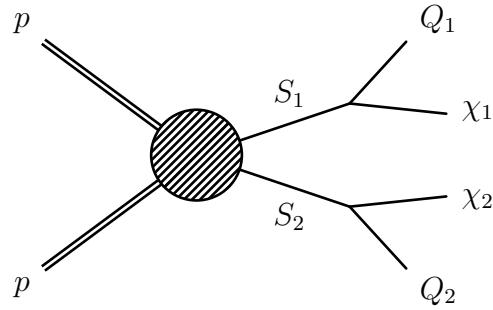

Figure 10.1: Canonical SUSY production scenario. Two massive particles, S_i, are produced in a pp collision and each decay to a SM particle Q_i, and a weakly interacting particle, χ_i.

Firstly, we identify the kinematical characteristics that are associated with events with $|\beta_R| \geq 1$. Using the notation of section 9.1.1, we recall that the decay angles in the S_i rest frames were denoted by unit vectors $\hat{u}_1$ and $\hat{u}_2$. It is these directions, relative to the S particles' axis of motion in the CM frame, $\vec{\beta}_{CM}$, that dictate whether $|\beta_R| < 1$. Setting $\gamma_{CM} = 1.1$, we scan over values for the unit vectors $\hat{u}_1$, $\hat{u}_2$ and $\hat{\beta}_{CM}$ in a toy simulation, noting for which values and with what frequency we find $|\beta_R| \geq 1$. In figure 10.2 we show the correlation between the normalized z-components of momenta of Q_1 and Q_2 in the rest frames of their respective parents S_i for events where the R-frame is ill-defined. We find, as perhaps one could infer from the expression of $\beta_R = \frac{|\vec{q}_1{}^l - \vec{q}_2{}^l|}{q_{1z}^l - q_{2z}^l}$, that these longitudinal momentum components tend to be equal in

both direction and magnitude. In fact, as γ_{CM} tends toward one, the distribution shown in figure 10.2 (left) tends toward a discrete line along the $\hat{u}_1 \cdot \hat{z} = \hat{u}_2 \cdot \hat{z}$ diagonal. Similarly, in figure 10.2 (right) we see the correlation of the difference in azimuthal angles between the momenta of Q_1 and Q_2 and the momenta of Q_1 and $\vec{\beta}_{CM}$. We find that events with $|\beta_R| \geq 1$ tend to have $\hat{u}_1$ and $\hat{u}_2$ pointing in the same direction in the transverse plane, with $\vec{\beta}_{CM}$ pointing in either the same or opposite direction.
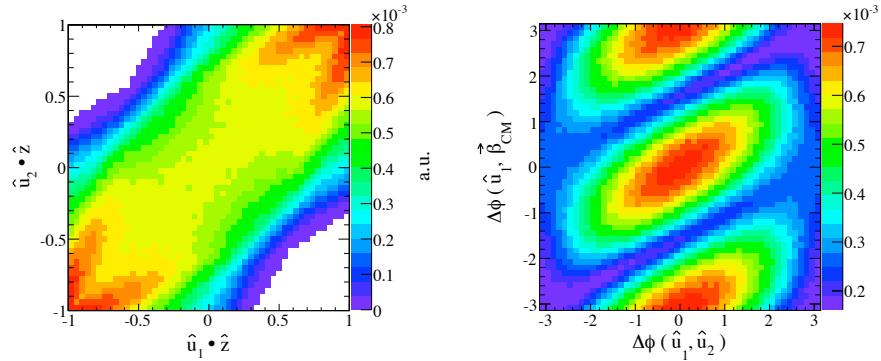

Figure 10.2: (Left) Correlation between $\hat{u}_1 \cdot \hat{z}$ and $\hat{u}_2 \cdot \hat{z}$ for events with $\gamma_{CM} = 1.1$ and $|\beta_R| \geq 1$. (Right) Correlation between $\Delta\phi(\hat{u}_1, \hat{u}_2)$ and $\Delta\phi(\hat{u}_1, \vec{\beta}_{CM})$ for events with $\gamma_{CM} = 1.1$ and $|\beta_R| \geq 1$. Distributions are normalized to unit volume.

These observations indicate that the cases where the R-frame is ill-defined result from the neglecting of the transverse component of $\vec{\beta}_{CM}$ in the approximations made in the derivation of M_R. We recall that in the R-frame approximation we assumed that $\vec{\beta}_{CM} \to 0$, and when the transverse components of $\vec{\beta}_{CM}$ are large, and point along the S_i decay axes, M_R can become imaginary. In order to derive a new variable that is always well defined we relax this assumption, instead assuming only that $\vec{\beta}_{CM} \cdot \hat{z} \to 0$. Now, there are two unknown boosts relating the lab frame and our rough approximations of the S_i rest frames: a longitudinal boost, β_{L^*}, which moves from the lab frame to an approximation of the CM frame (R-frame) and a transverse boost, $\vec{\beta}_T^{R^*}$, which acts in equal and opposite directions on the two particles Q_i to take them from the approximate CM frame to their respective S_i rest frames (R^*-frames). The series of Lorentz boosts taking Q_1 and Q_2 from the laboratory frame to their

respective R^*-frames can be summarized as:

$$q_1^l \xrightarrow{\beta_{L^*}} q_1^R \xrightarrow{\vec{\beta}_T^{R^*}} q_1^{R^*} \ ,$$
$$q_2^l \xrightarrow{\beta_{L^*}} q_2^R \xrightarrow{-\vec{\beta}_T^{R^*}} q_2^{R^*} \ . \tag{10.1}$$

There are three d.o.f. associated with these transformations but only one constraint: Since the decay chains have identical particle content the momenta of Q_1 and Q_2 must have the same value in their respective R^*-frames. This can be re-expressed as a constraint equation on the variables β_{L^*} and $\vec{\beta}_T^{R^*}$,

$$\gamma_{L^*}(|\vec{q}_1^{\ l}| - |\vec{q}_2^{\ l}|) - \gamma_{L^*}\beta_{L^*}(q_{1z}^l - q_{2z}^l) = \vec{\beta}_T^{R^*} \cdot (\vec{q}_{1T}^l + \vec{q}_{2T}^l) \ , \tag{10.2}$$

which can be used to solve for the magnitude of $\vec{\beta}_T^{R^*}$, $\beta_T^{R^*} \equiv |\vec{\beta}_T^{R^*}|$, in terms of $\hat{\beta}_T^{R^*}$ and β_{L^*},

$$\beta_T^{R^*} = \frac{\gamma_{L^*}(|\vec{q}_1^{\ l}| - |\vec{q}_2^{\ l}|) - \gamma_{L^*}\beta_{L^*}(q_{1z}^l - q_{2z}^l)}{\hat{\beta}_T^{R^*} \cdot (\vec{q}_{1T}^l + \vec{q}_{2T}^l)} \ . \tag{10.3}$$
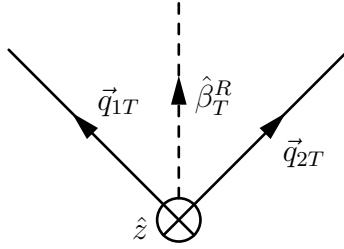
Just as we did in the R-frame, we will define the R^*-frame mass, M_{R^*}, as two times the magnitude of the momentum of Q_1 in its respective R^*-frame. M_{R^*} can be expressed as

$$
\begin{aligned}
M_{R^*} &\equiv 2|\vec{q}_1^{\ R^*}| = 2|\vec{q}_2^{\ R^*}| \\
&= \frac{2\gamma_{L^*}\hat{\beta}_T^{R^*} \cdot \left[(|\vec{q}_1^{\ l}|\vec{q}_{2T}^l + |\vec{q}_2^{\ l}|\vec{q}_{1T}^l) - \beta_{L^*}(q_{1z}^l\vec{q}_{2T}^l + q_{2z}^l\vec{q}_{1T}^l)\right]}{\sqrt{|\hat{\beta}_T^{R^*} \cdot (\vec{q}_{1T}^l + \vec{q}_{2T}^l)|^2 - \gamma_{L^*}\left[|\vec{q}_1^{\ l}| - |\vec{q}_2^{\ l}| - \beta_{L^*}(q_{1z}^l - q_{2z}^l)\right]^2}} \ .
\end{aligned} \tag{10.4}
$$

In order to calculate M_{R^*}, we must specify values of β_{L^*} and $\hat{\beta}_T^{R^*}$. We have allowed γ_{R*} to be nonzero but the considerations that led to the $\gamma_{R*} \to 1$ approximation still hold; instead of setting the under constrained d.o.f. to zero as we did for M_R we will minimize it away. We find that the choice of $\hat{\beta}_T^{R^*}$ which minimizes γ_{R*} is given by

$$\hat{\beta}_T^{R^*} = \frac{\vec{q}_{1T}^l + \vec{q}_{2T}^l}{|\vec{q}_{1T}^l + \vec{q}_{2T}^l|} \ , \tag{10.5}$$

illustrated in figure 10.3.

$$\hat{\beta}_T^R \equiv \frac{\vec{q}_{1T} + \vec{q}_{2T}}{|\vec{q}_{1T} + \vec{q}_{2T}|}$$

Figure 10.3: The particles' Q_1 and Q_2 momenta in the plane transverse to the beam axis. The direction of $\hat{\beta}_T^{R^*}$ which minimizes γ_{R*} points in the direction $\vec{q}_{1T}^{\,l} + \vec{q}_{2T}^{\,l}$.

For the remaining d.o.f. corresponding to β_{L*} we will take a similar approach, for slightly different reasons. We would like our final kinematic variable expressions (for M_{R*} and γ_{R*}) to be invariant under boosts along the beam axis. This property will be guaranteed by choosing the β_{L*} which satisfies $\frac{\partial M_{R*}}{\partial \beta_{L*}} = 0$.

With each unknown quantity now specified, M_{R*} can be expressed, event-by-event, as

$$M_{R*} = \sqrt{(|\vec{q}_1^{\,l}| + |\vec{q}_2^{\,l}|)^2 - (q_{1z}^l + q_{2z}^l)^2 - \frac{(|\vec{q}_{1T}^{\,l}|^2 - |\vec{q}_{2T}^{\,l}|^2)^2}{|\vec{q}_{1T}^{\,l} + \vec{q}_{2T}^{\,l}|^2}} \; . \tag{10.6}$$

$\gamma_{R*} = (1 - |\vec{\beta}_T^{R^*}|^2)^{-1/2}$ is given by

$$\gamma_{R*} = \sqrt{\frac{(|\vec{q}_1^{\,l}| + |\vec{q}_2^{\,l}|)^2 - (q_{1z}^l + q_{2z}^l)^2}{(|\vec{q}_1^{\,l}| + |\vec{q}_2^{\,l}|)^2 - (q_{1z}^l + q_{2z}^l)^2 - \frac{(|\vec{q}_{1T}^{\,l}|^2 - |\vec{q}_{2T}^{\,l}|^2)^2}{|\vec{q}_{1T}^{\,l} + \vec{q}_{2T}^{\,l}|^2}}} \; . \tag{10.7}$$

As is the case for M_R, M_{R*} is invariant under longitudinal boosts, as is γ_{R*}. Analogous to the R-frame, we define the R^*-frame razor, R^*, as the ratio of M_T^R and M_{R*}, with M_T^R given by equation (9.13).

To understand how the distribution of M_{R*} changes with γ_{CM}, we numerically integrate over all the decay angles assuming isotropic decays. The resulting M_{R*} and $\gamma_{R*} M_{R*}$ distributions are shown in figure 10.4. We find that the peak value of the

M_{R^*} distribution is at approximately M_Δ, regardless of γ_{CM}, while $\gamma_{R^*} M_{R^*}$ peaks at $\gamma_{CM} M_\Delta$.
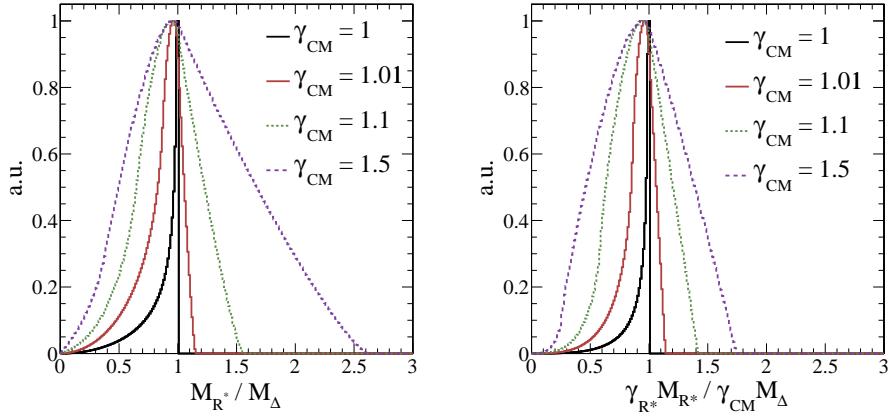

Figure 10.4: Distribution of M_{R^*} (Left) and $\gamma_{R^*} M_{R^*}$ (Right) for different values of γ_{CM}. Distributions are normalized such that their maximum value is equal to one.

Comparing figure 10.4 and figure 9.4, we see that the peak position of the $\gamma_{R^*} M_{R^*}$ distribution scales like the peak of the M_R distribution. M_R is a variable most useful for treating the case $\gamma_{CM} = 1$ which, in practice, is kinematically forbidden. The quantity $\gamma_{R^*} M_{R^*}$ reproduces the same peaking behavior, without ill-defined configurations and better resolution on the quantity $\gamma_{CM} M_\Delta$.

In fact, the variables M_{R^*}, $\gamma_{R^*} M_{R^*}$ and M_R share many properties. We consider two of the examples from section 9.1.4, now in the context of M_{R^*} and γ_{R^*}. The first scenario is of two massive particles, S_1 and S_2, with different masses decaying each to a visible particle and potentially massive weakly interacting particle, such that $M_\Delta^2 = M_\Delta^1 (1 + \delta) = M_\Delta (1 + \delta)$. Assuming $\gamma_{CM} = 1$, and numerically integrating over the decay angles, we calculate M_{R^*} as a function of δ, with the resulting distributions shown in figure 10.5 (left). We observe that M_{R^*}, like M_R, has a peak whose position scales with $\sqrt{1 + \delta}$.

The second example from section 9.1.4 involves two particles S_1 and S_2, with the same mass. S_1 undergoes a two-body decay to a visible particle, Q_3, and another particle, G_1, with mass $M_G = M_S (1 - \delta)$. The particles S_2 and G_1 then each decay to a weakly interacting particle and a visible particle, where the mass of the weakly
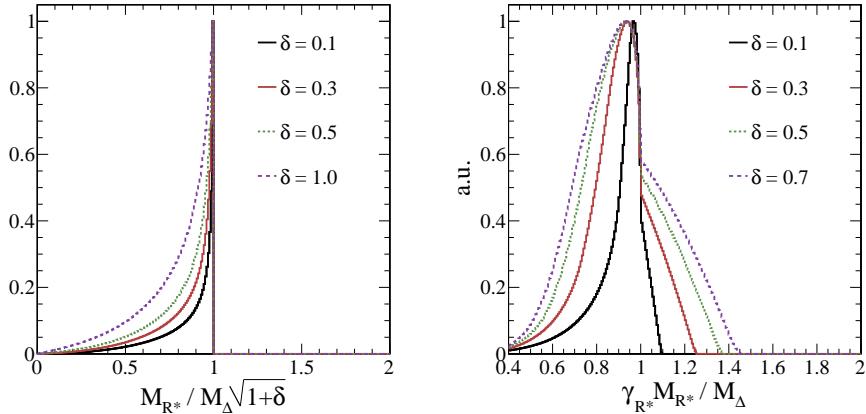
Figure 10.5: (Left) Distribution of M_{R^*}, in units of $M_\Delta \sqrt{1+\delta}$, for different values of δ. (Right) Distribution of $\gamma_{R^*} M_{R^*}$ when one of the particles S_i decays to an intermediate particle G_i with mass $M_G = M_S(1-\delta)$, for different values of δ. Distributions are normalized such that their maximum value is equal to one.

interacting particles is M_χ. The numerically integrated $\gamma_{R^*} M_{R^*}$ distributions, for different values of δ, are shown in figure 10.5 (right). We observe that, like M_R, the quantity $\gamma_{R^*} M_{R^*}$ peaks at $M_\Delta = \frac{M_S^2 - M_\chi^2}{M_S}$, regardless of the value of δ.

In general, we see that the quantity $\gamma_{R^*} M_{R^*}$ behaves almost identically to M_R. This correspondence also hold for SM background processes. This fact is illustrated in figure 10.6 with early 2011 CMS data, comparing the M_R and $\gamma_{R^*} M_{R^*}$ scaling behavior in QCD dijet events. We see that $\gamma_{R^*} M_{R^*}$ exhibits M_R/R^2 scaling qualitatively and qualitatively identical to M_R. The construction $\gamma_{R^*} M_{R^*}$ shares all the useful of properties of M_R and is guaranteed to be well-defined. In fact, for realistic γ_{CM} distributions, $\gamma_{R^*} M_{R^*}$ is a better indicator of the scale M_Δ than M_R.

Having found suitable replacements for the variables M_R and R in $\gamma_{R^*} M_{R^*}$ and R^* we take the notational liberty to *retire* the original expressions for these variables and replace them with those of $\gamma_{R^*} M_{R^*}$ and R^*. Having been superseded by superior concepts, the original M_R and R will not be discussed again; may they rest in peace.

$$
\begin{aligned}
M_R &\equiv \gamma_{R*} M_{R*} = \sqrt{(|\vec{q}_1^{\,l}| + |\vec{q}_1^{\,l}|)^2 - (q_{z1}^l + q_{z2}^l)^2} \\
R &\equiv R^* = \frac{M_T^R}{\gamma_{R*} M_{R*}} \ .
\end{aligned}
\tag{10.8}
$$

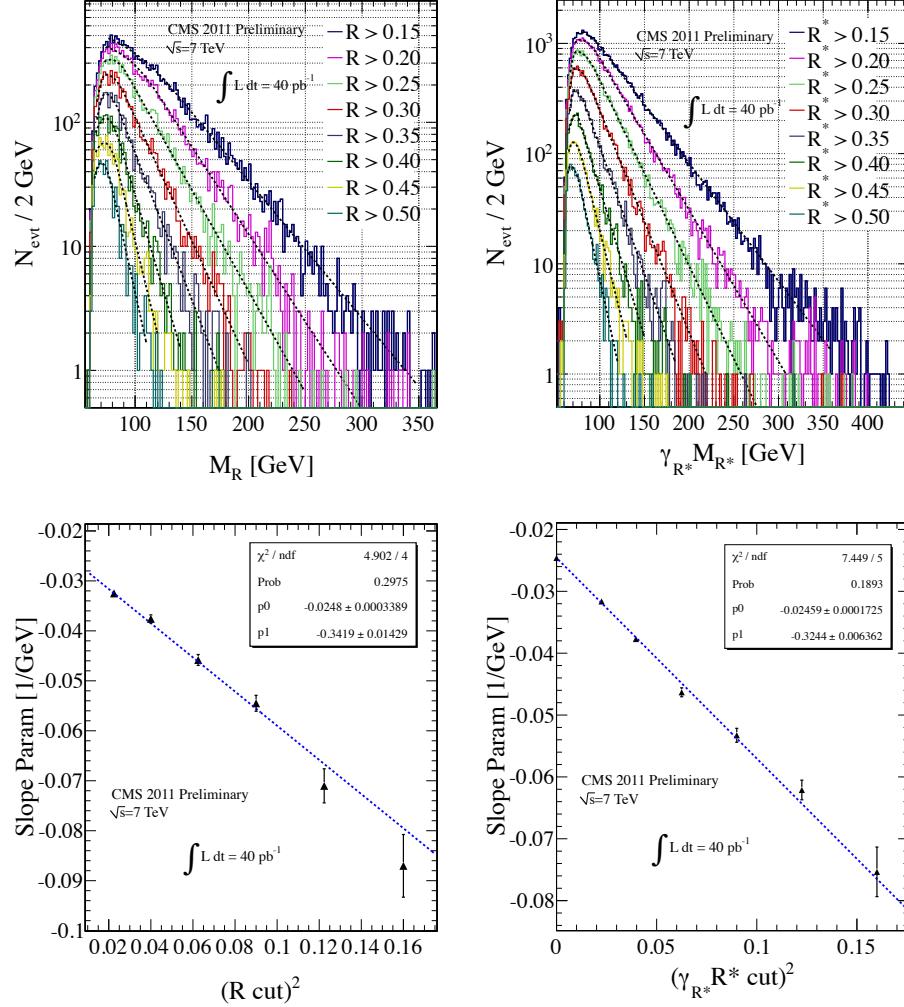Figure 10.6: Data selected with prescaled, low threshold dijet triggers in the all-hadronic final state from early 2011 $\sqrt{\hat{s}} = 7$ TeV LHC running. (Top) $M_R(\gamma_{R*}M_{R*})$ distribution as a function of $R(R^*)$ cut. (Bottom) Exponential slope of the $M_R(\gamma_{R*}M_{R*})$ as function of $R(R^*)$ cut. The quantitative and qualitative behavior of the variables M_R and $\gamma_{R*}M_{R*}$ are the same.

10.2   Phenomenology of the 2D Razor Plane

The razor variables M_R and R provide a distinctive description of SM background and SUSY signal events which allows us to identify a preferred region of kinematic phase space where we can search for signs of new physics. This kinematic separation is illustrated in the two dimensional M_R/R^2 distributions for simulated signal and background events shown in figure. 10.7.

The variable M_R is sensitive to the scale of each of these events and can distinguish SUSY events containing new sparticles with large masses from background events containing only lighter SM particles. Backgrounds like QCD multijets, whose events don't contain particles with large masses, have a preferred scale set by the convolution of minimal object momentum requirements in the selection of events (described in section 10.4.2) and steeply falling parton luminosities. R-parity conserving SUSY events, like the example of LM6 illustrated in figure 10.7 which features predominantly gluino pair production, have a peaking M_R distribution indicating the mass difference between the heavy pair-produced parent particles and escaping LSPs. The variable R is sensitive to the presence of weakly interacting particles, having larger values for events with at least two weakly interacting particles following from the decay of different particles, like our intended signal events. In the 2D razor plane, we expect the SM backgrounds to be restricted largely to the low M_R/low R region, while events with the pair production of heavy particles each decaying to visible and weakly interacting particles, our desired signal, will populate the high M_R/low R region.

For a given process, the events which populate the region of the razor plane above the characteristic scale of that process (set by the masses of particles in these events, or the center of mass energy threshold above which these events have a maximum production rate) follow hyperbolic constant-yield contours in the variables M_R and R. The reason for this can be understood in the construction of the razor variables. For a fixed center of mass (CM) energy in a hard collision, the energy of the event is shared between detected and undetected particles. The variable M_R will increase
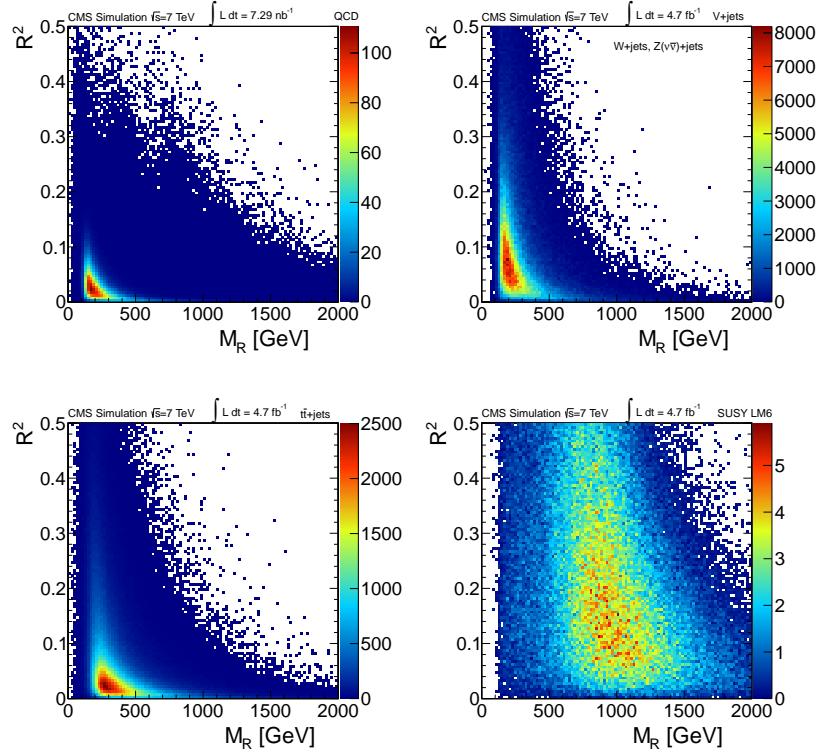
Figure 10.7: Scatter plot in the (R^2, M_R) plane for simulated events: (Top left) multijets, (Top right) W+jets and $Z(\nu\bar{\nu}$+jets, (Bottom left) $t\bar{t}$+jets, and (Bottom right) the SUSY benchmark model LM6 [173] with $M_\Delta = 831$ GeV. The yields are normalized to an integrated luminosity of ~ 4.7 fb^{-1}, except for the multijet background where the integrated luminosity of the generated sample is used. The bin size is (20 GeV $\times$ 0.005).

in value if more energy is contained in detected particles, which are included in the event mega-jets. On the other hand, R reflects the ratio of undetected to detected momentum in the event; its value will decrease with more detectable energy and increase with less. The product of M_R and R^2 is approximately constant for a fixed CM energy while the total energy can be shared differently between the two variables. The steeply falling distribution of CM energy, once it has exceeded the threshold value for a given process, results in the yields of each of these hyperbolic contours falling steeply towards the upper right hand corner of the razor plane.

Similarly, the interplay between these two variables can be used to suppress background events with spurious instrumental effects. Backgrounds with enormous rela-

tive cross sections, like QCD multijet production, can be especially pernicious since the large event yields give opportunities for rare and dramatic instances of detector noise and mis-reconstruction. For example, a large under-measurement of a single jet in a multijet event could make it incorrectly appear as if it has a large missing momentum component, and the tail of a missing transverse energy distribution will be populated with similar events. For the razor variables, the same concept of sharing a fixed CM energy between detected and undetected particles in the event applies here, accept now instead of weakly interacting particles, energy now escapes the event due to mis-measurements. An event containing an object whose energy is measured to be artificially small will be measured to have larger values of R, since the event will appear to have missing momentum, but will also decrease the measured value of M_R. With respect to over-measurements of objects' energies, the fact that the object resolution functions fall more steeply than the parton luminosity distributions with increasing over-measurement ensures that events at large values of M_R *and* R are made up predominantly of events with truly large CM energies, rather than mis-reconstructed examples of large backgrounds, which are resigned to the lower left corner of the razor plane

At larger values of R and M_R in excess of the SM particle masses, background events are comprised of processes with genuine missing transverse energy resulting from the decay of W and Z bosons to neutrinos, with CM energies in excess of the boson masses. In this hyperbolic regime of the razor plane we find that we can analytically model the shapes of each of the SM backgrounds by exploiting empirically observed M_R/R^2 scaling.

10.2.1   2D M_R/R^2 Scaling

The one dimensional consequence of the hyperbolic two dimensional correlation between M_R and R^2 illustrated in figure 10.7 for the variable M_R is that, for increasing values of R, the M_R distribution will fall more steeply. In section 9.2, we observed that this scaling behavior obeys a simple analytic model, which is demonstrated in a

sample multijet events selected from 2011 $\sqrt{s} = 7$ TeV CMS data shown in figure 10.8 (see section 10.4.2 and C.1 for details about the reconstruction and selection of this sample).
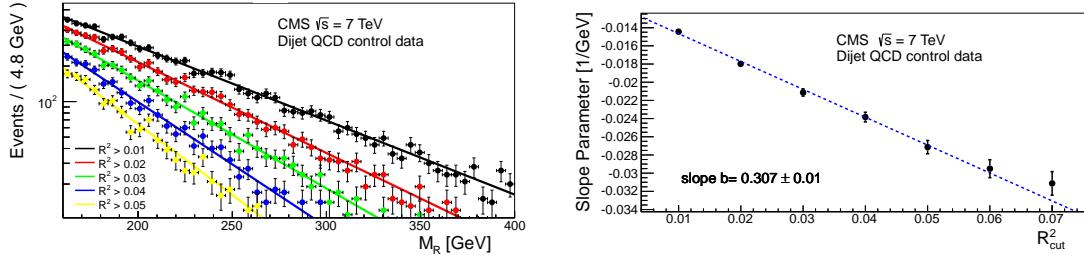

Figure 10.8: (Left) M_R distribution for different values of the R^2_{cut} for events in data selected in the HAD box with low threshold, prescaled jet triggers. Each distribution is fit with an exponential function. (Right) The exponential slope parameter, S, from a fit to the M_R distribution as a function of R^2_{cut}.

These events are selected with low-threshold jet triggers which ensures that, in the region of the razor plane illustrated in figure 10.8, they are comprised almost exclusively of multijet events because of relative production rates at low CM energy. With this single process isolated, we observe that the M_R distribution (integrated above some value of R) is well described by a single exponential function, e^{-SM_R}, where S is the exponential slope. Furthermore, we observe that the slope S exhibits a linear dependence on the value of the cut on R^2_{cut},

$$S = a + bR^2_{cut} \ . \tag{10.9}$$

This same hyperbolic behavior of backgrounds suggests that R^2 should behave similarly when applying cuts on M_R. In figure 10.9 we consider the same sample of multijet events selected in data, except now look at the R^2 distribution for different values of a cut on M_R. We observe that the R^2 distribution, like M_R, is well described by an exponential function and that the exponential slope scales linearly with M^{cut}_R,

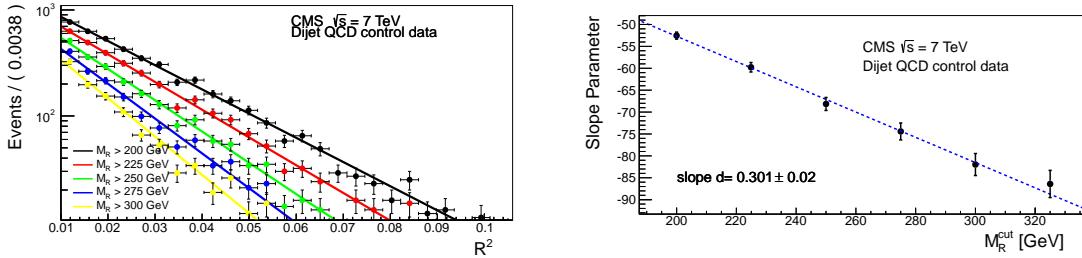$$S = c + dM^{cut}_R \ . \tag{10.10}$$

Figure 10.9: (Left) R^2 distribution for different values of the M_R^{cut} for events in data selected in the HAD box with low threshold, prescaled jet triggers. Each distribution is fit with an exponential function. (Right) the exponential slope parameter, S, from a fit to the R^2 distribution as a function of M_R^{cut}.

From linear fits to the measured values of M_R and R^2 slopes with increasing R_{cut}^2 and M_R^{cut} (figure 10.8 and 10.9 right, respectively) we find that the measured values of the scaling parameters b (from the M_R view) and d (from the R^2 view) are consistent. In fact, we note that in order for this event sample to follow the scaling behavior of equations (10.9) and (10.10) these two parameters must be the same, and the most general two-dimensional function of these variables that exhibits this scaling behavior is given by

$$F_j(M_R, R^2) = \left[k_j(M_R - M_{R,j}^0)(R^2 - R_{0,j}^2) - 1 \right] \times e^{-k_j(M_R - M_{R,j}^0)(R^2 - R_{0,j}^2)} , \quad (10.11)$$

where j indicates the background process the function describes, $k_j = b_j = d_j$ from the one-dimensional exponential views of the variables and $M_{R,j}^0$ and $R_{0,j}^2$ are constants specific to the process. Integrating this function over either R^2 or M_R, above a fixed value, recovers the one-dimensional exponential behavior in the other variable.

As was the case for the one-dimensional M_R scaling described in section 9.2, we observe that each SM background can be described in the razor plane by the function $F_j(M_R, R^2)$, with some backgrounds having distinct kinematic subcomponents that require two instances of F_j. One such background is $W(\ell\nu)$+jets, one of the largest SM backgrounds in the high M_R/R region of the razor plane events. These events can have large values of M_R when two or more associated jets are produced with large energies (and invariant mass), and large R when the these jets are recoiling against
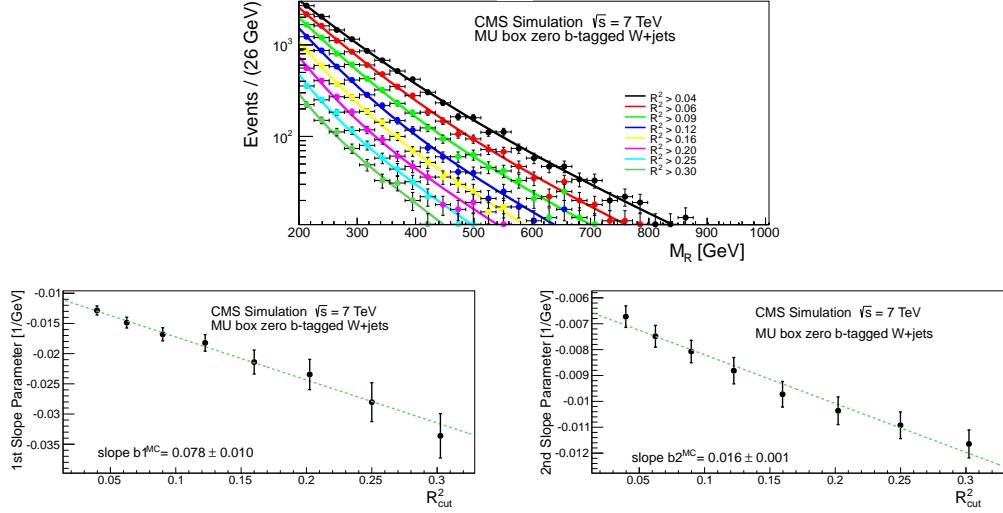
Figure 10.10: (Top) M_R distribution of simulated $W(\ell\nu)$+jets events as a function R^2_{cut}, requiring a reconstructed muon and an absence of b-tagged jets. Each distribution is fit independently with two exponential components. (Bottom left) value of exponential slope S of the first (steeper) component of M_R distribution, as a function R^2_{cut}. (Bottom right) value of exponential slope S of the second component as a function R^2_{cut}.

the weakly interacting neutrino from the W decay. The M_R distribution for simulated W+jets events in the muon final state is shown In figure 10.10, as a function of R^2_{cut}. We see that the distributions are well described by two exponential functions, and that the slope of each component scales linearly with M_R^{cut}, with different parameters.

Similarly, the analogous one-dimensional R^2 distributions for the same simulated W+jets data sample are shown in figure 10.11. We observe that the two exponential-component model describes the event sample well, and that the two exponential slopes evolve independently as a function of M_R^{cut}. Comparing the slope parameters $b1^{MC}$ and $b2^{MC}$ from the one-dimensional M_R view with $d1^{MC}$ and $d2^{MC}$ from the R^2 view we observe agreement, implying the two-dimensional distribution follows the functional form of equation (10.11).

In order to confirm this two dimensional scaling behavior in data for $W(\mu\nu)$+jets events, we select an event sample comprised almost entirely of this background process by requiring events have an isolated, well-identified muon (physics object reconstruc-
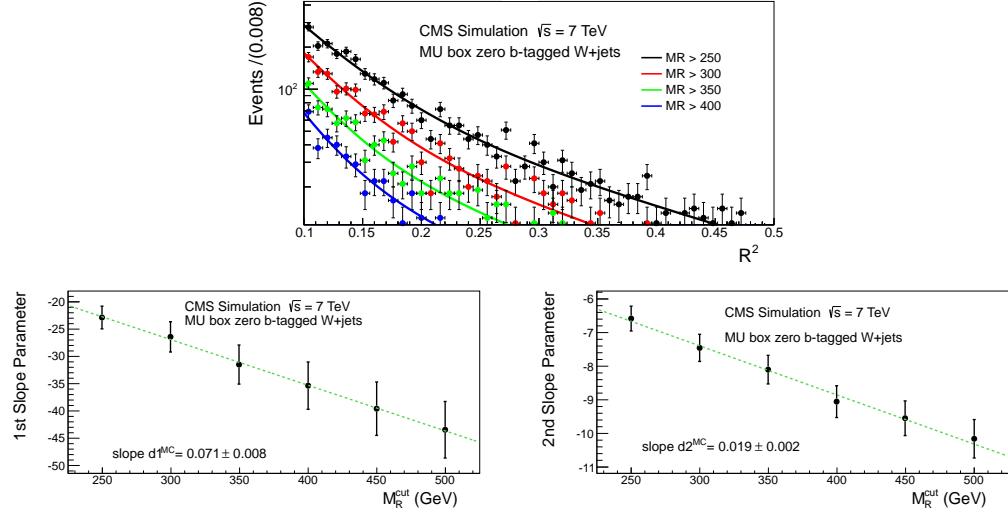
Figure 10.11: (Top) R^2 distribution of simulated $W(\ell\nu)$+jets events as a function $M_R^{\rm cut}$, requiring a reconstructed muon and an absence of b-tagged jets. Each distribution is fit independently with two exponential components. (Bottom left) value of exponential slope S of the first (steeper) component of R^2 distribution, as a function $M_R^{\rm cut}$. (Bottom right) value of exponential slope S of the second component as a function $M_R^{\rm cut}$.

Figure 10.12: (Top) M_R distribution of data events selected requiring a reconstructed muon and an absence of b-tagged jets as a function $R_{\rm cut}^2$. This event sample is composed almost exclusively of $W(\ell\nu)$+jets events. Each distribution is fit independently with two exponential components. (Bottom left) value of exponential slope S of the first (steeper) component of M_R distribution, as a function $R_{\rm cut}^2$. (Bottom right) value of exponential slope S of the second component as a function $R_{\rm cut}^2$.

tion and identification is described in section 10.4) and an absence on b-tagged jets, which rejects backgrounds with top quarks. With the selected event sample we perform two-exponential-component fits to the $M_R(R^2)$ distributions as a function of $R_{\text{cut}}^2(M_R^{\text{cut}})$, with the results shown in figure 10.12 (10.13). We observe that this data sample exhibits the two component scaling behavior in each of the one-dimensional M_R and R^2 distributions, and that the values of $b1^{\text{data}}$ and $b2^{\text{data}}$ are in agreement with $d1^{\text{data}}$ and $d2^{\text{data}}$, demonstrating that the sample can be described by two instances of equation (10.11). Additionally, we find that the values we measure for these parameters from data are in agreement with those extracted from our simulated $W(\ell\nu)$+jets sample.

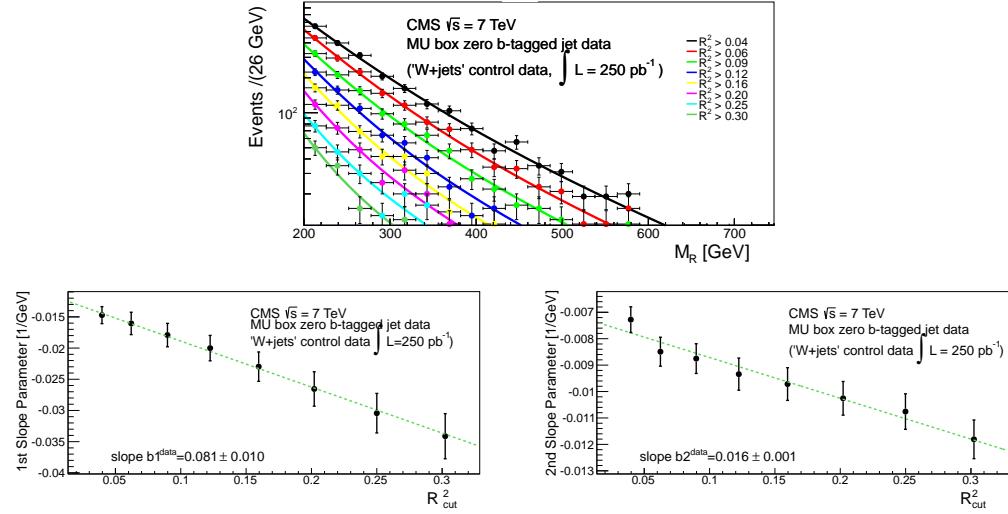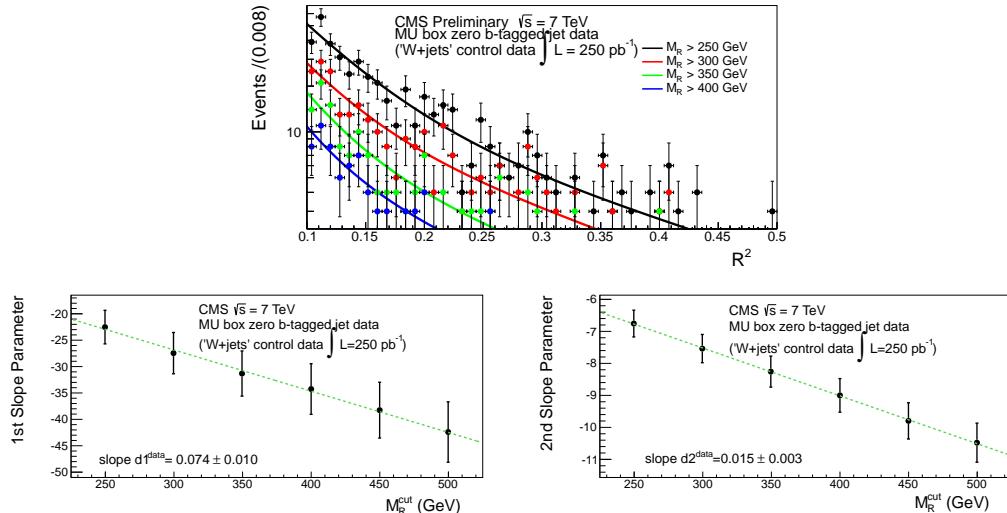Figure 10.13: (Top) R^2 distribution of data events selected requiring a reconstructed muon and an absence of b-tagged jets as a function M_R^{cut}. This event sample is composed almost exclusively of $W(\ell\nu)$+jets events. Each distribution is fit independently with two exponential components. (Bottom left) Value of exponential slope S of the first (steeper) component of R_2 distribution, as a function M_R^{cut}. (Bottom right) Value of exponential slope S of the second component as a function M_R^{cut}.

In the majority of final states in which we will search for evidence of SUSY, the other large background is comprised of $t\bar{t}$+jets production, particularly in final states with b-tagged jets. The corresponding one-dimensional M_R and R^2 distributions for simulated $t\bar{t}$ + jets events in final states with at least one b-tagged jet and no leptons

are are shown in figure 10.14-10.15. We see that this background is also well described by a two-component model.

Figure 10.14: (Top) M_R distribution of simulated $t\bar{t}$+jets events as a function $R^2_{\rm cut}$, requiring an absence of reconstructed leptons and at least one b-tagged jet. Each distribution is fit independently with two exponential components. (Bottom left) Value of exponential slope S of the first (steeper) componen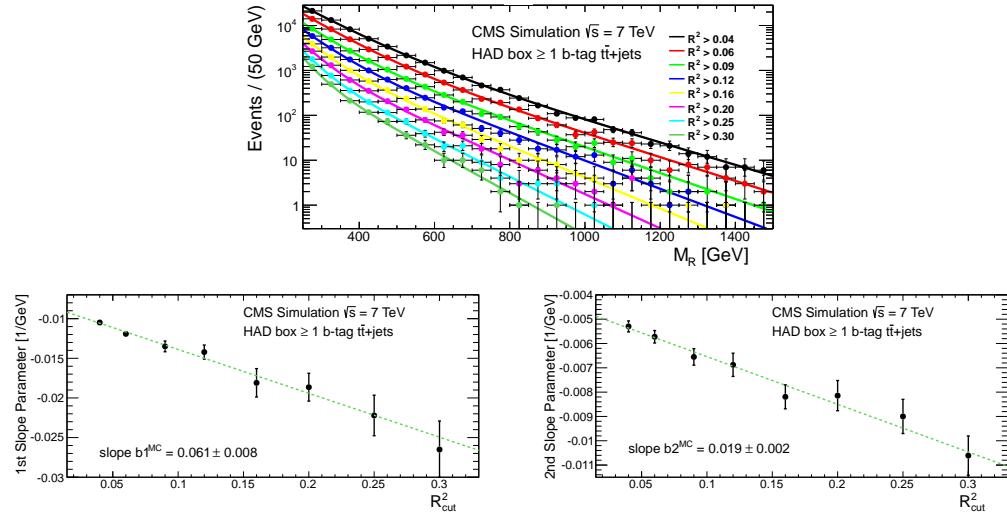t of M_R distribution, as a function $R^2_{\rm cut}$. (Bottom right) Value of exponential slope S of the second component as a function $R^2_{\rm cut}$.

In the search for SUSY described in section 10.3, each of the dominant background in various final states, $W(\ell\nu)$+jets, $Z(\ell\ell, \nu\nu)$+jets and $t\bar{t}$+jets, follow this two dimensional M_R/R^2 scaling behavior. For a given background type, we find that the parameters k_j, $M^0_{R,j}$ and $R^2_{0,j}$ are nearly identical between final states. We also observe that the parameters describing the second, or flatter, instances of F_j are nearly identical between different backgrounds, corresponding to a large initial-state radiation limit where we can no longer kinematically resolve the difference between, for example, semi-leptonic $t\bar{t}$+jets events and $W(\ell\nu)$+jets events as the $t\bar{t}$ and W systems recoil against hard jets. In this limit, the product of M_R and R^2 falls like the partonic luminosity as a function of $\sqrt{\hat{s}}$, with little sensitivity to the masses of the particles present in the event. All of the details of how the various backgrounds are modeled, and the observed relations between modeling parameters, are given in
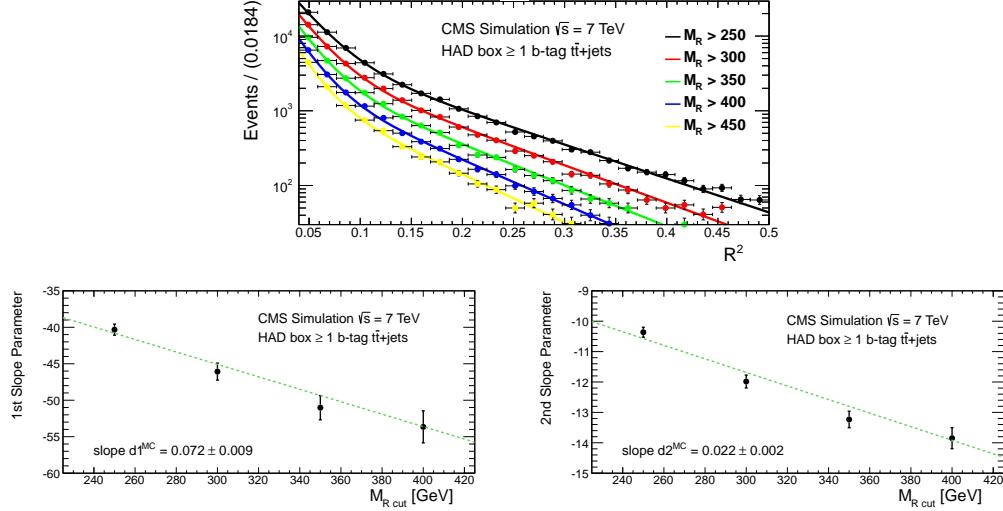
Figure 10.15: (Top) R^2 distribution of simulated $t\bar{t}$+jets events as a function M_R^{cut}, requiring an absence of reconstructed leptons and at least one b-tagged jet. Each distribution is fit independently with two exponential components. (Bottom left) Value of exponential slope S of the first (steeper) component of R_2 distribution, as a function M_R^{cut}. (Bottom right) Value of exponential slope S of the second component as a function M_R^{cut}.

section 10.5.2

It is important to note that the analytical description of the SM backgrounds in the razor plane given by equation (10.11) is empirical in nature, and will not hold to an arbitrary precision. Furthermore, the function F_j only describes backgrounds in a subset of the razor plane, at values of M_R in excess of the SM particles masses and intervals of R that can be process dependent. Despite these caveats, we find that this analytic approach provides an accurate description of SM backgrounds well within the precision we are sensitive to given the size of the dataset and the region of the razor plane we will search for SUSY; in fact, simulated events indicate that this background parameterization provides an adequate description for a dataset more than one hundred times as large as the one considered here.

With this 2D background prediction we can relate the shape of backgrounds in the low M_R/R^2 region to the shape at large values, meaning that we can measure the shape parameters of each background in a signal free region and extrapolate the full

analytic background prediction, with corresponding uncertainties, to the full razor plane. This procedure is described in detail in section 10.5.2. A fully continuous, 2D background predictions allows us to extract evidence of new physics not only through anomalously high event yields, but also through the *shape* of our selected data samples.

10.2.2   SUSY in the 2D Razor Plane

With the shape and yield of SM backgrounds in the razor plane understood, the task of identifying evidence of SUSY events amounts to identifying an excess of events that is *shaped* like SUSY. Since we have descriptions of both M_R and R, we can exploit our knowledge of both variables to identify and characterize signal events.

To illustrate the phenomenological properties of M_R and R^2 for SUSY events, we consider two example sparticle production and decay topologies: Di-squark production, where each squark decays to a quark (which hadronizes and is reconstructed as a jet) and an LSP, and di-gluino production, where the gluinos each decay to two quarks and an LSP. These example topologies are shown in figure 10.16. The details of event simulation for these samples can be found in section C.1.
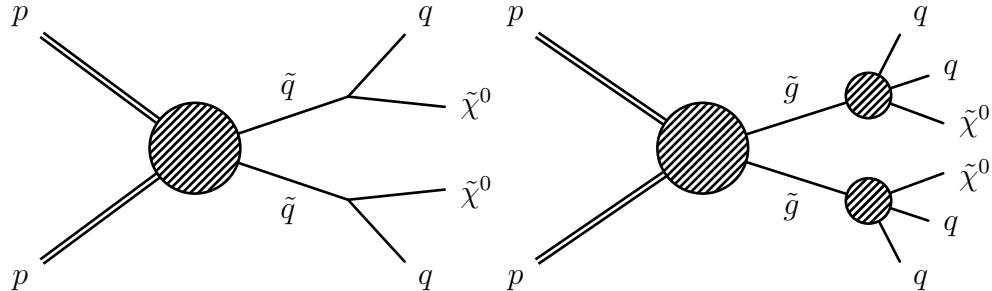
Figure 10.16: Example R-parity conserving SUSY production and decay topologies. (Left) Di-squark production. Each squark, $\tilde{q}$, decays directly to a quark, q. and a weakly interacting LSP, $\tilde{\chi}^0$. (Right) Di-gluino production with each gluino, $\tilde{g}$, decaying to two quarks, q, and an LSP, $\tilde{\chi}^0$.

For samples of each of these types of events, the M_R distribution is sensitive to the mass difference between the squarks/gluinos and the weakly interacting LSPs,

peaking at a characteristic scale

$$M_R^{peak} \sim M_\Delta = \frac{m_{\tilde{q}/\tilde{g}}^2 - m_{\tilde{\chi}^0}^2}{m_{\tilde{q}/\tilde{g}}} \ . \tag{10.12}$$

This peaking behavior is shown for our two example topologies in figure 10.17. Despite the fact that the di-gluino topology features three-body decays, and the quark-jets coming from those decays are not consistently paired in mega-jets with the correct gluino assignment, the M_R distribution still identifies with characteristic scale as if the decays were two-body to only one visible and invisible particle. This feature allows us to identify these events using their mass scale, even when decay topologies deviate form the simplest case illustrated by our d-squark example. It also means that, kinematically, the M_R distribution alone is not able to strongly distinguish between these two possibilities.

Figure 10.17: M_R distribution of simulated SUSY events as a function of different sparticle masses. (Left) Di-squark production where each squark decays to a quark and LSP. (Right) Di-gluino production where each gluino undergoes three-body decay to two quarks and an LSP.

The variable R can be used to distinguish between the two cases. R is sensitive to the partitioning of energy in the event between visible and invisible particles. Relative to the squark case, the di-gluino events have more visible, detectable jets in the final state and these objects carry more of the momentum on average than do the LSPs. The result is, on average, smaller values of R for di-gluino events than for

di-squarks, illustrated in the R^2 distributions shown in figure 10.18. In both cases, two weakly interacting particles in the final state result in larger values of R than the SM backgrounds, but more so for di-squarks. We also observe in figure 10.18 that the
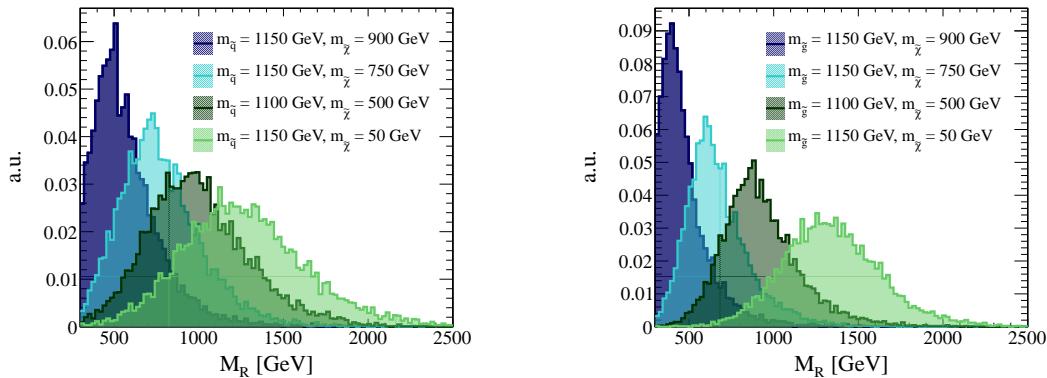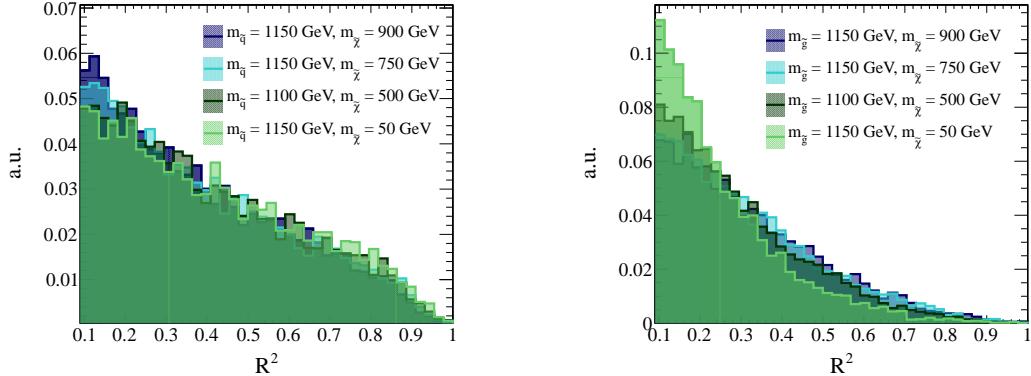

Figure 10.18: R^2 distribution of simulated SUSY events as a function of different sparticle masses. (Left) Di-squark production where each squark decays to a quark and LSP. (Right) Di-gluino production where each gluino undergoes three-body decay to two quarks and an LSP.

shape of the R^2 distribution for our examples is largely independent of the sparticle masses in these events. This implies an interesting phenomenological picture: The M_R distribution is sensitive to the masses of the particles in the event and largely insensitive to the decay topology. Conversely, the R^2 distribution is largely insensitive to the sparticle masses while sensitive to the decay topology. Hence, with information from both variables we can not only distinguish SUSY events from SM backgrounds, but also between different SUSY scenarios. Combined further with information from different SM particle content final states, the kinematic information contained in razor plane shapes can be used to characterize the properties of any potential observed excess.

The physics of the 2D razor plane can be understood pictorially in from figure 10.19: At large M_R the SM backgrounds (whose iso-yield contours are illustrated by orange lines) falls quickly towards the upper right, with the number of SM events getting exponentially smaller. On top of these steeply falling backgrounds we search for evidence of SUSY production by searching for deviations from this falling, or-
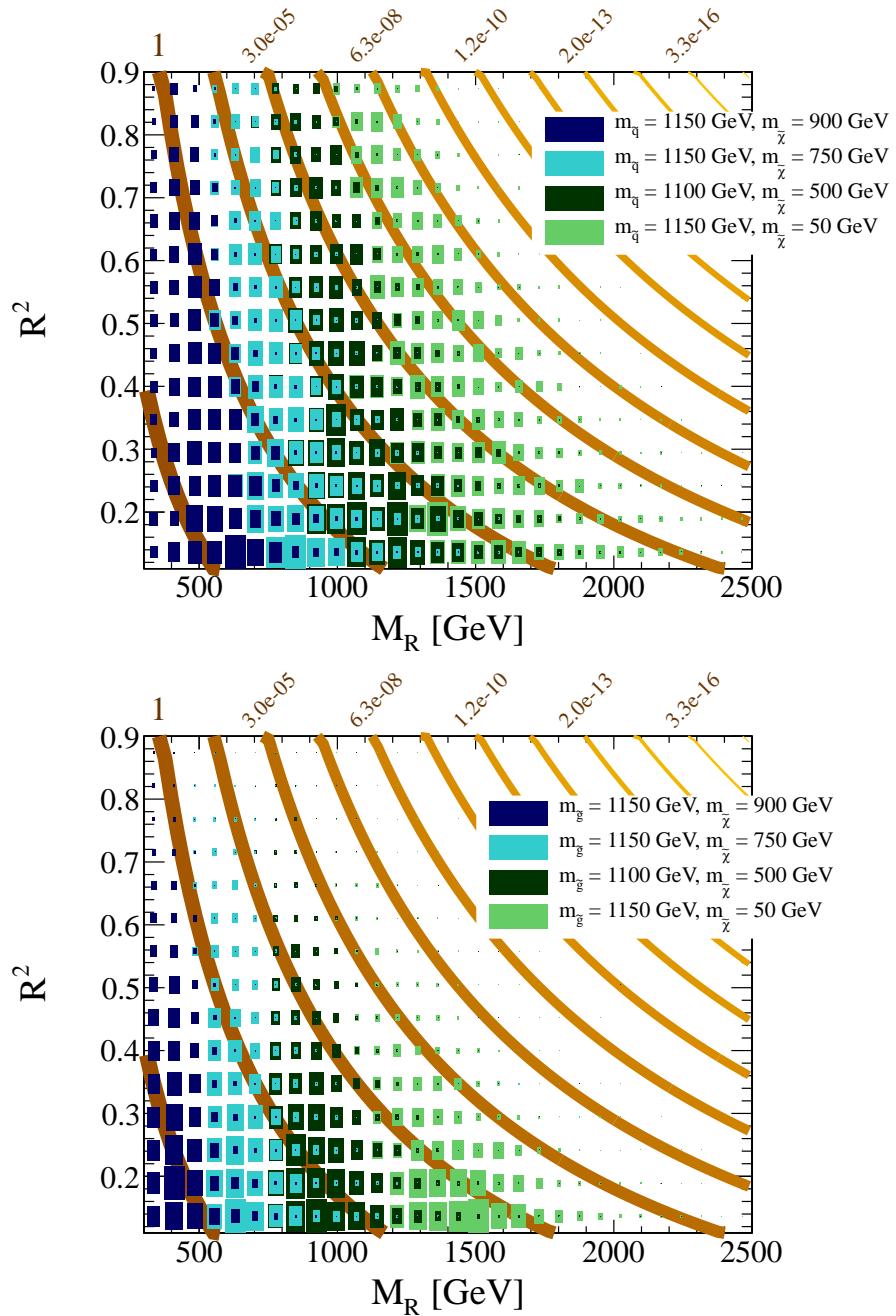
Figure 10.19: M_R vs. R^2 distribution of simulated SUSY events as a function of different sparticle masses. Orange contours indicate background iso-yield contours with number indicating yield relative to contour labelled **1**. (Top) Di-squark production where each squark decays to a quark and LSP. (Bottom) Di-gluino production where each gluino undergoes three-body decay to two quarks and an LSP.

derly behavior; blobs of new physics events with shapes set by the properties of those events, but unlike anything we have seen before.

10.3 The 5 fb^{-1} CMS Razor Search for SUSY

As discussed in section 8.2.1, models with softly broken supersymmetry (SUSY) [164–168] predict additional, undiscovered fundamental particles which correspond to the heavy superpartners of the SM particles. Experimental searches these for new particles in the context of R-parity [183] conserving SUSY at the Tevatron [184, 185] and the LHC [186–199] have focused on events signatures with energetic hadronic jets and leptons from the decays of pair-produced squarks and gluinos, with large missing transverse energy resulting from the two weakly interacting and stable sparticles produced in separate decay chains.

Of particular interest is the potential production of the superpartners of the third-generation quarks, stops and sbottoms. The quadratic divergences in the SM gauge sector are predominantly due to loop contributions with top quarks and their cancellation, as a result, is dependent on the masses of these superpartners, with a preference for light stops to avoid large fine-tuning. These considerations have motivated searches for the lightest allowed stops and sbottoms, whose decays would produce final states enriched in heavy flavor [200–204].
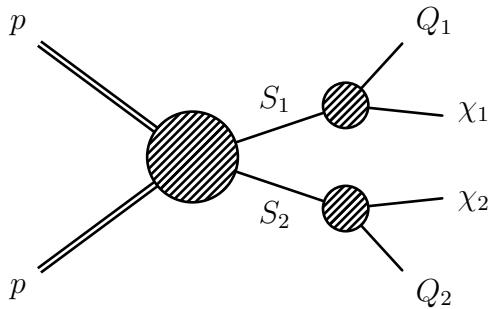
Figure 10.20: General R-parity conserving SUSY LHC event signature. Two massive sparticles, S_i, are produced in a pp collision and each decay to a system of detectable SM particles, Q_i, and a system of weakly interacting particles, χ_i.

In the remaining sections of this chapter we present an inclusive search for SUSY based on the razor kinematic variables [169, 205]. Motivated by the range of possibilities through which R-parity conserving SUSY could manifest itself at the LHC, we consider a variety of final states with leptons and b-tagged jets in the region of the

razor plane where we expect these events could appear.

This analysis follows the reconstruction approach introduced the search described in Chapter 9, where all the reconstructed final state objects in each event are grouped into two *mega-jets*. By casting each event into a two mega-jet final state, we interpret the event in the context of the canonical R-parity conserving signal topology of two pair-produced sparticles each decaying to a system of visible SM particles and one or more stable and weakly interacting LSPs, illustrated in figure 10.20. The razor variables, M_R and R, are derived in section 10.1 and can be calculated from the three momenta of these mega-jets event by event,

$$
\begin{aligned}
M_R &\equiv 2\sqrt{(|\vec{p}^{\,j1}| + |\vec{p}^{\,j2}|)^2 - (p_z^{j1} + p_z^{j2})^2} \;, \\
M_T^R &\equiv \sqrt{\frac{|\vec{M}|(|\vec{p}_T^{\,j1}| + |\vec{p}_T^{\,j2}|) - \vec{M}\cdot(\vec{p}_T^{\,j1} + \vec{p}_T^{\,j2})}{2}} \;, \\
R &\equiv M_T^R/M_R \;, \;,
\end{aligned}
\tag{10.13}
$$

where $\vec{p}^{\,j1}$ and $\vec{p}^{\,j2}$ are the three momenta of the two mega-jets and $\vec{M}$ is the missing transverse energy. As described in section 9.2, the phenomenological properties of these variables make them ideal for distinguishing between SM backgrounds and SUSY events reconstructed with the CMS detector.

The variable M_R is sensitive to the *scale* of events, such that for signal events of the type shown in figure 9.23 it will resolve the mass splittings between the parent sparticles S_i and the weakly interacting systems of particles, χ_i. The variable R is sensitive to the transverse imbalance of events, and requirements on its magnitude will suppress topologies that have less than two weakly interacting particles in the final state. This is accomplished without making strong assumptions about the missing transverse energy spectrum or any details of the intermediate decay chains. If the difference between the masses of new sparticles and those of the SM are resolvable, SUSY events will appear at larger values of M_R and R than the vast majority of SM background events. Using the phenomenological M_R scaling of these backgrounds, as described in section 9.2.1, we can make measurements of the background shapes and yields at low M_R and R, and use this information to predict these same backgrounds

in the high M_R and R razor plane where will search for evidence of signal events.

The strategy and execution of the 2011 CMS razor SUSY search, performed on 5 fb^{-1} of pp collision data at $\sqrt{s} = 7$ TeV, can be summarized as follows:

1. Events are selected by triggers which reconstruct the jets and leptons in the event at the higher level trigger (HLT) level. From these objects, the razor variables M_R and R are calculated and their values are used to decide whether to keep the event for study in the analysis, with looser kinematic requirements for events with electrons or muons. The excellent correspondence between the HLT and offline reconstruction makes this an efficient way to select events in interesting regions of the razor plane. These triggers are described in section 10.4.1.

2. Jets and leptons are reconstructed and identified in these events, and a b-tagging algorithm is applied to jets in order to identify those likely to correspond to b-quarks, as described in section 10.4.2.

3. The reconstructed objects in each event are combined into two mega-jets, which are used to calculate the variables M_R and R. Several baseline kinematic requirements are applied to clean the event samples of mis-reconstructed events and to ensure that we only consider regions of the razor plane that are efficiently selected by our triggers. This procedure is described in section 10.4.2

4. Events are assigned to *boxes* based on the presence or absence of a reconstructed lepton. This box partitioning scheme allows us to isolate individual background processes based on final state particle content and kinematic phase space in order to measure their shape and yield. Events with at least one b-tagged jet are considered in a parallel analysis focusing on searching for the superpartners of third generation quarks. In total we consider six different final state boxes: ELE-ELE, ELE-MU, MU-MU, ELE, MU and HAD. The details of the box classification scheme are provided in section 10.5.1.

5. In each box independently, we use the low M_R/R region of the razor plane to measure the shape and normalization of the various background contributions to

this final state. The analytic background model build from these measurements is used to extrapolate this background model from the signal-free region of the razor plane to the full region, where we can compare this independent prediction with the observed event sample. The details of this background model are explained in section 10.5.2.

6. Predictions for the shape and yield of SM background events in interesting regions of the razor plane are compared with observations, and used to make inferences about the presence or absence of SUSY events in these data samples. Section 10.7 explains how observations in each of the different final states considered in this analysis are used to constrain the parameter space of a collection of hypothetical SUSY models.

10.4 Event Selection and Reconstruction

There are two provenances for the event samples used throughout this analysis. The primary sample is events triggered and reconstructed during 2011 $\sqrt{s} = 7$ TeV pp collision running of the LHC using the CMS detector. A secondary sample is simulated Monte Carlo (MC) event samples, where event generators are used to simulate the particles produced in proton collisions, propagate these particles through a representation of the CMS detector, digitize the hypothetical signals these particles would leave in the detector and reconstruct the event as if it were recorded in data taking. This analysis is designed to minimally rely on information gleaned from simulated event samples. The different event samples used in this analysis are described in section C.1.

10.4.1 Trigger Selection

Events are recorded by the CMS detector if they satisfy one or more online trigger requirements. These triggers are based on fast, approximate reconstruction of the event that mimics later reconstruction requirements that will be applied in identifying

jets and leptons. For the 2011 CMS razor search data sample there were two distinct subsets of data, based on the triggers that were running online, that are used in two different ways in the analysis. After approximately the first 200 pb^{-1} of data was recorded in 2011 running, dedicated razor variable triggers, which calculate M_R and R from physics objects reconstructed at the HLT, were deployed to select events in interesting regions of the razor plane. This later dataset, with online razor triggers (R11 dataset), is what is used to search for evidence of SUSY in this analysis. The earlier dataset, without online razor triggers (NR11 dataset), does not necessarily provide an efficient selection of the razor plane in all final states. On the other hand, the lower instantaneous luminosity in this earlier running period allowed for generally lower threshold triggers, both with and without prescales, relative to later periods and is still useful for building an initial model of our SM backgrounds in the razor plane.

10.4.1.1 Razor Triggers

The razor triggers are implemented as a complementary suite, covering both different parts of the razor plane and different final states. For this analysis, we use a collection of hadronic triggers, which base their decisions on a calculation of M_R and R and the HLT level, and also leptonic triggers, which have lower M_R and R thresholds than their hadronic counterparts, but also require the presence of an electron or muon reconstructed at the trigger level, with further requirements on its momentum and quality.

HLT razor variables

The jets that go into the mega-jet calculation of M_R and R at the HLT are based on the trigger-level reconstruction of ECAL and HCAL energy measurements. After a full unpacking of the ECAL and HCAL, calorimeter towers are formed which represent massless particles with the measured energy of the ECAL and HCAL energies contained and with direction defined by the CMS projective geometry (see section 5.2 for details about calorimeter energy reconstruction and jet clustering).

These calorimeter towers are clustered into jets using an $R = 0.5$ (cone-size in azimuthal angle and pseudorapidity) anti-kT algorithm [57]. The reconstructed energies of these jets are corrected with on-line-environement-specific Jet-Energy-Scale (JES) corrections to ensure that these jets, and the variables calculated from them, have as close a correspondence to their off-line analogues which are used in the analysis. There must be at least two jets with $p_T > 56$ GeV$/c$ found in the central region of the detector ($|\eta| < 3$) in order for the event to pass. At the Level 1 (L1) trigger, razor triggers are seeded by `L1_DoubleJet36_CentralL1` seed, which requires two jets with a corrected E_T of 36 GeV and $|\eta| < 3$.

The jets that have $p_T > 40$ GeV$/c$ and $|\eta| < 3$ are further partitioned into two mega-jets which, along with MET determined from the same calorimeter towers, are used to calculate the variables M_R and R. This procedure is the same as for the offline analysis, described in section 10.4.3.

Hadronic razor triggers

HLT Path	Run Range	Int. Lumi.
HLT_R020_MR550	165088-183126	4.88 fb^{-1}
HLT_R025_MR450	165088-183126	4.88 fb^{-1}
HLT_R033_MR350	165088-183126	4.88 fb^{-1}
HLT_R038_MR250	165088-183126	4.88 fb^{-1}

Table 10.1: Evolution of the hadronic razor triggers. The luminosity quoted refers to the recorded value, before applying any data quality monitoring filter.

There are four hadronic razor triggers which base their firing decision solely on the variables M_R and R at the trigger level and are summarized in Table 10.1. Starting with run 165088, all of these triggers ran without a prescale and collected 4.9 fb^{-1} of integrated luminosity. Each trigger requires a different combination of minimum M_R and R^2 thresholds, indicated in the trigger name (HLT_R020_MR550 implies and $R > 0.2$ and $M_R > 550$ GeV).

Leptonic razor triggers

For final states with leptons, requiring a reconstructed lepton at the trigger level can reduce selection rates, permitting relaxed requirements on M_R and R. Events with electrons and muon are triggered using dedicated lepton razor triggers which include requirements on the momentum and identification quality of trigger level leptons, in addition to minimal values of M_R and R. These triggers, and the run periods there were deployed, are summarized in Tab. 10.2.

HLT Path	Run Range	Int. Lumi.
HLT_Mu8_R020_MR200	165088-166967	732 pb^{-1}
HLT_Mu8_R025_MR200	165088-173198	2.0 fb^{-1}
HLT_Mu10_R025_MR200	173236-180252	3.0 fb^{-1}
HLT_Ele10_CaloIdL_TrkIdVL_CaloIsoVL_TrkIsoVL_R020_MR200	165088-166967	732 pb^{-1}
HLT_Ele10_CaloIdL_CaloIsoVL_TrkIdVL_TrkIsoVL_R025_MR200	165088-173198	1.3 fb^{-1}
HLT_Ele12_CaloIdL_CaloIsoVL_TrkIdVL_TrkIsoVL_R025_MR200	173236-180252	3.0 fb^{-1}

Table 10.2: Evolution of the electron and muon razor triggers. The luminosity quoted refers to the recorded value, before applying any data quality monitoring.

The muons on which the lepton trigger decision is based follow from the L3 muon reconstruction, which follows from the addition of tracker information to L2 muon candidates, where a minimum p_T of 3 GeV/c is required. L2 muons are seeded from L1 through the `hltL1SingleMuOpenCandidate` module. In addition to the implicit identification requirement that there is an L3 muon, at least one muon must have a p_T of either 8 or 10 GeV/c, depending on the run range as shown in Table 10.2.

Due to higher rates of candidate electrons relative to muons (larger fake background), additional quality criteria are applied to the electrons used in the leptonic razor triggers and indicated in the trigger names. In general, these electron triggers use the loosest possible identification and isolation requirements. See section 4.3 for a description of the measured quantities on which these requirements are based.

- CaloIdL: loose calorimetric identification

  - $H/E < 0.15$ (0.1) for barrel (endcap) electrons (measure of hadronic energy near electron)

  - $\sigma_{i\eta,i\eta} < 0.014$ (0.035) for barrel (endcap) electrons (shower shape requirement)

- CaloIsoVL: very loose calorimetric isolation

  - relative ECAL isolation < 0.2

- TrkIdVL: very loose track-based electron identification

  - $\Delta\eta < 0.01$ between track and electron supercluster

  - $\Delta\phi < 0.15$ (0.1) between track and barrel (endcap) electrons

- TrkIsoVL: very loose track based isolation

  - ratio of track momenta to electron momenta in cone around electron less than 10%

Selected electrons must have a p_T of at least 10 or 12 GeV/c, depending on the run range as shown in Table 10.2. Increasing instantaneous luminosity through the 2011 run period resulted in tightening of lepton momentum and razor variable requirements with looser leptonic razor triggers being replaced by increasingly tighter ones. The combination of the events selected from each of them represents a complete dataset where events in the interesting region of the razor plane are efficiently selected in lepton final states.

Razor trigger performance

In addition to the razor triggers listed in Tab. 10.1 and 10.2 there were also prescaled, low-threshold monitoring triggers implemented in the trigger menu throughout running. These lower threshold triggers allow us to collect a small, unbiased sample with which we can evaluate the trigger turn-ons of the razor triggers relative the the values of the variables we compute offline and use in the analysis.

For the hadronic razor triggers there are loose threshold triggers `HLT_R014_MR150`, `HLT_R020_MR150` and `HLT_R025_MR150`. The turn-ons for the muon razor triggers can be evaluated using `HLT_IsoMu17` and `HLT_IsoMu24`, which each require an isolated

Figure 10.21: Turn-on curve on M_R (Left) and R^2 (Right) for events firing one of the prescaled loose monitoring triggers for hadronic (Top), muon (Center) and electron (bottom) events. The result of the fit of the plateaux to a constant function is shown.

L3 muon with transverse momentum of at least 17 and 24 GeV/c, respectively. The loose trigger HLT_Ele25_CaloIsoT_TrkIsoT_CaloIdT_DiCentralJet30 can be used to evaluate the electron razor triggers turn-ons.

The trigger turn-on functions for the suite of hadronic and leptonic razor triggers

are shown in figure 10.21. For each of these turn-on distributions, we fit a constant function to the plateaux region at high M_R and R. These fits indicate at which threshold the triggers become almost fully efficient and dictate the region of the razor plane where the analysis is performed. The parameters returned from these fits are interpreted as systematic uncertainties on the efficiency of these triggers for selecting signal events, as described in section 10.7.

An analogous set of turn-on curves and fits are created for the subset of events with at least one b-tagged jet in the event and are shown in figure 10.22. We observe that the additional b-tagging requirement is not strongly correlated with the small plateaux inefficiency of the razor triggers.

10.4.1.2   NR11 Control Dataset Triggers

The 200 pb^{-1} NR11 dataset trigger menu contains an array of prescaled and un-prescaled triggers which are useful for isolating particular background processes in the different final states considered in this analysis.

- `HLT_Ele8 and HLT_Ele17_CaloIdL_CaloIsoVL`: Two prescaled triggers in the DoubleElectronMu primary dataset (PD) with loose electron requirements and no constraints on hadronic activity in the event. Selected events result in an unbiased M_R/R distribution for electron final states.

- `HLT_JetX and HLT_DiJetAveX`: A collection of prescaled jet triggers requiring a jet with $p_T >$ X GeV/c or two jets with average $p_T >$ X GeV/c, respectively. Low values of X give an unbiased M_R distribution to very low values, the phase-space populated almost exclusively by QCD multijet events.

- `HLT_IsoMu17`: A low p_T threshold isolated muon trigger with no requirements on hadronic activity. The isolation requirement, applied to L3 muons, is looser than that applied for tight muons online (see section 10.4.2.2).

- `HLT_Ele17_CaloIdL_CaloIsoVL_Ele8_CaloIdL_CaloIsoVL`: Di-electron trigger with asymmetric electron minimum p_T thresholds. The identification, isolation
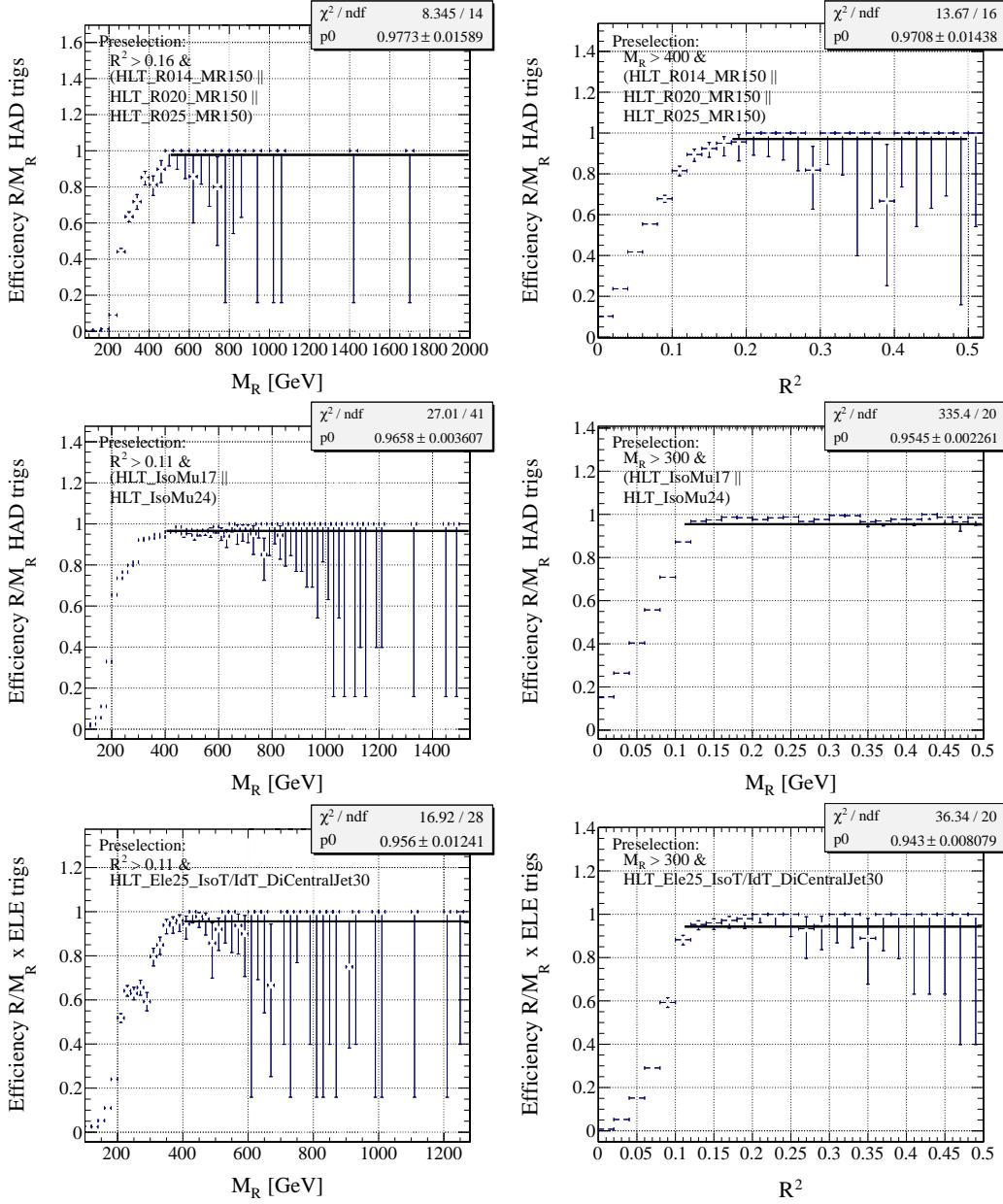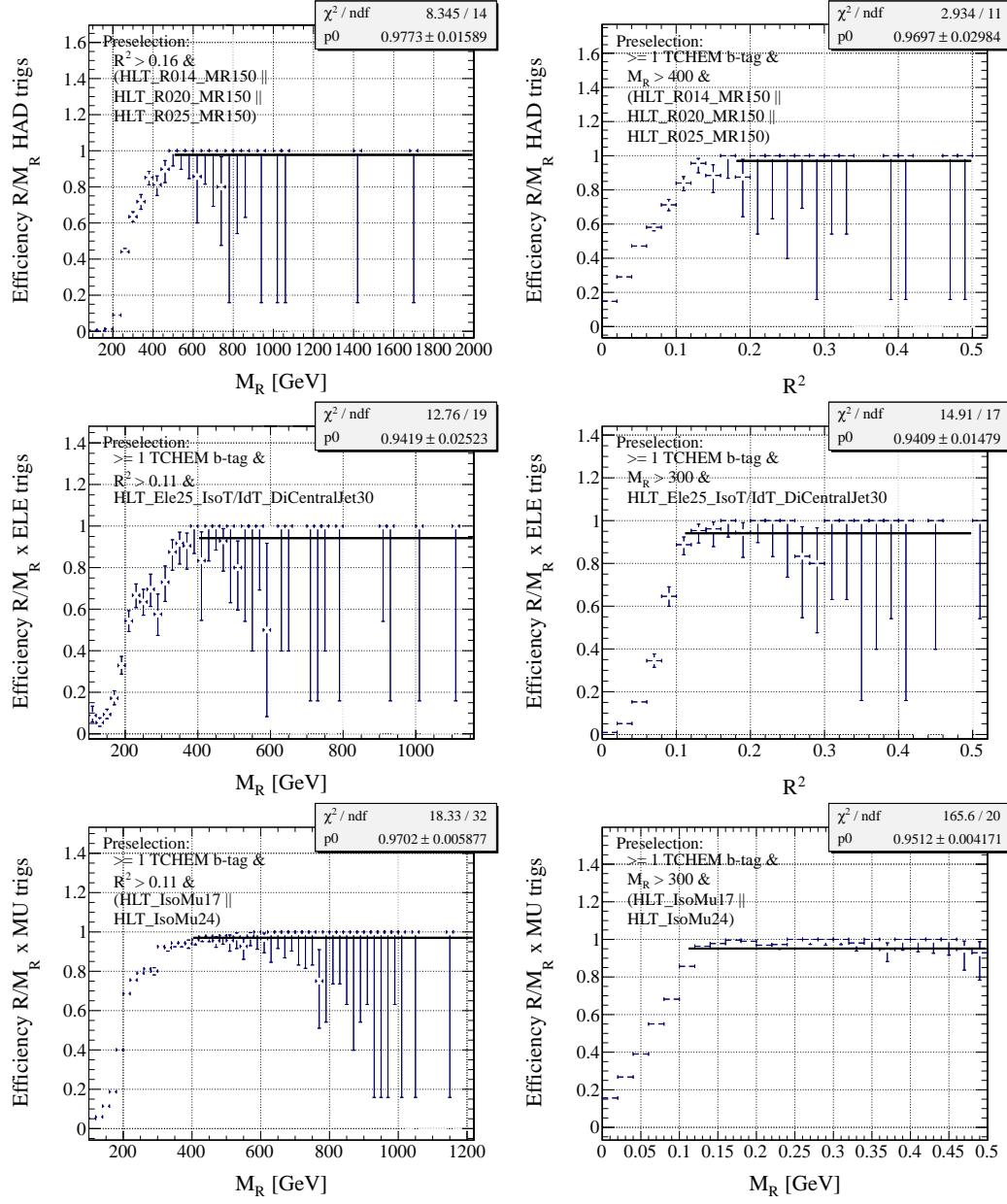
Figure 10.22: Turn-on curve on M_R (Left) and R^2 (Right) for events firing one of the prescaled loose monitoring triggers for hadronic (Top), muon (Center) and electron (Bottom) events, with the additional requirement of at least 1 b-tagged jet. The result of the fit of the plateaux to a constant function is shown.

and momentum requirements are tighter than the offline selection for di-electron final states and without hadronic activity requirements an unbiased razor plane selection can be applied.

- `HLT_DoubleMu7` and `HLT_Mu13_Mu8`: Di-muon triggers with symmetric and asymmetric p_T requirements for the two muons, respectively. Can be used to select di-muon final state events without biasing the distribution of razor variables.

- `HLT_Mu8_Ele17_CaloIdL`: Trigger requiring at least one L3 muon and one electron satisfying loose calorimetric identification with $p_T > 8$ and 17 GeV/c, respectively.

10.4.2 Physics Object Reconstruction and Identification

10.4.2.1 Primary Vertex Reconstruction

Each selected event is required to have at least one reconstructed Primary Vertex (PV) which satisfies several conditions. The vertex must be constructed with at least 13 associated degrees of freedom (at least 14 tracks matched to this vertex) and must be within a distance $|\Delta z| < 25$ cm from the beam spot along the beam axis. When multiple PV are reconstructed in an event the one with highest associated $\sum_{track} p_T$ is used to project physics object from when calculating their momenta.

10.4.2.2 Muon Identification and Selection

In order to minimize rates of identifying fake muons a series of quality requirements are enforced. In this analysis, we consider two different sets of criteria associated with *loose* and *tight* muons.

Tight muons must satisfy the following criteria:

- Muon is identified as both `TrackerMuon` (good candidate based on just inner tracker information with at least 10 valid hits) and `GlobalMuonPromptTtight` (global fit constraining tracker and muon hits with χ^2 per degree of freedom < 10 and transverse impact parameter ≤ 20 μm).

- At least one valid hit in the pixel layers of the tracker.

- Muon transverse momentum $p_T^\mu > 15$ GeV/c.

- Muon angular acceptance $|\eta^\mu| < 2.1$ GeV/c .

- Relative combined isolation $ISO(\mu, 0.3)/p_T^\mu < 0.15$, where $ISO(\mu, 0.3)$ is the sum of ECAL and HCAL energy deposits in an isolation cone of 0.3 $\Delta R = \sqrt{\Delta\eta^2 + \Delta\phi^2}$ and the sum of the momentum of tracks (excluding the muon) in the same isolation cone.

For the final combined isolation requirement, the momentum measured in the isolation cone is corrected for contributions from particles associated with PU interactions by subtracting the quantity $\pi(\Delta R)^2 \rho$ where $\Delta R = 0.3$ is the isolation cone size and ρ is the average energy per unit of area (in the $\Delta\eta$-$\Delta\phi$ plane) measured event by event using the FastJet package [57].

Whenever a muon is selected in the event, we look for a second muon satisfying the loose requirements:

- Muons are prompt with transverse impact parameter ≤ 10 μm.

- Muon transverse momentum $p_T^\mu > 10$ GeV/c.

- Muon angular acceptance $|\eta^\mu| < 2.1$ GeV/c.

- At least 10 valid hits in the strips layers of the tracker.

- At least 1 valid hits in the pixel layers of the tracke.

The tight muon requirements are designed to be strictly tighter than the muon legs of the triggers describe in section 10.4.1. More details about muon reconstruction can be found in section 3.5.

10.4.2.3 Electron Identification and Selection

Electrons are reconstructed in the CMS detector from clusters of ECAL energy deposits (superclusters) matched to tracks from the silicon tracker. The reconstruction of electrons is described in detail in section 4.3. In order to qualify as a good electron, candidates must satisfy a number of quality criteria based on the properties of

their ECAL clusters, their matched track and the inter-consistency of the energy and momentum measurements of the two, respectively. We define two different quality criterion for electrons: a tight selection denoted WP80, since it is tuned to be approximately 80% efficient for isolated electrons and a loose selection denoted WP95.

For an electron to be identified as WP80 is must satisfy the following requirements:

- Baseline transverse-momentum cuts $p_T^e > 10$ GeV$/c$.

- Super-cluster inside the fiducial region: $|\eta_{SC}| < 1.4442$ or $1.566 < |\eta_{SC}| < 2.5$.

- The lateral shower shape, $\sigma_{i\eta i\eta} < 0.01$ (0.031).

- The track-cluster matching in the ϕ-direction, $\Delta\phi_{\text{in}} < 0.027$ (0.021).

- The track-cluster matching in the η-direction, $\Delta\eta_{\text{in}} < 0.005$ (0.006).

- $ISO(e, 0.3)/p_T^e < 0.15$ (0.033) for barrel (endcap) electrons, where $ISO(e, 0.3)$ is the sum of ECAL and HCAL energy deposits in an isolation cone of 0.3 $\Delta R = \sqrt{\Delta\eta^2 + \Delta\phi^2}$ (excluding energies associated with the electrons) and the sum of the momentum of tracks (excluding the electrons') in the same isolation cone.

As for the muon, the momentum measured in the isolation cone is corrected for contributions from particles associated with PU interactions by subtracting the quantity $\pi(\Delta R)^2 \rho$ with ρ is the average energy density in the event.

Whenever an electron candidate passes this selection the event is queried for a second electron candidate satisfying the WP95 requirements:

- Baseline transverse-momentum cuts $p_T^e > 10$ GeV.

- Super-cluster inside the fiducial region: $|\eta_{SC}| < 1.4442$ or $1.566 < |\eta_{SC}| < 2.5$.

- The lateral shower shape, $\sigma_{i\eta i\eta} < 0.012$ (0.031).

- The track-cluster matching in the ϕ-direction, $\Delta\phi_{\text{in}} < 0.08$ (0.07).

- The track-cluster matching in the η-direction, $\Delta\eta_{\text{in}} < 0.007$ (0.011).

- $ISO(e, 0.3)/p_T^e < 0.40$ (0.1) for barrel (endcap) electrons.

Both WP80 and WP95 requirements are strictly tighter than the electron trigger requirements applied for the razor triggers described in section 10.4.1.

10.4.2.4   Jet Identification and Selection and Missing Transverse Energy Reconstruction

Jets are reconstructed from calorimeter towers, which are each associated with a 5x5 array of ECAL crystals and a HCAL module. Each of these towers is interpreted as a massless particle, with an energy equal to that measured by the calorimeters in the tower and momentum set by the projective geometry of the experiment. The towers are clustered into jets using the infrared and collinear-safe anti-kt jet algorithm [57] with a cone size of $R = 0.5$ in azimuthal angle and pseudorapidity-space. More details about jet reconstruction in CMS events can be found in section 5.2. Each of the reconstructed jets are corrected for noncompensating calorimetric energy response using Monte Carlo derived corrections (L2L3) and they are required to have a $p_T > 40$ GeV threshold and be within $|\eta| < 3.0$ in order to be considered in the analysis.

Jets passing this criteria are further required to have an electromagnetic fraction (the ratio of ECAL to HCAL energy deposits contributing to the jet) above 1% in order to reject HCAL noise.

Due to high the high number of interactions present in many of the events entering the analysis, we correct reconstructed jets' energy for contributions from particles not coming from the event primary vertex. This is accomplished by calculating an effective area for each jet and the average energy density per unit area, event by event using the `FastJet` software package [58, 59], with the input the same calorimeter towers used to create the jets. The jet area and energy density are multiplied and the sum removed for each jet.

The missing transverse energy is calculated as the negative transverse vectorial sum of all of the reconstructed particle flow (PF) candidates in the event. The PF reconstruction algorithm is described in section 5.2. Effectively, this collection of PF

candidates accounts for all of the reconstructed energy and momentum in each of the CMS subdetector systems.

10.4.2.5 Jet b-quark Tagging

The presence of b-quarks in final states can be an indication of events with top decays or new physics particles which couple to the third generation quarks. In order to select final states enriched in these events, a parallel event selection and analysis include the additional requirement of one b-tagged jet in the event (inclusive and b-tagged jet analyses, respectively).

To satisfy this b-tagged requirement, a jet must satisfy the *medium* working point of the Track-Count High-Efficiency (TCHE) tagging algorithm (TCHE discriminant > 3.3) [30]. The b-tagging discriminant is based on the tracks matched to a secondary vertex associated with the jet, with the likelihood of the observed configuration matching a b-quark-initiated jet being reflected in the discriminant magnitude. b-tagged jets must further satisfy $p_T > 40$ GeV and $|\eta| < 3.0$.

In order to test whether this b-tagging requirement will bias the shape of the M_R and R^2 distributions, we consider the shapes of event yields in samples selected from data and simulated events with different b-tagging requirements applied. These distributions, and the ratio of their shapes with respect to no b-tagging requirements, are shown in figure 10.23 (figure 10.24) for the M_R (R^2) distribution. The efficiency for tagging b-quark jets is dependent on the jets' momenta, particularly for low momenta jets. Similarly, if these jets represent a large fraction of the energy in an event then M_R can be correlated with those same jets' momenta, and the M_R distribution can be biased by tagging requirements.

We observe that the shapes of the M_R and R distribution are stable to within 10% for the simulated event samples considered in figure 10.23 and 10.24, for both M_R and R^2. The selected data sample exhibits are stronger M_R and R^2 shape dependence on the b-tagging requirement, although this is not necessarily indicative of kinematic bias. This data sample is made up of a combination SM backgrounds, the majority of events corresponding to $t\bar{t}$+jets and $W(\ell\nu)$+jets. Since the former is enriched

Figure 10.23: Distribution of M_R for different working points of the TCHE b-tagging algorithm, in the case of Data (Top left), QCD MC (Top right), W+jets MC (Bottom left), and $t\bar{t}$ MC (Bottom right). The ratio of these distributions to the inclusive distribution are shown in the bottom part of each plot.

in b-quarks while the latter is not, a b-tagged jet requirement changes the relative composition of these two background. Since they have different shapes, the overall shape of all SM backgrounds changes with the tagging requirement. Comparing, instead, the M_R and R^2 shapes of data with one and two b-tagged jet requirements
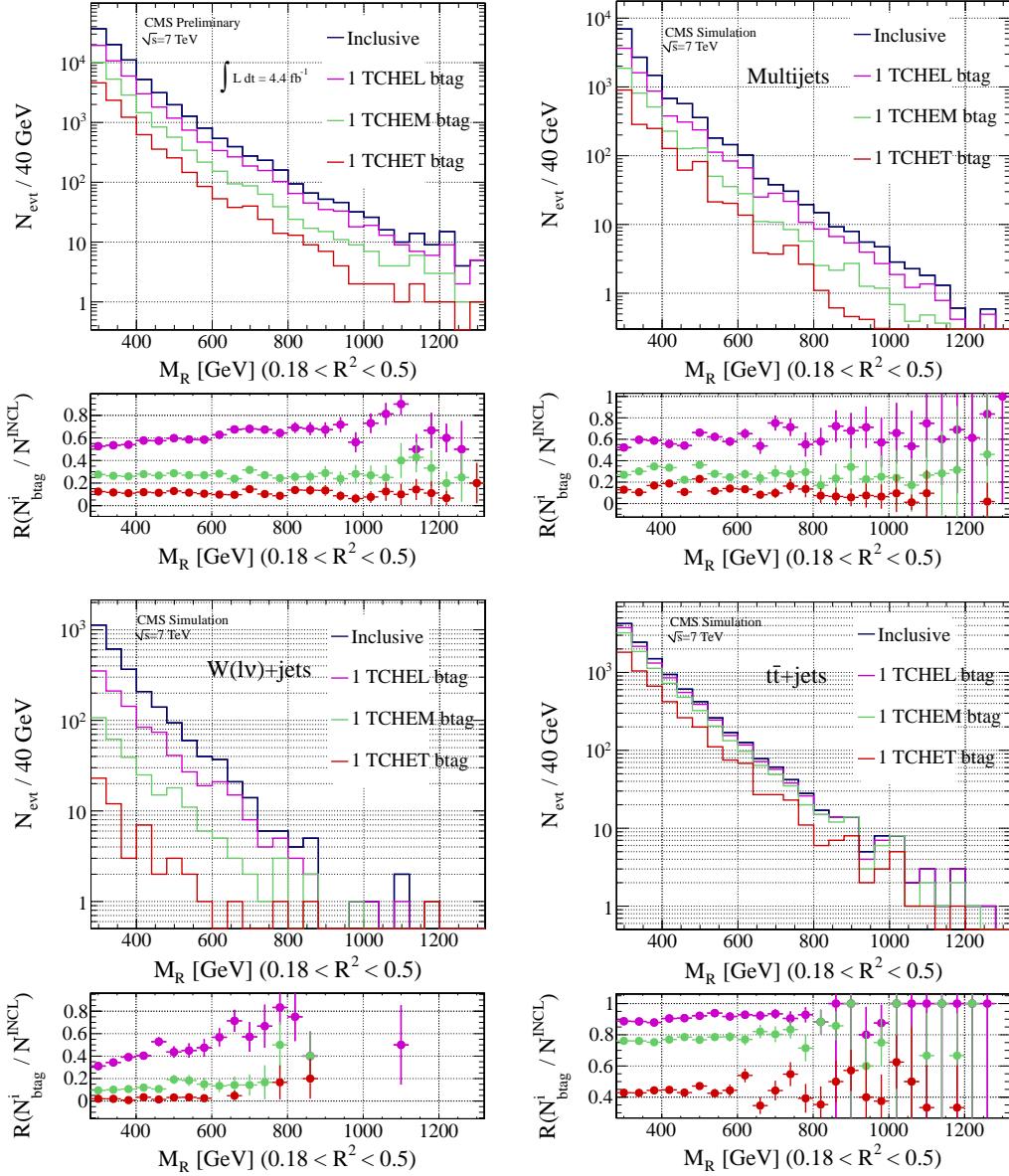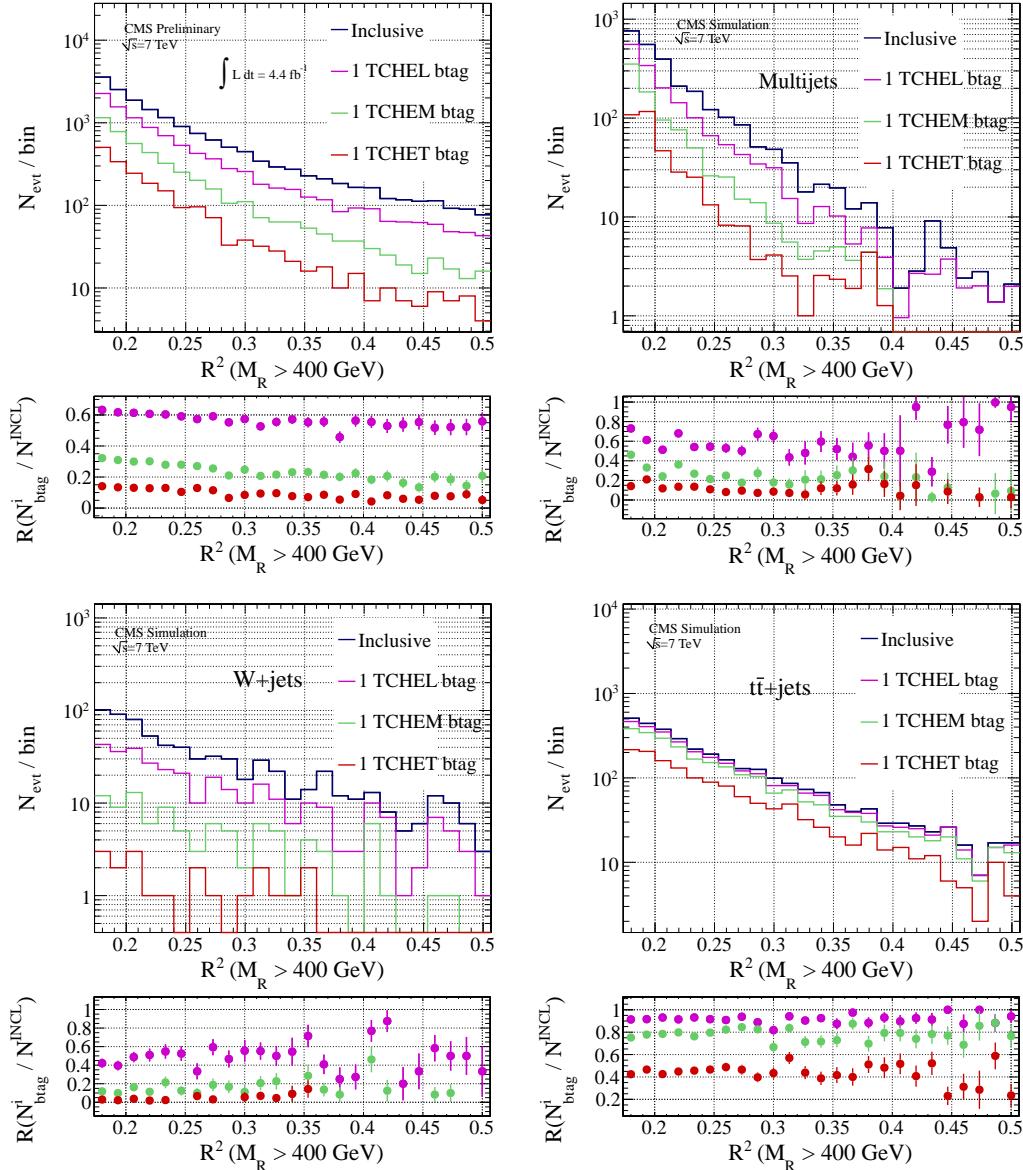
Figure 10.24: Distribution of R^2 for different working points of the TCHE b-tagging algorithm, in the case of Data (Top left), QCD MC (Top right), W+jets MC (Bottom left), and $t\bar{t}$ MC (Bottom right). The ratio of these distributions to the inclusive distribution are shown in the bottom part of each plot.

we observe agreement to a few percent, as illustrated in figure 10.25. In this case, both samples are made up predominantly of $t\bar{t}$ events and the additional b-tagged jet requirement is shown to not bias the shape of the kinematic variables for this isolated background.
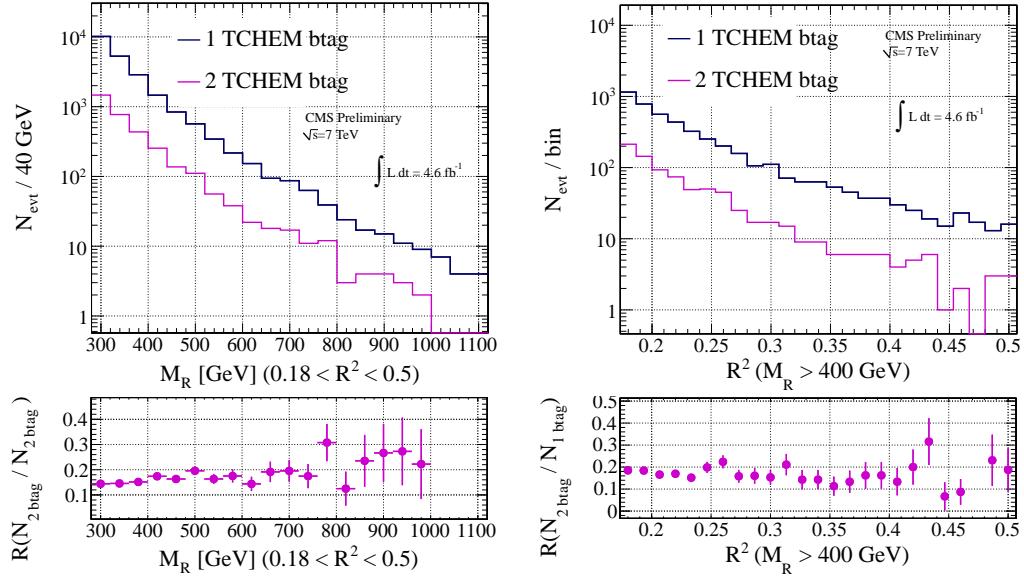
Figure 10.25: . Distribution of M_R (left) and R^2 (right) for data events with at least one and two b-tagged jets (TCHEM). The ratio of these distributions is shown in the bottom part of each plot.

10.4.3 Kinematic Requirements and Mega-Jet Reconstruction

The calculation of the variables M_R and R proceeds from events with two mega-jets, which are constructed from combinations of all of the visible objects in the event. In order to contain these mega-jets an event must have at least two reconstructed jets. If there are more than two then every partition of the jets into two nonempty groups is considered. A mega-jet is formed from the jets assigned in its partition by summing their four vectors

$$p_{MJ}^n = \sum_{i \in n} p_c^i \, , \tag{10.14}$$

where the index n runs over the two mega-jet four vectors p_{MJ}^n and the index i indicates the jets' four vectors p_c^i which are assigned to mega-jet n. Out of all the possible partitions of the reconstructed jets into two mega-jets the one which minimizes the mega-jet masses, $(p_{MJ}^n)^2$, summed in quadrature is chosen, an algorithm

adopted from [173].

Minimizing the mega-jets' invariant masses implicitly groups together jets traveling in similar directions, maximizing the balance of momentum in the event. It also maximizes the momenta of each of the mega-jets, yielding larger values of M_R and smaller R than other potential assignment choices. In practice, the chosen algorithm is found to be quite stable and effective at rejecting background with the kinematic variables of interest.

The mega-jets are built from a collection of jets which can contain energy from electrons, which is not explicitly removed from the calorimetric towers clustered into jets. Conversely, isolated muons will leave only a MIP energy deposit in the calorimeters such that their momentum is not accounted for in the collection of jets. This means that events with electrons can have subtle kinematic differences relative to those with muons, even if they both come from a W, in the context of the variables calculated from the mega-jets. This choice was made to maintain consistency with the reconstruction of the kinematic variables at the trigger level, which is based on calorimetric objects that are not sensitive to muons' momenta. While the backgrounds in muon final states can be artificially less constrained by the kinematic variables with this choice, this creates a natural control sample to study and constrain the shape of $Z(\nu\nu)$+jets events in the hadronic final state, using kinematically similar $W(\mu\nu)$+jets events.

When constructing mega-jets, one can choose whether to include reconstructed leptons in the final state among the visible objects used to build the mega-jets. Alternatively, the leptons can be treated as *invisible* and removed from consideration in the calculation of the kinematic variables (as if they were escaping weakly interacting particles). For backgrounds like $W(\ell\nu)$+jets, the former choice yields more transversely balanced mega-jets, and lower values of R, due to the fact that, since they come from the same decay, harder neutrinos (and hence larger MET) will also produce leptons. For this background, when leptons are treated as invisible, the MET corresponds to the entire W transverse momentum, similar to the case of $Z(\nu\nu)$+jets events.

The analysis has no other selection requirements, apart from restricting the portion of the razor plane considered to that efficiently selected by the razor triggers. Minimal M_R and R requirements are determined from the trigger turn-on curves presented in section 10.4.1 for each of the razor trigger combinations:

- $M_R > 300$ GeV and $0.11 < R^2 < 0.50$ for events selected with leptonic razor triggers.

- $M_R > 400$ GeV and $0.18 < R^2 < 0.50$ for events selected with hadronic razor triggers.

10.5 Box Classification and Background Modeling Strategy

The expected phenomenology of the sought after SUSY events suggests that the sparticles that could be produced at the LHC should be massive, maybe in excess of the SM particles, and that R-parity will guarantee that there are at least two weakly interacting particles in each event produced; these considerations motivate the region of the razor plane we examine for evidence of signal. Similarly, the types of SM particles we expect to see in these new physics events motivates which final states we consider, although in this case the unknown properties of potential signals compels us to search in a large collection of final states, including those with and without one or more leptons or b-tagged jets.

The selected data sample is classified according the the type and number of reconstructed objects into different *boxes*. If a potential signal happens to populate one or a subset of these boxes, then this partitioning enhances our ability to observe it. Similarly, comparing yields in different boxes provides information about the branching ratios of observed processes. The SM backgrounds are modeled independently in each box through fits to the low-M_R/low-R^2 region, where we expect negligible contributions from expected signals. Derived from the empirical M_R/R^2 scaling described in section 10.2, an analytic shape parameterization of M_R and R^2 is used to model

the shape and yield of backgrounds in this fit region and extrapolate the complete background model to the rest of the high-M_R/high-R^2 plane, where this prediction can be compared with observed event yields and the presence or absence of SUSY event contributions can be inferred.

10.5.1 Razor Box Definitions

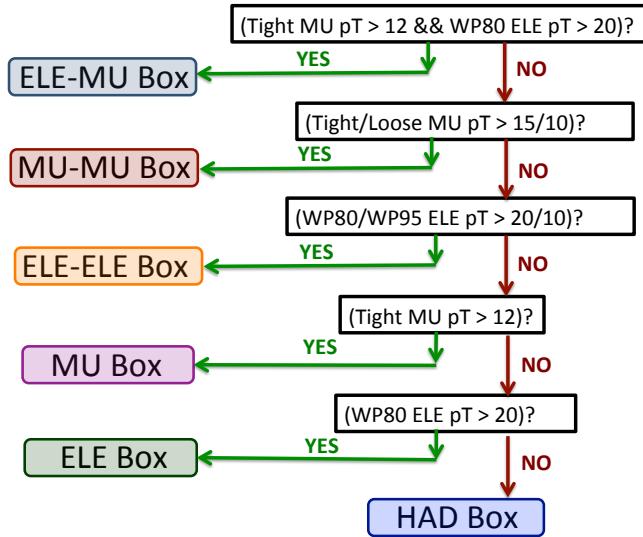
Figure 10.26: Flow diagram of box classification logic. The box selection proceeds hierarchially in order to ensure complete orthogonality of box selections by resolving ambiguities when an event satisfies more than one box's selection criteria.

Six different final state boxes are defined based on the number, type and quality of reconstructed electrons and muons in an event. These boxes are organized in a hierarchy, illustrated in Fig. 10.26, such that an event is consequently queried as to whether it satisfies each box's requirements until it is assigned to one and only one box. The box definitions, listed according to assignment hierarchy, are

- MU-ELE Box: Events must contain at least one WP80 electron with $p_T > 20$ GeV/c and at least one tight muon with $p_T > 10$ GeV/c.

- MU-MU Box: Events must contain at least one tight muon with $p_T > 15$ GeV/c and one additional loose muon with $p_T > 10$ GeV/c.

- ELE-ELE Box: Events must contain at least one WP95 electrons. with $p_T > 20$ GeV/c and one WP80 electron with $p_T > 10$ GeV/c.

- MU Box: Events must contain at least one tight muon with $p_T > 15$ GeV/c.

- ELE Box: Events must contain at least one WP80 electron with $p_T > 20$ GeV/c.

- HAD Box: Events must not satisfy any other box requirements.

Events contributing to these different boxes are originally selected in data using different triggers, depending on the box. The events assigned to the ELE, ELE-MU and ELE-ELE boxes are selected using the electron razor triggers; the MU and MU-MU box events are selected with the muon razor triggers; the HAD box is populated from the hadronic razor triggers.

The study of the razor trigger turn-on curves, described in section 10.4.1, indicates the trigger-dependent region of the razor plane where events are efficiently selected. Only these regions of the razor plane are considered in the analysis, and events not satisfying box-dependent minimal M_R and R^2 requirements are not considered in the analysis:

- $M_R > 300$ GeV and $0.11 < R^2 < 0.50$ for the leptonic boxes.

- $M_R > 400$ GeV and $0.18 < R^2 < 0.50$ for the HAD box.

10.5.2 Building the Background Model

The strategy for modeling the SM backgrounds over the high-M_R/high-R^2 region of the razor plane where signal events could appear begins in the low-M_R/low-R^2 region of the razor plane, above the minimal M_R and R^2 requirements set by trigger turn-ons described in section 10.5.1. Here, the yields of SM backgrounds are high, falling steeply with increasing M_R and R^2. The expected contribution of potential SUSY events is also small, particularly relative to SM backgrounds.

The M_R/R^2 scaling described in section 10.2 indicates that we can relate the shape and yields of backgrounds in this low-M_R/low-R^2 region to our expected backgrounds in the potentially signal populated part of the razor plane. Each SM background can be described by instances of the function $F_j(M_R, R^2)$,

$$F_j(M_R, R^2) = \left[k_j(M_R - M_{R,j}^0)(R^2 - R_{0,j}^2) - 1 \right] \times e^{-k_j(M_R - M_{R,j}^0)(R^2 - R_{0,j}^2)} \; , \quad (10.15)$$

where k_j, $M_{R,j}^0$ and $R_{0,j}^2$ are parameters specific to the component j of the background. By measuring the parameters that describe the SM backgrounds' shapes, and their normalizations, in the low-M_R/low-R^2 region a background model can be extrapolated to the full razor plane using these parameters and equation (10.15).

Each razor box is split into a *fit region*, located in the lower left corner of the razor plane, and a *signal sensitive region*, covering the remainder of the plane above the box minimum M_R and R requirements. The razor plane partitioning scheme for each box is illustrated in Fig. 10.27. The signal sensitive region is further split into six signal regions, uniform across the boxes except for the tighter minimum R^2 requirement in the definition of the HAD box. A comparison of the predicted and observed event yields in these signal regions is used to establish a signal-model-independent metric for assessing agreement between predictions and expectations. These signal regions are not used for model-dependent interpretations of the results.

The shape and normalization of the SM background for a given box can be described as a likelihood which incorporates the functions F_j which describe each component of the background. For the portion of the 2011 where razor triggers were deployed (R11 dataset) the likelihood describing the data can be written as [206]

$$\mathcal{L}^{\text{R11}} = \frac{e^{-(\sum_{i \in SM}^{N_{SM}} N_i)}}{N_{event}!} \prod_{j=1}^{N_{event}} \left[\sum_{i \in SM}^{N_{SM}} N_i F_i(M_R^j, R_j^2 | \vec{\alpha}_i) \right] \; , \quad (10.16)$$

where i indicates the background component among the N_{SM} modeled, the sum j is over all events observed in the box, N_i is the event yield of the ith background component and $\vec{\alpha}_i = [k_i, M_{R,i}^0, R_{0,i}^2]$ are the parameters describing the shape of the
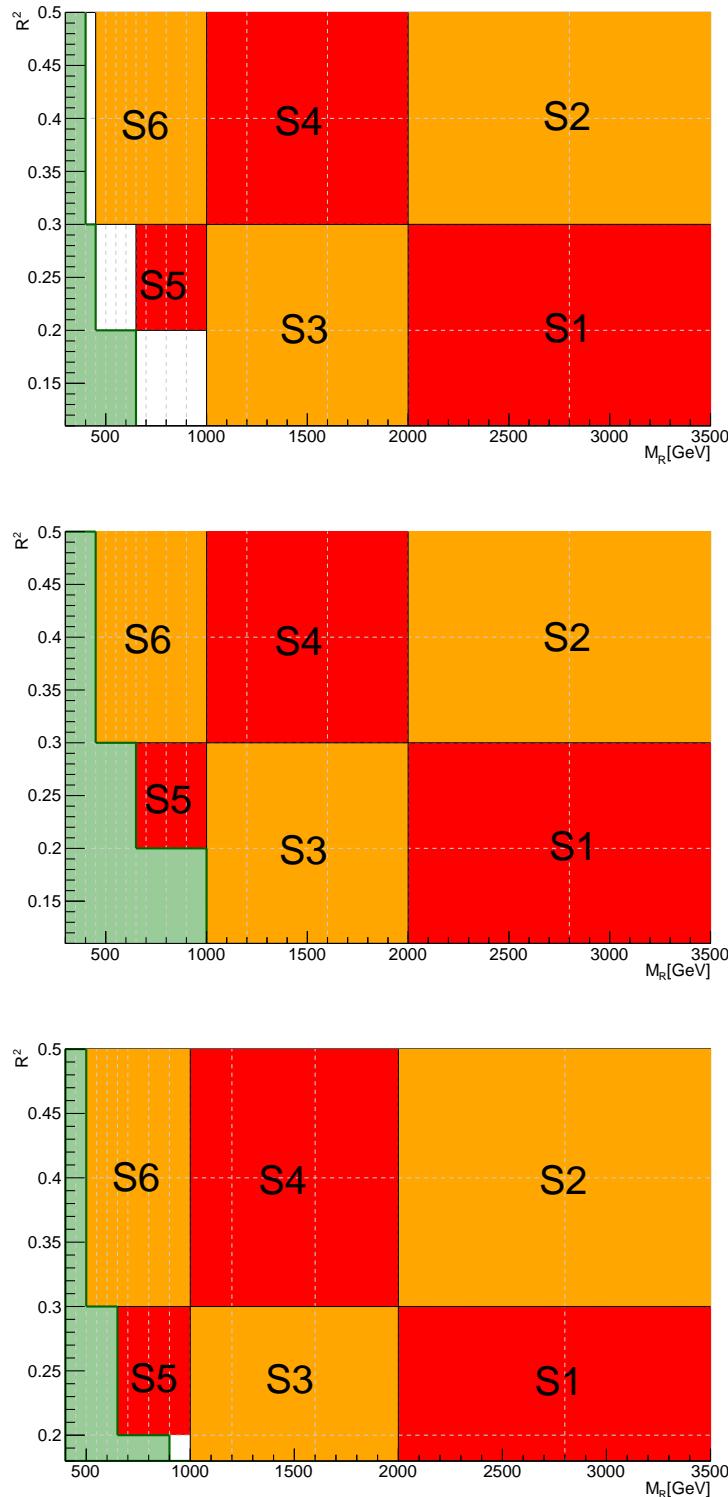
Figure 10.27: Definition of razor plane regions for double lepton (ELE-ELE, MU-MU, ELE-MU) (top), single lepton (ELE, MU) (middle), and HAD (bottom) boxes. The fit region for each box is indicated by a solid green contour while the signal sensitive regions, Si, are red and orange and labelled $i = 1, ..., 6$.

*i*th background. Independently in each box, a maximal likelihood fit is performed in the fit region on the R11 dataset to determine the parameters $\vec{\alpha}$ which describe each background. The extended term at the front of the likelihood indicates that the total sum of background events $\sum_i N_i$ does not need to add exactly to N_{event}, the total number of events in the sample, since these normalizations correspond to the true amplitude of each signal around which Poisson fluctuations can occur. With shape parameters and normalizations determined, the background model can then be extrapolated to the full signal sensitive region.

A challenge for this strategy is the large number of parameters contained in the likelihood $\mathcal{L}^{R11}$. Each pair $M_{R,i}^0, R_{0,i}^2$ describe translations of a given background component in the razor plane while the parameter k_i indicates the slope of the falling SM backgrounds moving towards the upper right hand corner. The determination of these parameters is often strongly correlated for individual backgrounds, and even between different backgrounds, particularly if they have similar shapes. In order to guide each ML fit to physically meaningful regions of parameter space, normally distributed penalty terms are added to the box likelihoods [207].

A penalty term for parameter α_{ik} can be written $G(\alpha_{ik}|\alpha_{ik}^{NR11}, \sigma(\alpha_{ik}^{NR11}))$, where G is a Gaussian (truncated at 0) with mean α_{ik}^{NR11} and standard deviation $\sigma(\alpha_{ik}^{NR11})$. With these terms added, the total likelihood used in each box's fit region ML fit can be expressed as

$$\mathcal{L}^{TOT} = \mathcal{L}^{R11} \times \prod_{ik} G(\alpha_{ik}|\alpha_{ik}^{NR11}, \sigma(\alpha_{ik}^{NR11})) \; , \tag{10.17}$$

where the sum over ik includes each parameter for which a penalty term is included.

The superscript $NR11$ hints at the provenance of the parameters associated with these penalty terms; they are derived from measurements from the NR11 dataset, the first 200 pb^{-1} of 2011 collision data taken when the razor triggers were not deployed. With low instantaneous luminosity, this early run period included low threshold lepton triggers with requirements looser than the lepton identification requirements of the razor box definitions, and without additional requirements on hadronic activity.

This removes the minimal M_R and R^2 requirements for considering the region of the razor plane that can be selected efficiently. In this more inclusive low-M_R/low-R^2 region control samples were defined in order to isolate events corresponding to the nonnegligible background contributions over all the boxes: $W(e\nu)$+jets, $Z(\ell\ell)$+jets and $t\bar{t}$+jets.

Requirements on the presence or absence of b-tagged jet are applied in order to isolate samples of $t\bar{t}$+jets and W/Z+jets, respectively, in each of the boxes. Using control samples composed almost exclusively of one background process, the shape parameters of each background component, $\vec{\alpha}_i$, is measured. Each of these shapes is then combined to form the full background description for each box, which is determined from the fit region ML fit to the R11 dataset in each box, with penalty terms derived from the NR11 control measurements multiplying the likelihood. The measurements performed on the NR11 data sample are described in section C.2.

10.6   The Razor Background Model

A full SM background model is built in each box from a ML fit performed in the low-M_R/low-R^2 fit regions illustrated in figure 10.27. There, the parameters describing the shape and yield of each SM background are measured, resulting in a background prediction that can be extrapolated to the full razor plane.

The SM backgrounds in each box are described by separate instances of the 2D razor function $F(M_R, R^2)$, corresponding to each background component present in the box. The fit measures the shape parameters $\vec{\alpha}_i = [k_i, M_{R,i}^0, R_{0,i}^2]$ associated with each background function F_i by finding the values which best fit the observed event yields in the fit region, maximizing the likelihood

$$\mathcal{L}^{\mathcal{SM}} = \frac{e^{-(\sum_{i \in SM}^{N_{SM}} N_i)}}{N_{event}!} \prod_{j=1}^{N_{event}} \left[\sum_{i \in SM}^{N_{SM}} N_i F_i(M_R^j, R_j^2 | \vec{\alpha}_i) \right] \times \prod_{ik} G(\alpha_{ik} | \alpha_{ik}^{NR11}, \sigma(\alpha_{ik}^{NR11}) \; ,$$

$$(10.18)$$

where i indicates the background component among the N_{SM} modeled, the sum j is over all events observed in the box and N_i is the number of events corresponding to the ith background component. The functions G describe normally distributed penalty terms for the parameters α_{ik} which reflect previous measurements performed in the independent NR11 dataset, as described in section 10.5.2.

For each box, the parameters describing the SM backgrounds are measured in an extended and unbinned ML fit using the likelihood of equation (10.18), performed in the fit region of that box. The measured parameter values, along with the covariance matrix of their errors, are then extrapolated to the whole razor plane, giving a full description of the SM backgrounds.

The choice of parameterization (number and type of background component functions F_i) follows from studies of the NR11 dataset and fits to earlier subsets of the R11 signal datasets. We observe that each of the SM backgrounds populating the final state boxes is well described by two components, and the shape of the second, flatter, component, is similar not only between final states but also between different background processes. The result is that fits with multiple second components

from different background processes are unable to distinguish between the multiple contributions. The shape parameters converge to indistinguishably similar values for the different second components and their yields are fully anti-correlated between the different instances. This motivates combining these components into one *universal effective component* (UEC) in the fits, representing the sum of all backgrounds' second components.

The UEC in each of the final states is observed to be nearly identical, and is understood to correspond to a hard initial-state-radiation limit where strong emissions, reconstructed as jets, recoil against the CM system of the hard interaction at such velocity that the differences between different background processes with weakly interacting particles are negligible at the precision relevant to the background model. As a result, the razor variable distributions for these events are nearly identical for the UEC portion of V+jets and top backgrounds, and adopt a shape characteristic on the partonic luminosities as a function of CM energy, $\sqrt{\hat{s}}$. The similarities of the UEC shapes between different boxes is not explicitly enforced in the background model fits. It is exploited through the penalty terms appearing in the likelihood of equation (10.18), with the NR11 values and errors for some parameters. Some of these values are replaced with those from the R11 background fits to higher event yield boxes, benefitting from increased precision as a result of a larger fit sample.

In each of the boxes there is no explicit component included to model QCD multi-jet backgrounds. The possibility for these events contributing to yields in the razor boxes is evaluated in section C.3, where it is demonstrated that the yields of these events is negligible and can be safely absorbed into other background components.

It is important to note that the veracity of this strategy for building an analytic description of the SM backgrounds in the razor plane is not self-evident from the results of these predictions. Deviations between background predictions and observed yields are interpreted in the context of BSM physics event contributions, not as a confirmation or rejection of the background model. Rather, the accuracy of the background model and its corresponding errors is demonstrated through a series of self-contained studies which evaluate each step of the procedure for possible biases or

systematic uncertainties.

The razor phenomenology which motivates the analytic functional description of backgrounds, described in section 10.2, follows from measurements to a large collection of simulated event samples, with a complete list provided in section C.1. Each of these samples indicates that the two-dimensional background model used in this analysis should provide an excellent description of these background. This is confirmed in measurements to data control samples in both the NR11 and R11 datasets. While this model will not hold to arbitrary precision (particularly in the consolidation of multiple backgrounds into single components), simulated data samples with more than 2000 times the integrated luminosity of those considered here indicate that it is accurate well within the required tolerance, as explained in section C.4.

Potential biases in the shapes of the razor variables following from lepton and b-tagging identification requirements are found to be negligible, as explained in section C.5, as are similar effects from PU interactions (section C.6) and trigger inefficiencies or choices for the fit region boundaries (section C.7). Possible deviations from exponential scaling behavior are discussed in section C.8 and the systematic uncertainty associated to the choice of functional form is found to be small.

Finally, the full background model procedure is performed on a cocktail of simulated background events with the expected composition and integrated luminosity of that selected in data. This exercise is described in section C.9 and constitutes a closure test of the procedure by successfully predicting the correct background shape and yield in the signal sensitive region using the subset of events in the fit region.

10.6.1 Background Model of Inclusive Samples

In the following sections, the composition and results of each fit region ML background fit are described for each of the inclusive dataset boxes. A more expansive discussion, with additional figures, are included for the MU box. A complete collection of auxiliary plots detailing these fits is included in section C.10 while the list of fit parameters determined to describe each of the boxes is tabulated in section C.11.

10.6.1.1 Inclusive MU Box

The background model in the MU box has contributions from three distinct compo-
nents (instances of $F_i(M_R, R^2)$). Two separate first components model the two largest
processes in this final state, $W(\mu\nu)$+jets and $t\bar{t}$+jets, while one UEC models the sum
of all backgrounds' second components.

Penalty terms derived from the NR11 dataset are added to the likelihood for each
of the shape parameters corresponding to each of the background components, with
the exception of the UEC exponential slope parameter k. For this parameter, we
instead float its value in the ML fit freely. The parameters corresponding to the
normalization of each background component are also floated without constraints.
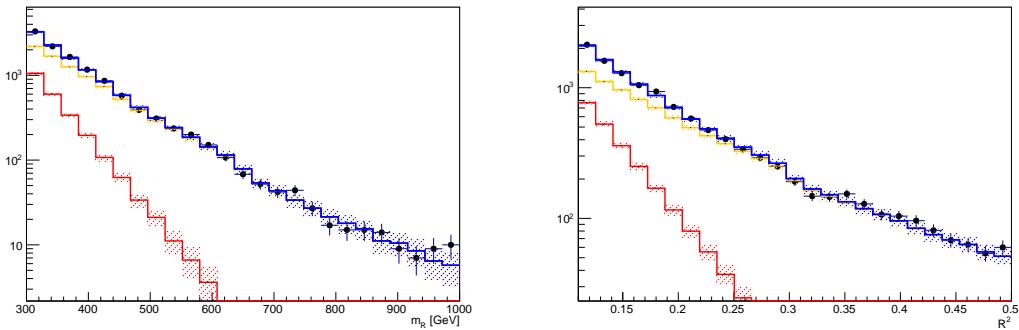

Figure 10.28: 1D projections of 2D ML fit in the inclusive MU box for M_R (Left) and
R^2 (Right) in the fit region. The total background prediction is indicated in blue.
The yellow contour corresponds to the UEC component combined with the first $t\bar{t}$
component. The V+jets first component contribution is indicated in red. Uncer-
tainty bands on each contribution only include statistical fluctuations, not systematic
uncertainties on the function parameters.

The 1D M_R and R^2 projections of the 2D in the MU box fit region are shown
In figure 10.28. Subtle discontinuities in the data sample and background prediction
reflect the saw-like shape of the fit region. We observe that the background model is
able to describe the event yields in the fit region. As a metric to identify potential
biases in the fit region description of the background model we consider two dimen-
sional binned histograms calculated to give the difference, bin-by-bin, between the
integral of the background model and the observed yield, shown for the MU box in

figure 10.29. Large areas with predicted/observed differences all in the same direction can indicate a systematic bias of the shape in the fit region. For the samples considered in this analysis we observe no such bias.
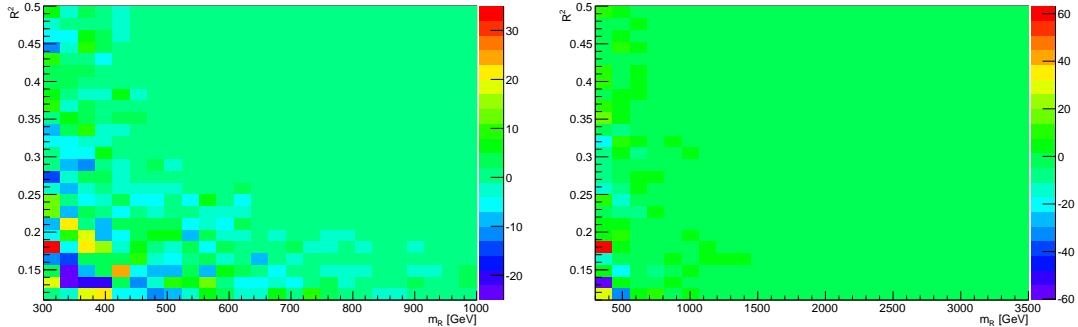

Figure 10.29: Bin-by-bin difference between the number of observed data events and the integral of the background model in the inclusive MU box for the fit region (Left) and the full razor plane (Right). Difference is in absolute number of events.

The full background model is extrapolated from the fit region to the full razor plane. This extrapolation includes not only the central values of the shape parameters and normalizations (adjusted to correspond to the full signal sensitive region) but also the covariance matrix from the ML fit which includes the errors for each parameter and the correlations between them. The one-dimensional M_R and R^2 projections of the full MU box background model are shown in figure 10.30.

In order to quantify the agreement between the data and the background prediction in the MU box, we use the background model to generate a set of toy pseudodatasets, corresponding to hypothetical outcomes of our experiment, with the frequency predicted by the background model. The covariance matrix is interpreted as a multinormal distribution corresponding to the pdf of all the background model parameters. For each toy experiment, a new set of parameters is chosen from this pdf. From this new background model, a random pseudodataset of background events is generated. This ensemble of pseudodatasets are used to calculate the expected yield distribution for each of the signal regions in the MU box, incorporating errors on the background model through this marginalization procedure.
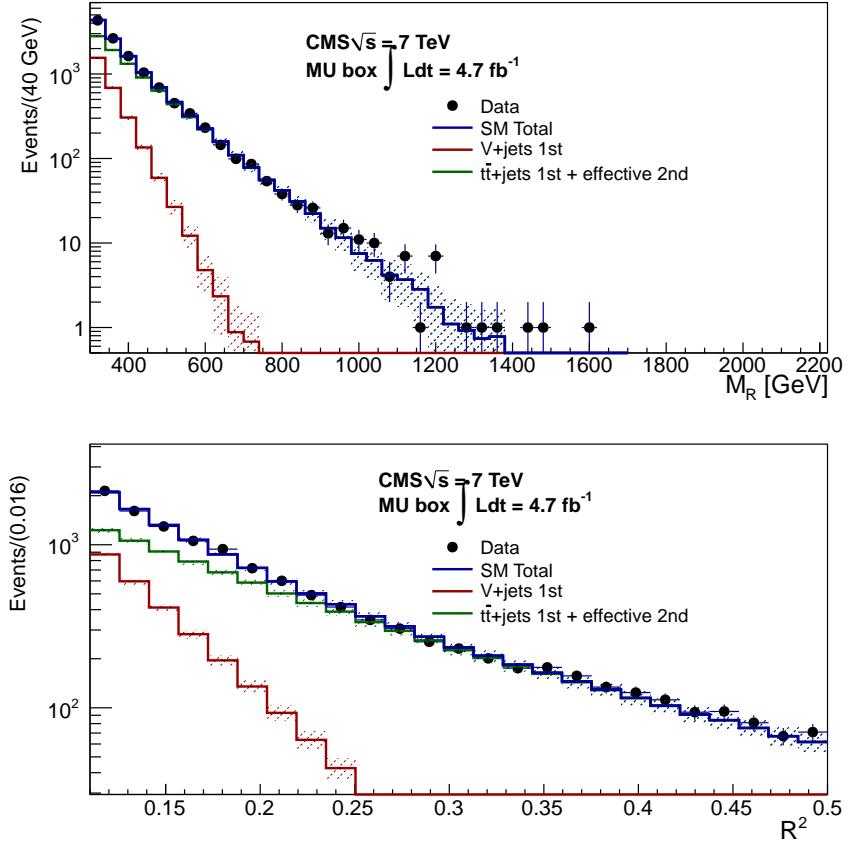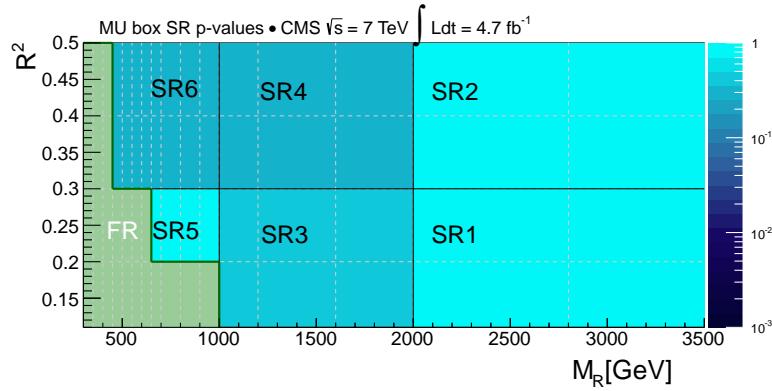
Figure 10.30: Projection of the 2D fit result on M_R (Top) and R^2 (Bottom) for the inclusive MU box. The blue histogram is the total Standard Model prediction from the shape parameter central values from the 2D fit. The green contour represents the combined UEC and $t\bar{t}$+jets first component. The V+jets first component contribution is shown red. The fit is performed the R^2-M_R fit region and extrapolated into the full razor plane. The error band on the contours includes only contributions from statistical fluctuations.

These distributions of expected yields for each of the signal regions is used to quantify the probability of observing an outcome less probable than that which was observed in the actual data, or a *p-value*. For each signal region, a two-sided *p*-value is calculated as

$$p - \text{value} = \sum_{p(n) \leq p(\text{data})} p(n) , \qquad (10.19)$$

where the sum over probability includes only the yield outcomes less likely than the one observed. The *p*-values calculated for the observed yields in each of the MU box signal regions is shown in figure 10.31.

MU	68% range	mode	median	observed	p-value
SR1	(0, 0.7)	0.5	0.5	0	0.99
SR2	(0, 0.7)	0.5	0.5	0	0.99
SR3	(19, 36)	30	29	35	0.44
SR4	(0, 1.8)	0.5	1.5	2	0.26
SR5	(67, 95)	88	82	74	0.76
SR6	(240, 318)	288	281	314	0.30

Figure 10.31: The *p*-values indicating the statistical consistency of the number of observed events in each of the inclusive MU box signal regions (SR_i) with predictions from the background model. The fit region is shown in green while z-axis color scale indicates *p*-value. The median and the mode of the yield distribution for each SR is quoted, together with the observed yield. A 68% probability interval for the expected yield is calculated, using the probability associated with each yield outcome as the ordering principle.

The *p*-values in these signal regions reflect a *model independent* interpretation of these results in that they only indicate the probability of this observed outcome in the context of the background model, without interpreting any deviations as other

than statistical fluctuations. The implications of these results for the parameters of hypothetical SUSY models are discussed in section 10.7.

10.6.1.2 Inclusive ELE Box

The SM background model in the inclusive ELE box, like the MU box, includes three contributions: a V+jets first component (predominantly $W(e\nu)$+jets), a $t\bar{t}$+jets first component and a UEC reflecting the combination of each background's second component. Penalty terms are included in the fit likelihood for each of the background shape parameters except for the slope parameter k of the UEC, which is floated without constraint.

The 1D M_R and R^2 projections of the 2D fit in the full razor plane are shown in figure 10.32. The p-values quantifying the agreement between the background model prediction and observed event yields in the ELE box signal regions are shown in figure 10.33.

The UEC slope parameter k is floated in the MU and ELE boxes since they are well-suited for better constraining its value than the NR11 dataset, with large statistics samples and a larger fit region than the HAD box. The values of the UEC component fit parameters for the MU and ELE boxes are compared in Tab. 10.3.

UEC Parameter	MU box Fit	ELE box Fit
k [GeV^{-1}]	0.0189 ± 0.0020	0.0138 ± 0.0030
M_R^0 [GeV]	43 ± 29	40 ± 28
R_0^2	-0.275 ± 0.028	-0.274 ± 0.027

Table 10.3: Comparison of the shape parameters for the second component from the ELE and the MU fit.

While the parameters show reasonable agreement, a difference between the two slope parameters k values is observed at the level of one standard deviation. This difference could result from a statistical fluctuation but, to be conservative, it is interpreted as a residual consequence of the different treatment of electrons and muons in the razor mega-jets construction. The MU-like and ELE-like UEC shapes, as
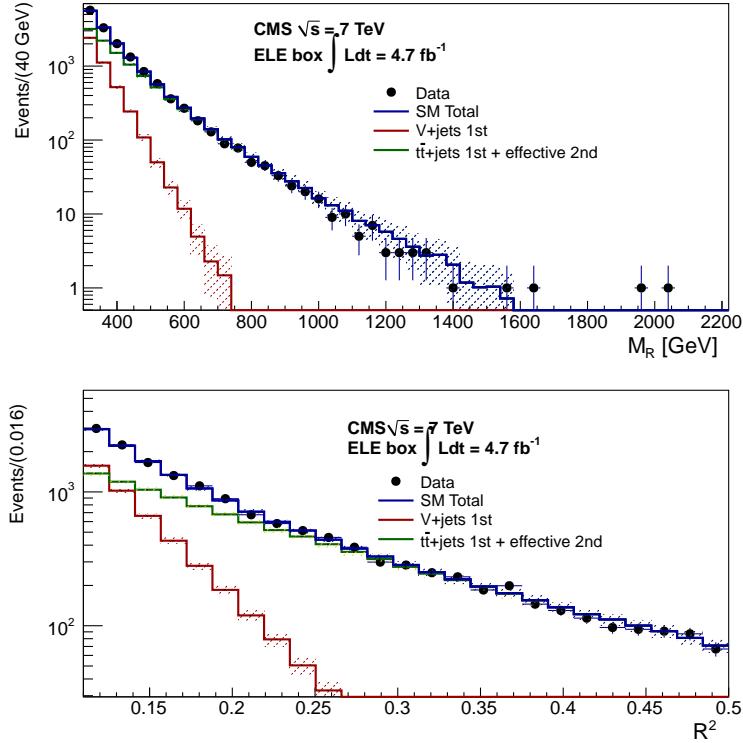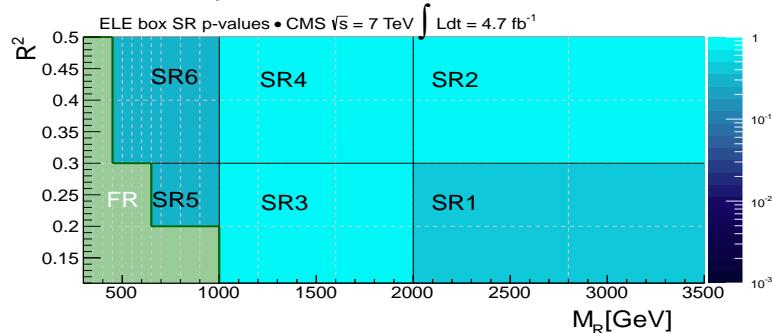
Figure 10.32: Projection of the 2D fit result on M_R (Top) and R^2 (Bottom) for the inclusive ELE box. The blue contour indicates total SM background predictions from fit model. The green contour represents the combined UEC and $t\bar{t}$+jets first component. The V+jets first component contribution is shown red. The error band on the contours includes only contributions from statistical fluctuations.

ELE	68% range	mode	median	observed	p-value
SR1	(0, 1.2)	0.5	0.5	1	0.37
SR2	(0, 0.7)	0.5	0.5	0	0.99
SR3	(36, 94)	50	66.5	51	0.74
SR4	(0, 8.5)	2.5	5.5	3	0.79
SR5	(98, 141)	122	119	98	0.29
SR6	(308, 412)	370	359	316	0.33

Figure 10.33: The p-values for the SRs of the inclusive ELE box. z-axis color scale indicates p-value. The median, mode and 68% probability interval of the yield distribution for each SR is quoted, together with the observed yield.

determined from these fits, are used in the penalty terms multiplying the likelihood for the other boxes.

10.6.1.3 Inclusive MU-MU Box

The SM background model in the inclusive MU-MU box includes three contributions: a V+jets first component (predominantly $Z(\mu\mu)$+jets), a $t\bar{t}$+jets first component and a UEC reflecting the combination of each background's second component. Penalty terms are included in the fit likelihood for each of the background shape parameters, with the UEC component penalty parameters are taken from the MU box fit.

The 1D M_R and R^2 projections of the 2D fit in the full razor plane are shown in figure 10.34. The p-values quantifying the agreement between the background model prediction and observed event yields in the ELE box signal regions are shown in figure 10.35.

10.6.1.4 Inclusive MU-ELE Box

The SM background model in the inclusive ELE-MU box includes two contributions: an effective first component and a UEC. Here, the background is composed almost exclusively of $t\bar{t}$+jets events, such that only two components are required. Penalty terms are included in the fit likelihood for each of the background shape parameters, with the UEC component penalty parameters taken from the MU box fit.

The 1D M_R and R^2 projections of the 2D fit in the full razor plane are shown in figure 10.36. The p-values quantifying the agreement between the background model prediction and observed event yields in the ELE box signal regions are shown in figure 10.37.

10.6.1.5 Inclusive ELE-ELE Box

The SM background model in the inclusive ELE-ELE box includes three contributions: a V+jets first component (predominantly $Z(ee)$+jets), a $t\bar{t}$+jets first component and a UEC reflecting the combination of each background's second component.
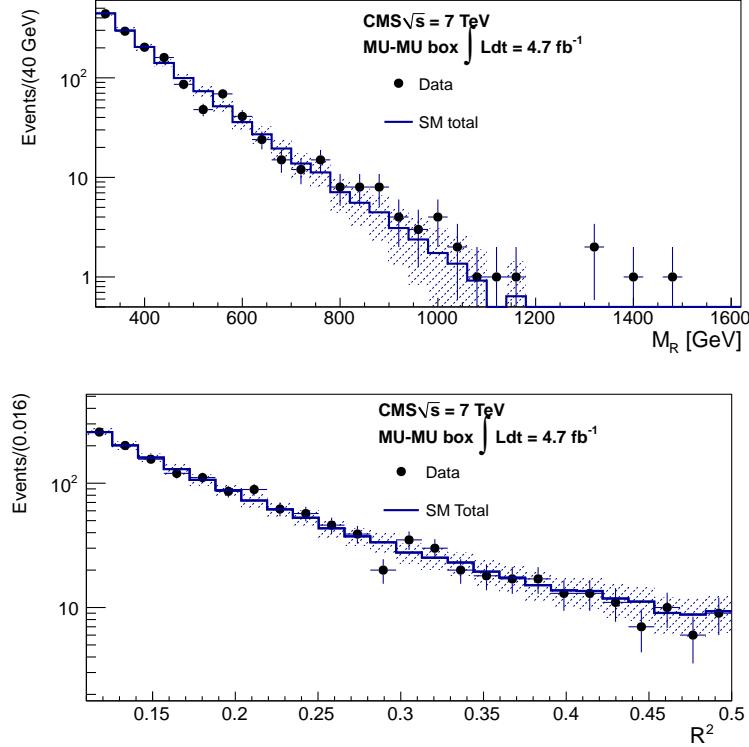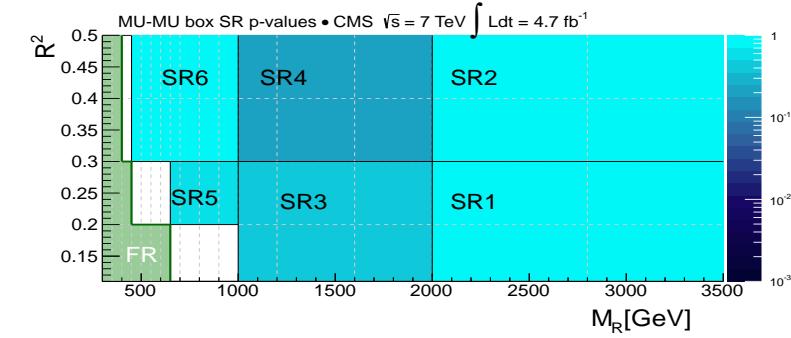
Figure 10.34: Projection of the 2D fit result on M_R (Top) and R^2 (Bottom) for the MU-MU box. The blue contour indicates total SM background predictions from fit model. The error band on the contours includes only contributions from statistical fluctuations.

MU-MU	68% range	mode	median	observed	p-value
SR1	(0, 0.7)	0.5	0.5	0	0.99
SR2	(0, 0.7)	0.5	0.5	0	0.99
SR3	(1, 10)	5.5	7.5	10	0.36
SR4	(0, 0.9)	0.5	0.5	1	0.20
SR5	(7.5, 20)	13.5	14.5	16	0.61
SR6	(26, 56)	44.5	42.5	37	0.79

Figure 10.35: The p-values for the SRs of the MU-MU box. z-axis color scale indicates p-value. The median, mode and 68% probability interval of the yield distribution for each SR is quoted, together with the observed yield.
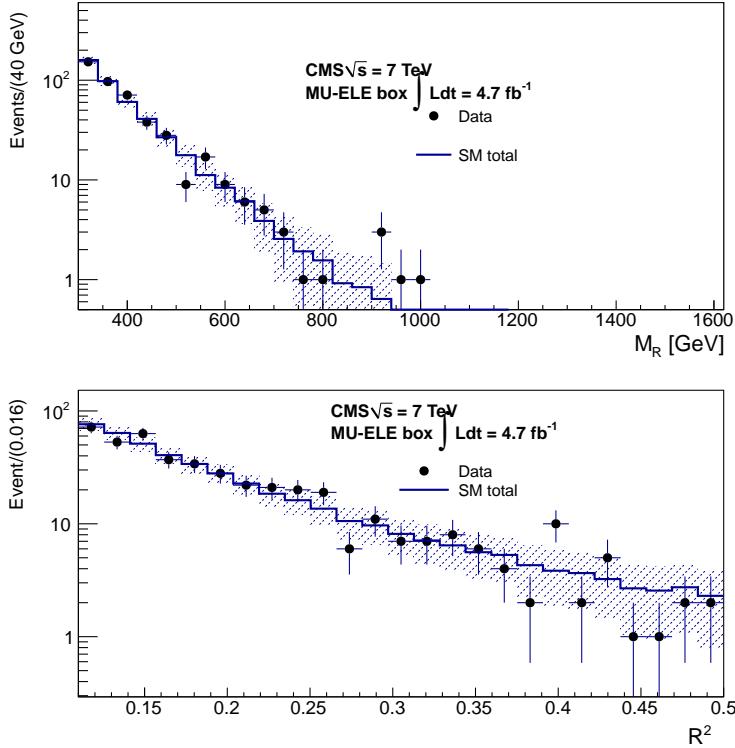
Figure 10.36: Projection of the 2D fit result on M_R (Top) and R^2 (Bottom) for the ELE-MU box. The blue contour indicates total SM background predictions from fit model. The error band on the contours includes only contributions from statistical fluctuations.

MU-ELE	68% range	mode	median	observed	p-value
SR1	(0, 0.7)	0.5	0.5	0	0.99
SR2	(0, 0.7)	0.5	0.5	0	0.99
SR3	(0, 2.1)	0.5	1.5	1	0.66
SR4	(0, 0.7)	0.5	0.5	0	0.99
SR5	(0.6, 5)	2.5	3.5	5	0.33
SR6	(5.2, 15.2)	10.5	11.5	11	0.82

Figure 10.37: The p-values for the SRs of the ELE-MU box. z-axis color scale indicates p-value. The median, mode and 68% probability interval of the yield distribution for each SR is quoted, together with the observed yield.
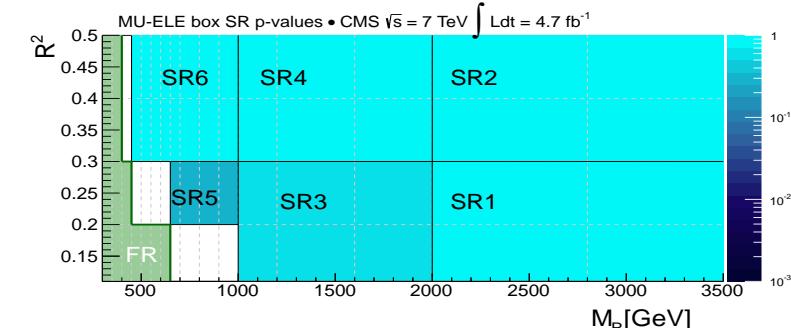
Penalty terms are included in the fit likelihood for each of the background shape parameters, with the UEC component penalty parameters taken from the ELE box fit.

The 1D M_R and R^2 projections of the 2D fit in the full razor plane are shown in figure 10.38. The p-values quantifying the agreement between the background model prediction and observed event yields in the ELE box signal regions are shown in figure 10.39.

10.6.1.6 Inclusive HAD Box

The background model for the inclusive HAD box is unique relative to the other boxes due to the larger variety of backgrounds appearing. $Z(\nu\nu)$+jets events constitute a relatively large contribution, with kinematics similar to the $W(\mu\nu)$+jets events which also appear when the muon is not identified as such. There are backgrounds with unidentified electrons and hadronic taus resulting from leptonic W decays, with kinematics like the backgrounds of the ELE box. Of course, there are also analogous backgrounds with top quarks, with similar W decays.

These many background contributions are consolidated into several components which are used to model the HAD box background shape, combining kinematically compatible backgrounds. These components include

- An effective first component, whose shape parameters are floated without constraint in the fit, which combines the residual first components of each of the SM backgrounds. This single component provides an adequate description of several contributions which cannot be individually resolved in the more restrictive fit region of the HAD box, resulting from tighter razor trigger requirements.

- A MU-like UEC, with shape parameters constrained with penalty terms in the likelihood with values from the fit to the MU box.

- An ELE-like UEC, with shape parameters constrained with penalty terms in the likelihood with values from the fit to the ELE box.
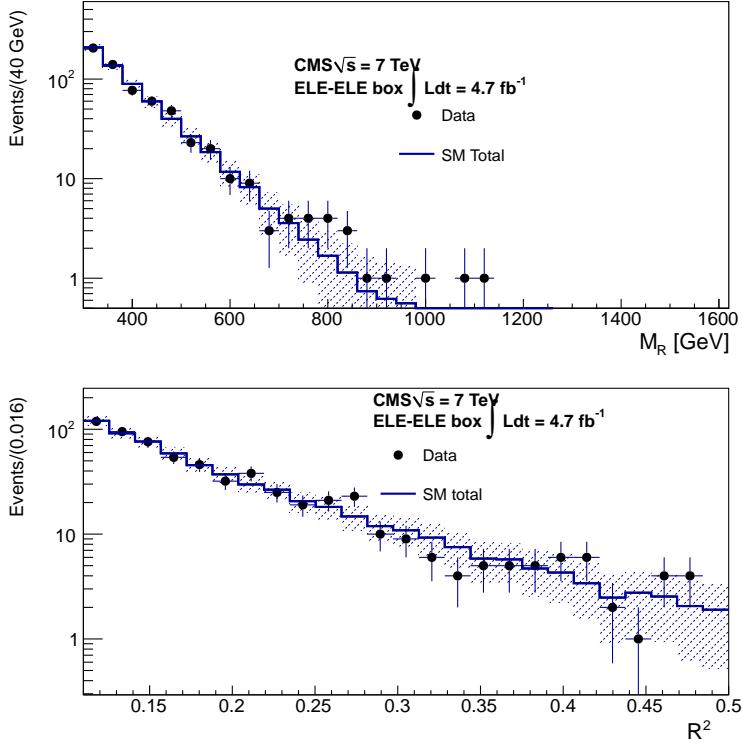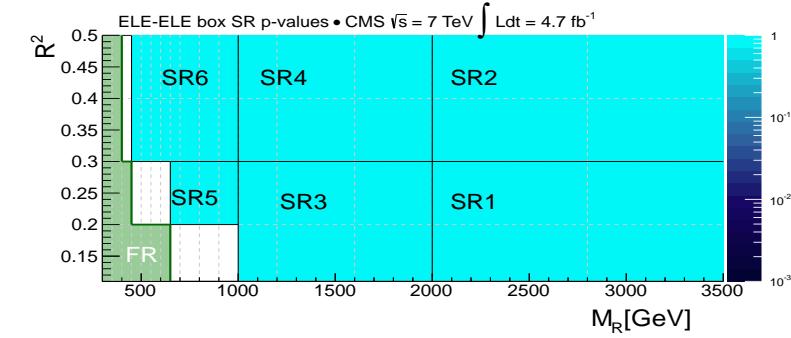
Figure 10.38: Projection of the 2D fit result on M_R (Top) and R^2 (Bottom) for the ELE-ELE box. The blue contour indicates total SM background predictions from fit model. The error band on the contours includes only contributions from statistical fluctuations.

ELE-ELE	68% range	mode	median	observed	p-value
SR1	(0, 0.7)	0.5	0.5	0	0.99
SR2	(0, 0.7)	0.5	0.5	0	0.99
SR3	(0, 6.8)	1.5	4.5	2	0.86
SR4	(0, 1)	0.5	0.5	0	0.99
SR5	(0.7, 11)	3.5	7.5	5	0.83
SR6	(4.4, 32)	15.5	20.5	17	0.79

Figure 10.39: The p-values for the SRs of the ELE-ELE box. z-axis color scale indicates p-value. The median, mode and 68% probability interval of the yield distribution for each SR is quoted, together with the observed yield.

All the components' normalizations are floated in the fit without external constraints.

The 1D M_R and R^2 projections of the 2D fit in the full razor plane are shown in figure 10.40. The p-values quantifying the agreement between the background model prediction and observed event yields in the ELE box signal regions are shown in figure 10.41.

10.6.2 Background Model of b-tagged Samples

In the following sections, the composition and results of each fit region ML background fit are described for each of the b-tagged dataset boxes. A complete collection of auxiliary plots detailing these fits is included in section C.10.

10.6.2.1 b-tagged MU Box

As for the inclusive dataset analysis, the MU box fit is performed first in order to derive new, more precise constraints for the UEC penalty terms for the other boxes (relative to the NR11 measurements).

The background model in the MU box has contributions from three distinct instances of $F_i(M_R, R^2)$. Two separate first components model the two largest processes in this final state, $W(\mu\nu)$+jets and $t\bar{t}$+jets, while one UEC models the sum of all backgrounds' second components.

Penalty terms derived from the NR11 dataset are added to the likelihood for each of the shape parameters corresponding to each of the background components, with the exception of the UEC exponential slope parameter k. For this parameter, we instead float its value in the ML fit freely. The parameters corresponding to the normalization of each background component are also floated without constraints.

The 1D M_R and R^2 projections of the 2D fit in the full razor plane are shown in figure 10.42. The p-values quantifying the agreement between the background model prediction and observed event yields in the ELE box signal regions are shown in figure 10.45.

An excess of events is observed in the S_6 region, corresponding to about two
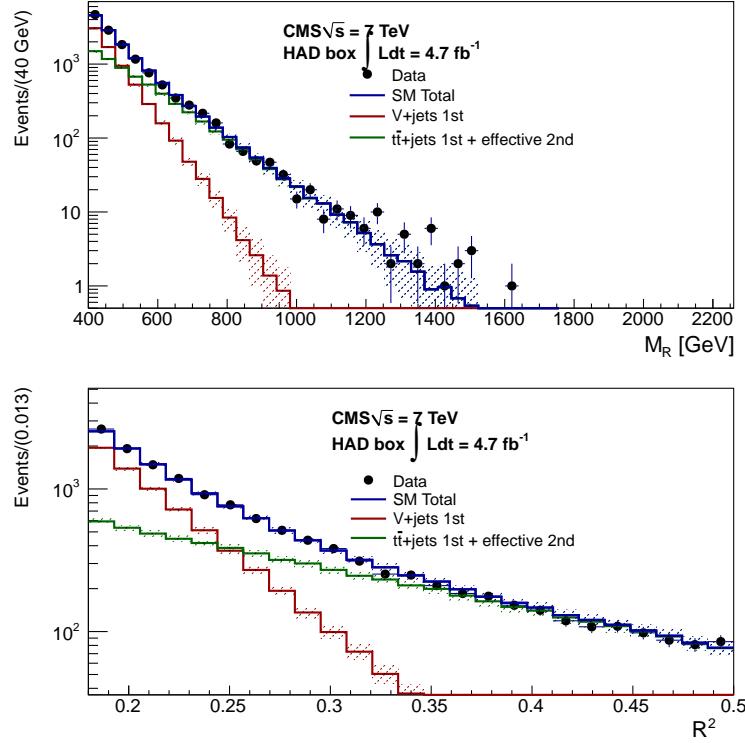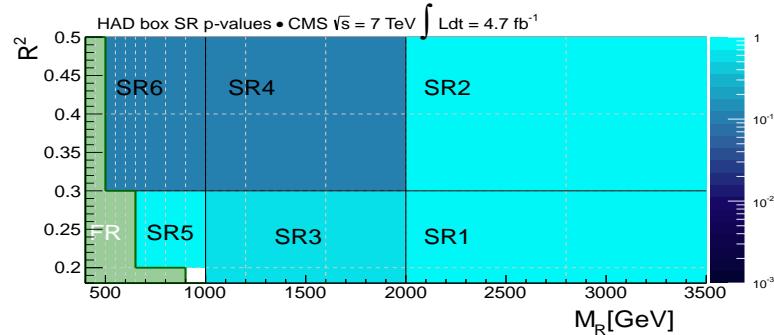
Figure 10.40: Projection of the 2D fit result on M_R (Top) and R^2 (Bottom) for the inclusive HAD box. The blue contour indicates total SM background predictions from fit model. The green contour represents the combined ELE and MU-like UECs. The effective first component contribution is shown red. The error band on the contours includes only contributions from statistical fluctuations.

HAD	68% range	mode	median	observed	p-value
SR1	(0, 0.7)	0.5	0.5	0	0.99
SR2	(0, 0.7)	0.5	0.5	0	0.99
SR3	(7, 62)	13.5	36.5	22	0.84
SR4	(0, 16)	3.5	9.5	6	0.65
SR5	(140, 224.5)	190.5	183.5	170	0.99
SR6	(198, 397)	311.5	290.5	186	0.14

Figure 10.41: The p-values for the SRs of the inclusive HAD box. z-axis color scale indicates p-value. The median, mode and 68% probability interval of the yield distribution for each SR is quoted, together with the observed yield.

Figure 10.42: Projection of the 2D fit result on M_R (Top) and R^2 (Bottom) for the b-tagged MU box. The blue contour indicates total SM background predictions from fit model. The yellow contour represents the combined UEC and $t\bar{t}$+jets first component. The V+jets first component contribution is shown red. The error band on the contours includes only contributions from statistical fluctuations.

MU	68% range	mode	median	observed	p-value
SR1	(0, 0.7)	0.5	0.5	0	0.99
SR2	(0, 0.7)	0.5	0.5	0	0.99
SR3	(3, 11.1)	8.5	8.5	9	0.61
SR4	(0, 0.9)	0.5	0.5	0	0.99
SR5	(12.3, 26.5)	19.5	20.5	19	0.99
SR6	(46, 78.8)	58.5	64.5	95	0.07

Figure 10.43: The p-values for the SRs of the b-tagged MU box. z-axis color scale indicates p-value. The median, mode and 68% probability interval of the yield distribution for each SR is quoted, together with the observed yield.
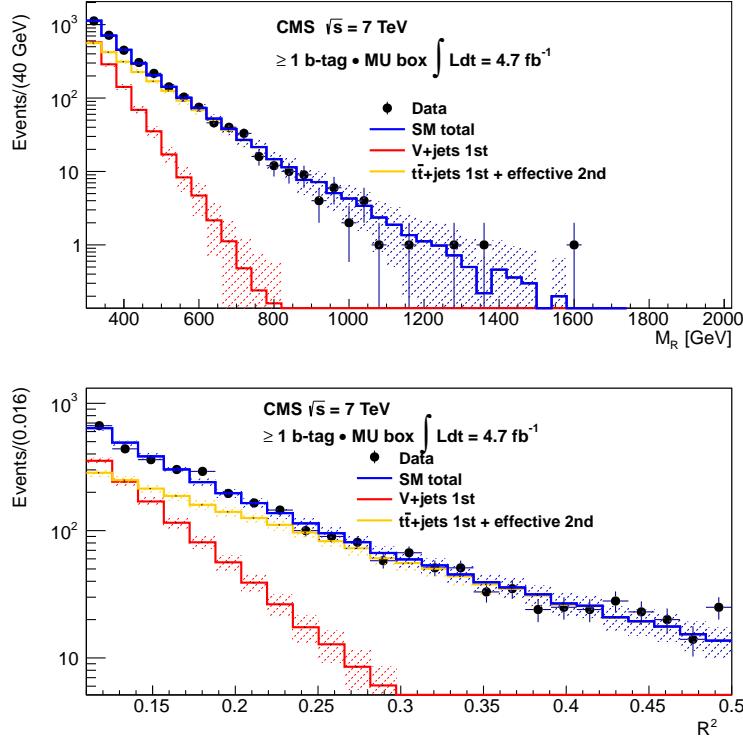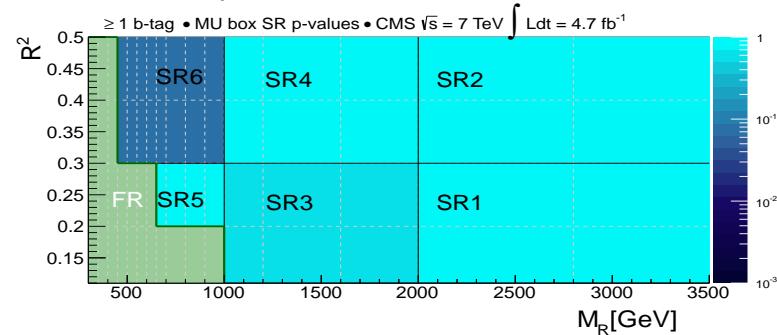
standard deviations relative to the background mode. This is the smallest p-value observed in the leptonic boxes and does not coincide with similar observations in the S_6 region of any other box, indicating that it is consistent with a statistical fluctuation.

10.6.2.2 b-tagged ELE Box

The SM background model in the b-tagged ELE box, like the MU box, in includes three contributions: a V+jets first component (predominantly $W(e\nu)$+jets), a $t\bar{t}$+jets first component and a UEC reflecting the combination of each background's second component. Penalty terms are included in the fit likelihood for each of the background shape parameters, with penalty parameters for the UEC coming from the MU box fit values.

The1D M_R and R^2 projections of the 2D fit in the full razor plane are shown in figure 10.44. The p-values quantifying the agreement between the background model prediction and observed event yields in the ELE box signal regions are shown in figure 10.45.

10.6.2.3 b-tagged MU-ELE Box

The SM background model in the b-tagged MU-ELE box includes two contributions: an effective first component and UEC second component. The majority of background events are $t\bar{t}$, allowing us to model the total background as we would an isolated top sample. Penalty terms are included in the fit likelihood for each of the background shape parameters, with penalty parameters for the UEC coming from the MU box fit values.

The 1D M_R and R^2 projections of the 2D fit in the full razor plane are shown in figure 10.46. The p-values quantifying the agreement between the background model prediction and observed event yields in the ELE box signal regions are shown in figure 10.47.
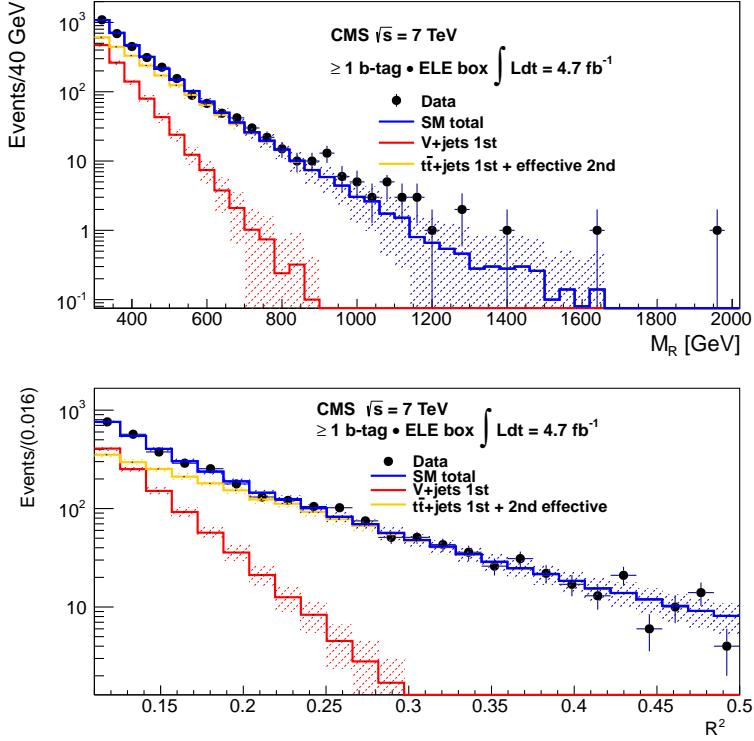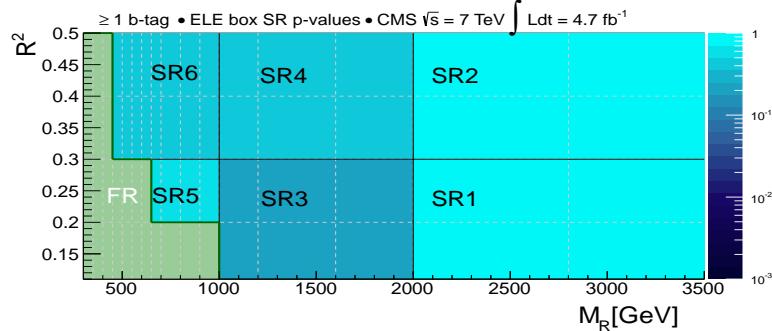
Figure 10.44: Projection of the 2D fit result on M_R (Top) and R^2 (Bottom) for the b-tagged ELE box. The blue contour indicates total SM background predictions from fit model. The yellow contour represents the combined UEC and $t\bar{t}$+jets first component. The V+jets first component contribution is shown red. The error band on the contours includes only contributions from statistical fluctuations.

ELE	68% range	mode	median	observed	p-value
SR1	(0, 1.2)	0.5	0.5	1	0.37
SR2	(0, 0.7)	0.5	0.5	0	0.99
SR3	(3.1, 19)	9.5	12.5	22	0.25
SR4	(0, 2.0)	0.5	0.5	1	0.46
SR5	(10, 25.3)	16.5	19.5	20	0.62
SR6	(24.8, 60)	35.5	46.5	55	0.41

Figure 10.45: The p-values for the SRs of the b-tagged ELE box. z-axis color scale indicates p-value. The median, mode and 68% probability interval of the yield distribution for each SR is quoted, together with the observed yield.

Figure 10.46: Projection of the 2D fit result on M_R (Top) and R^2 (Bottom) for the b-tagged MU-ELE box. The blue contour indicates total SM background predictions from fit model. The error band on the contours includes only contributions from statistical fluctuations.

MU-ELE	68% range	mode	median	observed	p-value
SR1	(0, 0.8)	0.5	0.5	0	0.99
SR2	(0, 0.7)	0.5	0.5	0	0.99
SR3	(0, 3.8)	0.5	1.5	1	0.68
SR4	(0, 1.0)	0.5	0.5	0	0.99
SR5	(0, 4)	1.5	2.5	3	0.45
SR6	(0.7, 10.4)	3.5	7.5	5	0.82

Figure 10.47: The p-values for the SRs of the b-tagged MU-ELE box. z-axis color scale indicates p-value. The median, mode and 68% probability interval of the yield distribution for each SR is quoted, together with the observed yield.
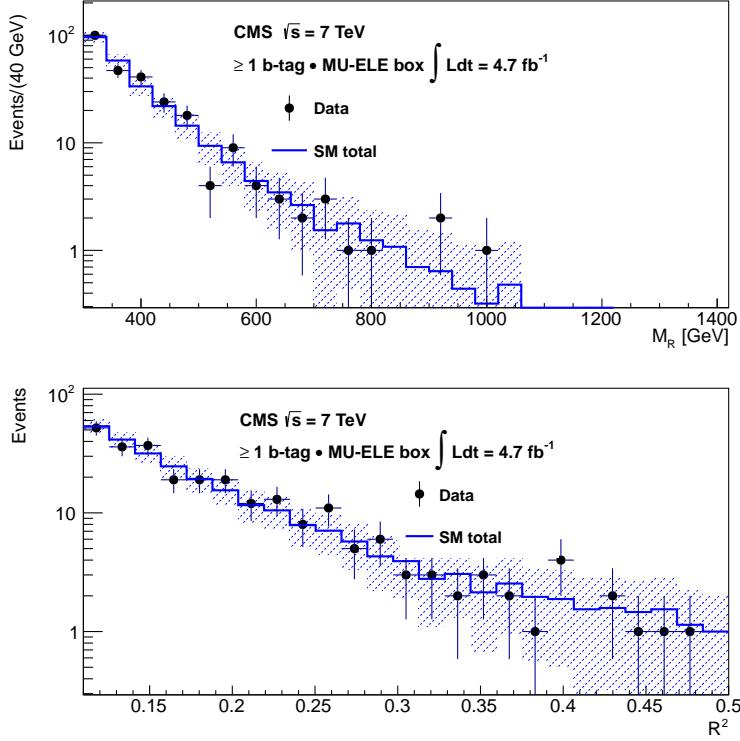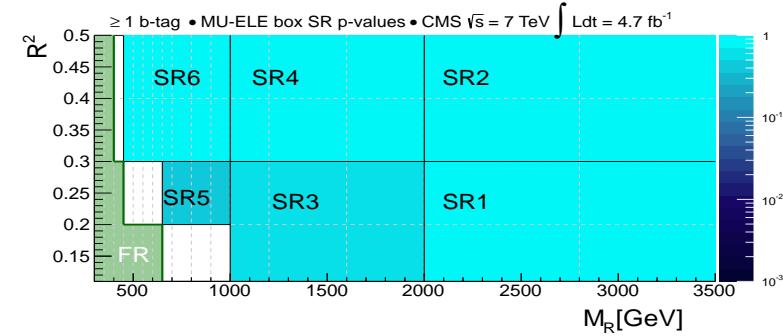
10.6.2.4 b-tagged MU-MU Box

The SM background model in the b-tagged MU-MU box is simplified relative to the inclusive analysis since the background from $Z(\mu\mu)$+jets events is effectively removed with the b-tagged jet requirement. As a result, the background model includes two contributions: an effective first component and UEC second component. The majority of background events are $t\bar{t}$, allowing us to model the total background as we would an isolated top sample. Penalty terms are included in the fit likelihood for each of the background shape parameters, with penalty parameters for the UEC coming from the MU box fit values.

The 1D M_R and R^2 projections of the 2D fit in the full razor plane are shown in figure 10.48. The p-values quantifying the agreement between the background model prediction and observed event yields in the ELE box signal regions are shown in figure 10.49.

10.6.2.5 b-tagged ELE-ELE Box

The SM background model in the b-tagged ELE-ELE box, like the MU-MU, has a simplified configuration due to the abundance of a single background: an effective first component and UEC second component. Penalty terms are included in the fit likelihood for each of the background shape parameters, with penalty parameters for the UEC coming from the MU box fit values.

The 1D M_R and R^2 projections of the 2D fit in the full razor plane are shown in figure 10.50. The p-values quantifying the agreement between the background model prediction and observed event yields in the ELE-ELE box signal regions are shown in figure 10.51.

10.6.2.6 b-tagged HAD Box

The SM background model in the b-tagged HAD box includes three contributions: a V+jets first component, a $t\bar{t}$+jets first component and a UEC reflecting the combination of each background's second component. Penalty terms are included in the fit
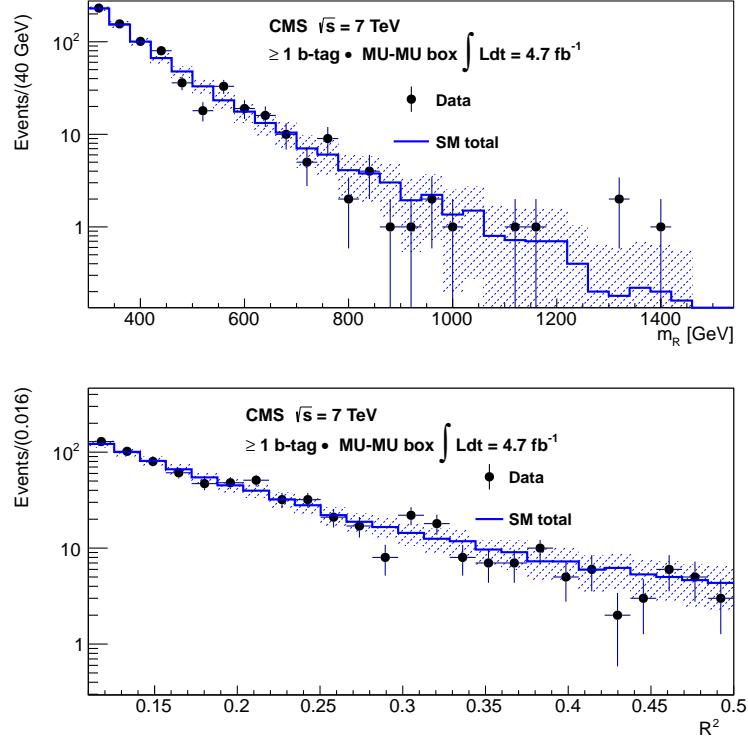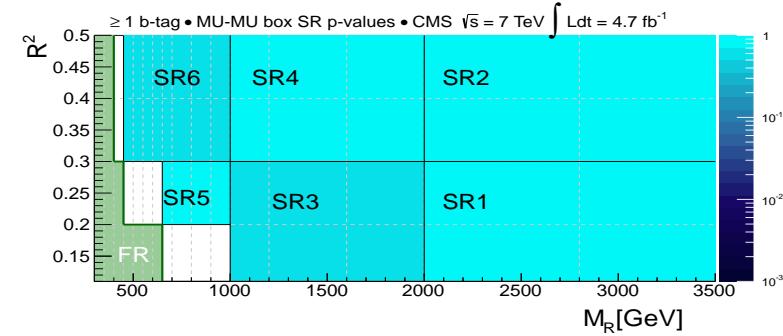
Figure 10.48: Projection of the 2D fit result on M_R (Top) and R^2 (Bottom) for the b-tagged MU-MU box. The blue contour indicates total SM background predictions from fit model. The error band on the contours includes only contributions from statistical fluctuations.

MU-MU	68% range	mode	median	observed	p-value
SR1	(0, 0.7)	0.5	0.5	0	0.99
SR2	(0, 0.7)	0.5	0.5	0	0.99
SR3	(0, 12.5)	1.5	6.5	5	0.36
SR4	(0, 2.0)	0.5	0.5	0	0.99
SR5	(2.2, 16)	7.5	10.5	7	0.99
SR6	(10, 47)	19.5	28.5	13	0.53

Figure 10.49: The p-values for the SRs of the b-tagged MU-MU box. z-axis color scale indicates p-value. The median, mode and 68% probability interval of the yield distribution for each SR is quoted, together with the observed yield.

Figure 10.50: Projection of the 2D fit result on M_R (Top) and R^2 (Bottom) for the b-tagged ELE-ELE box. The blue contour indicates total SM background predictions from fit model. The error band on the contours includes only contributions from statistical fluctuations.

ELE-ELE	68% range	mode	median	observed	p-value
SR1	(0, 0.8)	0.5	0.5	0	0.99
SR2	(0, 0.7)	0.5	0.5	0	0.99
SR3	(0, 2)	0.5	1.5	1	0.54
SR4	(0, 1)	0.5	0.5	0	0.99
SR5	(0, 2.4)	0.5	1.5	1	0.69
SR6	(0, 5.2)	2.5	3.5	6	0.26

Figure 10.51: The p-values for the SRs of the b-tagged ELE-ELE box. z-axis color scale indicates p-value. The median, mode and 68% probability interval of the yield distribution for each SR is quoted, together with the observed yield.
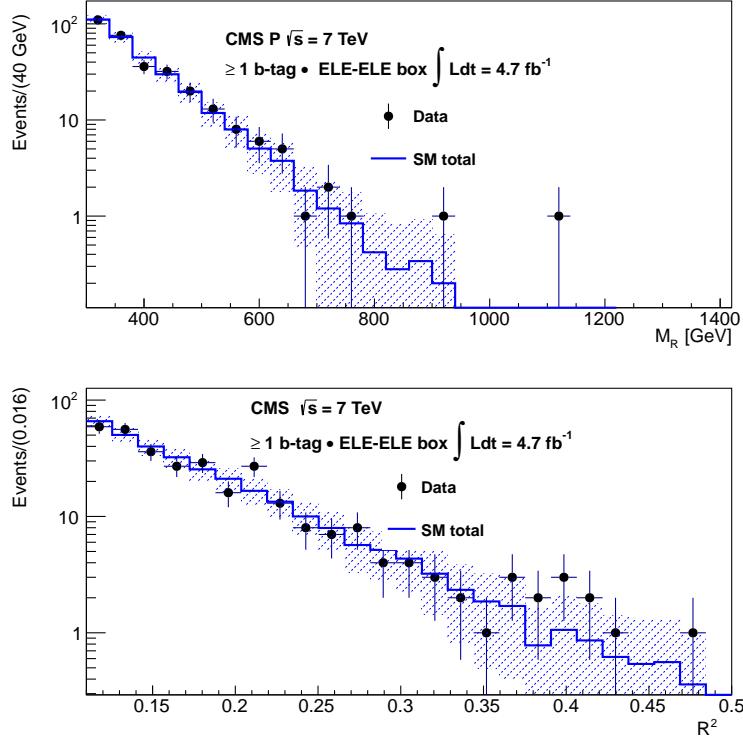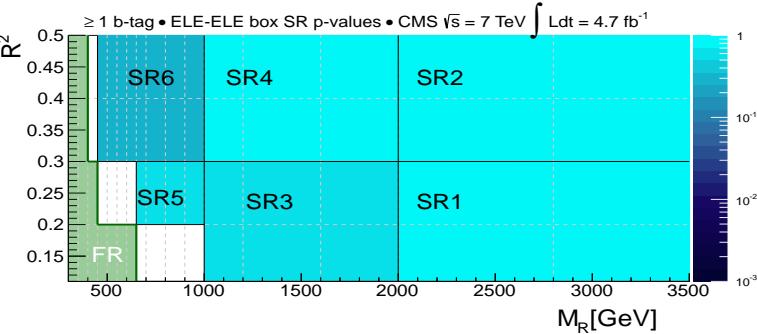
likelihood for each of the background shape parameters, with penalty parameters for the UEC coming from the MU box fit values, except for the effective first component which is floated freely in the fit, as in the inclusive HAD box.

The 1D M_R and R^2 projections of the 2D fit in the full razor plane are shown in figure 10.52. The p-values quantifying the agreement between the background model prediction and observed event yields in the HAD box signal regions are shown in figure 10.53. The component labeled V+jets is actually made of predominantly of $t\bar{t}$+jets events and, since it is floated freely, plays the role of the effective first component analogous to that in the inclusive HAD box. In fact, this component is largely redundant in shape with that labelled $t\bar{t}$+jets first component, such that this second component is largely absorbed into the former, with near zero event yield.
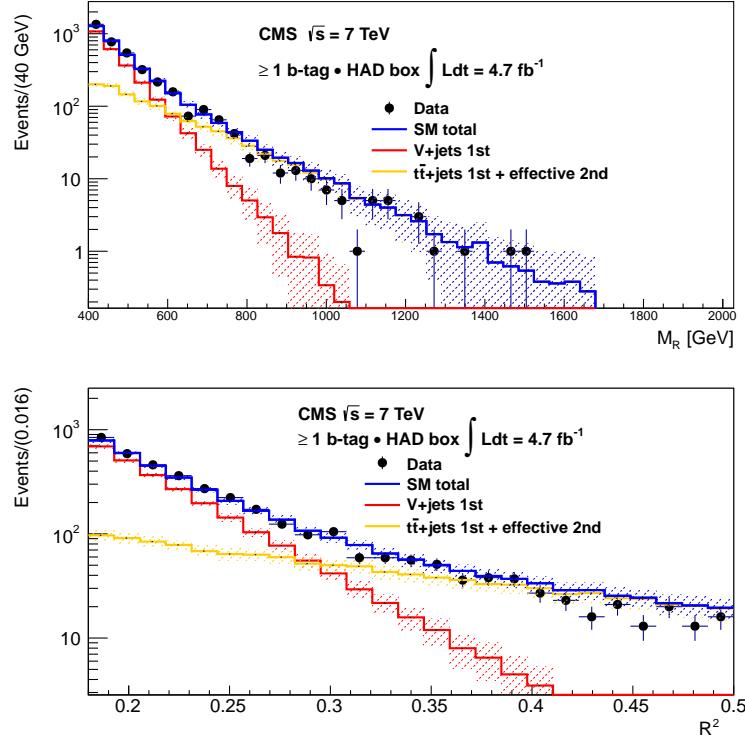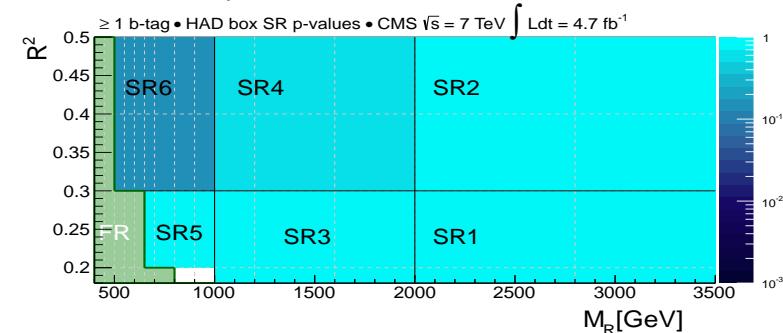
Figure 10.52: Projection of the 2D fit result on M_R (Top) and R^2 (Bottom) for the b-tagged HAD box. The blue contour indicates total SM background predictions from fit model. The yellow contour represents the combined UEC and $t\bar{t}$+jets first component. The V+jets first component contribution is shown red. The error band on the contours includes only contributions from statistical fluctuations.

HAD	68% range	mode	median	observed	p-value
SR1	(0, 0.7)	0.5	0.5	0	0.99
SR2	(0, 0.7)	0.5	0.5	0	0.99
SR3	(7, 62)	13.5	36.5	22	0.84
SR4	(0, 16)	3.5	9.5	6	0.65
SR5	(140, 224.5)	190.5	183.5	170	0.99
SR6	(198, 397)	311.5	290.5	186	0.14

Figure 10.53: The p-values for the SRs of the b-tagged HAD box. z-axis color scale indicates p-value. The median, mode and 68% probability interval of the yield distribution for each SR is quoted, together with the observed yield.

10.7 Model-Dependent Results Interpretation in SUSY Parameter Space

The model-dependent interpretation of the background model results, p-values reflecting the agreement between the predictions of the background model and event yields observed in data, indicate that predictions are in good agreement with the data collected in the full razor plane, in all of the final state boxes. This absence of a large deviation from expectations indicates that there is not significant evidence of BSM physics events in the R11 4.7 fb^{-1} 7 TeV data sample. Conversely, the agreement between predictions and expectations can be used to constrain the total number of new physics events in the dataset which are unaccounted for by the SM background model.

As described in section 10.2, signal events can appear in different places throughout the razor plane and in different final states, depending on the masses, decay interactions and branching fractions of new particles. For each new physics scenario, the data is queried to see if can accommodate its expected signature by comparing the likelihood of two hypotheses: the data is best described by the SM background model or the data is more similar to the expected background with a contribution from this new physics model. A strong preference for the latter could indicate the presence of new physics consistent with this particular scenario, while data consistent with the former can be used to exclude the presence of new physics events, and hence the possibility of that particular model describing nature.

The likelihood-based hypothesis testing procedure used to place constraints on hypothetical SUSY models is described in section 10.7.1. The results from the inclusive and b-tagged searches are interpreted in both the CMSSM (see section 8.2.1) and in a range of simplified models which include only specific sparticle production and decays, with constraints described in sections 10.7.3 and 10.7.4, respectively.

10.7.1 Model-Dependent Limit Constraints Setting Procedure

For each hypothetical SUSY scenario a conceptually simple question is asked: does the observed data spectrum fit better to a background only model, or does it indicate that a combination of background and signal events provides a better description? The answer is determined through a statistical analysis of the observed data yields based on evaluating the mathematical likelihoods of each of these two possibilities.

For each model for which these results are interpreted the two hypotheses, that signal events consistent with the expectations of this model are present in the collected data sample (denoted $\mathbb{H}_{s+b}$) or there is an absence of signal-like events (denoted $\mathbb{H}_b$), can be expressed as two different likelihoods. The likelihood that the SM background model alone describing the data in a particular box is given by

$$\mathcal{L}_b = \frac{e^{-(\sum_{i \in SM}^{N_{SM}} N_i)}}{N_{event}!} \prod_{j=1}^{N_{event}} \left[\sum_{i \in SM}^{N_{SM}} N_i F_i(M_R^j, R_j^2) \right] , \tag{10.20}$$

where the sum over N_{SM} corresponds to the different background components in the box, the sum over j is over all events in the signal region of the box and the shapes for the functions F_i are taken from the background model. Similarly, the likelihood that there are also signal events for a particular model in the data sample can be written

$$\mathcal{L}_{s+b}(\mathcal{M}) = \frac{e^{-(\sum N_i + N_S(\mathcal{M}))}}{N_{event}!} \prod_{j=1}^{N_{event}} \left[\sum_{i \in SM}^{N_{SM}} N_i F_i(M_R^j, R_j^2) + N_S(\mathcal{M}) P_S(M_R, R^2 | \mathcal{M}) \right] , \tag{10.21}$$

where N_S is the number of expected signal events for the model, P_S is the pdf describing the shape of these events in the razor plane, and $\mathcal{M}$ indicates the particular model being tested. The explicit dependence on the model $\mathcal{M}$ in the likelihood indicates that each is specific to a particular model. N_S is fixed according to the expected production cross section and branching ratios for the model $\mathcal{M}$ while $P_S(M_R, R^2)$ is also specific to that model. As a result, the conclusions drawn from each hypothesis

test are only relevant for a single model $\mathcal{M}$; excluding one SUSY scenario has no direct implication for the existence of another.

The specificity of each hypothesis test is a computational shortcoming but also a blessing; since the background is described analytically and continuously, these likelihood-ratio based hypothesis test will only consider events falling in the region of the razor plane consistent with events from the model $\mathcal{M}$, ignoring other background events. This can be understood from the likelihoods from the two hypotheses: When taking the ratio of the likelihoods, events that fall in regions of M_R/R^2 space where $P_S(M_R, R^2|\mathcal{M})$ is small will cancel in the two likelihood event products, automatically optimizing the region of the razor plane queried for the existence of that particular model. This means that only events which *look* like a particular model are considered as background to it, while the yields in the rest of the razor plane are ignored.

Signal pdfs P_S are described numerically using a binned two-dimensional histogram in the razor variables built from simulated signal events for a given model. In order to avoid discontinuities in the pdfs resulting from limited simulated event statistics, a variable binning scheme is implemented. In M_R, 50 GeV wide bins are defined starting from the minimal M_R for a given box, up to 700 GeV. The following bin edges are then defined,

$$[700, 800, 900, 1000, 1200, 1600, 2000, 2800, 3500] \ . \tag{10.22}$$

For R^2 bin edges are defined as,

$$[R^2_{min}, 0.2, 0.3, 0.4, 0.5] \ , \tag{10.23}$$

where R^2_{min} is the minimum value of R^2 in the box (0.11 for the leptonic boxes, 0.18 for the HAD boxes).

Constraints are placed on a hypothetical SUSY model $\mathcal{M}$ with the following procedure, performed for each box:

- Simulated signal events from $\mathcal{M}$ are used to build the nominal signal pdf P_S to

be used in the likelihood $\mathcal{L}_{s+b}(\mathcal{M})$.

- Systematic uncertainties on the shape and yield P_S and N_S, respectively, are included by building an ensemble of alternative signal pdfs, each including random variations reflecting these uncertainties. The type and magnitude of these systematic variations are described in section C.14.2. Log-normal distributions are used to model the expected event yield in each signal pdf bins, and random values are generated for each bin in each alternative pdf. For systematic error correlated across the R^2 vs. M_R plane these variations are performed coherently for all of the bins. Otherwise, variations are implemented independently.

- The covariance matrix from the background model fit is interpreted as a multi-normal pdf of the background shape parameters and yields. For each alternative signal pdf in the generated ensemble, an alternative background model pdf is created by sampling the background parameter pdf randomly. This procedure takes into account not only the errors for each parameter in the background model but also their correlations.

- Alternative signal and background pdfs are combined to create two collections of pseudodatasets, one corresponding to outcomes where there are only background events in the pseudodata sample, and the other with both signal and background events. The first is created by randomly generating a pseudodataset from each alternative background pdf. The second collection is created similarly, except adding signal events randomly generated from an alternative signal pdf to each pseudodataset.

- The likelihoods of the two different hypotheses, $\mathcal{L}_{s+b}$ and $\mathcal{L}_b$, are evaluated for each pseudodataset. From these values, the *test-statistic* $\lambda = \log Q = \log(\mathcal{L}_{s+b}/\mathcal{L}_b)$ is calculated. The ensemble of pseudodatasets with and without signal events added is used to calculate expected distributions of λ for each of the hypotheses being tested. By comparing the value of λ evaluated with the actual data sample observed in the experiment with the expected distributions

from each hypothesis, quantitative statements about the preference of data for one or the other can be made.

The distributions of the test statistic λ for a sample model, assuming the two hypotheses being considered, are shown in figure 10.54 for the HAD box. For pseudo-datasets with only background events λ tends to smaller values, reflecting the larger likelihood of hypothesis $\mathbb{H}_b$ relative to $\mathbb{H}_{s+b}$. Conversely, λ becomes large for events with signal. The value of λ observed in data, along with the expected distributions
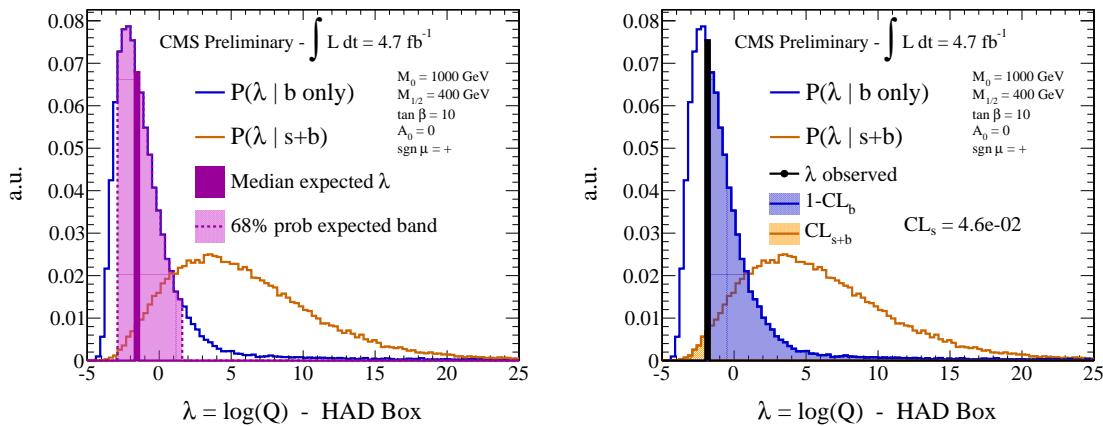

Figure 10.54: Distributions of λ for the background-only and signal+background hypotheses for a CMSSM model point with $M_0 = 1000$ GeV, $M_{1/2} = 400$ GeV, $\tan \beta = 10$, $A_0 = 0$ and sgn $\mu = +$. The blue distribution corresponds to the background-only hypothesis while the orange assumes signal and background events in the data sample. (Left) The purple line indicates the expected median value of λ for $\mathbb{H}_b$, while the shaded magenta region indicates the 68% probability interval centered on this value. (Right) The black line indicates the observed value of λ in data. This is used, with the expected λ distributions for each hypothesis, to calculate $\mathrm{CL}_S = \mathrm{CL}_{s+b}/(1\text{-}\mathrm{CL}_b)$.

of λ for each hypothesis, are used to calculate two probabilities: CL_{s+b} is defined as the integral of the expected $\mathbb{H}_{s+b}$ λ pdf below the observed value of λ, indicated by the shaded orange region in the example shown in the rightmost plot of figure 10.54. This is the probability that a hypothetical dataset with signal events from model $\mathcal{M}$ would have a smaller, or more $\mathbb{H}_b$-like, value of λ. Similarly, 1-CL_b (shaded blue in figure 10.54) is defined as the integral of the expected $\mathbb{H}_b$ λ pdf above the observed

value of λ, and represents the probability of the SM background producing a more $\mathbb{H}_{s+b}$-like configuration than that observed in data, assuming that $\mathbb{H}_b$ is true.

The test statistic distributions from each box are combined to form a more powerful test statistic, defined as the sum of the test statistics from each of the boxes. Since λ is defined as the natural logarithm of the likelihood ratio, this combined test statistic is equivalent to the logarithm of the ratios of the full likelihoods of each hypothesis, combining the information from each box final state

$$\lambda_{TOT} = \sum_i^{boxes} \lambda_i = \sum_i \log \frac{\mathcal{L}_{s+b}^i}{\mathcal{L}_b^i} = \log \frac{\prod_i^{boxes} \mathcal{L}_{s+b}^i}{\prod_i^{boxes} \mathcal{L}_b^i} = \log \frac{\mathcal{L}_{s+b}^{TOT}}{\mathcal{L}_b^{TOT}} \ . \qquad (10.24)$$

This combination is performed for each pseudodataset and the actual observed data sample separately.

The metric by which a models are judged to be excluded is called CL_S [208, 209], and is defined as

$$\mathrm{CL}_S = \frac{\mathrm{CL}_{s+b}}{(1 - \mathrm{CL}_b)} \ . \qquad (10.25)$$

While not a probability, CL_S is reflective of the p-value CL_{s+b} while being prevented from becoming artificially large in potentially pathological situations where the observed value of λ is inconsistent with both hypotheses. A model $\mathcal{M}$ is said to be excluded at 95% confidence level (C.L.) if $\mathrm{CL}_S < 0.05$ for its hypothesis test.

10.7.2 Systematic Errors in Interpretations

The systematic errors associated with the models of background and signal event yields are incorporated into the interpretation procedure by variations of the shapes and yields used to generated events for pseudodatasets. For background events, these systematic effects are reflected by sampling background shape and yield parameters, for each pseudo-experiment independently, from the multinormal covariance matrix.

The shape and normalization of the numerical 2D models used to generate signal events is systematically varied to account for a range of possible mis-modeling effects.

This effects are integrated into the toys in several different ways. Each pseudodataset includes samples from each of the boxes, such that systematic variations can be applied coherently for several boxes to account for effects that apply to each similarly, like the uncertainty of the selection efficiency of the razor triggers (see section 10.4.1) or the uncertainty on signal normalization following from the measurement of integrated luminosity in the data samples [210].

Some effects are taken into account by considering signal templates derived from simulated event samples where properties of the objects in these events are varied according to their respective uncertainties. Lepton identification uncertainties, derived from $Z(\ell\ell)$ data control samples as described in section 6.1, are used to derive event-by-event variations of reconstructed leptons' momenta. The application of analogous variations for jets and MET following from jet energy scale uncertainties [211, 212] are explained in section C.14.2.

Uncertainties following from PDFs can result in changes to the shape of signal yield templates, and are evaluated according to the procedure described in section C.13. These variations are performed bin by bin in the numerical 2D signal templates, allowing for these expected shape variations. Similarly, the uncertainty in b-tagging efficiency is reflected by variations, dependent on the p_T of a given jet [30], which can change the signal yield shape.

Each of these systematic variations are sampled from log-normal distributions with width set by the estimated magnitude of each of these potential effects. Variations modeled with normal distributions yield compatible results. The full list of the systematic effects associated with modeling the signal is summarized in Tab. 10.4. For effects labeled bin by bin, their size depends on the bin in the R^2 vs. M_R plane as well as on the model and box. Effects labelled point by point are evaluated for each model independently.

yield systematics	
$\mathcal{L}$	6.5%
cross section	point by point
trigger efficiency R^2-M_R	2%
trigger efficiency lepton	3% (lepton, di-lepton boxes)
shape systematics	
2D signal shape	20% (bin by bin)
PDF	point by point (up to 30%)
JES	point by point (up to 1%)
lepton-id (tag and probe)	1% (per lepton)

Table 10.4: Summary of the systematic uncertainties on the signal yield and shape.

10.7.3 Results Interpretation in the CMSSM

The null results of the inclusive razor search boxes are interpreted in the context of CMSSM models with varying values $m_{1/2}$ vs. m_0 and with $\tan \beta = 10$, $A_0 = 0$ and positive μ. The 2D numerical M_R/R^2 templates for these models, and corresponding systematic uncertainties, are determined from simulated CMSSM event samples generated with the PYTHIA6 [213] Monte Carlo program and analyzed using the full GEANT4-based [214] detector simulation, with the same reconstruction chain performed on events collected from real collisions. The sparticle mass spectrum for these models is calculated with SOFTSUSY [215] and decays with SUSYHIT [216]. The events are generated with PYTHIA6 through the SLHA interface [217]. Next-to-leading order (NLO) plus next-to-leading-logarithm (NLL) cross section calculations [218–223] are used to normalize the different production channels independently.

CL$_S$ is calculated through the likelihood ratio hypothesis test for each model, and a model is considered excluded at 95% confidence level if $CL_S < 0.05$. The result of interpreting the inclusive analysis in the CMSSM according to this metric is shown in figure 10.55. Models falling below the observed limit lines are excluded at a C.L. of at least 95%. The total observed limit is less constraining than the median expected limit at lower m_0 due to squark-squark production having a harder R^2 distribution (relative to gluino-gluino) and a local excess of events at large R^2 in the inclusive HAD box. Cascading decays of gluinos yield more leptons than squark decays, resulting in the contribution of the lepton final state boxes increasing in magnitude, relative to the HAD box, with increasing m_0.
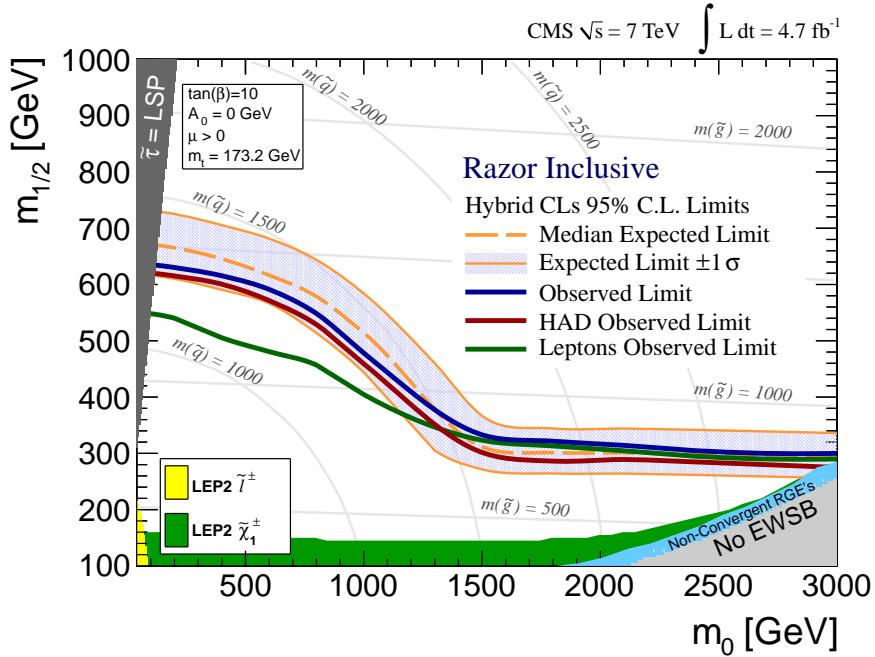
Figure 10.55: Observed (solid blue curve) and median expected (dot-dashed curve) 95% C.L. limits in the $(m_0, m_{1/2})$ CMSSM plane [224] with $\tan\beta = 10$, $A_0 = 0$, $\text{sgn}(\mu) = +1$ from the inclusive razor analysis. The $\pm$ one standard deviation equivalent variations in the uncertainties are shown as a band around the median expected limit. The observed limits resulting from an analysis of only the HAD or leptonic boxes independently are indicated as solid crimson and green lines, respectively.

The signal-like fluctuation of the observed limit relative to the expected in the HAD box is illustrated in figure 10.56. Despite the fact that the CMSSM favors the hadronic final state with strong production of squarks and gluinos, an important fraction of events are selected in the lepton boxes, as indicated in the selection efficiency figures contained in section C.14.1. Overall, the observed limits are in agreement with expectations, reflecting the lack of deviations between the background predictions and observed data consistent with signal in each of the boxes.
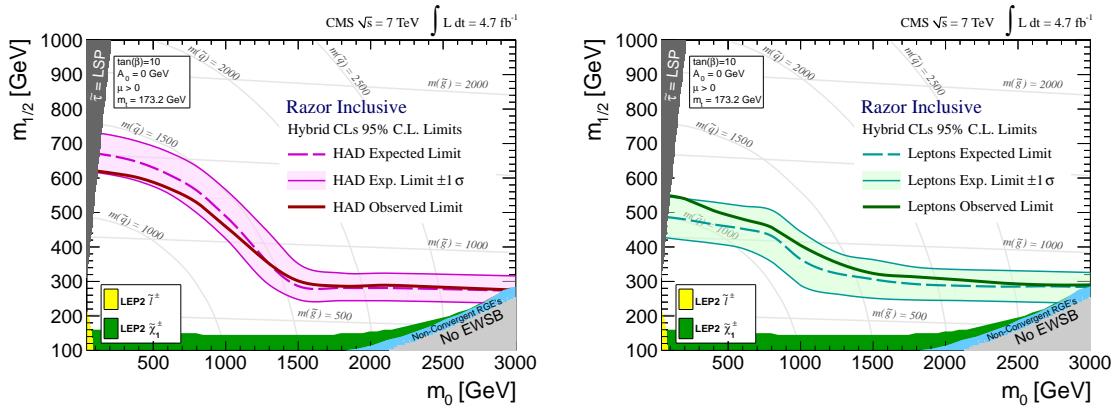
Figure 10.56: (Left) Observed (solid curve) and median expected (dot-dashed curve) 95% CL limits in the $(m_0, m_{1/2})$ CMSSM plane for the HAD box with $\pm$ one standard deviation uncertainties around the median expected limit. (Right) Observed and expected limits for the combined lepton boxes.

10.7.4 Simplified Model Interpretation

The null search results are also interpreted as cross section upper limits on a number of simplified models [225–227] where a limited set of hypothetical particles and decay chains are introduced to produce a given topological signature. Specific applications of results according to these Simplified Mass Spectra (SMS) models have appeared in Refs. [180, 181, 226, 227]. For each model studied, a cross section at 95% C.L. is derived, as a function of the mass of the produced particles (gluinos or squarks, depending on the model) and a single neutralino mass. These cross section upper limits are compared with NLL-NLO calculations of the production cross sections of these sparticles, and corresponding uncertainties [223], in order to identify regions of excluded sparticle mass-parameter space. These cross sections are summarized in figure 10.57.
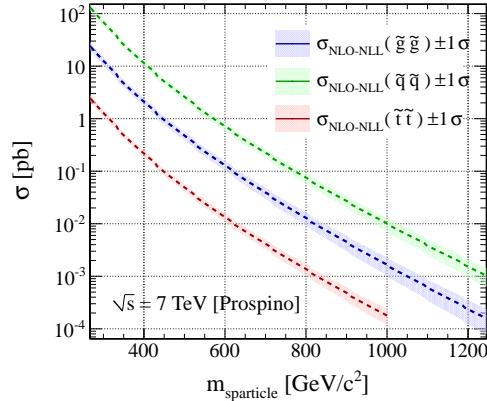

Figure 10.57: Theoretical cross sections for gluino, squark and stop pair production, calculated with Prospino2.0 [223, 228].

10.7.4.1 Models with Squarks and Gluinos Decaying to Jets

We interpret the results of the inclusive search analysis for two SMS models with decays to light-quark jets, illustrated in figure 10.58. The model T1 describes gluino pair production, with each gluino undergoing a three-body decay to two light quarks and a neutralino. Similarly, the model T2 has events with pair-produced squarks,

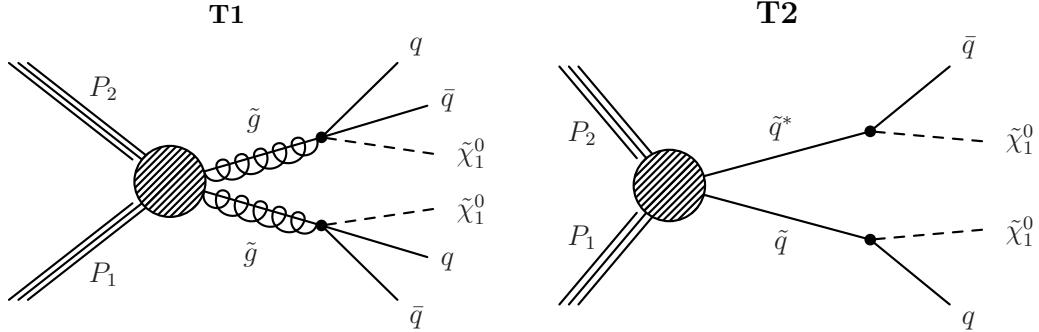each decaying directly to a light quark and neutralino.

Figure 10.58: SMS models T1 and T2 with pair-produced gluinos or squarks decaying to two or one light quarks and a neutralino, respectively.

The phenomenological properties of these models are discussed in section 10.2, where the sparticle mass dependence on the shape of signal yields in the razor plane is explained. This migration of signal events with changing masses is reflected in the derived cross section upper limits for these models, shown in figure 10.59. Models whose events populate the razor plane in regions with more background events (low M_R/low-R^2) tend to have larger cross section upper limits, with a strong correspondence with the squark/gluino and neutralino mass difference due to the M_R peak position varying with its value.

In order to evaluate potential systematic uncertainties due to modeling of initial state radiation (ISR), particularly for models with compressed squark/gluino and neutralino spectra, we consider variations of the scale parameters which dictate ISR behavior in simulated events. Model points which are found to have a large dependence are excluded from the results, as described in section C.14.2.2, and are indicated in grey. The majority of the events produced in these models populate the HAD box, with the selection efficiency as a function of sparticle masses shown in section C.14.1. When interpreted in the context of the theoretical squark and gluino production cross sections, the excluded parameter space corresponds closely to the expected sensitivity of the analysis, with variations around the median expected limit of small magnitude (relative to experimental uncertainties) and consistent with the agreement of observed
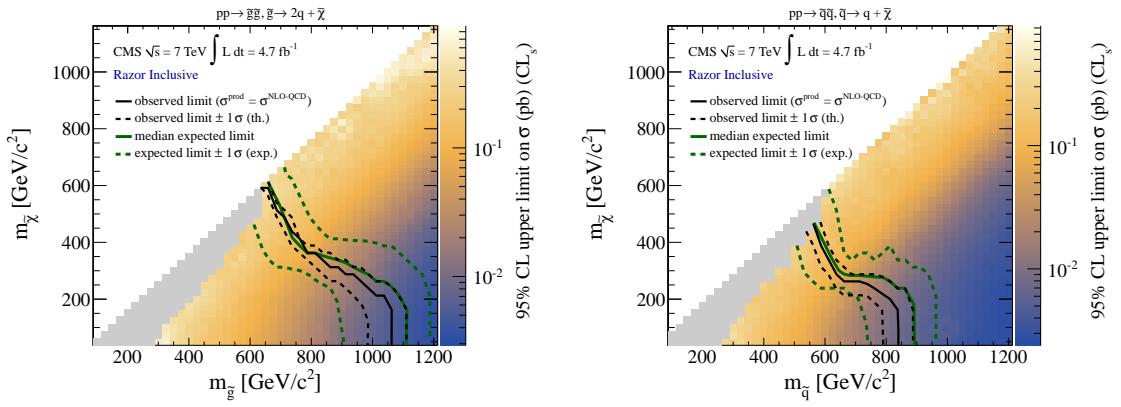
Figure 10.59: Cross section upper limit as a function of model mass spectrum for T1 (Left) and T2 (Right) from the inclusive razor analysis. The color scale (z-axis) indicates the observed cross section upper limit for each model point. The solid black line indicates the observed exclusion region, assuming nominal NLO+NLL SUSY production cross sections for squarks and gluinos, as a function of their masses with dotted black lines reflecting the $\pm 1\,\sigma$ theory errors around the nominal cross section. The solid green line marks the median expected exclusion region, with dotted green lines indicating the expected exclusion with $\pm 1\,\sigma$ experimental uncertainties. The solid grey region indicates model points where the analysis was found to have dependence on ISR modeling in simulation of signal events above a predefined tolerance; no interpretation is presented for these model points.

inclusive HAD box yields with expectations from the background prediction.

10.7.4.2 Models with Sbottoms and Gluinos Decaying to b-quarks

Motivated by the importance of the superpartners to the third generation quarks in the cancellation of quadratic divergences to the Higgs mass, we consider models with real and virtual sbottom production with decays to b-quarks. The results of the b-tagged analysis is used to constrain these models, which are shown pictorially in figure 10.60. The model T1bbbb describes gluino pair production, with each gluino undergoing a three-body decay to two b-quarks and a neutralino. Similarly, the model T2bb has events with pair-produced sbottoms, each decaying directly to b-quark and neutralino.
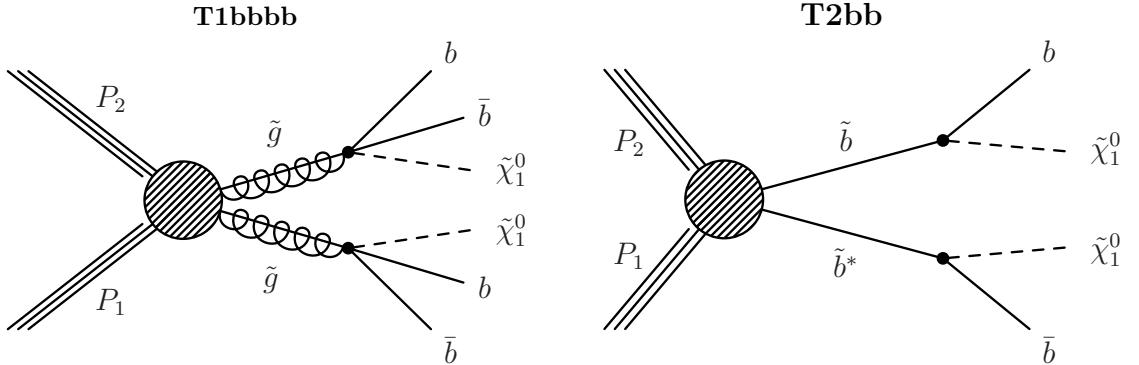

Figure 10.60: SMS models T1bbbb and T2bb with pair-produced gluinos or sbottoms decaying to two or one b-quarks and a neutralino, respectively.

Cross section upper limits for T1bbbb and T2bb are shown in figure 10.61. The events from these b-tagged-jet enriched models have kinematics very similar to the light-quark models T1 and T2, with cross section upper limits increasing with larger gluino/sbottom and neutralino mass differences. Contributions from the b-tagged HAD box provide the most sensitivity to these models, with some sensitivity coming from the MU and MU-MU boxes due to muons being reconstructed from semi-leptonic b-meson decays. As a result, the agreement between the observed exclusion line and expected is a convolution of the background observation and prediction differences in

the HAD box with the expected location of signal events in the razor plane. From the b-tagged background prediction results described in section 10.6, we observe that the data shows a background-like deviation in the high R^2 region of the HAD box, resulting in models with larger values of R^2 being more tightly constrained than expected. For sbottom pair production, the di-stop production cross section of figure 10.57 is used for comparison with cross section upper limits.
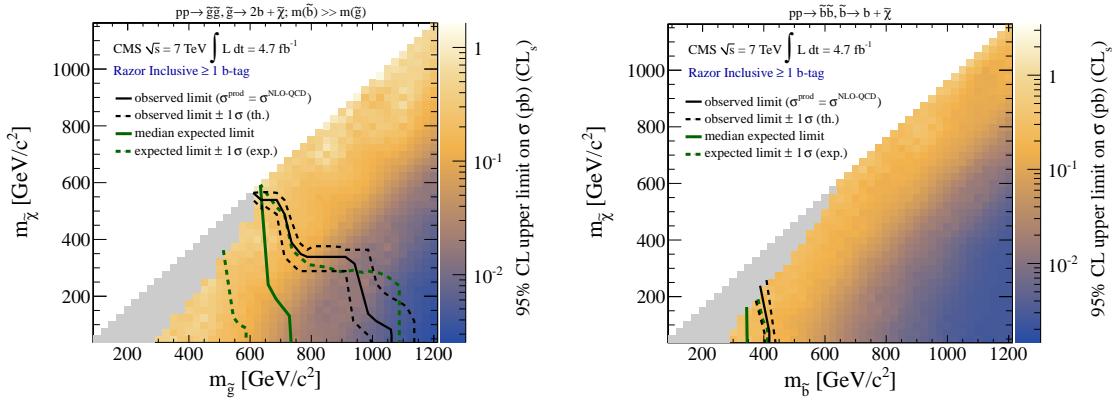

Figure 10.61: Cross section upper limit as a function of model mass spectrum for T1bbbb (Left) and T2bb (Right) following from the b-tagged razor analysis. The color scale (z-axis) indicates the observed cross section upper limit. The solid black line indicates the observed exclusion region with $\pm 1\,\sigma$ theory errors. The solid green line marks the median expected exclusion region, with $\pm 1\,\sigma$ experimental uncertainties.

10.7.4.3   Models with Sparticles Decaying to Top Quarks

The results of the inclusive and b-tagged jet searches are interpreted in two simplified scenarios where sparticles decay to SM top quarks. Model T1tttt includes gluino pair production, where each gluino decays through a virtual (very heavy) stop to two top quarks and a neutralino. The second model, T2tt, has pair-produced stops decaying directly to a top quark and a neutralino. Illustrations of these models are shown in figure 10.62.

The large number of top quarks produced in these events implies that there will also be W bosons and b-quarks from decays. This results in events being democratically distributed throughout the different final state boxes according to the number

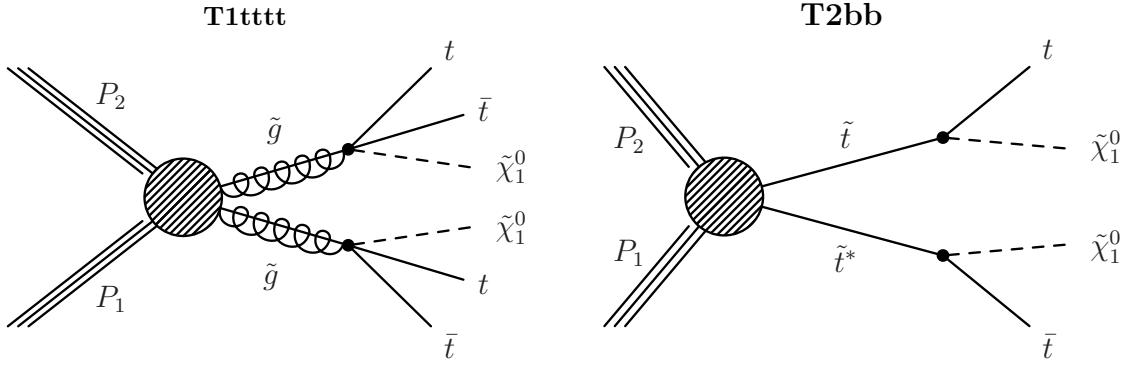**T1tttt**   **T2bb**

Figure 10.62: SMS models T1tttt and T2tt with pair-produced gluinos or stops decaying to two or one top quarks and a neutralino, respectively.

and decay type of the W bosons and an enhancement in signal-to-background for the b-tagged analysis. The cross section upper limits following from both search analyses for these top-enriched models are presented in figure 10.63. We observe that the b-tagged analysis has increased sensitivity to these models, relative to the inclusive search, because of reduced backgrounds. The sensitivity to these models is achieved through a combination of the results from all of the final state boxes which, in this case, contribute with comparable magnitudes

10.8   Conclusion

We have performed a search for evidence of sparticle production using a data sample of 4.7 fb^{-1} integrated luminosity from pp collisions at $\sqrt{s} = 7$ TeV, recorded by the CMS detector at the LHC. The search was based on the razor variables, M_R and R, which are used to distinguish between events containing two or more weakly interacting particles resulting from the decays of new, heavy sparticles and the SM backgrounds in final states with jets, missing transverse energy and with and without leptons and b-tagged jets.

Over many different final states, in both the inclusive and b-tagged jet analyses, no significant deviation from the predicted background shapes and yields is observed. This null result is interpreted as 95% C.L. limits in the $(m_0, m_{1/2})$ CMSSM parameter
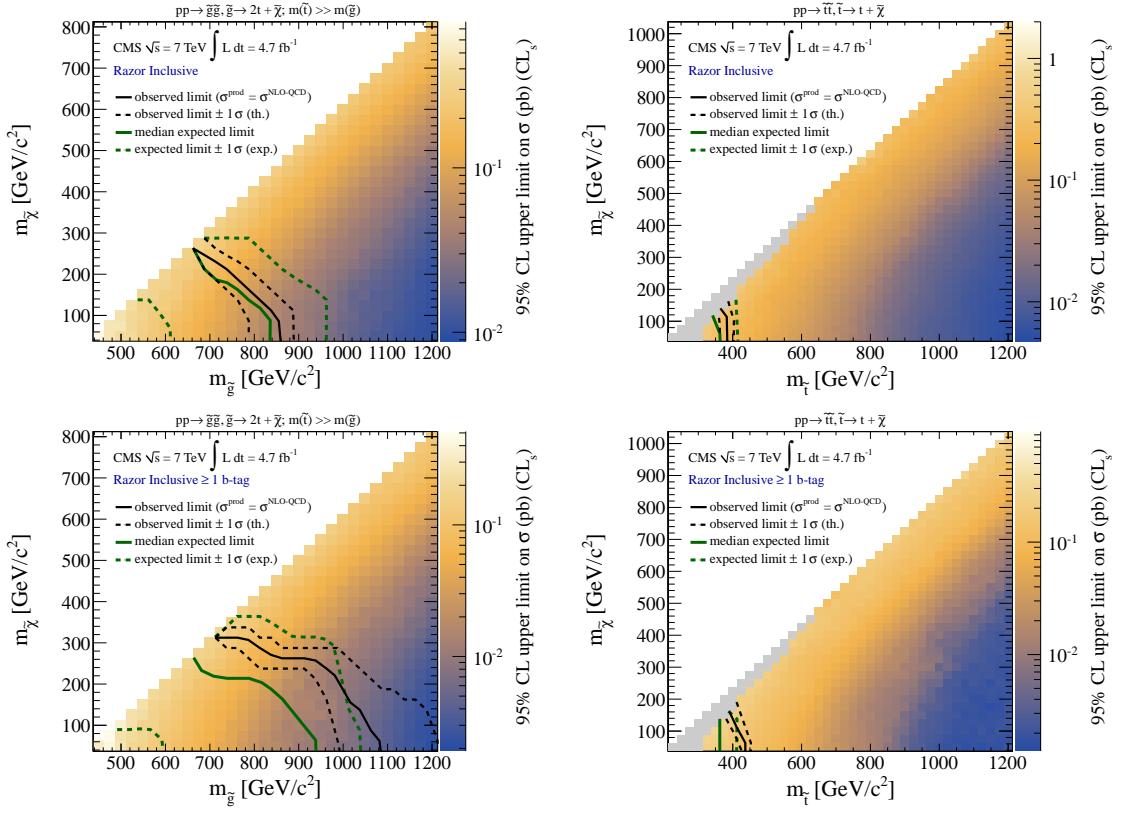
Figure 10.63: Cross section upper limit as a function of model mass spectrum for T1tttt (Left) and T2tt (Right) interpreted with the inclusive (Top) and b-tagged (Bottom) razor analyses. The color-scale (z-axis) indicates the observed cross section upper limit. The solid black line indicates the observed exclusion region with $\pm 1\ \sigma$ theory errors. The solid green line marks the median expected exclusion region, with $\pm 1\ \sigma$ experimental uncertainties.

space, where for $m_{\tilde{q}} \sim m_{\tilde{q}}$ we exclude squarks and gluinos up to 1.35 TeV in mass, and for $m_{\tilde{q}} > m_{\tilde{g}}$ gluinos up to 800 GeV.

The results are also interpreted in a collection of simplified models, with excluded sparticle masses summarized in figure 10.64. For these models, gluinos are excluded with masses as large as 1.1 TeV, for small neutralino masses, and first-two generation squarks, stops and sbottoms are excluded up to about 800, 425 and 400 GeV, respectively.

The strength of these constraints on hypothetical SUSY parameter space highlights the features of the razor analysis approach; the continuous modeling of the
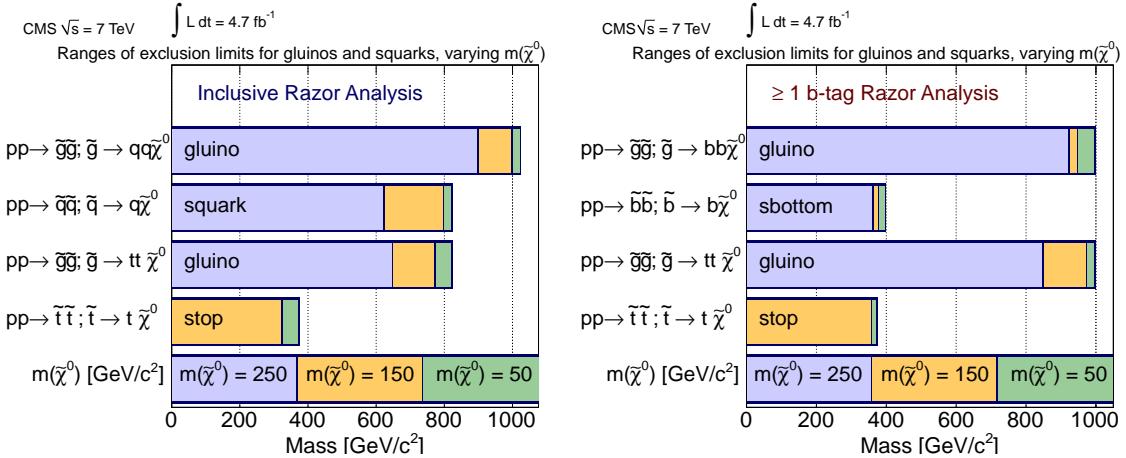
Figure 10.64: Summary of largest excluded sparticle masses in simplified models for the inclusive (left) and *b*-tagged jet (right) razor searches, for different neutralino masses.

background ensures that the knowledge of signal and background shapes in the razor plane is exploited maximally. Partitioning the event sample according to final state increases sensitivity to models with an enrichment in lepton final states, as does the *b*-tagged jet selection for those with many *b*-quarks. The highly efficient razor triggers allow these searches to extend to large regions of phase space, correspondingly constraining large volumes of SUSY model parameter-space with accuracy.

Chapter 11

Future Searches

In chapters 9 and 10 we introduced the event kinematic variables M_R and R and described how they could be used to search for SUSY and other BSM theories in inclusive final states with jets, leptons and missing transverse energy. The approach involved grouping all the particles measured in the detector into two distinct mega-jets based on their kinematics, effectively casting each event into a two visible object topology where the razor variables could be calculated and used to study the event. This approach was necessary due to the combinatorial ambiguity of the final state; when searching inclusively for a general class of phenomena it is not possible to use the specific details of a particular model to help assign particles to one decay chain or another. The result is that jets from initial state radiation are included in these mega-jets, leading to the M_R/R phenomenological scaling behavior which is exploited in these searches.

To complement these inclusive searches we have developed a suite of *exclusive* search strategies. These are appropriate for looking for specific, well-motivated signal topologies in final states where, based on the type of particles reconstructed, one knows which decay chain to assign them to. In the following chapter we describe a new strategy for studying inclusive and exclusive final states and explain a systematic approach for deriving appropriate razor variables for each case. The phenomenology of these variables is demonstrated in the example of $H \to WW \to (\ell\nu)(\ell\nu)$ where we see that we are able to extract information about the H mass and the angular properties of its decays. Finally, we introduce a new set of kinematic variables designed to study

events with pairs two-step decay chains, like $t \to Wb \to (\ell\nu)b$, and independently extract the characteristic scale of *both* decays. The possible application of these variables to a search for new-symmetry partners of the top quark is described.

11.1　The Generalized Approach to Razor Variables

We return once again to the canonical event topology of the pair production of new particles S_i, each decaying to a reconstructable SM particle, Q_i and a weakly interacting, potentially massive, particle χ_i, illustrated in figure 11.1. In the previous derivations of razor variables, we made several assumptions about the kinematics of these events; we assumed that the transverse momentum of the CM system was negligible, the masses of the visible particles are negligible and the particles S_i are produced with energy as close to the $2M_S$ threshold as kinematically possible. The generalized approach to these topologies involves only one assumption: that the two decay chains are *symmetric*, in that the particle content of each decay is identical. In this section we see how this single principle can be used to approximately reconstruct events of this type.
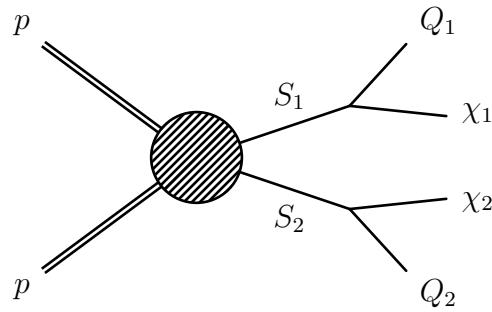
Figure 11.1: Canonical event topology illustrating pair production of particles S_i in a pp collision, with each S_i decaying as $S_i \to Q_i\chi_i$.

Using the notation of section 9.1.1, we recall that the particles S_i, Q_i and χ_i are represented by four vectors s_i, q_i and ν_i, respectively, and that the superscripts on these quantities indicate which reference frame they are being evaluated in. The

condition of symmetry between the two decay chains can be summarized as

$$s_1^2 = s_2^2 = M_S^2, \quad \nu_1^2 = \nu_2^2 = M_\chi^2, \quad q_1^2 = q_2^2 = M_Q^2 . \tag{11.1}$$

With this symmetry condition, the particles Q_i and χ_i have a fixed, characteristic momentum P^Δ in their respective S_i rest frames, which can be expressed as

$$M_\Delta \equiv |\vec{P}_1^\Delta| = |\vec{P}_2^\Delta| = \frac{\sqrt{(M_S^2 - (M_\chi - M_Q)^2)(M_S^2 - (M_\chi + M_Q)^2)}}{2M_S} . \tag{11.2}$$

In order to reconstruct the scale M_Δ we must move the visible particles Q_i from the lab frame to approximations of their respective S_i rest frames. The series of transformations these approximations are meant to represent are illustrated in figure 11.2. There are two unknown velocities $\vec{\beta}_{CM}$ and $\vec{\beta}_L = \{\vec{\beta}_T, \beta_z\}$, the first of which is applied asymmetrically (opposite directions) to Q_1 and Q_2 while the second acts symmetrically on the two particles. In order to derive expressions for these velocities we factorize the problem. We first work in the CM frame to derive an approximation for β_{CM}. Subsequently, we demonstrate how these CM quantities can used to approximate $\vec{\beta}_L$.
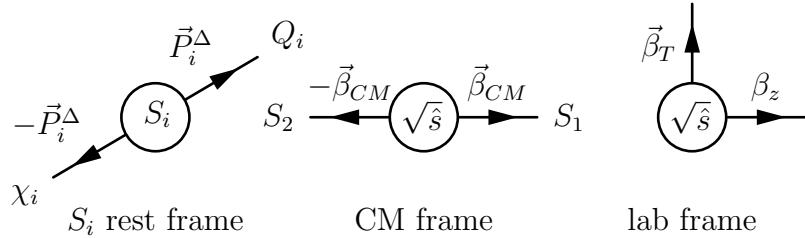

$$\begin{array}{ccc} S_i \text{ rest frame} & \text{CM frame} & \text{lab frame} \end{array}$$

Figure 11.2: The four relevant reference frames of the canonical topology. (Left) In the S_i rest frames the particles Q_i and χ_i recoil against each other with characteristic momentum P_i^Δ. (Center) In the CM frame S_1 and S_2 travel in opposite directions with velocity β_{CM}. (Right) In the lab frame, the CM system is traveling with a longitudinal velocity β_z and transverse velocity $\vec{\beta}_T$.

11.1.1 $\tilde{\beta}_{\mathbf{R}}$ in the CM Frame

The CM frame is related to the S_i rest frames by the asymmetric boost $\vec{\beta}_{CM}$, with three unknown d.o.f.. There is only one absolute constraint on $\vec{\beta}_{CM}$: after its application, the particles Q_i must have the same energy in their respective S_i rest frames. This symmetry condition can be expressed as a constraint on our expression for $\vec{\beta}_{CM}$, which we denote $\vec{\beta}_R$

$$E_{Q_1}^S = E_{Q_2}^S \quad \Rightarrow \quad \gamma_R(E_{Q_1}^{CM} - \vec{\beta}_R \cdot \vec{q}_1^{\,CM}) = \gamma_R(E_{Q_2}^{CM} + \vec{\beta}_R \cdot \vec{q}_2^{\,CM}) \quad (11.3)$$

$$\Rightarrow \quad \vec{\beta}_R \cdot (\vec{q}_1^{\,CM} + \vec{q}_2^{\,CM}) = E_{Q_1}^{CM} - E_{Q_2}^{CM} \ .$$

We see that this condition dictates the magnitude of $\vec{\beta}_R$ in the direction $(\vec{q}_1^{\,CM} + \vec{q}_2^{\,CM})$. What remains is to determine $\vec{\beta}_R$ in the perpendicular directions. We expand $\vec{\beta}_R$ in an orthonormal basis, taking $(\vec{q}_1^{\,CM} + \vec{q}_2^{\,CM})$ as the first direction. The choice of basis is illustrated in figure 11.3.
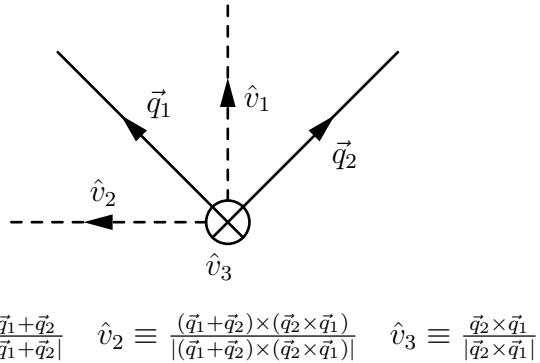

$$\hat{v}_1 \equiv \frac{\vec{q}_1 + \vec{q}_2}{|\vec{q}_1 + \vec{q}_2|} \quad \hat{v}_2 \equiv \frac{(\vec{q}_1 + \vec{q}_2) \times (\vec{q}_2 \times \vec{q}_1)}{|(\vec{q}_1 + \vec{q}_2) \times (\vec{q}_2 \times \vec{q}_1)|} \quad \hat{v}_3 \equiv \frac{\vec{q}_2 \times \vec{q}_1}{|\vec{q}_2 \times \vec{q}_1|}$$

Figure 11.3: Q_1 and Q_2 in the CM frame. The plane defined by the momentum $\vec{q}_1$ and $\vec{q}_2$ is illustrated. An orthonormal basis $\hat{v}_i$ is defined based on these momenta.

We can write $\vec{\beta}_R$ in this basis as

$$\vec{\beta}_R = \beta_1 \hat{v}_1 + \beta_2 \hat{v}_2 + \beta_3 \hat{v}_3 \ , \qquad (11.4)$$

and the condition equation (11.3) requires that

$$\beta_1 = \frac{E_{Q_1}^{CM} - E_{Q_2}^{CM}}{|\vec{q_1}^{CM} + \vec{q_2}^{CM}|} \ .$$ (11.5)

Now we must assign values to β_2 and β_3, the latter of which is a particularly difficult d.o.f.. Since it is a boost perpendicular to both $\vec{q_1}^{CM}$ and $\vec{q_2}^{CM}$ it will only multiply $E_{Q_i}^S$ by a constant factor. Additionally, it will have the same effect regardless of its sign. This is an indication of the fact that, without a priori knowledge of the new particle masses in these events, we have no constraining information about β_3 whatsoever. Any variable, V, we construct that depends on $\vec{\beta}_R$ will satisfy

$$\left. \frac{\partial V}{\partial \beta_3} \right|_{\beta_3 = 0} = 0 \ .$$ (11.6)

Given this consideration, we set $\beta_3 = 0$.

In the derivation of M_R from chapter 10 we removed a d.o.f. by appealing to the fact that, for non-resonant S pair production, γ_{CM} will generally be close to one. We could take the same approach here and set $\beta_2 = 0$, arguing that this choice minimizes γ_R. Let us denote this solution as $\gamma_{\min}$, such that $\vec{\beta}_{\min} = \beta_1 \hat{v}_1$. We will see later that this choice will result in the variables γ_{R*} and M_{R*} derived in section 10.1. The shortcoming of this choice is that $\gamma_{\min}$ will not scale with the true value of γ_{CM}. If the particles S_i are produced through a heavy s-channel resonance γ_{CM} will be fixed at a value greater than one. Ideally, we want our approximation γ_R to contain information about γ_{CM}, particularly if it is concentrated at nontrivial values.

A choice for β_2 that will give sensitivity to the true value of γ_{CM} can be derived by appealing to a different principle. Rather than minimizing γ_R, we can minimize $(E_{Q_1}^S + E_{Q_2}^S)$. Conceptually, this choice will turn $(E_{Q_1}^S + E_{Q_2}^S)$ into a quantity that will be invariant under asymmetric boosts in the direction $\hat{v}_2$. This choice yields

$$\frac{\partial (E_{Q_1}^S + E_{Q_2}^S)}{\partial \beta_2} = 0 \quad \Rightarrow \quad \beta_2 = \frac{\hat{v}_2 \cdot (\vec{q_1}^{CM} - \vec{q_2}^{CM})}{E_{Q_1}^{CM} + E_{Q_2}^{CM}} \ .$$ (11.7)

Curiously, with above choice for β_2 the expression for $\vec{\beta}_R$ becomes simpler,

$$\vec{\beta}_R = \beta_1 \hat{v}_1 + \beta_2 \hat{v}_2 = \frac{\vec{q}_1^{\,CM} - \vec{q}_2^{\,CM}}{E_{Q_1}^{CM} + E_{Q_2}^{CM}} \;. \tag{11.8}$$

In fact, our solution for $\vec{\beta}_R$ satisfies

$$\left.\frac{\partial(E_{Q_1}^S + E_{Q_2}^S)}{\partial\beta_1}\right|_{\vec{\beta}_{CM} = \vec{\beta}_R} = 0 \;, \tag{11.9}$$

even with β_1 chosen for seemingly independent reasons. We have found that the principle of choosing a $\vec{\beta}_R$ which minimizes $(E_{Q_1}^S + E_{Q_2}^S)$ implies the original symmetry constraint from equation (11.3).

To summarize, we have derived two possible solutions for $\vec{\beta}_{CM}$, which we denote $\vec{\beta}_{\min}$ and $\vec{\beta}_R$. In the following sections we will see how these variables behave, particularly in the context of moving from the CM frame to the lab frame.

11.1.2 From the Lab Frame to the R Frame

We have derived two solutions to our problem in the CM, but we must now try to express our CM frame variables in terms of the particles measured in the detector, which corresponds to the lab frame. The two reference frame are related by a boost, $\vec{\beta}_L$, corresponding to the velocity of the CM system. In this section we derive an approximation of this boost denoted $\vec{\beta}_{L_R}$, which will move from the lab frame to an approximation of the CM frame, denoted the R-frame.

Now, there are several constraints that $\vec{\beta}_{L_R}$ must satisfy. Firstly, the expressions for the momenta of the particles Q_i in the R frame must be invariant under longitudinal boosts. Secondly, the transverse component of $\vec{\beta}_{L_R}$ must point in a specific direction. Appealing to conservation of transverse momentum, we can interpret the missing transverse energy in the event $(\vec{M})$ as the vectorial sum of the transverse momenta of the χ_i particles. Denoting the momentum of the CM system in the lab

frame $\vec{P}^{CM}$, we observe that

$$\vec{\beta}_{L_R} = \frac{\vec{P}^{CM}}{\sqrt{|\vec{P}^{CM}|^2 + \hat{s}}} = \frac{\{\vec{M} + \vec{q}_{1T}^{\,l} + \vec{q}_{2T}^{\,l}, P_z^{CM}\}}{\sqrt{|\vec{P}^{CM}|^2 + \hat{s}}} \ . \tag{11.10}$$

Since an expression for P_z^{CM} can be derived by requiring invariance of our final observables under longitudinal boosts, if we knew the value of $\hat{s}$ then the boost would be completely determined.

Returning to our expressions for the asymmetric CM frame boosts $\vec{\beta}_R$ we construct an approximation for $\sqrt{\hat{s}}$. We cannot calculate this quantity in the CM frame, even if our approximations for $\vec{\beta}_{CM}$ hold exactly, because we do not know the mass M_χ. As we have previously seen, our sensitivity to the actual masses M_S and M_χ is quite poor; only the expression of mass differences M_Δ can be extracted from events of this type in general. Given this situation, we choose to *interpret* this mass difference as an actual mass. Specifically, if we assume that $M_\chi = M_Q$ then exactly one half of the energy of each particle S_i is given to each particle Q_i in the S_i rest frame, such that $M_S = 2E_{Q_i}^S$. Using our expressions for E_S and γ_R in terms of the CM frame observables, and recalling that $\sqrt{\hat{s}} = 2\gamma_{CM} M_S$, we can construct an approximation of $\sqrt{\hat{s}}$, denoted $\sqrt{\hat{s}_R}$, given by

$$\sqrt{\hat{s}_R} = 2\gamma_R(E_{Q_1}^S + E_{Q_2}^S) = 2(E_{Q_1}^{CM} + E_{Q_2}^{CM}). \tag{11.11}$$

Hence $\sqrt{\hat{s}_R}$ is simply two times the sum of the Q_i energies in the CM frame. Interestingly, choosing $\vec{\beta}_{\min}$ instead of $\vec{\beta}_R$ as an approximation of $\vec{\beta}_{CM}$ results in the same expression for $\sqrt{\hat{s}_R}$.

We can now solve for $\sqrt{\hat{s}_R}$ in terms of $\vec{P}^{CM}$ and our lab frame observables by relating $\hat{s}_R$ to the energies of the visible objects in their respective S frames, which depend on the boosts from the CM frame

$$\sqrt{\hat{s}_R} = 2(E_{Q_1}^{CM} + E_{Q_2}^{CM}) = 2\gamma_{L_R}(E_{Q_1}^l + E_{Q_2}^l - \vec{\beta}_{L_R} \cdot (\vec{q}_1^{\,l} + \vec{q}_2^{\,l})) \ . \tag{11.12}$$

Interpreting $\sqrt{\hat{s}_R}$ as the mass of the particles S, this relation results in a quadratic constraint on $\sqrt{\hat{s}_R}$,

$$\sqrt{\hat{s}_R} = 2 \frac{\sqrt{\hat{s}_R + |\vec{P}^{CM}|^2}}{\sqrt{\hat{s}_R}} (E^l_{Q_1} + E^l_{Q_2}) - 2 \frac{\vec{P}^{CM} \cdot (\vec{q}_1^{\;l} + \vec{q}_2^{\;l}))}{\sqrt{\hat{s}_R}} \; , \qquad (11.13)$$

which gives the solution

$$\begin{aligned}
\hat{s}_R \;\; &= \;\; 2(E^l_{Q_1} + E^l_{Q_2})^2 - 2\vec{P}^{CM} \cdot (\vec{q}_1^{\;l} + \vec{q}_2^{\;l}) & (11.14) \\
&\pm \;\; 2(E^l_{Q_1} + E^l_{Q_2})\sqrt{(E^l_{Q_1} + E^l_{Q_2})^2 + |\vec{P}^{CM}|^2 - 2\vec{P}^{CM} \cdot (\vec{q}_1^{\;l} + \vec{q}_2^{\;l})} \; .
\end{aligned}$$

The two-fold sign ambiguity is resolved by choosing the positive solution, which ensures that $\sqrt{\hat{s}_R}$ is real and positive. There is now only one remaining unspecified d.o.f., which corresponds to P_z^{CM}. We can assign a value to P_z^{CM} and ensure that the subsequent kinematic variables are invariant under longitudinal boosts simultaneously by requiring

$$\begin{aligned}
\frac{\partial \sqrt{\hat{s}_R}}{\partial P_z^{CM}} \;\; &= \;\; 0 & (11.15) \\
\Rightarrow \;\; P_z^{CM} \;\; &= \;\; (q^l_{1z} + q^l_{2z}) \left[1 + \sqrt{1 + \frac{|\vec{P}_T^{CM}|^2 - 2\vec{P}_T^{CM} \cdot (\vec{q}_{1T}^{\;l} + q_{2T}^l)}{(E^l_{Q_1} + E^l_{Q_2})^2 - (q^l_{1z} + q^l_{2z})^2}} \right] \; ,
\end{aligned}$$

where $\vec{P}_T^{CM} = \vec{M} + \vec{q}_1^{\;l} + \vec{q}_2^{\;l}$. We note the appearance of the factor $[(E^l_{Q_1} + E^l_{Q_2})^2 - (q^l_{1z} + q^l_{2z})^2]$, which is equivalent to our expression for M_R from section 10.1. Defining

$$M_R^0 \equiv \sqrt{(E^l_{Q_1} + E^l_{Q_2})^2 - (q^l_{1z} + q^l_{2z})^2} \; , \qquad (11.16)$$

we can express $\sqrt{\hat{s}_R}$ as

$$\begin{aligned}
\sqrt{\hat{s}_R} \;\; &= \;\; 2(M_R^0)^2 - 2\vec{P}_T^{CM} \cdot (\vec{q}_1^{\;l} + \vec{q}_2^{\;l}) & (11.17) \\
&+ \;\; 2M_R^0 \sqrt{(M_R^0)^2 + |\vec{P}_T^{CM}|^2 - 2\vec{P}_T^{CM} \cdot (\vec{q}_{1T}^{\;l} + q_{2T}^l)} \; .
\end{aligned}$$

We observe that as $P_T^{CM} \to 0$ equations (11.16) and (11.18) simplify to

$$\sqrt{\hat{s}_R}\Big|_{P_T^{CM} \to 0} = 2M_R^0 \quad , \quad P_z^{CM}\Big|_{P_T^{CM} \to 0} = 2(q_1 z^l + q_2 z^l) \; , \tag{11.18}$$

which corresponds to the original solution to this event topology derived in section 10.1. In fact, we see that the longitudinal and transverse parts of the boost $\vec{\beta}_{L_R}$ can be factorized. Applying the boost $\vec{\beta}_{L_R}$ with P_z^{CM} taken from equation (11.16) is equivalent to first applying a longitudinal boost $\beta_{L_R z}$ followed by the transverse boost $\vec{\beta}_{L_R T}$ with values

$$\beta_{L_R z} = \frac{q_{1z}^l + q_{2z}^l}{E_{Q_1}^l + E_{Q_2}^l} \quad , \quad \vec{\beta}_{L_R T} = \frac{\vec{M} + \vec{q}_{1T}^{\,l} + \vec{q}_{2T}^{\,l}}{\sqrt{|\vec{P}_T^{CM}|^2 + \hat{s}_R}} \; . \tag{11.19}$$

Both sets of transformations result in identical R-frames and the longitudinal boost $\beta_{L_R z}$ is identical to the analogous β_{L*} from section 10.1. We see that we have re-derived the original razor variables for the case $P_T^{CM} \to 0$ but have also developed an effective p_T^{CM}-correction through the boost $\vec{\beta}_{L_R T}$, which turns $2M_R^0$ into $\sqrt{\hat{s}_R}$ in from equation (11.18).

11.1.3 Calculation of $\mathbf{M_R}$ and $\tilde{\vec{\beta}}_{\mathbf{R}}$

We can now calculate the transformations through each of the reference frames of interest in terms of the observables we measure in the laboratory frame: the four vectors of the two visible particles, q_1^l and q_2^l, and the missing transverse energy, $\vec{M}$. We interpret the vectorial sum of the missing transverse energy and the visible particles', q_i, transverse momentum as the transverse momentum of the CM system, $\vec{P}_T^{CM}$, such that

$$\vec{P}_T^{CM} = \vec{M} + \vec{q}_{1T}^{\,l} + \vec{q}_{2T}^{\,l} \; . \tag{11.20}$$

Our estimate for the invariant mass of the CM system, $\sqrt{\hat{s}_R}$, is given by equation (11.18) in terms of M_R^0, defined in equation (11.16).

First, we move the four vectors q_i^l from the lab frame to our approximation of the CM frame, which we denote the R-frame. This is accomplished through two Lorentz transformations: a longitudinal boost, $\beta_{L_R z}$ followed by a transverse boost, $\vec{\beta}_{L_R T}$, both defined in equation (11.19). The energies and momenta of the particles q_i can be expressed in the R-frame as

$$
\begin{aligned}
E_{Q_i}^R &= \gamma_{L_R T}\left(\gamma_{L_R z}\left(E_{Q_i}^l - \beta_{L_R z}q_{iz}^l\right) - \vec{\beta}_{L_R T}\cdot\vec{q}_{iT}^{\,l}\right) , \qquad (11.21)\\
\vec{q}_{iT}^{\,R} &= \vec{q}_{Ti}^{\,l} + (\gamma_{L_R T}-1)\,\vec{q}_{Ti}^{\,l}\cdot\hat{\beta}_{L_R T}\hat{\beta}_{L_R T} - \gamma_{L_R T}\gamma_{L_R z}\left(E_{Q_i}^l - \beta_{L_R z}q_{zi}^l\right)\vec{\beta}_{L_R T} ,\\
q_{iz}^R &= \gamma_{L_R z}\left(q_{iz}^l - \beta_{L_R z}E_{Q_i}^l\right) .
\end{aligned}
$$

From the R-frame, we next move each particle q_i to their respective approximate S_i rest frames. This is accomplished through our approximation of $\vec{\beta}_{CM}$, denoted $\vec{\beta}_R$ and given by

$$
\vec{\beta}_R = \frac{\vec{q}_1^{\,R} + \vec{q}_2^{\,R}}{E_{Q_1}^R + E_{Q_2}^R} . \qquad (11.22)
$$

This boost is applied asymmetrically to q_1 and q_2, since S_1 should be recoiling against S_2 in the R-frame. The energies and momenta of the particles q_i can be expressed in their respective S_i rest frames as

$$
\begin{aligned}
E_{Q_1}^S &= \gamma_R\left(E_{Q_1}^R - \vec{\beta}_R\cdot\vec{q}_1^{\,R}\right) , \qquad (11.23)\\
E_{Q_2}^S &= \gamma_R\left(E_{Q_2}^R + \vec{\beta}_R\cdot\vec{q}_2^{\,R}\right) .\\
\vec{q}_1^{\,S} &= \vec{q}_1 + (\gamma_R-1)\,\vec{q}_1^{\,R}\cdot\hat{\beta}_R\hat{\beta}_R - \gamma_R E_{Q_1}^R\vec{\beta}_R ,\\
\vec{q}_2^{\,S} &= \vec{q}_2 + (\gamma_R-1)\,\vec{q}_2^{\,R}\cdot\hat{\beta}_R\hat{\beta}_R + \gamma_R E_{Q_2}^R\vec{\beta}_R .
\end{aligned}
$$

The R frame mass, M_R, is our estimator of M_Δ and is defined as

$$
M_R = (E_{Q_1}^S + E_{Q_2}^S) = \sqrt{(E_{Q_1}^R + E_{Q_2}^R)^2 - |\vec{q}_1^{\,R} - \vec{q}_2^{\,R}|^2} , \qquad (11.24)
$$

and is related to $\sqrt{\hat{s}_R}$, which is our estimator of $2\gamma_{CM}M_\Delta$, by

$$\sqrt{\hat{s}_R} = 2\gamma_R M_R \ . \tag{11.25}$$

11.1.4 Properties of $\mathbf{M_R}$ and $\tilde{\beta}_{\mathbf{R}}$

In order to see how M_R and $\vec{\beta}_R$ behave under different conditions, we consider a toy simulation of the production and decays shown in figure 11.1. We assume trivial matrix elements for each of the decays, and allow the possibility to either fix γ_{CM} at a particular value (corresponding to resonant S pair production through a heavy s-channel particle) or vary it according to a predefined distribution. Non-trivial γ_{CM} distributions are derived by considering the toy two scalar model from section 9.1.1 and using PDF parameterizations corresponding to gluon-gluon interactions, which are expected to dominate in the production of increasingly heavy particles, relative to quark interactions, at the LHC.

Using this toy simulation, we first study the dependence of the reconstructed γ_R distribution as a function of the true value of γ_{CM}. From figure 11.4 we observe that the peak position of the γ_R distribution scales with γ_{CM}. This is compared with the variable $\gamma_{\min}$ which, since it was derived to minimize γ, peaks at one for all γ_{CM}. Similarly, in figure 11.5 we see how the quantities $\gamma_R M_R$ and M_R are sensitive to the characteristic scale M_Δ regardless of γ_{CM}. As γ_{CM} increases, the resolution of M_Δ degrades, but only to an asymptotic maximum. We find that, even for very large γ_{CM}, we are still able to resolve both γ_{CM} and M_Δ.

This fact is illustrated succinctly in figure 11.6, where we see the two dimensional $\gamma_R M_R/M_R$ peak. By allowing γ_R to scale with the true value of γ_{CM} we now have sensitivity to two scales in these events, γ_{CM} and M_Δ. The lack of strong correlations between $\gamma_R M_R$ and M_R (indicated by a peak in figure 11.6 rather than a long ridge) indicate that we are resolving these scales largely independent of each other. Furthermore, figure 11.7 demonstrates that these scales can be resolved regardless of the true value of M_Δ.
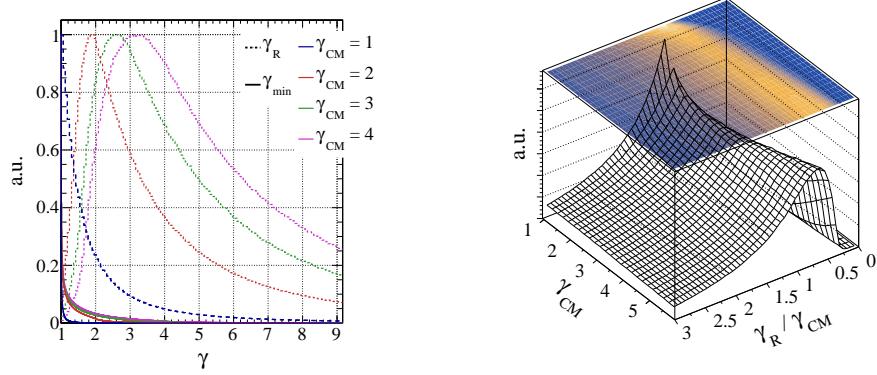
Figure 11.4: (Left) Distribution of γ_R and $\gamma_{\min}$ for different values of γ_{CM}. (Right) Distribution of γ_R/γ_{CM} as a function of γ_{CM}. Toy simulations are performed with $M_S = 1$ TeV, $M_\chi = 500$ GeV and $M_Q = 0$.
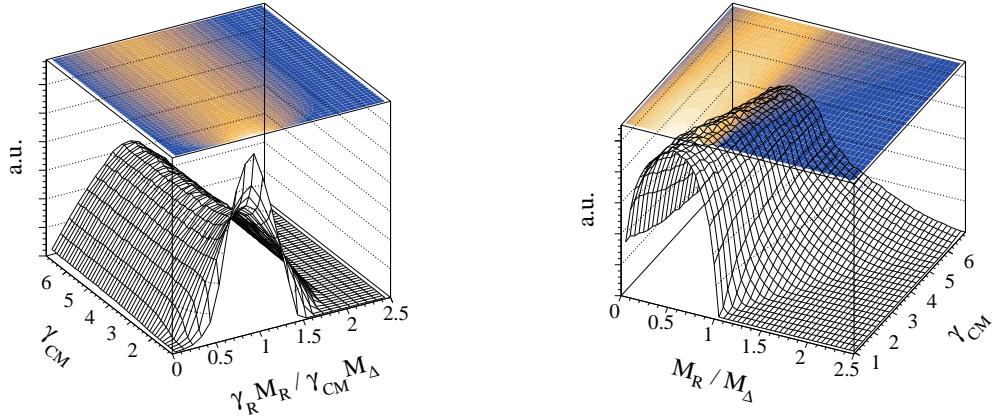

Figure 11.5: (Left) Distribution of $\gamma_R M_R/\gamma_{CM} M_\Delta$ as a function of γ_{CM}. (Right) Distribution of M_R/M_Δ as a function of γ_{CM}. Toy simulations are performed with $M_S = 1$ TeV, $M_\chi = 500$ GeV and $M_Q = 0$.
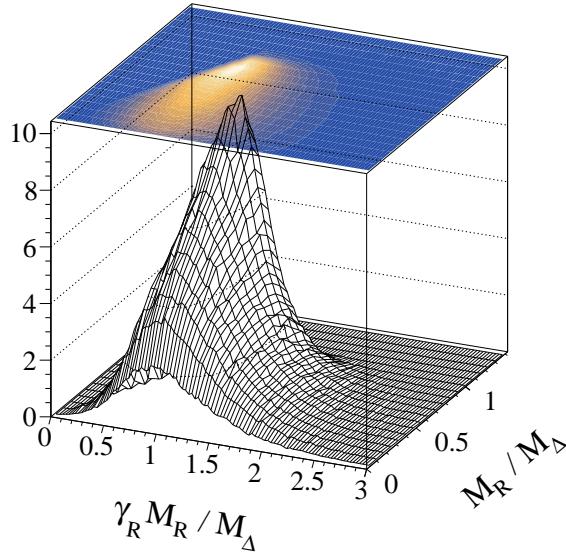
Figure 11.6: Distribution of $\gamma_R M_R / M_\Delta$ vs. M_R / M_Δ. Both variables are independently sensitive to the same characteristic scale. Toy simulations are performed with $M_S = 1$ TeV, $M_\chi = 500$ GeV and $M_Q = 0$ with γ_{CM} taken from a nontrivial distribution.
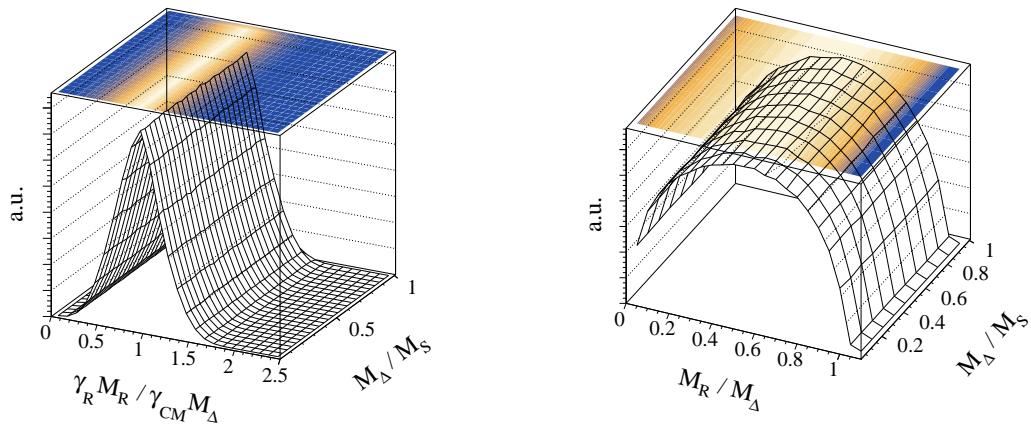

Figure 11.7: (Left) Distribution of $\gamma_R M_R / \gamma_{CM} M_\Delta$ as a function of M_Δ for fixed M_S. (Right) Distribution of M_R / M_Δ as a function of M_Δ for fixed M_S. Toy simulations are performed with $M_S = 1$ TeV and $M_Q = 0$ with γ_{CM} taken from a nontrivial, realistic distribution. M_χ is varied in order to vary M_Δ.

Finally, we evaluate the effect of non-zero P_T^{CM} on our kinematic variables. Without the p_T correction corresponding to $\vec{\beta}_{L_R T}$ the expressions for M_R and γ_R are not invariant under transverse boosts, meaning that their ability to resolve their respective scales will degrade with increasing P_T^{CM}. This fact is illustrated in the figure 11.8, along with the remarkable observation that, with the transverse boost $\vec{\beta}_{L_R T}$ applied, the variables M_R and γ_R are nearly independent of P_T^{CM}. This means that the derived expressions are, in a sense, nearly invariant under *all* Lorentz transformations, a striking property for a mass-sensitive variable that is shared only by the canonical invariant mass.
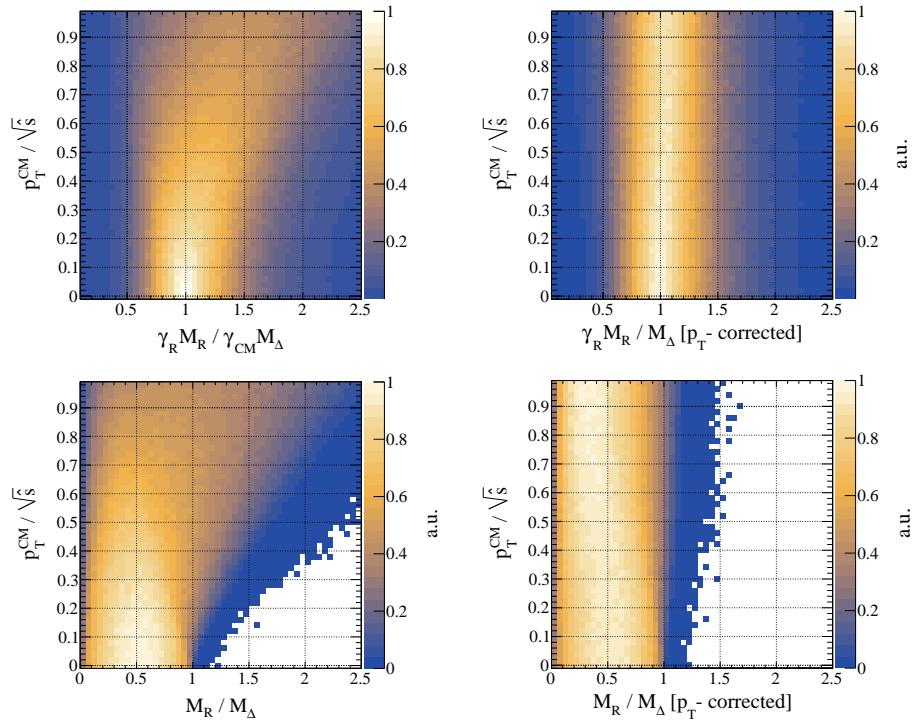

Figure 11.8: (Top) $\gamma_R M_R / \gamma_{CM} M_\Delta$ as a function of p_T^{CM}. (Bottom) M_{R/M_Δ} as a function of p_T^{CM}. (Left) Not transverse p_T correction is applied. (Right) The boost $\vec{\beta}_{L_R T}$ to move events to the R frame. Toy simulations are performed with $M_S = 1$ TeV, $M_\chi = 500$ GeV and $M_Q = 0$ with γ_{CM} taken from a nontrivial distribution.

Having returned a final time to the general razor topology we have re-derived the original razor variables in a simpler, systematic way. In the process, we have discovered sensitivity to a new scale in these events: γ_{CM}, which is approximated by

γ_R. We have also developed a procedure for making these variables almost completely invariant under all possible transformations, yielding robust indicators of the scales of new physics.

11.2   The Razor for $\mathbf{H \rightarrow WW \rightarrow (\ell\nu)(\ell\nu)}$

As an example application of the variables γ_R and M_R we consider the case of a Higgs boson decaying to two W bosons, each then decaying leptonically. This type of event is Illustrated in figure 11.9. The similarity between this decay topology and the one used to derive the razor variables (figure 11.1) is clear. The massive W bosons undergo two-body decays to leptons and neutrinos, the latter escaping detection. The two leptons are reconstructed in the detector and can be used to calculate the razor kinematic variables. In these events, the characteristic scale M_Δ is given by

$$M_\Delta = \frac{M_W^2 - M_\nu^2}{M_W} \sim M_W \ . \tag{11.26}$$

Since the Higgs mass is fixed, γ_{CM} takes a fixed value equal to $\gamma_{CM} = M_H/2M_W$, assuming the Higgs is sufficiently massive to produce two on-shell Ws. If this is the case, the construction $2\gamma_R M_R$ will peak at the Higgs mass. For convenience, we define $M_R^{p_T-\mathrm{corr}} \equiv 2\gamma_R M_R$ for this discussion.
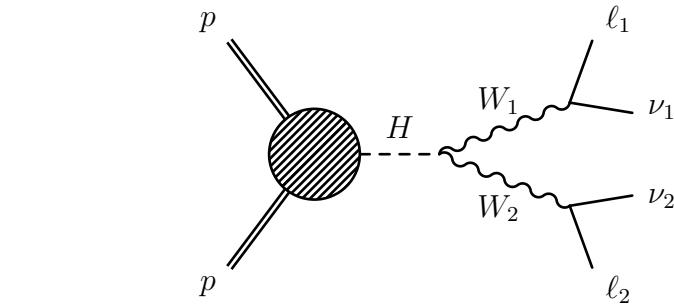

Figure 11.9: Illustration of a $H \rightarrow WW \rightarrow (\ell\nu)(\ell\nu)$ event.

Using the CMS full simulation, we examine $H \rightarrow WW$ and SM background events in the di-lepton final state. Leptons are identified using the same algorithms as the

CMS SUSY search described in Chapter 10. We select events from these simulated samples requiring that each has two identified leptons with opposite charges and that one has $p_T > 20$ GeV/c while the second has $p_T > 10$ GeV/c. These requirements mimic part of the selection for a real analysis and ensure that we could trigger on these events in collision data.
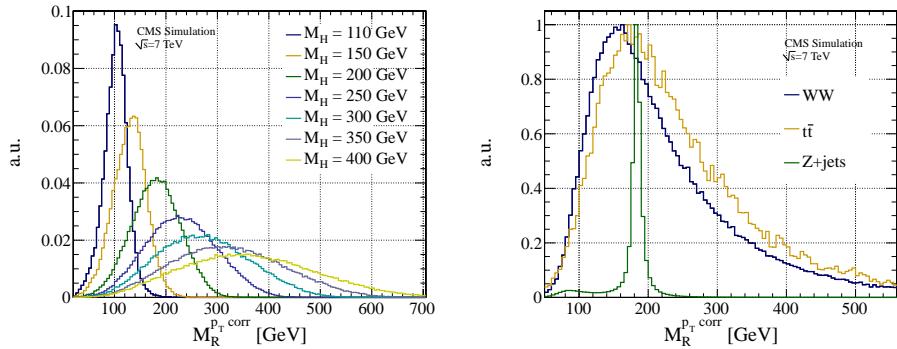

Figure 11.10: (Left) Distribution of $M_R^{p_T-\mathrm{corr}}$ for simulated $H \to WW$ events with varying Higgs mass. (Right) Distribution of $M_R^{p_T-\mathrm{corr}}$ for SM backgrounds in the di-lepton final state.

With this selection applied, we calculate $M_R^{p_T-\mathrm{corr}}$ for signal samples with different Higgs masses and SM background samples. The resulting distributions are shown in figure 11.10. We see that the variable $M_R^{p_T-\mathrm{corr}}$ peaks at the Higgs mass, as expected. The resolution of this mass peak degrades in both relative and absolute terms with increasing M_H, as predicted from previous studies of the resolution dependence of $M_R^{p_T-\mathrm{corr}}$ with increasing γ_{CM}, and also from an increasing intrinsic Higgs mass width. Similarly, the SM backgrounds have features at their own respective mass scales. The $M_R^{p_T-\mathrm{corr}}$ distribution for non-resonant $q\bar{q} \to WW$ events reflects the differential distribution of $\sqrt{\hat{s}}$ for this process (it is topologically identical to the signal process), while the $t\bar{t}$ background distribution mimics the invariant mass distribution of the Ws in these events. Z+jets events, which comprise an especially difficult experimental background to this final state when the leptons have the same flavor due to a large cross section and the potential for missing transverse momentum from jet mis-measurements, peak a the $2M_Z$ pole, with a low tail with the shape of the
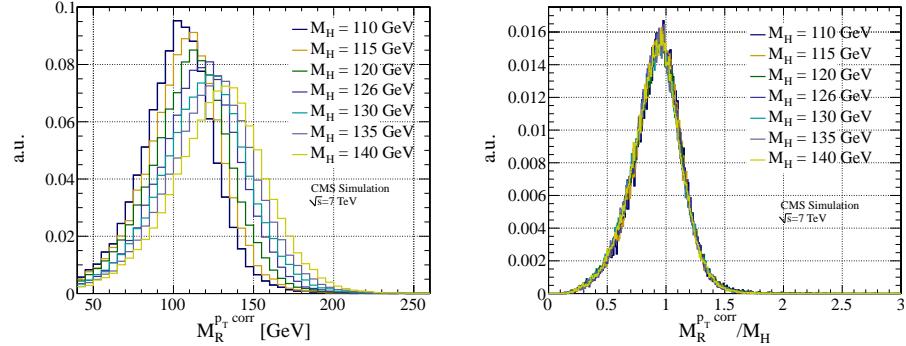
Drell-Yan di-lepton invariant mass distribution.

Figure 11.11: Distribution of $M_R^{p_T-\mathrm{corr}}$ for simulated $H \to WW$ events with varying Higgs mass. (Left) $M_R^{p_T-\mathrm{corr}}$ shown in absolute scale, with peak position equal to the Higgs mass. (Right) $M_R^{p_T-\mathrm{corr}}$ distribution normalized to Higgs mass.

We also consider cases where the Higgs mass is below the $2M_W$ threshold, such that at least one of the W bosons is produced off-shell. Despite violating the decay chain symmetry requirements that were used to derive $M_R^{p_T-\mathrm{corr}}$, we observe in figure 11.11 that $M_R^{p_T-\mathrm{corr}}$ still peaks at the Higgs mass for these events. This property can be understood from the toy studies described in section 9.1.1 and 10.1 where it was demonstrated that razor mass variables are sensitive to an effective *average* characteristic scale of the two decay chains. In this scenario, the average, multiplied by the factor γ_R, is equivalent to one half the Higgs mass. In fact, we see that the variable $M_R^{p_T-\mathrm{corr}}$ can resolve the Higgs mass *better* when it is below the $2M_W$ threshold compared to when it is above. The relative resolution for these off-shell cases is insensitive to the actual value of the Higgs mass, as demonstrated in the rightmost plot in Fig 11.11.

In addition to the mass scale, angular information can be used to discriminate between signal events and non-resonant SM WW production. For example, the azimuthal angle between the two leptons, $\Delta\phi_{\ell\ell}$, is sensitive to the helicities and production mechanism of the W pair and takes smaller values, on average, for signal events (assuming a $J^{PC} = 0^{++}$ Higgs). A challenge when using angles like this to discriminate between processes is that they are not invariant under transverse boosts.

If the WW system has nonzero transverse momentum then the two leptons will become more collimated, biasing this angular distribution to smaller values. This is demonstrated in the leftmost plots of figure 11.12, where we see that an increasing number of jets (indicative of increasing Higgs p_T) implies smaller values of $\Delta\phi_{\ell\ell}$, on average.
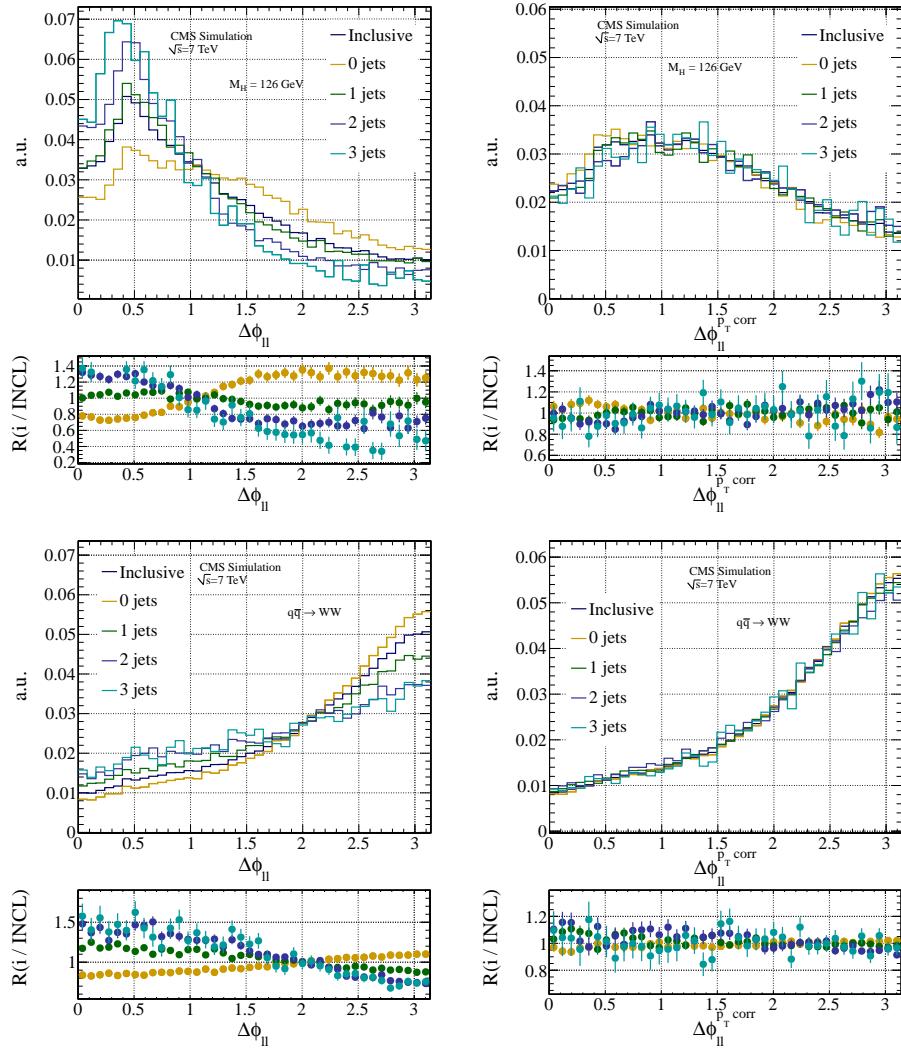

Figure 11.12: Distribution of $\Delta\phi_{\ell\ell}$ for signal and background events in the di-lepton final state. (Top) Simulated signal events with $M_H = 126$ GeV. (Bottom) Non-resonant SM WW events. (Left) $\Delta\phi_{\ell\ell}$ evaluated in the laboratory frame. (Right) $\Delta\phi_{\ell\ell}$ evaluated in the R-frame.

To correct for this effect, we find that we can evaluate angles like $\Delta\phi_{\ell\ell}$ in the

R-frame rather than the lab frame, effectively removing the dependence on the p_T of the WW system. This is demonstrated in the rightmost plots of figure 11.12. $\Delta\phi_{\ell\ell}$ is not the only angle that can be calculated in the R-frame. Another interesting variable is the azimuthal angle between direction of the transverse boost from the lab frame to the R-frame, $\vec{\beta}_{L_RT}$, and the direction of the di-lepton system $(\vec{\ell}_1^R + \vec{\ell}_2^R)$. Since $\vec{\beta}_{L_RT}$ points in the direction of the Higgs candidate momentum, we denote this angle $\Delta\phi_{\ell\ell,\vec{H}}$. This second angle is particularly useful for discriminating between signal events and $Z(\ell\ell)$+jets, demonstrated by the 2D $\Delta\phi_{\ell\ell}$ vs. $\Delta\phi_{\ell\ell,\vec{H}}$ distributions shown in figure 11.13. The reason for this strong discrimination is the fact that, for $Z(\ell\ell)$
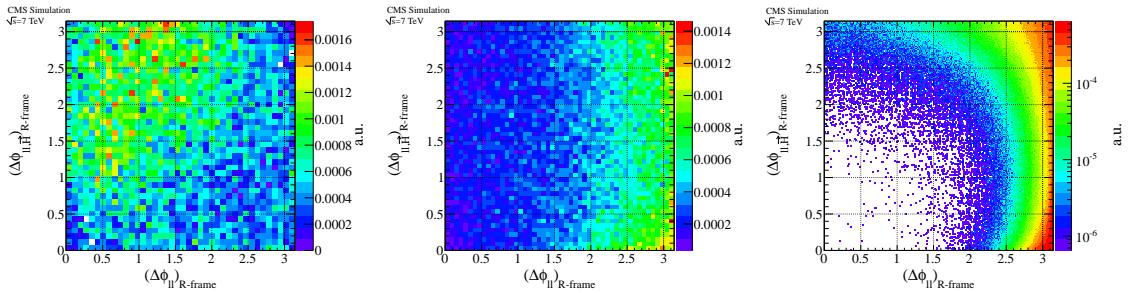

Figure 11.13: Two-dimensional distributions of $\Delta\phi_{\ell\ell}$ and $\Delta\phi_{\ell\ell,\vec{H}}$ evaluated in the R-frame for simulated events in the di-lepton final state. (Left) $H \to WW$ events with $M_H = 126$. (center) Non-resonant SM WW production. (Right) $Z(\ell\ell)$+jets events.

events, both leptons come from the same decay rather than separate decay chains. This means that the R-frame will generally be a good approximation to the Z rest frame, where the two leptons will be back to back with $\Delta\phi_{\ell\ell}$ equal to π. This feature is demonstrated in figure 11.13. A small fraction of the time, fake missing transverse energy can be measured erroneously in events with Z bosons, resulting in a topology faking that of the signal. In these cases, the p_T of the CM system will be incorrectly assigned from the anomalous missing transverse momentum and, as a result, so will the R-frame. For these events, the boost relating the incorrect R-frame to the actual CM will generally point in the direction of the mis-measured transverse momentum, as will $(\vec{\ell}_1^R + \vec{\ell}_2^R)$ from the R-frame boost. The result is that the angle $\Delta\phi_{\ell\ell,\vec{H}}$ is forced to be close to zero when $\Delta\phi_{\ell\ell}$ deviates from π for events where two leptons are back to back in the true CM frame, as we observe in figure 11.13.

Using the interplay between $\Delta\phi_{\ell\ell}$ and $\Delta\phi_{\ell\ell,\vec{H}}$ to reject specific types of backgrounds is not exclusive to this di-lepton case. In final states with jets, such as those studied in the searches for SUSY described in Chapters 9 and 10, events with QCD multi-jet production behave similarly to the $Z(\ell\ell)$ background events here. SM backgrounds without high momentum weakly interacting particles in the final state can mimic signals with missing transverse momentum through mis-measurements of particles and limited detector acceptance. The angles $\Delta\phi_{\ell\ell}$ and $\Delta\phi_{\ell\ell,\vec{H}}$ can be used to suppress this entire class of backgrounds, exploiting the relationship between the R-frame and the true CM frame when these mis-measurements occur.

In general, we observe that variables like $\Delta\phi_{\ell\ell}$ and $\Delta\phi_{\ell\ell,\vec{H}}$ can be evaluated in the R-frame independently of the scale $M_R^{p_T-\text{corr}}$, in that variations of their values around the those evaluated in the true CM frame are uncorrelated with variations of $M_R^{p_T-\text{corr}}$ around the true characteristic scale. This makes an approach based on the razor variables well suited to cases where we need to identify both the scale an angular properties of events, such as for quantum number measurements of the newly discovered boson resonance in this final state. Which variables are most useful to calculate in the R-frame for this case are the subject of future studies.

11.3 Perfect Pairs: Symmetric Two Step Decays

In the previous two sections we have seen that we can extract the characteristic mass scale from events where massive particles are pair produced and each decays to visible and invisible particles. This topology manifests itself in a general class of new physics scenarios and also in the SM, through processes like $H \to WW \to (\ell\nu)(\ell\nu)$ and top quark pair production. The latter case is particularly interesting. When top quarks are pair produced they decay to a b quark and W bosons (no other decays have been observed to date). If the two W bosons from the tops each decay leptonically as $W \to \ell\nu$ then these events result in a striking topology, shown in figure 11.14. In each event, there are two two-body decays in each decay chain, each with identical particle content. Each two body decay produces a visible, detectable particle and the particles coming from the first decays are distinguishable by type from those coming from the second. As was the case for the simple one step decays of the previous sections, we will see that the visible particles in these events contain information about the mass splitting between the other particles participating in their decay, or the characteristic scale of each two-body decay. We denote the class of events described in figure 11.14 the *perfect pairs* topology, such that all of the particles in this final state can have nonzero masses (more general than just top pair production), requiring only that the decay chains are symmetric in particle content. In the following section we derive razor kinematic variables for perfect pairs events, resulting in independent estimators M_{R1} and M_{R2} of the two characteristic scales in these events and the variable γ_R, which scales with the true value of $\gamma_{CM} = \sqrt{\hat{s}}/2M_T$. The variables we propose are not only useful for studying $t\bar{t}$ production, but also a wide class of signal topologies including models with new symmetry partners of top quarks, like stops in SUSY. We examine the properties of these variables in SUSY-like events with stop pair production with decays to b-quarks and charginos.
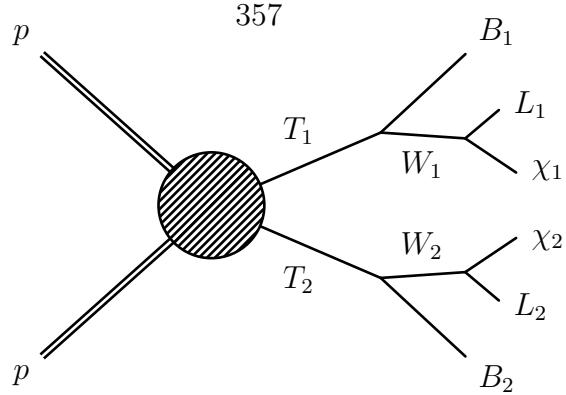
Figure 11.14: The canonical perfect pairs topology. Two massive particles, T_i, are pair produced and each decays as $T_i \rightarrow B_i W_i$, where the particles B_i are reconstructed in the detector. The massive particles W_i decay as $W_i \rightarrow L_i \chi_i$, where the particles L_i are reconstructed in the detector while the particles χ_i are weakly interacting and escape detection.

11.3.1 Derivation of Perfect Pairs Razor Variables

For convenience, we will adopt the naming conventions of di-leptonic $t\bar{t}$ for the particles in the general perfect pairs topology. In this case two top quarks (T_1 and T_2), with equal mass M_T, are produced. Each top quark then decays to a massive W boson, W_1 and W_2 respectively, and a b-quark, B_i. Subsequently, each W boson then decays to a lepton, L_I, and a weakly interacting neutrino, N_i. For the derivation of general perfect pairs variables we allow each particle in the final state to have nontrivial masses (despite our naming conventions).

We denote the center of mass frame of the $T_1 T_2$ system the CM-frame, the rest frames of the top quarks the T_i frames and the rest frames of the W bosons the W_i frames. The four vectors of the b-quarks are denoted by b_i, with three momenta $\vec{b}_i$ and energy B_i. Similarly, we denote the four vectors of the leptons (neutrinos) $\ell_i(\nu_i)$, with three momenta $\vec{\ell}_i(\vec{\nu}_i)$ and energy $L_i(N_i)$. Throughout this discussion the superscripts of these four and three vectors will indicate which reference frame the object is in. For example, L_1^W refers to the energy of lepton 1 in its respective W_1-frame. There will be no instances of any leptons' or b-quarks' energy or momentum being evaluated in the T or W frame of the top or W that it did not come from (no B_1 in the T_2-frame,

etc.). The b and lepton four vectors satisfy the following relations

$$b_i^2 = M_B \quad , \quad \ell_i^2 = M_L \quad , \quad \nu_i^2 = M_\chi \ , \tag{11.27}$$

$$(\nu_i + l_i)^2 = M_W \quad , \quad (b_i + \nu_i + l_i)^2 = M_T \ .$$

The $\sqrt{\hat{s}}$ for this process is defined as $(b_i + \ell_i + \nu_i)^2 = \hat{s} = 4\gamma_{CM} M_T^2$, such that γ_{CM} indicates how far off-threshold the tops are produced. The boost β_{CM} is defined such that $\gamma^2 = 1/(1 - \beta_{CM}^2)$ and $\vec{\beta}_{CM}$ indicates the direction each top is moving in the CM frame (T_1 is moving in direction $\vec{\beta}_{CM}$ while T_2 is moving in direction $-\vec{\beta}$). Let us also define two masses, m_i, such that

$$m_i^2 = (b_i + \ell_i)^2 \ . \tag{11.28}$$

The characteristic scales in perfect pairs events that we would like to measure are related to the momenta of the visible particles, evaluated in rest frames of the particles that produced them. The reference frames relevant to the perfect pairs topology are described in figure 11.15. The decay chain symmetry, and the fact that each of the decays is a two-body, monochromatic decay will lead to two important constraints. Firstly, in its respective T_i-frame each b-quark is recoiling against a W boson. Since the T mass and W mass are fixed, we know that

$$B_1^T = B_2^T = \frac{M_T^2 - M_W^2 + M_B^2}{2M_T} \ . \tag{11.29}$$

Similarly
$$L_1^W = L_2^W = \frac{M_W^2 - M_\chi^2 + M_L^2}{2M_W} \ . \tag{11.30}$$

Hence, if we were able to find the T_i and W_i rest frames, event by event, the magnitude of the b-quark and lepton three momenta in these frame would be equal to these mass differences. In order to find approximations of these reference frames, we will attempt to reconstruct the series of boosts that relate the reference frames shown in figure 11.15, using the principles of the general razor approach described in
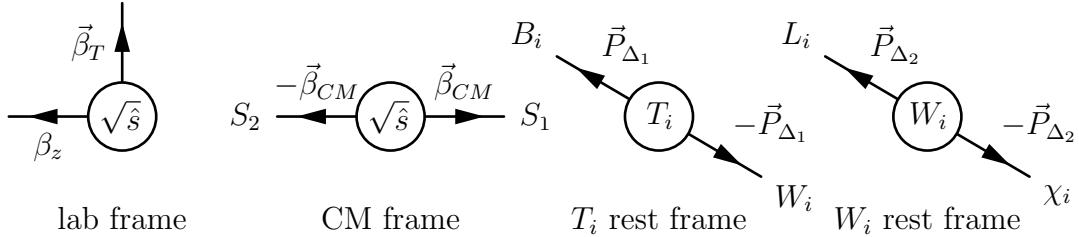
section 11.1.

Figure 11.15: The six reference frames appearing in the perfect pairs topology. (Far left) In the lab frame, the CM system is traveling with a longitudinal velocity β_z and transverse velocity $\vec{\beta}_T$. (Center left) In the CM frame, the two particles T_1 and T_2 are traveling with equal and opposite velocities β_{CM}, with $\sqrt{\hat{s}}$ representing the $T_1 + T_2$ CM object. (Center right) In each of the respective T_i rest frames the particles W_i and B_i are traveling with equal and opposite momentum, with the magnitude of their momentum set by the particle mass differences. (Far right) In the W_i rest frames the lepton and neutrino recoil against each other with a second characteristic momentum.

We begin the derivation in the CM frame, and assume that we have measured the three momenta $\vec{b}_i^{\,CM}$ and $\vec{l}_i^{\,CM}$, along with these particles' masses. We know that each T_i frame is traveling with equal and opposite velocity in this frame, so one way to view the problem is to determine the asymmetric boost, $\vec{\beta}_R$, that will move the b_i^{CM} into their respective T_i frames. Exploiting the symmetry of the decay chains, we also have a constraint for one of the three degrees of freedom of $\vec{\beta}_R$,

$$B_1^T = B_2^T \quad \Rightarrow \quad \gamma_R(B_1^{CM} - \vec{\beta}_R \cdot \vec{b}_1^{CM}) = \gamma_R(B_2^{CM} + \vec{\beta}_R \cdot \vec{b}_2^{CM}) \quad (11.31)$$

$$\Rightarrow \quad \vec{\beta}_R \cdot (\vec{b}_1^{CM} + \vec{b}_2^{CM}) = B_1^{CM} - B_2^{CM} \ .$$

With the value of $\vec{\beta}_R$ parallel to $(\vec{b}_1^{CM} + \vec{b}_2^{CM})$ specified, what remains is to assign values to the other two d.o.f.. We postpone this for the moment and leave $\vec{\beta}_R$ unspecified. Instead, we embark on a short *Gedankenexperiment* to the approximations of the T_i frames that our unknown $\vec{\beta}_R$ will take us to. Let us denote these reference frames the T_i^R frames, for the rough-approximation T frames, noting that $B_1^{T^R} = B_2^{T^R} \equiv B^{T^R}$ is guaranteed from the part of $\vec{\beta}_R$ that we do know.

Now, if $\vec{\beta}_R = \vec{\beta}_{CM}$, then each b quark should be traveling with equal and opposite

momenta to its respective W in its T frame. This implies that we know the direction of the boost that will take the leptons from their T frames to their W frames. Specifically, denoting the boosts that take ℓ_i^T to their respective W frames as $\vec{\xi}_i$ we can express them as

$$\Gamma_i = \sqrt{\frac{1}{1 - |\vec{\xi}_i|^2}} \qquad , \qquad \Gamma_1 = \Gamma_2 \equiv \Gamma \ , \tag{11.32}$$

$$L_i^W = \Gamma(L_i^T - \vec{\xi}_i \cdot \ell_i^T) \qquad , \qquad \vec{\xi}_i = -\frac{\vec{b}_i^T}{\sqrt{(B^T)^2 + M_W^2}} \ .$$

Symmetry requires that $L_1^W = L_2^W$, which gives us the following constraint
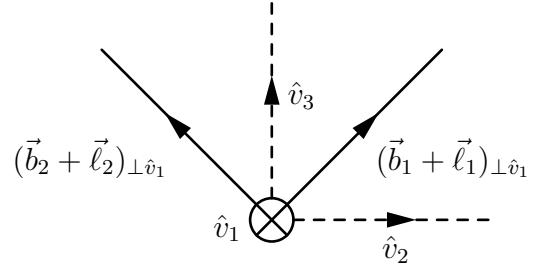
$$
\begin{aligned}
L_1^W \;=\; L_2^W \quad &\Rightarrow \quad \Gamma(L_1^{T_R} - \vec{\xi}_1 \cdot \vec{\ell}_1^{\,T_R}) = \Gamma(L_2^{T_R} - \vec{\xi}_2 \cdot \vec{\ell}_2^{\,T_R}) \\
&\Rightarrow \quad L_1^{T_R} - L_2^{T_R} = \vec{\xi}_1 \cdot \vec{\ell}_1^{\,T_R} - \vec{\xi}_2 \cdot \vec{\ell}_2^{\,T_R} = \frac{\vec{b}_2^{\,T_R} \cdot \vec{\ell}_2^{\,T_R} - \vec{b}_1^{\,T_R} \cdot \vec{\ell}_1^{\,T_R}}{\sqrt{(B^{T_R})^2 + M_W^2}} \\
&\Rightarrow \quad \sqrt{(B^{T_R})^2 + M_W^2}(L_1^{T_R} - L_2^{T_R}) = \frac{m_1^2 - m_2^2}{2} + L_2^{T_R} B^{T_R} - L_1^{T_R} B^{T_R} \\
&\Rightarrow \quad \frac{m_1^2 - m_2^2}{2} = (B^{T_R} + \sqrt{(B^{T_R})^2 + M_W^2})(L_1^{T_R} - L_2^{T_R}) = M_T(L_1^{T_R} - L_2^{T_R}) \\
&\Rightarrow \quad M_T = \frac{m_1^2 - m_2^2}{2(L_1^{T_R} - L_2^{T_R})} \ .
\end{aligned}
\tag{11.33}
$$

Having peered down the road assuming we knew the other two d.o.f. of $\vec{\beta}_R$ we see that we could calculate B^{T_R}, L^W, γ_R, M_T, and, subsequently from the relations we already know, the masses M_W and M_χ – all of the unknown parameters in this event. To understand how to assign values to these d.o.f. we first expand $\vec{\beta}_R$ in a basis of orthonormal vectors $\hat{v}_i$ defined in figure 11.16.

In this basis, $\vec{\beta}_R$ can be expressed as

$$\vec{\beta}_R = \beta_1 \hat{v}_1 + \beta_2 \hat{v}_2 + \beta_3 \hat{v}_3 \ , \tag{11.34}$$

with $\beta_1 = \frac{B_1^{CM} - B_2^{CM}}{|\vec{b}_1^{\,CM} + \vec{b}_2^{\,CM}|}$ following from equation (11.32). To assign β_2 and β_3 we appeal to a slightly different problem. If we view each pair $(L_i + B_i)$ as a single particle, then the perfect pairs event can be interpreted as the canonical razor topology of

$$\hat{v}_1 \equiv \frac{\vec{b}_1 + \vec{b}_2}{|\vec{b}_1 + \vec{q}_b|} \quad \hat{v}_2 \equiv \frac{(\vec{b}_1 + \vec{\ell}_1)_{\perp \hat{v}_1} - (\vec{b}_2 + \vec{\ell}_2)_{\perp \hat{v}_1}}{|(\vec{b}_1 + \vec{\ell}_1)_{\perp \hat{v}_1} - (\vec{b}_2 + \vec{\ell}_2)_{\perp \hat{v}_1}|} \quad \hat{v}_3 \equiv \hat{v}_2 \times \hat{v}_1$$

Figure 11.16: The momenta of the b-quarks and leptons in the CM frame. The diagram shows the plane perpendicular to $(\vec{b}_1 + \vec{b}_2)$. An orthonormal basis $\hat{v}_i$ is defined using the b's and leptons' momenta. The subscript $\perp \hat{v}_1$ indicates the component of a vector perpendicular to $\hat{v}_1$.

section 11.1, with two massive tops decaying to a a visible particle $(L_i + B_i)$ and an invisible neutrino. Even though our reinterpreted event lacks the decay chain symmetry usually required, there is a different property we would like to borrow from the canonical razor derivation. We observe that if we want γ_R to scale with the true value of γ_{CM} then we should choose β_2 and β_3 which minimize the quantity $(B_1^{T_R} + B_2^{T_R} + L_1^{T_R} + L_2^{T_R})$, corresponding to our estimator for the razor characteristic scale of the asymmetric simplified toplology. Requiring

$$\frac{\partial (B_1^{T_R} + B_2^{T_R} + L_1^{T_R} + L_2^{T_R})}{\partial \beta_2} = \frac{\partial (B_1^{T_R} + B_2^{T_R} + L_1^{T_R} + L_2^{T_R})}{\partial \beta_3} = 0 , \quad (11.35)$$

results in $\beta_3 = 0$ and β_2 such that β_R expressed as

$$\vec{\beta}_R = \frac{B_1^{CM} - B_2^{CM}}{|\vec{b}_1^{CM} + \vec{b}_2^{CM}|} \hat{v}_1 + \frac{[(\vec{b}_1^{CM} + \vec{\ell}_1^{CM}) - (\vec{b}_2^{CM} + \vec{\ell}_2^{CM})] \cdot \hat{v}_2}{B_1^{CM} + B_2^{CM} + L_1^{CM} + L_2^{CM}} \hat{v}_2 . \quad (11.36)$$

To calculate our approximation of the boost $\vec{\beta}_L$ which relates the lab frame and the CM frame we reuse another concept from the general razor derivation. By interpreting two times the sum of the b-quark and lepton energies in the CM frame as the effective $\sqrt{\hat{s}}$ for the event, $\sqrt{\hat{s}_R}$, we can solve for the boosts $\beta_{L_R z}$ and $\vec{\beta}_{L_R z}$ which move from

the lab frame along the beam axis and in the transverse direction, respectively, to our approximation of the CM frame, denoted the R frame. For perfect pairs events, we find that these boosts can be expressed as

$$\beta_{L_R z} = \frac{b_{1z}^l + b_{2z}^l + \ell_{1z}^l + \ell_{2z}^l}{B_1^l + B_2^l + L_1^l + L_2^l} \quad , \quad \vec{\beta}_{L_R T} = \frac{\vec{P}_T^{CM}}{\sqrt{|\vec{P}_T^{CM}|^2 + \hat{s}_R}} \ , \qquad (11.37)$$

where $\vec{P}_T^{CM} = \vec{b}_{1T}^{\,l} + \vec{b}_{2T}^{\,l} + \vec{\ell}_{1T}^{\,l} + \vec{\ell}_{2T}^{\,l} + \vec{M}$. Defining

$$\sum E = B_1^l + B_2^l + L_1^l + L_2^l \quad , \quad \sum \vec{P} = \vec{b}_1^{\,l} + \vec{b}_2^{\,l} + \vec{\ell}_1^{\,l} + \vec{\ell}_2^{\,l} \ , \qquad (11.38)$$

$\hat{s}_R$ is given by

$$
\begin{aligned}
\hat{s}_R \ &= \ 2\left(\sum E\right)^2 - 2\vec{P}^{CM} \cdot \left(\sum \vec{P}\right) \qquad\qquad\qquad (11.39)\\
&\pm \ 2\left(\sum E\right) \sqrt{\left(\sum E\right)^2 + |\vec{P}^{CM}|^2 - 2\vec{P}^{CM} \cdot \left(\sum \vec{P}\right)} \ .
\end{aligned}
$$

We can now summarize the complete definitions of our kinematic variables. The R frame is now defined as the reference frame resulting from the longitudinal boost $\beta_{L_R z}$ from the lab frame followed by the transverse boost $\vec{\beta}_{L_R T}$. The T_R frames are defined as those resulting from an asymmetric boost β_R applied to the b-quarks and leptons in the R frame, given by

$$\vec{\beta}_R = \frac{B_1^R - B_2^R}{|\vec{b}_1^{\,R} + \vec{b}_2^{\,R}|} \hat{v}_1 + \frac{[(\vec{b}_1^{\,R} + \vec{\ell}_1^{\,R}) - (\vec{b}_2^{\,R} + \vec{\ell}_2^{\,R})] \cdot \hat{v}_2}{B_1^R + B_2^R + L_1^R + L_2^R} \hat{v}_2 \ . \qquad (11.40)$$

The energies of the visible particles in their respective T_R frames are defined as

$$B_1^{T_R} = \gamma_R(B_1^R - \vec{\beta}_R \cdot \vec{b}_1^{\,R}) \quad , \quad B_2^{T_R} = \gamma_R(B_2^R + \vec{\beta}_R \cdot \vec{b}_2^{\,R}) \ , \qquad (11.41)$$

$$L_1^{T_R} = \gamma_R(L_1^R - \vec{\beta}_R \cdot \vec{\ell}_1^{\,R}) \quad , \quad L_2^{T_R} = \gamma_R(L_2^R + \vec{\beta}_R \cdot \vec{\ell}_2^{\,R}) \ .$$

The first characteristic scale variable, M_{R1}, is designed to be sensitive to the mass

splitting of the first decay in each chain, and is equal to

$$M_{R1} = B_1^{T_R} + B_2^{T_R} = 2B^{T_R} \sim M_{\Delta 1} = \frac{M_T^2 - M_W^2 + M_B^2}{M_T} \; . \tag{11.42}$$

Our estimates of the top and W masses, M_T^R and M_W^R, are given by

$$M_T^R = \frac{m_1^2 - m_2^2}{2(L_1^{T_R} - L_2^{T_R})} \quad , \quad M_W^R = \sqrt{(M_T^R)^2 + M_B^2 - 2M_T^R B^{T_R}} \; . \tag{11.43}$$

In order to calculate the second scale variable, M_{R2}, we must evaluate the lepton energies in our approximations of the W rest frames, denoted the W_R frames. M_{R2} can be expressed as

$$\begin{aligned}
M_{R2} &= L_1^{W_R} + L_2^{W_R} = 2L^{W_R} \\
&= \frac{1}{2M_W} \left[\frac{(m_1^2 - m_2^2)(L_1^{T_R} + L_2^{T_R})}{L_1^{T_R} - L_2^{T_R}} + M_B^2 + M_L^2 - m_1^2 - m_2^2 \right] \\
&\sim M_{\Delta 2} = \frac{M_W^2 - M_\chi^2 + M_L^2}{M_W} \; .
\end{aligned} \tag{11.44}$$

In the following sections we will study the phenomenology of the variables M_{R1}, M_{R2} and γ_R in perfect pairs events.

11.3.2   The $t\bar{t}$ Pole

As a first test of the variables M_{R1}, M_{R2} and γ_R we examine their behavior in events with di-leptonic $t\bar{t}$ decays. The toy event simulation described in section 11.1 is used to generate sample $t\bar{t}$ events, using the SM values for the masses of each of the particles in the final state and flat matrix elements for each of the decays ($t\bar{t}$ spin correlations are neglected). We also compare the results from this toy simulation to events simulated with the full CMS detector. For the latter, we select simulated $t\bar{t}$ events with two reconstructed leptons with $p_T > 20$ GeV/c and lepton identification identical to that used in the CMS SUSY search described in Chapter 10. We also use the b-tagging discriminant from that search to identify our two b-quark candidates in

these samples. Selected events are required to have at least two reconstructed jets with $p_T > 30$ GeV/c. We choose our two b-jets from the collection of jets satisfying this requirement by picking the two with the largest values of the b-tag discriminant. In both our toy simulation and CMS simulated events there is an ambiguity as to which b-jet to pair with each lepton. We choose the combination that minimizes $m_1^2 + m_2^2$, with these masses defined in equation (11.28). As β_{CM} increases, this choice will result in the correct pairing a larger fraction of the time.

In figure 11.17 we compare the distributions for M_{R1} and M_{R2} for $t\bar{t}$ events from the toy simulation and CMS simulation. In both cases we observe the expected peaking behavior at the characteristic scales $M_{\Delta 1} = \frac{M_T^2 - M_W^2}{M_T}$ and $M_{\Delta 2} = \frac{M_W^2 - M_\nu^2}{M_W} \sim M_W$. The M_R peak is biased slightly low in CMS simulated events, most likely due to imperfect jet energy scale calibrations for b-tagged jets. The resolution of both peaks is quite similar between the two simulations.
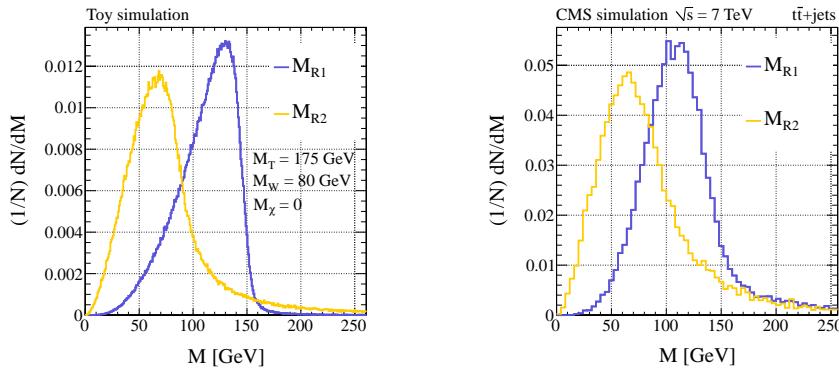

Figure 11.17: Distributions of M_{R1} and M_{R2} for simulated di-leptonic $t\bar{t}$ events. (Left) A toy simulation with only kinematic effects taken into account. (Right) Events produced with CMS full simulation.

We observe that a shortcoming to this perfect pairs razor approach is that M_{R2} is not always well-defined. The fraction of the time this occurs can vary between 30% and 60%, depending and the kinematics of the event, with the fraction increasing for larger γ_{CM}. M_{R1} is well defined for almost all events ($\lesssim 2\%$ ill-defined). An analysis based on these variables would most likely require different treatments of events with and without well-defined M_{R2}.
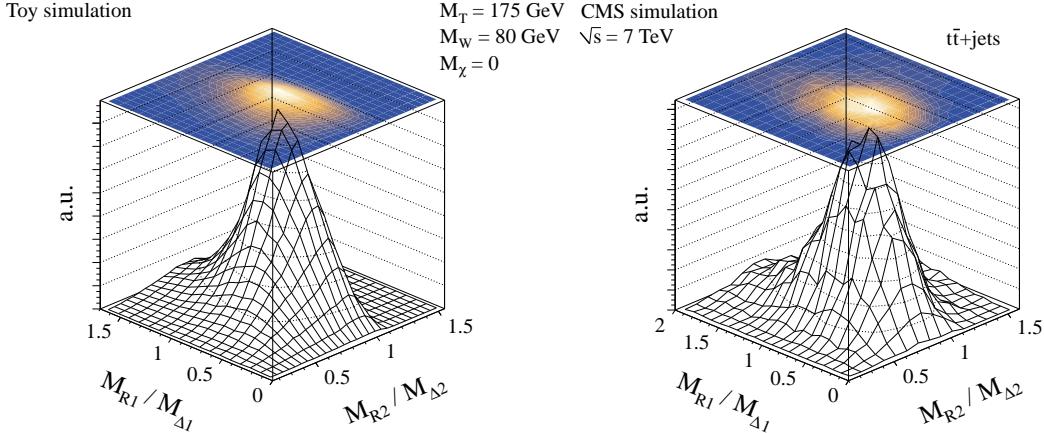
Figure 11.18: Distributions of $M_{R1}/M_{\Delta 1}$ vs. $M_{R2}/M_{\Delta 2}$ for simulated di-leptonic $t\bar{t}$ events. (Left) A toy simulation with only kinematic effects taken into account. (Right) Events produced with CMS full simulation.

The next property to understand is the correlation between the calculated values of M_{R1} and M_{R2}. We see in figure 11.18 that the two quantities are nearly completely uncorrelated, resulting in a prominent two dimensional peak. Additionally, we observe that our simple toy simulation successfully reproduces the qualitative features of the fully simulated events. This is not surprising, since the intrinsic resolution of these peaks are larger than the resolution of the measured particles' momentum used to calculate them. As a result, the overall performance is dictated predominantly by just the underlying kinematics. This gives us some confidence in the predictions of the toy simulation when we systematically vary the masses in the following section.

The last property we would like to examine in $t\bar{t}$ events is the scaling of γ_R with the true value of γ_{CM}. If the two top quarks are produced in the decay of a heavy resonance, G, then γ_{CM} will have a fixed value at a scale related to the top and G mass splitting, which we would like to be sensitive to. We consider a scenario where two tops are produced in the decay of a particle, G, with mass M_G is chosen such that $\gamma_{CM} = 3$. The distributions of $M_{R1}/M_{\Delta 1}$, $M_{R2}/M_{\Delta 2}$ and γ_R/γ_{CM} for these events are shown in figure 11.19. We observe that we can resolve all three quantities in nearly uncorrelated independent measurements with relatively good resolution. The result is a three-dimensional peak in these three variables, which would be a striking signal
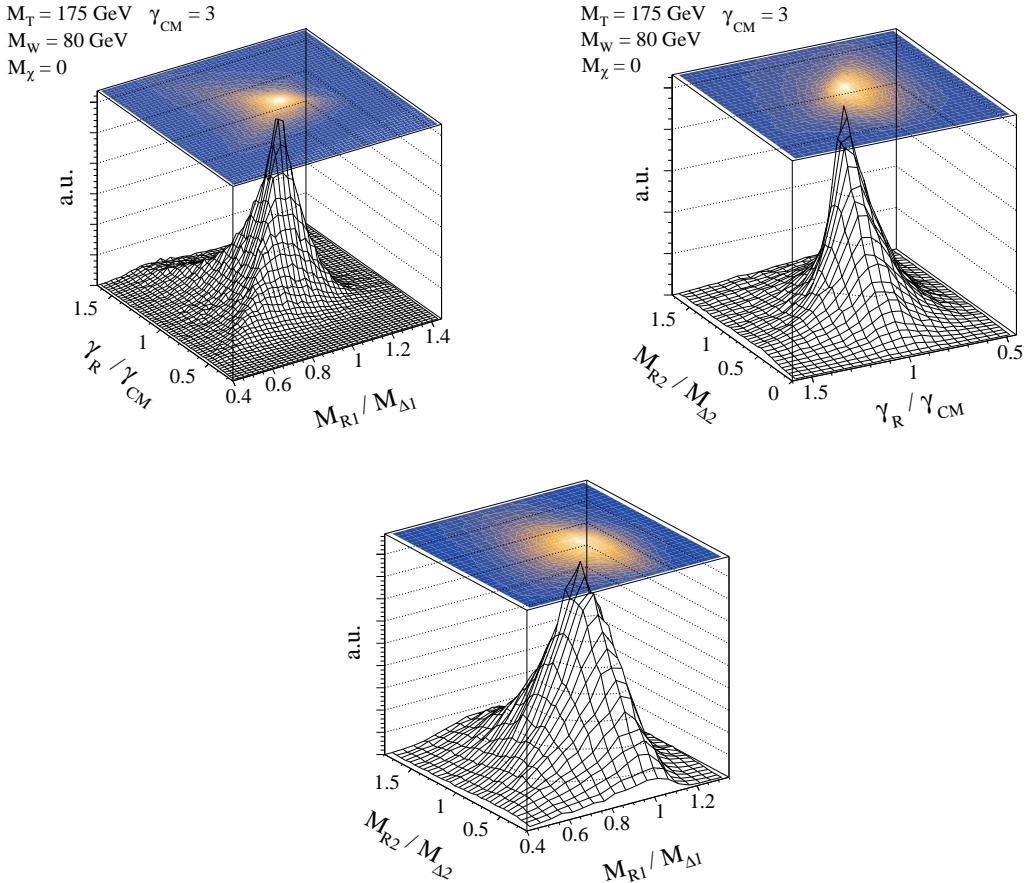
Figure 11.19: Distributions of $M_{R1}/M_{\Delta 1}$, $M_{R2}/M_{\Delta 2}$ and γ_R/γ_{CM} for simulated dileptonic $t\bar{t}$ events produced through a heavy resonance G with $M_G/2M_T = \gamma_{CM} = 3$. (Upper left) $M_{R1}/M_{\Delta 1}$ vs. γ_R/γ_{CM}. (Upper right) $M_{R2}/M_{\Delta 2}$ vs. γ_R/γ_{CM}. (Bottom) $M_{R1}/M_{\Delta 1}$ vs. $M_{R2}/M_{\Delta 2}$.

if events of this type appear in data.

11.3.3 Perfect Pairs for Signal Events

In models that attempt to address the hierarchy problem there appear many new particle production possibilities that satisfy the perfect pairs topology. One example is the pair production of SUSY stops which decay to a chargino, snuetrino and neutralino, illustrated in figure 11.20. Each of the decays is two-body, with the b-quarks and leptons produced in the first and second decay steps, respectively, just as for di-leptonic $t\bar{t}$ events. In this case both of the sneutrino decay products are weakly interacting and escape detection. As a result, the sneutrino plays the role of a heavy neutrino in the perfect pairs topology.
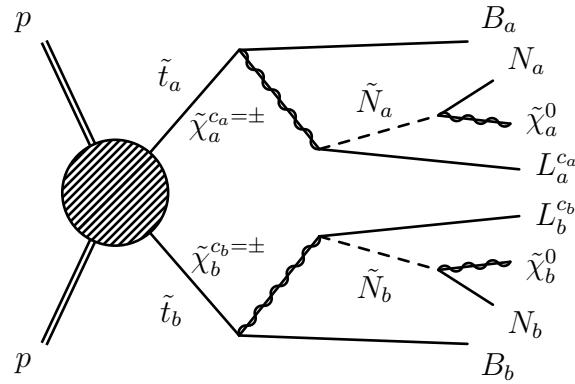
Figure 11.20: A perfect pair signal topology. Here, a pair of stops are pair produced (labelled a and b to avoid confusion with the different stop mass eigenstates often denoted as 1 and 2) and undergo two-body decays to b-quarks and charginos. The charginos then decay to leptons and sneutrinos. There are two additional particles in the final state from each subsequent sneutrino decay (a neutralino and neutrino) which are both weakly interacting. As a result, the sneutrinos play the role of the neutrinos in the analogous perfect pairs case for $t\bar{t}$ events.

For these di-stop events the two characteristic scales which we are sensitive to are

$$M_{\Delta 1} = \frac{M_{\tilde{t}}^2 - M_{\tilde{\chi}^+}^2}{M_{\tilde{t}}} \quad , \quad M_{\Delta 2} = \frac{M_{\tilde{\chi}^+}^2 - M_{\tilde{N}}^2}{M_{\tilde{\chi}^+}} \; . \tag{11.45}$$

Whether or not an analysis using these variables would be sensitive to a particular signal model depends on how different the scales in the model's mass spectrum are

from those of the SM and how well we are able to resolve those scales. In the case of di-lepton and di-b-tagged jet final states, this is reduced to a question of whether it is possible to resolve the peaking signal in the presence of the multidimensional top peak. In this section we examine a wide range of perfect pairs topologies by systematically varying the parameter γ_{CM} and the mass spectrum, observing how well the razor variables can resolve the quantities that they approximate.

We first test the dependence of the variables M_{R1} and M_{R2} on γ_{CM}, shown in figure 11.21. The variables are able to successfully resolve the correct scales up to large values of γ_{CM}, with the performance of the variables degrading with increasing γ_{CM}. We find that, as γ_{CM} gets bigger, the information about the characteristic scales is progressively washed-out.
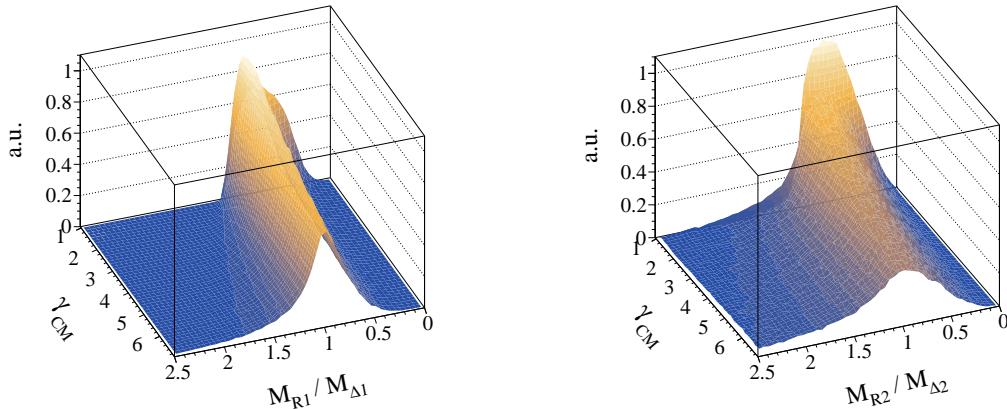

Figure 11.21: (Left) Distribution of $M_{R1}/M_{\Delta 1}$ as a function of γ_{CM}. (Right) Distribution of $M_{R2}/M_{\Delta 2}$ as a function of γ_{CM}. Events are generated from a toy simulation with $M_T = 400$ GeV, $M_W = 200$ GeV and $M_\chi = M_B = M_L = 0$.

Next, we test what happens as each of the two mass splittings becomes large or small relative to the other. With M_T and M_χ masses fixed, we vary M_W between the two values. The resulting distributions for M_{R1} are shown in figure 11.22. We consider two cases: one where $M_\chi = 0$ and the other with $M_\chi = 200$ GeV $= M_T/2$. We observe that for small first mass splitting (M_T-M_W) the two cases behave almost identically. M_{R1} is able to successfully resolve the correct scale unless the mass splitting approaches zero and the correspondence with $M_{\Delta 1}$ is lost. When this first

mass splitting is large, the two cases behave slightly differently, with the nonzero M_χ case resulting in a distortion of M_{R1} that does not appear when $M_\chi = 0$.
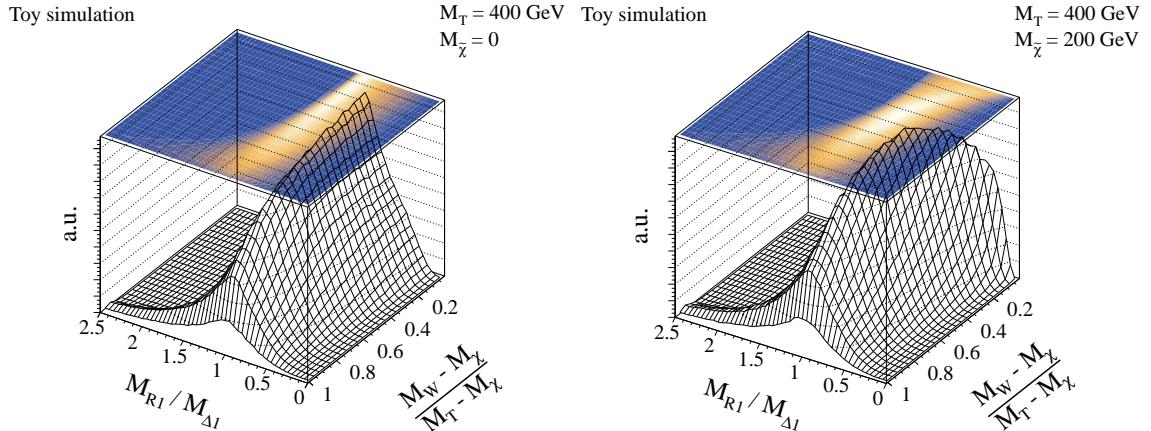

Figure 11.22: Distribution of $M_{R1}/M_{\Delta 1}$ as a function of M_W for two different scenarios. (Left) $M_\chi = 0$. (Right) $M_\chi = 200$ GeV and other masses fixed. Events are generated from a toy simulation with $M_T = 400$ GeV and $M_B = M_L = 0$.

Performing the same exercise for M_{R2}, we see in figure 11.23 that the two M_χ cases yield effectively identical results when varying M_W and the mass splittings. When the second mass splitting $(M_W - M_\chi)$ is large M_{R2} is able to identify the scale $M_{\Delta 2}$, peaking at the right value. As the second mass splitting gets smaller (and the leptons get softer relative to the b-quarks) information about this scale gets increasingly worse, similarly to M_{R1} and the first mass splitting. We see that the behavior is as expected: For the majority of possible scenarios for these masses, the variables M_{R1} and M_{R2} are able to resolve the correct characteristic scales. Naturally, their ability to do so degrades as the corresponding mass splittings get smaller.

Finally, we evaluate our procedure for correcting for nonzero p_T^{CM}. In figure 11.24 we compare the distributions of M_{R1} and M_{R2} as a function of p_T^{CM}, with and without p_T corrections using the transverse boost $\vec{\beta}_{L_R T}$. We find that the procedure works as designed. Rather than being washed-out with increasing p_T^{CM} as is the case without a p_T correction, M_{R1} and M_{R2} are nearly independent of p_T^{CM}.

The perfect pairs razor variables are a potentially powerful tool for studying events consistent with this topology for a wide range of mass splittings and kine-
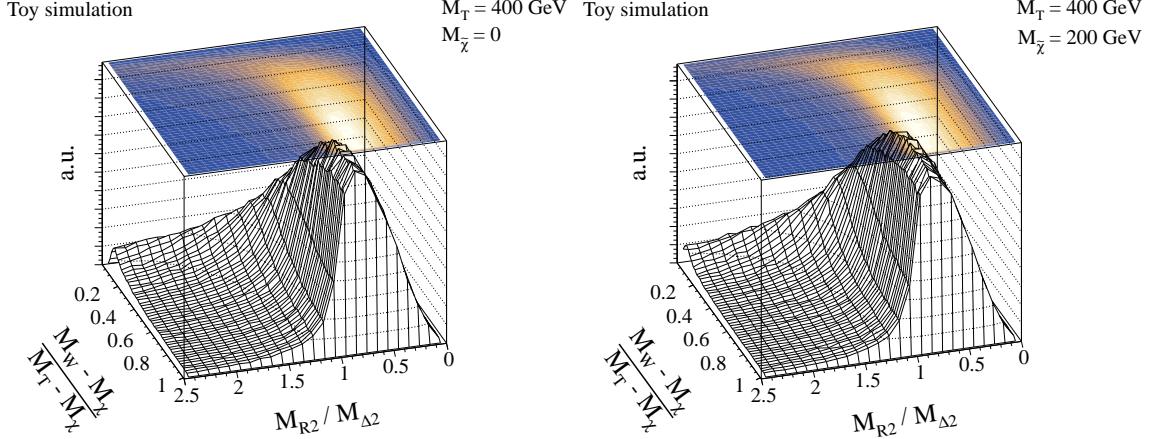
Figure 11.23: Distribution of $M_{R2}/M_{\Delta 2}$ as a function of M_W for two different scenarios. (Left) $M_\chi = 0$. (Right) $M_\chi = 200$ GeV and other masses fixed. Events are generated from a toy simulation with $M_T = 400$ GeV and $M_B = M_L = 0$.
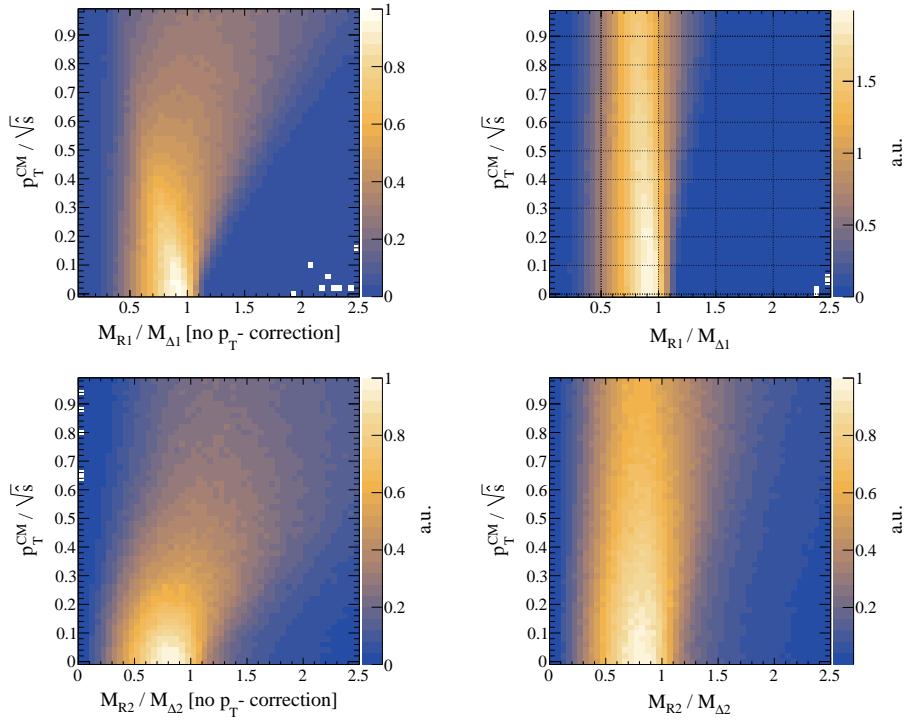

Figure 11.24: (Top) Distribution of M_{R1} as a function of the transverse momenta of the CM frame, p_T^{CM}. (Bottom) M_{R2} as a function of p_T^{CM}. (Left) The transverse boost $\vec{\beta}_{L_RT}$ is not applied to correct for nonzero p_T. (Right) $\vec{\beta}_{L_RT}$ p_T corrections are applied. Events are generated with $M_T = 400$ GeV, $M_W = 200$ GeV and $M_\chi = M_B = M_L = 0$.

matic regimes. In the following section we will see that they can also be useful for cases which aren't perfect pairs.

11.3.4 Almost Perfect Pairs

In addition to the many possible signal processes that satisfy the perfect pair symmetry requirements there is a collection that violates one or more symmetry conditions but for which the perfect pairs razor variables can still be useful. An example is $\tilde{t} \rightarrow b\tilde{\chi}^+$ pair production, where the charginos decay to W bosons and neutralinos, illustrated in figure 11.25. If the Ws are on-shell and decay hadronically, the dijet W resonances can be treated as single objects in a perfect pairs approach. If the Ws both decay leptonically the presence of two extra neutrinos in the final state breaks the perfect pairs symmetry. The lepton will no longer carry exactly half of the momentum away from the chargino decay. Furthermore, if the chargino-neutralino mass splitting is smaller than M_W the W will be produced off-shell, resulting in potentially different masses between the two decay chains. In this section, we study the behavior of the perfect pair razor variables for these nonperfect cases.
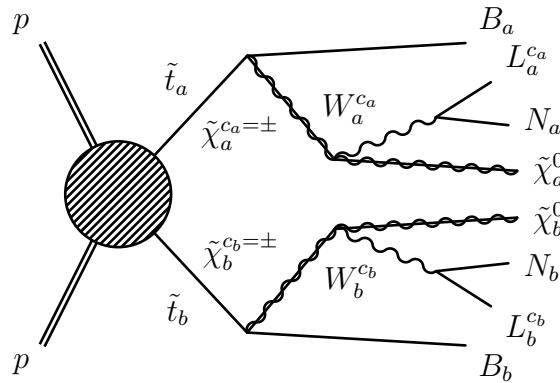

Figure 11.25: An almost perfect pair signal topology. Here, a pair of stops are pair produced and undergo two-body decays to b-quarks and charginos. The charginos then decay to W bosons and neutralinos. Leptons and neutrinos are produced in the decay of the W bosons. If the $\tilde{\chi}^+ - \tilde{\chi}^0$ mass splitting is not sufficiently large to accommodate an on-shell W they can be produced off-shell, with masses varying event by event. The presence of two extra weakly interacting particles breaks the perfect pairs symmetry.

We consider a toy simulation of these $\tilde{t}$ pair production events, adapted to account for off-shell decays of Ws. In these cases, each W mass is independently sampled from a nontrivial distribution modeling M_{W*} in the kinematically allowed phase space afforded by the chargino-neutralino mass splitting. Results using several different probability distribution functions for M_{W*} are compared, with the distributions considered shown in figure 11.27. We observe no perceivable differences between the behavior of the perfect pairs razor variables when varying the M_{W*} distribution. In fact, we see almost identical behavior between off-shell and on-shell W cases in that the variables are sensitive to *effective* scales, independent of the W mass.
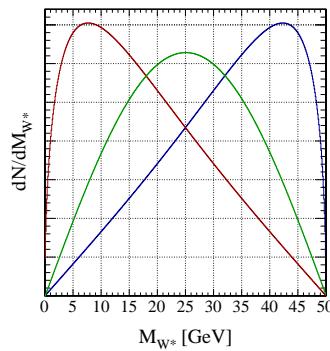

Figure 11.26: Distributions of the off-shell W mass for events with a chargino-neutralino mass splitting of 50 GeV. Three different types of probability distribution functions are considered for the off-shell W mass, represented by the different colored curves. We observe that the behavior of the perfect pairs razor variables are indistinguishable between the three distributions for M_{W*}.

For the case we are discussing, we define effective mass scales $M_{\Delta 1}$ and $M_{\Delta 2}^{eff}$ in terms of the SUSY particle masses

$$M_{\Delta 1} = \frac{M_{\tilde{t}}^2 - M_{\tilde{\chi}^+}^2}{M_{\tilde{t}}} \quad , \quad M_{\Delta 2}^{eff} = \frac{M_{\tilde{\chi}^+}^2 - M_{\tilde{\chi}^0}^2}{2M_{\tilde{\chi}^+}} \, , \tag{11.46}$$

with the scale of the second mass splitting taking one half the value for the canonical perfect pairs case. The reason is that the momentum from the chargino decay is shared evenly between the Ws and neutralinos. This momentum, along with energy from the W mass, is then split further between the lepton and neutrino.

The distributions of γ_R, M_{R1} and M_{R2} are evaluated while varying γ_{CM} for a benchmark signal case, shown in figure 11.21. We observe that, even at large γ_{CM}, each of the variables is sensitive to its corresponding effective scale. M_{R1} has a kinematic edge at $M_{\Delta 1}$ while M_{R2} peaks at $M_{\Delta 2}^{eff}$, both with better resolution of their respective scales at smaller γ_{CM}. Similarly, the γ_R distribution scales with γ_{CM}, with an increasing bias at larger values.
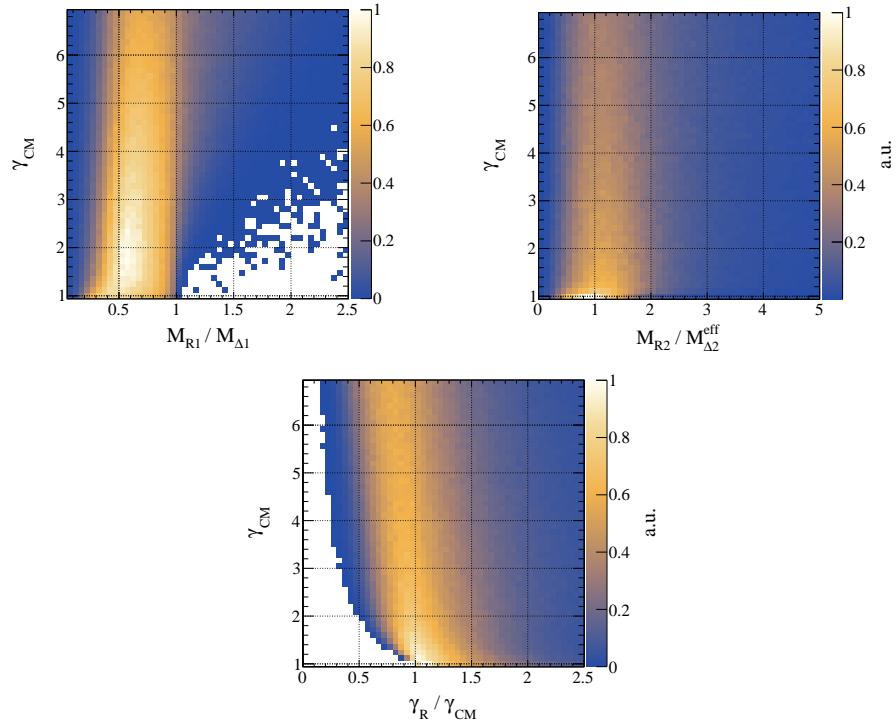

Figure 11.27: (Left) Distribution of $M_{R1}/M_{\Delta 1}$ as a function of γ_{CM}. (Right) Distribution of $M_{R2}/M_{\Delta 2}^{eff}$ as a function of γ_{CM}. Events are generated from a toy simulation with $M_{\tilde{t}} = 400$ GeV, $M_{\tilde{\chi}^+} = 150$ GeV and $M_{\tilde{\chi}^0} = 100$ GeV.

In the distributions of M_{R1} shown in figure 11.28 the value of the chargino mass is varied while keeping the other masses fixed, effectively probing the whole range of effective mass splitting for the decays in these events. The variables behave qualitatively in the same way for these events as for perfect pair topologies under the same variation. The M_{R1} edge at $M_{\Delta 1}$ persists for the whole range of mass splittings, losing resolution when the first mass splitting ($\tilde{t} - \tilde{\chi}^+$) approaches zero. In figure 11.29 we observe that M_{R2} behaves similarly, losing focus of the effective scale $M_{\Delta 2}^{eff}$ when the
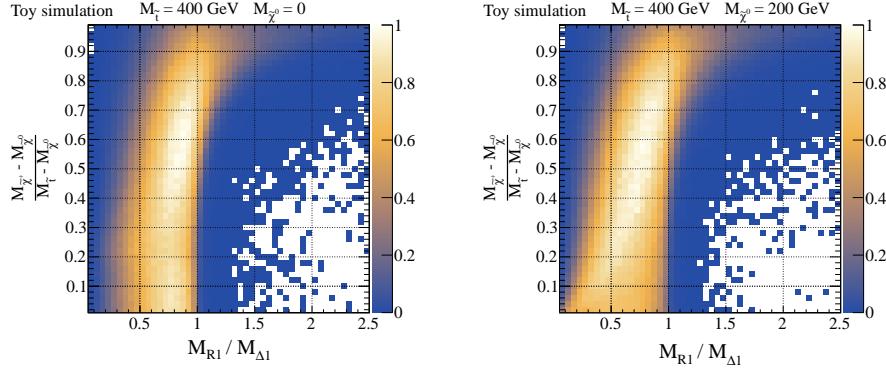
Figure 11.28: Distribution of $M_{R1}/M_{\Delta1}$ as a function of $M_{\tilde{\chi}^+}$ for two different scenarios. (Left) $M_{\tilde{\chi}^0} = 0$. (Right) $M_{\tilde{\chi}^0} = 200$ GeV and other masses fixed. Events are generated from a toy simulation with $M_{\tilde{t}} = 400$ GeV.

second mass splitting $(\tilde{\chi}^+ - \tilde{\chi}^0)$ gets small. We also see that nonzero $M_{\tilde{\chi}}^0$ can change $M_{\Delta2}^{eff}$ for the peak position of M_{R2} and also induce small effects in M_{R1} at small second mass splittings.
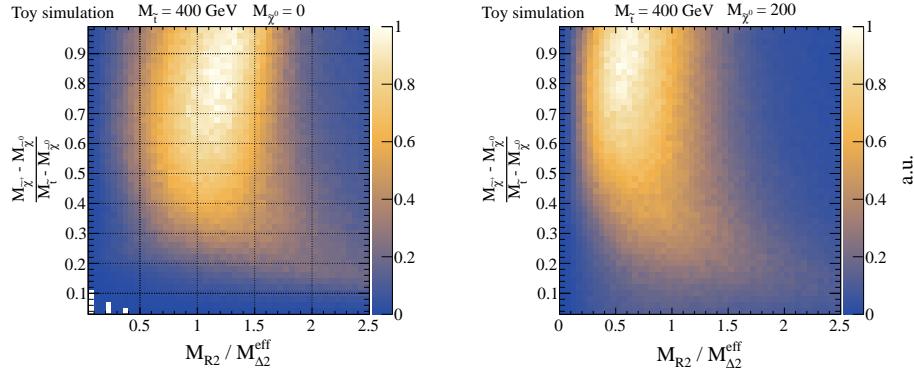

Figure 11.29: Distribution of $M_{R2}/M_{\Delta1}^{eff}$ as a function of $M_{\tilde{\chi}^+}$ for two different scenarios. (Left) $M_{\tilde{\chi}^0} = 0$. (Right) $M_{\tilde{\chi}^0} = 200$ GeV and other masses fixed. Events are generated from a toy simulation with $M_{\tilde{t}} = 400$ GeV.

In summary, we see that perfect pairs for nonperfect events can still be useful. Even though the chargino decays from our example nonperfect case are effectively three body decays to leptons and a weakly interacting system of particles, we find that by imposing the perfect-pair symmetry conditions causes the variations in decays to decrease in magnitude by effectively averaging the scale of the two decay chains. As a result, these events will peak in the three-dimensional M_{R1}, M_{R2}, γ_R space

of observables with its own characteristic scales, albeit with worse resolution than perfect pair events. Regardless, gaining sensitivity to these scales could facilitate the discovery of these new phenomena, if they exist and if our lenses are sharp enough.

Appendix A

Chapter 7 Appendix: Spinning the Higgs

A.1   Simulation, Detector Modeling and Selection

A.1.1   Event Generation

The knowledge of the four-momenta of the leptons fully specifies the information needed in this analysis. We generate the four-momenta of the leptons from the five- or six-dimensional probability density functions (*pdfs*) of

$$
\begin{aligned}
\vec{X} &\equiv \{\vec{\omega}, \vec{\Omega}\} \ \text{ for } ZZ \ , \\
\vec{X} &\equiv \{\vec{\omega}, \ \vec{\Omega}, \ M_{Z^*}\} \ \text{ for } ZZ^* \ ,
\end{aligned}
\tag{A.1}
$$

where $\vec{\Omega}, \vec{\omega}$ are given in equation (7.4). The $\vec{X}$ quantities are generated in the rest frame of the decaying resonance. The muons are then boosted to the laboratory frame, and the detector effects (acceptance, efficiency and resolution) are applied to the boosted momenta. We use the azimuthal symmetry of the LHC detectors to reduce the remaining kinematic degrees of freedom to the knowledge of the p_T, η and the invariant mass $m_{4\mu}$ of the 4μ system. The p_T, η for the signal is taken from a two-dimensional *pdf* generated using `MC@NLO` [229]. We consider proton-proton collisions at $\sqrt{s}$=10 TeV, and we model the parton distribution functions (PDFs)

using CTEQ5L [230].

In this analysis we do not assume a specific signal production mechanism and cross section, instead relying on the discrimination provided by the angular distributions of the leptons in the final state. For the SM ZZ background the p_T, η and $m_{4\mu}$ are taken from a three-dimensional *pdf* generated using the PYTHIA [213] leading-order MC generator. The momenta of the four muons in the rest frame of the $ZZ^{(*)}$ system as a function of $m_{4\mu}$ are generated according to the theoretical distributions.
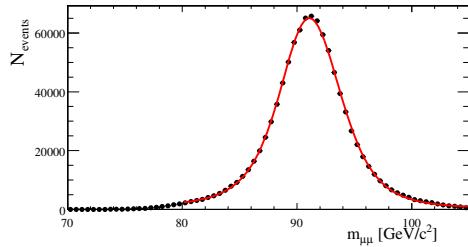

Figure A.1: Distribution of the dimuon invariant mass for a sample of signal $H \rightarrow ZZ$ events, generated using our very-fast muon simulation. The parameters of the superimposed fit are extracted from [231].

A.1.2   Detector Emulation and Event Selection

Muon reconstruction efficiency and resolution are parameterized as a function of the muon p_T and η according to [232], where the muon reconstruction efficiency is close to 100% for muons with $p_T \geq 10$ GeV/c and $|\eta| \leq 2.3$, corresponding to the event selection in our analysis. The reconstruction efficiency is applied through a hit-or-miss technique. For muon candidates accepted by the efficiency filter, the reconstructed momentum is determined by applying Gaussian smearing functions to the true p_T, η and ϕ with p_T- and η-dependent resolutions. We verified the goodness of our *very-fast* muon simulation by comparing the parameters of the fit of the Z invariant mass distribution obtained in our analysis, see figure A.1, with the corresponding ones from a published full-simulation analysis [231].

A number of detector related effects can modify the $\vec{X}$ observables' *pdfs*. The resolution of the observables used in the analysis is shown in figure A.2 and is found

to be small independent of the HLL resonance mass and quantum numbers. The systematic bias in the reconstruction of the same variables is shown in figure A.3 and is found to be negligible. This shows that the sculpting of the observables' *pdfs* is not a result of reconstruction resolution or bias. Rather, it depends on the simulated kinematics of the HLL resonance, including its mass, and on the particular model considered (0^+, 0^-, etc.). Specifically, the overall phase space acceptance, implemented in the signal selection by means of the p_T and η requirements, produces the largest effects on the observables. This is shown in figure A.4 for a resonance of mass 145 GeV/c^2 generated with no explicit angular correlations. Adding the angular correlations can enhance or reduce the overall selection efficiency depending on the
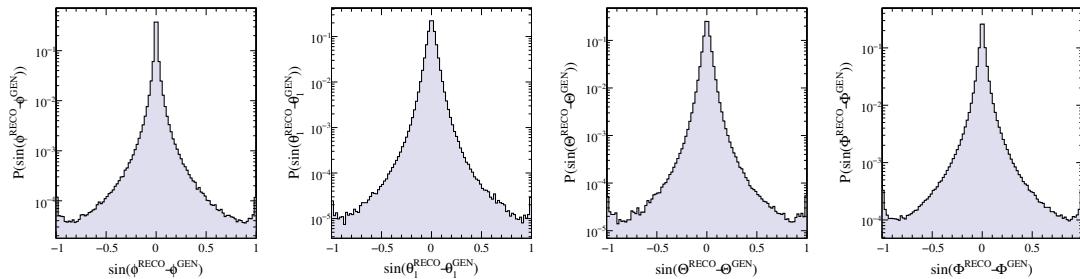

Figure A.2: Reconstruction resolution for the angular variables of $\vec{X}$ shown here for a resonance with mass 145 GeV/c^2. The $\cos\theta_2$ and $\cos\theta_1$ distributions are very similar in this case. Only events surviving the signal selection are included. All distributions are normalized to unit integral.

details of the multidimensional *pdf*. Our selection is 60% (74%) efficient for a 0^+ resonance of mass 200 GeV/c^2 (350 GeV/c^2) as shown in figure A.5. The same figure demonstrates that the efficiency has a nontrivial dependence on the nature of the spin correlations. Specifically, for a 0^- resonance of 200 GeV/c^2 (350 GeV/c^2) the efficiency is 60% (69%). With an absence of explicit spin correlations the efficiency for a 350 GeV/c^2 resonance is 71%.

We find that changes in the $\vec{X}$ distributions are strongly correlated with the kinematics of the off-shell Z, e.g., for $\cos\theta_2$ the largest inefficiencies correspond to the kinematic configurations where at least one of the muons is soft. When the correlations between the variables $\vec{\omega}$ and $\vec{\Omega}$ appear explicitly in the differential cross
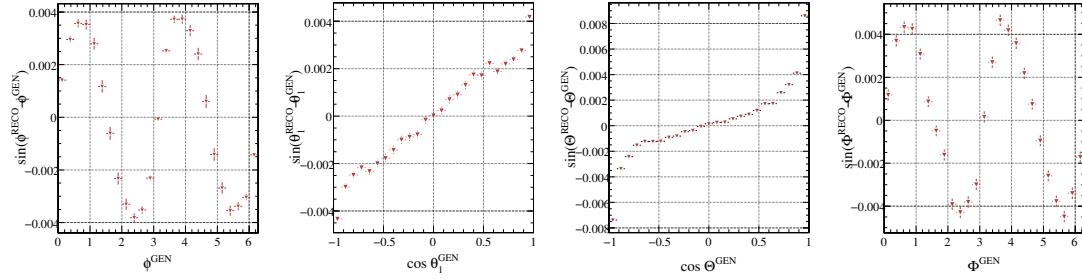
Figure A.3: Distributions showing systematic biases for a subset of the reconstructed variables $\vec{X}$ for a resonance with mass 145 GeV/c^2. Only events that survive the signal selection are included. All biases are negligible.

sections, as is the case for $J=1^\pm$, the phase space acceptance effects are amplified.

The shapes of the reconstructed $\vec{\omega}$ and $\vec{\Omega}$ distributions depend on the phase space acceptance both for electron and muon final states ($H \to ZZ \to 2e2\mu$ or $4e$). Figure A.6 shows the relevant kinematic distributions. All the results concerning model discrimination, as a function of the number of observed signal events, will be nearly identical when the additional final states are included ($2e2\mu$, $4e$), especially when the off-shell Z mass is not used as an observable. This is not necessarily the case for results concerning the discovery of a resonance in these final states with respect to the background-only hypothesis, since different backgrounds need to be considered for electron and muon final states.

A.1.3   Fit Definition and Signal Extraction

The $H \to ZZ$ signal events can be discriminated from SM backgrounds using an extended and unbinned ML fit. Since there is no resonant 4μ background in the SM, the fit can use as a discriminating variable the 4μ mass distribution. In the presence of a sizable background due to fake Z candidates (such as top decays) the 2μ mass distributions can be included in the likelihood. Since this is not a conceptually different situation, we ignore this possibility and assume for simplicity that the only relevant background is given by events with two real Z candidates. We write the
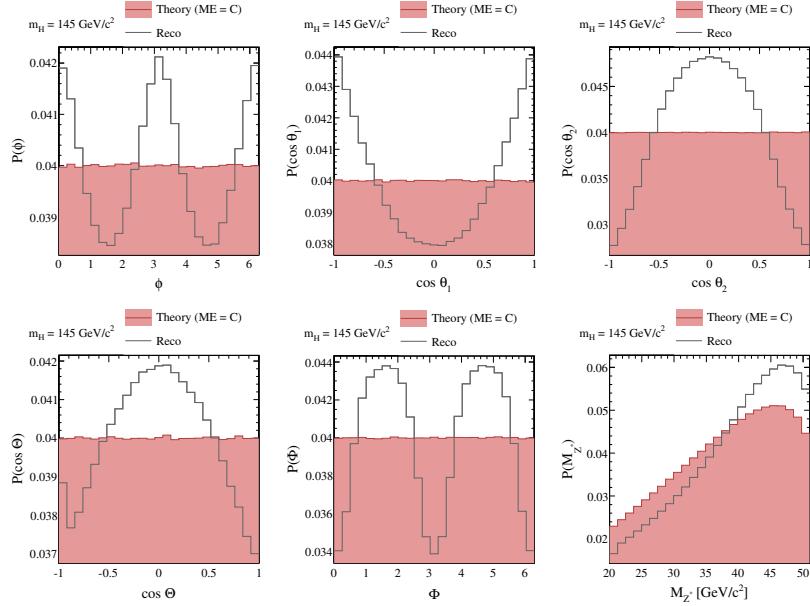
Figure A.4: The variables $\vec{X}$ used in this analysis for a 145 GeV/c^2 resonance. The off-shell M_{Z^*} is required to lie a window between 20 and 50 GeV/c^2. The shaded histograms are the 1D distributions using a constant matrix element (i.e., no angular correlations included). The overlaid histograms show the same distributions for reconstructed events passing the p_T and η signal selection after the detector parameterization. All distributions are normalized to unit integral.

likelihood function as:

$$\mathcal{L} = \frac{1}{N!} \exp\left(-\sum_j N_j\right) \tag{A.2}$$
$$\prod_{i=1}^{N}\left(N_S P_S[m_{4\mu}^i] + N_B P_B[m_{4\mu}^i]\right),$$

where N_j ($j=S,B$) represents the yield of each component, $m_{4\mu}^i$ is the 4μ candidate mass for the event i, and $P_S[m]$ ($P_B[m]$) is the signal (background) distribution for the variable m. The *pdfs* for the signal and background components are described using the template distributions from the simulation, as shown in figure A.7 for $m_H=250$ GeV/c^2. This fit configuration is appropriate for the HLL characterization.
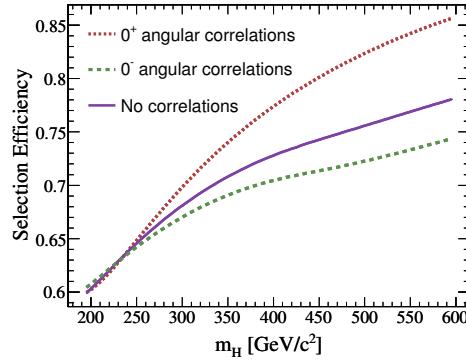
Figure A.5: The analysis efficiency for 0^+, 0^- as a function of the resonance mass. The case with no correlations is also shown for comparison.
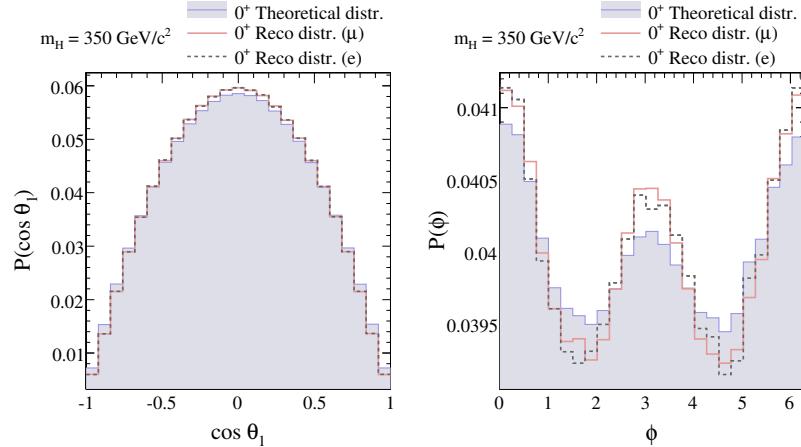

Figure A.6: Kinematic distributions for the variables $\cos\theta_1$ (Left) and ϕ (Right) for a 0^+ resonance with mass $350\ \mathrm{GeV}/c^2$. The shaded histograms show the 1D projections of the variables as described by the analytic *pdfs*. The overlaid histograms (blue, red) show the same 1D projections for reconstructed events passing the p_T and η signal selection after the detector parameterization for 4μ and $4e$ final states. All distributions are normalized to unit integral.
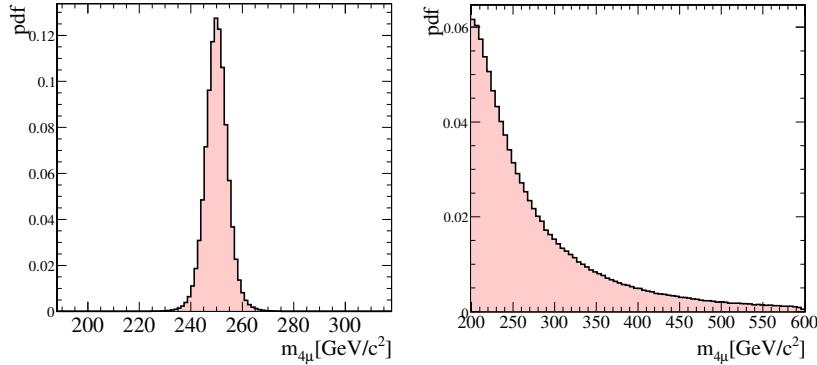
Figure A.7: Distribution of the 4μ invariant mass for a sample of signal with m_H=250 GeV/c^2 (Left), and background (Right) ZZ events.

A.1.4   Background Subtraction

In order to establish if a newly discovered resonance is indeed the Higgs boson or not, a hypothesis test is performed (see section 7.3.2). In this context, a tool to disentangle signal and background events from the selected dataset is an important prerequisite. We use the $_sWeight$ [233] technique and re-weight the selected dataset according to *how likely* each event is considered to be signal by the fit. The $_sWeight$
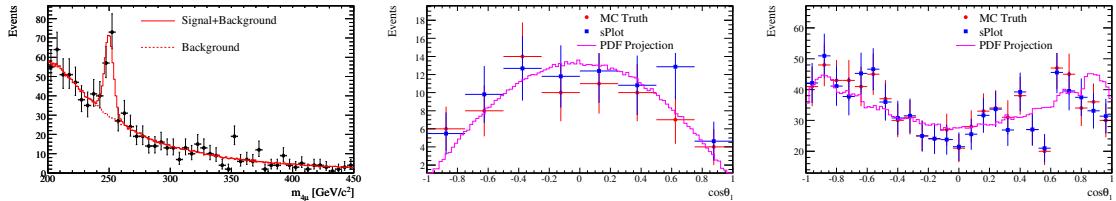

Figure A.8: The 4μ invariant mass distribution for a sample of N_S=70 $H \rightarrow ZZ$ events with m_H=250 GeV/c^2 and N_B=1000 ZZ background events. The superimposed curves represent the likelihood function returned by an ML fit, with N_S, N_B, and $m_{4\mu}$ as free parameters (Left). Comparison of the signal-only MC distribution of $\cos\theta_1$, with the background-subtracted distribution obtained with the $_sWeight$ technique (Center). Comparison of the background-only MC distribution of $\cos\theta_1$, with the signal-subtracted distribution obtained with the $_sWeight$ technique (Right).

technique is statistically optimal when the discriminating variable ($m_{4\mu}$ in our case) in the fit is uncorrelated with the subsequently used variables ($\vec{X}$ in our case). On the upper plot of figure A.8, the 4μ invariant mass distribution is shown for a sample

of N_S=70 $H \to ZZ$ events (with m_H=250 GeV/c^2) on top of N_B=1000 continuum ZZ background events, corresponding to a $\simeq 5\sigma$ deviation from the background-only hypothesis. The superimposed curves represent the likelihood function returned by an ML fit (with N_S, N_B, and $m_{4\mu}$ as free parameters). The middle plot shows the signal $_sWeighted \cos \theta_1$ distribution. Similarly, the bottom plot shows the background $_sWeighted \cos \theta_1$ distribution. The comparison of the two sets of points shows how the background (signal) subtraction allows one to recover the signal (background) distribution for the considered variable in the given sample, the deviation from the expected *pdfs* being due to statistical fluctuations already present at the MC level.

A.2   SU(2)$_\mathbf{L}$ × U(1)$_\mathbf{Y}$ Gauge-Invariant Couplings

To write Lagrangians generating the couplings of section 7.2.1 and respecting the electroweak gauge symmetry one must specify the electroweak charges of the Higgs look-alikes. Consider the example of HLLs that are "neutral," i.e., are weak singlets and have zero hypercharge. For the scalar case, in a conventional notation for isovector and isoscalar gauge fields, the lowest-dimensionality Lagrangian density is

$$
\begin{aligned}
L = \frac{1}{\Lambda} \, H \, (A_1 \, \vec{W}_{\mu\alpha} \, \vec{W}^{\mu\alpha} + A_2 \, B_{\mu\alpha} B^{\mu\alpha}) \\
+ \frac{1}{\Lambda} \, H \, i \, \epsilon^{\mu\alpha\sigma\tau} (A_3 \, \vec{W}_{\mu\alpha} \vec{W}_{\sigma\tau} + A_4 \, B_{\mu\alpha} B_{\sigma\tau}) \,,
\end{aligned}
\tag{A.3}
$$

with A_i arbitrary constants and Λ a mass parameter. This object generates, amongst others, the couplings of equation (7.9). The "true" dimensionality of the operators in equation (7.9) is that of the ones appearing in equation (A.3), that is, dimension five.

The form of equation (A.3) results in a coupling $HZ_{\mu\alpha} \, Z^{\mu\alpha} \to 2 \, p_1 \cdot p_2 \, g_{\mu\alpha} - 2 \, k_\mu k_\alpha$, establishing a relation between X and $Y + i \, Z$ in equation (7.9). We do not impose it, for it is not general even at tree level. Consider, for instance, a model with a conventionally-charged but otherwise nonstandard HLL, dubbed Φ before the spontaneous symmetry breaking. Call $V_{\mu\nu}$ any of the field tensors in equation (A.3).

The operators in this Lagrangian could be "descendants" of dimension 6 operators of the form $\Phi^\dagger \Phi\, V^2$, with $\Phi \to H + v$, see e.g. [99]. In such a case there would be a standard-like $g_{\mu\nu}$ coupling plus the one induced by the higher-dimensional operators.

For a canonical-dimension spin-2 neutral HLL, $H_{\mu\nu}$, the lowest-dimension gauge-invariant Lagrangian has couplings of dimension 5:

$$
\begin{aligned}
L = & \frac{1}{\Lambda}\, H_{\mu\nu}\, (A_1\, \vec{W}^\mu_\alpha\, \vec{W}^{\nu\alpha} + A_2\, B^\mu_\alpha B^{\nu\alpha}) \\
& + \frac{1}{\Lambda}\, H^{\nu\rho}\, i\, \epsilon_{\mu\nu\alpha\beta}(A_3\, \vec{W}^{\mu\alpha}\vec{W}^{\rho\beta} + A_4\, B^{\mu\alpha}B^{\rho\beta})\,.
\end{aligned}
\tag{A.4}
$$

The consideration of gauge-invariant constructions for HLLs with nontrivial electroweak charges would take us well beyond the scope of this work.

A.3 Phase Space for ZZ*

In the case in which one of the two Z bosons is off-shell, the dependence on its mass (M_{Z^*}, either m_1 or m_2) is an extra handle in determining the shapes of signal and backgrounds. Let $p_{\rm cms} \equiv |\vec{p}\,[Z]| = m_1\, \gamma_1\, \beta_1 = m_2\, \gamma_2\, \beta_2$ be the momentum of one or the other Z in the H center-of-mass system:

$$
\begin{aligned}
p_{\rm cms} = & \frac{1}{2\, m_H}\, \Theta[m_H - (M_Z + M_{Z^*})] \\
& \times \sqrt{m_H^2 - (M_Z - M_{Z^*})^2}\, \sqrt{m_H^2 - (M_Z + M_{Z^*})^2}\,.
\end{aligned}
\tag{A.5}
$$

Let $\mathcal{M}$ be the matrix element for the process. The expectation for the rate of events, including the dependence on M_{Z^*}, is

$$
\frac{dN}{d\cos\theta_1\, d\cos\theta_2\, d\phi\, d\cos\Theta\, d\Phi\, dM_{Z^*}} \\
\propto |\mathcal{M}|^2\, \frac{M_{Z^*}\, p_{\rm cms}}{(M_{Z^*}^2 - M_Z^2)^2 + M_{Z^*}^2\, \Gamma_Z^2}\,,
\tag{A.6}
$$

with $|\mathcal{M}|^2$ an explicit function of c_1, c_2, ϕ, Θ, Φ and M_{Z^*} for each specific case to be

discussed.

A.4 General Results for $J = 0$ Coupled to ZZ^*

In Section 7.2 we have already written the angular distributions $d\Gamma[0^+]$ and $d\Gamma[0^-]$ for the pure scalar and pseudoscalar cases, see equations (7.13) and (7.14). We also discussed the T-odd and C-odd interferences between the standard coupling, which is proportional to X in equation (7.9), and the P and Q terms of the same equation. Thus we defined $d\Gamma[0, \text{Todd}]$ and $d\Gamma[0, \text{Codd}]$ in equations (7.16) and (7.17). Similarly we discussed the complete result for the composite case with $X \neq 0$ and $Y \neq 0$, defining $d\Gamma_{XY}$ and $d\Gamma_{YY}$ in equations (7.19) and (7.20). This allows us to gather the results corresponding to the most general deviations from the SM Higgs couplings:

$$
\begin{aligned}
d\Gamma[0] \quad &= X^2 \, d\Gamma[0^+] + (P^2 + Q^2) \, d\Gamma[0^-] \\
&+ X \, P \, d\Gamma[0, \text{ Todd}] + X \, Q \, d\Gamma[0, \text{ Codd}] \\
&+ X \, Y \, d\Gamma_{XY} + (Y^2 + Z^2) \, d\Gamma_{YY} \, .
\end{aligned}
\tag{A.7}
$$

To obtain the complete spin 0 result one must add to equation (A.7) the interferences between the nonstandard terms themselves:

$$
\begin{aligned}
\Delta d\Gamma[0] \quad &= \quad X Z \, d\Gamma_{XZ} + Y P \, d\Gamma_{YP} \\
&+ \quad Y Q \, d\Gamma_{YQ} + Z P \, d\Gamma_{ZP} + Z Q \, d\Gamma_{ZQ} \, ,
\end{aligned}
\tag{A.8}
$$

where

$$
d\Gamma_{XZ} = 2 \, \eta \, m_1^3 \, m_2^3 \, m_H^2 \, \gamma_b^2 \, (c_1 + c_2) \, s \, s_1 \, s_2 \, ,
\tag{A.9}
$$

$$
d\Gamma_{YP} = d\Gamma_{ZQ} = -2 \, m_1^4 \, m_2^4 \gamma_b^3 \, s \, s_1 s_2 (c_1 c_2 + \eta^2) \, ,
\tag{A.10}
$$

$$
d\Gamma_{YQ} = -d\Gamma_{ZP} = 2 \, \eta \, m_1^4 \, m_2^4 \, \gamma_b^3 \, c \, (c_1 + c_2) s_1 s_2 \, .
\tag{A.11}
$$

Appendix B

Chapter 9 Appendix: Razor Searches for Supersymmetry

B.1 Event Samples

B.1.1 Simulated Signal and SM Background Samples

All simulated samples used in this work are processed with the full CMS detector simulation and standard `CMSSW` reconstruction chain.

The W($\to\ell\nu$) + n-jets events ($\ell = $ e, μ, τ), Z($\to\ell\ell$) + n-jets events and $\gamma + n$-jets events are produced with Monte Carlo simulation, using the `MadGraph` [234] event generator, based on a leading-order calculation of the matrix element (ME). ME calculation is performed for final states with at most four primary partons, requiring that the parton p_{T} exceeds 10 GeV/c. `PYTHIA` [213] is used for parton showering, hadronization and the underlying event description. Parton shower matching is applied to avoid double counting of emissions in overlapping phase space regions. The MLM [235] matching algorithm with k_{T}. The lepton clustering is used with matching threshold 15 GeV/c. pair invariant mass is required to be $m_{\ell\ell} > 50$ GeV/c^2 at the generator level. The `CTEQ6L1` [236] parton distribution functions are used. The $t\bar{t}$ + jets and single top (s-channel, t-channel and tW) backgrounds are generated with `MadGraph` interfaced with `PYTHIA` with the associated parton $p_{\mathrm{T}} > 20$ GeV/c and matching threshold 30 GeV/c.

Backgrounds from QCD multijet processes are studied with Monte Carlo sim-

ulation using the `ALPGEN` event generator based on a leading-order calculation of the matrix element for final states with at most six primary partons, requiring that the parton p_T exceeds 15 GeV/c. `PYTHIA` [213] is again used for parton showering, hadronization and the underlying event description, and MLM parton shower matching is again applied to avoid double counting of emissions in overlapping phase space regions. `PYTHIA` is used to generate di-photon QCD and diboson events (W^+W^-, $W^\pm Z$, ZZ). The heavy flavor QCD component is studied using a sample of simulaated events generated with `PYTHIA` and a filter that selects electron and muon enriched multijet samples. The generation includes b$\bar{\text{b}}$, c$\bar{\text{c}}$ and decays of long-lived light mesons as sources of muons and loosely isolated hadrons or jets with an increased electromagnetic fraction as a source of the electrons. The filter also requires an outgoing parton with $p_T > 20$ GeV/c.
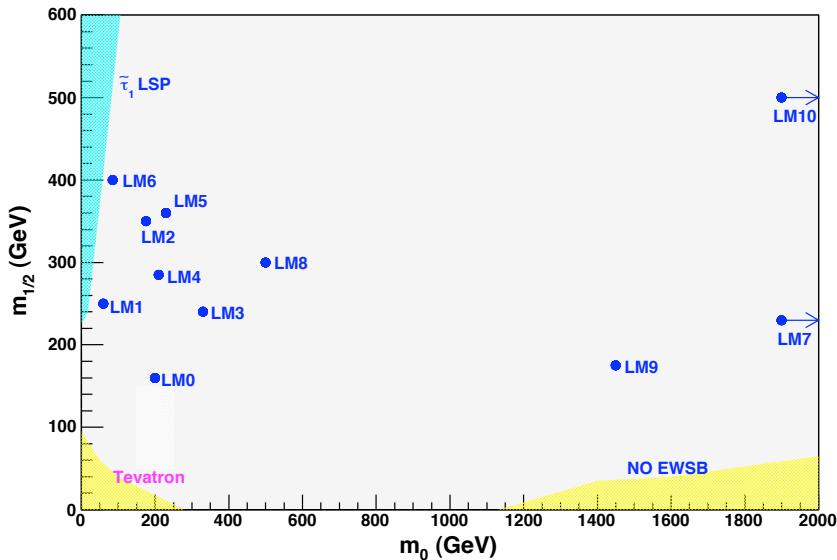

Figure B.1: The CMS benchmark mSUGRA points. Also shown is the estimated Tevatron exclusion with 10 fb^{-1} (lower left shading), and a region that is theoretically excluded by the absence of electroweak symmetry breaking (lower right shading).

To generate simulated samples for SUSY signal models the mass spectrum is first calculated with `SOFTSUSY` [215] and the decays with `SUSYHIT` [216]. `PYTHIA` is then used with the SLHA interface [217] to generate the events. The generator level cross

section and the k-factor for the NLO calculation computed using Prospino 2 [228]. Note however that In this analysis the LO cross section values are used for mSUGRA simulated signal events to extract the results; this facilitates comparison with results from previous searches.

The mSUGRA benchmarks used in CMS are illustrated in Figure B.1, along with the expected Tevatron exclusion reach with 10 fb^{-1} integrated luminosity.

The SM background simulated event samples are summarized in table B.1. The effective integrated luminosities of the samples are computed using the NLO cross section where available.

sample	cross section		integrated luminosity (pb^{-1})
ZJets-madgraph	2.4 (2.95)	nb	350
WJets-madgraph	24.2 (31.0)	nb	315
ttbarJets-madgraph	95 (162)	pb	$7.9{\cdot}10^3$
QCD-BCtoE-Pt20to30	0.108	μb	20.9
QCD-BCtoE-Pt30to80	0.138	μb	6.34
QCD-BCtoE-Pt80to170	9.46	nb	0.92
QCD-EMEnriched-Pt20to30	1.72	μb	17.6
QCD-EMEnriched-Pt30to80	3.48	μb	9.80
QCD-EMEnriched-Pt80to170	0.135	μb	37.3
PhotonJet-Pt0to15	$8.446{\cdot}10^7$	pb	$1.36{\cdot}10^{-3}$
PhotonJet-Pt15to20	$1.147{\cdot}10^5$	pb	0.944
PhotonJet-Pt20to30	$5.718{\cdot}10^4$	pb	1.05
PhotonJet-Pt30to50	$1.652{\cdot}10^4$	pb	6.66
PhotonJet-Pt50to80	$2.723{\cdot}10^3$	pb	40.6
PhotonJet-Pt80to120	$4.462{\cdot}10^2$	pb	151
PhotonJet-Pt120to170	84.43	pb	$1.46{\cdot}10^3$
PhotonJet-Pt170to300	22.55	pb	$5.44{\cdot}10^3$
PhotonJet-Pt300to500	1.545	pb	$69.4{\cdot}10^3$
PhotonJet-Pt500toInf	$9.230{\cdot}10^{-2}$	pb	$0.618{\cdot}10^6$

Table B.1: Cross sections and integrated luminosities for simulated samples used in this analysis. Cross sections are quoted at leading order, NLO cross sections are given in parentheses where available. The integrated luminosity of each sample is computed using the NLO cross section where applicable. For samples that are enriched in certain final states (QCD-BCtoE and QCD-EMEnriched), the cross section has been corrected for the enrichment factor.

	DATA	$W(\ell\nu)$+jets	$Z(\nu\nu)$+jets
N_{TOT}	5.8e+07	1.1e+06 $\pm$ 3.5e+02	2.1e+05 $\pm$ 1.5e+02
trigger	2.3e+07	3e+04 $\pm$ 57	3.5e+03 $\pm$ 19
Good PV	2.2e+07 (97%)	3e+04 $\pm$ 57 (1e+02%)	3.5e+03 $\pm$ 19 (1e+02%)
$\geq$ 2 jet	2e+07 (88%)	8.7e+03 $\pm$ 31 (29%)	1.5e+03 $\pm$ 12 (43%)
$|\beta_R| \leq 1$	1.9e+07 (84%)	7.5e+03 $\pm$ 29 (25%)	1.2e+03 $\pm$ 11 (35%)
$|\beta_R| \leq 0.99$	1.9e+07 (82%)	7.1e+03 $\pm$ 28 (24%)	1.2e+03 $\pm$ 11 (34%)
$\Delta\phi(H_1, H_2) < 2.8$ rad	3.4e+06 (15%)	4.4e+03 $\pm$ 22 (15%)	7.9e+02 $\pm$ 9 (23%)
$R > 0.45$	5.9e+03 (0.026%)	9.2e+02 $\pm$ 10 (3.1%)	2.4e+02 $\pm$ 5 (7%)
$M_R > 400$ GeV		22 $\pm$ 1.6 (0.073%)	16 $\pm$ 1.3 (0.43%)
$M_R > 500$ GeV		5.8 $\pm$ 0.8 (0.019%)	6.1 $\pm$ 0.79 (0.18%)
$R > 0.5$	2.8e+03 (0.012%)	6.3e+02 $\pm$ 8.3 (2.1%)	1.8e+02 $\pm$ 4.3 (5.1%)
$M_R > 400$ GeV		8.7 $\pm$ 1.0 (0.029%)	8.5 $\pm$ 0.9 (0.29%)
$M_R > 500$ GeV		2.3 $\pm$ 0.5 (0.0076%)	2.7 $\pm$ 0.53 (0.078%)
$R > 0.55$	1.3e+03 (0.0058%)	4.2e+02 $\pm$ 6.8 (1.4%)	1.2e+02 $\pm$ 3.6 (3.6%)
$M_R > 400$ GeV		3.1 $\pm$ 0.6 (0.01%)	5.3 $\pm$ 0.74 (0.15%)
$M_R > 500$ GeV		0.76 $\pm$ 0.29 (0.0025%)	1.7 $\pm$ 0.41 (0.048%)

Table B.2: HAD box event yields selected with trigger HLT_100U. Efficiencies evaluated w.r.t. yields after trigger.

B.1.2   Data Samples

The data sample consists of the events collected by the CMS experiment at $\sqrt{s}$=7 TeV during the Run2010A CMS data-taking period, starting in March 2010.

We use the electron primary datasets (PDs) (`EG` and `Electron`) the jet triggers (`JetMETTau`, `JetMET`, `Jet`, and `MultiJet`). The `DiJetAve` trigger (with different L1 pre-scale factors per luminosity era) was included in the `JetMETTau`, `JetMET`, and `Jet` PDs.

B.1.3   Analysis Cut-Flow for Data and Simulated Events

In Tables B.2-B.14 we present the yields for data and simulated event samples after each of the requirements in the cut-based selection of the 2010 search analysis, for the HAD, ELE and MU signal boxes. Yields are normalized to 35 pb^{-1} for simulated events and include both SM background processes and sample signal models, defined in figure B.1.

	tt	Single t s-chan	Single t t-chan
N$_{\text{TOT}}$	5.5e+03 ± 4.9	1.5e+02 ± 0.26	2.3e+03 ± 3.3
trigger	3.2e+03 ± 3.7	27 ± 0.11	3.3e+02 ± 1.3
Good PV	3.2e+03 ± 3.7 (100%)	27 ± 0.11 (100%)	3.3e+02 ± 1.3 (100%)
≥ 2 jet	3.1e+03 ± 3.7 (99%)	26 ± 0.11 (97%)	3.2e+02 ± 1.2 (95%)
$|\beta_R| \leq 1$	3.1e+03 ± 3.7 (98%)	23 ± 0.1 (86%)	2.9e+02 ± 1.2 (89%)
$|\beta_R| \leq 0.99$	3.1e+03 ± 3.6 (97%)	22 ± 0.1 (83%)	2.9e+02 ± 1.2 (86%)
$\Delta\phi(H_1, H_2) < 2.8$ rad	9.5e+02 ± 2 (30%)	11 ± 0.073 (42%)	1.6e+02 ± 0.87 (48%)
$R > 0.45$	84 ± 0.6 (2.7%)	1.4 ± 0.026 (5.2%)	13 ± 0.25 (3.8%)
$M_R > 400$ GeV	5.6 ± 0.16 (0.18%)	0.052 ± 0.005 (0.2%)	0.69 ± 0.057 (0.21%)
$M_R > 500$ GeV	1.4 ± 0.079 (0.046%)	0.011 ± 0.0023 (0.04%)	0.16 ± 0.027 (0.047%)
$R > 0.5$	53 ± 0.48 (1.7%)	0.75 ± 0.019 (2.8%)	6.3 ± 0.17 (1.9%)
$M_R > 400$ GeV	2.5 ± 0.1 (0.078%)	0.02 ± 0.0031 (0.079%)	0.16 ± 0.027 (0.0047%)
$M_R > 500$ GeV	0.5 ± 0.046 (0.016%)	0.0057 ± 0.0016 (0.021%)	0.028 ± 0.012 (0.0086%)
$R > 0.55$	34 ± 0.38 (1.1%)	0.42 ± 0.014 (1.5%)	3.1 ± 0.12 (0.95%)
$M_R > 400$ GeV	1.2 ± 0.072 (0.038%)	0.009 ± 0.0021 (0.035%)	0.057 ± 0.016 (0.017%)
$M_R > 500$ GeV	0.24 ± 0.032 (0.0077%)	0.0014 ± 0.00082 (0.0053%)	0.014 ± 0.0082 (0.0043%)

Table B.3: HAD box event yields selected with trigger HLT_100U. Efficiencies evaluated w.r.t. yields after trigger.

	LM0	LM1	LM2	LM3
N$_{\text{TOT}}$	1.4e+03 ± 3	1.7e+02 ± 0.38	21 ± 0.048	1.2e+02 ± 0.25
trigger	9e+02 ± 2.4	1.2e+02 ± 0.31	15 ± 0.04	79 ± 0.2
Good PV	9e+02 ± 2.4 (100%)	1.2e+02 ± 0.31 (100%)	14 ± 0.04 (100%)	79 ± 0.2 (100%)
≥ 2 jet	8.9e+02 ± 2.4 (99%)	1.2e+02 ± 0.31 (97%)	14 ± 0.039 (97%)	78 ± 0.2 (99%)
$|\beta_R| \leq 1$	8.5e+02 ± 2.4 (94%)	1e+02 ± 0.29 (86%)	12 ± 0.037 (84%)	74 ± 0.2 (93%)
$|\beta_R| \leq 0.99$	8.3e+02 ± 2.3 (92%)	1e+02 ± 0.29 (85%)	12 ± 0.037 (84%)	73 ± 0.2 (92%)
$\Delta\phi(H_1, H_2) < 2.8$ rad	5.2e+02 ± 1.9 (58%)	79 ± 0.26 (66%)	9.6 ± 0.033 (66%)	47 ± 0.16 (59%)
$R > 0.45$	1.5e+02 ± 1 (17%)	35 ± 0.17 (29%)	4.4 ± 0.022 (31%)	15 ± 0.088 (19%)
$M_R > 400$ GeV	98 ± 0.8 (11%)	31 ± 0.16 (26%)	4.2 ± 0.022 (29%)	13 ± 0.084 (17%)
$M_R > 500$ GeV	57 ± 0.61 (6.3%)	26 ± 0.15 (22%)	4 ± 0.021 (27%)	12 ± 0.078 (15%)
$R > 0.5$	1.1e+02 ± 0.84 (12%)	26 ± 0.15 (22%)	3.3 ± 0.019 (23%)	10 ± 0.074 (13%)
$M_R > 400$ GeV	63 ± 0.64 (7%)	23 ± 0.14 (19%)	3.1 ± 0.019 (21%)	9.2 ± 0.07 (12%)
$M_R > 500$ GeV	33 ± 0.47 (3.7%)	18 ± 0.12 (15%)	2.9 ± 0.018 (20%)	7.9 ± 0.064 (10%)
$R > 0.55$	74 ± 0.7 (8.2%)	19 ± 0.12 (16%)	2.4 ± 0.016 (17%)	7.2 ± 0.062 (9.1%)
$M_R > 400$ GeV	40 ± 0.51 (4.5%)	16 ± 0.12 (13%)	2.2 ± 0.016 (15%)	6.3 ± 0.057 (7.9%)
$M_R > 500$ GeV	19 ± 0.35 (2.1%)	13 ± 0.1 (10%)	2.1 ± 0.015 (14%)	5.2 ± 0.052 (6.6%)

Table B.4: HAD box event yields selected with trigger HLT_100U. Efficiencies evaluated w.r.t. yields after trigger.

	LM4	LM5	LM6	LM7
N_{TOT}	66 ± 0.13	17 ± 0.034	11 ± 0.021	42 ± 0.081
trigger	47 ± 0.11	12 ± 0.029	5.6 ± 0.015	11 ± 0.041
Good PV	47 ± 0.11 (100%)	12 ± 0.029 (100%)	5.6 ± 0.015 (100%)	11 ± 0.041 (100%)
≥ 2 jet	46 ± 0.11 (99%)	12 ± 0.029 (98%)	5.4 ± 0.015 (97%)	11 ± 0.041 (98%)
$|\beta_R| \leq 1$	43 ± 0.1 (91%)	11 ± 0.028 (90%)	4.6 ± 0.014 (83%)	11 ± 0.04 (94%)
$|\beta_R| \leq 0.99$	42 ± 0.1 (90%)	11 ± 0.027 (89%)	4.6 ± 0.013 (82%)	10 ± 0.04 (93%)
$\Delta\phi(H_1, H_2) < 2.8$ rad	29 ± 0.087 (62%)	7.5 ± 0.023 (64%)	3.7 ± 0.012 (66%)	4.9 ± 0.028 (44%)
$R > 0.45$	11 ± 0.053 (23%)	2.9 ± 0.014 (25%)	1.8 ± 0.0085 (33%)	0.89 ± 0.012 (8%)
$M_R > 400$ GeV	10 ± 0.051 (21%)	2.6 ± 0.014 (22%)	1.7 ± 0.0083 (31%)	0.39 ± 0.0078 (3.5%)
$M_R > 500$ GeV	9.2 ± 0.049 (20%)	2.5 ± 0.013 (21%)	1.7 ± 0.0081 (30%)	0.31 ± 0.0069 (2.7%)
$R > 0.5$	7.7 ± 0.045 (16%)	2.1 ± 0.012 (18%)	1.4 ± 0.0074 (25%)	0.57 ± 0.0094 (5.1%)
$M_R > 400$ GeV	7.1 ± 0.043 (15%)	1.9 ± 0.012 (16%)	1.3 ± 0.007 (23%)	0.22 ± 0.0059 (2%)
$M_R > 500$ GeV	6.5 ± 0.041 (14%)	1.8 ± 0.011 (15%)	1.3 ± 0.007 (23%)	0.17 ± 0.0051 (1.5%)
$R > 0.55$	5.4 ± 0.037 (12%)	1.5 ± 0.01 (13%)	1 ± 0.0063 (18%)	0.37 ± 0.0076 (3.4%)
$M_R > 400$ GeV	5.0 ± 0.036 (11%)	1.3 ± 0.0098 (11%)	0.96 ± 0.0062 (17%)	0.13 ± 0.0045 (1.2%)
$M_R > 500$ GeV	4.4 ± 0.034 (9.5%)	1.3 ± 0.0095 (11%)	0.92 ± 0.006 (17%)	0.094 ± 0.0038 (0.84%)

Table B.5: HAD box event yields selected with trigger HLT_100U. Efficiencies evaluated w.r.t. yields after trigger.

	LM8	LM9	LM10	LM11
N_{TOT}	26 ± 0.055	2.5e+02 ± 0.53	1.7 ± 0.0037	29 ± 0.056
trigger	14 ± 0.04	49 ± 0.23	0.81 ± 0.0026	19 ± 0.046
Good PV	14 ± 0.04 (100%)	48 ± 0.23 (100%)	0.81 ± 0.0026 (100%)	19 ± 0.046 (100%)
≥ 2 jet	14 ± 0.04 (99%)	47 ± 0.23 (97%)	0.8 ± 0.0026 (98%)	19 ± 0.045 (97%)
$|\beta_R| \leq 1$	13 ± 0.039 (95%)	45 ± 0.23 (92%)	0.73 ± 0.0024 (90%)	17 ± 0.043 (87%)
$|\beta_R| \leq 0.99$	13 ± 0.039 (93%)	44 ± 0.22 (91%)	0.72 ± 0.0024 (89%)	17 ± 0.042 (86%)
$\Delta\phi(H_1, H_2) < 2.8$ rad	7.9 ± 0.03 (57%)	23 ± 0.16 (47%)	0.53 ± 0.0021 (66%)	12 ± 0.037 (65%)
$R > 0.45$	2.2 ± 0.016 (16%)	4.3 ± 0.07 (8.9%)	0.21 ± 0.0013 (26%)	5.1 ± 0.024 (27%)
$M_R > 500$ GeV	1.7 ± 0.014 (12%)	1.5 ± 0.041 (3.1%)	0.018 ± 0.00038 (2.2%)	4.5 ± 0.022 (23%)
$R > 0.5$	1.5 ± 0.013 (11%)	2.8 ± 0.056 (5.7%)	0.16 ± 0.0011 (20%)	3.7 ± 0.02 (20%)
$M_R > 500$ GeV	1.1 ± 0.011 (7.8%)	0.75 ± 0.029 (1.5%)	0.0081 ± 0.00026 (1%)	3.2 ± 0.019 (17%)
$R > 0.55$	0.98 ± 0.011 (7.1%)	1.7 ± 0.044 (3.6%)	0.11 ± 0.00097 (14%)	2.7 ± 0.017 (14%)
$M_R > 500$ GeV	0.7 ± 0.0091 (5.1%)	0.36 ± 0.02 (0.73%)	0.0039 ± 0.00018 (0.48%)	2.3 ± 0.016 (12%)

Table B.6: HAD box event yields selected with trigger HLT_100U. Efficiencies evaluated w.r.t. yields after trigger.

	DATA	$W(\ell\nu)$+jets	$t\bar{t}$
N_{TOT}	7.6e+07	1.1e+06 ± 3.5e+02	5.5e+03 ± 4.9
trigger	3.9e+05	1.6e+05 ± 1.3e+02	6.1e+02 ± 1.6
Good PV	3.9e+05 (100%)	1.6e+05 ± 1.3e+02 (100%)	6.1e+02 ± 1.6 (100%)
≥ 2 jet	1.3e+05 (34%)	3e+04 ± 57 (18%)	6.1e+02 ± 1.6 (99%)
$|\beta_R| \leq 1$	1.1e+05 (29%)	2.5e+04 ± 52 (15%)	6e+02 ± 1.6 (98%)
$|\beta_R| \leq 0.99$	1.1e+05 (28%)	2.4e+04 ± 52 (15%)	5.9e+02 ± 1.6 (96%)
$\Delta\phi(H_1, H_2) < 2.8$ rad	4.5e+04 (12%)	1.7e+04 ± 43 (11%)	3.1e+02 ± 1.2 (50%)
$R > 0.45$	5.9e+03 (1.5%)	6.6e+03 ± 27 (4.1%)	49 ± 0.46 (8%)
$M_R > 500$ GeV		1.1 ± 0.35 (0.00068%)	0.29 ± 0.035 (0.047%)
$R > 0.5$	4.4e+03 (1.1%)	5e+03 ± 23 (3.1%)	31 ± 0.37 (5.1%)
$M_R > 500$ GeV		0.76 ± 0.29 (0.00048%)	0.12 ± 0.022 (0.019%)
$R > 0.55$	3.2e+03 (0.82%)	3.6e+03 ± 20 (2.3%)	20 ± 0.29 (3.3%)
$M_R > 500$ GeV		0.44 ± 0.22 (0.00027%)	0.03 ± 0.011 (0.0049%)

Table B.7: ELE box event yields selected with electron triggers. Efficiencies evaluated w.r.t. yields after trigger.

	LM0	LM1	LM2	LM3
N_{TOT}	1.4e+03 ± 3	1.7e+02 ± 0.38	21 ± 0.048	1.2e+02 ± 0.25
trigger	1e+02 ± 0.82	10 ± 0.093	1.2 ± 0.012	9.1 ± 0.069
Good PV	1e+02 ± 0.82 (100%)	10 ± 0.093 (100%)	1.2 ± 0.012 (100%)	9.1 ± 0.069 (100%)
≥ 2 jet	99 ± 0.81 (97%)	9.3 ± 0.088 (89%)	1.1 ± 0.011 (87%)	8.8 ± 0.068 (96%)
$|\beta_R| \leq 1$	94 ± 0.78 (92%)	8.2 ± 0.082 (78%)	0.91 ± 0.01 (76%)	8.2 ± 0.065 (90%)
$|\beta_R| \leq 0.99$	92 ± 0.78 (90%)	8 ± 0.081 (77%)	0.9 ± 0.01 (74%)	8 ± 0.065 (88%)
$\Delta\phi(H_1, H_2) < 2.8$ rad	59 ± 0.62 (57%)	6.2 ± 0.072 (60%)	0.69 ± 0.0088 (58%)	5.3 ± 0.053 (58%)
$R > 0.45$	15 ± 0.32 (15%)	2.6 ± 0.046 (25%)	0.3 ± 0.0058 (25%)	1.6 ± 0.029 (18%)
$M_R > 500$ GeV	3.4 ± 0.15 (3.3%)	1.5 ± 0.035 (14%)	0.2 ± 0.0047 (17%)	0.89 ± 0.022 (9.8%)
$R > 0.5$	11 ± 0.26 (10%)	1.8 ± 0.039 (18%)	0.21 ± 0.0049 (18%)	1.1 ± 0.024 (13%)
$M_R > 500$ GeV	1.7 ± 0.11 (1.7%)	0.96 ± 0.028 (9.2%)	0.14 ± 0.0039 (11%)	0.56 ± 0.017 (6.1%)
$R > 0.55$	7.3 ± 0.22 (7.2%)	1.3 ± 0.033 (13%)	0.15 ± 0.0041 (13%)	0.79 ± 0.02 (8.7%)
$M_R > 500$ GeV	0.89 ± 0.076 (0.87%)	0.63 ± 0.023 (6%)	0.093 ± 0.0032 (7.7%)	0.35 ± 0.014 (3.8%)

Table B.8: ELE box event yields selected with electron triggers. Efficiencies evaluated w.r.t. yields after trigger.

	LM4	LM5	LM6	LM7
N_{TOT}	66 ± 0.13	17 ± 0.034	11 ± 0.021	42 ± 0.081
trigger	4.6 ± 0.034	1.2 ± 0.0093	1.4 ± 0.0074	3.7 ± 0.024
Good PV	4.6 ± 0.034 (100%)	1.2 ± 0.0093 (100%)	1.4 ± 0.0074 (100%)	3.7 ± 0.024 (100%)
≥ 2 jet	4.4 ± 0.034 (96%)	1.2 ± 0.0091 (97%)	1.1 ± 0.0067 (82%)	3.3 ± 0.022 (88%)
$|\beta_R| \leq 1$	4.1 ± 0.033 (89%)	1.1 ± 0.0087 (88%)	0.98 ± 0.0062 (70%)	2.9 ± 0.021 (78%)
$|\beta_R| \leq 0.99$	4.1 ± 0.032 (88%)	1 ± 0.0086 (87%)	0.96 ± 0.0062 (69%)	2.8 ± 0.021 (76%)
$\Delta\phi(H_1, H_2) < 2.8$ rad	2.8 ± 0.027 (60%)	0.75 ± 0.0073 (62%)	0.76 ± 0.0055 (55%)	1.9 ± 0.017 (51%)
$R > 0.45$	0.91 ± 0.015 (20%)	0.28 ± 0.0045 (23%)	0.34 ± 0.0037 (24%)	0.63 ± 0.0098 (17%)
$M_R > 500$ GeV	0.54 ± 0.012 (12%)	0.15 ± 0.0033 (13%)	0.22 ± 0.0029 (16%)	0.037 ± 0.0024 (0.98%)
$R > 0.5$	0.63 ± 0.013 (14%)	0.2 ± 0.0038 (17%)	0.25 ± 0.0032 (18%)	0.45 ± 0.0083 (12%)
$M_R > 500$ GeV	0.35 ± 0.0094 (7.5%)	0.1 ± 0.0027 (8.5%)	0.16 ± 0.0025 (11%)	0.018 ± 0.0016 (0.47%)
$R > 0.55$	0.43 ± 0.011 (9.4%)	0.14 ± 0.0031 (12%)	0.18 ± 0.0027 (13%)	0.32 ± 0.007 (8.5%)
$M_R > 500$ GeV	0.22 ± 0.0075 (4.7%)	0.066 ± 0.0022 (5.5%)	0.11 ± 0.0021 (7.7%)	0.0094 ± 0.0012 (0.25%)

Table B.9: ELE box event yields selected with electron triggers. Efficiencies evaluated w.r.t. yields after trigger.

	LM8	LM9	LM10	LM11
N_{TOT}	26 ± 0.055	2.5e+02 ± 0.53	1.7 ± 0.0037	29 ± 0.056
trigger	2.7 ± 0.018	12 ± 0.12	0.16 ± 0.0012	2.4 ± 0.016
Good PV	2.7 ± 0.018 (100%)	12 ± 0.12 (100%)	0.16 ± 0.0012 (100%)	2.4 ± 0.016 (100%)
≥ 2 jet	2.5 ± 0.017 (96%)	9.3 ± 0.1 (76%)	0.15 ± 0.0011 (95%)	2.3 ± 0.016 (94%)
$|\beta_R| \leq 1$	2.4 ± 0.017 (90%)	8 ± 0.096 (66%)	0.13 ± 0.001 (83%)	2 ± 0.015 (84%)
$|\beta_R| \leq 0.99$	2.3 ± 0.017 (88%)	7.9 ± 0.095 (65%)	0.13 ± 0.001 (81%)	2 ± 0.015 (83%)
$\Delta\phi(H_1, H_2) < 2.8$ rad	1.5 ± 0.013 (58%)	5.1 ± 0.076 (42%)	0.1 ± 0.00092 (63%)	1.5 ± 0.013 (62%)
$R > 0.45$	0.46 ± 0.0073 (17%)	1.5 ± 0.041 (12%)	0.048 ± 0.00063 (30%)	0.57 ± 0.0078 (24%)
$M_R > 500$ GeV	0.22 ± 0.0051 (8.4%)	0.092 ± 0.01 (0.75%)	0.002 ± 0.00013 (1.2%)	0.36 ± 0.0062 (15%)
$R > 0.5$	0.32 ± 0.0061 (12%)	1.1 ± 0.035 (8.6%)	0.037 ± 0.00055 (23%)	0.4 ± 0.0066 (17%)
$M_R > 500$ GeV	0.14 ± 0.0041 (5.3%)	0.051 ± 0.0076 (0.42%)	0.00095 ± 8.8e-05 (0.59%)	0.24 ± 0.0051 (9.9%)
$R > 0.55$	0.21 ± 0.005 (8.1%)	0.73 ± 0.029 (6%)	0.028 ± 0.00048 (17%)	0.28 ± 0.0055 (12%)
$M_R > 500$ GeV	0.089 ± 0.0032 (3.3%)	0.027 ± 0.0056 (0.22%)	0.00053 ± 6.6e-05 (0.33%)	0.16 ± 0.0041 (6.5%)

Table B.10: ELE box event yields selected with electron triggers. Efficiencies evaluated w.r.t. yields after trigger.

	DATA	$W(\ell\nu)$+jets	$t\bar{t}$
N_{TOT}	4.9e+07	1.1e+06 $\pm$ 3.5e+02	5.5e+03 $\pm$ 4.9
trigger	2.4e+05	1.6e+05 $\pm$ 1.3e+02	5.7e+02 $\pm$ 1.6
Good PV	2.4e+05 (100%)	1.6e+05 $\pm$ 1.3e+02 (100%)	5.7e+02 $\pm$ 1.6 (100%)
≥ 2 jet	6.3e+04 (26%)	3.1e+04 $\pm$ 59 (19%)	5.7e+02 $\pm$ 1.6 (99%)
$|\beta_R| \leq 1$	5.2e+04 (21%)	2.5e+04 $\pm$ 53 (16%)	5.6e+02 $\pm$ 1.5 (98%)
$|\beta_R| \leq 0.99$	5e+04 (21%)	2.5e+04 $\pm$ 52 (15%)	5.4e+02 $\pm$ 1.5 (96%)
$\Delta\phi(H_1, H_2) < 2.8$ rad	2.8e+04 (11%)	1.8e+04 $\pm$ 44 (11%)	2.9e+02 $\pm$ 1.1 (51%)
$R > 0.45$	6.7e+03 (2.8%)	7.1e+03 $\pm$ 28 (4.4%)	45 $\pm$ 0.44 (8%)
$M_R > 500$ GeV		1.5 $\pm$ 0.41 (0.00094%)	0.27 $\pm$ 0.034 (0.048%)
$R > 0.5$	5.1e+03 (2.1%)	5.4e+03 $\pm$ 24 (3.3%)	29 $\pm$ 0.35 (5.1%)
$M_R > 500$ GeV		0.55 $\pm$ 0.24 (0.00033%)	0.069 $\pm$ 0.017 (0.012%)
$R > 0.55$	3.7e+03 (1.5%)	4e+03 $\pm$ 21 (2.5%)	18 $\pm$ 0.28 (3.2%)
$M_R > 500$ GeV		0.22 $\pm$ 0.15 (0.00013%)	0.026 $\pm$ 0.011 (0.0045%)

Table B.11: MU box event yields selected with muon triggers. Efficiencies evaluated w.r.t. yields after trigger.

	LM0	LM1	LM2	LM3
N_{TOT}	1.4e+03 $\pm$ 3	1.7e+02 $\pm$ 0.38	21 $\pm$ 0.048	1.2e+02 $\pm$ 0.25
trigger	91 $\pm$ 0.77	8.8 $\pm$ 0.085	1 $\pm$ 0.011	7.8 $\pm$ 0.064
Good PV	91 $\pm$ 0.77 (100%)	8.8 $\pm$ 0.085 (100%)	1 $\pm$ 0.011 (100%)	7.8 $\pm$ 0.064 (100%)
≥ 2 jet	87 $\pm$ 0.76 (96%)	7.8 $\pm$ 0.08 (88%)	0.89 $\pm$ 0.0099 (86%)	7.5 $\pm$ 0.063 (96%)
$|\beta_R| \leq 1$	83 $\pm$ 0.74 (92%)	6.8 $\pm$ 0.075 (77%)	0.77 $\pm$ 0.0092 (74%)	7 $\pm$ 0.06 (89%)
$|\beta_R| \leq 0.99$	81 $\pm$ 0.73 (90%)	6.7 $\pm$ 0.074 (76%)	0.75 $\pm$ 0.0091 (73%)	6.8 $\pm$ 0.06 (88%)
$\Delta\phi(H_1, H_2) < 2.8$ rad	52 $\pm$ 0.58 (57%)	5.3 $\pm$ 0.066 (60%)	0.59 $\pm$ 0.0081 (58%)	4.6 $\pm$ 0.049 (58%)
$R > 0.45$	14 $\pm$ 0.3 (15%)	2.2 $\pm$ 0.043 (26%)	0.25 $\pm$ 0.0053 (25%)	1.4 $\pm$ 0.027 (18%)
$M_R > 500$ GeV	3.2 $\pm$ 0.14 (3.5%)	1.3 $\pm$ 0.033 (15%)	0.17 $\pm$ 0.0044 (17%)	0.72 $\pm$ 0.019 (9.3%)
$R > 0.5$	9.3 $\pm$ 0.25 (10%)	1.6 $\pm$ 0.037 (18%)	0.18 $\pm$ 0.0045 (18%)	0.98 $\pm$ 0.023 (12%)
$M_R > 500$ GeV	1.6 $\pm$ 0.1 (1.8%)	0.89 $\pm$ 0.027 (10%)	0.12 $\pm$ 0.0036 (12%)	0.46 $\pm$ 0.015 (5.8%)
$R > 0.55$	6.3 $\pm$ 0.2 (7%)	1.2 $\pm$ 0.031 (13%)	0.13 $\pm$ 0.0037 (12%)	0.67 $\pm$ 0.019 (8.5%)
$M_R > 500$ GeV	0.75 $\pm$ 0.07 (0.83%)	0.59 $\pm$ 0.022 (6.7%)	0.077 $\pm$ 0.0029 (7.5%)	0.28 $\pm$ 0.012 (3.6%)

Table B.12: MU box event yields selected with muon triggers. Efficiencies evaluated w.r.t. yields after trigger.

	LM4	LM5	LM6	LM7
N_{TOT}	66 $\pm$ 0.13	17 $\pm$ 0.034	11 $\pm$ 0.021	42 $\pm$ 0.081
trigger	4 $\pm$ 0.032	1.1 $\pm$ 0.0087	1.2 $\pm$ 0.007	3.6 $\pm$ 0.023
Good PV	4 $\pm$ 0.032 (100%)	1.1 $\pm$ 0.0087 (100%)	1.2 $\pm$ 0.007 (100%)	3.6 $\pm$ 0.023 (100%)
≥ 2 jet	3.8 $\pm$ 0.031 (95%)	1 $\pm$ 0.0085 (96%)	0.98 $\pm$ 0.0062 (80%)	3.1 $\pm$ 0.022 (87%)
$|\beta_R| \leq 1$	3.5 $\pm$ 0.03 (88%)	0.92 $\pm$ 0.0081 (88%)	0.83 $\pm$ 0.0057 (67%)	2.7 $\pm$ 0.02 (76%)
$|\beta_R| \leq 0.99$	3.5 $\pm$ 0.03 (87%)	0.91 $\pm$ 0.008 (86%)	0.82 $\pm$ 0.0057 (66%)	2.6 $\pm$ 0.02 (74%)
$\Delta\phi(H_1, H_2) < 2.8$ rad	2.4 $\pm$ 0.025 (60%)	0.66 $\pm$ 0.0069 (63%)	0.65 $\pm$ 0.0051 (53%)	1.8 $\pm$ 0.017 (51%)
$R > 0.45$	0.83 $\pm$ 0.015 (21%)	0.25 $\pm$ 0.0042 (23%)	0.3 $\pm$ 0.0034 (24%)	0.63 $\pm$ 0.0098 (18%)
$M_R > 500$ GeV	0.47 $\pm$ 0.011 (12%)	0.13 $\pm$ 0.003 (12%)	0.19 $\pm$ 0.0027 (15%)	0.03 $\pm$ 0.0021 (0.83%)
$R > 0.5$	0.59 $\pm$ 0.012 (15%)	0.17 $\pm$ 0.0035 (17%)	0.22 $\pm$ 0.0029 (18%)	0.46 $\pm$ 0.0084 (13%)
$M_R > 500$ GeV	0.3 $\pm$ 0.0088 (7.6%)	0.082 $\pm$ 0.0024 (7.7%)	0.13 $\pm$ 0.0023 (11%)	0.015 $\pm$ 0.0015 (0.41%)
$R > 0.55$	0.4 $\pm$ 0.01 (10%)	0.12 $\pm$ 0.003 (12%)	0.16 $\pm$ 0.0025 (13%)	0.32 $\pm$ 0.007 (9.1%)
$M_R > 500$ GeV	0.19 $\pm$ 0.007 (4.8%)	0.055 $\pm$ 0.002 (5.2%)	0.09 $\pm$ 0.0019 (7.3%)	0.0083 $\pm$ 0.0011 (0.23%)

Table B.13: MU box event yields selected with muon triggers. Efficiencies evaluated w.r.t. yields after trigger.

	LM8	LM9	LM10	LM11
N_{TOT}	26 ± 0.055	2.5e+02 ± 0.53	1.7 ± 0.0037	29 ± 0.056
trigger	2.1 ± 0.016	11 ± 0.11	0.15 ± 0.0011	2.1 ± 0.015
Good PV	2.1 ± 0.016 (100%)	11 ± 0.11 (100%)	0.15 ± 0.0011 (100%)	2.1 ± 0.015 (100%)
≥ 2 jet	2 ± 0.015 (95%)	8.3 ± 0.097 (73%)	0.14 ± 0.0011 (95%)	2 ± 0.015 (94%)
$\lvert\beta_R\rvert \le 1$	1.9 ± 0.015 (88%)	6.9 ± 0.089 (61%)	0.13 ± 0.001 (83%)	1.8 ± 0.014 (84%)
$\lvert\beta_R\rvert \le 0.99$	1.9 ± 0.015 (87%)	6.7 ± 0.087 (59%)	0.12 ± 0.001 (81%)	1.8 ± 0.014 (82%)
$\Delta\phi(H_1, H_2) < 2.8$ rad	1.2 ± 0.012 (58%)	4.4 ± 0.071 (39%)	0.098 ± 0.0009 (64%)	1.3 ± 0.012 (62%)
$R > 0.45$	0.41 ± 0.0069 (19%)	1.3 ± 0.039 (12%)	0.045 ± 0.00061 (30%)	0.51 ± 0.0074 (24%)
$M_R > 500$ GeV	0.18 ± 0.0046 (8.3%)	0.07 ± 0.0089 (0.62%)	0.0017 ± 0.00012 (1.1%)	0.31 ± 0.0058 (14%)
$R > 0.5$	0.28 ± 0.0057 (13%)	0.94 ± 0.033 (8.3%)	0.035 ± 0.00053 (23%)	0.35 ± 0.0062 (17%)
$M_R > 500$ GeV	0.11 ± 0.0035 (5%)	0.041 ± 0.0068 (0.36%)	0.00086 ± 8.4e-05 (0.56%)	0.2 ± 0.0047 (9.4%)
$R > 0.55$	0.2 ± 0.0048 (9.2%)	0.63 ± 0.027 (5.5%)	0.026 ± 0.00046 (17%)	0.25 ± 0.0052 (12%)
$M_R > 500$ GeV	0.07 ± 0.0029 (3.3%)	0.022 ± 0.0049 (0.19%)	0.00048 ± 6.2e-05 (0.31%)	0.13 ± 0.0038 (6.3%)

Table B.14: MU box event yields selected with muon triggers. Efficiencies evaluated w.r.t. yields after trigger.

B.2 HT Trigger Turn-On Efficiency Convolution

The shape of the M_R distribution for different background processes in the signal HAD box is dictated not only by the intrinsic exponentially falling M_R scaling behavior but is also sculpted by the H_T triggers used to select these events. The reason is that an event's efficiency for passing these H_T triggers is correlated with its reconstructed M_R value. In order to recover the exponential M_R scaling behavior in the HAD box, even in ranges of M_R where these H_T triggers are not fully efficient, we must understand and model these trigger efficiencies for each background process.

In order to study how to model these efficiency turn-on curves, we implement an emulation of the `HLT` level H_T, which is the used to make the trigger decisions, using uncorrected calorimetric jets and defining H_T as the scalar sum of the jet p_Ts for jets with $p_T > 20\,\mathrm{GeV}/c$ (matching the HLT requirements). We check the accuracy of this approximation by examining events selected using the prescaled `HLT_DiJetAve15U` trigger and comparing the M_R trigger turn-on curves. Results of this exercise are shown in figures B.2 and B.3 for H_T thresholds of 100 and 140 GeV, respectively. We observe excellent agreement between the actual and emulated trigger turn-ons. This is expected, since the M_R turn-on is more sensitive to the differences between the quantities M_R and H_T, rather than the differences between H_T reconstructed at the HLT and that calculated using offline reconstructed calorimetric jets.

Using the emulated H_T triggers, we evaluate M_R trigger turn-on curves for different background processes to the HAD box at different values of the R cut, using simulated events. We observe that for sufficiently high R cuts ($\sim R > 0.3$) the shape of these trigger turn-on curves is well described by an error function:

$$f^{TRIG}(M_R \mid \mu, \lambda) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\sqrt{2}\lambda(M_R-\mu)} e^{-t^2/2} dt = \frac{1}{2}[1 + \mathrm{erf}(\lambda(M_R - \mu))] \ . \quad \text{(B.1)}$$

We fit these turn-on curves derived from simulated events, for different values of the R cut and different H_T trigger thresholds, in a binned maximum likelihood fit
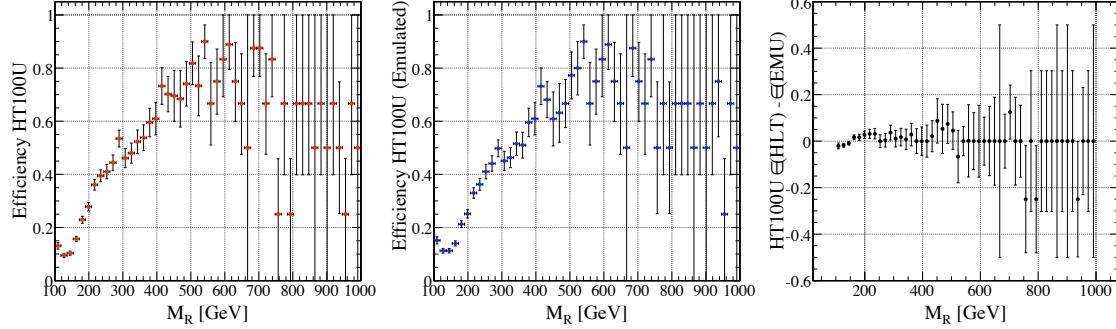
Figure B.2: Events selected using the HLT_DiJetAve15U and satisfying the HAD Box requirements without an R cut. (Left) M_R turn-on curves for the HLT_HT100U HLT trigger bit. (center) M_R turn-on curves for the emulated HLT_HT100U trigger decision. (Right) Comparison of the HLT trigger bit and emulated M_R turn-on curves.
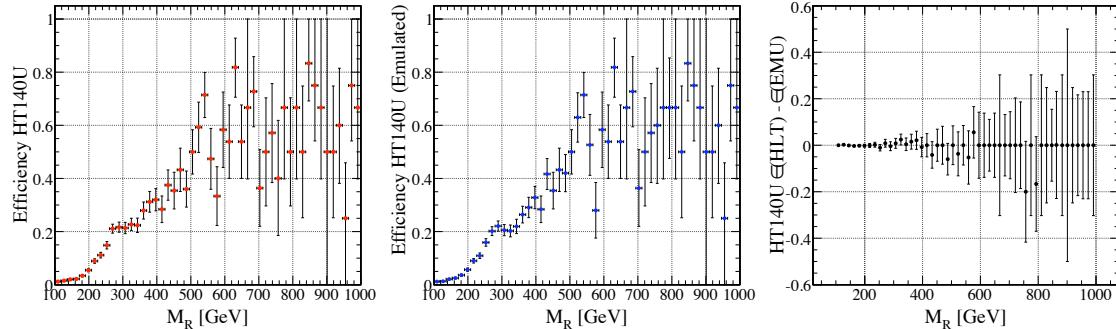
Figure B.3: Events selected using the HLT_DiJetAve15U and satisfying the HAD Box requirements without an R cut. (Left) M_R turn-on curves for the HLT_HT140U HLT trigger bit. (center) M_R turn-on curves for the emulated HLT_HT140U trigger decision. (Right) Comparison of the HLT trigger bit and emulated M_R turn-on curves.

where the likelihood is constructed as the product of bin-by-bin binomial probabilities. Fits of these curves are shown in figures B.7-B.10. We also observe that these trigger turn-on curves are similar between different background processes, but not necessarily identical (due to differences in kinematic correlations between M_R and H_T). We measure, in simulated events, the parameters μ and λ for each trigger, each process and each physics object box, independently. These parameter values derived from simulated events are used to build the initial trigger turn-on curves for each process, with additional parameters introduced to account for differences between simulation

and data for these curves, as described describe below. These parameter values are summarized in Tables B.15-B.18.

The values of the parameters μ and λ, for a given process, have a clear dependence on the H_T threshold being considered. We denote $\mu^{ijk}(\lambda^{ijk})$ as the $\mu(\lambda)$ H_T trigger turn-on parameter for the ith H_T threshold, the jth R cut value and the kth background process. We observe that μ is linearly correlated with the H_T threshold, which is demonstrated in figure B.4 (left) by plotting the values of μ for the various different background processes and R cuts, scaling by the H_T threshold and normalizing for each process separately. Similarly, we observe that λ is linearly correlated with the inverse of the square root of the H_T threshold, illustrated in figure B.4 (right). Considering the precision with which we know these parameters from the fits to simulated data, and by fitting the distributions with a Normal function, we estimate that this H_T threshold correspondence holds within a few percent deviation. This H_T dependence for these two parameters can be understood in in considering M_R as an estimator of the H_T. The functional form which we are fitting suggests that M_R has an approximately Gaussian response relative to H_T. The linear H_T scaling of the μ parameter results from this response being approximately linear (at least in the region from H_T 100 to 150 GeV). The H_T threshold dependence of the λ parameters suggests that the resolution of this response is inversely proportional to the inverse of the H_T threshold.
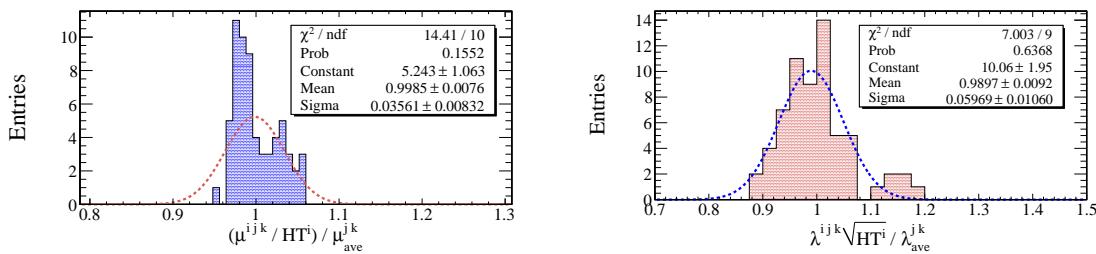
Figure B.4: (Left) H_T trigger turn-on μ parameters for different background processes in the hadronic box, normalized by the H_T threshold and for each process separately. (Right) H_T trigger turn-on λ parameters for different background processes in the hadronic box, scaled by the square-root of the H_T threshold and normalized for each process separately.

These observations dictate the strategy for unfolding the H_T trigger efficiency in the HAD box. The selected events populating the HAD box are selected using three different H_T triggers, with thresholds of 100, 140 and 150 GeV. As a result, the trigger turn-on for a given process is not one instance of the function f^{TRIG} from equation (B.1), but rather an integrated-luminosity-fraction-weighted sum of the turn-on functions for the three different triggers used. Explicitly, $f^{TRIG}(x \mid \mu, \lambda)$ is generalized to $F^{TRIG}(x \mid \vec{\mu}, \vec{\lambda})$, which is defined as

$$F^{TRIG}(x \mid \vec{\mu}, \vec{\lambda}) = \sum_{j}^{N^{TRIG}} \epsilon^j f^{TRIG}(x \mid \mu^j, \lambda^j) , \qquad (B.2)$$

where $N^{TRIG} = 3$ and ϵ^j is the fraction of the total integrated luminosity for which trigger j was the lowest threshold, un-prescaled H_T trigger in the `HLT` menu. The H_T threshold dependence of the parameters μ and λ which we observe in simulated events is assumed for the QCD multijet background to the hadronic box, such that we constrain the ratio parameters μ and λ for different H_T trigger thresholds to those described above, floating only one universal parameter for μ and λ, respectively.

For the non-QCD multijet backgrounds, we assume that the ratios of the parameters μ and λ between difference background processes to be those measured in simulated events. Two parameters are floated in the final fit in the hadronic box, corresponding to universal scaling factors μ^{EWK} and λ^{EWK} which are applied as an overall normalization of the μ and λ MC values. Allowing these parameters to vary effectively takes into account the possibility of the jet energy scale difference between corrected calorimetric jets (used to calculate M_R) and trigger level jets or the resolutions of these jets being different between simulation and data.

In Figure B.5 and B.6 we present examples of data/simulation comparisons for the `HLT_HT100U` and `HLT_HT140U` trigger turn-on efficiency curves. We observe agreement within large statistical uncertainties.
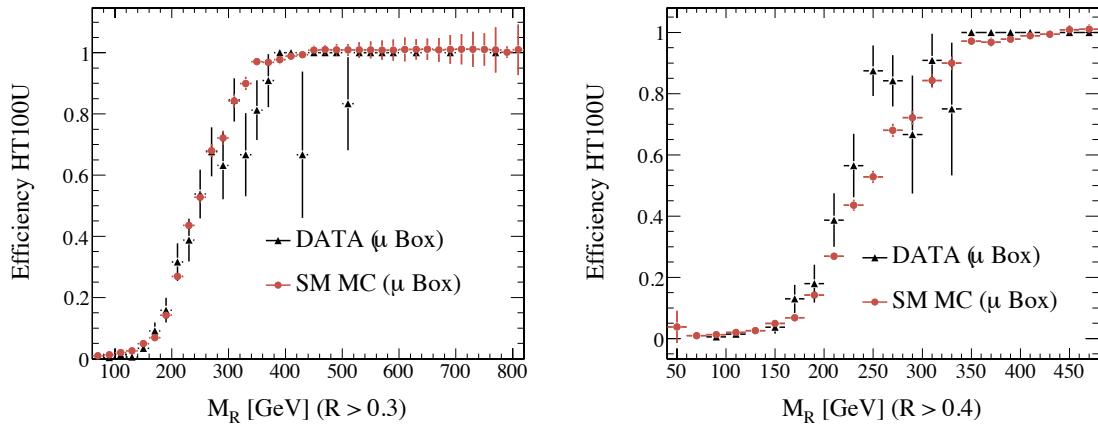
Figure B.5: M_R turn-on curves for the HLT_HT100U trigger. Events selected in data by the MU Box are compared to a cocktail of simulated SM background satisfying the MU* Box selection, with relative yields set by MU* Box expectations.
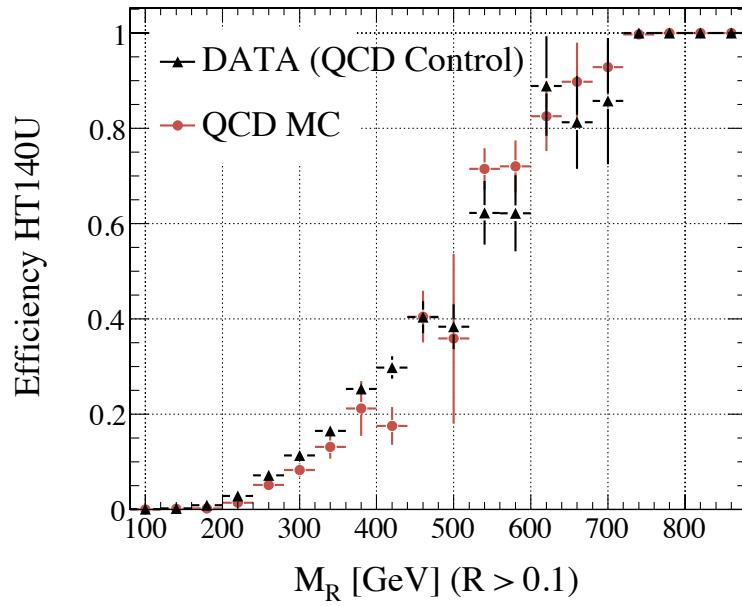
Figure B.6: M_R turn-on curves for the HLT_HT140U trigger for events selected in data by the QCD control box and simulated QCD ALPGEN events satisfying the HAD Box selection.

HT threshold / R cut	μ	λ
$W(\mu\nu)$+jets		
HT100U / $R > 0.40$	238 ± 3 GeV	(120 ± 5) x 10^{-4} GeV^{-1}
HT140U / $R > 0.40$	323 ± 8 GeV	(80 ± 5) x 10^{-4} GeV^{-1}
HT150U / $R > 0.40$	342 ± 10 GeV	(75 ± 5) x 10^{-4} GeV^{-1}
HT100U / $R > 0.45$	244 ± 2 GeV	(117 ± 4) x 10^{-4} GeV^{-1}
HT140U / $R > 0.45$	324 ± 6 GeV	(82 ± 4) x 10^{-4} GeV^{-1}
HT150U / $R > 0.45$	342 ± 7 GeV	(78 ± 4) x 10^{-4} GeV^{-1}
HT100U / $R > 0.50$	249 ± 2 GeV	(112 ± 3) x 10^{-4} GeV^{-1}
HT140U / $R > 0.50$	334 ± 5 GeV	(79 ± 3) x 10^{-4} GeV^{-1}
HT150U / $R > 0.50$	355 ± 6 GeV	(74 ± 3) x 10^{-4} GeV^{-1}

Table B.15: HT trigger turn-on parameters for simulated $W(\mu\nu)$+jets events satisfying the HAD Box requirements for different R cuts and HT thresholds.
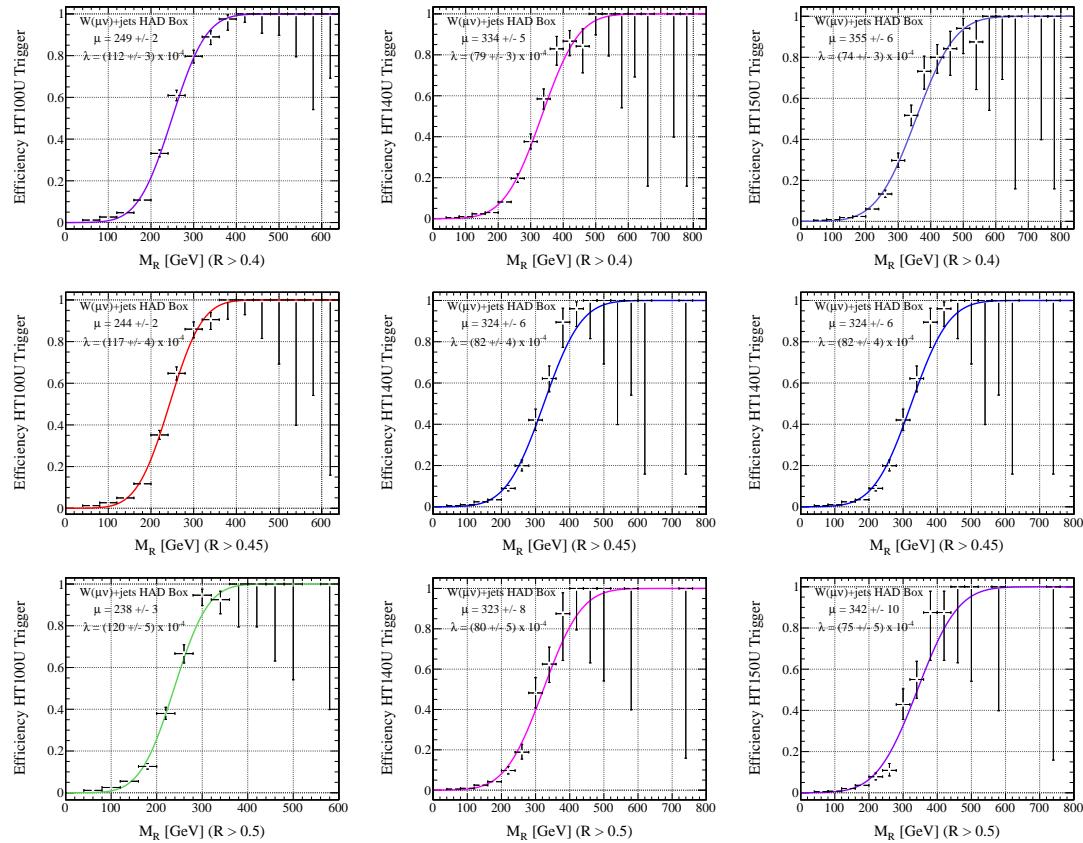
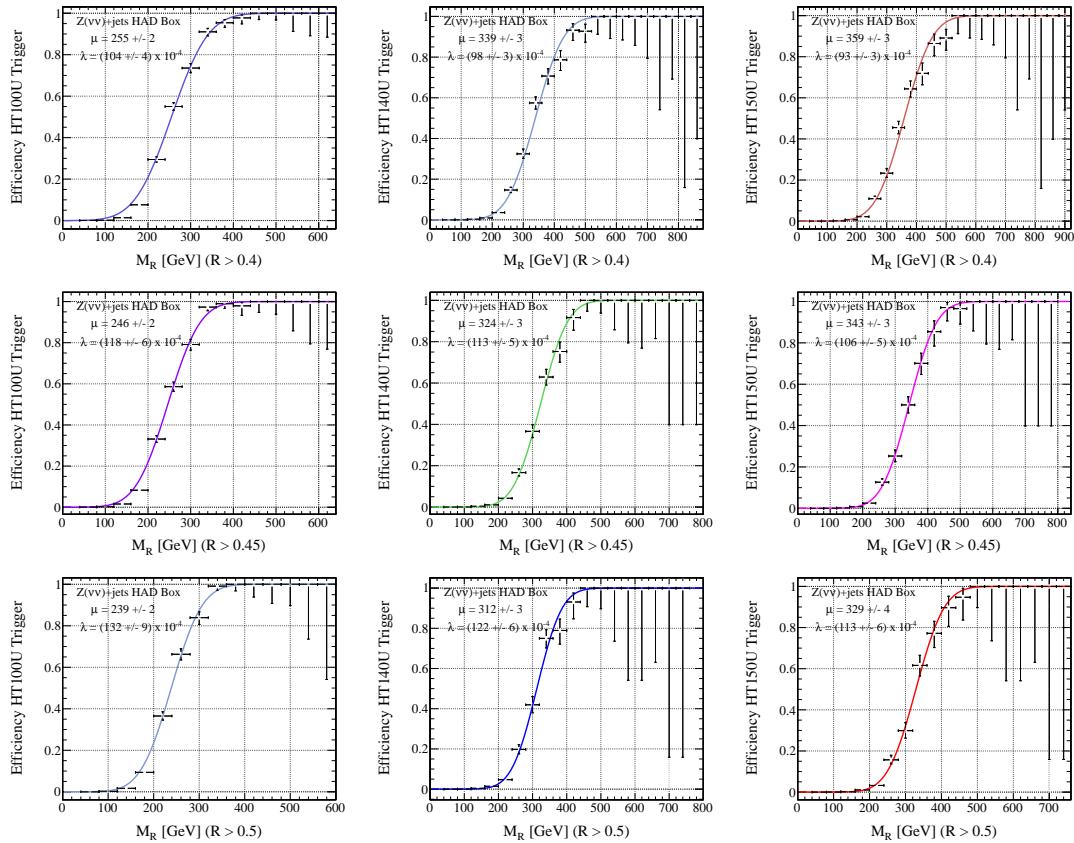Figure B.7: $W(\mu\nu)$+jets emulated HT trigger turn-ons in the HAD Box.

HT threshold / R cut	μ	λ
$t\bar{t}(\mu + X)$+jets		
HT100U / $R > 0.40$	200 ± 1 GeV	(94 ± 1) x 10^{-4} GeV^{-1}
HT140U / $R > 0.40$	284 ± 1 GeV	(83 ± 1) x 10^{-4} GeV^{-1}
HT150U / $R > 0.40$	302 ± 1 GeV	(82 ± 1) x 10^{-4} GeV^{-1}
HT100U / $R > 0.45$	198 ± 1 GeV	(101 ± 2) x 10^{-4} GeV^{-1}
HT140U / $R > 0.45$	277 ± 1 GeV	(88 ± 2) x 10^{-4} GeV^{-1}
HT150U / $R > 0.45$	295 ± 2 GeV	(86 ± 2) x 10^{-4} GeV^{-1}
HT100U / $R > 0.50$	195 ± 1 GeV	(110 ± 2) x 10^{-4} GeV^{-1}
HT140U / $R > 0.50$	269 ± 2 GeV	(93 ± 2) x 10^{-4} GeV^{-1}
HT150U / $R > 0.50$	286 ± 2 GeV	(91 ± 2) x 10^{-4} GeV^{-1}

Table B.16: HT trigger turn-on parameters for simulated $t\bar{t}(\mu + X)$+jets events satisfying the HAD Box requirements for different R cuts and HT thresholds.
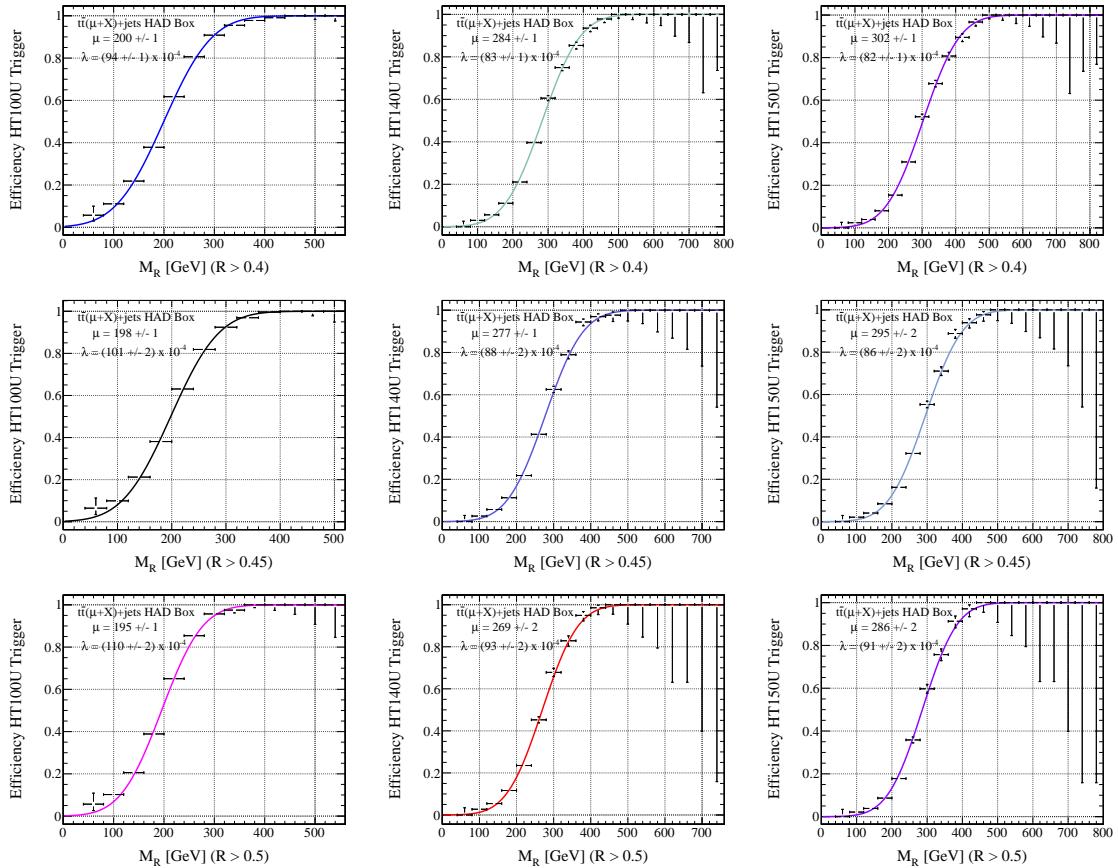
Figure B.8: $t\bar{t}(\mu + X)$+jets emulated HT trigger turn-ons in the HAD Box.

HT threshold / R cut	μ	λ
$W(\tau\nu)$+jets		
HT100U / $R > 0.40$	243 ± 2 GeV	(105 ± 5) x 10^{-4} GeV^{-1}
HT140U / $R > 0.40$	315 ± 3 GeV	(98 ± 4) x 10^{-4} GeV^{-1}
HT150U / $R > 0.40$	336 ± 3 GeV	(90 ± 3) x 10^{-4} GeV^{-1}
HT100U / $R > 0.45$	237 ± 2 GeV	(108 ± 7) x 10^{-4} GeV^{-1}
HT140U / $R > 0.45$	308 ± 3 GeV	(98 ± 5) x 10^{-4} GeV^{-1}
HT150U / $R > 0.45$	331 ± 4 GeV	(86 ± 5) x 10^{-4} GeV^{-1}
HT100U / $R > 0.50$	227 ± 3 GeV	(124 ± 11) x 10^{-4} GeV^{-1}
HT140U / $R > 0.50$	293 ± 4 GeV	(108 ± 8) x 10^{-4} GeV^{-1}
HT150U / $R > 0.50$	312 ± 5 GeV	(96 ± 7) x 10^{-4} GeV^{-1}

Table B.17: HT trigger turn-on parameters for simulated $W(\tau\nu)$+jets events satisfying the HAD Box requirements for different R cuts and HT thresholds.
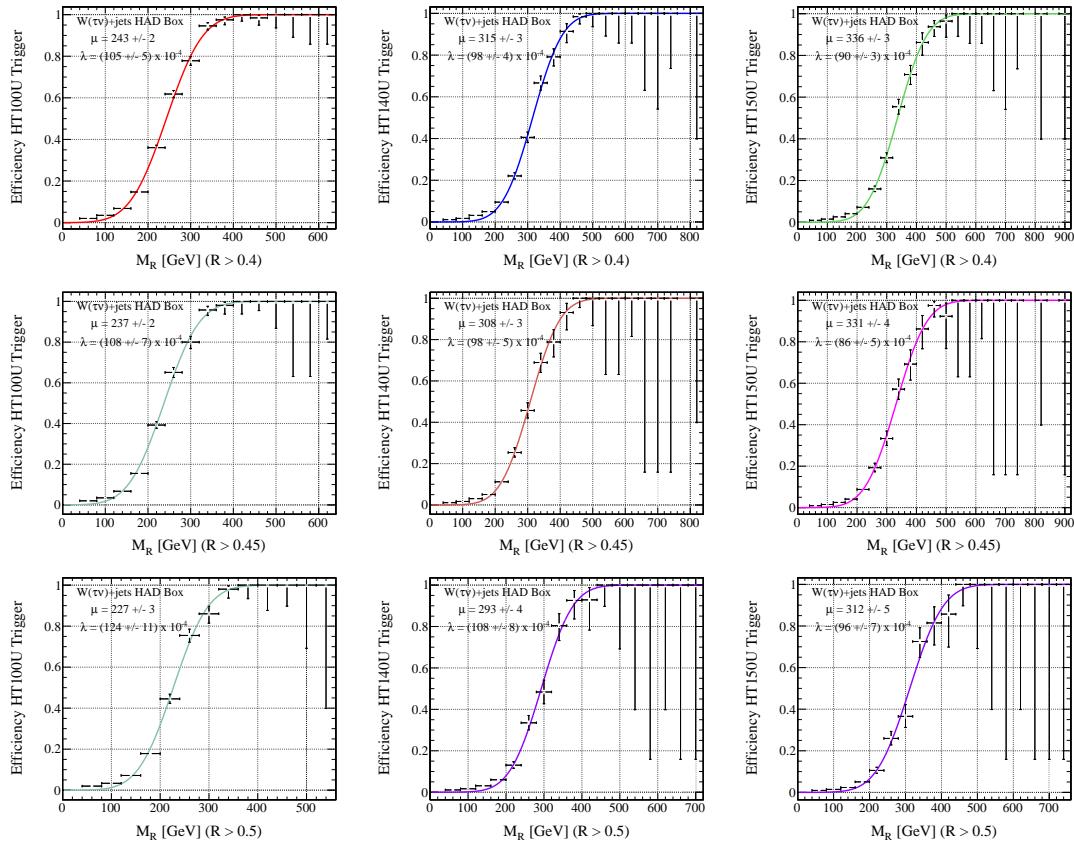
Figure B.9: $W(\tau\nu)$+jets emulated HT trigger turn-ons in the HAD Box.

HT threshold / R cut	μ	λ
$Z(\nu\nu)$+jets		
HT100U / $R > 0.40$	255 ± 2 GeV	$(104 \pm 4) \times 10^{-4}$ GeV^{-1}
HT140U / $R > 0.40$	339 ± 3 GeV	$(98 \pm 3) \times 10^{-4}$ GeV^{-1}
HT150U / $R > 0.40$	359 ± 3 GeV	$(93 \pm 3) \times 10^{-4}$ GeV^{-1}
HT100U / $R > 0.45$	246 ± 2 GeV	$(118 \pm 6) \times 10^{-4}$ GeV^{-1}
HT140U / $R > 0.45$	324 ± 3 GeV	$(113 \pm 5) \times 10^{-4}$ GeV^{-1}
HT150U / $R > 0.45$	343 ± 3 GeV	$(106 \pm 5) \times 10^{-4}$ GeV^{-1}
HT100U / $R > 0.50$	239 ± 2 GeV	$(132 \pm 9) \times 10^{-4}$ GeV^{-1}
HT140U / $R > 0.50$	312 ± 3 GeV	$(122 \pm 6) \times 10^{-4}$ GeV^{-1}
HT150U / $R > 0.50$	329 ± 4 GeV	$(113 \pm 6) \times 10^{-4}$ GeV^{-1}

Table B.18: HT trigger turn-on parameters for simulated $Z(\nu\nu)$+jets events satisfying the HAD Box requirements for different R cuts and HT thresholds.

Figure B.10: $Z(\nu\nu)$+jets emulated HT trigger turn-ons in the HAD Box.

B.3 Evaluation of Potential M_R Bias from Lepton Identification

The lepton boxes in this analysis are used to infer the shapes of the M_R distribution of backgrounds to the HAD box. In order for this strategy to be reliable, we must first demonstrate that lepton identification, for both electrons and muons, does not bias the shape of a given sample's M_R distribution from, for example, a strong correlation between the probability of a lepton being successfully identified and the value of M_R.

In order to evaluate any potential bias in the distribution of M_R from muon reconstruction or identification we consider simulated $W(\mu\nu)$+jets events and define a generator level analogue of the MU Box selection. We denote this selection the A^μ Box; it consists of the requirement that there is at least one generator level muon with $p_T > 20$ GeV/c and $|\eta| < 2.1$, mirroring the kinematic acceptance requirements on the reconstructed muon in the MU Box selection. With respect to the construction of the variables R and M_R, the same reconstructed quantities are used in the A^μ and MU* Boxes, factorizing the effects of muon identification and reconstruction. The M_R distribution of $W(\mu\nu)$+jets events, shown in figure B.11, demonstrates that, despite a difference in the absolute normalization (due to non-unity muon reconstruction and identification efficiencies), the shapes of the M_R distributions for A^μ and MU* Box selected events are the same. This indicates that muon reconstruction and ID is uncorrelated with the behavior of these variables.

We also define a generator level analogue to the ELE Box selection, denoted A^e, which consists of the requirement that there is at least one generator level electron incident in the ECAL barrel or endcap fiducial region with $E_T > 20$ GeV. The only difference here in the treatment of the electron, with respect to the ELE Box approach, is that any reconstructed jets matching the electron are not replaced with the electron momentum. This is done in order to isolate the effect of the electron ID from other kinematic effects. Comparing the R^2 scaling behavior of M_R between simulated $W(e\nu)$+jets events satisfying the ELE selection with those passing the A^e requirements indicates that the electron reconstruction and ID is uncorrelated with
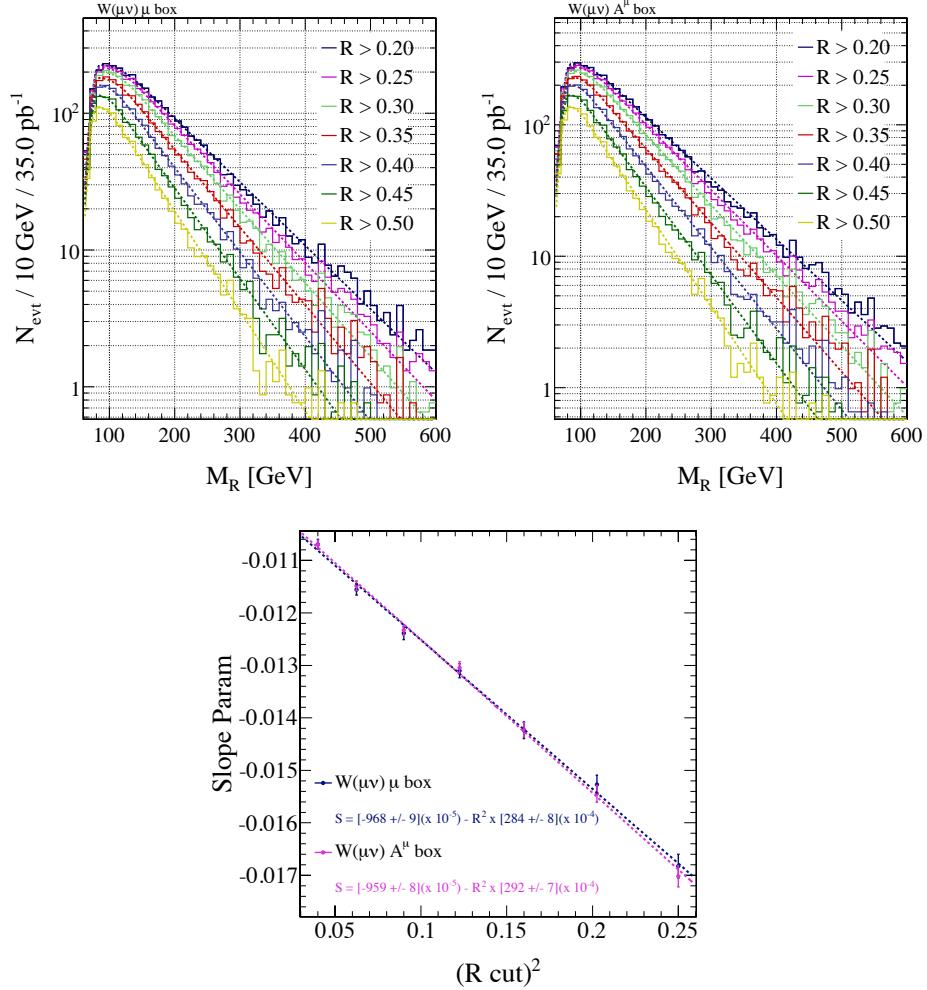
Figure B.11: (Left) M_R distributions for different R selections for simulated $W(\mu\nu)$+jets events satisfying the MU box selection. Fits to the exponential part of the M_R distribution are shown as dotted colored lines. (Center) M_R distributions for different R selections for simulated $W(\mu\nu)$+jets events satisfying the generator level muon acceptance (A^μ) selection. (Right) Value of the exponential slope S from fits to the M_R distribution, as a function of R cut for simulated $W(\mu\nu)$+jets events satisfying the MU box selection, and independently those that satisfy the A^μ selection. Here the muons are treated as neutrinos.

the behavior of R and M_R. The exponential fits for these distributions are shown in figure B.12.
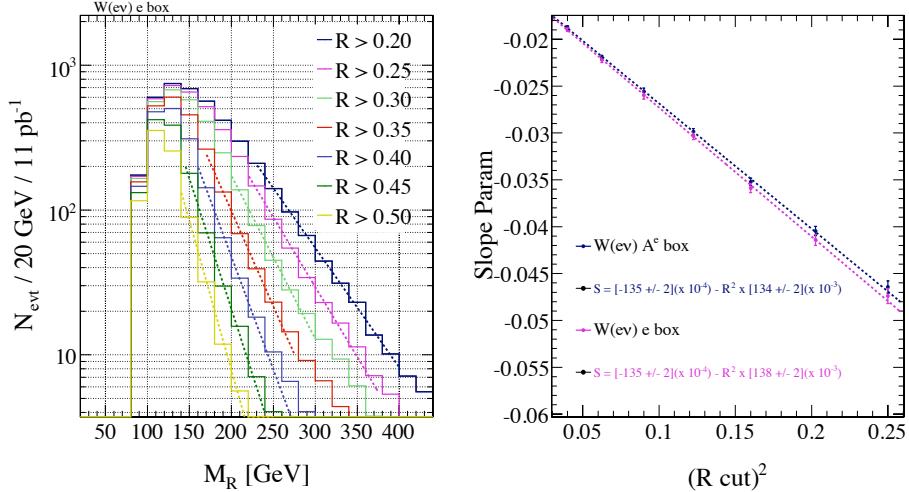
Figure B.12: (Left) Here, we do not use the reconstructed electron to replace the corresponding reconstructed jet in order evaluate any bias in M_R introduced by electron ID. M_R distributions for different values of the cut on R for simulated $W(e\nu)$+jets events satisfying the ELE Box selection. Fits to the exponential part of the M_R distribution are shown as dotted colored lines. (Right) Value of the exponential slope S from fits to the M_R distribution, as a function of R cut for simulated $W(e\nu)$+jets events satisfying the ELE Box selection and independently those that satisfy the generator level electron acceptance (A^e) selection.

B.4 Study of Correlations Between Different R Requirement Fits

The background distribution for M_R is described with a set of exponential functions $f(M_R) = Ne^{-i\kappa M_R}$. Each function, associated to a specific background (QCD multijets, $W(\mu\nu)$+jets, etc.) has a different slope parameters κ which we observe to exhibit M_R scaling with the requirement placed on the variable R, R_{min}, according to

$$\kappa(R_{min}) = a + b \cdot R_{min}^2. \tag{B.3}$$

In the SUSY search, control region measurements at low values of R_{min} are used to measure a and b for different processes in order to predict $\kappa(R_{min})$ in the signal region at large values of R_{min}. In the fits of $\kappa(R_{min})$ values measured in control regions used to measure a and b each point is considered as independent from the others, ignoring the fact that the fitted samples are correlated (each sample being fully included in all the samples corresponding to looser values of R_{min}). We use a set of toy Monte Carlo experiments to determine the effect of the neglected correlation on the determination of a and b for each sample, evaluating whether this procedure incurs a systematic bias and, if so, how large. For each toy experiment proceed as follows:

- We generate exclusive samples in bins of R. For the range $[R^a, R^b]$ we use the distribution $f(M_R) = N_{R^a}e^{-i\kappa_a M_R} - N_{R^b}e^{-i\kappa_b M_R}$, enforcing the relation $\kappa(R_{min}) = a_{true} + b_{true}R_{min}^2$ and taking the number of events in the M_R distribution for each R_{min} value from that observed in data. The last bin is generated inclusively, according to an exponential function.

- The exclusive samples are combined into inclusive samples, equivalent to those used in the analysis.

- A fit in each sample, according to an exponential function, is performed, to determine the value of $\kappa(R_{min})$ in each fit.

- As in the analysis, the values of $\kappa(R_{min})$ are fitted as if the points were independent, determining a and b.

As a result, we obtain a distribution of $a_{fit} - a_{gen}$ and $b_{fit} - b_{gen}$. The distribution is used to quantify the bias on a and b due to neglecting the correlation (the mean of the distribution) and the systematic error associated to it (the RMS of this distribution).

In Tab. B.19 the results of the toys are summarized.

In principle, the procedure suffers of circularity dependence since the test starts from the result of the uncorrelated fit in the analysis, which is itself affected by the bias. As long as the bias is small, this should not result in a significant error. To verify this, we repeated the set of toy experiments of Tab. B.19, shifting the generated values

Data Sample	BOX	a_{true}	b_{true}	$a_{fit} - a_{gen}$	$b_{fit} - b_{gen}$
QCD data	QCD prescaled box	-0.0192	-0.387	-0.0061 ± 0.0010	-0.024 ± 0.005
QCD data	MU control box	-0.0106	-0.117	-0.0005 ± 0.0006	0.000 ± 0.005
QCD data	ELE control box	-0.0208	-0.180	0.0008 ± 0.0001	-0.016 ± 0.003
$W(\mu\nu)$+jets	MU Box	-0.0096	-0.026	0.0008 ± 0.0007	-0.004 ± 0.007
$W(e\nu)$+jets	ELE Box	-0.0102	-0.028	0.0001 ± 0.0022	-0.028 ± 0.030

Table B.19: Values of $a_{fit} - a_{gen}$ and $b_{fit} - b_{gen}$ determined from toy Monte Carlo experiments for each background sample in the analysis. Results are quoted as (mean value $\pm$ RMS) of the distribution.

by minus the bias of the tables. The result obtained in this case is the same as for Tab. B.19, such that this fit gives, on average, the result we find when the correlations are neglected. This proves that the bias does not depend on the central value used in the toy generation in the small, relevant range of a and b we are considered.

B.5 ELE, MU, and HAD Box Background Predictions by Process

Tables B.20, B.21, B.22 show the expected background composition for the three signal-candidate regions of the ELE, MU and HAD boxes, taking into account all nonnegligible backgrounds.

Background Process	Prediction
$W(\ell\nu)$+jets	0.31 ± 0.17
$t\bar{t}$+jets	0.32 ± 0.20
$Z(\ell\ell)$+jets	0.001 ± 0.001

Table B.20: Background breakdown for ELE Box background prediction with $R > 0.45$ and $M_R > 500$ GeV.

Background Process	Prediction
$W(\ell\nu)$+jets	0.20 ± 0.11
$t\bar{t}$+jets	0.31 ± 0.19
$Z(\ell\ell)$+jets	0.002 ± 0.002

Table B.21: Background breakdown for MU Box background prediction with $R > 0.45$ and $M_R > 500$ GeV.

Background Process	Prediction
$W(\ell\nu)$+jets	3.2 ± 1.3
$t\bar{t}$+jets	0.70 ± 0.35
$Z(\nu\nu)$+jets	1.7 ± 0.51

Table B.22: Background break-down for HAD Box background prediction with $R > 0.5$ and $M_R > 500$ GeV.

B.6   Background Prediction Coverage Tests

For the MU, ELE and HAD Box background predictions, measurements of the shape and normalization of the M_R distribution for different processes is used to calculate the integral of the M_R distribution over each signal region, corresponding to the total number of expected SM background events. In addition, the input measurements and assumptions that go into these predictions have associated errors, both systematic and statistical in nature, that are propagated to these integrals and background predictions. In order to check that our procedure for predicting these background yields is unbiased and that the errors we quote correspond to 68% probability interval, we perform toy Monte Carlo experiments where we repeat the fitting procedure on trial pseudodata. The procedure for generating and analyzing these toy Monte Carlo experiments is summarized as follows:

- Perform background estimation with data. This provides the central values for any parameters included in the background estimation (normalizations, shape parameters).

- Generate a *trial* background by randomly generating trial values for any parameters with systematic errors. Trial values are drawn from normal distributions

with central values equal to the values observed in data and standard deviation equal to the systematic error on the parameter.

- From the trial background calculate the M_R integral in the signal region. This corresponds to the true background yield for the trial.

- From the trial background, generate a trial pseudodata set by drawing bin-by-bin event yields from Poisson distributions with peak value equal to the integral of the trial background distribution in each bin.

- Perform the background estimation procedure, treating the trial pseudodata in the same way as the observed data.

We use the MU Box observed M_R yields as an example to test the background prediction. For each value of the M_R/R cut, we generate 100K trial M_R background distributions and repeat the above procedure, arriving at 100K alternate background predictions. These background predictions correspond to the central value of the true background, whereas the observed yields are distributed as a Poisson around this central value. In order to check whether these background predictions are an unbiased estimator of the "true" background, and that the corresponding error on that prediction covers a 68% probability interval, we compare the predictions with the corresponding true values, shown in figure B.13. We do not observe any significant biases for any M_R/R cut values, and in each case the $\pm 1\ \sigma$ interval covers approximately 68%, which implies the quoted systematic errors have the intended meaning. From these observations we concluded that the background prediction procedure is unbiased, and that systematic errors on the different parameters entering the background prediction are correctly propagated to the final prediction.

B.6.1 Compatibility of Predicted and Observed Yield in the MU Box

Using the same framework as for the coverage tests, we can predict the distribution of the number of observed events, for a given set of R/M_R cuts, in the MU Box. This
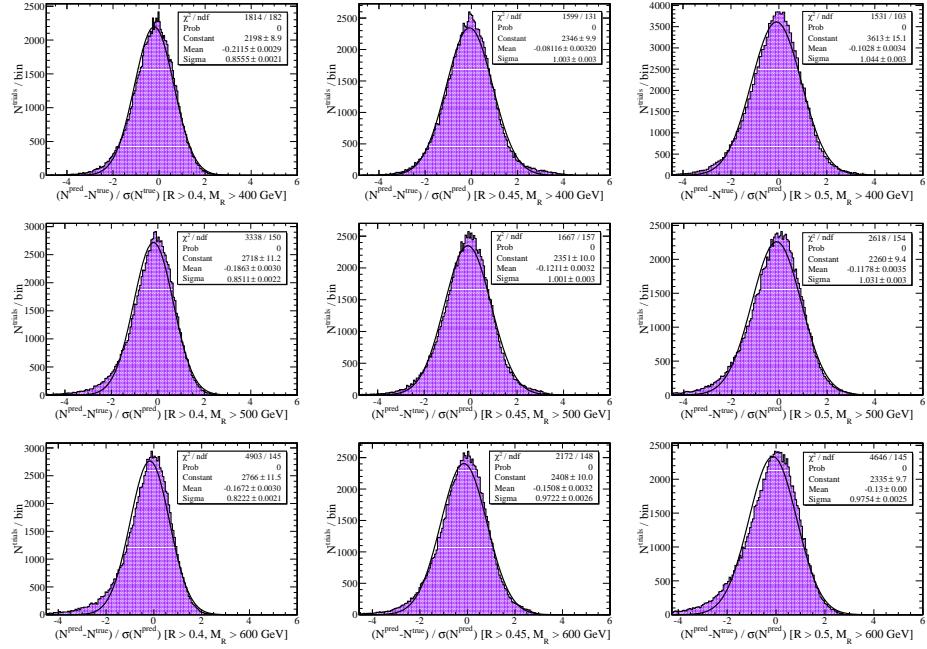
Figure B.13: Background prediction pulls for the MU Box, for different R/M_R cut values. Shown for M_R cuts of (top) 400 GeV, (middle) 500 GeV, (bottom) 600 GeV and R cuts of (left) 0.4, (middle) 0.45, (right) 0.5 .

distribution can then be used to assign probabilities (p-values) to the observations we have made in the data. These toy MC background predictions contain all of the elements of the analysis background prediction:

- systematic uncertainties on parameters used to determine the background Poisson smearing of the "true" expected yields

- Data-driven normalizations and parameter determination from M_R sideband

The resulting distributions of the observed yields, for different M_R/R cuts, are the most accurate predictions of these distributions available within the analysis, such that toys modeling the systematic error of the background predictions as normal distributed or log-normal are essentially approximations to this fully-detailed approach.

We use the expected *pdf* $N_{observed}$ distribution and the number of actual observed events, to calculate the probability of observing an equal or less probable configuration (Z likelihood). The results are shown for different R/M_R boxes, inclusive and

exclusive in M_R in figures B.14-B.16. We don't observe any significant deviations from expectations.

These tests demonstrate that the configuration of observed event yields in the MU Box are not that improbable, even though it may not appear to be the case by eye. These same tests prove that the background prediction procedure is unbiased, and that the quoted systematic errors mean what they are intended to. Both these points are inter-consistent; the toys tell us that the Poissonian contribution to these yields, in the regime we are working at (low expected yields, nontrivial relative systematic uncertainties) are highly nontrivial and dominate the resulting distribution of observed yields. In the MU Box the largest effect is 1.8σ, corresponds to a nonoptimal set of cuts.
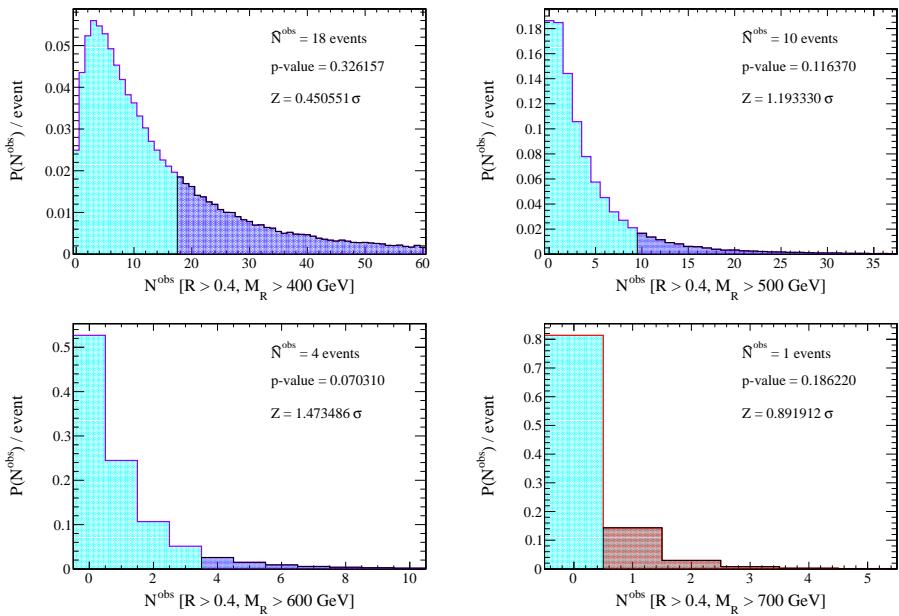
Figure B.14: Expected *pdf* for N_{obs}, for MU Box selected events with $R > 0.4$ and different inclusive M_R cuts as indicated.
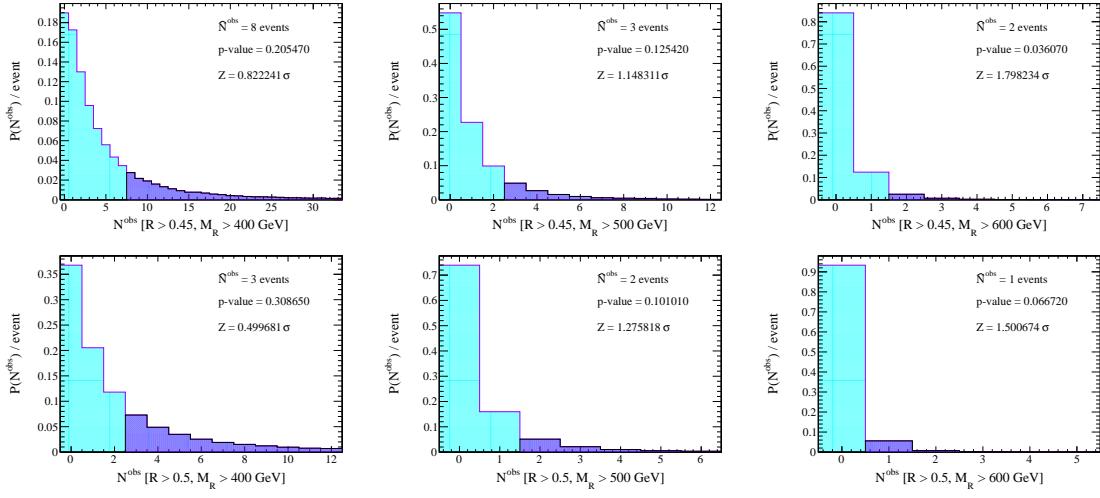
Figure B.15: Expected *pdf* for N_{obs}, for MU Box selected events with $R > 0.45$ and $R > 0.5$ and different inclusive M_R cuts as indicated.
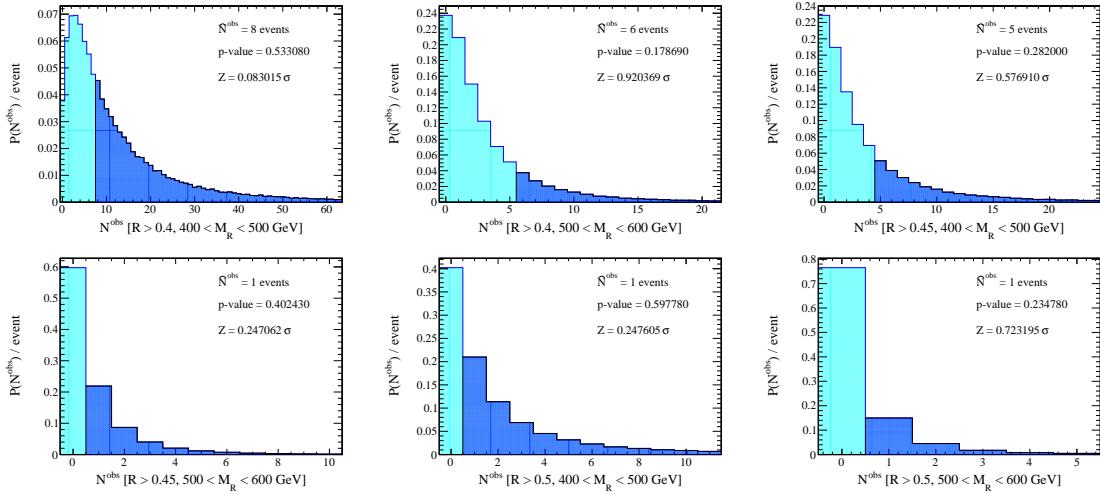
Figure B.16: Expected *pdf* for N_{obs}, for MU Box selected events with $R > 0.4$, $R > 0.45$ and $R > 0.5$ and different exclusive M_R bins as indicated.

B.7 M_R as a Function of Instantaneous Luminosity

During the course of data taking, the instantaneous luminosity and noise profiles in different subdetector is changing. In order to evaluate the effect of these variations we compare the distribution of M_R in the QCD control box between different run ranges and as a function of the number of reconstructed primary vertices, which is indicative of the instantaneous luminosity. These comparisons are shown in figure B.17. With the 35 pb^{-1} data, we observe no significant changes in the shape of the M_R distribution between the different primary datasets used or as a function of the number of reconstructed primary vertices.
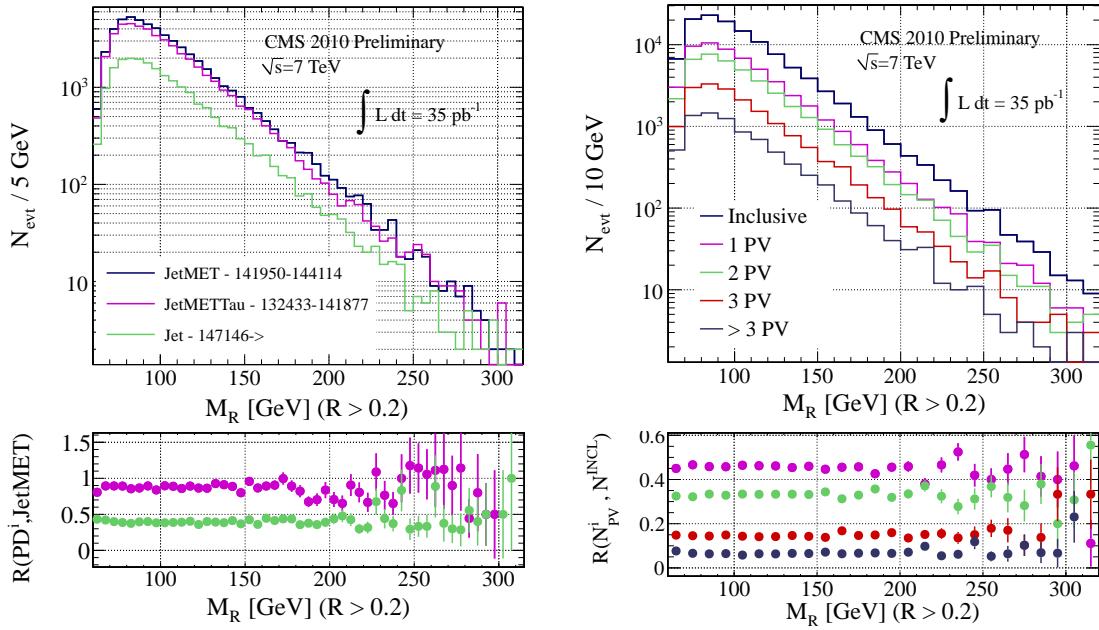
Figure B.17: (Left) M_R distribution for the QCD control box, split into different primary datasets, and the ratio of each distribution relative to PD JetMET. (Right) M_R distribution for QCD control box, shown as a function of the number of reconstructed primary vertices. The ratio of each M_R distribution to the inclusive distribution is shown.

B.8   EWK Processes in the QCD Control Box

The QCD control box provides the opportunity to select an unbiased M_R distribution (w.r.t. and H_T trigger), such that in the low M_R, QCD-dominated, region we can measure the evolution of the QCD M_R shape as a function of the R cut. Since the slope of the QCD M_R distribution is significantly steeper than for the EWK and top backgrounds these other backgrounds become relatively larger with increasing M_R.

We quantify the presence of non-QCD multi-jet processes in the QCD control box in order to distinguish between these events and the possibility of an additional background. To do this, we follow the same procedure as for the hadronic box background prediction, taking the normalizations from a measurement in the MU Box. Specifically, we measure the yield in the MU Box for $100 < M_R < 250$ GeV, requiring that the `HLT_DiJetAve15U` trigger also fired. This allows us to measure the effect of any prescales on the `HLT_DiJetAve15U` selected events. This measurement is shown in figure B.18.
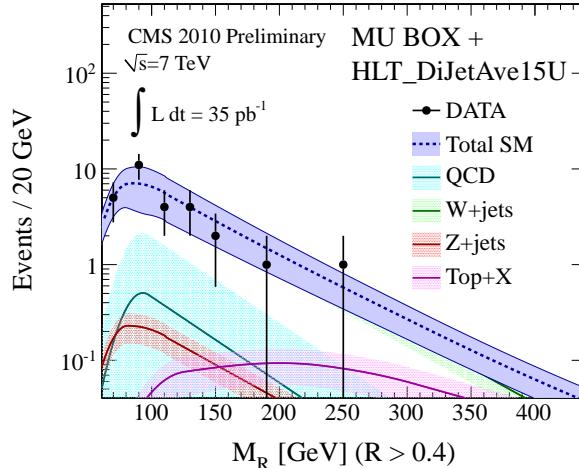
Figure B.18: M_R distribution for the MU Box with $R > 0.2$ and the additional requirement that the prescaled HLT_DiJetAve15U trigger was fired. The event yields here are used to estimate the EWK contribution to the QCD control box, which also includes the HLT_DiJetAve15U requirement.

Using these selected MU Box events we measure an effective integrated luminosity of 67 ± 31 nb^{-1}. If we assume that the total dataset corresponds to an integrated lumi-

nosity of 36.1 pb^{-1} this corresponds to an effective prescale for the `HLT_DiJetAve15U` trigger path of $\sim$500$\pm$250. This measurement is then used to normalize the EWK backgrounds in the QCD control box. These background predictions are shown in figure B.19, for different R cuts.
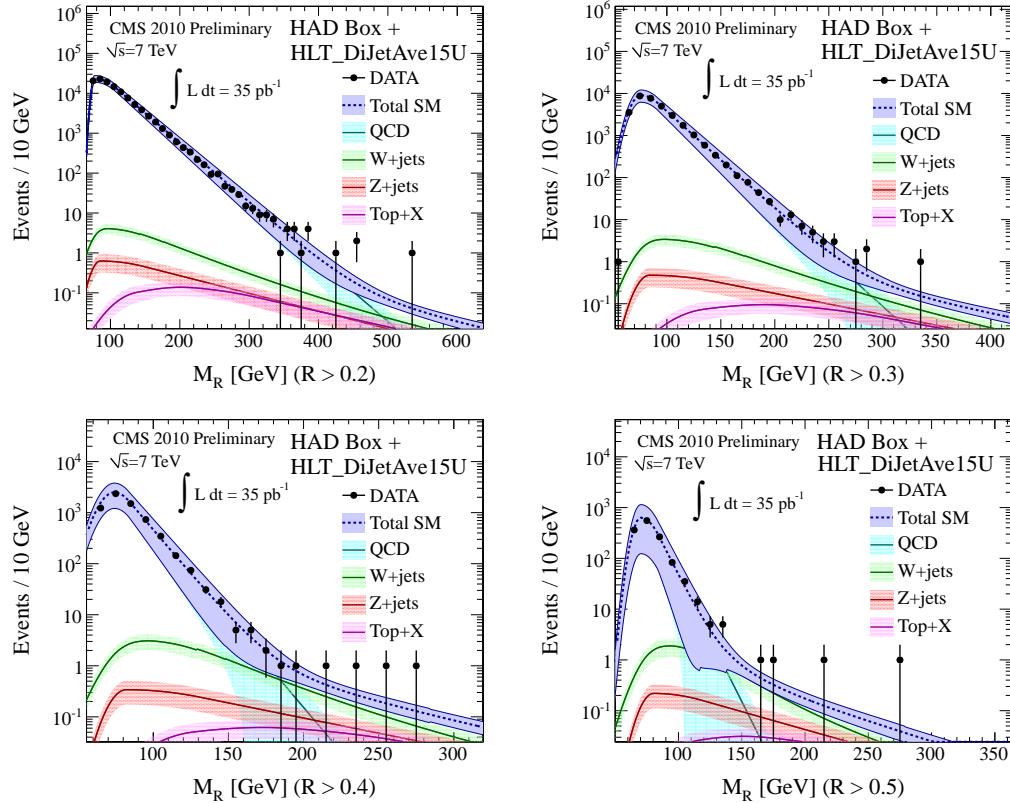
Figure B.19: M_R distribution for the MU Box with the additional requirement that the prescaled HLT_DiJetAve15U trigger was fired. (Top left) $R > 0.2$. (Top right) $R > 0.3$. (Bottom left) $R > 0.4$. (Bottom right) $R > 0.5$. The event yields here are used to estimate the EWK contribution to the QCD control box, which also includes the HLT_DiJetAve15U requirement.

We find that the combination of a single QCD exponential component with the EWK and top backgrounds provides a good description of the observed data. Predicted and observed yields for different values of R/M_R cuts are summarized in Table B.23.

R cut $/M_R$ cut	Predicted BKG	Observed
$R > 0.2$ / $M_R > 200$ GeV	1621 ± 350	1528
$R > 0.2$ / $M_R > 300$ GeV	56 ± 13	56
$R > 0.2$ / $M_R > 400$ GeV	3.07 ± 0.81	4
$R > 0.3$ / $M_R > 200$ GeV	39 ± 9	45
$R > 0.3$ / $M_R > 300$ GeV	1.87 ± 0.48	1
$R > 0.3$ / $M_R > 400$ GeV	0.50 ± 0.15	0
$R > 0.4$ / $M_R > 200$ GeV	2.99 ± 0.74	4
$R > 0.4$ / $M_R > 300$ GeV	0.56 ± 0.17	0
$R > 0.4$ / $M_R > 400$ GeV	0.13 ± 0.05	0
$R > 0.5$ / $M_R > 200$ GeV	0.87 ± 0.24	2
$R > 0.5$ / $M_R > 300$ GeV	0.14 ± 0.05	0
$R > 0.5$ / $M_R > 400$ GeV	0.025 ± 0.009	0

Table B.23: Predicted and observed yields for the QCD Control box for different values of the R/M_R cut

B.9 Background Prediction Closure Tests in Simulated Events

In order to test the background prediction procedure for the different boxes we perform closure tests using simulated events. For each box we take the binned M_R distribution predicted from simulated events, normalized to 35 pb^{-1} of integrated luminosity using the theoretical (N)LO cross sections for each background process, respectively. These M_R distributions are added together, yielding the MC background prediction for a given set of R/M_R cuts and box requirements. Since the MC background M_R distribution is made up of weighted events, there are cases where single bins have less than 1 event. In order to ensure that the background prediction procedure is done in the same way as for data, we take the simulation predicted binned M_R distribution (1 bin / GeV) and set each bin value to a number drawn from a Poisson distribution centered at the simulation predicted value.

With these simulation pseudo-data sets we then repeat the same background prediction procedure, for each box, as is done for events selected in data. The background predictions for the MU, ELE and HAD boxes, using simulation pseudo-data, are shown in figures B.20 and B.21, respectively.
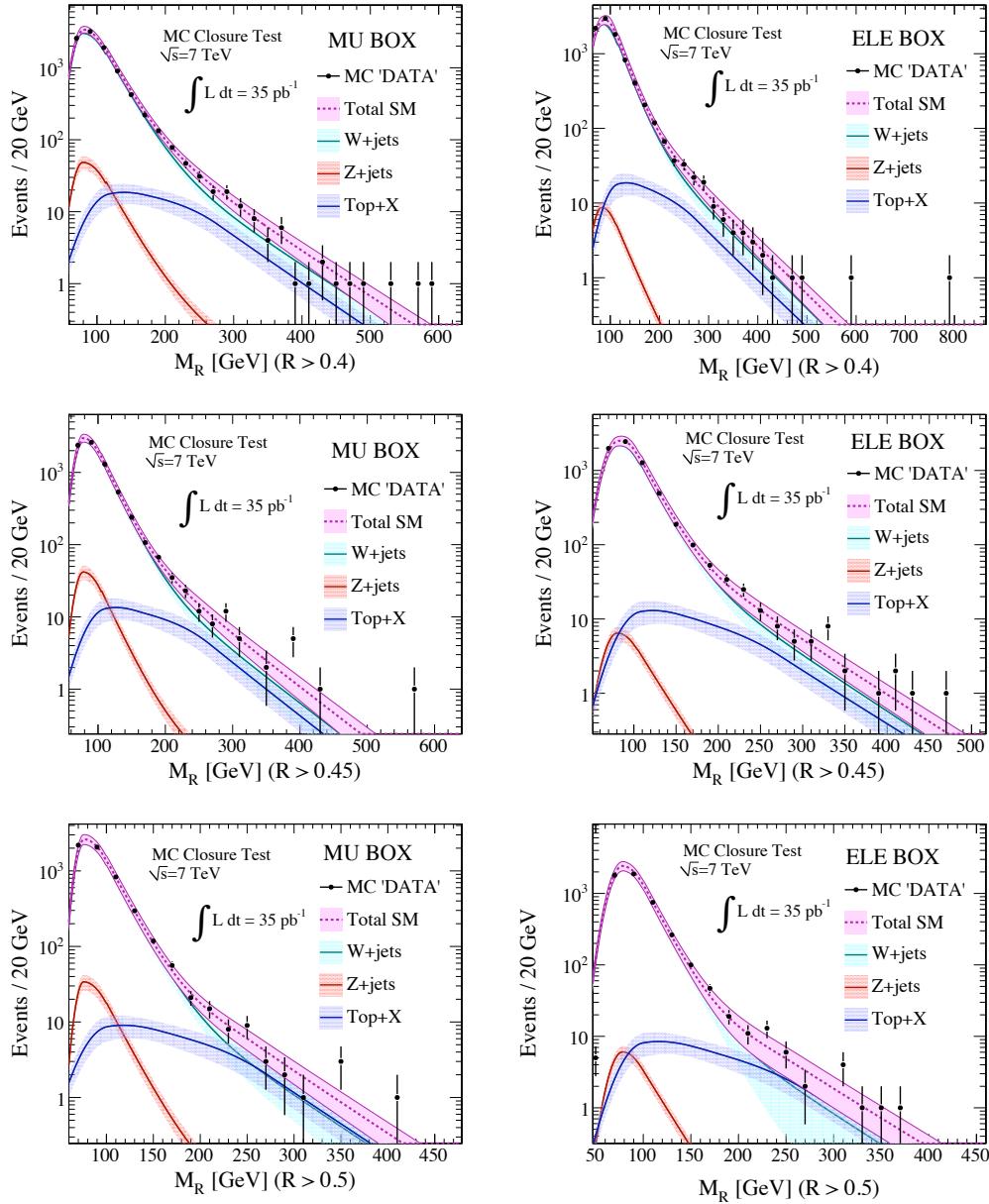
Figure B.20: (Left) M_R distribution for the MU Box for different R cuts from the background prediction closure test using simulated events. (Right) M_R distributions for the ELE Box from closure tests.
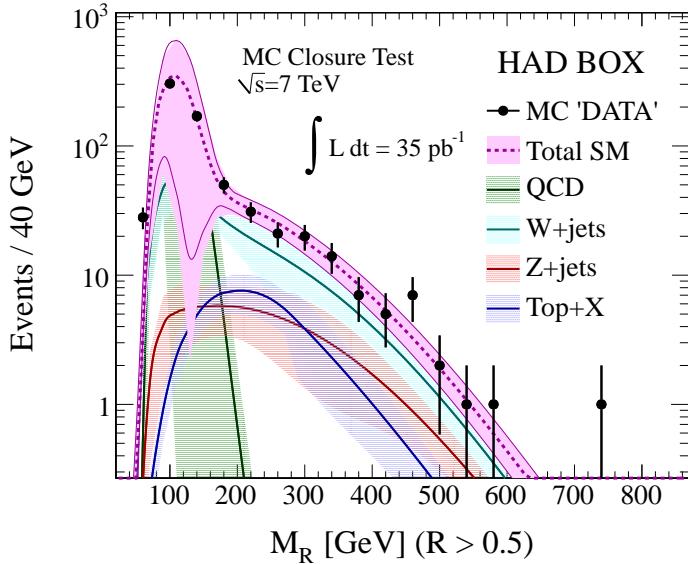
Figure B.21: M_R distribution for the HAD Box for $R > 0.5$ from the background prediction closure test.

As is done for data, the normalizations of backgrounds in the hadronic box are taken from measurements in the MU and ELE boxes. Table B.24 summarizes the values of these normalizations as measured from the simulation pseudo-data. We find agreement with the value used to generate the pseudo-data, 35 pb^{-1}.

Box / R cut	$A_{corr}^{\ell \text{ Box}}$
MU Box $R > 0.40$	34.5 ± 1.2 pb^{-1}
MU Box $R > 0.45$	34.3 ± 1.6 pb^{-1}
MU Box $R > 0.50$	35.4 ± 2.2 pb^{-1}
ELE Box $R > 0.40$	35.1 ± 1.6 pb^{-1}
ELE Box $R > 0.45$	34.5 ± 1.9 pb^{-1}
ELE Box $R > 0.50$	35.5 ± 2.4 pb^{-1}

Table B.24: Lepton efficiency-corrected normalizations, $A_{corr}^{\ell \text{ Box}}$, from MC closure tests in the lepton boxes (35 pb^{-1} simulated).

The background predictions for the MU, ELE and HAD boxes, for the simulation pseudo-data, are summarized in Tables B.25, B.26 and B.27, respectively. In all cases we observe agreement between the predicted and observed yields.

R cut /M_R cut	Predicted BKG	Observed
$R > 0.40$ / $M_R > 400$ GeV	9.9 ± 3.6	9
$R > 0.40$ / $M_R > 500$ GeV	2.33 ± 1.00	3
$R > 0.40$ / $M_R > 600$ GeV	0.55 ± 0.28	0
$R > 0.45$ / $M_R > 400$ GeV	3.25 ± 1.31	2
$R > 0.45$ / $M_R > 500$ GeV	0.62 ± 0.29	1
$R > 0.45$ / $M_R > 600$ GeV	0.12 ± 0.07	0
$R > 0.50$ / $M_R > 400$ GeV	0.89 ± 0.38	1
$R > 0.50$ / $M_R > 500$ GeV	0.14 ± 0.07	0
$R > 0.50$ / $M_R > 600$ GeV	0.021 ± 0.011	0

Table B.25: Predicted and observed yields for MU Box from background prediction closure test.

R cut /M_R cut	Predicted BKG	Observed
$R > 0.40$ / $M_R > 400$ GeV	9.1 ± 2.9	7
$R > 0.40$ / $M_R > 500$ GeV	2.08 ± 0.76	2
$R > 0.40$ / $M_R > 600$ GeV	0.48 ± 0.20	1
$R > 0.45$ / $M_R > 400$ GeV	2.92 ± 1.10	4
$R > 0.45$ / $M_R > 500$ GeV	0.55 ± 0.23	0
$R > 0.45$ / $M_R > 600$ GeV	0.10 ± 0.05	0
$R > 0.50$ / $M_R > 400$ GeV	0.76 ± 0.38	1
$R > 0.50$ / $M_R > 500$ GeV	0.12 ± 0.07	0
$R > 0.50$ / $M_R > 600$ GeV	0.018 ± 0.009	0

Table B.26: Predicted and observed yields for ELE Box from background prediction closure test.

M_R cut cut	Predicted	Observed
$M_R > 400$ GeV	13.8 ± 3.4	17
$M_R > 500$ GeV	3.15 ± 0.92	4
$M_R > 600$ GeV	0.66 ± 0.21	1

Table B.27: Predicted and observed yields for different M_R cuts with $R > 0.5$ in the HAD Box from background prediction closure test.

B.10   Signal Efficiency Maps

B.10.1   CMSSM

Efficiency maps (covering the grid of model points considered in the interpretation of search results) are included for each step in the analysis selection and for each final state box.
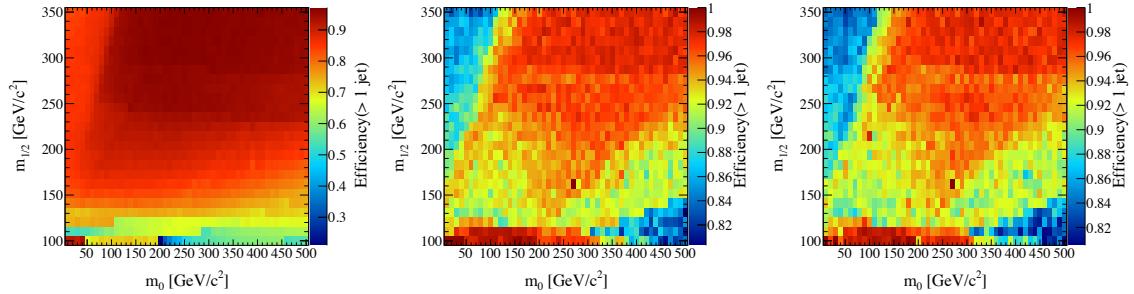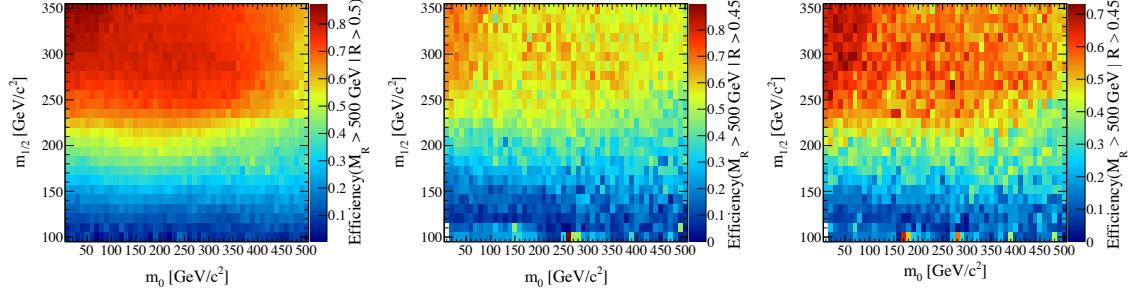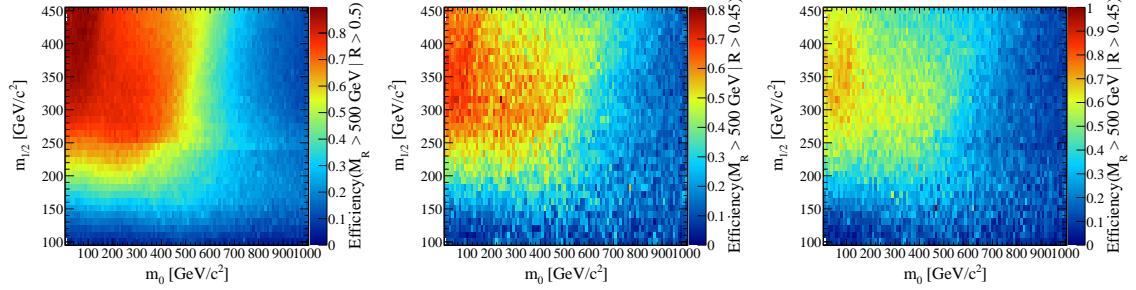

Figure B.22: Efficiency of ≥ 2 jet requirement for CMSSM models ($\tan\beta = 3$, $A_0 = 0$, $\mathrm{sgn}\mu = +$) as a function of the parameters m_0 and $m_{1/2}$. Results are shown for the (Left) HAD Box, (Center) ELE Box and (Right) MU Box selections.
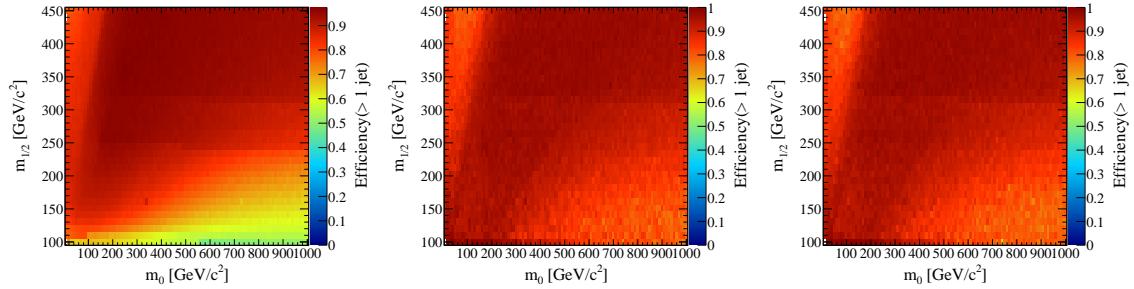
Figure B.23: Efficiency of ≥ 2 jet requirement for CMSSM models ($\tan\beta = 10$, $A_0 = 0$, $\mathrm{sgn}\mu = +$) as a function of the parameters m_0 and $m_{1/2}$. Results are shown for the (Left) HAD Box, (Center) ELE Box and (Right) MU Box selections.

B.10.2   Simplified Models

Efficiency maps (covering the grid of model points considered in the interpretation of search results) are included for each step in the analysis selection and for each final state box.
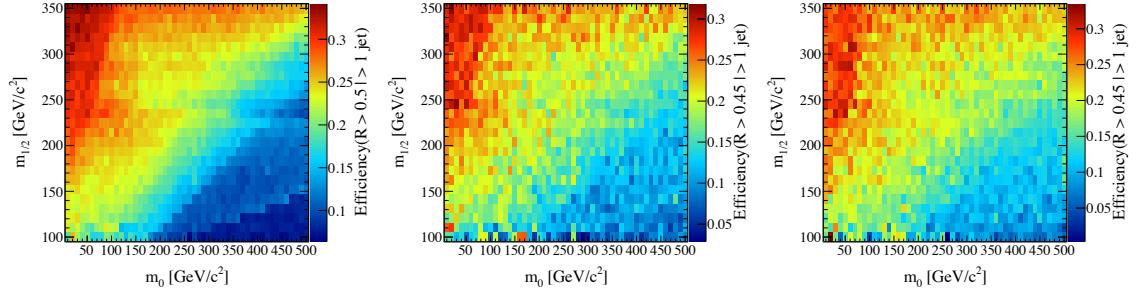
Figure B.24: Efficiency of R cut requirement for CMSSM models ($\tan \beta = 3$, $A_0 = 0$, sgn$\mu = +$) as a function of the parameters m_0 and $m_{1/2}$. Cut efficiencies are calculated using events that satisfy the ≥ 2 jet requirement. Results are shown for the (Left) HAD Box ($R > 0.5$), (Center) ELE Box ($R > 0.45$) and (Right) MU Box ($R > 0.45$) selections.
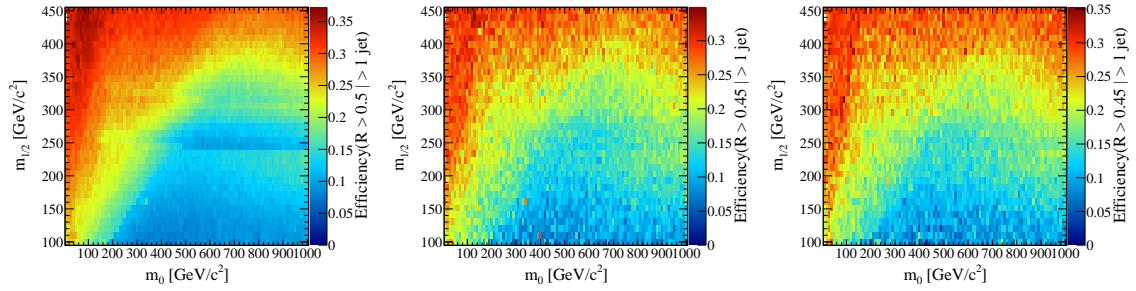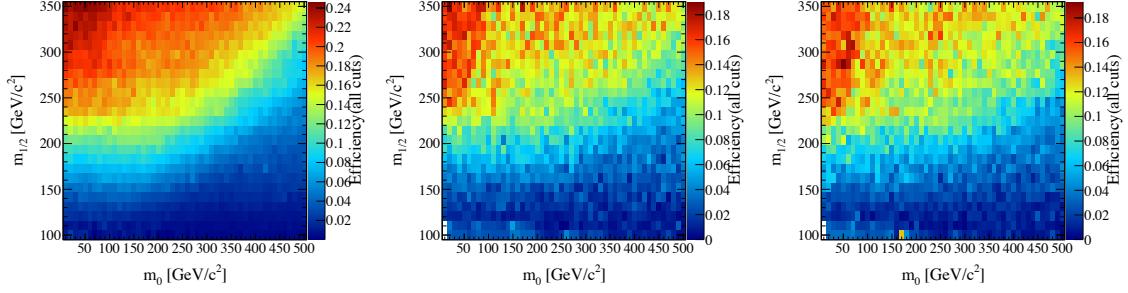

Figure B.25: Efficiency of R cut requirement for CMSSM models ($\tan \beta = 10$, $A_0 = 0$, sgn$\mu = +$) as a function of the parameters m_0 and $m_{1/2}$. Cut efficiencies are calculated using events that satisfy the ≥ 2 jet requirement. Results are shown for the (Left) HAD Box ($R > 0.5$), (Center) ELE Box ($R > 0.45$) and (Right) MU Box ($R > 0.45$) selections.

B.11   NLO Factorization and Renormalization Scale Uncertainties

The systematic uncertainty on signal yields resulting from PDF uncertainties on NLO cross sections is evaluated, using the final signal region definitions for each box. The analysis selection is repeated for events normalized using scale "up" and "down" cross sections where the effective value of α_S has been varied up and down by its uncertainty in these samples, respectively. The percent deviation of the final signal yield for these two scenarios is used to assess a systematic uncertainty on the signal yield prediction

Figure B.26: Efficiency of $M_R > 500$ GeV cut requirement for CMSSM models ($\tan\beta = 3$, $A_0 = 0$, $\mathrm{sgn}\mu = +$) as a function of the parameters m_0 and $m_{1/2}$. Cut efficiencies are calculated using events that satisfy the R cut requirement. Results are shown for the (Left) HAD Box ($R > 0.5$), (Center) ELE Box ($R > 0.45$) and (Right) MU Box ($R > 0.45$) selections.

Figure B.27: Efficiency of $M_R > 500$ GeV cut requirement for CMSSM models ($\tan\beta = 10$, $A_0 = 0$, $\mathrm{sgn}\mu = +$) as a function of the parameters m_0 and $m_{1/2}$. Cut efficiencies are calculated using events that satisfy the R cut requirement. Results are shown for the (Left) HAD Box ($R > 0.5$), (Center) ELE Box ($R > 0.45$) and (Right) MU Box ($R > 0.45$) selections.

accounting for these PDF uncertainties.

Figure B.28: Total razor efficiency for CMSSM models with ($\tan\beta = 3$, $A_0 = 0$, sgnμ = +) in the m_0-$m_{1/2}$ plane. The efficiency maps are for the (Left) HAD Box, (Center) ELE Box and (Right) MU Box selections.
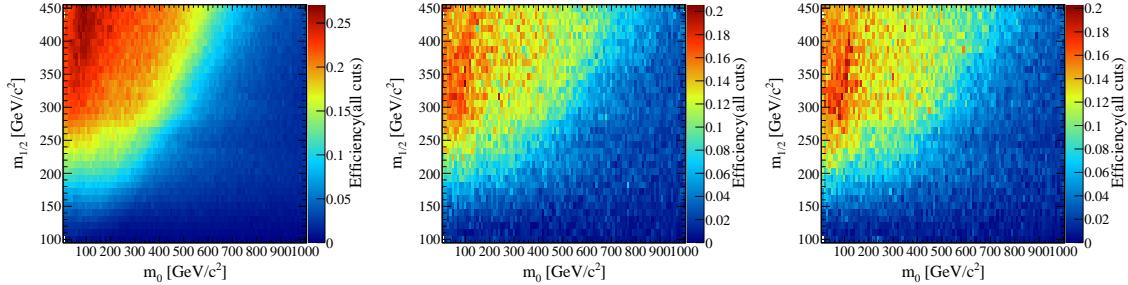
Figure B.29: Total razor efficiency for CMSSM models with ($\tan\beta = 10$, $A_0 = 0$, sgnμ = +) in the m_0-$m_{1/2}$ plane. The efficiency maps are for the (Left) HAD Box, (Center) ELE Box and (Right) MU Box selections.
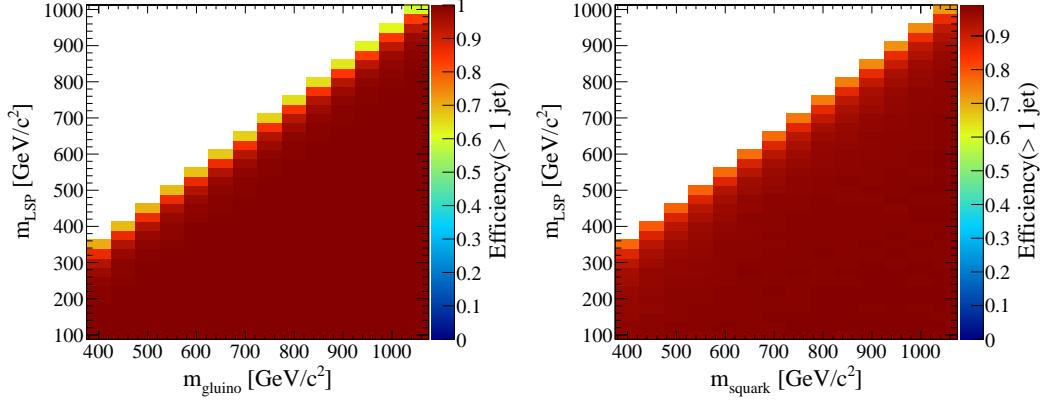
Figure B.30: Efficiency of ≥ 2 jet requirement in the HAD Box, as a function of LSP and gluino/squark mass. Results are shown for the (Left) di-gluino production simplified model and (Right) di-squark production simplified model.
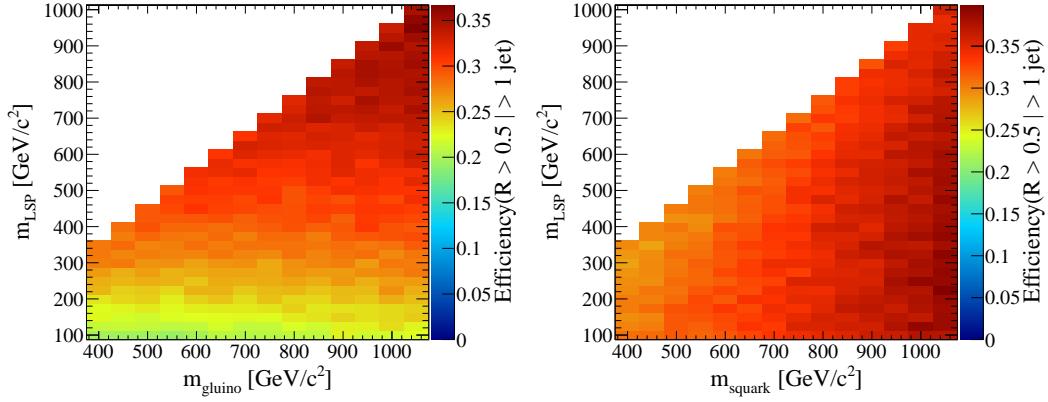
Figure B.31: Efficiency of $R > 0.5$ cut requirement in the HAD Box, as a function of LSP and gluino/squark mass. Cut efficiencies are calculated using events that satisfy the ≥ 2 jet requirement. Results are shown for the (Left) di-gluino production simplified model and (Right) di-squark production simplified model.
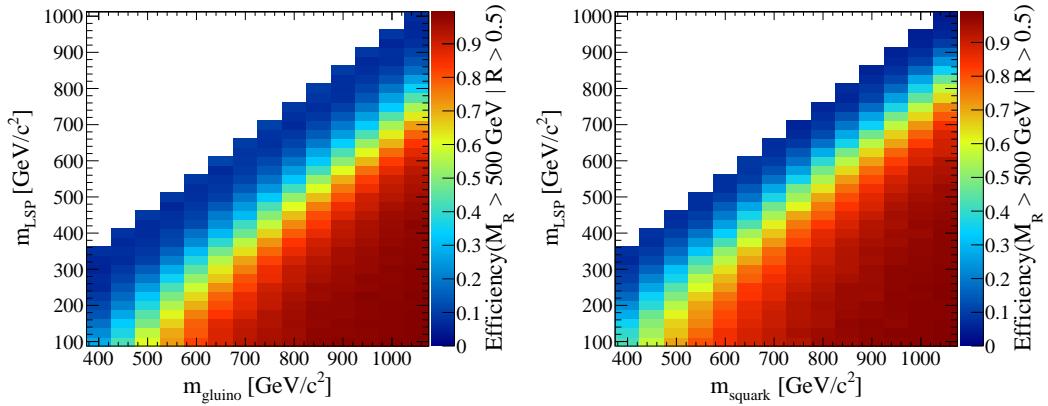
Figure B.32: Efficiency of $M_R > 500$ GeV cut requirement in the HAD Box, as a function of LSP and gluino/squark mass. Cut efficiencies are calculated using events that satisfy the $R > 0.5$ cut requirement. Results are shown for the (Left) di-gluino production simplified model and (Right) di-squark production simplified model.
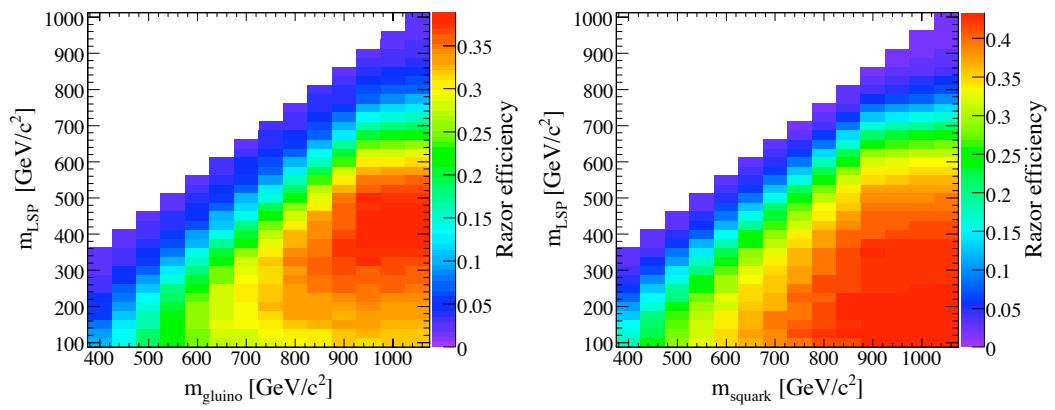
Figure B.33: Razor efficiency in the HAD Box, as a function of the LSP and gluino/squark mass. Results are shown for the (Left) di-gluino production simplified model and (Right) di-squark production simplified model.
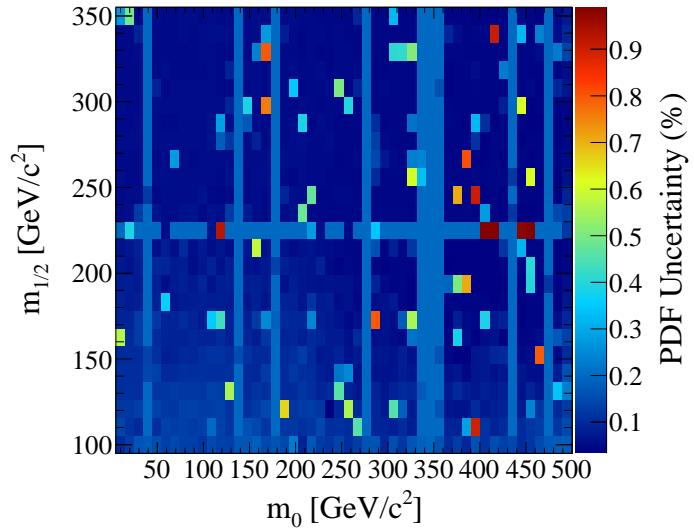
B.11.1   CMSSM
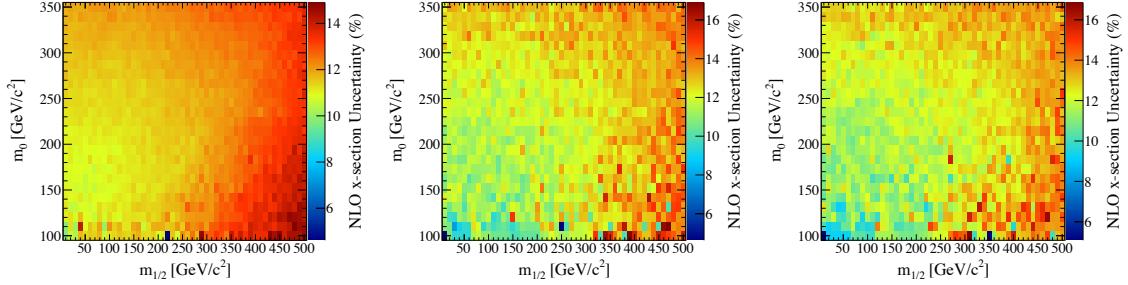
Figure B.34: NLO factorization and renormalization scale uncertainties on the expected signal yield in the signal region for CMSSM models ($\tan\beta = 3$, $A_0 = 0$, sgnμ = +), as a function of the parameters m_0 and $m_{1/2}$. Results are shown for the (Left) HAD Box ($R > 0.5$, $M_R > 500$ GeV), (Center) ELE Box ($R > 0.45$, $M_R > 500$ GeV) and (Right) MU Box ($R > 0.45$, $M_R > 500$ GeV) selections.
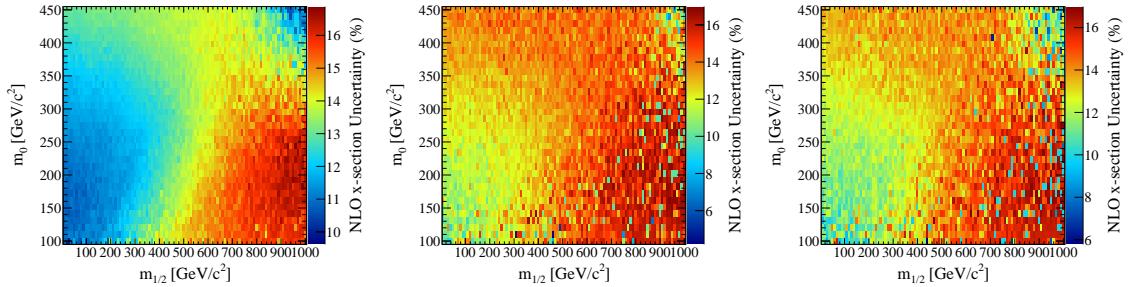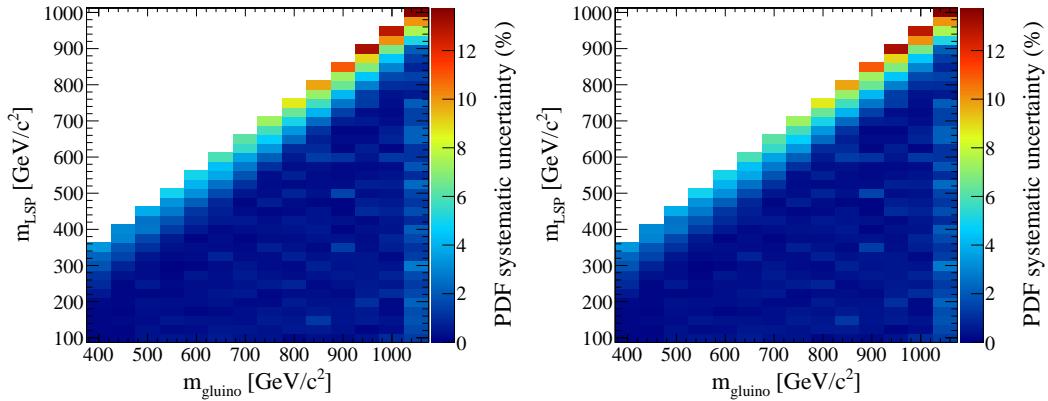

Figure B.35: NLO factorization and renormalization scale uncertainties on the expected signal yield in the signal region for CMSSM models ($\tan\beta = 10$, $A_0 = 0$, sgnμ = +), as a function of the parameters m_0 and $m_{1/2}$. Results are shown for the (Left) HAD Box ($R > 0.5$, $M_R > 500$ GeV), (Center) ELE Box ($R > 0.45$, $M_R > 500$ GeV) and (Right) MU Box ($R > 0.45$, $M_R > 500$ GeV) selections.

Figure B.36: PDF uncertainties on the expected signal yield in the signal region for CMSSM models ($\tan \beta = 3$, $A_0 = 0$, sgn$\mu = +$), as a function of the parameters m_0 and $m_{1/2}$. Results are shown for the (Left) HAD Box ($R > 0.5$, $M_R > 500$ GeV), (Center) ELE Box ($R > 0.45$, $M_R > 500$ GeV) and (Right) MU Box ($R > 0.45$, $M_R > 500$ GeV) selections.

B.11.2 Simplified Models

Figure B.37: PDF uncertainties on the expected signal yield in the signal region for simplified models. (Left) di-gluino production. (Right) di-squark production. Results are shown for the HAD Box ($R > 0.5$, $M_R > 500$ GeV) selection.

The ISR uncertainty is $\mathcal{O}(1\%)$ except close to the diagonal in the SMS models where it is up to 10%. We have performed the PDF and ISR scans of the unceratinties point by point and use them in the final limit.

B.12   JES Uncertainties

The systematic uncertainty on signal yields resulting from jet energy scale (JES) uncertainties [212] is evaluated, using the final signal region definitions for each box. These are uncertainties are determined using p_T and η dependent JES uncertainty maps, where each event is considered twice, once where jets' momentum are systematically increased by one standard deviation and then again where they are instead decreased by the same factor. The analysis selection is repeated for events with these two treatments, and the percent deviation of the final signal yield from the nominal JES for these two scenarios is used to assess a systematic uncertainty on the signal yield prediction accounting for these JES uncertainties.
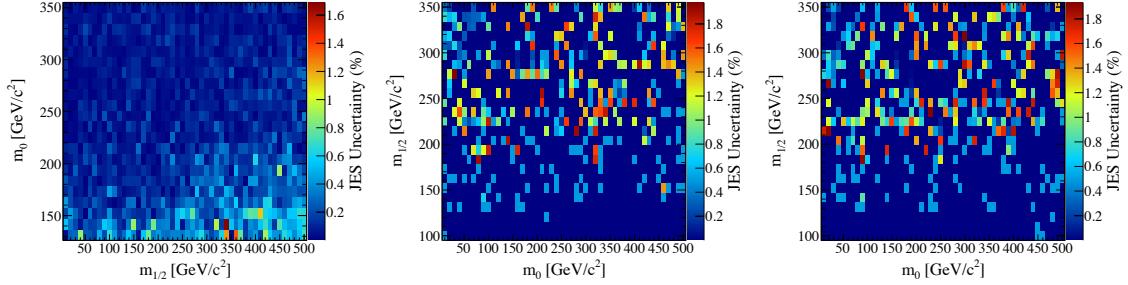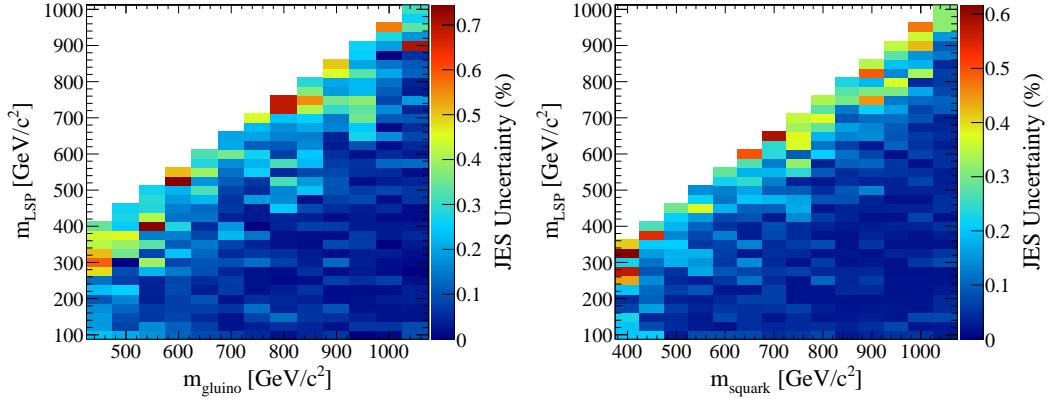
B.12.1 CMSSM

Figure B.38: Systematic uncertainty on the expected signal yield in the signal region due to JES uncertainties for CMSSM models ($\tan \beta = 3$, $A_0 = 0$, sgn$\mu = +$), as a function of the parameters m_0 and $m_{1/2}$. Results are shown for the (Left) HAD Box ($R > 0.5$, $M_R > 500$ GeV), (Center) ELE Box ($R > 0.45$, $M_R > 500$ GeV) and (Right) MU Box ($R > 0.45$, $M_R > 500$ GeV) selections.

B.12.2 Simplified Models

The resulting JES systematic uncertainty on the signal yield, for each simplified model point, are shown in figure C.49. A flat systematic uncertainty of 1% is assigned to account for this effect.

Figure B.39: Systematic uncertainty on the expected signal yield in the signal region due to JES uncertainties in the HAD Box, as a function of LSP and gluino/squark mass. Results are shown for the (Left) di-gluino production simplified model and (Right) di-squark production simplified model.

B.13 Potential Signal Contamination in M_R Control Regions

.

B.13.1 CMSSM

In the background prediction procedure, low and intermediate M_R sidebands are used as control regions. These control regions are used to measure background normalizations and shapes. Here, we quantify the possibility for signal events to contaminate these measurements through their presence in these control regions.

In the lepton boxes, an the region 125 GeV $< M_R <$ 175 GeV is used to measure

the normalization of the $W + jets$ background, and this normalization is ultimately propagated to the other SM backgrouns. This region is dominated by $W + jets$ production, such that the contamination from other SM backgrounds ($Z(\ell\ell)$+jets and $t\bar{t}$+jets) are small. Similarly, any potential contribution from CMSSM signal events is completely negligible, for all of the CMSSM parameter space examined.

The control region that can potentially be polluted with signal events is the intermediate M_R sideband (200 GeV$< M_R < 400$ GeV) which is used to determine the normalization of the W+jets 2nd component, in each of the three boxes. In order to evaluate the possibility of signal contamination in this control region, we calculate the expected background yield, for each of the 3 boxes and each point in the CMSSM parameter space considered.

The relevant metric for assessing the impact of this potential signal contamination is not the absolute expected signal yield in these control regions, but rather the *fractional* yield, relative to the expected SM contribution in the same region. In each of the three boxes, we calculate the expected SM contribution in these control regions, based on our 35 pb^{-1} background predictions. We find that we expect {287, 118, 95} events in this control region for the HAD, ELE and MU Boxes, respectively. The expected signal yields for these three control regions, relative to the expected SM contribution, are shown in figure B.40 for each of the three boxes and each point in the CMSSM parameter space considered. The only cases where there is potential for significant signal contamination is in the CMSSM parameter space that is already ruled out. We would observe this signal before it the background prediction would go off.

We observe that the expected fractional signal contamination in the regions of the CMSSM which this analysis is sensitive to is negligible, relative to the corresponding uncertainties on the background yields in this region, and as a result does not affect the background prediction in the high M_R region of the HAD, ELE and MU Boxes. (up to 5% in the HAD box and a couple percent in the lepton boxes).
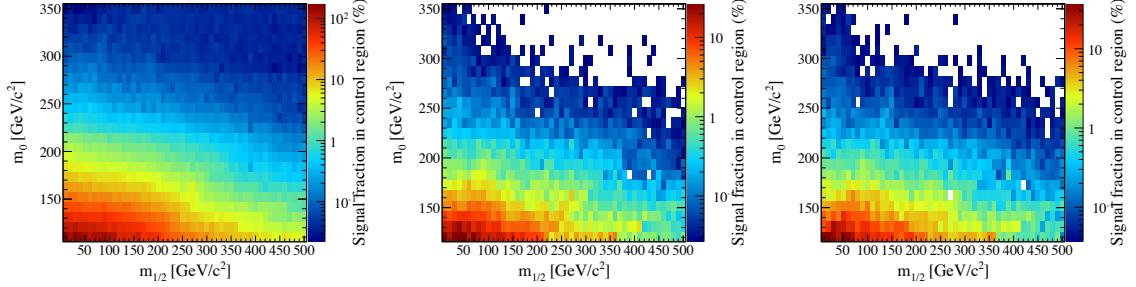
Figure B.40: Expected fractional signal contribution to M_R control regions for CMSSM models ($\tan\beta = 3$, $A_0 = 0$, $\text{sgn}\mu = +$), as a function of the parameters m_0 and $m_{1/2}$. Results are shown for the (Left) HAD Box, (Center) ELE Box and (Right) MU Box selections.

B.13.2 Simplified Models

.

The relevant discussion here is the same as for the CMSSM signal contamination and the conclusion is the same. The expected fractional signal contamination in the HAD box M_R sideband is shown in figure B.41 for the di-gluino and di-squark simplified signal models, and is negligible for the models we are sensitive to.
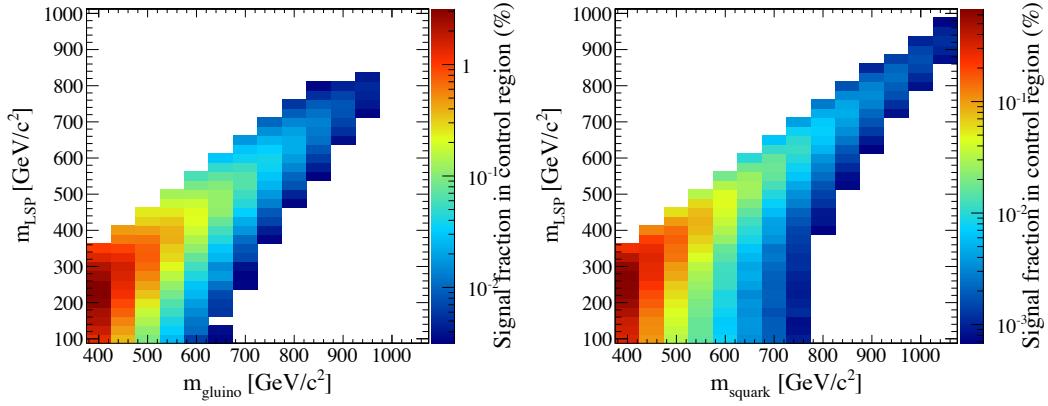
Figure B.41: Expected fractional signal contribution to the M_R control region in the HAD Box, as a function of LSP and gluino/squark mass for simplified models. Results are shown for the (Left) di-gluino production simplified model and (Right) di-squark production simplified model.

Appendix C

Chapter 10 Appendix: Searching Through Razor Space

C.1   Event Samples

C.1.1   Data Samples

The data sample consists of the events collected by the CMS experiment at $\sqrt{s}$=7 TeV during the Run2011A and Run2011B CMS data-taking period. We consider final states including 2μ, 1μ1e, 2e, 1μ, 1e and hadronic. The analysis uses three different Primary Datasets (PDs), HT, MuHad, and ElectronHad, in which the events firing the dedicated razor triggers are included (see section 10.4.1). A summary of the datasets used in this analysis is given in Tab. C.1.

In addition to these PDs, control measurements are made on data samples in May10ReReco dataset corresponding to the Jet, SingleElectron, SingleMuon, DoubleElectron, DoubleMuon and MuEle PDs, which contain events triggered by the NR11 dataset control triggers described in section 10.4.1.

C.1.2   Simulated Signal and SM Background Samples

Simulated event samples are used in this analysis to characterize the agreement between data and expectations from simulation (section C.12) and to test the closure of our background modeling strategy (section C.9). All simulated samples used in this work are processed with the full CMS detector simulation and standard `CMSSW`

Run Range	Dataset Name
multi-box razor data samples	
165071 - 168437	/ElectronHad/Run2011A-PromptReco-v4/AOD
165071 - 168437	/MuHad/Run2011A-PromptReco-v4/AOD
165071 - 168437	/HT/Run2011A-PromptReco-v4/AOD
170053 - 172619	/ElectronHad/Run2011A-05Aug2011-v1/AOD
170053 - 172619	/MuHad/Run2011A-05Aug2011-v1/AO
170053 - 172619	/HT/Run2011A-05Aug2011-v1/AOD
172620 - 175770	/ElectronHad/Run2011A-PromptReco-v6/AOD
172620 - 175770	/MuHad/Run2011A-PromptReco-v6/AOD
172620 - 175770	/HT/Run2011A-PromptReco-v6/AOD
> 175832	/ElectronHad/Run2011B-PromptReco-v1/AOD
> 175832	/MuHad/Run2011B-PromptReco-v1/AOD
> 175832	/HT/Run2011B-PromptReco-v1/AOD

Table C.1: Summary of 2011 datasets.

reconstruction chain.

The W($\rightarrow \ell \nu$) + n-jets events (ℓ = e, μ, τ), Z($\rightarrow \ell \ell$) + n-jets events and γ + n-jets events are produced with Monte Carlo simulation, using the MadGraph v4.22 event generator, based on a leading-order calculation of the matrix element (ME). ME calculation is performed for final states with at most four primary partons, requiring that the parton p_T exceeds 10 GeV/c. PYTHIA [213] is used for parton showering, hadronization and the underlying event description. Parton shower matching is applied to avoid double counting of emissions in overlapping phase space regions. The MLM [235] matching algorithm with k_T. The lepton clustering is used with matching threshold 15 GeV/c. pair invariant mass is required to be $m_{\ell\ell} > 50$ GeV/c^2 at the generator level. The CTEQ6L1 [236] parton distribution functions are used. The $t\bar{t}$ + jets and single top (s-channel, t-channel and tW) backgrounds are generated with MadGraph interfaced with PYTHIA with the associated parton $p_T > 20$ GeV/c and matching threshold 30 GeV/c. Samples are considered from the *Summer11* and *Spring11* Monte Carlo productions, as summarized in Tab. C.2.

To generate simulated samples for SUSY signal models the mass spectrum is first calculated with SOFTSUSY [215] and the decays with SUSYHIT [216]. PYTHIA is then used with the SLHA interface [217] to generate the events. The generator level cross

With Pileup: Processed dataset name is always /Spring11-PU_S1_START311_V1G1-v*/AODSIM		
Dataset Description	Generator Details	cross section (pb)
$qq \rightarrow WW$	madgraph-tauola	43.0
$gg \rightarrow WW \rightarrow 2l2\nu$	pythia6	0.153
$t\bar{t}$	tauola	157.5
single top s-channel	madgraph	1.4
single top t-channel	madgraph	20.9
tW	madgraph	10.6
Z[20-inf] $\rightarrow ee$	powheg-pythia	1666.0
Z[20-inf] $\rightarrow \mu\mu$	powheg-pythia	1666.0
Z[20-inf] $\rightarrow \tau\tau$	powheg-pythia-tauola	1666.0
Z[10-20] $\rightarrow ee$	powheg-pythia	3892.9
Z[10-20] $\rightarrow \mu\mu$	powheg-pythia	3892.9
Z[10-20] $\rightarrow \tau\tau$	powheg-pythia-tauola	3892.9
W/Z+γ	madgraph	165.0
W $\rightarrow \ell\nu$ + jets	madgraph-tauola	31314.0
Z[50-inf] + jets	madgraph-tauola	3048
WZ	pythia6-tauola	18.2
ZZ	pythia6-tauola	5.9
QCD di-jets, $50 < H_T < 100$ GeV	madgraph	$30 \cdot 10^6$
QCD di-jets, $100 < H_T < 250$ GeV	madgraph	$7 \cdot 10^6$
QCD di-jets, $250 < H_T < 500$ GeV	madgraph	171000
QCD di-jets, $500 < H_T < 1000$ GeV	madgraph	5200
QCD di-jets, $H_T > 1000$ GeV	madgraph	83
Z$\rightarrow \nu\nu$	madgraph	4500

Table C.2: Summary of Monte Carlo datasets.

section and the k-factor for the NLO calculation computed using Prospino 2 [228].

C.2 NR11 Control Sample Measurements

The NR11 data set is used to isolate the different SM contributions to the event yield of each of the final state boxes. Using these control samples, the parameters describing the M_R/R^2 shapes of the backgrounds are measured to be used in the fit region ML fits with the R11 data sample.

During the runs corresponding to the NR11 data set low threshold, inclusive (w.r.t. requirements on hadronic activity) lepton triggers were deployed which allows us to perform measurements in a more expansive low-M_R/low-R^2 region, relative to the R11 ML fit region. For each isolated SM background, a two-dimensional maximum likelihood fit is performed modeling the events as two instances of the function

$F_i(M_R, R^2)$ (equation (10.15)), corresponding to the first and second component of that background and with the total background pdf given by

$$F_i^{TOT}(M_R, R^2) = N_i \left[F_i^{1st}(M_R, R^2) + f_i F_i^{2nd}(M_R, R^2) \right] \ , \qquad (C.1)$$

where the parameter f_i describes the relative normalization of the two components and the functions F^{1st} and F^{2nd} each have independent parameters k, M_R^0 and R_0^2.

The selection used to isolate each of the nonnegligible backgrounds ($t\bar{t}$+jets, $W(\ell\nu)$+jets and $Z(\ell\ell)$+jets) in each of the boxes is described below, along with the two-dimensional fits which measure the shape parameters.

C.2.1   W($\ell\nu$)+jets in the ELE Box

A control sample for $W(\ell\nu)$+jets events is defined by combining the ELE Box requirements with a veto on the presence of b-tagged jets. The latter requirement reduces the $t\bar{t}$+jets contribution to the ELE Box, resulting in a high purity $W(\ell\nu)$+jets sample.

Events are selected from the DoubleElectron PD of the May10 ReReco data sample, requiring the OR of `HLT_Ele8` and `HLT_Ele17_CaloId_CaloIsoVL` (both prescaled triggers). This trigger is fully efficient for events satisfying the offline ELE Box electron requirements.
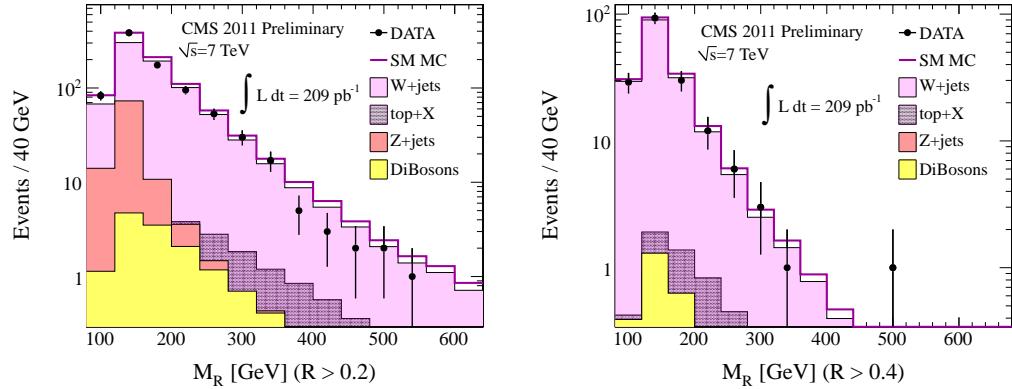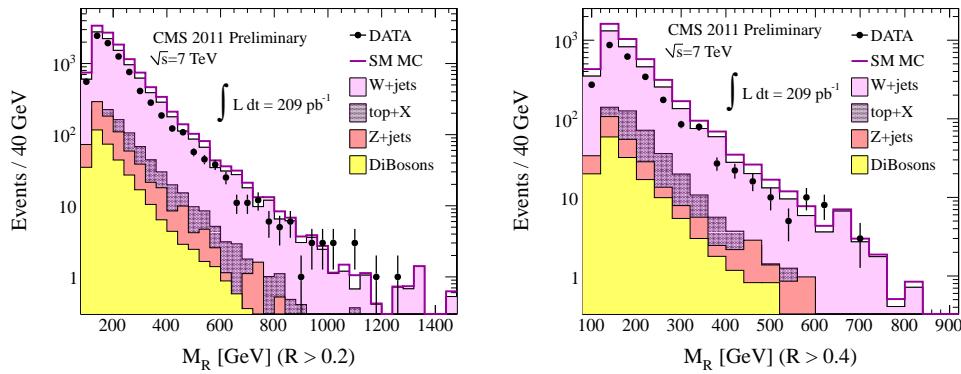

Figure C.1: M_R distribution for events selected in the ELE Box requiring that the event has no b-tagged jets, for an R cut of (Left) $R > 0.2$ (Right) $R > 0.4$.

A comparison of the M_R distribution for data and simulated background events satisfying this control sample selection is shown in figure C.1. The simulated backgrounds are normalized relatively using N^{vertex} event re-weighting for pileup and theoretical cross sections. In order to account for data being selected with a prescaled trigger, the simulated backgrounds are scaled so that the total background prediction agrees with the data in the largest bin.

Box	Sample		
	$W(\ell\nu)$+jets 0 b-tagged jets	$t\bar{t}$+jets 0 b-tagged jets	$Z(\ell\ell)$+jets 0 b-tagged jets
ELE	5825 (91.3%)	428 (6.7%)	112 (1.8%)
MU	4002 (84.8%)	351 (7.4%)	191 (4.1%)

Table C.3: Yields normalized to 1 fb^{-1} with 0 b-tag requirement and $M_R > 300$ GeV, $R > 0.3$ ($R^2 > 0.09$)

Yields estimated from simulated events for the control selection are listed in Table C.3, indicating that more than 90% of this sample should be comprised of $W(e\nu)$+jets events.
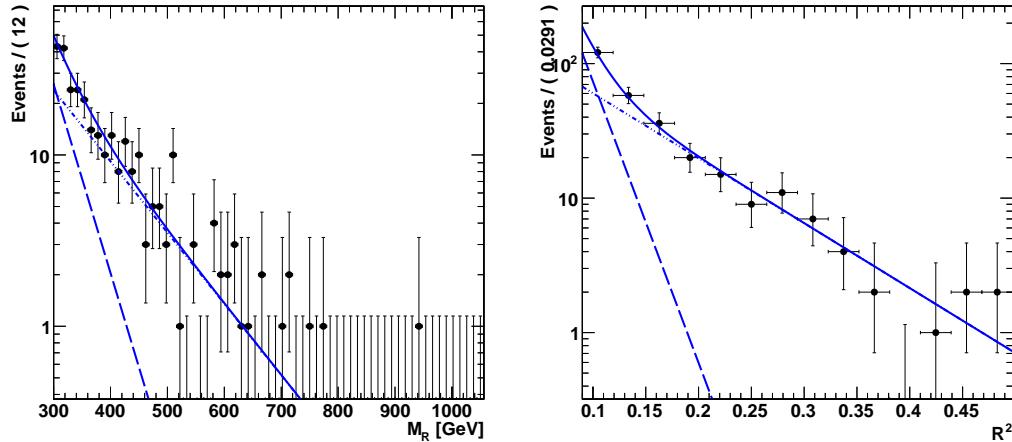
Figure C.2: 1D projections of a 2D ML fit in M_R (Left) and R^2 (Right) for W+jets data control samples in the ELE box. The sample is described by two F_i components.

C.2.2   W($\mu\nu$)+jets in the MU Box

A control sample for $W(\mu\nu)$+jets events in the MU box is defined by combining the MU box requirements with a veto on the presence of b-tagged jets, which significantly reduces $t\bar{t}$+jets contamination. Events are selected from the SingleMu PD for May10 ReReco data, requiring `HLT_IsoMu17` to fire. The muon p_T requirement is increased from 15 to 20 GeV/c in order to be efficient with respect to this trigger and to make a consistent requirement on simulated events for comparison. Studies of simulated event shapes as a function of lepton p_T requirement indicate that the shape parameters measured from this control sample are insensitive to this change.

Figure C.3: M_R distribution for events selected in the MU Box ($p_T^\mu > 20$ GeV/c) requiring that the event has no b-tagged jets, for an R cut of (Left) $R > 0.2$ (Right) $R > 0.4$.

A comparison of the M_R distribution for data and simulated background events satisfying this control sample selection is shown in figure C.3. Yields estimated from simulated events for the control selection are listed also in Table C.3, quantifying the purity of this control sample. The two-dimensional ML fit to this control sample is shown in figure C.2 and the measured shape parameters are listed in Tab. C.6.

C.2.3   Z($\ell\ell$)+jets in the ELE-ELE and MU-MU Boxes

Isolated samples of $Z(\ell\ell)$+jets events are selected by combining the ELE-ELE and MU-MU box requirements with a b-tagged jet veto, respectively. The later require-

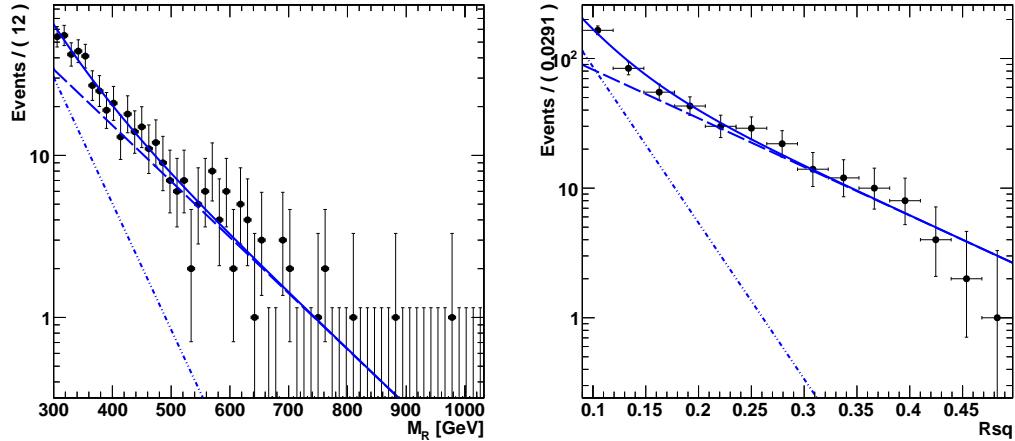Figure C.4: 1D projections of a 2D ML fit in M_R (Left) and R^2 (Right) for W+jets data control samples in the MU box. The sample is described by two F_i components.

ment reduces the $t\bar{t}$+jets background contribution while maintaining high efficiency for $Z(\ell\ell)$+jets events. Similarly, the invariant mass of the di-lepton system is required to exceed 60 GeV, further reducing non-Z contributions.
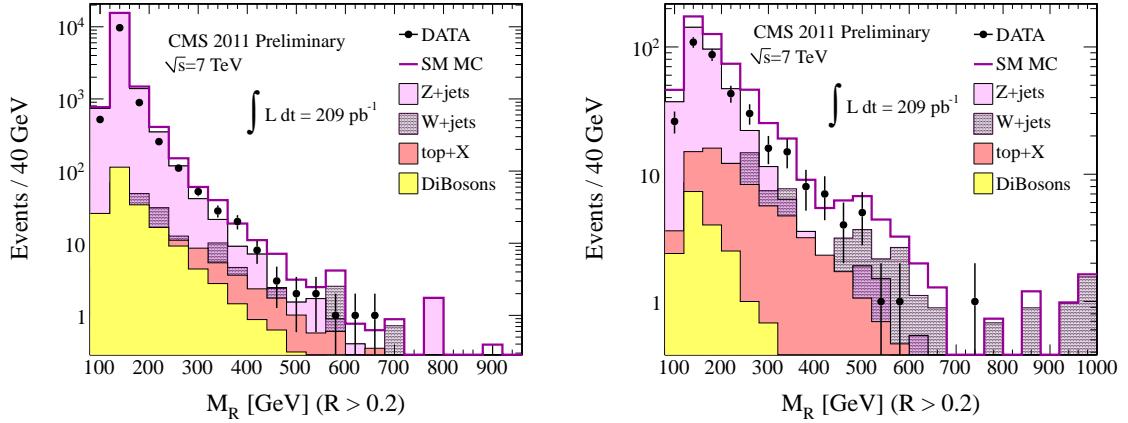

Figure C.5: M_R distribution for events selected in the ELE-ELE box (Left) and MU-MU box (Right) requiring that the event has no b-tagged jets and satisfies a cut of $R > 0.2$.

For the ELE-ELE box, Events are selected from the DoubleElectron PD for May10 ReReco data requiring `HLT_Ele17_CaloIdL_CaloIsoVL_Ele8_CaloIdL_CaloIsoVL` fires. MU-MU box events are selected by requiring the OR of `HLT_DoubleMu7` and `HLT_Mu13_Mu8`

in the DoubleMu PD.

Box	Sample		
	$t\bar{t}$+jets 0btag	$W(\ell\nu)$+jets 0btag	$Z(\ell\ell)$+jets 0btag
MU-ELE	231 (54.0%)	103 (24.0%)	94 (21.9%)
MU-MU	169 (18.0%)	300 (32.0%)	469 (50.0%)
ELE-ELE	204 (6.4%)	344 (10.7%)	2651 (82.9%)

Table C.4: Yields are normalized to 1 fb^{-1}. Requirements include $M_{\ell\ell} > 60$ GeV cut, 0 b-tagged jets, $M_R > 200$ GeV and $R > 0.2$ ($R^2 > 0.04$)

A comparison of the M_R distribution for data and simulated background events satisfying these control sample selections are shown in figure C.5. Yields estimated from simulated events for the control selections are listed in Table C.4.

C.2.4 $t\bar{t}$+jets NR11 Control Samples

An isolated sample of $t\bar{t}$+jets events can be selected in each of the final state boxes by requiring the presence of at least one b-tagged jet. Tab. C.5 summarizes the expected event composition in the final state boxes considered for $t\bar{t}$+jets control samples. An ELE box sample is not considered since there is not an acceptable trigger available in the NR11 menu with a high enough event yield.

Box	Sample		
	$t\bar{t}$+jets ≥ 1 btag	$W(\ell\nu)$+jets ≥ 1 btag	$Z(\ell\ell)$+jets ≥ 1 btag
MU-ELE	474 (95.6%)	10 (2.1%)	12 (2.3%)
MU-MU	915 (77.6%)	94 (8.0%)	29 (2.5%)
ELE-ELE	468 (81.4%)	29 (5.1%)	77 (13.4%)
MU	4019 (61.1%)	1487 (22.6%)	65 (1.0%)

Table C.5: Yields normalized to 1 fb^{-1} with ≥ 1 b-tag requirement, $M_R > 200$ GeV, and $R > 0.2$ ($R^2 > 0.04$)

A comparison of the M_R distribution for data and simulated background events satisfying these control sample selections are shown in figure C.6 for the ELE-MU, MU-MU, ELE-ELE and MU boxes, respectively. The SM background simulated events are normalized using N^{vertex} event re-weighting for pileup and theoretical cross

sections without trigger requirements. The two-dimensional ML fits to these samples are shown in figure C.7, with shape parameters from those fits summarized in Tab. C.8.

C.2.5 Summary of Shape Parameters from NR11 Measurements

The background shape parameters measured from the NR11 data set are summarized for $W(\ell\nu)$+jets, $Z(\ell\ell)$+jets and $t\bar{t}$+jets events in Tab. C.6, C.7 and C.8, respectively.

Sample	Box	First Component			
		M_R^0	R_0^2	k	
W+jets	MU	-51 ± 40	-0.137 ± 0.029	0.079 ± 0.014	
W+jets	ELE	-21 ± 22	-0.079 ± 0.015	0.150 ± 0.020	
Sample	Box	Second Component			
		M_R^0	R_0^2	k	f
W+jets	MU	-32 ± 29	-0.216 ± 0.033	0.026 ± 0.003	0.72 ± 0.04
W+jets	ELE	-37 ± 36	-0.200 ± 0.040	0.033 ± 0.004	0.71 ± 0.04

Table C.6: Fit results for a set of $W + jets$ data control samples (no b-tagged jets).

Sample	Box	First Component			
		M_R^0	R_0^2	k	
Z+jets	MU-MU	116 ± 19	0.000 ± 0.001	0.400 ± 0.030	
Z+jets	ELE-ELE	-324 ± 427	-0.15 ± 0.16	0.11 ± 0.09	
Sample	Box	Second Component			
		M_R^0	R_0^2	k	f
Z+jets	MU-MU	-164 ± 56	-0.281 ± 0.004	0.021 ± 0.003	0.94 ± 0.06
Z+jets	ELE-ELE	-140 ± 151	-0.15 ± 0.09	0.05 ± 0.02	1.00 ± 0.11

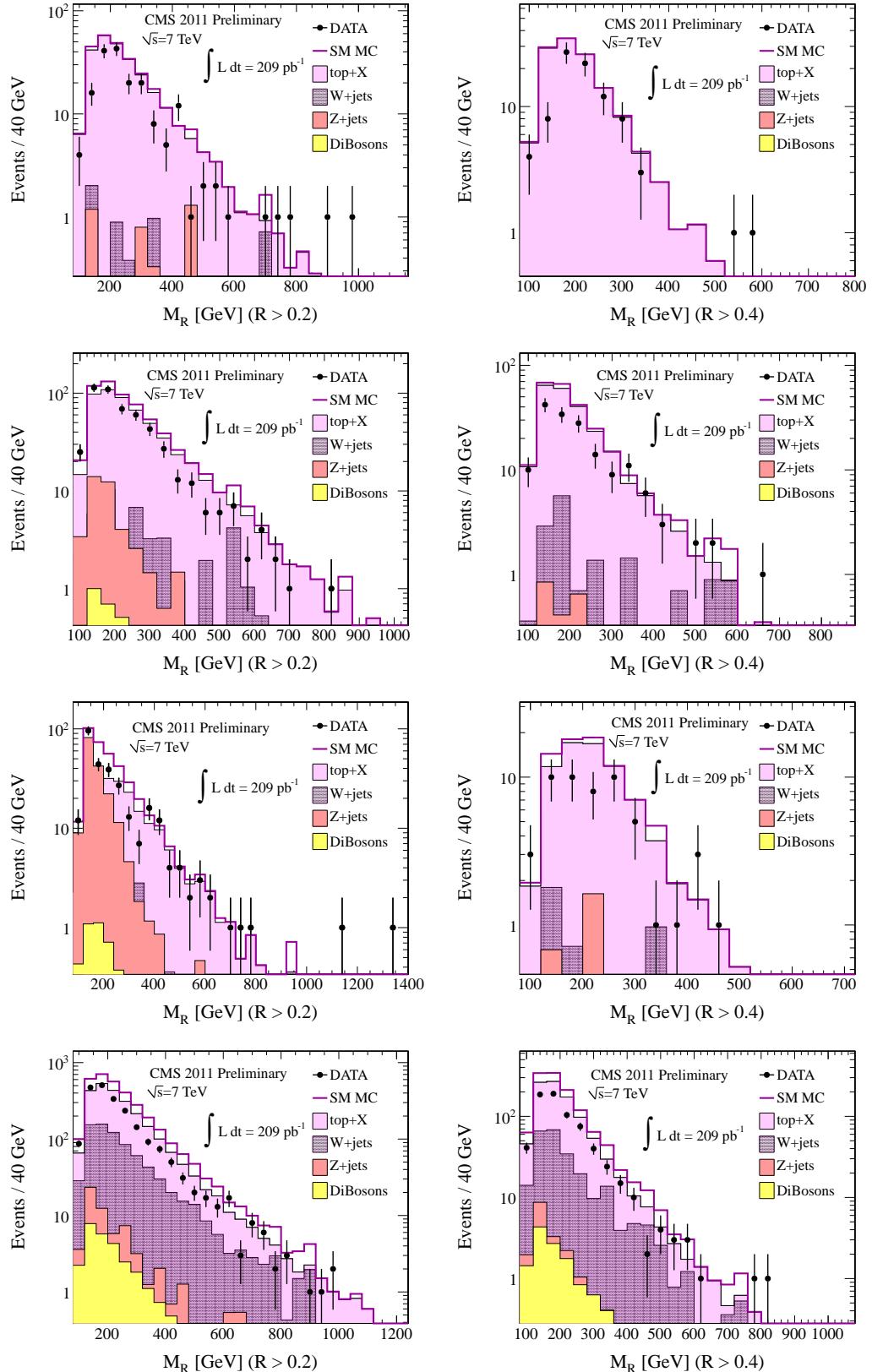Table C.7: Fit results for a set of Z+jets data control samples (no b-tagged jets and $m_{ll} > 60$).

Figure C.6: M_R distribution for events selected requiring ≥ 1 b-tags, for an R cut of (Left) $R > 0.2$ and (Right) $R > 0.4$. (First row) ELE-MU box. (Second row) MU-MU box. (Third row) ELE-ELE box. (Fourth row) MU box.

Sample	Box	First Component			
		M_R^0	R_0^2	k	
$t\bar{t}$+jets	MU	-30 ± 44	-0.137 ± 0.035	0.063 ± 0.013	
$t\bar{t}$+jets	ELE	-44 ± 43	-0.080 ± 0.021	0.073 ± 0.013	
$t\bar{t}$+jets	MU-MU	57 ± 72	-0.055 ± 0.047	0.130 ± 0.070	
$t\bar{t}$+jets	ELE-ELE	63 ± 47	-0.050 ± 0.020	0.140 ± 0.030	
$t\bar{t}$+jets	MU-ELE	63 ± 47	-0.140 ± 0.060	0.060 ± 0.020	
Sample	Box	Second Component			
		M_R^0	R_0^2	k	f
$t\bar{t}$+jets	MU	-194 ± 190	-0.310 ± 0.180	0.017 ± 0.004	1.00 ± 0.01
$t\bar{t}$+jets	ELE	-40 ± 40	-0.08 ± 0.03	0.017 ± 0.004	0.38 ± 0.08
$t\bar{t}$+jets	MU-MU	48 ± 67	-0.06 ± 0.05	0.018 ± 0.003	0.74 ± 0.10
$t\bar{t}$+jets	ELE-ELE	-55 ± 40	-0.04 ± 0.02	0.022 ± 0.004	0.76 ± 0.15
$t\bar{t}$+jets	MU-ELE	66 ± 50	-0.14 ± 0.05	0.012 ± 0.005	0.51 ± 0.22

Table C.8: Fit results for a set of $t\bar{t}$+jets data control samples (≥ 1 b-tagged jets)
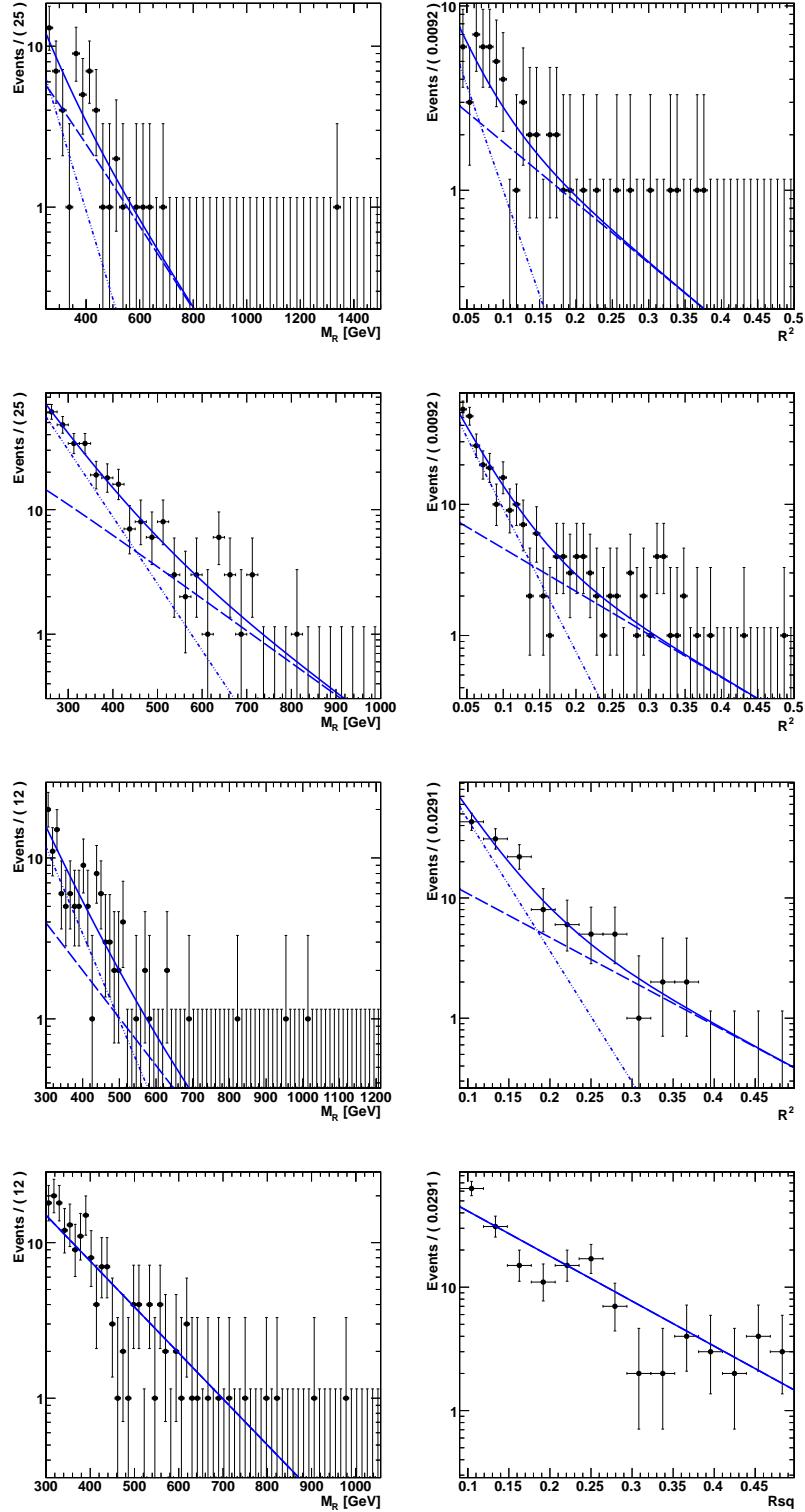
Figure C.7: 1D projections of a 2D ML fit in M_R (Left) and R^2 (Right) for $t\bar{t}$ NR11 control samples. (First row) ELE-MU box. (Second row) MU-MU box. (Third row) ELE-ELE box. (Fourth row) MU box.

C.3   Study of Residual QCD Background

Without heavy bosons decaying to neutrinos, QCD multijet events have no intrinsic source of high momentum weakly interacting particles. This results in these events having naturally small values of R and an M_R distribution which falls more quickly than other SM backgrounds with increasing values of R.

The di-lepton boxes have negligible contributions from QCD multijet events due to both suppression from M_R/R^2 minimum requirements and the need to have two reconstructed and high quality leptons. For the other boxes, it is expected for some multijet events to appear at low-M_R/low-R^2. In order to estimate this contribution test are performed for each of the boxes with two parts:

- Using control samples from data, measure the shape of QCD multijet backgrounds

- Repeat the 2D background model fits including an additional multijet background component, and test whether its shape is consistent with event yields observed in data

At low M_R and R^2 the vast majority of events correspond to QCD multijet background because the large relative production cross section. In order to isolate a QCD multijet sample events are selected in this phase-space (which is not accessible using the higher-threshold razor triggers) by applying requirements on prescaled, low-threshold jet and lepton triggers.

From the NR11 data sample, which is independent of the sample used in the final background model fits, multijet control samples are defined for each of the HAD, MU and ELE boxes.

HAD box multijet sample

The QCD multijet control sample for the HAD box is defined by combining the HAD box selection with prescaled jet trigger requirements, with thresholds low enough to give an unbiased M_R distributions to low values, where the expected contribution

from non-multijet events is negligible. We consider two samples: a low-statistics sample selecting events by requiring that both `HLT_Jet30` and `HLT_DiJet30` fire and a larger statistics sample (with less available M_R phase-space due to higher thresholds) which requires the OR of `HLT_Jet30`, `HLT_DiJet30`, `HLT_Jet60`, `HLT_DiJet60`, `HLT_Jet80` and `HLT_DiJet80`. For each of these triggers, the number indicates the minimum jet p_T required at the HLT to fire the trigger, with "DiJet" referring to the average of the two hardest jets.

ELE and MU box multijet samples

In order to select a sample of multijet events in lepton final states we alter the ELE and MU box requirements by *inverting* the lepton isolation criteria when identifying those leptons. This selects a sample composed predominantly of non-isolated leptons coming from hadronic decays within jets which, at low-M_R/low-R, is consistent with a pure QCD multijet sample. Multi-jet ELE and MU events are selected from the DoubleEle and SingleMu PDs, respectively, using the same triggers as for the NR11 measurements described in section C.2.

The 2D ML fits to each of these samples are shown in figure C.8, where each sample is modeled as a single instance of the function $F_j(M_R, R^2)$ (equation (10.15)). The shape parameters measured from these fits are summarized in Tab. C.9.

Sample	M_R^0 [GeV]	R_0^2	k [GeV^{-1}]
Multi-jet HAD box (HLT_DiJetAve30)	3 ± 8	-0.046 ± 0.003	0.269 ± 0.014
Multi-jet HAD box (HLT_DiJetAve80)	6 ± 3	-0.043 ± 0.001	0.271 ± 0.004
Multi-jet ELE Box	-190 ± 90	-0.14 ± 0.05	0.11 ± 0.03
Multi-jet MU Box	-1200 ± 600	-1 ± 1	0.017 ± 0.009

Table C.9: Example fit results for data events selected by the HAD box selection for the QCD control box, for two different choices of the HLT requirement. The data is modeled as a single-component 2D pdf.

Background components with the multijet shapes measured from control samples are added to the background model for fits in the HAD, ELE and MU boxes with the
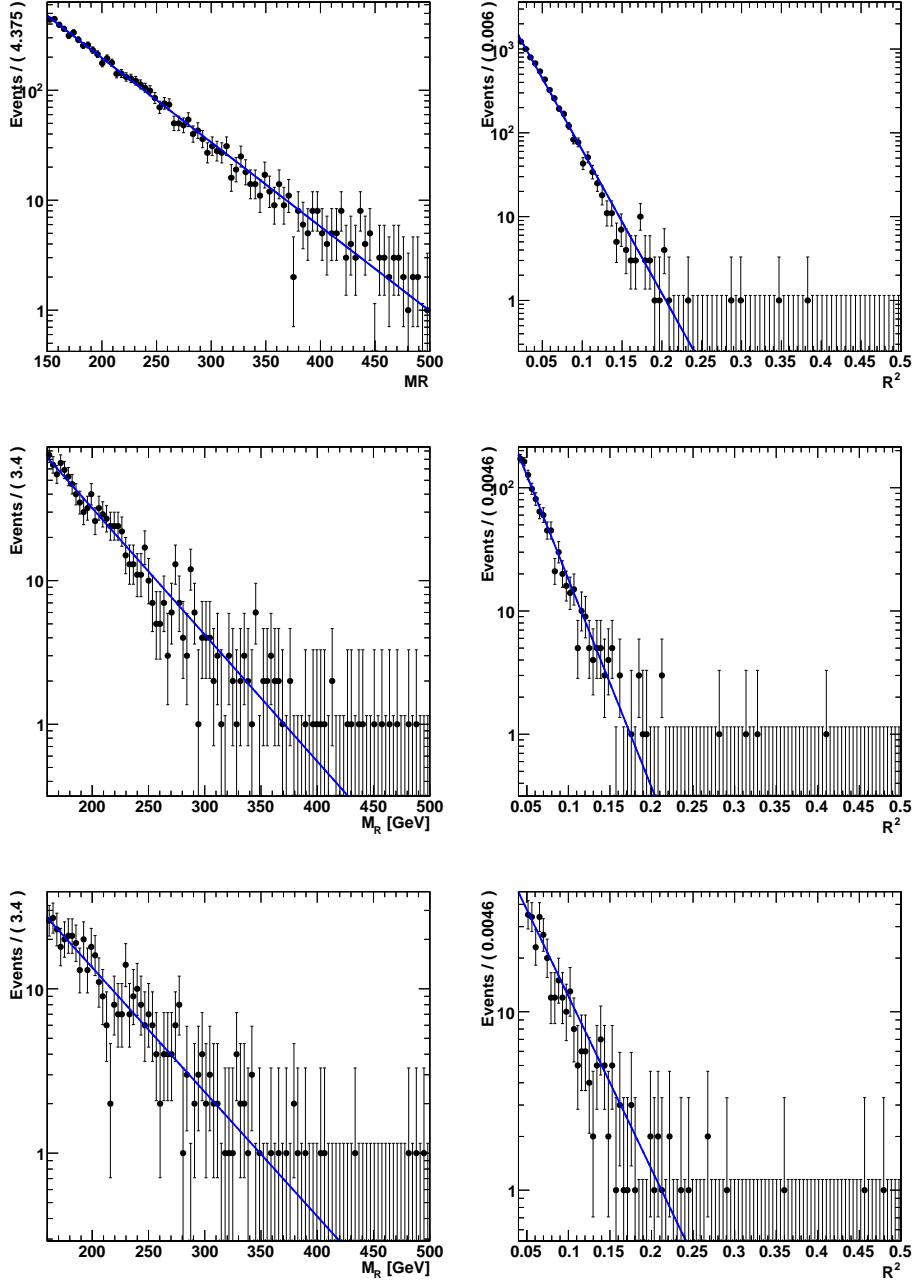
Figure C.8: Projection of the 2D ML fits results on M_R (Left) and R^2 (Right) compared with data samples resulting from multijet control selections. Figures correspond to the HAD (Top), ELE (Center), and MU (Bottom) boxes, with the corresponding trigger requirements.

inclusive data sample. The yield parameter corresponding to the multijet background from each of these fits is given in Tab. C.10. We observe that the fits prefer negligible

multijet yields, indicating that any potential multijet contamination is small and adequately described by the other components of the background model.

Box	N_{QCD}
HAD	-70 ± -40
MU	-10 ± -100
ELE	-10 ± 40

Table C.10: Fit results for the yield of the QCD background in the HAD, MU, ELE boxes

C.4   High-Statistics Simulated $t\bar{t}$ Shape Study

In order to study the robustness of the two dimensional background model we consider a high-statistics sample of simulated $t\bar{t}$+jets events, corresponding to an integrated luminosity of 10 ab^{-1}, about 2000 times larger than the data set collected from 2011 collisions. With this high-statistics sample a number of tests are performed:

- Test the background model to a greater precision than possible with 2011 data sets.

- Test to which extent the second (flatter) component, which is included in the UEC for the background model, is stable across boxes (noting that this fact is not enforced as an assumption in the analysis)

- Quantify the impact of a p_T-dependence of the b-tagging discriminant by comparing background shapes between inclusive and b-tagged subsamples

- Quantify the agreement between simulated events and data for $t\bar{t}$+jets final states

Two-dimensional ML fits are performed on the full 10 ab^{-1} simulated $t\bar{t}$+jets sample in each of the six final state boxes. Each of these fits models the sample with two instances of $F_j(M_R, R^2)$ (Eq. 10.15), a first and second component. The latter is part of the UEC component which combines the contributions of the similar $t\bar{t}$+jets and V+jets second components for the background model used in the analysis. The fits are performed in the full razor plane for both inclusive event samples and the subsample with at least one b-tagged jet.

The inclusive HAD box simulated event sample, with projections of the ML fit result, are shown in figure C.9. We observe that the 2D background model gives an excellent description of the data sample in the razor plane, over many decades of data. This background model agreement, with comparable results for each of the selections considered, indicates that the 2D background model should provide an adequate description of SM backgrounds selected in collision events for samples significantly larger than the 4.7 fb^{-1} sample considered for this analysis.
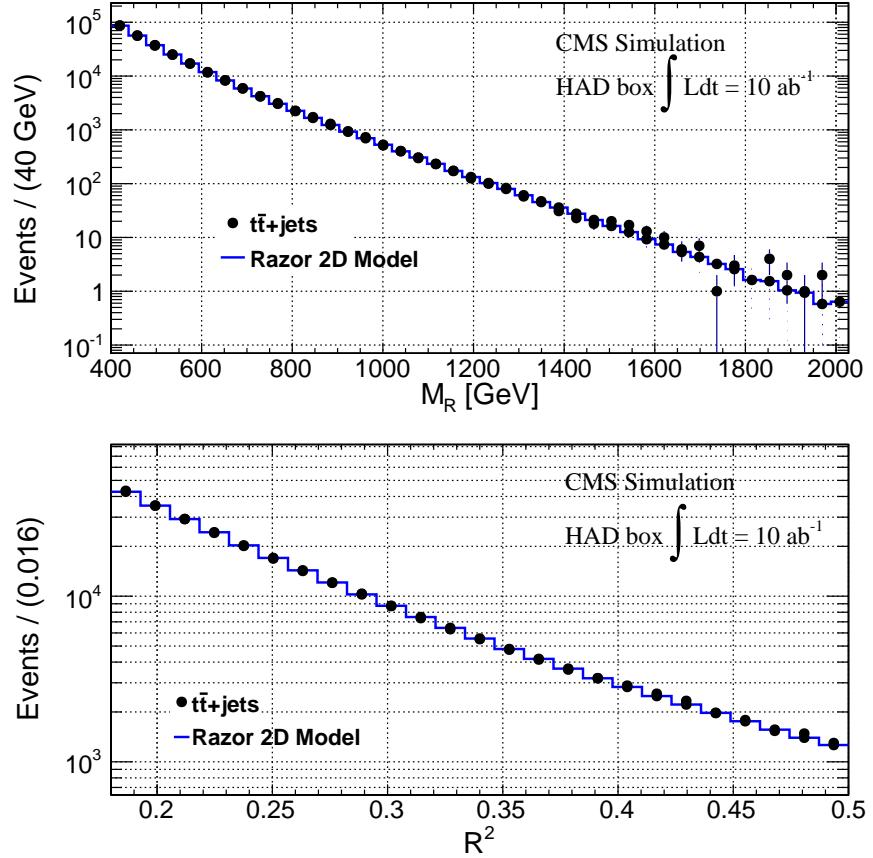
Figure C.9: Distribution of M_R (Top) and R^2 (Bottom) for 10 ab^{-1} of simulated $t\bar{t}$+jets events satisfying the HAD box selection. Blue line indicates the central value of the 2D ML fit to this sample with the background model for $t\bar{t}$+jets, projected onto each variable. Fits are performed in the full razor plane region.

The shape parameters for each of the boxes, as determined by ML fits, are summarized in Tab. C.11. From these results, two conclusions can be drawn. Firstly, a comparison of the shape parameters for a singe box between the inclusive and b-tagged samples indicates that the values are consistent. This implies that the b-tagged jet requirement does not significantly change the shape the razor variables. Secondly, comparing the the shape parameters between different boxes we observe that they are very similar, with the second component consistent throughout the boxes, with small numerical differences resulting from the difference in the treatment of electrons and muons when constructing the variables, particularly for k.

HAD box				
Parameter	1st Component		2nd Component	
	Inclusive	≥ 1 b-tagged	Inclusive	≥ 1 b-tagged
M_R^0	-61 $\pm$ 37	-91 $\pm$ 23	-53 $\pm$ 27	-69 $\pm$ 18
R_0^2	-0.164 $\pm$ 0.028	-0.171 $\pm$ 0.017	-0.231 $\pm$ 0.016	-0.228 $\pm$ 0.012
k	0.0366 $\pm$ 0.0037	0.0349 $\pm$ 0.0020	0.0155 $\pm$ 0.0004	0.0152 $\pm$ 0.0004

MU box				
Parameter	1st Component		2nd Component	
	Inclusive	≥ 1 b-tagged	Inclusive	≥ 1 b-tagged
M_R^0	53 $\pm$ 24	52 $\pm$ 26	-21 $\pm$ 17	8 $\pm$ 18
R_0^2	-0.108 $\pm$ 0.021	-0.102 $\pm$ 0.022	-0.264 $\pm$ 0.016	-0.261 $\pm$ 0.017
k	0.0495 $\pm$ 0.0041	0.0501 $\pm$ 0.0045	0.0179 $\pm$ 0.0006	0.0178 $\pm$ 0.0006

ELE box				
Parameter	1st Component		2nd Component	
	Inclusive	≥ 1 b-tagged	Inclusive	≥ 1 b-tagged
M_R^0	-140 $\pm$ 32	-121 $\pm$ 31	-55 $\pm$ 21	-55 $\pm$ 20
R_0^2	-0.154 $\pm$ 0.020	-0.159 $\pm$ 0.020	-0.220 $\pm$ 0.017	-0.229 $\pm$ 0.015
k	0.0349 $\pm$ 0.0025	0.0360 $\pm$ 0.0025	0.0161 $\pm$ 0.0007	0.0165 $\pm$ 0.0006

ELE-ELE box				
Parameter	1st Component		2nd Component	
	Inclusive	≥ 1 b-tagged	Inclusive	≥ 1 b-tagged
M_R^0	-58 $\pm$ 38	-72 $\pm$ 34	-35 $\pm$ 22	-20 $\pm$ 23
R_0^2	-0.138 $\pm$ 0.026	-0.131 $\pm$ 0.025	-0.239 $\pm$ 0.029	-0.268 $\pm$ 0.030
k	0.0413 $\pm$ 0.0041	0.0404 $\pm$ 0.0037	0.0144 $\pm$ 0.0014	0.0133 $\pm$ 0.0012

ELE-MU box				
Parameter	1st Component		2nd Component	
	Inclusive	≥ 1 b-tagged	Inclusive	≥ 1 b-tagged
M_R^0	-34 $\pm$ 43	-49 $\pm$ 35	-30 $\pm$ 21	-34 $\pm$ 22
R_0^2	-0.108 $\pm$ 0.028	-0.106 $\pm$ 0.028	-0.247 $\pm$ 0.022	-0.246 $\pm$ 0.025
k	0.0516 $\pm$ 0.0063	0.0499 $\pm$ 0.0065	0.0197 $\pm$ 0.0012	0.0193 $\pm$ 0.0014

MU-MU box				
Parameter	1st Component		2nd Component	
	Inclusive	≥ 1 b-tagged	Inclusive	≥ 1 b-tagged
M_R^0	-14 $\pm$ 36	-23 $\pm$ 38	-20 $\pm$ 20	-29 $\pm$ 20
R_0^2	-0.110 $\pm$ 0.026	-0.117 $\pm$ 0.027	-0.256 $\pm$ 0.019	-0.243 $\pm$ 0.019
k	0.0429 $\pm$ 0.0046	0.0434 $\pm$ 0.0047	0.0164 $\pm$ 0.0007	0.0166 $\pm$ 0.0008

Table C.11: Comparison of fit output for the the inclusive and b-tagged simulated event samples for each of the boxes. Fits are performed on the full R^2 vs. M_R plane after the box requirements are applied.

The shape parameters of the HAD box measured from the b-tagged simulated sample are compared with the shapes measured from b-tagged jet data in Tab. C.12. The sample from data has contamination from W+jets events, which generally has a different first component shape than $t\bar{t}$+jets. The second components are modeled with one UEC, which is in agreement with the shape derived from the simulated event sample.

Parameter	1st Component	
	≥ 1 b-tagged data	≥ 1 b-tagged simulation
M_R^0	-89 $\pm$ 146	-91 $\pm$ 23
R_0^2	-0.097 $\pm$ 0.083	-0.171 $\pm$ 0.017
k	0.050 $\pm$ 0.0016	0.0349 $\pm$ 0.0020
Parameter	2nd Component	
	≥ 1 b-tagged data	≥ 1 b-tagged simulation
M_R^0	-35 $\pm$ 22	-69 $\pm$ 18
R_0^2	-0.239 $\pm$ 0.029	-0.228 $\pm$ 0.012
k	0.0144 $\pm$ 0.0014	0.0152 $\pm$ 0.0004

Table C.12: Comparison of fit parameters for the the b-tagged event sample in the HAD box for data and simulated $t\bar{t}$+jets events. Fits for simulated events are performed on the full R^2 vs. M_R plane after the box requirements are applied, while fits to data are constrained to the fit region.

C.5 Evaluation of M_R Dependence on Lepton Identification

In order to perform the background extrapolation described in section 10.6 we must ensure that the shapes of the razor variables for SM backgrounds are not significantly biased by our lepton identification. To estimate the size of any potential bias, we consider simulated event samples of $t\bar{t}$+jets and W+jets with electrons and muons from W decays. We define a truth level lepton identification, where we require that the electrons and muons have momentum and direction within the phase-space acceptance cuts of our offline reconstruction algorithms. By comparing the M_R shapes of these truth selected events with the subset of events that are identified by our offline lepton reconstruction algorithms, we can isolate any potential bias of the M_R distribution due solely to lepton identification. The M_R distributions for these simulated events, for different identification requirements, are shown in figures C.10 and C.11 for muons and electrons, respectively.
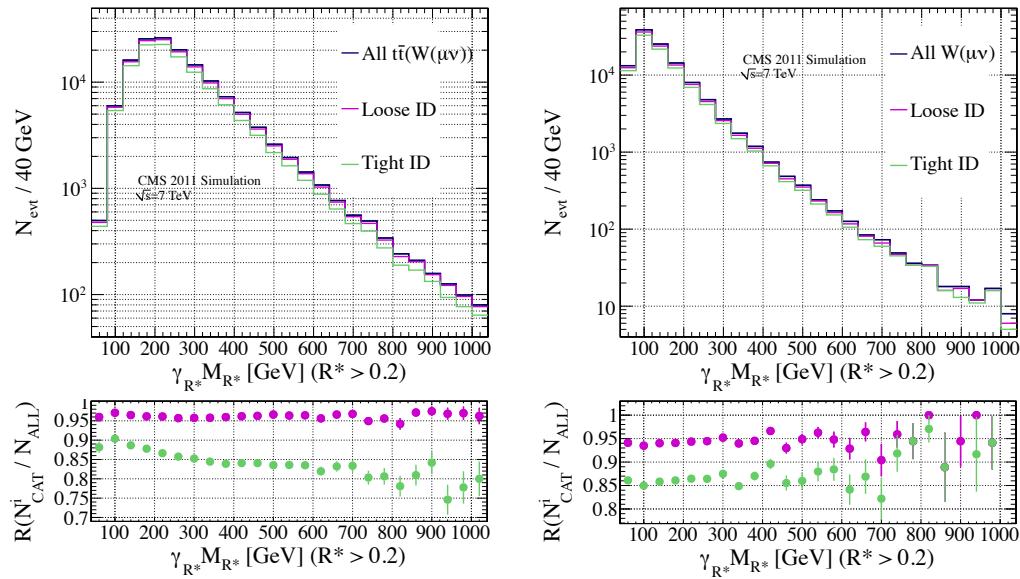
Figure C.10: M_R distribution for simulated events with different muon identification requirements applied. (Left) $t\bar{t}$+ jets events with truth level muon phase-space requirements. (Right) $W(\mu\nu)$+jets events. Lower plots show ratio of event yield for different identifications with respect to total truth sample.
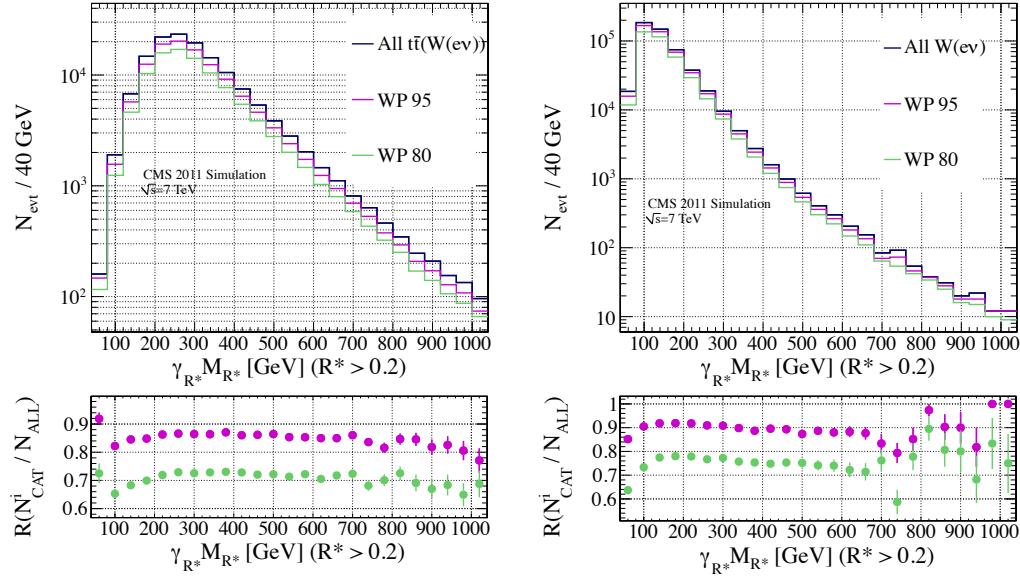
Figure C.11: M_R distribution for simulated events with different electron identification requirements applied. (Left) $t\bar{t}+$ jets events with truth level electron phase-space requirements. (Right) $W(e\nu)+$jets events.

The ratios of shapes in these figures indicate that there is some dependence of the shape of the M_R distribution on lepton identification requirements. We observe that this dependence is strongest at low M_R (below the range considered for this analysis) and that in the relevant M_R range the bias is a few percent over an interval where the background yields decrease by more than two orders of magnitude. We conclude that any potential bias due to lepton identification is negligible.

C.6 Pile-up Dependence of Kinematic Observables

The instantaneous luminosity of LHC collisions within CMS varied during 2011 data-taking, meaning that the number of interactions in each bunch crossing could range from a few to tens. In order to ensure that the shapes of razor variables M_R and R^2 for SM backgrounds and potential signals are stable throughout this period, we evaluate the dependence of the M_R and R^2 distributions on the observed number of vertices, both on data and simulated events. The ratios of the M_R (R^2) distributions for different number of vertices are shown in figure C.12 (C.13).

As demonstrated by these figures, the shape of the kinematic variables is stable for different running conditions, once the pile-up corrections to the jets and lepton isolation variables are applied. Analogous figures for the subset of events with at least one b-tagged jet are shown in figures C.14 and C.15. The stability of the M_R and R^2 shapes despite changing numbers of interactions, with and without a b-tagged jet requirement, indicates that the b-tagging efficiency is also sufficiently stable over changing run conditions for the purposes of this analysis.

It is also necessary to ensure that our electron identification algorithms are stable under changing run conditions. The electron selection efficiency as a function of the number of reconstructed primary vertices, integrated in p_T and η for barrel-only, endcap-only, and all the electrons is shown in figure C.16 for $Z(ee)$ events selected in 2011 data. The efficiency is determined using a tag-and-probe technique (see section 6.1). In these figures we observe that PU energy subtraction applied to electron isolation criteria reduces the dependence of the selection efficiency on the PU conditions to within a few %. The same conclusions applies to the case of muons, for which the isolation variables and the PU subtraction being computed and applied in the same way.
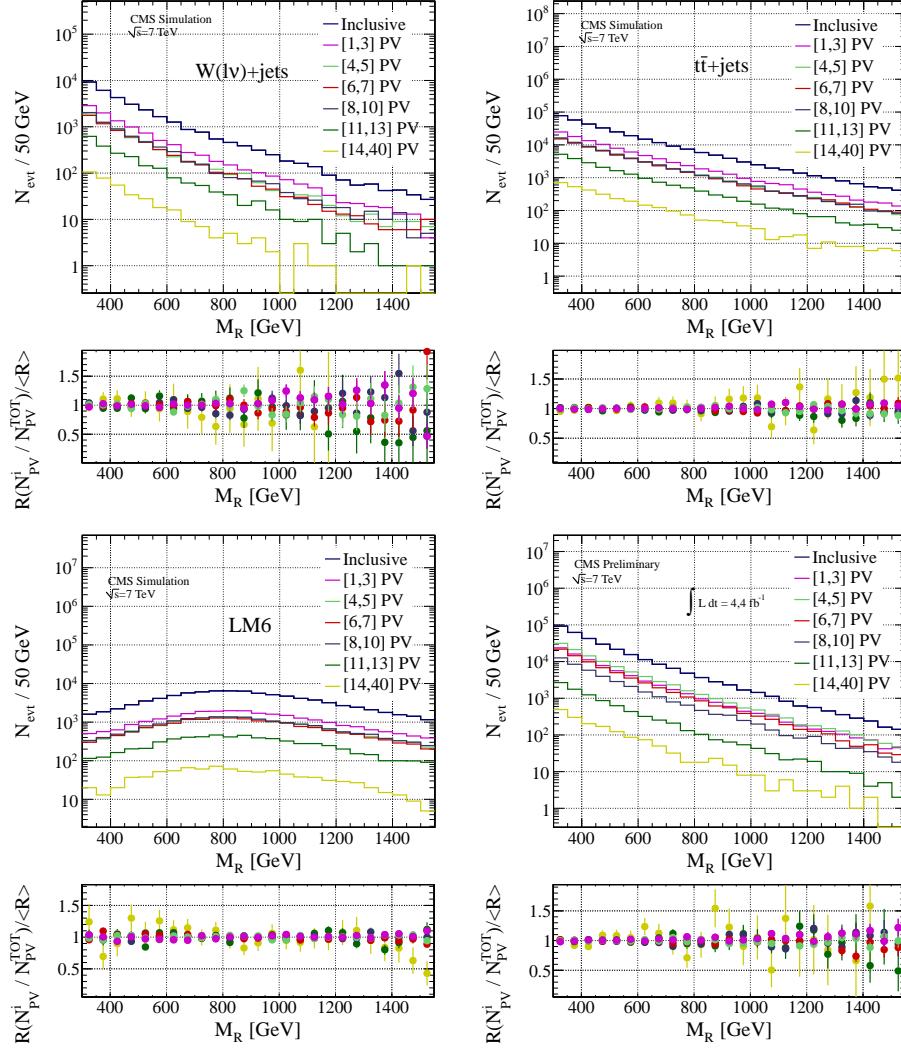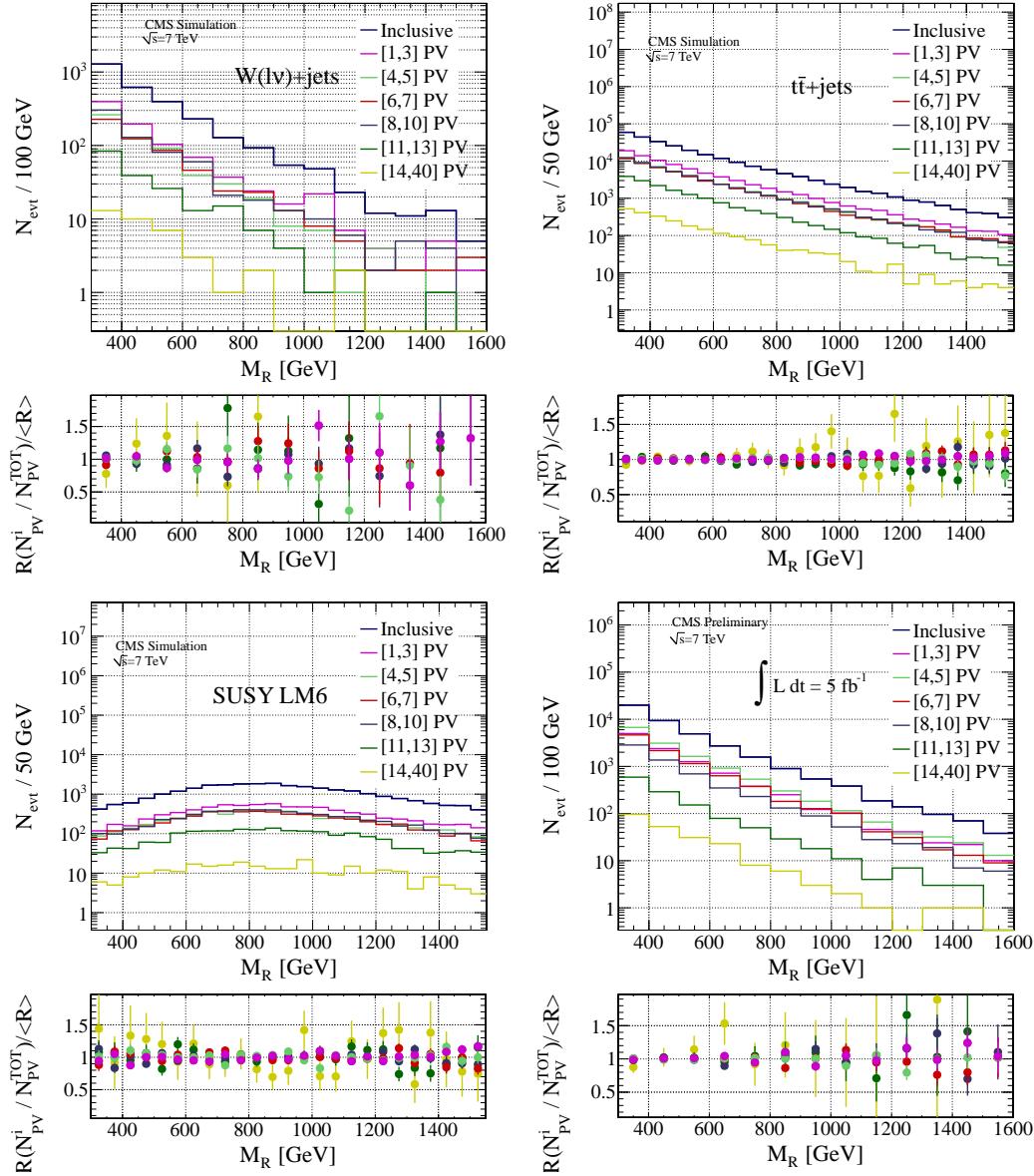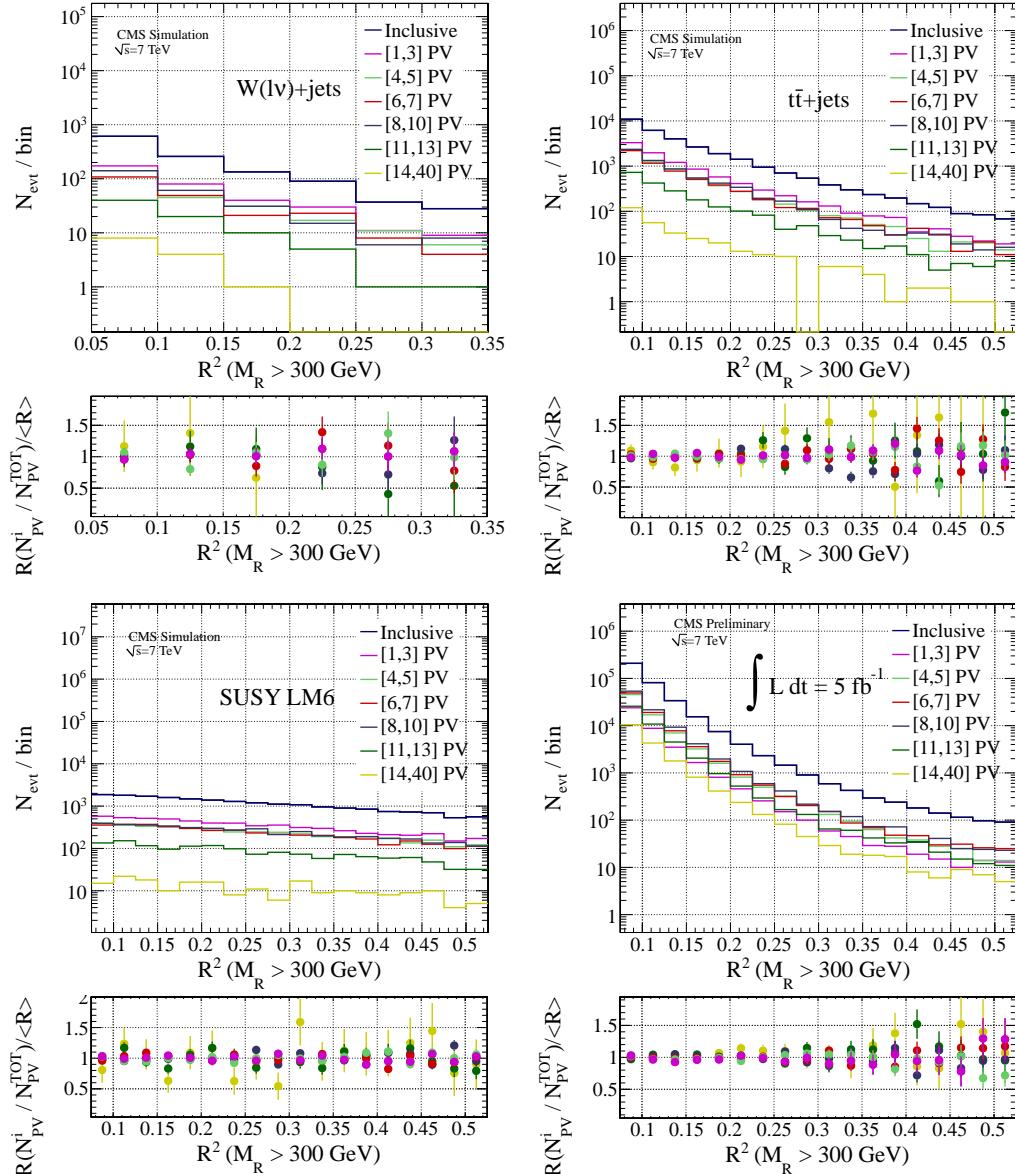
Figure C.12: Distribution of M_R in different bins of the number of vertices, for $W +$ $jets$ (Top left), $t\bar{t}$ (Top right), and SUSY LM6 (Bottom left) simulated samples, and for data in the HAD box (Bottom right). On the bottom part of each plot, the ratio of each distribution to the inclusive distribution is also shown.
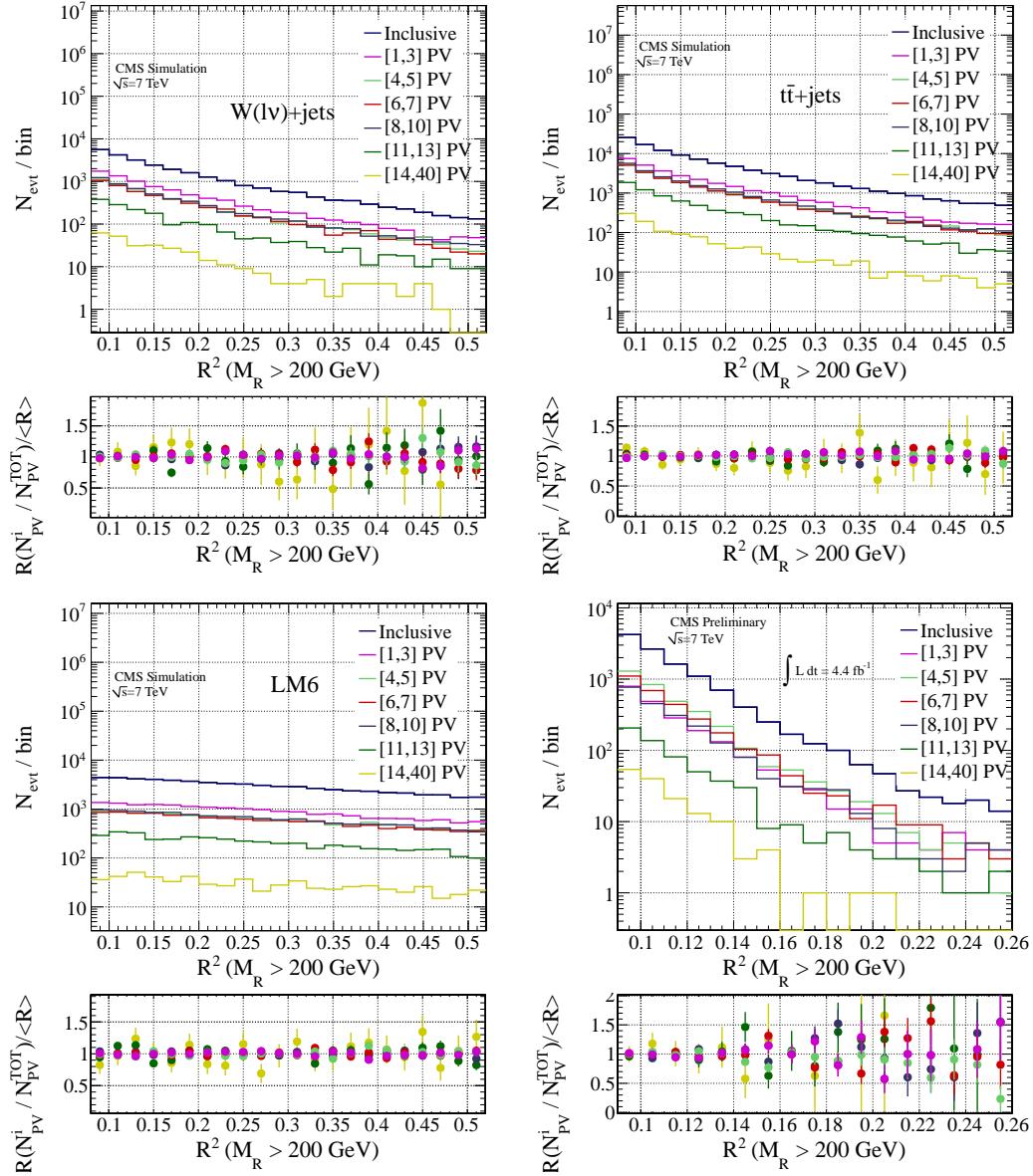
Figure C.13: Distribution of R^2 in different bins of the number of vertices, for $W+jets$ (Top left), $t\bar{t}$ (Top right), and SUSY LM6 (Bottom left) Monte Carlo samples, and for data in the HAD box (Bottom right). A baseline selection ($m_R > 200$ GeV) is applied. On the bottom part of each plot, the ratio of each distribution to the inclusive distribution is also shown.

Figure C.14: Distribution of M_R in different bins of the number of vertices, for $W + jets$ (Top left), $t\bar{t}$ (Top right), and SUSY LM6 (Bottom left) Monte Carlo samples, and for data in the HAD box (Bottom right). A baseline selection ($R > 0.2$) is applied. On the bottom part of each plot, the ratio of each distribution to the inclusive distribution is also shown.

C.7 Shape Dependence of Fit Region Variations

The trigger turn-on curves for the razor triggers, presented in section 10.4.1, indicate that events through the razor plane are efficiently selected. At low M_R, there are

Figure C.15: Distribution of R^2 in different bins of the number of vertices, for $W+jets$ (Top left), $t\bar{t}$ (Top right), and SUSY LM6 (Bottom left) Monte Carlo samples, and for data in the HAD box (Bottom right). A baseline selection ($M_R > 300$ GeV) is applied. On the bottom part of each plot, the ratio of each distribution to the inclusive distribution is also shown.

small inefficiencies which do not enter in the signal sensitive region but potentially do at the smallest M_R part of the fit region. In order to test for any potential shape bias resulting from these inefficiencies we repeat the ML fits for the HAD, MU and ELE, in both the inclusive and b-tagged analysis, varying the lower M_R bound of the
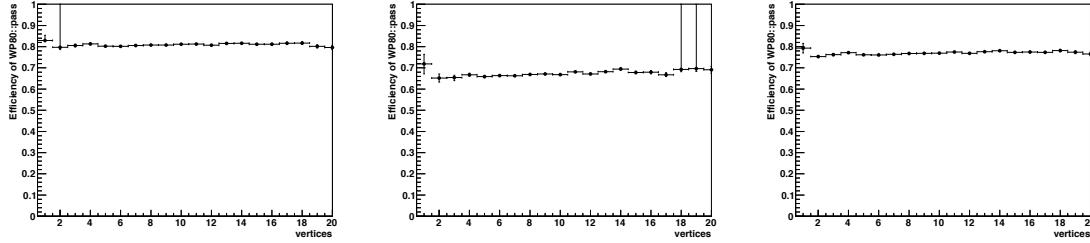
Figure C.16: Electron identification efficiency as a function of the number of reconstructed vertices in $Z(ee)$ events selected from 2011 data. Efficiencies are determined using a tag-and-probe fit to the di-electron invariant mass distribution. The efficiency is integrated in p_T and shown separately for barrel-only (Top), endcap-only (Center), and total acceptance (Bottom).

fit region by 50 and 100 GeV.

Inefficiencies of the razor triggers at small values of M_R could bias the fit region ML fit performed in each of the boxes. The shape parameters measured in these fits are summarized in Tab. C.13. We observe that the shape parameters are consistent between the fits with varying M_R lower bounds for each box, indicating that there is no resolvable bias introduced in the fits from trigger inefficiencies.

C.8 Systematic Uncertainties from Choice of Background Function

In order to quantify the systematic error associated to our choice of background function, $F(M_R, R^2)$, we consider a more general function, $F_{SYS}(M_R, R^2)$ which deviates from exponentially falling behavior. $F_{SYS}(M_R, R^2)$ can be expressed as

$$F_{SYS}(M_R, R^2) = \left[b(M_R - M_R^0)^{1/n}(R^2 - R_0^2)^{1/n} - n \right] e^{-b(M_R-M_R^0)^{1/n}(R^2-R_0^2)^{1/n}} , \quad (C.2)$$

where the added parameter n changes the functional dependence on the razor variables. How the function changes can be understood in terms of the one dimensional projections of M_R and R^2. Integrating out each variable, independently, above some

HAD box (inclusive)			
Parameter	Nominal	+50 GeV	+ 100 GeV
M_R^0	19 ± 24	43 ± 26	41 ± 27
R_0^2	-0.258 ± 0.024	-0.271 ± 0.030	-0.276 ± 0.026
k	0.0169 ± 0.0016	0.0137 ± 0.0032	0.0136 ± 0.0032
MU box (inclusive)			
Parameter	Nominal	+50 GeV	+ 100 GeV
M_R^0	104 ± 29	49 ± 30	100 ± 51
R_0^2	-0.275 ± 0.028	-0.281 ± 0.027	-0.281 ± 0.026
k	0.0188 ± 0.0020	0.0178 ± 0.0022	0.0182 ± 0.0019
ELE box (inclusive)			
Parameter	Nominal	+50 GeV	+ 100 GeV
M_R^0	43 ± 28	37 ± 30	41 ± 27
R_0^2	-0.274 ± 0.027	-0.269 ± 0.029	-0.271 ± 0.030
k	0.0138 ± 0.0030	0.0105 ± 0.0047	0.0112 ± 0.0050
HAD box (b-tagged)			
Parameter	Nominal	+50 GeV	+ 100 GeV
M_R^0	-7 ± 23	-2 ± 23	-2 ± 23
R_0^2	-0.294 ± 0.030	-0.302 ± 0.030	-0.300 ± 0.030
k	0.0123 ± 0.00296	0.0122 ± 0.0036	0.0110 ± 0.0038
MU box (b-tagged)			
Parameter	Nominal	+50 GeV	+ 100 GeV
M_R^0	-1 ± 23	-2 ± 23	-4 ± 23
R_0^2	-0.304 ± 0.030	-0.305 ± 0.029	-0.303 ± 0.030
k	0.0137 ± 0.0050	0.0163 ± 0.0050	0.0144 ± 0.0074
ELE box (b-tagged)			
Parameter	Nominal	+50 GeV	+ 100 GeV
M_R^0	-2 ± 23	-1 ± 23	-1 ± 23
R_0^2	-0.301 ± 0.030	-0.303 ± 0.030	-0.304 ± 0.030
k	0.0126 ± 0.0050	0.0128 ± 0.0053	0.0137 ± 0.0051

Table C.13: Determination of the UEC parameters in the HAD, MU and ELE boxes, for the inclusive and b-tagged data samples. Different baseline cuts on M_R are considered and the fits are repeated, giving new shape parameters.

cut value yields one dimensional functions

$$\int_{R_{\text{cut}}^2}^{\infty} F_{SYS}(M_R, R^2) dR^2 \sim e^{-k_{M_R}(M_R - M_R^0)^{1/n}} \; , \tag{C.3}$$

$$\int_{M_R^{\text{cut}}}^{\infty} F_{SYS}(M_R, R^2) dM_R \sim e^{-k_{R^2}(R^2 - R_0^2)^{1/n}} \; , \tag{C.4}$$

where

$$k_{M_R} = k^0_{M_R} + k(R^2_{\text{cut}})^{1/n} \; , \tag{C.5}$$

$$k_{R^2} = k^0_{R^2} + k(M^{\text{cut}}_R)^{1/n} \; , \tag{C.6}$$

with R^2_{cut} and M^{cut}_R indicating the lower bounds for integrating over the two variables, respectively. We observe that the addition of the parameter n results in *bending* of the one dimensional functional forms, with the deviation from an exponential slope progressively increasing with larger M_R and R^2 values.

We consider potential deviations from exponential behavior as a possible systematic uncertainty and, using instances of the function F_{SYS}, evaluate the magnitude of this uncertainty as follows:

- For each box, the functional form of the UEC component (which is the only relevant component over the majority of the razor plane in each box) is replaced with the new function $F_{SYS}(M_R, R^2)$. The fit in the fit region of each box is repeated with the nominal configuration, except now floating n. From this fit we determine $n_{fit} \pm \sigma_n$.

- A range of n variations from one is defined by considering the largest examples of $|n_{fit} - 1|$ and σ_n, building an interval $[n_{min}, n_{max}]$ which reflects the preference of the data.

- The fit is repeated in the fit region of each box, now *fixing* n to n_{min} and n_{max} in independent fits. The background models corresponding to these fits are considered as alternative background models which include potential systematic variations due to deviations from exponential behavior.

- The signal sensitive region of the razor plane is partitioned into a collection of subregions, and the yield prediction is calculated for each background model in each subregion.

The fit result with n floated in the HAD box is shown in figure C.17. We see that
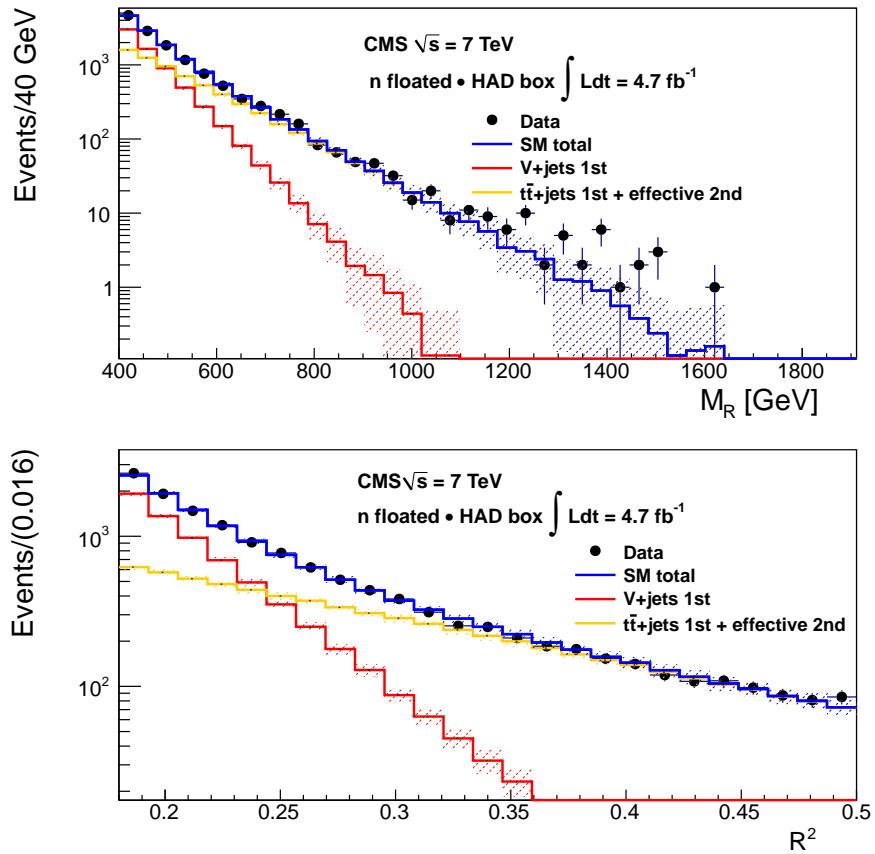
Figure C.17: Projection of the fit result on the M_R (Top) and R^2 (Bottom) axis for the HAD Box, where a background model is used with a floating n parameter and function F_{SYS} for the ELE-like UEC.

the extrapolation of the fit over the full razor plane is in agreement with observed event yields and with the nominal fit. We observe $n_{fit} = 0.96 \pm 0.04$, which motivates the choice of $n_{max} = 1.04$ and $n_{min} = 0.96$ for building our alternative background models.

Each alternative background model is used to determine a prediction of the expected yield in an array of subregions of the razor plane. The background predictions for each of the subregions for each of the alternative background models is summarized in Tab. C.14 for the HAD box. We observe that systematic differences between the predictions from each background model considered are small, negligible relative to the other uncertainties included in the background model. The same conclusion follows from the analogous results from each box, indicating that this systematic can be safely ignored in the analysis.

C.9   Background Prediction Closure Tests for Simulated Events

In order to test the closure of the background prediction method we repeat the exercise on a cocktail of simulated SM background events. Due to the different effective integrated luminosities between simulated samples, each process is weighted so that its total yield matches that expected for 4.7 fb^{-1}, assuming NLO theoretical cross sections. This cocktail is used to generate a pseudodata sample on which the analysis can be repeated.

The closure test procedure follows several steps:

- The weighted SM background simulated event cocktail is used to create a 2D background template in M_R/R^2 for each box. A random data set of events is drawn from each boxes's template, resulting in a pseudodata sample.

- The ML fit procedure is applied to the pseudodataset and the background model, with associated errors, is determined for the signal sensitive region of the razor plane .

SubRegion (SR)	$n = 1$	$n = n_{min}$	$n = n_{max}$	n floated
HAD $SR_{1,1}$	1558.2 ± 69.0	1527.3 ± 108.8	1509.0 ± 111.0	1510.8 ± 126.2
HAD $SR_{1,2}$	2898.4 ± 79.9	2887.5 ± 88.9	2868.4 ± 97.9	2865.9 ± 99.2
HAD $SR_{1,3}$	710.8 ± 35.0	728.9 ± 45.0	713.5 ± 43.2	726.4 ± 49.3
HAD $SR_{1,4}$	328.5 ± 36.5	337.7 ± 31.1	328.1 ± 31.5	336.9 ± 33.6
HAD $SR_{2,1}$	1785.2 ± 63.6	1786.7 ± 74.9	1758.6 ± 68.9	1774.3 ± 67.1
HAD $SR_{2,2}$	3301.3 ± 82.0	3335.8 ± 103.5	3313.3 ± 111.5	3348.5 ± 117.9
HAD $SR_{2,3}$	944.6 ± 45.7	957.0 ± 46.7	956.7 ± 47.1	964.0 ± 48.1
HAD $SR_{2,4}$	432.2 ± 36.1	422.5 ± 35.1	453.8 ± 37.3	424.1 ± 37.5
HAD $SR_{3,1}$	251.0 ± 26.0	262.7 ± 27.7	258.7 ± 30.6	259.7 ± 29.1
HAD $SR_{3,2}$	536.8 ± 46.6	543.5 ± 45.4	561.0 ± 49.7	550.0 ± 48.9
HAD $SR_{3,3}$	172.9 ± 36.1	156.6 ± 28.7	181.6 ± 32.9	161.8 ± 34.1
HAD $SR_{3,4}$	57.9 ± 18.0	51.5 ± 16.8	66.2 ± 18.7	50.5 ± 17.8
HAD $SR_{4,1}$	38.5 ± 8.7	37.4 ± 10.6	43.1 ± 9.2	37.7 ± 8.9
HAD $SR_{4,2}$	86.4 ± 22.7	73.7 ± 17.1	90.0 ± 24.3	75.6 ± 20.8
HAD $SR_{4,3}$	19.5 ± 7.4	14.3 ± 5.7	21.7 ± 9.0	14.4 ± 7.0
HAD $SR_{4,4}$	4.2 ± 2.9	2.7 ± 2.3	4.9 ± 3.1	2.4 ± 2.4
HAD $SR_{5,1}$	4.7 ± 2.8	3.9 ± 2.5	5.3 ± 3.1	4.1 ± 2.9
HAD $SR_{5,2}$	8.3 ± 4.7	6.0 ± 3.7	9.5 ± 4.7	5.9 ± 4.0
HAD $SR_{5,3}$	1.2 ± 1.2	0.8 ± 0.8	1.5 ± 1.5	0.8 ± 0.8
HAD $SR_{5,4}$	0.4 ± 0.4	0.4 ± 0.4	0.5 ± 0.5	0.4 ± 0.4
HAD $SR_{6,1}$	0.8 ± 0.8	0.6 ± 0.6	0.9 ± 0.9	0.6 ± 0.6
HAD $SR_{6,2}$	1.0 ± 1.0	0.7 ± 0.7	1.2 ± 1.2	0.8 ± 0.8
HAD $SR_{6,3}$	0.4 ± 0.4	0.3 ± 0.3	0.4 ± 0.4	0.4 ± 0.4

Table C.14: Background yield predictions for different subregions of the razor plane from several alternative background models.

- The background model is used to estimate the distribution of expected yields in subregions of the signal sensitive razor plane. These predictions are compared with the yields of the original weighted template to test for closure.

This entire procedure is completed for both the inclusive analysis and b-tagged jet subsample analysis.

The one-dimensional projections in M_R and R^2 of the 2D ML fits to the pseudo-datasets in each box are shown for the inclusive analysis in figures C.18 and C.19, respectively. The p-values corresponding to the resulting background prediction in the signal regions of each box are given in figure C.20 and summarized in figure C.21. The procedure successfully closes, correctly predicting the shape and yield of the

pseudodata sample in the signal regions. Analogous results for the b-tagged analysis closure test are presented in figures C.18-C.21, and also indicate closure of the method.
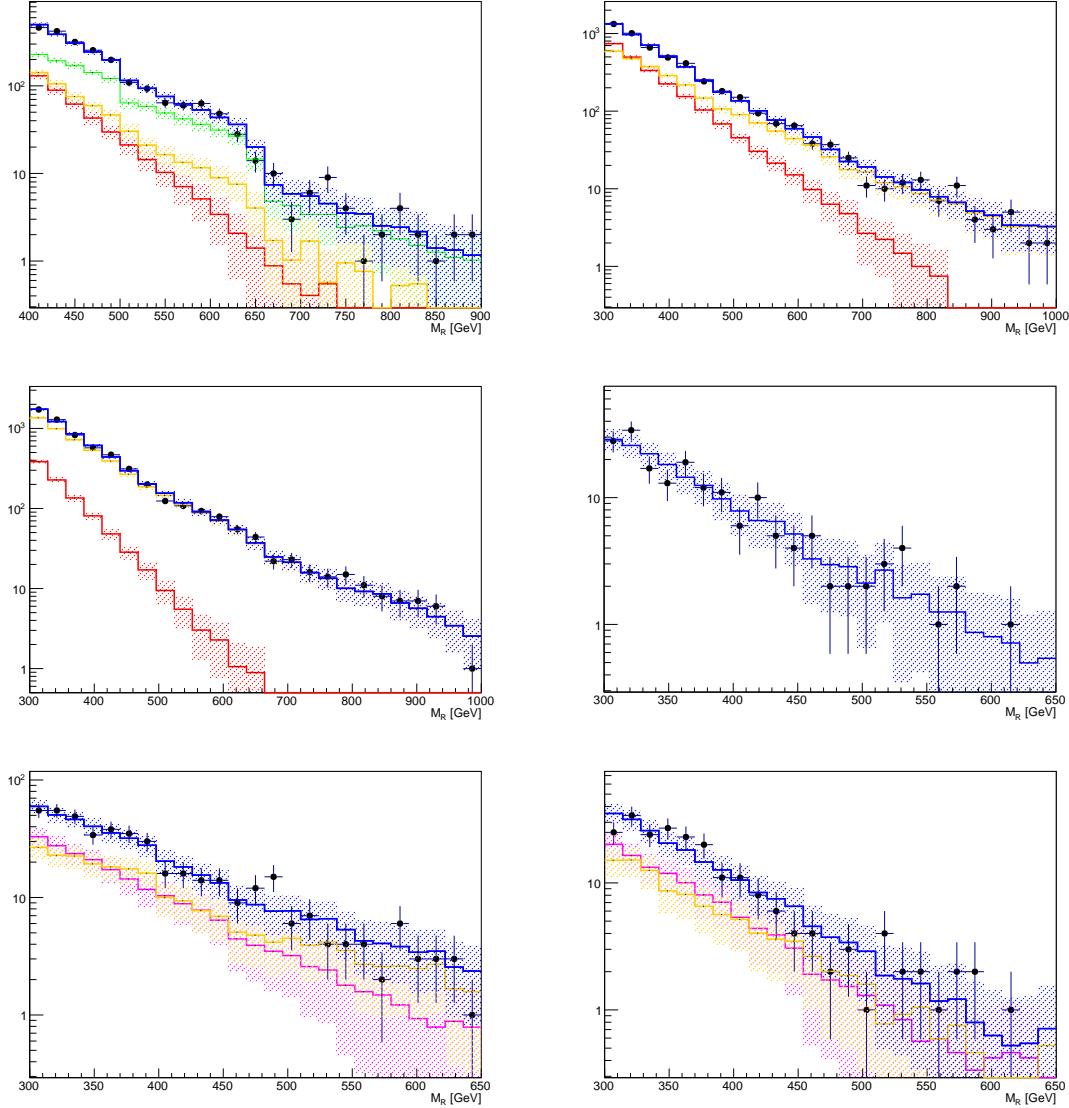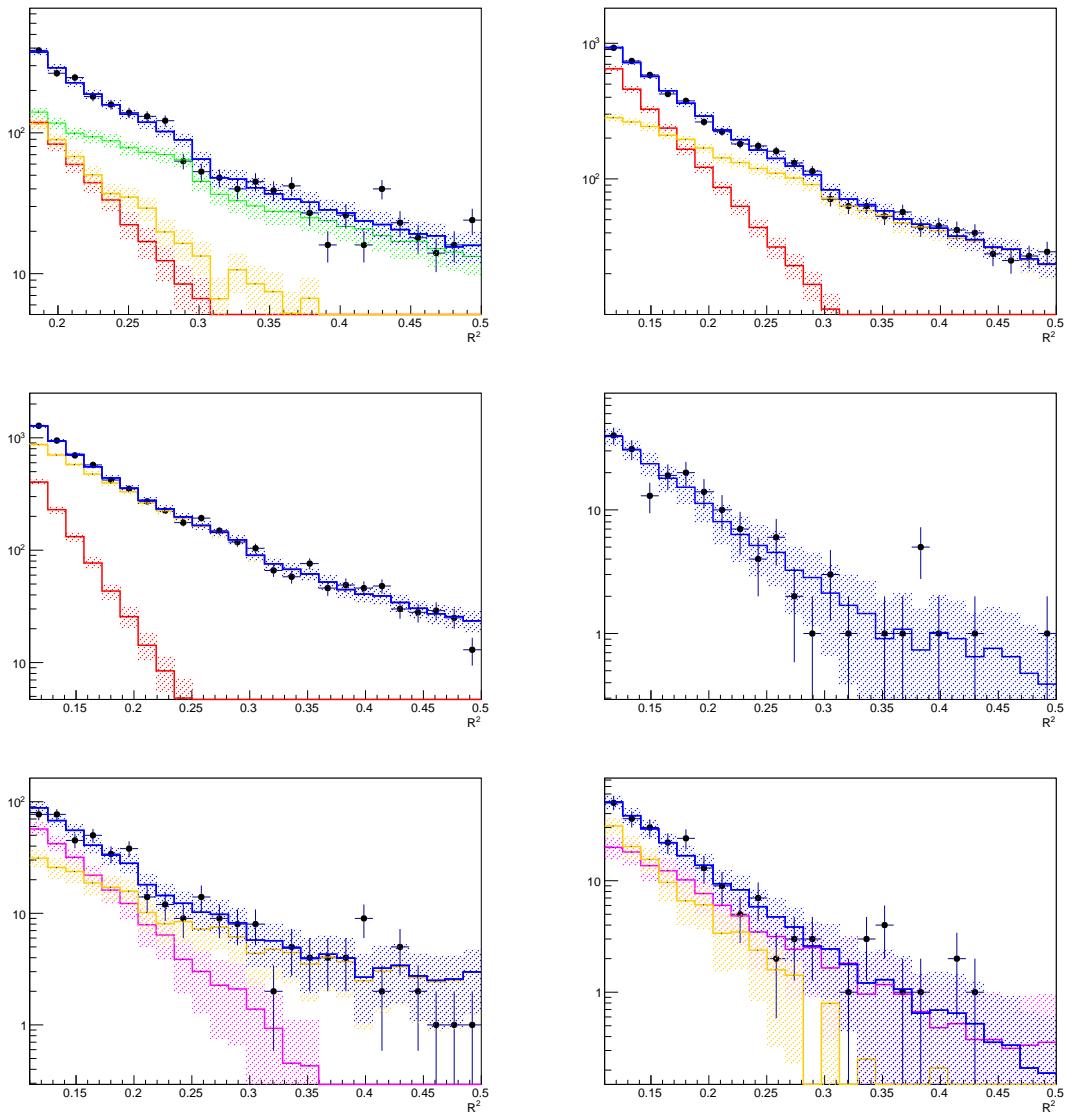

Figure C.18: Projection of the 2D ML fit on M_R for the pseudodata sample in the fit region for the HAD (Top left), MU (Top right), ELE (Center left), ELE-MU (Center right), MU-MU (Bottom left) and ELE-ELE (Bottom right) boxes. Inclusive pseudodata used. Different colors correspond to different background components.
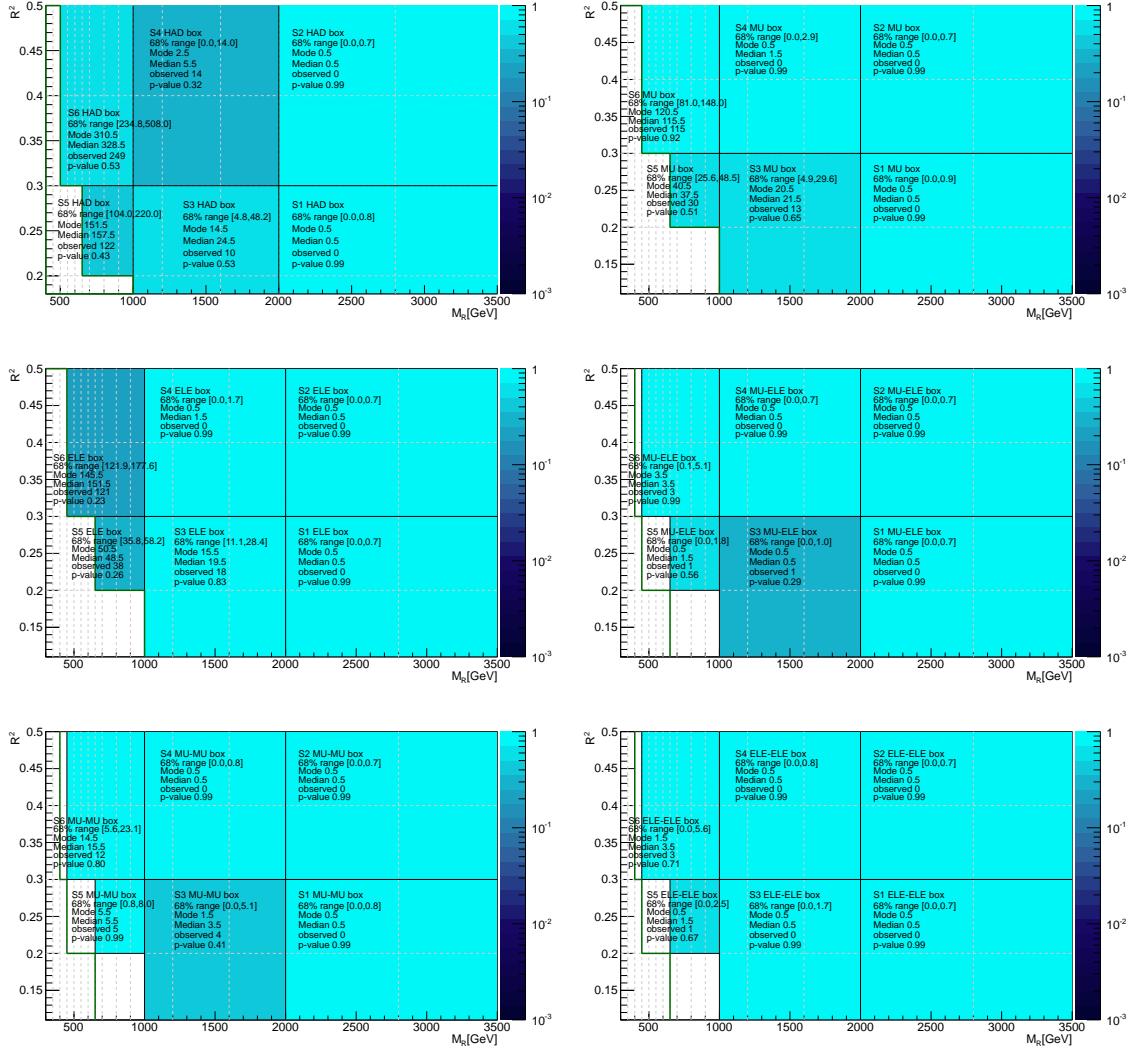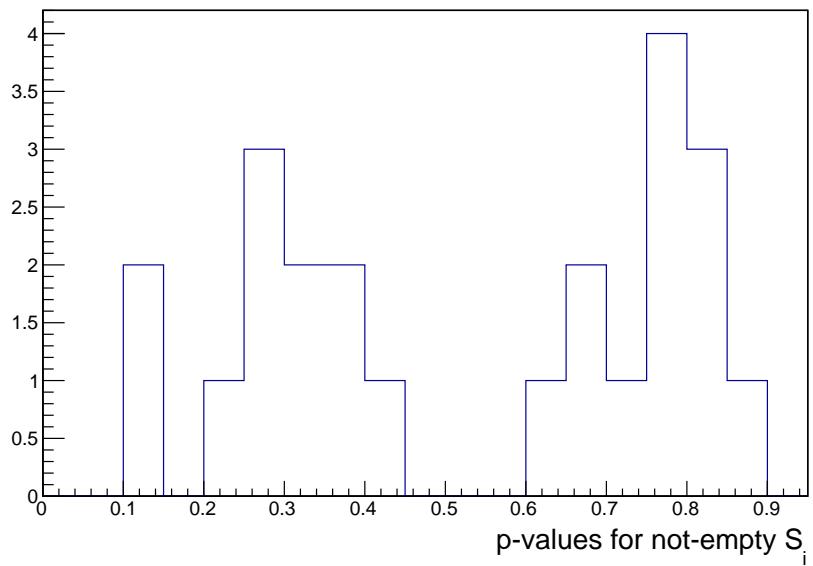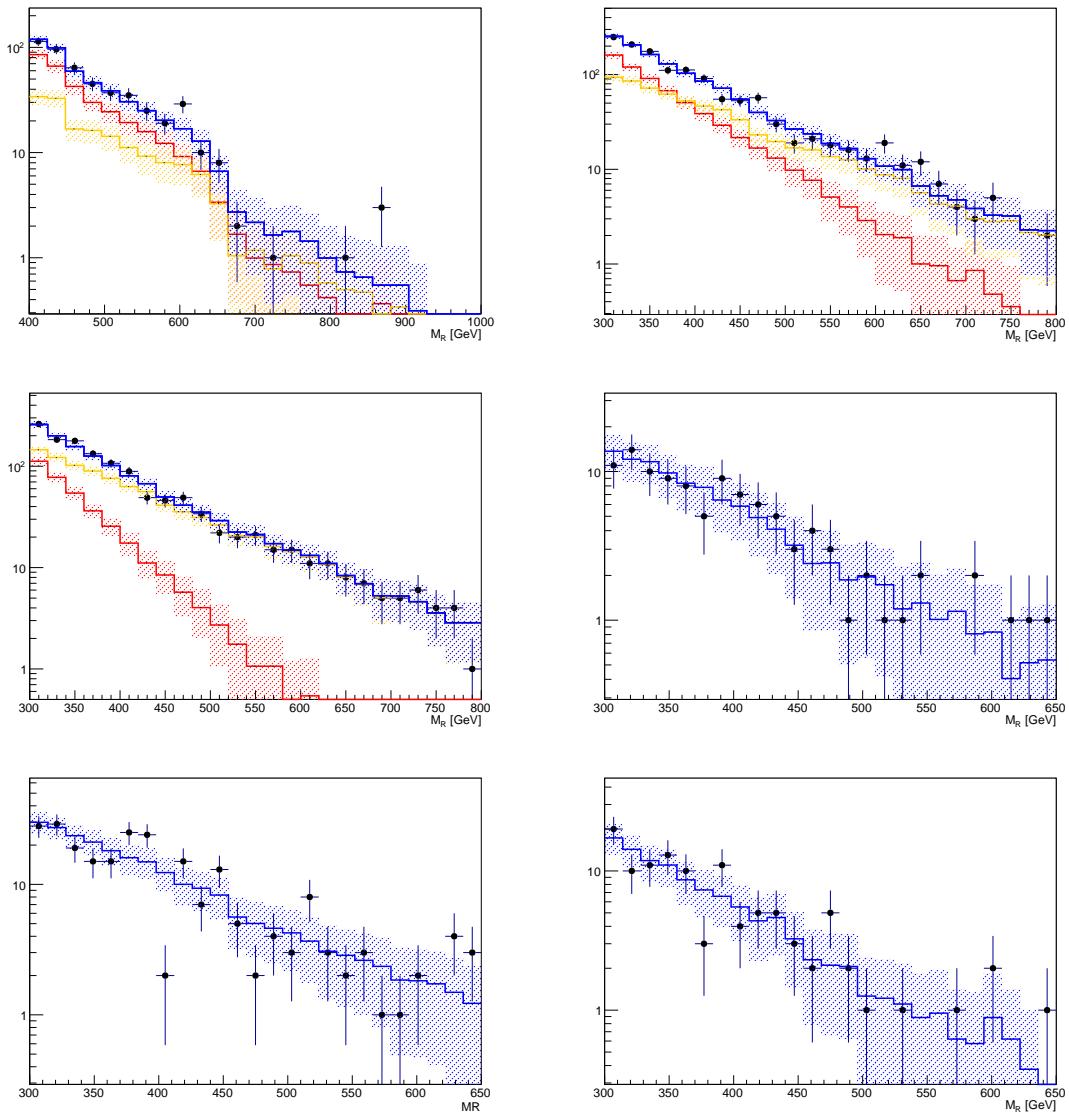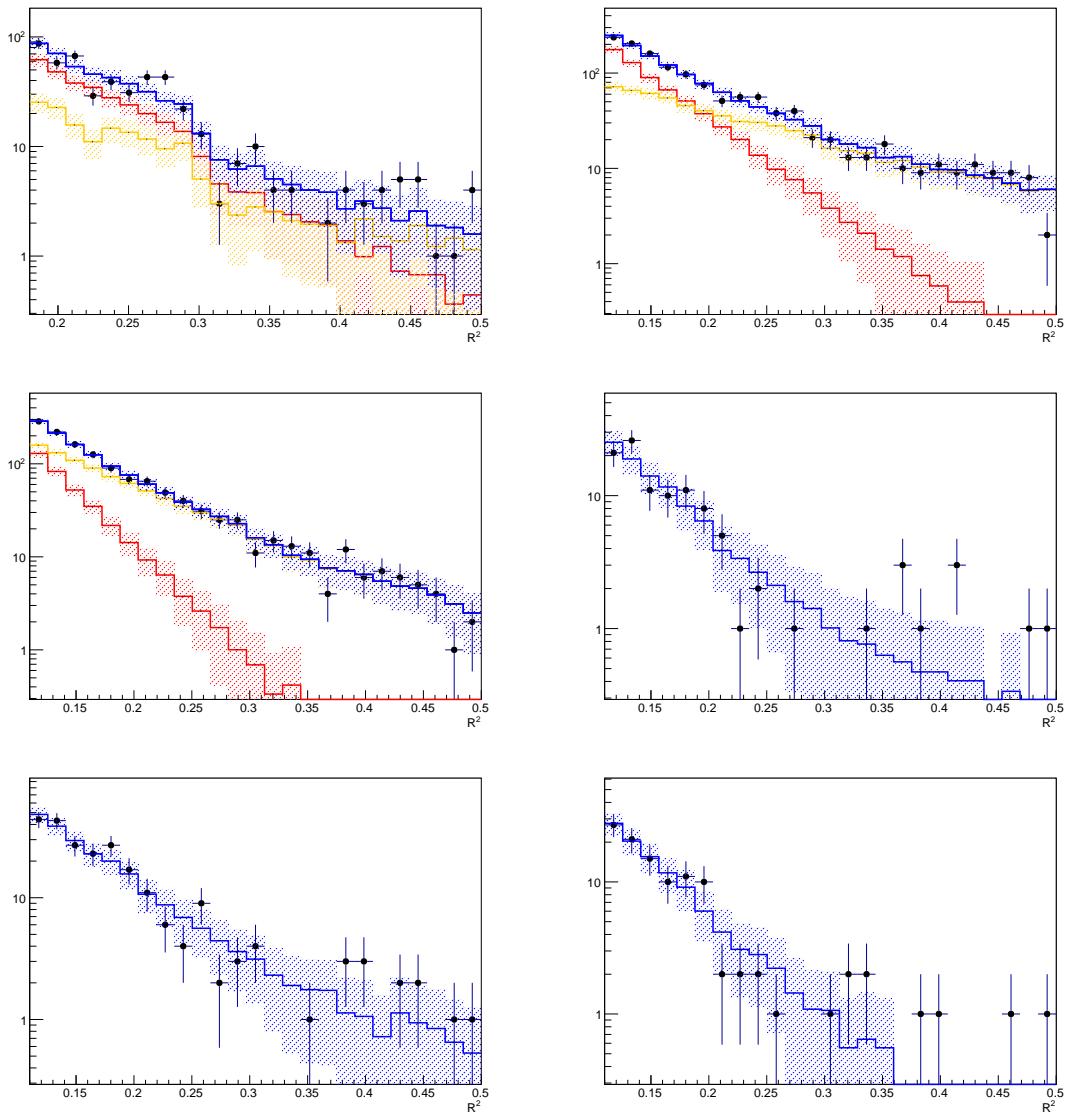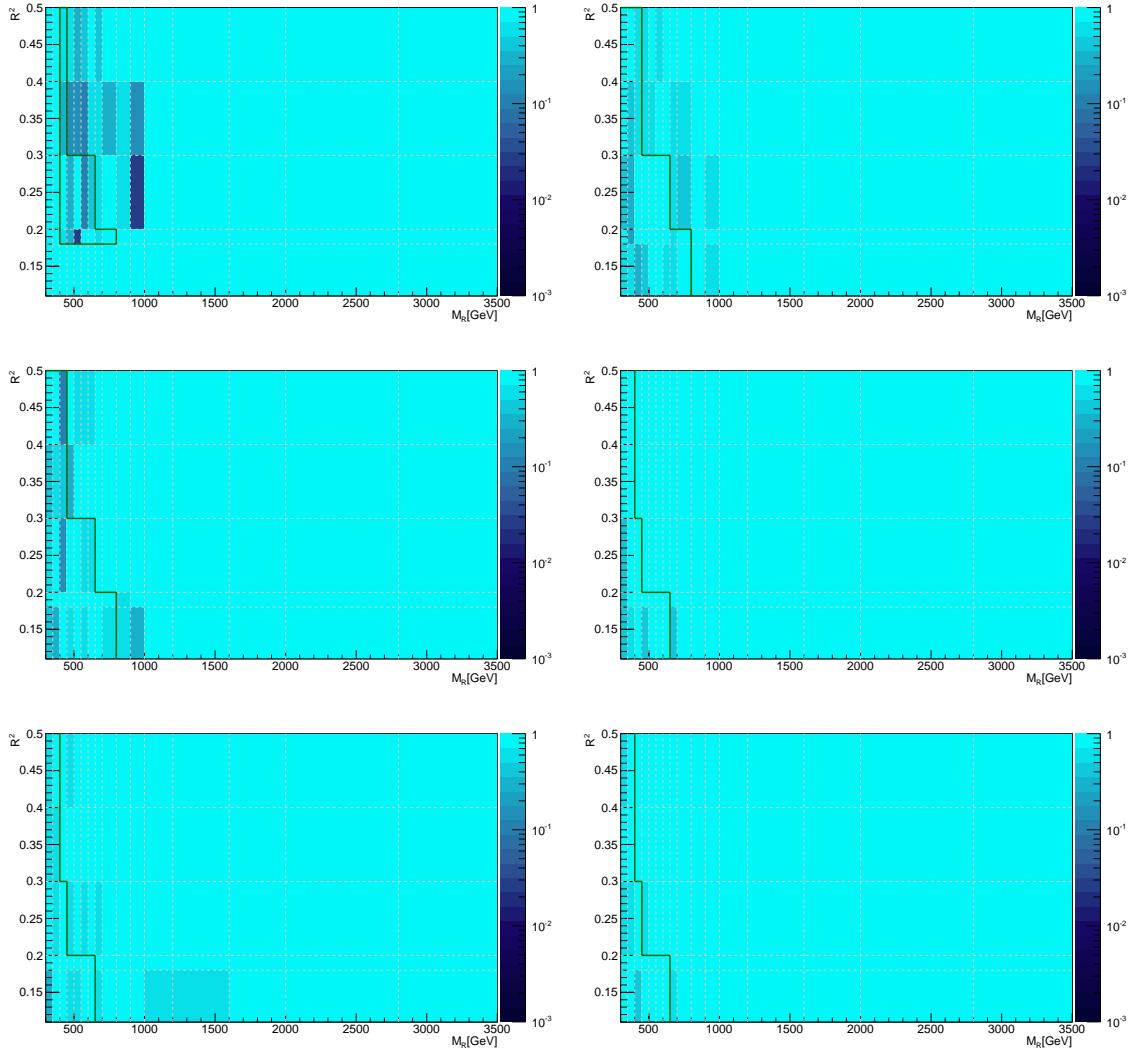
Figure C.19: Projection of the 2D ML fit on R^2 for the pseudodata sample in the fit region for the HAD (Top left), MU (Top right), ELE (Center left), ELE-MU (Center right), MU-MU (Bottom left) and ELE-ELE (Bottom right) boxes. Inclusive pseudodata used. Different colors correspond to different background components.

Figure C.20: P-values for the yields in the signal regions of each box relative to the closure test background prediction for the HAD (Top left), MU (Top right), ELE (Center left), ELE-MU (Center right), MU-MU (Bottom left) and ELE-ELE (Bottom right) boxes. Inclusive pseudodata used.

Figure C.21: *P*-values for the non-empty signal regions from the background closure test for the inclusive analysis.
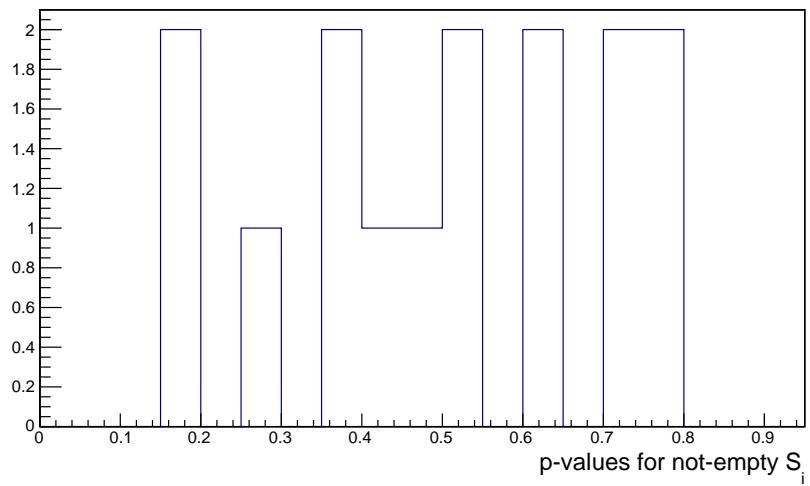
Figure C.22: *P*-values for the non-empty signal regions from the background closure test for the *b*-tagged jet analysis.

Figure C.23: Projection of the 2D ML fit on M_R for the pseudodata sample in the fit region for the HAD (Top left), MU (Top right), ELE (Center left), ELE-MU (Center right), MU-MU (Bottom left) and ELE-ELE (Bottom right) boxes. *b*-tagged jet pseudodata used. Different colors correspond to different background components.

Figure C.24: Projection of the 2D ML fit on R^2 for the pseudodata sample in the fit region for the HAD (Top left), MU (Top right), ELE (Center left), ELE-MU (Center right), MU-MU (Bottom left) and ELE-ELE (Bottom right) boxes. b-tagged jet pseudodata used. Different colors correspond to different background components.

Figure C.25: P-values for the yields in the signal regions of each box relative to the closure test background prediction for the HAD (Top left), MU (Top right), ELE (Center left), ELE-MU (Center right), MU-MU (Bottom left) and ELE-ELE (Bottom right) boxes. b-tagged jet pseudodata used.

C.10 Additional Information on Background Fits

C.10.1 Inclusive Data Sample
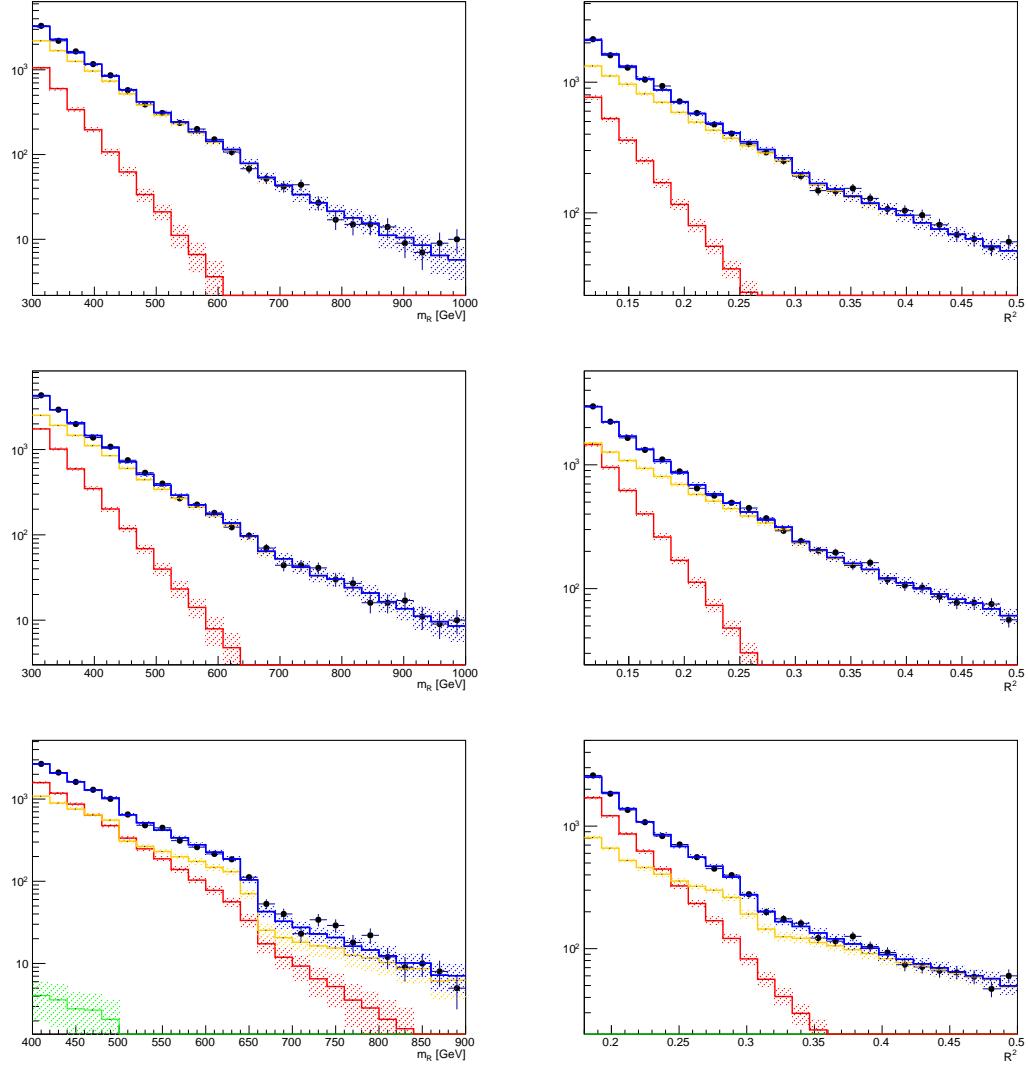
Figure C.26: 1D projections of 2D ML fits in M_R (Left) and R^2 (Right). Samples correspond to the inclusive MU (Top), ELE (middle) and HAD (Bottom) boxes. The total background prediction is indicated in blue. The yellow contour corresponds to the UEC component combined with the first $t\bar{t}$ component. The V+jets first component contribution is indicated in red. For the HAD box, the second UEC component (MU-like) contribution is shown in green. Uncertainty bands on each contribution only include statistical fluctuations, not systematic uncertainties on the function parameters.
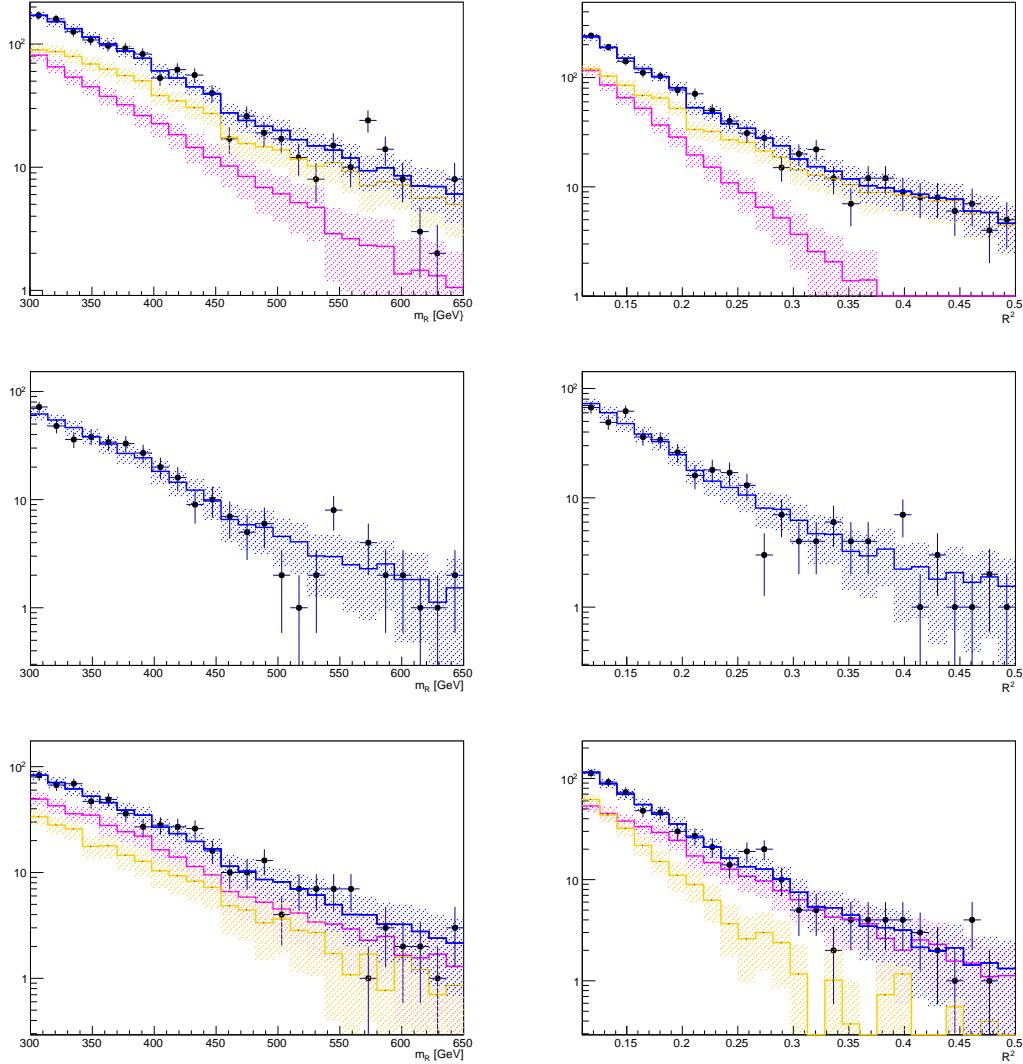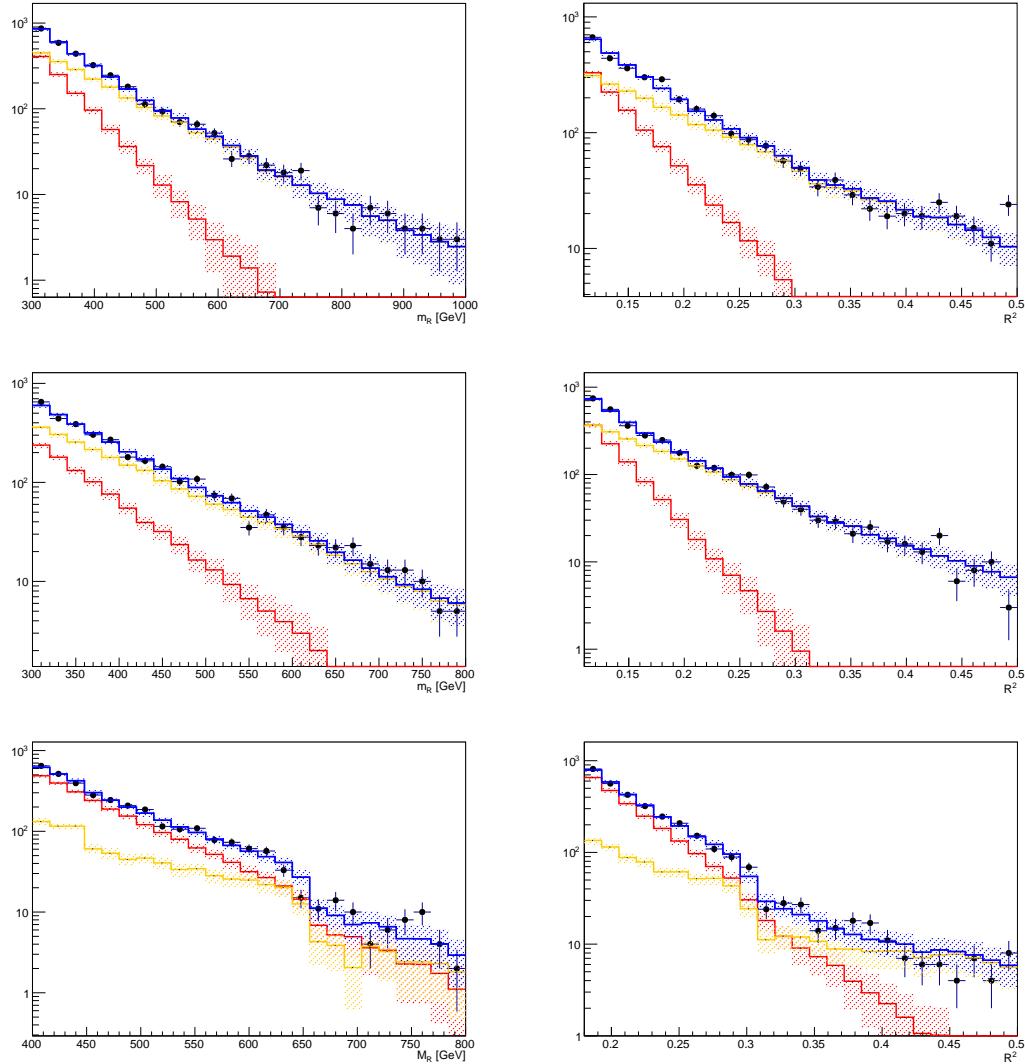
Figure C.27: 1D projections of 2D ML fits in M_R (Left) and R^2 (Right). Samples correspond to the inclusive MU-MU (Top), ELE-MU (middle) and ELE-ELE (Bottom) boxes. The total background prediction is indicated in blue. The yellow contour corresponds to the UEC component combined with the first $t\bar{t}$ component. The V+jets first component contribution is indicated in magenta. The ELE-MU box is modeled without a V+jets contribution. Uncertainty bands on each contribution only include statistical fluctuations, not systematic uncertainties on the function parameters.
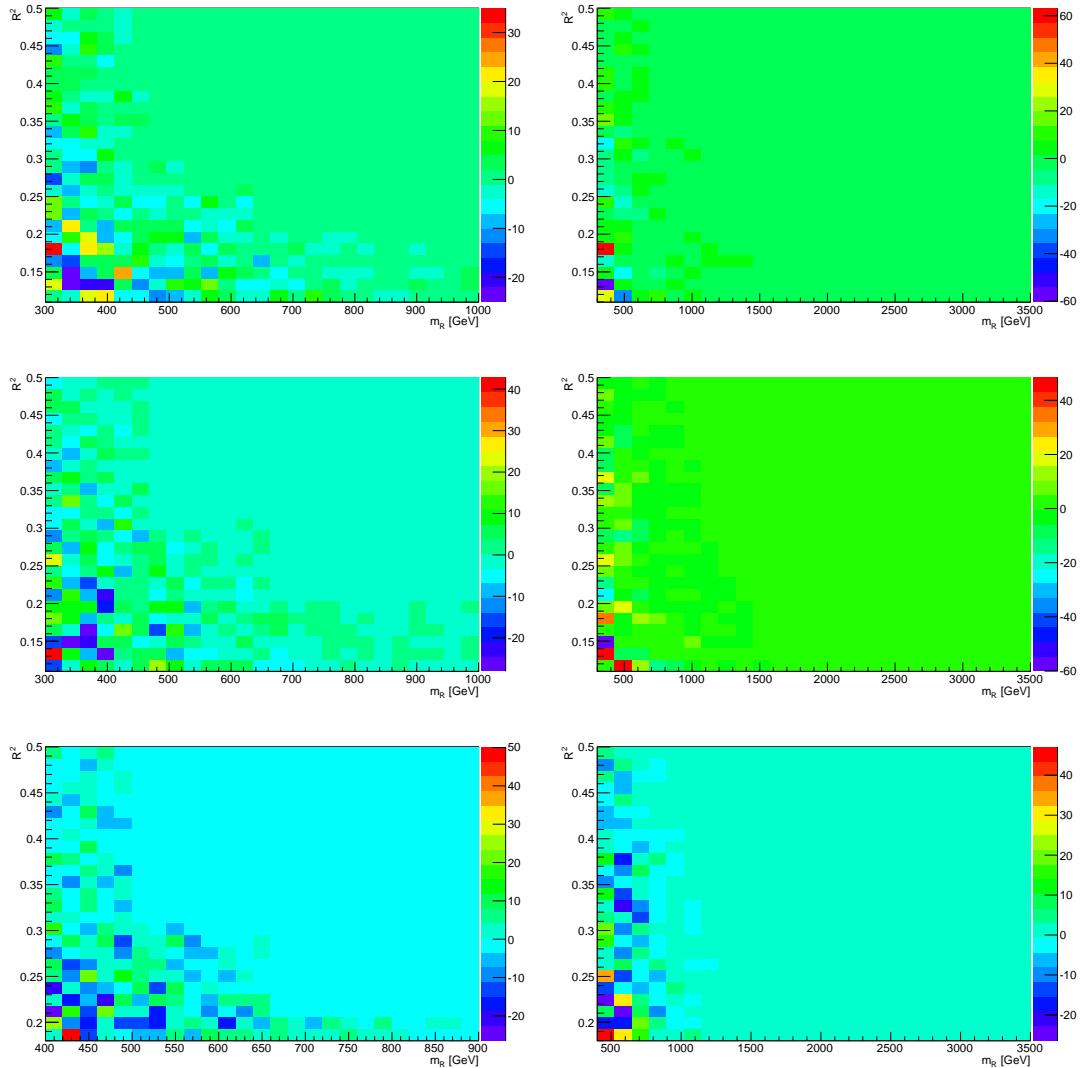
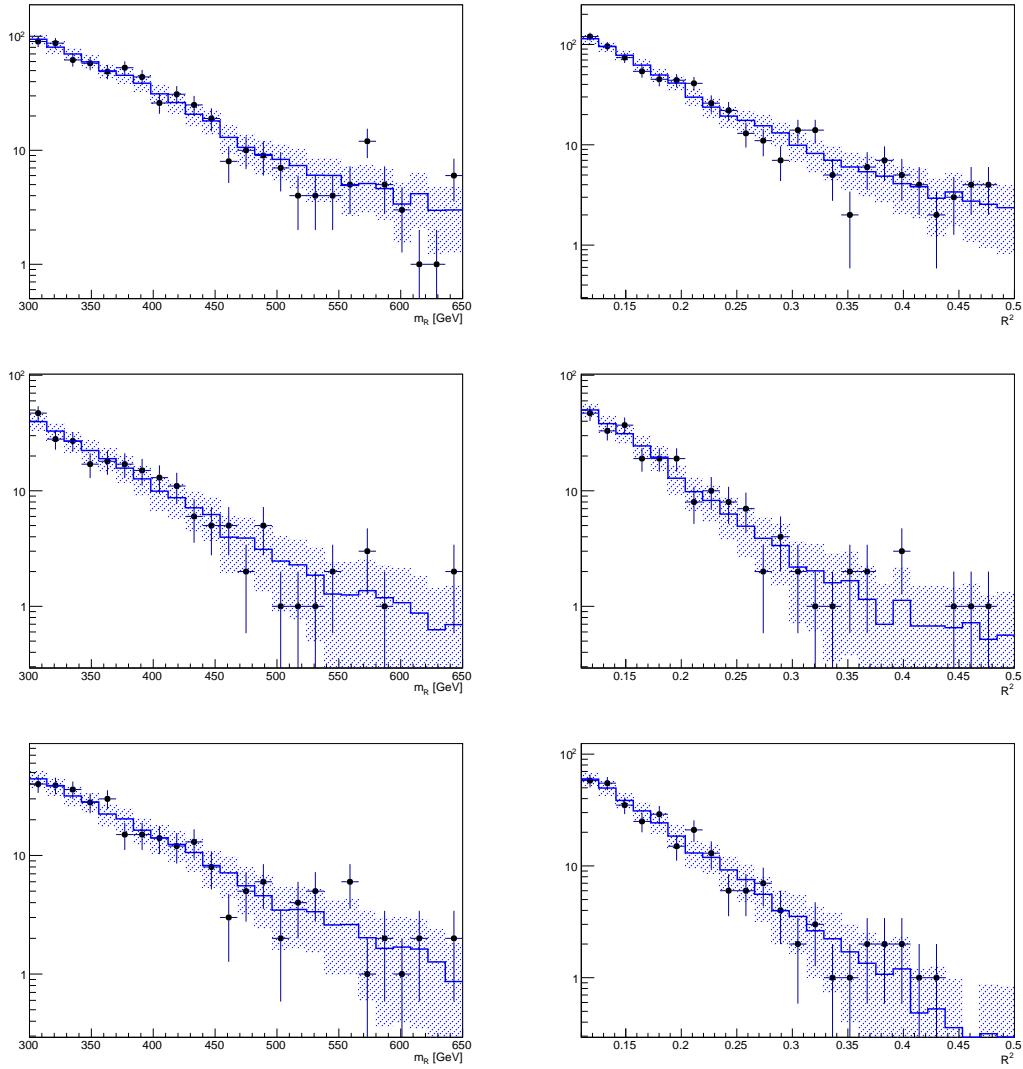Figure C.28: Bin-by-bin difference between the number of observed data events and the integral of the background model for the fit region (Left) and the full razor plane (Right). Samples correspond to the inclusive MU (Top), ELE (middle) and HAD (Bottom) boxes. Difference is in absolute number of events.
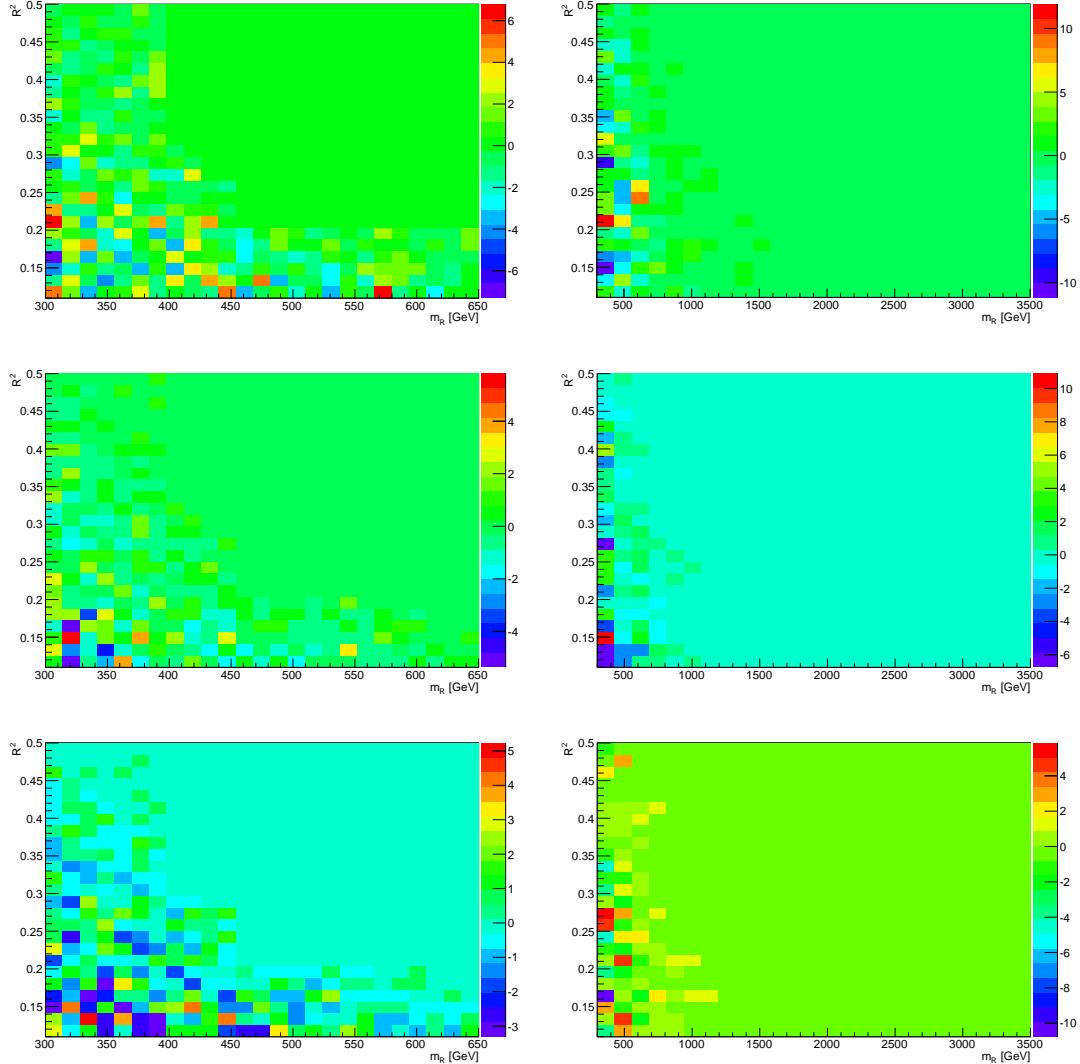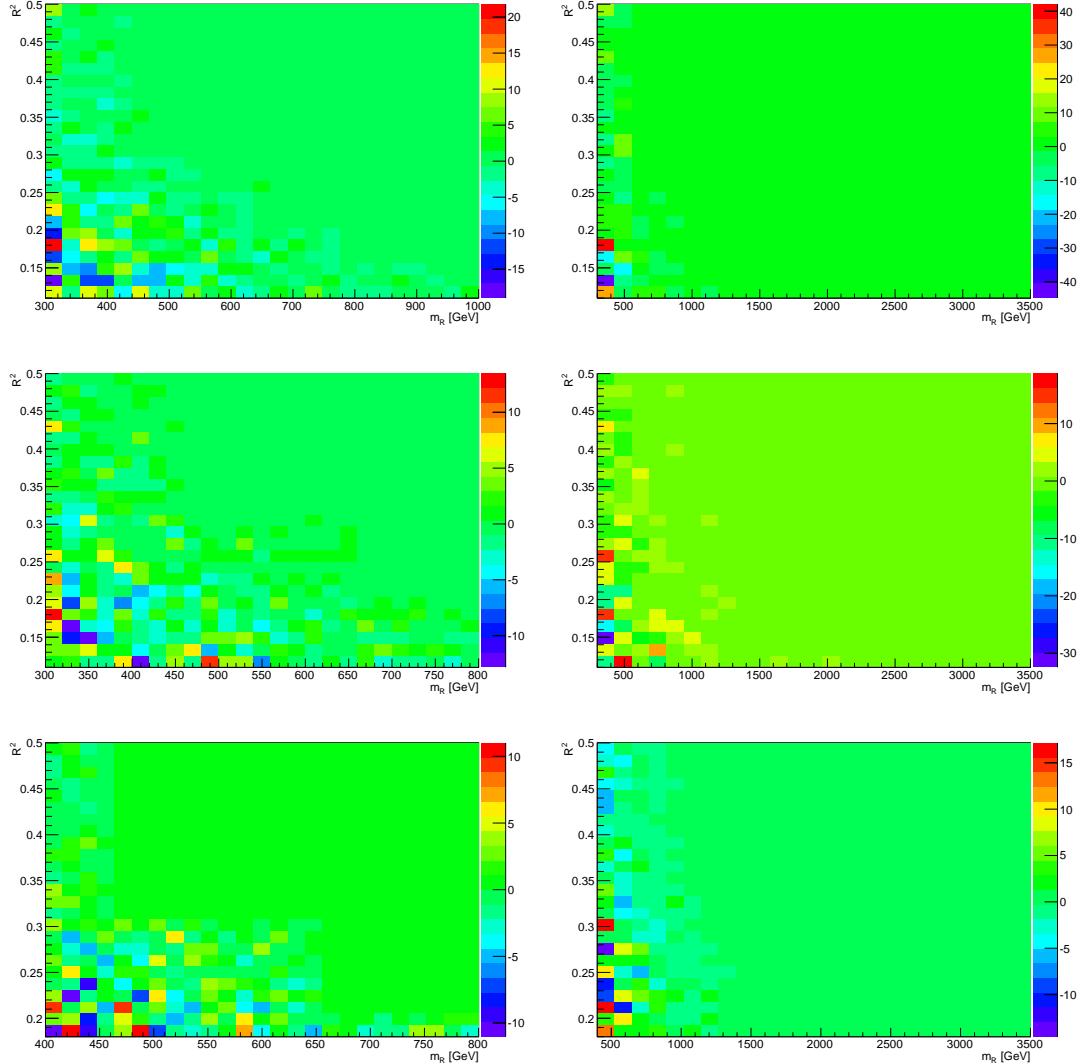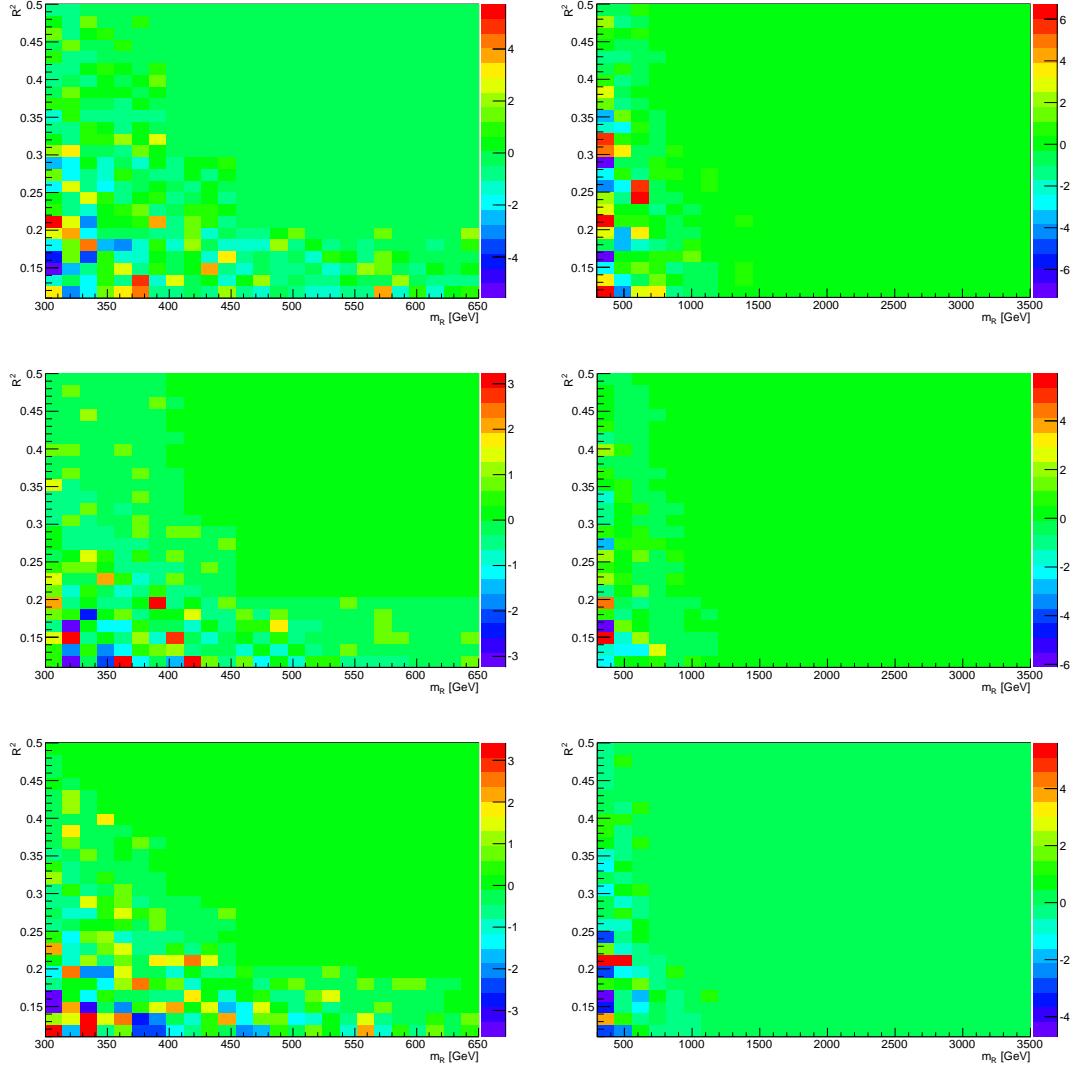
Figure C.29: Bin-by-bin difference between the number of observed data events and the integral of the background model for the fit region (Left) and the full razor plane (Right). Samples correspond to the inclusive MU-MU (Top), ELE-MU (middle) and ELE-ELE (Bottom) boxes. Difference is in absolute number of events.

C.10.2   b-tagged Data Sample

Figure C.30: 1D projections of 2D ML fits in M_R (Left) and R^2 (Right). Samples correspond to the b-tagged MU-MU (Top), ELE-MU (middle) and ELE-ELE (Bottom) boxes. The total background prediction is indicated in blue. The yellow contour corresponds to the UEC component while the first $t\bar{t}$ component is indicated in red. Uncertainty bands on each contribution only include statistical fluctuations, not systematic uncertainties on the function parameters.

Figure C.31: 1D projections of 2D ML fits in M_R (Left) and R^2 (Right). Samples correspond to the b-tagged MU-MU (Top), ELE-MU (middle) and ELE-ELE (Bottom) boxes. The total background prediction is indicated in blue, including an effective first component and UEC. Uncertainty bands on each contribution only include statistical fluctuations, not systematic uncertainties on the function parameters.

Figure C.32: Bin-by-bin difference between the number of observed data events and the integral of the background model for the fit region (Left) and the full razor plane (Right). Samples correspond to the *b*-tagged MU (Top), ELE (middle) and HAD (Bottom) boxes. Difference is in absolute number of events.

Figure C.33: Bin-by-bin difference between the number of observed data events and the integral of the background model for the fit region (Left) and the full razor plane (Right). Samples correspond to the *b*-tagged MU-MU (Top), ELE-MU (middle) and ELE-ELE (Bottom) boxes. Difference is in absolute number of events.

C.11 Fit Parameters

MU Box			
component / parameter	k [GeV^{-1}]	M_R^0 [GeV]	R_0^2
$t\bar{t}$+jets 1st component	0.061 ± 0.009	100 ± 30	-0.08 ± 0.03
$W(\ell\nu)/Z(\ell\ell)$+jets 1st component	0.081 ± 0.010	280 ± 30	-0.14 ± 0.02
UEC 2nd component	0.019 ± 0.002	40 ± 30	-0.27 ± 0.02
ELE Box			
component / parameter	k [GeV^{-1}]	M_R^0 [GeV]	R_0^2
$t\bar{t}$+jets 1st component	0.44 ± 0.007	80 ± 20	-0.13 ± 0.03
$W(\ell\nu)/Z(\ell\ell)$+jets 1st component	0.067 ± 0.014	-100 ± 70	-0.17 ± 0.05
UEC 2nd component	0.014 ± 0.003	430 ± 30	-0.27 ± 0.03
MU-MU Box			
component / parameter	k [GeV^{-1}]	M_R^0 [GeV]	R_0^2
$t\bar{t}$+jets 1st component	0.40 ± 0.015	-460 ± 80	-0.09 ± 0.06
$W(\ell\nu)/Z(\ell\ell)$+jets 1st component	0.078 ± 0.016	60 ± 20	-0.05 ± 0.03
UEC 2nd component	0.018 ± 0.002	400 ± 30	-0.27 ± 0.03
ELE-ELE Box			
component / parameter	k [GeV^{-1}]	M_R^0 [GeV]	R_0^2
$t\bar{t}$+jets 1st component	0.60 ± 0.014	-80 ± 80	-0.09 ± 0.05
$W(\ell\nu)/Z(\ell\ell)$+jets 1st component	0.02 ± 0.01	-20 ± 20	-0.03 ± 0.02
UEC 2nd component	0.014 ± 0.003	430 ± 30	-0.27 ± 0.02
ELE-MU Box			
component / parameter	k [GeV^{-1}]	M_R^0 [GeV]	R_0^2
$t\bar{t}$+jets 1st component	0.56 ± 0.008	10 ± 40	-0.14 ± 0.03
UEC 2nd component	0.019 ± 0.002	430 ± 30	-0.28 ± 0.03
HAD Box			
component / parameter	k [GeV^{-1}]	M_R^0 [GeV]	R_0^2
effective 1st component	0.056 ± 0.007	-60 ± 50	-0.09 ± 0.03
UEC MU-like component	0.019 ± 0.002	370 ± 30	-0.27 ± 0.03
UEC ELE-like component	0.017 ± 0.002	20 ± 20	-0.26 ± 0.03

Table C.15: ML fit region shape parameters for all of the boxes

C.12 Comparison of M_R and R^2 Distributions Between Data and Simulation

The distribution of the data in each box to the prediction from the Monte Carlo simulation, normalized to a luminosity of 4.6 fb^{-1} are presented in figures C.34-C.37. The inclusive (w.r.t. b-tagging requirements) event sample is shown in figures C.34 and C.35 for M_R and R^2, respectively, while the analogous plots for the subsample with at least one b-tagged jet are included in figures C.36 and C.37. The details of the simulation of these SM backgrounds can be found in section C.1.

While the normalization is as expected not well predicted, the shape is found in reasonable agreement. The disagreement at low R^2 and low M_R is due to the (not modeled) turn-on effects of the razor triggers. It is noted that the analysis does not depend on these simulated shapes. The agreement on the shapes in the region far from the trigger turn-on supports our understanding of the 2D R^2 vs. M_R distribution.
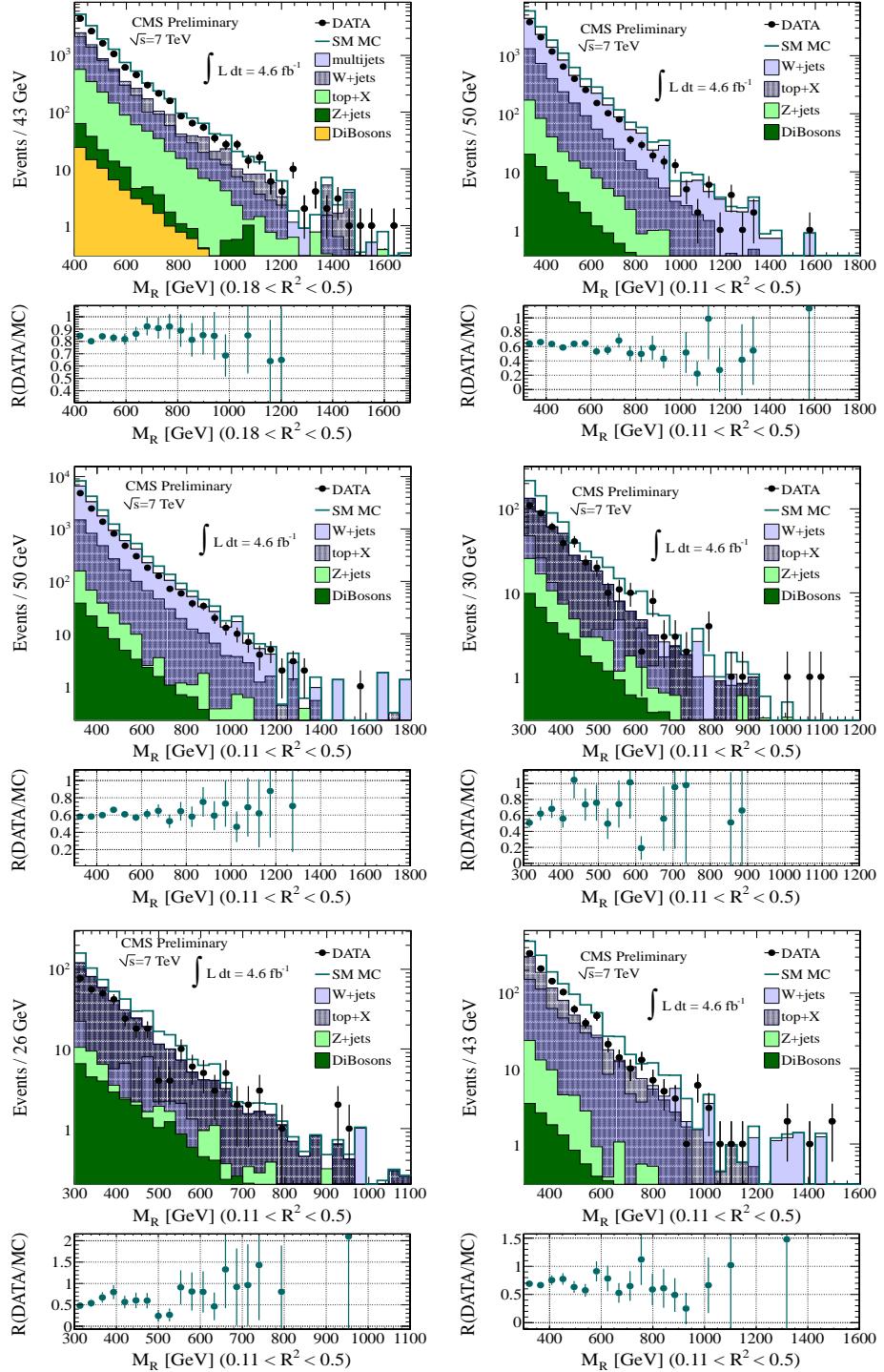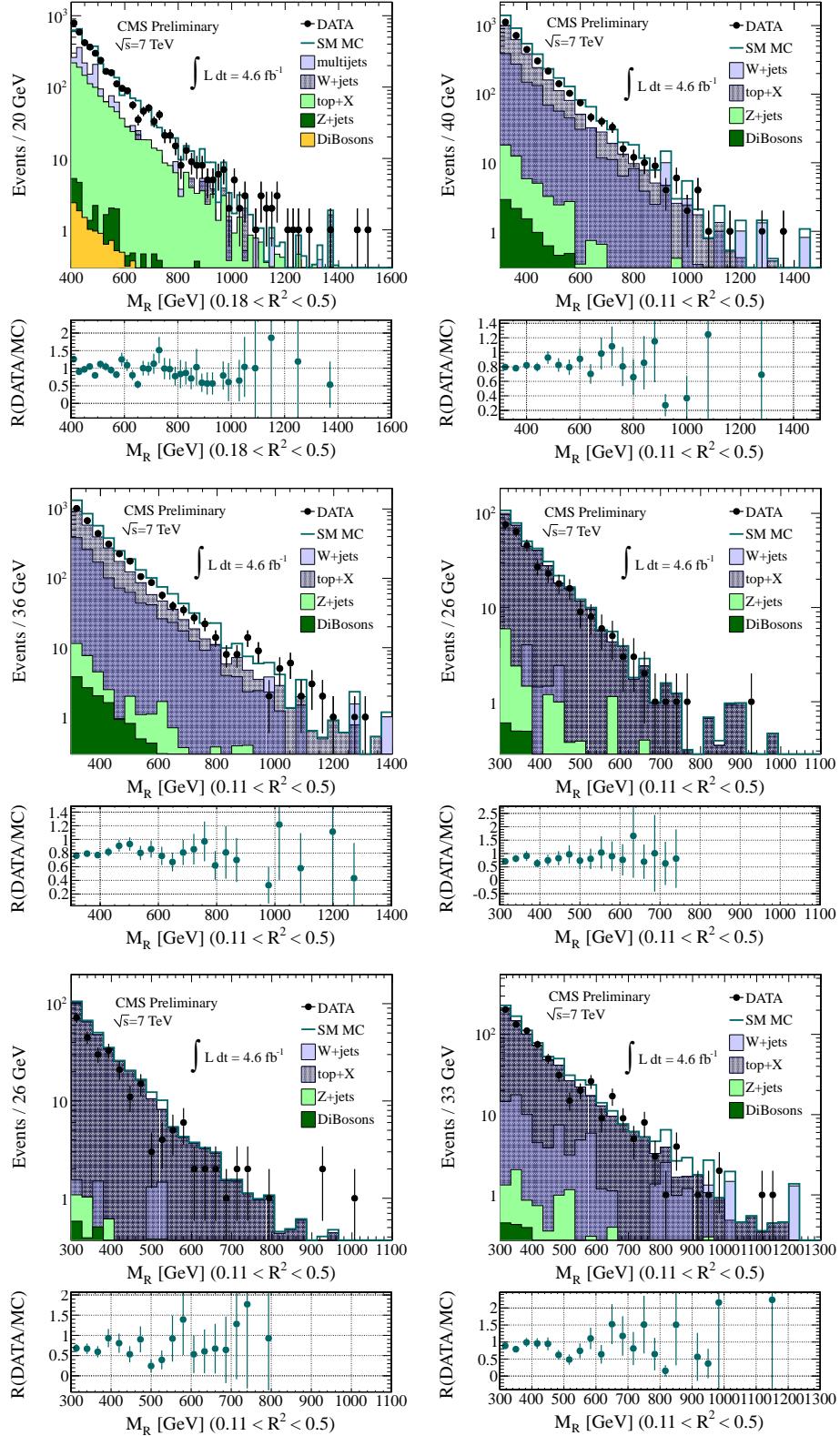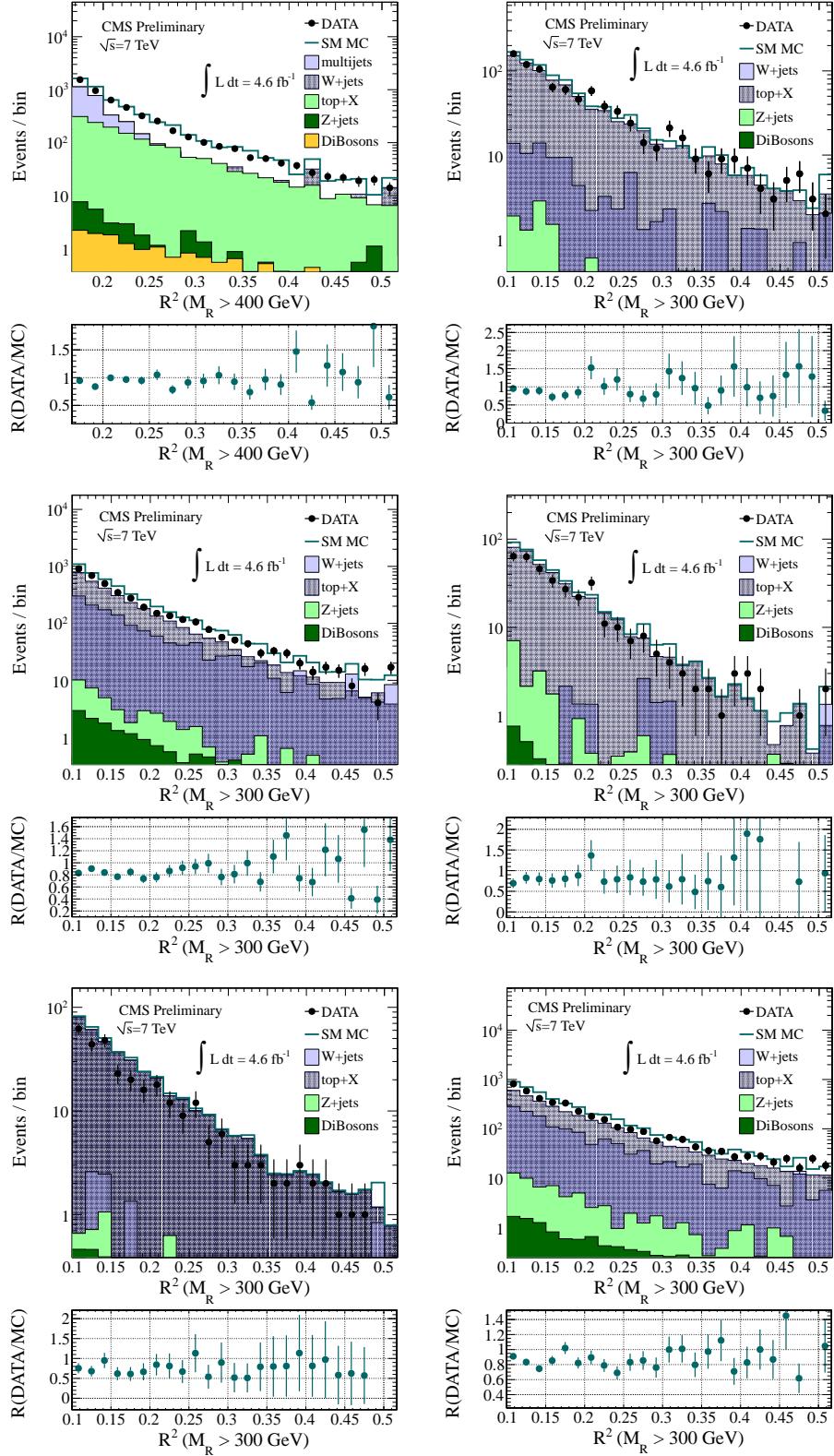
Figure C.34: Data/simulated event comparisons of the M_R distribution in the HAD (Top left), MU (Top right), ELE (center left), ELE-ELE (center right), MU-MU (Bottom left) and MU-ELE (Bottom right) boxes. A baseline selection requirement of $R^2 > 0.11$ ($R^2 > 0.18$) is applied to the leptonic boxes (HAD box).
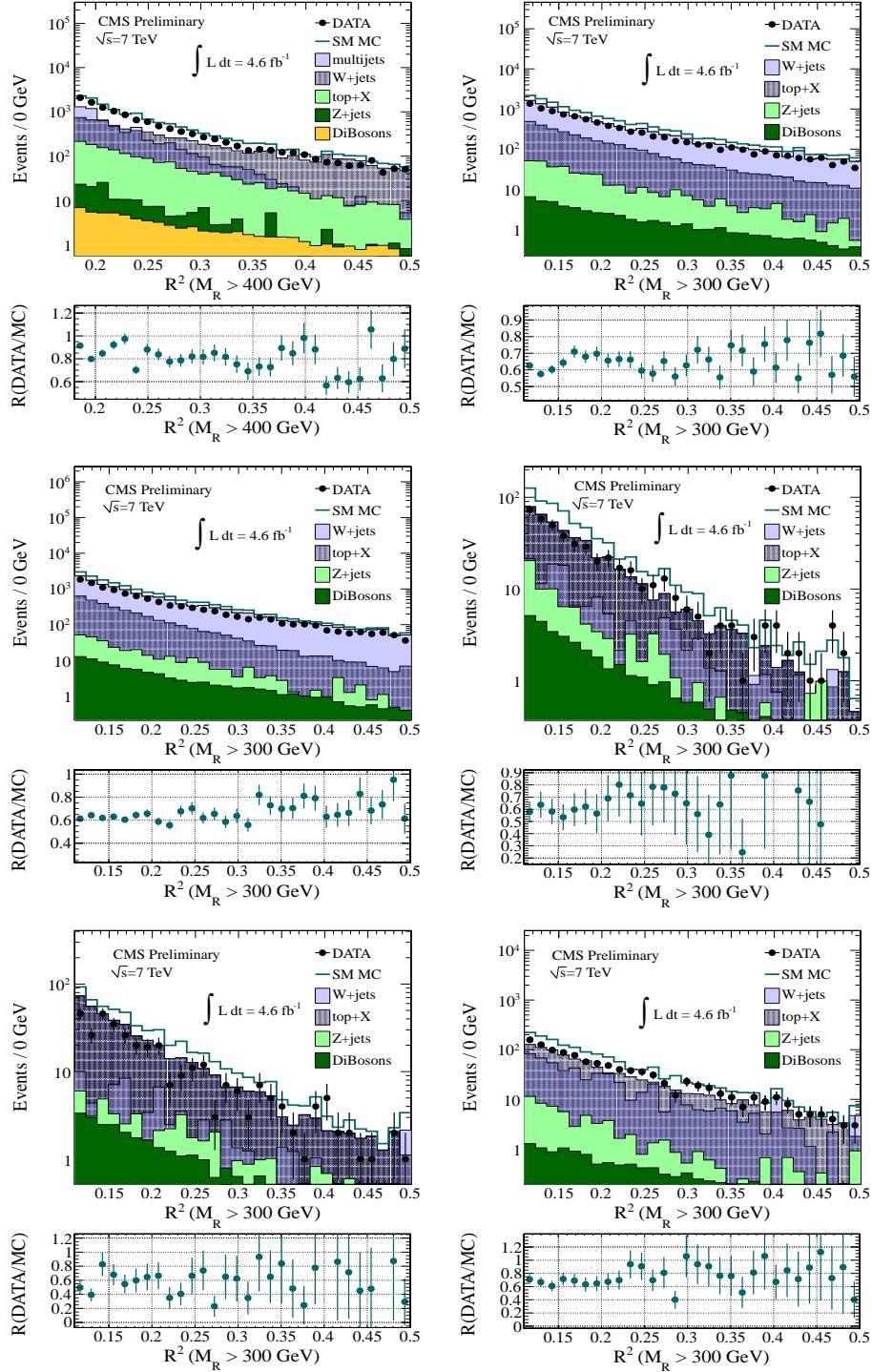
Figure C.35: Data/simulated event comparisons for the R^2 distribution in the Had (Top left), MU (Top right), ELE (center left), ELE-ELE (center right), MU-MU (Bottom left) and MU-ELE (Bottom right) boxes. A baseline selection requirement $M_R > 300$ GeV ($M_R > 400$ GeV) is applied to the leptonic boxes (HAD box).

Figure C.36: Data/simulated event comparisons of the M_R distribution in the HAD (Top left), MU (Top right), ELE (Center left), ELE-ELE (Center right), MU-MU (Bottom left) and MU-ELE (Bottom right) boxes. A baseline selection requirement of $R^2 > 0.11$ ($R^2 > 0.18$) is applied to the leptonic boxes (HAD box). Events are required to have at least one b-tagged jet.

Figure C.37: Data/simulated event comparison for the R^2 distribution in the HAD (Top left), MU (Top right), ELE (Center left), ELE-ELE (Center right), MU-MU (Bottom left) and MU-ELE (Bottom right) boxes. A baseline selection requirement $M_R > 300$ GeV ($M_R > 400$ GeV) is applied to the leptonic boxes (HAD box). Events are required to have at least one b-tagged jet.

C.13   Calculation of PDF Uncertainties for Signal Models

The systematic uncertainty on signal yields for hypothetical SUSY models is estimated by comparing three different PDF sets with the default used in sample generation, CTEQ6L1 PDF. For each PDF set, an uncertainty is obtained by evaluating the uncertainty of sub PDF sets. The uncertainty contribution for each of the alternative PDFs considered is calculated in a different way.

The cross section for an alternative PDF set is estimated by multiplying the nominal reference cross section by a weighting factor, R, calculated as

$$R = \frac{1}{N} \sum_{i=1}^{N} w_i(x_1, x_2, Q) \ , \tag{C.7}$$

where N is the number of events in the generated sample. The weight $w_i(x_1, x_2, Q)$ of the event i for a given PDF set is calculated from the formula

$$w_i(x_1, x_2, Q) = \frac{f_1^{(P)}(x_1, Q) \times f_2^{(P)}(x_2, Q)}{f_1^{(ref)}(x_1, Q) \times f_2^{(ref)}(x_2, Q)} \ , \tag{C.8}$$

where

- P is the alternative PDF being considered, MSTW2008, CTEQ6.6 or MRST2007.lomod,

- ref corresponds to the nominal PDF set, CTEQ6L1,

- x_j = Bjorken x for parton j ($j = 1, 2$),

- Q is the factorization scale,

- $f_i^{(J)}(x_j, Q)$ = PDF value for the jth parton and PDF set J.

The systematic uncertainty on the expected signal yield attributed to PDFs is calculated from the variations in R between the different alternative PDF set.

The CTEQ group [230] has developed a method in order to provide the error associated with the estimation of each of the free parameters describing PDF sets.

Additional information can be extracted from the χ^2 derivatives, in particular the Hessian matrix can give an estimation of the uncertainties coming from the error on the parameters estimation. In order to have independent errors the Hessian matrix is diagonalized and a new set of parameters is computed in this new orthogonal basis. The up-down variations of the new parameters are then computed by independently varying them in this new basis and calculating the relative $\Delta\chi^2$ variation. The main PDF groups (CTEQ, MRST, MSTW...) work in this basis to provide the central values of the d parameters describing the PDF and $2d$ parameter variations.

C.13.1 CTEQ6.6 PDF

The CTEQ6.6 PDF [230] set consists of 45 subsets, one central set and 22 sets where the eigenvector values have been varied around their best fit values. These values are varied to fall into the 90% CL. The uncertainty obtained with this set is rescaled by a factor of 1.645 to approximate the band corresponding to the desired 68% CL. The symmetric deviation of 22 sets from central value of α_s=0.118 are calculated using the following relation,

$$(\Delta R_{PDF})_{\pm,CL_{90}} = \frac{1}{2}\sqrt{\sum_{i=1}^{22}(R_i^+ - R_i^-)^2} \ . \tag{C.9}$$

C.13.2 MSTW2008 PDF

The MSTW PDF sets [163] (including NLO and NNLO sets) consists of 41 subsets, one central set and 20 sets where the eigenvector values have been varied around their best fit values. The asymmetric deviations of 20 sets from central value can be evaluated with the "modified tolerance method":

$$\Delta X_{max}^+ = \sqrt{\sum_{i=1}^{N}[max(X_i^+ - X_0, X_i^- - X_0, 0)]^2} \ , \tag{C.10}$$

$$\Delta X^-_{max} = \sqrt{\sum_{i=1}^{N} [max(X_0 - X_i^+, X_0 - X_i^-, 0)]^2} \ . \qquad (C.11)$$

C.13.3 PDF Weight Technique

Since often it is not possible to generate the desired simulated event sample many times in order to obtain the uncertainty on an observable due to the PDF set, the most commonly used method is the one of the 'PDF weights'. This method consists of assigning, for each event generated with the central PDF from the set, a PDF weight

$$W_n^0 = 1, W_n^i = \frac{f(x_1, Q; S_i) f(x_2, Q; S_i)}{f(x_1, Q; S_0) f(x_2, Q; S_0)} \ , \qquad (C.12)$$

where $n = 1...N_{events}, i = 1...N_{PDF}$ and S_i indicates the PDF set. These weights can be calculated for multiple PDF sets in single simulated events, permitting a rudimentary re-weighting of the sample to correspond to a different PDF than the one used to generate it.

C.14 SUSY Signal Models for Results Interpretation

Several different SUSY model scenarios are considered in the interpretation of the search results.

CMSSM

A scan of benchmark scenarios for the CMSSM model with varying values $m_{1/2}$ vs m_0 and with $\tan\beta = 10$, $A_0 = 0$ and positive μ.

Simplified models (SMS)

A collection of simplified models which each include only one sparticle production and decay topology with only two sparticles in the SUSY spectrum, for a variety of values for the masses of these sparticles. The SMS models for which results are presented are illustrated in figure C.38 and described below:
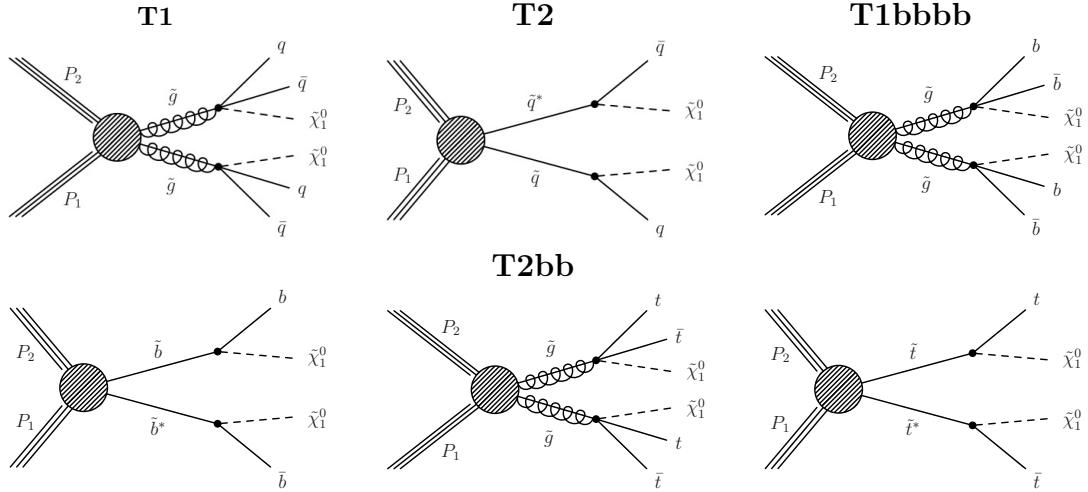
Figure C.38: SMS models considered in this analysis, with accompanying labels.

- **T1:** Di-gluino production, with each gluino undergoing a three-body decay to two light quarks and a neutralino.

- **T2:** Di-squark production, with each squark decaying to a light quark and a neutralino.

- **T1bbbb:** Di-gluino production, with each gluino undergoing a three-body decay to two b-quarks and a neutralino.

- **T2bb:** Di-sbottom production, with each sbottom decaying to a b-quark and a neutralino.

- **T1tttt:** Di-gluino production, with each gluino undergoing a three-body decay to two top quarks and a neutralino.

- **T2tt:** Di-stop production, with each stop decaying to a top quark and a neutralino.

C.14.1 Selection Efficiencies for Signal Models

C.14.1.1 CMSSM
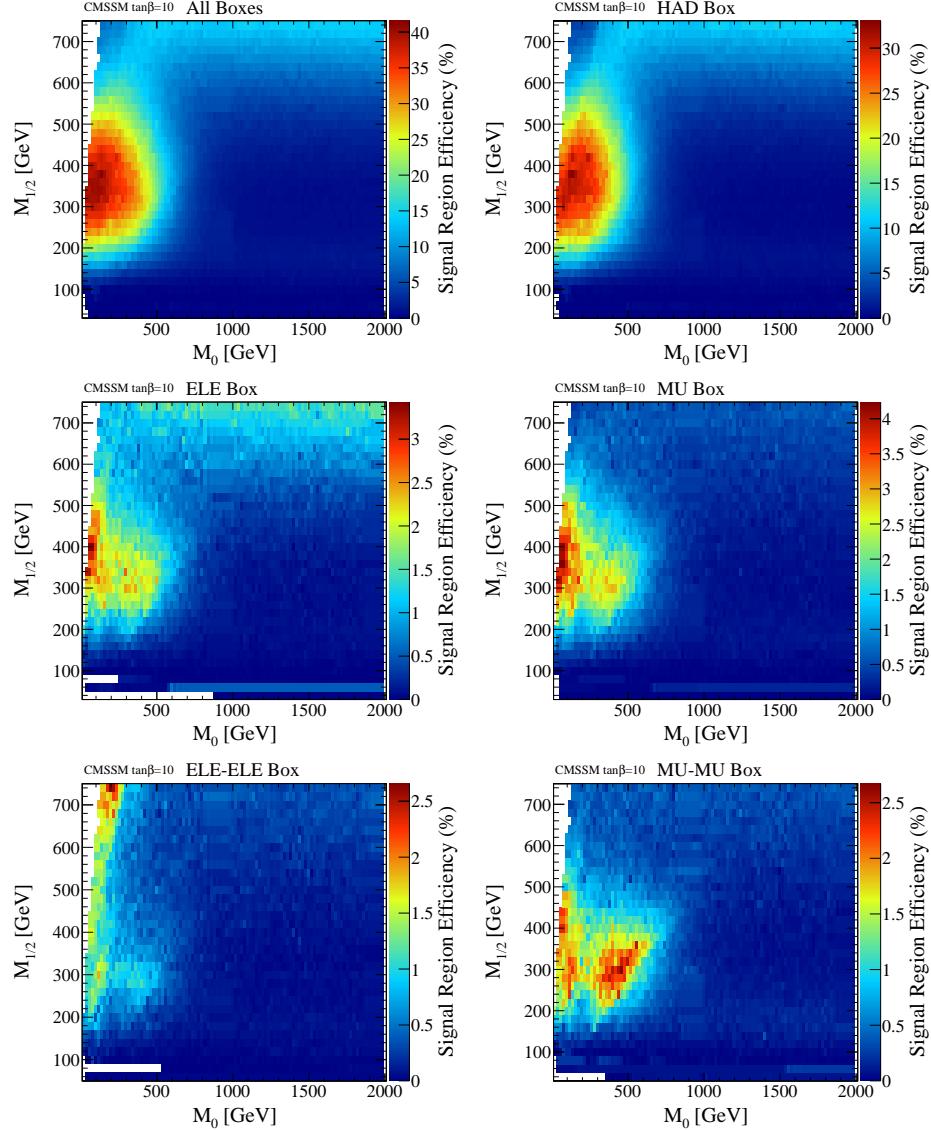
Figure C.39: Signal selection efficiency across the $m_{1/2}$ vs m_0 plane in the CMSSM for the sum of the six inclusive analysis boxes (Top left), the Had box alone (Top right), the ELE box (Middle left), the MU box (Middle right), the ELE-ELE box (Bottom left) and the MU-MU box (Bottom right). The efficiency is computed with respect to an inclusive sample of simulated events for a given CMSSM point and it includes the events populating the signal regions of the six boxes. Despite the fact that the CMSSM favors the hadronic final state, by adding the leptonic boxes an important fraction of events is recovered, particularly due to looser fit region definitions.
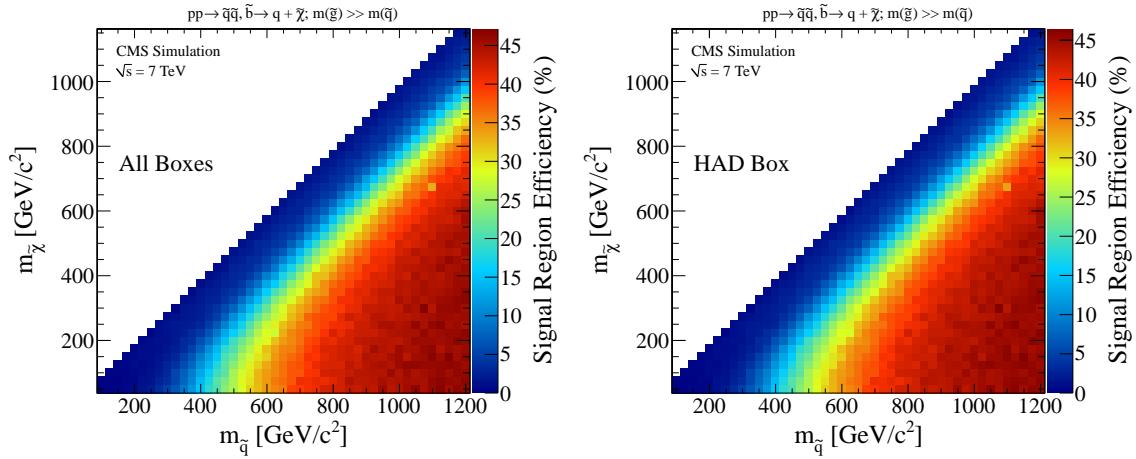
C.14.1.2 Simplified Models
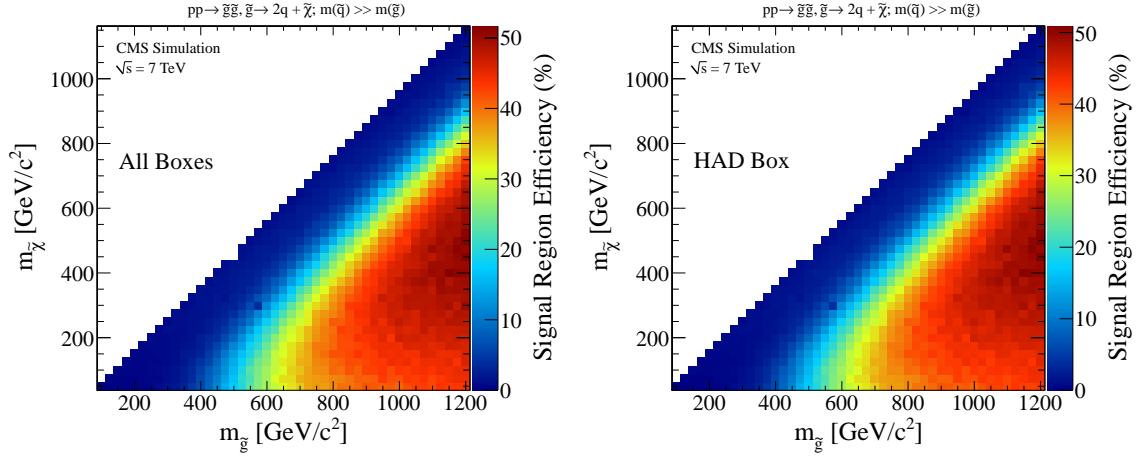
Figure C.40: Signal selection efficiency for SMS T1 in the inclusive analysis for all of the six boxes (Left) and the HAD box (Right). In this model, gluinos ($\tilde{q}$) are pair-produced and decay into two light quark jet sand a neutralino ($\tilde{\chi}$). Efficiency contours closely follow contours of $M_\Delta = [m_{\tilde{g}}^2 - m_{\tilde{q}}^2]/2m_{\tilde{g}}$ due to its correspondence with M_R and the minimum requirements on that variable in the definition of the signal regions.

Figure C.41: Signal selection efficiency for SMS T2 in the inclusive analysis for all of the six boxes (Left) and the HAD box (Right). In this model, gluinos ($\tilde{q}$) are pair-produced and decay into a light quark jet and a neutralino ($\tilde{\chi}$). Efficiency contours closely follow contours of $M_\Delta = [m_{\tilde{q}}^2 - m_{\tilde{q}}^2]/2m_{\tilde{g}}$ due to its correspondence with M_R and the minimum requirements on that variable in the definition of the signal regions.
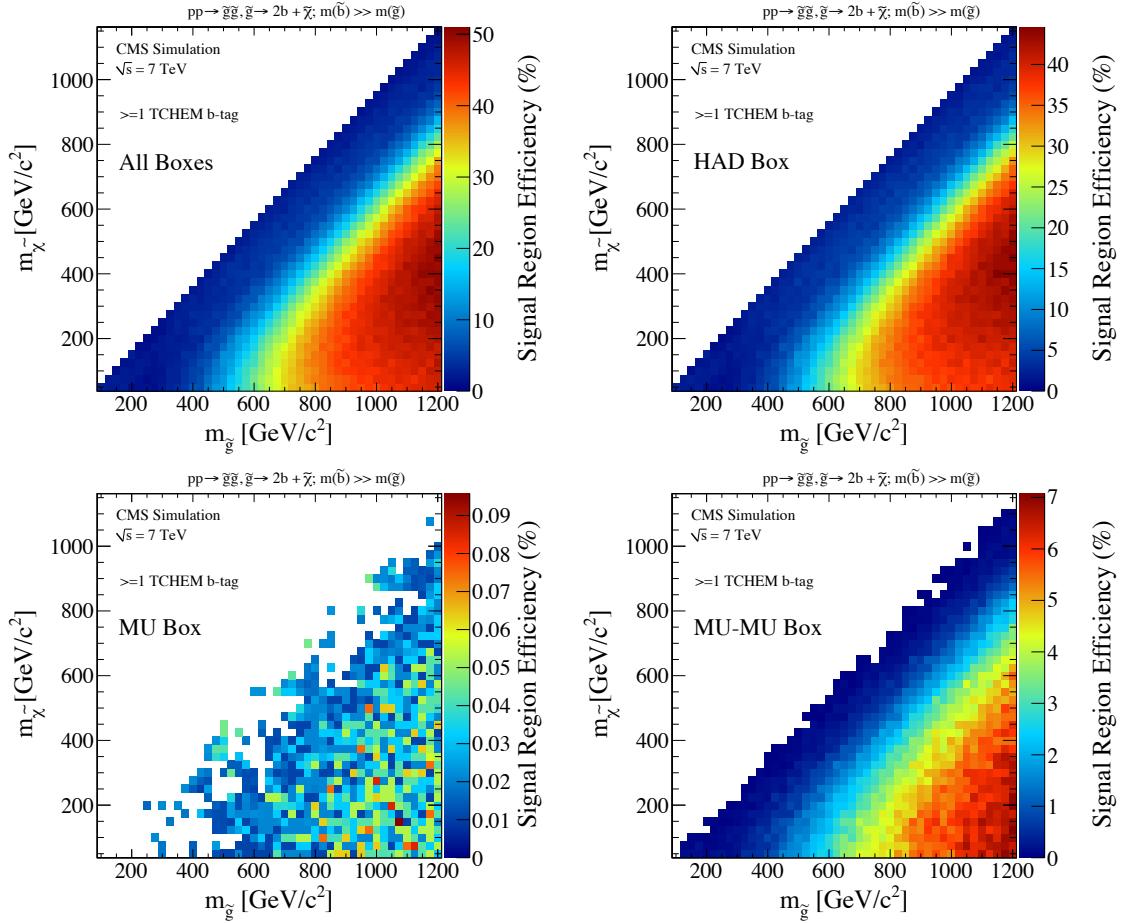
Figure C.42: Signal selection efficiency for SMS T1bbbb in the *b*-tagged jet analysis for all of the six boxes (Top left), HAD box (Top right), MU box (Bottom left) and MU-MU box (Bottom right). In this model, gluinos are pair-produced and decay into two *b*-quarks and a neutralino. An abundance of *b*-tagged jets in these events results in many non-isolated muons coming from *b*-quark decays which satisfy the MU-MU box loose second muon requirement.
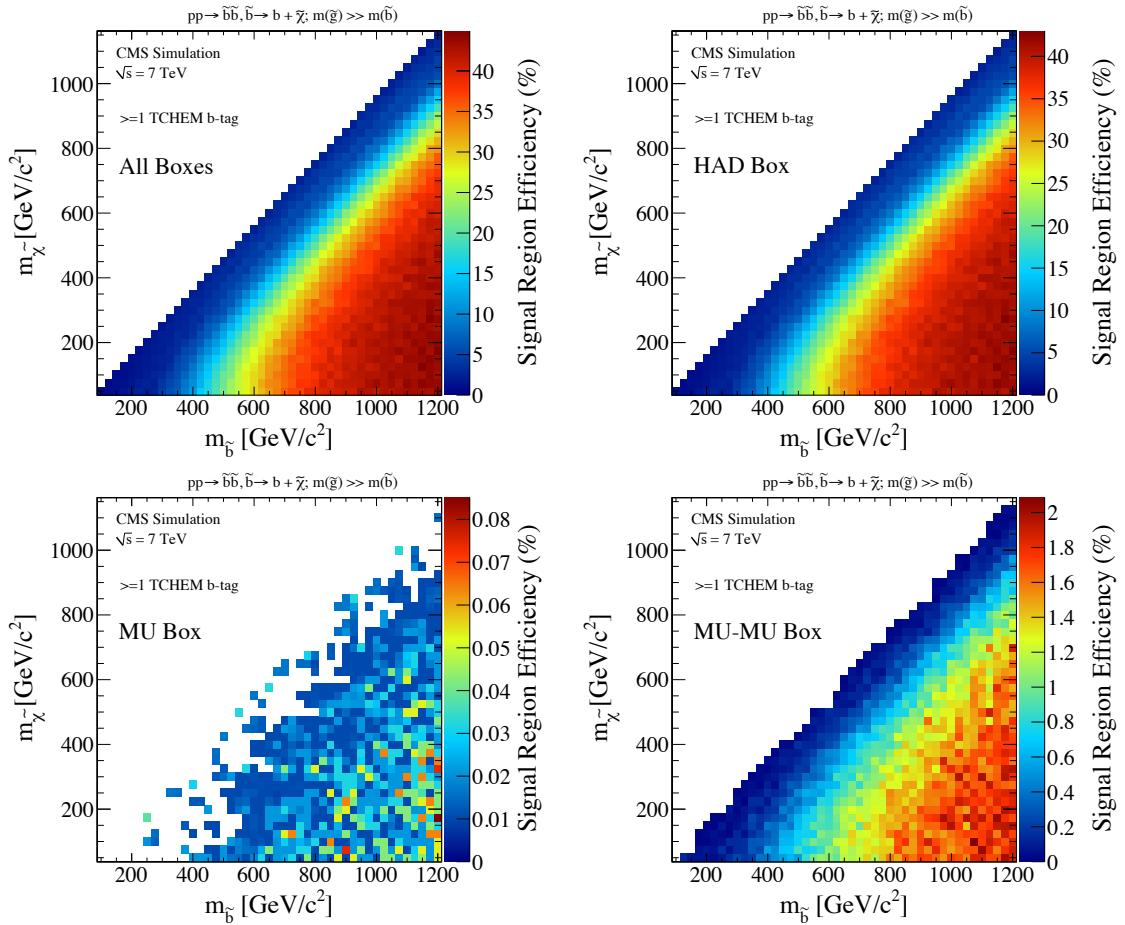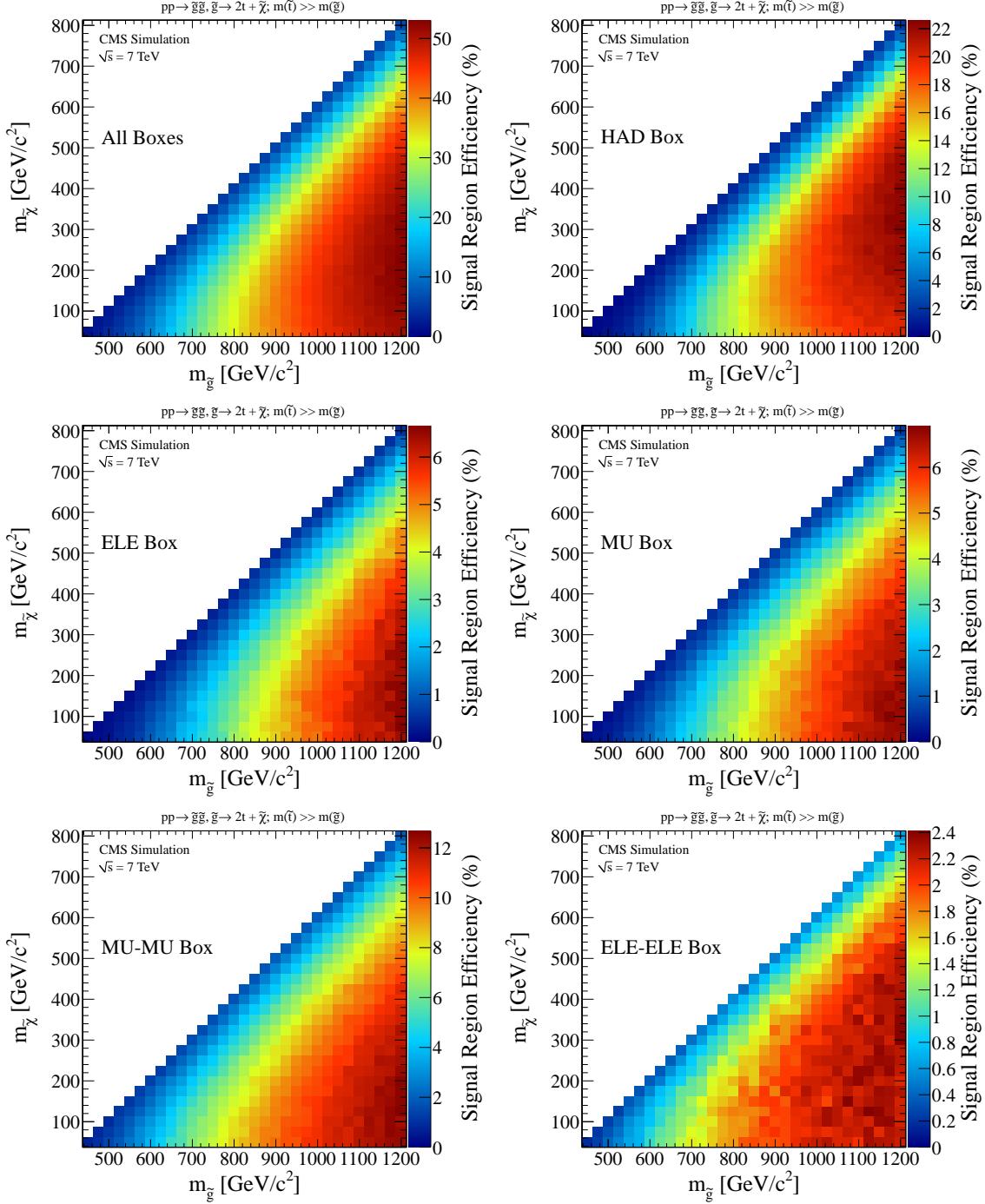
Figure C.43: Signal selection efficiency for SMS T2bb in the *b*-tagged jet analysis for all of the six boxes (Top left), HAD box (Top right), MU box (Bottom left) and MU-MU box (Bottom right). In this model, sbottoms are pair-produced and each decay into a *b*-quark and a neutralino. An abundance of *b*-tagged jets in these events results in many non-isolated muons coming from *b*-quark decays which satisfy the MU-MU box loose second muon requirement.
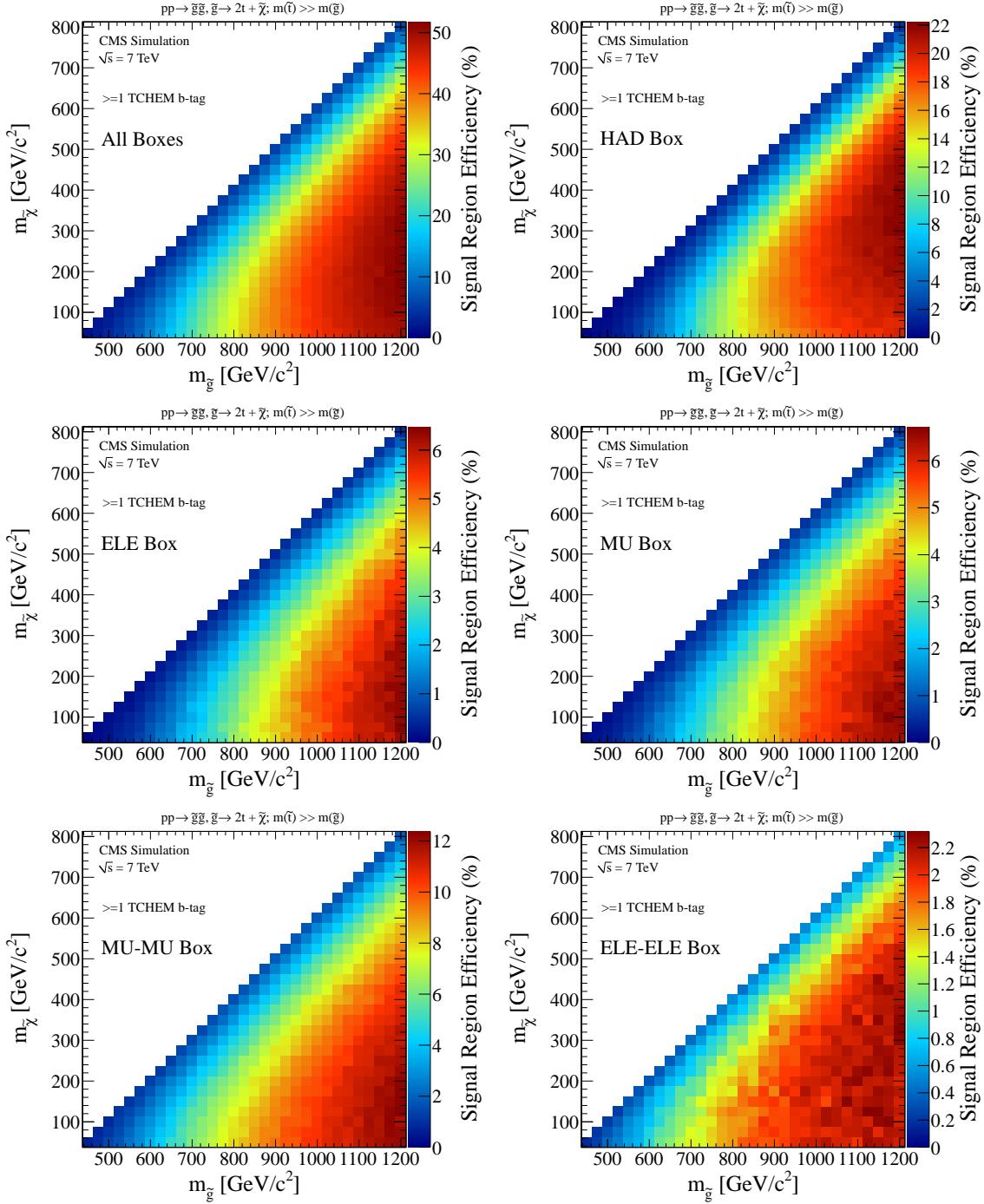
Figure C.44: Signal selection efficiency for SMS T1tttt in the inclusive jet analysis for all of the six boxes (Top left), the HAD box (Top right), the ELE box (Middle left), the MU box (Middle right), the MU-MU box (Bottom left) and the ELE-ELE box (Bottom right). In this model, gluinos are pair-produced and decay into two top quarks and a neutralino. The many tops in the final state of these events each decays to a W boson, whose leptonic decays can produce multiple leptons in the final state. Here, the box classification scheme enhances the sensitivity to this model.

Figure C.45: Signal selection efficiency for SMS T1tttt in the *b*-tagged jet analysis for all of the six boxes (Top left), the HAD box (Top right), the ELE box (Middle left), the MU box (Middle right), the MU-MU box (Bottom left) and the ELE-ELE box (Bottom right). The many tops in the final state of these events each decays to a *W* boson, whose leptonic decays can produce multiple leptons in the final state. Here, the box classification scheme enhances the sensitivity to this model. Each top also decays to a *b*-quark, meaning that the *b*-tagged jet requirement will be almost completely efficient for these events, while rejecting backgrounds with fewer sources of *b*-quarks.
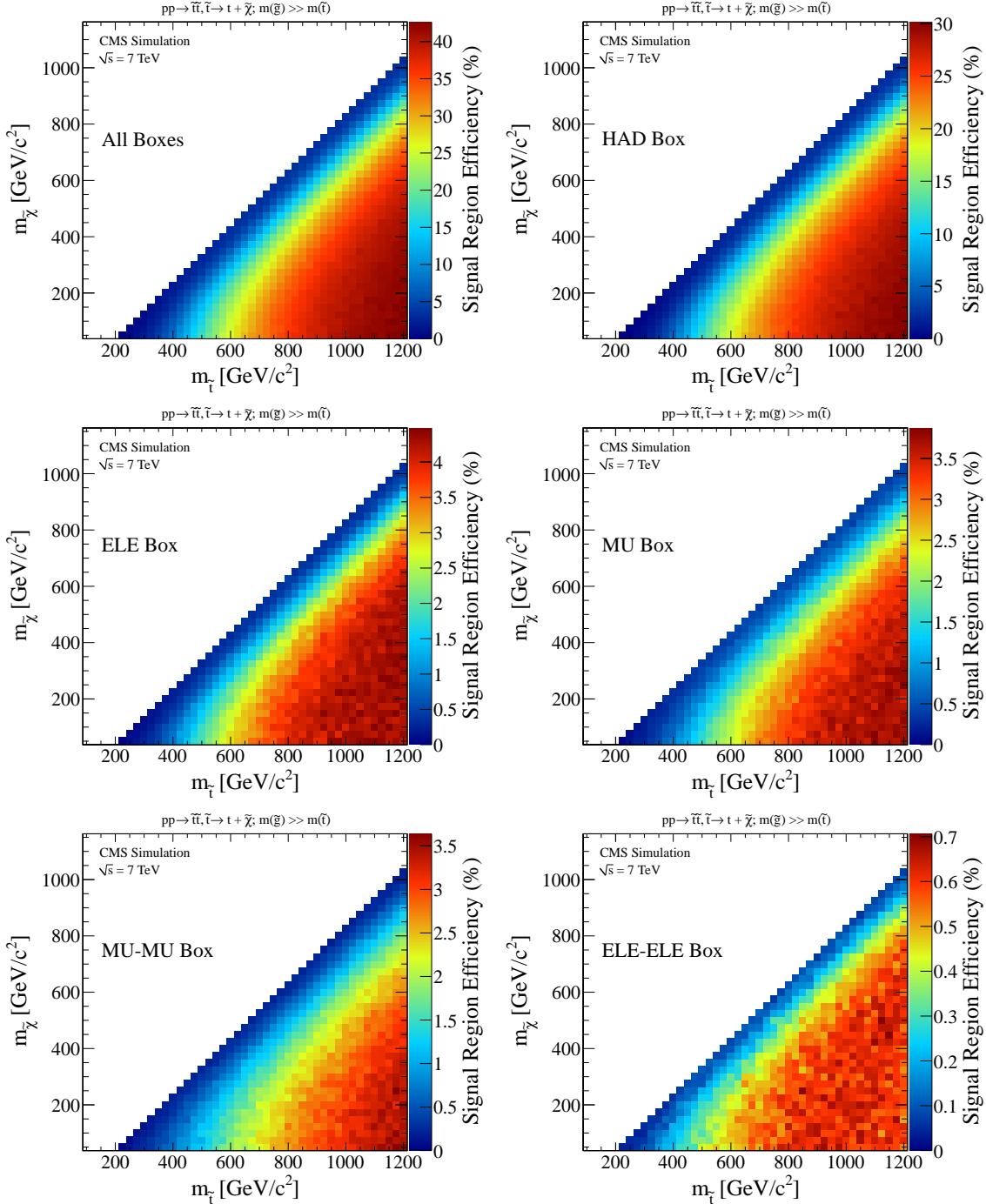
Figure C.46: Signal selection efficiency for SMS T2tt in the inclusive jet analysis for all of the six boxes (Top left), the HAD box (Top right), the ELE box (Middle left), the MU box (Middle right), the MU-MU box (Bottom left) and the ELE-ELE box (Bottom right). In this model, stops are pair-produced and decay into a top quark and a neutralino. The many two tops in the final state of these events each decays to a W boson, whose leptonic decays can result in di-leptonic final states. An abundance of b-quarks results in many non-isolated muons from semi-leptonic b-decays.
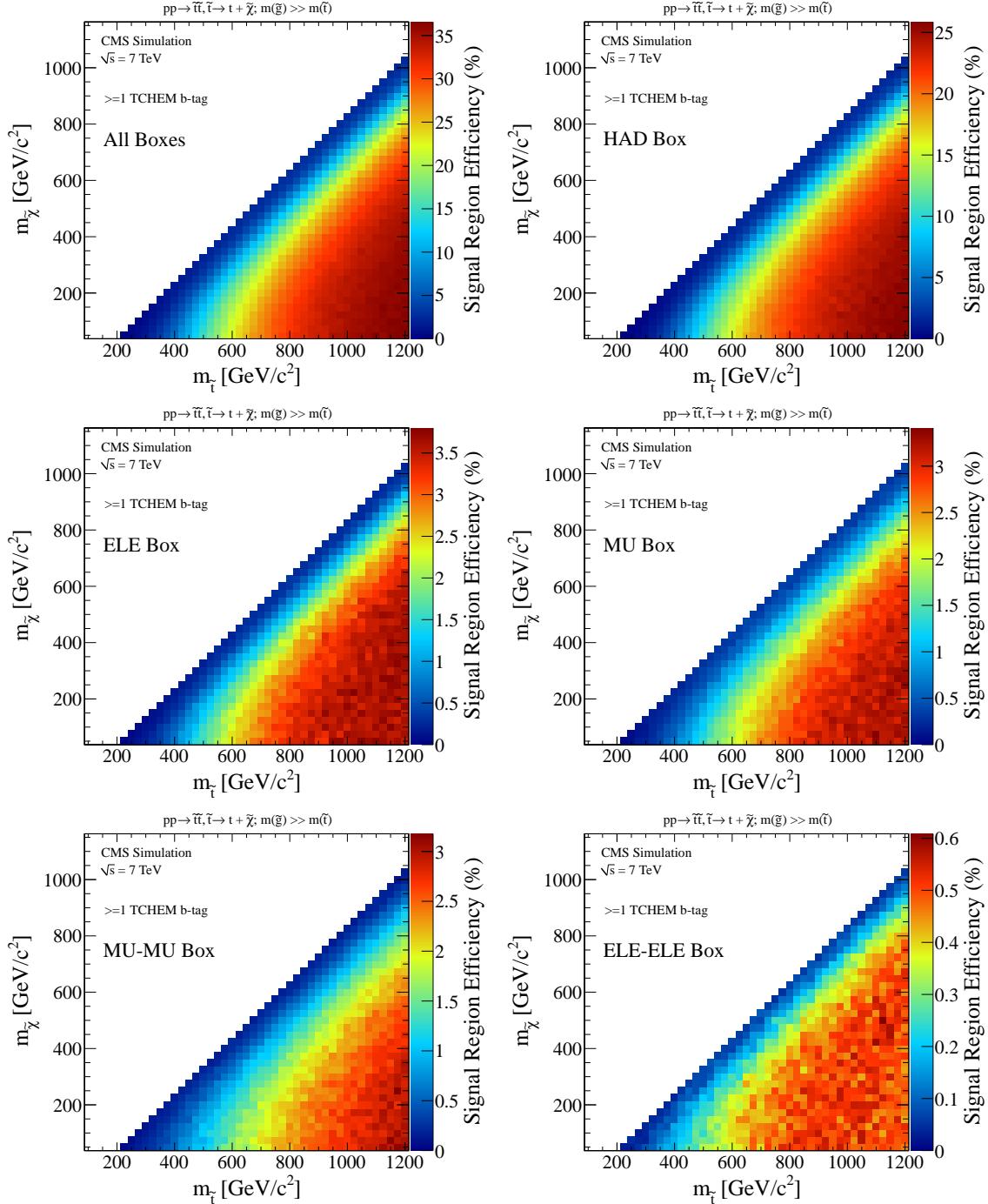
Figure C.47: Signal selection efficiency for SMS T2tt in the b-tagged jet analysis for all of the six boxes (Top left), the HAD box (Top right), the ELE box (Middle left), the MU box (Middle right), the MU-MU box (Bottom left) and the ELE-ELE box (Bottom right). The many two tops in the final state of these events each decays to a W boson, whose leptonic decays can result in di-leptonic final states. Each top also decays to a b-quark, meaning that the b-tagged jet requirement will be almost completely efficient for these events, while rejecting backgrounds with fewer sources of b-quarks.

C.14.2   Systematic Uncertainties for Signal Models

C.14.2.1   CMSSM

The size of the JES correction as well as the error associated to the model cross section (not included in the limit as a systematic) are shown in figure C.48. The JES correction error is particularly small in this analysis, due to the loose jet thresholds used with respect to the typical jet p_T at large M_R.
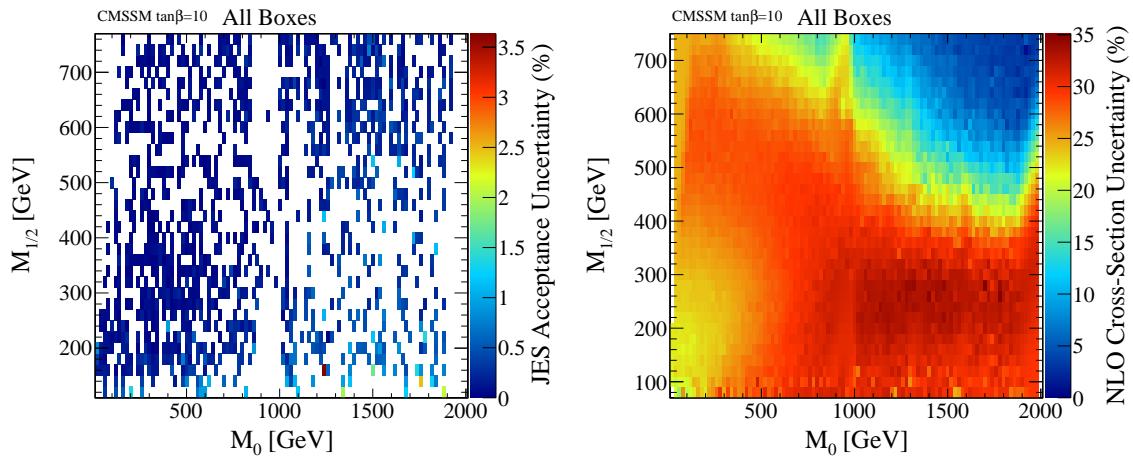

Figure C.48: (Left) Magnitude of JES systematic uncertainty on CMSSM model event yields in inclusive analysis signal regions. This is calculated by reconsidering each event with the jets' momenta and measured MET adjusted up and down according to jet-by-jet uncertainties. Total uncertainty represents the maximum changes in signal selection efficiency as a result of these variations; white indicates a negligible change. (Right) Magnitude of uncertainty on yields in signal sensitive regions from NLO cross section uncertainties.
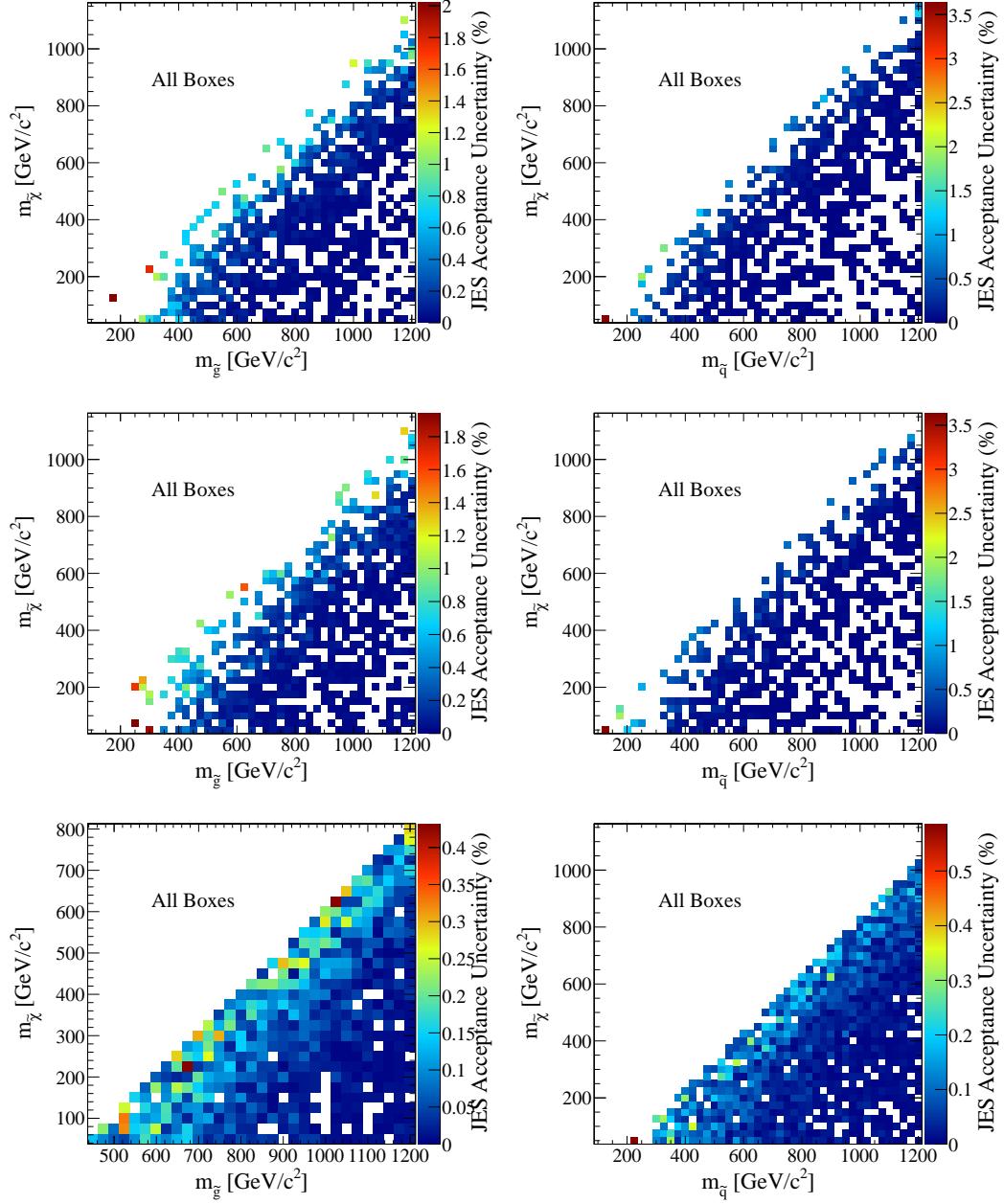
C.14.2.2 Simplified Models

Figure C.49: Magnitude of JES systematic uncertainty on SMS model event yields in inclusive analysis signal regions. (Top left) model T1, (Top right) model T2, (Middle left) model T1bbbb, (Middle right) model T2bb, (Bottom left) model T1tttt and (Bottom left) model T2tt. Uncertainty is calculated by reconsidering each event with the jets' momenta and measured MET adjusted up and down according to jet-by-jet uncertainties. Total uncertainty represents the maximum changes in signal selection efficiency as a result of these variations; white indicates a negligible change.
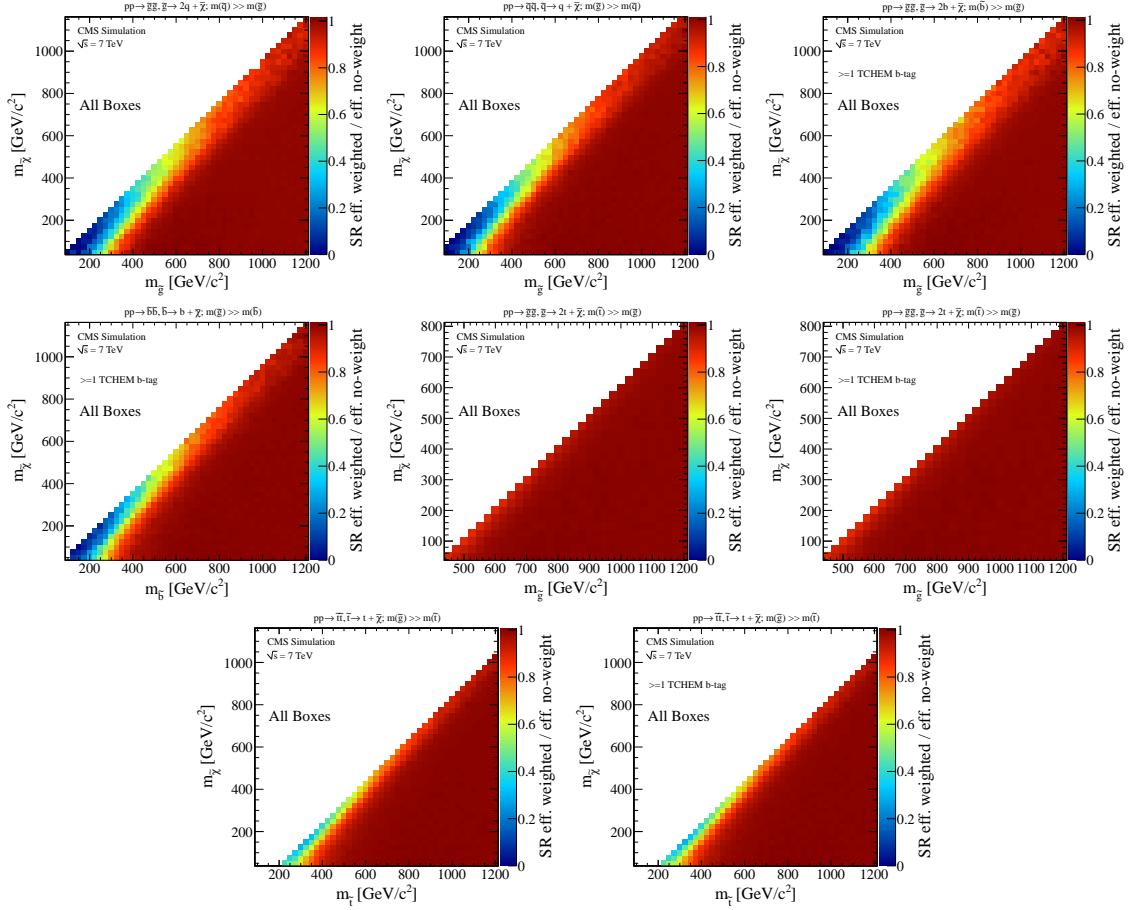
Figure C.50: Evaluation of SMS model selection efficiency sensitivity to modeling of ISR. The selection efficiency is compared between events generated with different variations of ISR magnitude parameters. The ratio of these efficiencies for models left to right and top to bottom: T1, T2, T1bbbb, T2bb, T1tttt (inclusive analysis), T1tttt (*b*-tagged analysis), T2tt (inclusive analysis) and T2tt (*b*-tagged) analysis). Cross section upper limits are not quoted if ratio of efficiencies is below 0.75.

C.14.3 Evaluation of Potential Signal Contamination in Fit Region
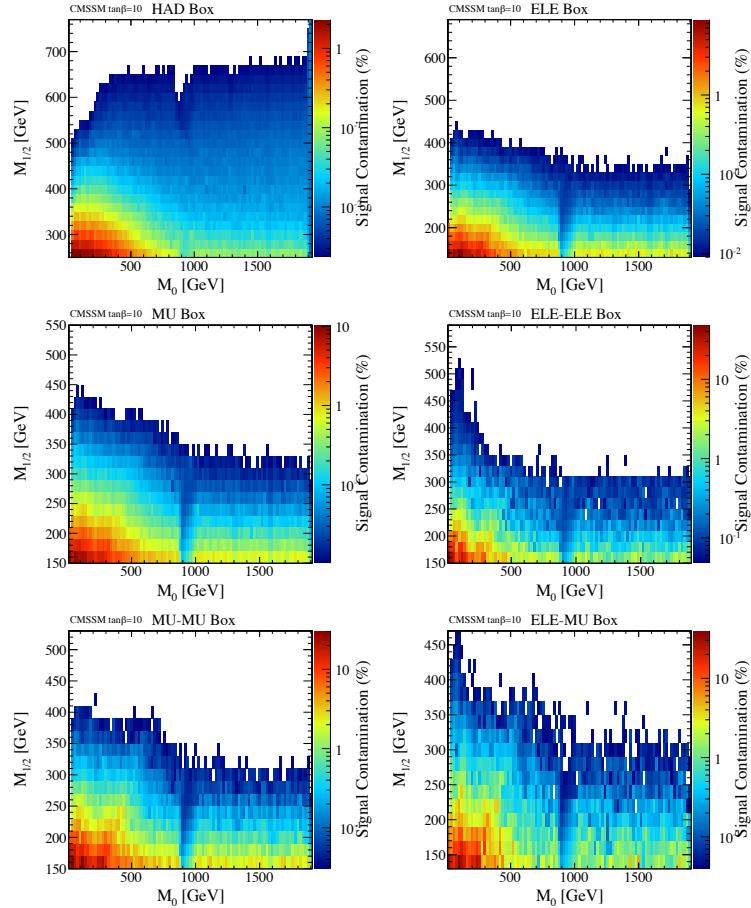
C.14.3.1 CMSSM

Figure C.51: If too many signal events are selected in the fit regions of the various boxes, they could potentially bias the SM background shape and yield estimation performed through fits. In order to quantify the size of any potential bias, we consider the magnitude of expected contamination for each of the boxes for CMSSM models, using expected signal model cross sections. The size of this signal event contamination is shown for the HAD box (Top left), the ELE box (Top right), the MU box (Top right), the ELE-ELE box (Bottom left), the MU-MU box (Bottom center) and the ELE-MU box (Bottom right). White spaces indicate precent contamination is smaller than axis interval. The expected contamination is observed to be negligible in the region of parameter space near the observed limit. The lower $m_{1/2}/m_0$ region is excluded from previous searches (see chapter 10).
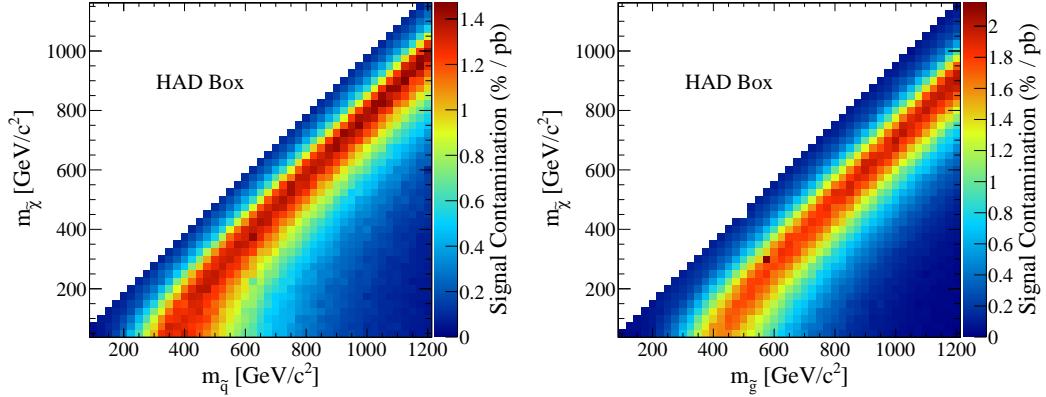
C.14.3.2   Simplified Models

Figure C.52: Magnitude of expected signal contamination in the fit region for SMS models T2 (Left) and T1 (Right), expressed as a percent w.r.t. the number of observed events in the fit region in data, per pb of signal cross section. White spaces indicate precent contamination is smaller than axis interval. For cross sections near the observed upper limits for these models (fractions of a pb at maximum) the potential signal contamination is negligible.
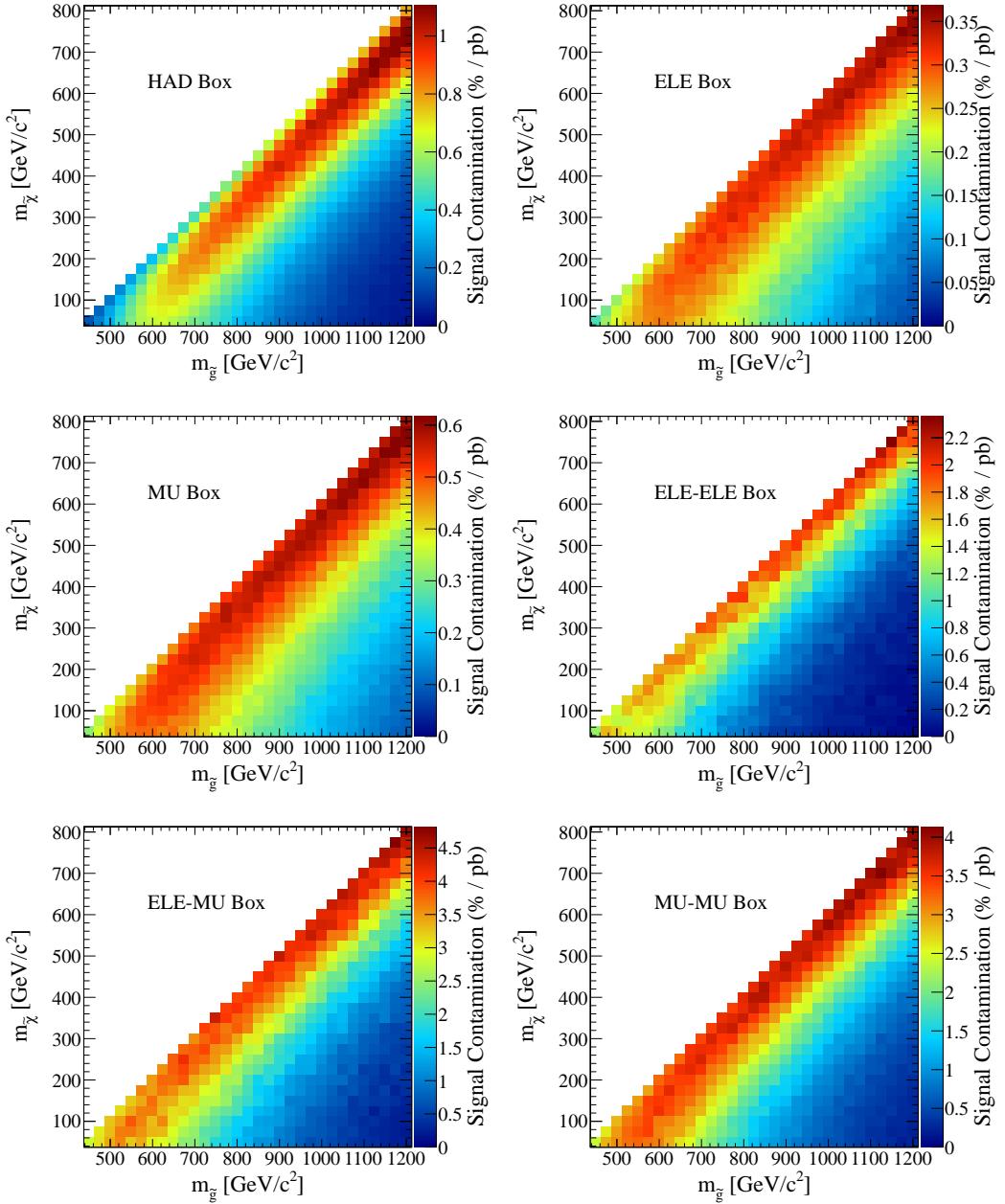
Figure C.53: Magnitude of expected signal contamination in the fit region for SMS model T1tttt analyzed with the inclusive analysis, expressed as a percent w.r.t. the number of observed events in the fit region in data, per pb of signal cross section, for the HAD (Top left) ELE (Top right), MU (Center left), ELE-ELE (Center right), ELE-MU (Bottom left), and MU-ME (Bottom right) boxes. White spaces indicate precent contamination is smaller than axis interval. For cross sections near the observed upper limits for these models (fractions of a pb at maximum) the potential signal contamination is negligible.
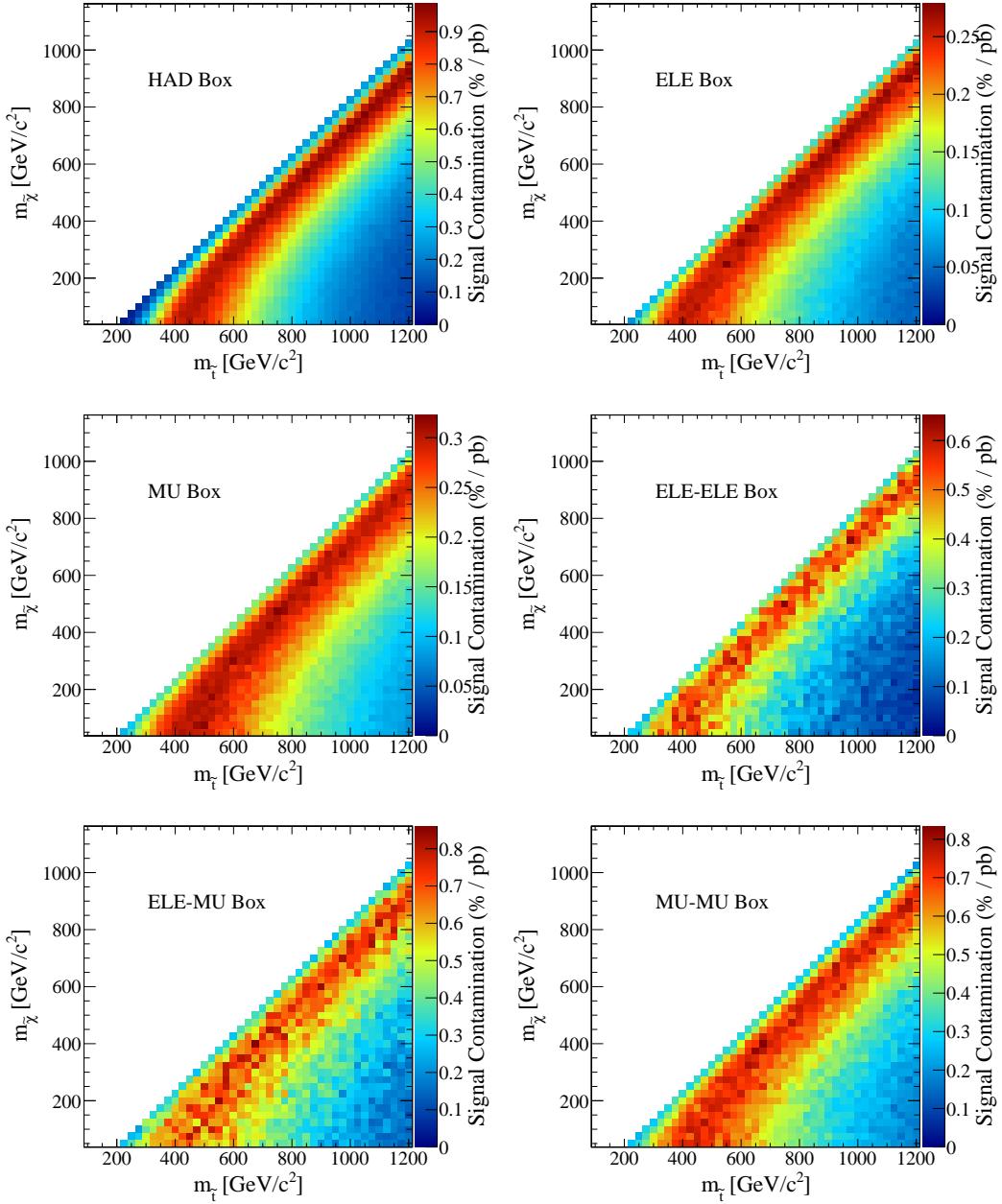
Figure C.54: Magnitude of expected signal contamination in the fit region for the SMS model T2tt, expressed as a percent w.r.t. the number of observed events in the fit region in inclusive analysis data, per pb of signal cross section, for the HAD (Top left) ELE (Top right), MU (Center left), ELE-ELE (Center right), ELE-MU (Bottom left), and MU-ME (Bottom right) boxes. White spaces indicate precent contamination is smaller than axis interval. For cross sections near the observed upper limits for these models (fractions of a pb at maximum) the potential signal contamination is negligible.

Bibliography

[1] C. D. Anderson, *Am. J. Phys.*, **29** (1961) 825.

[2] M. Gell-Man and Y. Ne'eman, *The Eightfold Way*, (1964).

[3] H. Politzer, *Phys. Rev. Lett.*, **30** (1973) 1346.

[4] D. Gross and F. Wilczek, *Phys. Rev. Lett.*, **30** (1973) 1343.

[5] S. L. Glashow, *Nucl. Phys.*, **22** (1961) 579.

[6] S. Weinberg, *Phys. Rev. Lett.*, **19** (1967) 1264.

[7] CMS Collaboration, *Phys. Lett. B* **716** (2012) 30.

[8] ATLAS Collaboration, *Phys. Lett. B* **716** (2012) 1.

[9] M. E. Peskin and D. V. Schroeder, *Introduction to quantum field theory*, (1995).

[10] P. W. Anderson, *Phys. Rev.*, **130** (1963) 439.

[11] G. S. Guralnik, C. R. Hagen, and T. W. B. Kibble, *Phys. Rev. Lett.*, **13** (1964) 585.

[12] T. W. B. Kibble, *Phys. Rev.*, **155** (1967) 1554.

[13] P. W. Higgs, *Phys. Rev. Lett.* **13** (1964) 508.

[14] F. Englert and R. Brout, *Phys. Rev. Lett.*, **13** (1964) 321.

[15] P. W. Higgs, *Phys. Rev.*, **145** (1966) 1156.

[16] CMS Collaboration, *J. Instrum.* **3** (2008) 208004.

[17] L. Evans, (ed.) and P. Bryant (ed.), *J. Instrum.* **3** (2008) S08001.

[18] CMS Collaboration, *CERN-LHCC* **97** (1997) 010.

[19] A. Herve et al., *IEEE Trans. Appl. Supercond.* **10** (2000) 389.

[20] A. Herve et al., *IEEE Trans. Appl. Supercond.* **14** (2004) 524.

[21] F. Kircher et al., *IEEE Trans. Appl. Supercond.* **10** (2000) 407.

[22] CMS Collaboration, *CERN-LHCC* **2006** 001.

[23] CMS Collaboration, *CERN-LHCC* **98** 006.

[24] CMS Collaboration, *CERN-LHCC* **2000** 016.

[25] CMS Collaboration, *Physics Analysis Summary* **CMS-PAS**-TRK-11-001.

[26] CMS Collaboration, *CERN-LHCC* **97** 032.

[27] CMS Collaboration, *Physics Analysis Summary* **CMS-PAS**-MUO-11-001.

[28] G. Charpak et al., *Nucl. Instrum. Meth.* **167** (1979) 455.

[29] CMS Collaboration, *CMS Physics Analysis Summary* **CMS-PAS**-MUO-10-004.

[30] CMS Collaboration, *CMS Physics Analysis Summary* **CMS-PAS**-BTV-11-004.

[31] CMS Collaboration, *CERN-LHCC* **97** 033.

[32] CMS Collaboration, *JINST* **5** (2010) T03001.

[33] S. Baccaro et al., *Nucl. Instrum. Meth. A* **426** (1999) 206-211.

[34] Z. Antunovic et al., *Nucl. Instr. Meth. A* **537** (2005) 379.

[35] K. W. Bell et al., *IEEE Trans. Nucl. Sci.* **51** (2004) 2284.

[36] Particle Data Group Collaboration, *Phys. Lett. B* **592** (2004) 1.

[37] P. Lecoq et al., *Nucl. Instrum. Meth. A* **365** (1995) 291.

[38] A. A. Annenkiv, M. V. Korzhik, and P. Lecoq, *Nucl. Instrum. Meth. A* **490** (2002) 30.

[39] X. D. Qu, L. Y. Zhang, and R. Y. Zhu, *IEEE Trans. Nucl. Sci.* **47** (2000) 1741.

[40] P. Adzic et al., *JINST* **2** (2007) P04004.

[41] CMS Collaboration, *CMS Physics Analysis Summary* **CMS-PAS**-EGM-10-003.

[42] CMS Collaboration, *CMS Physics Analysis Summary* **CMS-PAS**-EGM-11-001.

[43] CMS Collaboration, *JINST* **5** (2010) P03010.

[44] A. Van Lysebetten and P. Verrecchia, *CMS Note* **2004-001**.

[45] A. Ghezzi et al., *CMS Note* **2006-038**.

[46] P. Bonamy et al., *CMS Note* **1998-013**.

[47] G. Daskalakis et al., *CMS Note* **2006-037**.

[48] Particle Data Group Collaboration, *J. Phys. G* **33** (2006) 258-267.

[49] A. Benaglia et al., *CMS Note* **2007-001**.

[50] X. Qu et al., *CMS Note* **2007-028**.

[51] M. Raymond et al., *Proc. of the 9th workshop on electronics for the LHC experiments* CERN-LHCC-2003-005 (2003).

[52] P. Adzic et al., *Eur. Phys. J* **C46S1** (2006) 23.

[53] CMS Collaboration, *CMS Physics Analysis Summary* **CMS-PAS**-EGM-10-004.

[54] P. Cushman, A. Heering and A. Ronzhin, *Nucl. Instrum. Meth. A* **442** (2000) 289.

[55] CMS Collaboration, *CERN-LHCC* **97** 031.

[56] N. Akchurin and R. Wigmans, *Rev. Sci. Instrum.* **74** (2002) 2955.

[57] M. Cacciari, G. P. Salam and G. Soyez, *JHEP* **0804:063** (2008).

[58] M. Cacciari and G. P. Salam, *Phys. Lett. B* **659** (2007) 119.

[59] M. Cacciari, G. P. Salam and G. Soyez, *JHEP* **04** (2007).

[60] CMS Collaboration, *CMS Physics Analysis Summary* **CMS-PAS**-JME-10-001.

[61] CMS Collaboration, *Physics Analysis Summary* **CMS-PAS**-PFT-09-001.

[62] CMS Collaboration, *Physics Analysis Summary* **CMS-PAS**-PFT-10-002.

[63] CMS Collaboration, *J. Instrum* **6** (2011) 09001.

[64] CDF Collaboration, *Phys. Rev. Lett.* **88** (2002) 041801.

[65] G. Daskalakis et al., *CMS Note* **2004-024**.

[66] G. Grindhammer et al., *SLAC-PUB* **5072** (1989).

[67] H. Ita et al., arXiv:1108.2229v2 [hep-ph].

[68] Berger, C. F. et al., *Phys. Rev. Lett.* **106** (2011) 092001.

[69] Berger, C. F. et al., *Phys. Rev. D* **80** (2009) 31.

[70] R. K. Ellis, K. Melnikov, and G. Zanderighi, *Phys. Rev. D.* **80** (2009) 094002.

[71] S. D. Ellis, R. Kleiss, and W. J. Stirling, *Phys. Lett. B* **154** (1985) 435.

[72] F. A. Berends et al., *Phys. Lett. B* **224** (1989) 237.

[73] F. A. Berends, H. Kuijf, H. B. Tausk, and W. T. Giele, *Nucl. Phys. B* **357** (1991) 32.

[74] CDF Collaboration, *Phys. Rev. D* **77** (2008) 011108.

[75] CDF Collaboration, *Phys. Rev. Lett.* **100** (2008) 102001.

[76] D0 Collaboration, arXiv:1106.1457v1 [hep-ph].

[77] D0 Collaboration, *Phys. Lett. B* **678** (2009) 45.

[78] ATLAS Collaboration, *Phys. Lett. B* **698** 0370-2693 (2011) 325.

[79] CMS Collaboration, *J. High Energy Phys.* **01** (2012) 010.

[80] J. Alwall et al., *JHEP* **06** (2011) 128.

[81] A. De Rujula, J. Lykken, M. Pierini, C. Rogan, and M. Spiropulu, *Phys. Rev. D* **82** (2010) 013003.

[82] L. D. Landau, *Dokl. Akad. Nauk* **60** (1948) 207.

[83] C. N. Yanh, *Phys. Rev.* **77** (1950) 242.

[84] H. Davoudiasl, S. Gopalakrishna, E. Ponton and J. Santiago, arXiv:0908.1968 [hep-ph].

[85] A. Djouadi, *Phys. Rept.* **457** (2008) 1.

[86] A. Datta, G. L. Kane, and M. Toharia, arXiv:hep-ph/0510204.

[87] J. Hubisz, J. Lykken, M. Pierini, and M. Spiropulu, *Phys. Rev. D* **78** (2008) 075008.

[88] M. Burns, K. Kong, K. T. Matchev, and M. Park, *JHEP* **0810** (2008) 081.

[89] J. Alwall, P. Schuster, and N. Toro, *Phys. Rev. D* **79** (2009) 075020.

[90] G. Hallenbeck, M. Perelstein, C. Spethmann, J. Thom, and J. Vaughan, *Phys. Rev. D* **79** (2009) 075024.

[91] J. R. Dell'Aquila and C. A. Nelson, *Phys. Rev. D* **33** (1986) 80.

[92] C. A. Nelson, *Phys. Rev. D* **37** (1988) 1220.

[93] T. Matsuura and J. J. van der Bij, *Z. Phys. C* **51** (1991) 259.

[94] A. Soni and R. M. Xu, *Phys. Rev. D* **48** (1993) 5259.

[95] D. Chang, W. Y. Keung, and I. Phillips, *Phys. Rev. D* **48** (1993) 3225.

[96] V. D. Barger, K. M. Cheung, A. Djouadi, B. A. Kniehl, and P. M. Zerwas, *Phys. Rev. D* **49** (1994) 79.

[97] A. Skjold and P. Osland, *Phys. Lett. B* **311** (1993) 261.

[98] B. Grzadkowski and J. F. Gunion, *Phys. Lett. B* **350** (1995) 218.

[99] T. Plehn, D. L. Rainwater, and D. Zeppenfeld, *Phys. Rev. Lett.* **88** (2002) 051801.

[100] S. Y. Choi, D. J. Miller, M. M. Muhlleitner, and P. M. Zerwas, *Phys. Lett. B* **553** (2003) 61.

[101] S. S. Biswal, R. M. Godbole, R. K. Singh, and D. Choudhury, *Phys. Rev. D* **73** (2006) 035001.

[102] R. M. Godbole, D. J. Miller, and M. M. Muhlleitner, *JHEP* **0712** (2007) 031.

[103] E. Accomando et al., arXiv:hep-ph/0608079.

[104] C. P. Buszello and P. Marquard, arXiv:hep-ph/0603209.

[105] C. P. Buszello, P. Marquard, and J. J. van der Bij, arXiv:hep-ph/0406181.

[106] C. P. Buszello, I. Fleck, P. Marquard, and J. J. van der Bij, *Eur. Phys. J. C* **32** (2004) 209.

[107] W. Bernreuther, P. Gonzalez, and M. Wiebusch, arXiv:0909.3772 [hep-ph].

[108] K. Hagiwara, Q. Li, and K. Mawatari, *JHEP* **0907** (2009) 101.

[109] R. Lafaye, T. Plehn, M. Rauch, D. Zerwas, and M. Duhrssen, arXiv:0904.3866 [hep-ph].

[110] Q. H. Cao, C. B. Jackson, W. Y. Keung, I. Low, and J. Shu, arXiv:0911.3398 [hep-ph].

[111] N. Cabibbo and A. Maksymowicz, *Phys. Rev. B* **137** (1965) 438.

[112] T. Han, J. D. Lykken, and R. J. Zhang, *Phys. Rev. D* **59** (1999) 105006.

[113] A. De Rújula, J. M. K. Kaplan, and E. de Rafael, *Nuc. Phys.* **B35** (1971) 365.

[114] K. Hagiwara, T. Hatsukano, S. Ishihara, and R. Szalapski, *Nucl. Phys. B* **496** (1997) 66.

[115] L. J. Hall, and C. F. Kolda, *Phys. Lett. B* **459** (1999) 213.

[116] R. Barbieri, and A. Strumia, *Phys. Lett. B* **462** (1999) 144.

[117] G. J. Feldman and R. D. Cousins, *Phys. Rev. D* **57** (1998) 3873.

[118] J. Neyman and E. S. Pearson, *Phil. Trans. of the Royal Society of London* **A 231** (1933) 289-337.

[119] V. Ahrens, T. Becher, M. Neubert, and L. L. Yang, *Eur. Phys. J. C* **62** (2009) 333.

[120] D. de Florian and M. Grazzini, *Phys. Lett. B* **674** (2009) 291.

[121] M. Grazzini, *JHEP* **0802** (2008) 043.

[122] A. Bredenstein, A. Denner, S. Dittmaier, and M. M. Weber, *Phys. Rev. D* **74** (2006) 013004.

[123] CMS Collaboration, *CMS Physics Analysis Summary* **CMS-PAS**-HIG-12-045.

[124] D. Dicus and V. Mathur, *Phys. Rev.* **D7** (1973) 3111.

[125] B. Lee, C. Quigg and H. Thacker, *Phys. Rev.* **D16** (1977) 1519.

[126] J. Wess and B. Zumino, *Phys. Lett. B* **49** (1974) 52.

[127] P. Ramond, *Phys. Rev. D* **3** (1971) 2415.

[128] N. Arkani-Hamed, A. G. Cohen, and H. Georgi, *Phys. Lett, B* **513** (2001) 232.

[129] N. Arkani-Hamed et al., *JHEP* **0208** (2002) 021.

[130] N. Arkani-Hamed, S. Dimopoulos, and G. R. Dvali, *Phys. Lett. B* **429** (1998) 263.

[131] L. Randall and R. Sundrum, *Phys. Rev. Lett.* **83** 3370.

[132] S. Dimopoulos, H. Georgi, *Nucl. Phys. B* **193** (1981) 150.

[133] A. Chamseddine, R. Arnowitt, P. Nath, *Phys. Rev. Lett.* **49** (1982) 970-974.

[134] N. Arkani-Hamed, A. G. Cohen, E. Katz, and A. E. Nelson, *JHEP* **0207** (2002) 034.

[135] M. Schamltz and D. Tucker-Smith, *Ann. Rev. Nucl. Part. Sci.* **55** (2005) 229.

[136] J. Hubisz, S. J. Lee, and G. Paz, *JHEP* **0606** (2006) 041.

[137] M. Asano, S. Matsumoto, N. Okada, and Y. Okada, *Phys. Rev. D.* **75** (2007) 063506.

[138] M. S. Carena, J. Hubisz, M. Perelstein, and P. Verdier, *Phys. Rev. D* **75** (2007) 091701.

[139] T. Appelquist, H. C. Cheng, and B. A. Dobrescu, *Phys. Rev. D* **64** (2001) 035002.

[140] H. C. Cheng, K. T. Matchev, and M. Schmaltz, *Phys. Rev. D* **66** (2002) 036005.

[141] G. Servant and T. M. P. Tait, *Nucl. Phys. B* **650** (2003) 391.

[142] H. C. Cheng, K. T. Matchev, and M. Schmaltz, *Phys. Rev. D* **66** (2002) 056006.

[143] A. Datta, K. Kong, and K. T. Matchev, *Phys. Rev. D* **72** (2005) 096006.

[144] B. A. Dobrescu, K. Kong, and R. Mahbubani, *JHEP* **0707** (2007) 012.

[145] B. A. Dobrescu, D. Hooper, K. Kong, and R. Mahbubani, *JCAP* **0710** (2007) 012.

[146] L. Randall and R. Sundrum, *Phys. Rev. Lett.* **83** (1999) 4690.

[147] I. Antoniadis, *Phys. Lett. B* **246** (1990) 377.

[148] D. E. Kaplan, G. D. Kribs, and M. Schmaltz, *Phys. Rev. D* **62** (2000) 035010.

[149] L. Randall and D. Tucker-Smith, *Phys. Rev. Lett.* **101** (2008) 221803. [arXiv:0806.1049 [hep-ph]].

[150] C. G. Lester and D. J. Summers, *Phys. Lett. B* **463** (1999) 99.

[151] B. C. Allanach, C. G. Lester, M. A. Parker, and B. R. Webber, *JHEP* **0009** (2000) 004.

[152] A. Barr, C. Lester, and P. Stephens, *J. Phys. G* **29** (2003) 2343.

[153] C. Lester and A. Barr, *JHEP* **0712** (2007) 102.

[154] B. Gripaios, *JHEP* **0802** (2008) 053.

[155] W. S. Cho, K. Choi, Y. G. Kim, and C. B. Park, *Phys. Rev. Lett.* **100** (2008) 171801.

[156] W. S. Cho, K. Choi, Y. G. Kim, and C. B. Park, *JHEP* **0802** (2008) 035.

[157] P. Konar, K. Kong, and K. T. Matchev, *JHEP* **0903** (2009) 085.

[158] K. T. Matchev, F. Moortgat, L. Pape, and M. Park, *JHEP* **0908** (2009) 104.

[159] K. T. Matchev and M. Park, arXiv:0910.1584 [hep-ph].

[160] M. M. Nojiri, Y. Shimizu, S. Okada, and K. Kawagoe, *JHEP* **0806** (2008) 035.

[161] G. Polesello and D. R. Tovey, *JHEP* **1003** (2010) 030.

[162] I. W. Kim, *Phys. Rev. Lett.* **104** (2010) 081601. [arXiv:0910.1149 [hep-ph]].

[163] A. D. Martin, W. J. Stirling, R. S. Thorne, and G. Watt, *Eur. Phys. J.* **C63** (2009) 189-285.

[164] P. Ramond, *Phys. Rev. D* **3** (1971) 2415-2418.

[165] Y. Golfand and E. Likhtman, *JHEP Lett.* **13** (1971) 323.

[166] D. Volkov and V. Akulov, *JETP Lett.* **16** (1972) 438.

[167] J. Wess and B. Zumino, *Nucl. Phys. B* **70** (1974) 39.

[168] P. Fayet, *Nucl. Phys. B* **90** (1975) 104.

[169] C. S. Rogan, arXiv:1006.2727 [hep-ph].

[170] CMS Collaboration, *CMS Physics Analysis Summary* **CMS-PAS**-MUO-10-002.

[171] CMS Collaboration, *CMS Physics Analysis Summary* **CMS-PAS**-EGM-10-004.

[172] CMS Collaboration, *PoS* **2008LHC** (2008) 104.

[173] CMS Collaboration, *J. Phys.* **G34** (2007) 995-1579.

[174] CMS Collaboration, *CMS Physics Analysis Summary* **CMS-PAS**-EWK-09-006.

[175] CMS Collaboration, *CMS Physics Analysis Summary* **CMS-PAS**-EWK-08-006.

[176] CMS Collaboration, *CMS Physics Analysis Summary* **CMS-PAS**-EWK-10-005.

[177] L. Demortier, S. Jain, and H. B. Prosper, *Phys. Rev. D* **82** (2010) 034002.

[178] D. Bourilkov, R. C. Group, and M. R. Whalley, arXiv:hep-ph/0605240.

[179] CMS Collaboration, *CMS DPS* **CERN-CMS-DP-2011-002** (2011).

[180] J. Alwall, P. Schuster, and N. Toro, *Phys. Rev.* **D79** (2009) 075020.

[181] J. Alwall, M. P. Le, M. Lisanti et al., *Int. J. Mod. Phys.* **A23** (2008) 4637.

[182] D. Alves et al., lhcnewphysics.org (2011).

[183] G. R. Farrar and P. Fayet, *Phys. Lett. B* **76** (1978) 575.

[184] D0 Collaboration, *Phys. Lett.* **B660** (2008) 449.

[185] CDF Collaboration, *Phys. Rev. Lett.* **102** (2009) 121801.

[186] ATLAS Collaboration, arXiv:1103.6208 [hep-ex].

[187] ATLAS Collaboration, arXiv:1102.5290 [hep-ex].

[188] ATLAS Collaboration, arXiv:1103.6214 [hep-ex].

[189] ATLAS Collaboration, arXiv:1102.2357 [hep-ex].

[190] CMS Collaboration, *CMS Physics Analysis Summary* **CMS-PAS**-SUS-10-005.

[191] CMS Collaboration, *Phys. Lett. B* **698** (2011) 196.

[192] CMS Collaboration, arXiv:1104.3168 [hep-ex].

[193] CMS Collaboration, arXiv:1103.1348 [hep-ex].

[194] CMS Collaboration, *Phys. Lett.* **B716** (2012) 260.

[195] CMS Collaboration, *JHEP* **1206** (2012) 169.

[196] CMS Collaboration, *Phys. Rev. Lett.* **109** (2012) 071803.

[197] CMS Collaboration, *JHEP* **1210** (2012) 018.

[198] ATLAS Collaboration, *Phys. Rev. Lett.* **108** (2012) 241802.

[199] ATLAS Collaboration, *Phys. Rev. Lett.* **108** (2012) 261804.

[200] CMS Collaboration, *JHEP* **1208** (2012) 110.

[201] ATLAS Collaboration, *Phys. Rev. D* **85** (2012) 112006.

[202] ATLAS Collaboration, *Phys. Lett. B* **714** (2012) 180.

[203] ATLAS Collaboration, *Phys. Lett. B* **714** (2012) 197.

[204] ATLAS Collaboration, *Phys. Lett. B* **715** (2012) 44.

[205] CMS Collaboration, *Phys. Rev. D* **85** (2012) 012004.

[206] R. J. Barlow, *Nucl. Instrum. Meth. A* **297** (1990) 496.

[207] G. F. de Montricher, R. A. Tapia, and J. R. Thompson, *Ann. Statist.* **3** (1975) 1329.

[208] A. L. Read, *J. Phys. G* **28** (2002) 2693.

[209] T. Junk, *Nucl. Insr. and Meth. A* **434** (1999) 435.

[210] CMS Collaboration, *CMS Physics Analysis Summary* **CMS-PAS**-SMP-12-008.

[211] CMS Collaboration, *CMS Physics Analysis Summary* **CMS-PAS**-PFT-09-001.

[212] CMS Collaboration, *CMS Physics Analysis Summary* **CMS-PAS**-JME-10-010.

[213] T. Sjostrand, S. Mrenna, and P. Skands, *JHEP* **05** (2006) 026.

[214] GEANT4 Collaboration, *Nucl. Instrum. Meth. A* **506** (2003) 250.

[215] B. C. Allanach, *Comput. Phys. Commun.* **143** (2002) 305-331.

[216] A. Djouadi, M. M. Muhlleitner, and M. Spira, *Acta Phys. Polon.* **B38** (2007) 635.

[217] P. Z. Skands et al., *JHEP* **07** (2004) 036.

[218] W. Beenakker et al., *Nucl. Phys. B* **492** (1997) 51.

[219] A. Kulesza and L. Motyka, *Phys. Rev. Lett.* **102** (2009) 111802.

[220] A. Kulesza and L. Motyka, *Phys. Rev. D* **80** (2009) 095004.

[221] W. Beenakker et al., *JHEP* **12** (2009) 041.

[222] W. Beenakker et al., *Int. J. Mod. Phys. A* **26** (2011) 2637.

[223] M. Kramer et al., arXiv:1206.2892 [hep-ph].

[224] K. Matchev and R. Remington, arXiv:1202.6580 [hep-ph].

[225] N. Arkani-Hamed et al., arXiv:hep-ph/0703088.

[226] J. Alwall et al., *Phys. Rev. D* **79** (2009) 015005.

[227] LHC New Physics Working Group Collaboration, *J. Phys. G* **39** (2012) 105005.

[228] W. Beenakker, R. Hopker, and M. Spira, "PROSPINO: A program for the PROduction of Supersymmetric Particles In Next-to-leading Order QCD," [arXiv:hep-ph/9611232].

[229] S. Frixione and B. R. Webber, *JHEP* **0206** (2002) 029.

[230] H. L. Lai et al., *Eur. Phys. J. C* **12** (2000) 375.

[231] The CMS Collaboration, *CMS Physics Analysis Summary* **CMS-PAS**-EWK-08-006.

[232] CMS Collaboration, *CMS* **TDR** 008-1.

[233] M. Pivk and F. R. Le Diberder, *Nucl. Instrum. Meth. A* **555** (2005) 356.

[234] F. Maltoni and T. Stelzer, *JHEP* **02** (2003) 027.

[235] M. L. Mangano, M. Moretti, F. Piccinini et al., *JHEP* **01** (2007) 013.

[236] S. Kretzer, H. L. Lai, F. I. Olness et al., *Phys. Rev. D* **69** (2004) 114005.