Determining SUSY model parameters and masses at the LHC using cross-sections, kinematic edges and other observables.

Christopher G. Lester, Michael A. Parker and Martin J. White

Cavendish Laboratory
Madingley Road, Cambridge CB3 0HE, U.K.
E-mail: lester@hep.phy.cam.ac.uk, parker@hep.phy.cam.ac.uk, white@hep.phy.cam.ac.uk

ABSTRACT: We address the problem of mass measurements of supersymmetric particles at the Large Hadron Collider, using the ATLAS detector as an example. By using Markov Chain sampling techniques to combine standard measurements of kinematic edges in the invariant mass distributions of decay products with a measurement of a missing p_T cross-section, we show that the precision of mass measurements at the LHC can be dramatically improved, even when we do not assume that we have measured the kinematic endpoints precisely, or that we have identified exactly which particles are involved in the decay chain causing the endpoints. The generality of the technique is demonstrated in a preliminary investigation of a non-universal SUGRA model, in which we relax the requirements of mSUGRA by breaking the degeneracy of the GUT scale gaugino masses. The model studied is compatible with the WMAP limits on dark matter relic density.

Contents

1. Introduction

1.1 Background and motivation

The recent data from the WMAP satellite [1] have allowed the matter density of the Universe to be quantified with greater precision than ever before, whilst also strongly disfavouring warm dark matter. With the relative[1] matter density Ω_m constrained by the measurement $\Omega_m h^2 = 0.135^{+0.008}_{-0.009}$ and the relative baryon density Ω_b constrained by $\Omega_b h^2 = 0.0224 \pm 0.0009$, one can infer the following 2-σ contraint on the relative density of cold dark matter: $\Omega_{CDM} h^2 = 0.1126^{+0.0161}_{-0.0181}$, where the reduced Hubble constant h is measured to be 0.73 ± 0.03.

In R-parity conserving supersymmetric (SUSY) models, the lightest supersymmetric particle (LSP) is stable and is therefore an ideal candidate for non-baryonic cold dark matter. Past studies in the context of the minimal supergravity (mSUGRA) model have identified regions of the five dimensional mSUGRA parameter space in which the relic density of the LSP (usually the lightest neutralino $\tilde{\chi}_1^0$) is consistent with dark matter constraints [2], and recent studies carried out post-WMAP have narrowed these regions further [3]. There has been much recent interest in examining the phenomenology of SUSY models that are consistent with the WMAP results in preparation for the arrival of the LHC.

The aim of this paper is to use the study of one such model to demonstrate a new approach to mass measurements at the LHC. In present analyses, inclusive signatures are rarely used to constrain SUSY models, despite the fact that they are straightforward to define and measure at the LHC. This is almost certainly due to the difficulty associated with calculating the expected values of these signatures at many points in parameter space, a process that requires a large amount of computing power. Nevertheless, we demonstrate that inclusive signatures contain a great deal of information, using as an example the cross-section of events with missing p_T greater than $500\,\text{GeV}$.

The standard technique for analysis of mSUGRA models is to look for kinematic endpoints, and use these to measure the masses of particles involved in cascade decays. These can then be used to obtain the values of the GUT scale mSUGRA parameters. The

---

[1]That is the density divided by the universe's critical density

problem, however, is that such an analysis is often loaded with assumptions. Although endpoint measurements are in principle model independent, it is usually assumed that one has correctly identified the particles in the decay chain, giving unreasonably good precision on the measured masses. Furthermore, it is inevitable that models more general than mSUGRA will be able to reproduce the endpoints seen in cascade decays, and hence it is important to develop techniques that allow one to investigate other possibilities.

Our approach is to combine endpoint measurements with inclusive signatures through the use of Markov chain sampling techniques, a method that can in principle be applied to any parameter space, with any information we happen to have obtained experimentally. The advantage of Markov chain techniques is their efficiency; a crude scan of 100 points per axis in a 3 dimensional parameter space would require one million points, whereas obtaining the same useful information with our choice of sampling algorithm required only 15,000 points. Even so, in order to evaluate inclusive signatures at many points in the parameter space within a sensible period of time, it was necessary to develop an MPI adaptation of the `HERWIG 6.5` Monte Carlo event generator [4−6] for use on a supercomputer with parallel processing.

Throughout this paper, we use a particular mSUGRA model as a description of nature, but it is important to realise that we could in principle have chosen any SUSY model that fits with current observations; the techniques described here rely only on the fact that we have observed endpoints in invariant mass distributions and are able to measure other well-defined observables. Indeed, given enough inclusive observables, one would not even need to have observed endpoints in order to obtain precise results.

Section 2 demonstrates the successful application of kinematic edge analysis to the chosen mSUGRA point before section 3 reviews Metropolis sampling and applies the technique to the reconstruction of the masses involved in a squark decay chain. This differs from current techniques only in the choice of the method used to fit the masses, as it is assumed in sections 2 and 3 that we have correctly identified the particles in the decay chain. In section 4, we introduce a method by which we can combine the endpoint data with a cross-section measurement in order to tighten the precision on the masses, using the sampling techniques reviewed in section 3. For the sake of clarity, this is introduced in the familiar context of an mSUGRA analysis where it is assumed that the particles in the decay chain have been identified correctly, and we merely wish to fit the endpoints and obtain masses and mSUGRA parameters.

Finally, in section 5 we admit that we do not know which particles are in the decay chain, and we also start to relax the conditions of the mSUGRA model by having non-universal gaugino masses at the GUT scale. These are both powerful extensions of the current analysis, and as far as the authors are aware have only rarely been looked at before (e.g. [7, 8] for consideration of particle ambiguity). We also investigate the effect of a jet energy scale error, in order to demonstrate how one might include systematic experimental effects in our technique.

The method developed in sections 4 and 5 can easily be generalised to include other inclusive signatures, and to explore larger parameter spaces, and it can be used in future as a basis for obtaining precise measurements in general SUSY models.

Particle	Mass (GeV)
$\tilde{\chi}_1^0$	137
$\tilde{\chi}_2^0$	264
$\tilde{e}_L$	255
$\tilde{e}_R$	154
$\tilde{g}$	832
$\tilde{u}_L$	760
$\tilde{u}_R$	735
$\tilde{d}_L$	764
$\tilde{d}_R$	733
$\tilde{b}_1$	698
$\tilde{b}_2$	723
$\tilde{t}_1$	574
$\tilde{t}_2$	749
$\tilde{\tau}_1$	147
$\tilde{\tau}_2$	257
h	116

Table 1: The most important sparticle masses at the coannihilation point.

Process	Branching Ratio
$\tilde{\chi}_2^0 \to \tilde{e}_R e$	2%
$\tilde{\chi}_2^0 \to \tilde{e}_L e$	29%
$\tilde{\chi}_2^0 \to \tilde{\tau}_1 \tau$	18%
$\tilde{\chi}_2^0 \to \tilde{\tau}_2 \tau$	2%
$\tilde{\chi}_2^0 \to \tilde{\chi}_1^0 h$	48%
Process	Cross-Section
SUSY (Total, HERWIG)	9.3 pb
SUSY (After ATLFAST missing $p_T > 500\,\mathrm{GeV}$ cut)	2.0 pb

Table 2: Branching ratios and cross-sections for important processes at the coannihilation point.

1.2 Definition of model

This paper describes an analysis carried out on a point consistent with the WMAP data, described by the following set of mSUGRA parameters:

$$m_0 = 70\,\mathrm{GeV}, \qquad m_{1/2} = 350\,\mathrm{GeV}$$
$$\tan\beta = 10, \qquad A_0 = 0, \qquad \mu > 0$$

The values of the universal scalar and gaugino masses at the GUT-scale (respectively m_0 and $m_{1/2}$) are chosen such that the point lies in the coannihilation region in which the LSP's annihilate with sleptons, thus reducing the LSP relic density to a value within the range consistent with WMAP. Henceforth we will refer to this model as the 'coannihilation point'.

The masses of the most relevant particles are contained in table 1, whilst branching ratios for some of the most significant decay processes are given in table 2, generated using ISAJET 7.69. Cross-sections in table 2 were calculated with HERWIG 6.5 and with fortran ATLFAST-2.16. Although similar to the point 5 analysed in the ATLAS Physics TDR [9], this particular model differs by having small mass differences between the $\tilde{\chi}_1^0$ and the $\tilde{e}_R$ and between the $\tilde{\chi}_2^0$ and the $\tilde{e}_L$, leading to the production of soft leptons that may be missed in the detector thereby reducing the efficiency with which we are able to select relevant SUSY decays.

2. Kinematic edge analysis

We begin by demonstrating that standard edge analysis techniques work (within their limitations) for the chosen coannihiliation point.

2.1 Search for squark decay

Previous studies (for example [9, 10, 7]) have illustrated the procedure of searching for kinematic edges in the various invariant mass distributions resulting from a given event. By isolating exclusive decay processes, one can use these kinematic edges to obtain measurements of the masses of the sparticles that participate in the decay chain. The procedure is used here in conjunction with the decay:

$$\tilde{q} \to q\tilde{\chi}_2^0 \to q l_2^\pm \tilde{l}_L^\mp \to q l_2^\pm l_1^\mp \tilde{\chi}_1^0$$

This is an excellent starting point for analysis due to the clear signature provided by the two opposite-sign, same-flavour (OSSF) leptons. The left-handed slepton is considered here rather than the right-handed slepton due to the much greater branching ratio $\mathrm{BR}(\tilde{\chi}_2^0 \to \tilde{e}_L e)$. The following endpoints are expected to be observed in invariant mass spectra associated with this decay chain ($\tilde{\psi} = m_{\tilde{\chi}_2^0}^2, \tilde{q} = m_{\tilde{q}}^2, \tilde{l} = m_{\tilde{e}_L}^2, \tilde{\chi} = m_{\tilde{\chi}_1^0}^2$):

$$(m_{ll}^2)^{\text{edge}} = \frac{(\tilde{\psi} - \tilde{l})(\tilde{l} - \tilde{\chi})}{\tilde{l}} \tag{2.1}$$

$$(m_{llq}^2)^{\text{edge}} = \begin{cases} \max\left[\frac{(\tilde{q}-\tilde{\psi})(\tilde{\psi}-\tilde{\chi})}{\tilde{\psi}}, \frac{(\tilde{q}-\tilde{l})(\tilde{l}-\tilde{\chi})}{\tilde{l}}, \frac{(\tilde{q}\tilde{l}-\tilde{\psi}\tilde{\chi})(\tilde{\psi}-\tilde{l})}{\tilde{\psi}\tilde{l}} \right] \\ \text{except when } \tilde{l}^2 < \tilde{q}\tilde{\chi} < \tilde{\psi}^2 \text{and } \tilde{\psi}^2\tilde{\chi} < \tilde{q}\tilde{l}^2 \\ \text{where one must use} (m_{\tilde{q}} - m_{\tilde{\chi}_1^0})^2. \end{cases} \tag{2.2}$$

$$(m_{lq}^2)_{\max}^{\text{edge}} = \max\left[\frac{(\tilde{q} - \tilde{\psi})(\tilde{\psi} - \tilde{l})}{\tilde{\psi}}, \frac{(\tilde{q} - \tilde{\psi})(\tilde{l} - \tilde{\chi})}{\tilde{l}} \right] \tag{2.3}$$

$$(m_{lq}^2)_{\min}^{\text{edge}} = \min\left[\frac{(\tilde{q} - \tilde{\psi})(\tilde{\psi} - \tilde{l})}{\tilde{\psi}}, \frac{(\tilde{q} - \tilde{\psi})(\tilde{l} - \tilde{\chi})}{(2\tilde{l} - \tilde{\chi})} \right] \tag{2.4}$$

$$(m_{llq}^2)^{\text{thres}} = \frac{2\tilde{l}(\tilde{q} - \tilde{\psi})(\tilde{\psi} - \tilde{\chi}) + (\tilde{q} + \tilde{\psi})(\tilde{\psi} - \tilde{l})(\tilde{l} - \tilde{\chi}) - (\tilde{q} - \tilde{\psi})\sqrt{(\tilde{\psi} + \tilde{l})^2(\tilde{l} + \tilde{\chi})^2 - 16\tilde{\psi}\tilde{l}^2\tilde{\chi}}}{4\tilde{l}\tilde{\psi}} \tag{2.5}$$

where "min" and "max" refer to minimising and maximising with respect to the choice of lepton. In addition, "thres" refers to the threshold that appears in the m_{llq} distribution when events are chosen such that $m_{ll}^{\text{edge}}/\sqrt{2} < m_{ll} < m_{ll}^{\text{edge}}$, corresponding to the angle between the two lepton momenta exceeding $\pi/2$ in the slepton rest frame (see [7]).

2.2 Monte Carlo event simulation

Monte Carlo simulations of SUSY production at the above mass point have been performed using HERWIG 6.5 [4−6], with the particles subsequently passed through the fortran ATLFAST-2.16 detector simulation [11]. A HERWIG input file was generated using ISAJET v7.69 [12] in conjunction with the Herwig-Isajet interface ISAWIG which converts the ISAJET output into HERWIG input format. A sample corresponding to 100fb^{-1} has been generated (being one year of design luminosity in the high luminosity mode).

2.3 Invariant mass distributions

2.3.1 Cuts

In order to see the above edges clearly, one must apply various cuts to the event data in order to isolate a clean sample of the squark decay chain. Here, one can select events with the OSSF lepton signature described above, and one can also exploit the fact that the required events have a large amount of missing energy (due to the departure from the detector of two invisible $\tilde{\chi}_1^0$'s). Furthermore, one expects to obtain hard jets in SUSY events, resulting from the decay of gluinos and squarks. All plots are obtained through the use of the following cuts:

- $E_T^{\text{miss}} > 300$ GeV;

- exactly two opposite-sign leptons with $p_T > 5$ GeV and $|\eta| < 2.5$;

- at least two jets with $p_T > 150$ GeV;

Although the cuts chosen are similar to those used for point 5 in the ATLAS Physics TDR, there are some exceptions. For example, one needs to impose a p_T cut on the leptons in the event due to the fact that `ATLFAST` is not parametrised properly for low p_T leptons, and yet it is essential to pick up soft leptons due to the small mass differences that crop up in the decay chain. Hence, a compromise between these two factors must be chosen. Some plots are the result of additional cuts, and these are given below.

The SM background for dilepton processes is generally negligible once the missing p_T cut has been applied, though the OSSF lepton signature can be produced by SUSY processes other than the decay of the $\tilde{\chi}_2^0$. One would expect these to produce equal amounts of opposite-sign opposite-flavour (OSOF) leptons and hence one can often subtract the dilepton background by producing "flavour subtracted plots" in which one plots the combination $e^+e^- + \mu^+\mu^- - e^+\mu^- - e^-\mu^+$. This is only useful in cases where there are sufficient statistics, and was not done for every plot below.

2.3.2 m_{ll} plot

As seen in figure 1, a sharp edge is produced in the spectrum at ≈ 58 GeV, and this is a very clear signature. The second edge visible at ≈ 98 GeV results from the right-handed selectron. In practise, it will be very difficult to assign these two edges correctly, and this problem is revisited in section 5.

2.3.3 m_{llq} plot

This is produced by selecting events with exactly two leptons, and forming the invariant mass m_{llq} first with the jet q_1 with the largest transverse momentum, and then with the jet q_2 with the second largest transverse momentum. As the hardest jets in events containing cascade decays of squarks and gluinos usually come from the processes $\tilde{q} \to \tilde{\chi}_2^0 q$ and $\tilde{q} \to \tilde{\chi}_1^0 q$, the lower of the two invariant masses formed in this way should lie below the llq edge defined by equation (2.2), and so m_{llq} is defined by $m_{llq} = \min(m_{llq_1}, m_{llq_2})$. Figure 2 shows a clear endpoint in the m_{llq} distribution at ≈ 600 GeV.
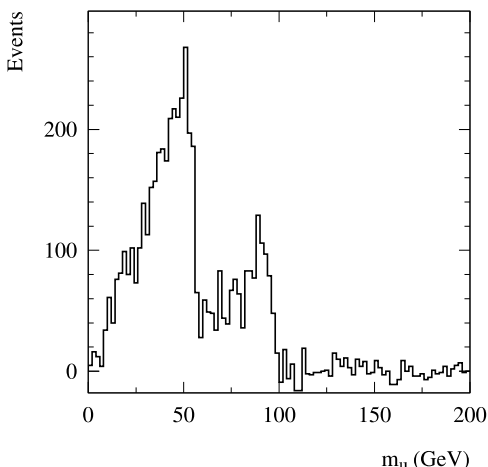
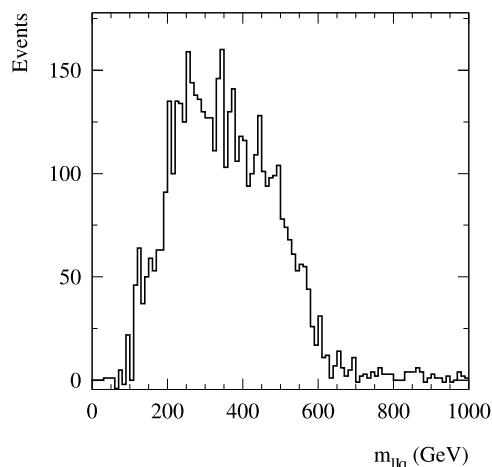**Figure 1:** The flavour-subtracted dilepton invariant mass plotted with the cuts described in the text.

Figure 2: The llq invariant mass plot.

2.3.4 $m_{llq}^{\rm thres}$ plot

The m_{llq} variable plotted in order to measure $m_{llq}^{\rm thres}$ is defined almost in the same way as the m_{llq} variable defined in section 2.3.3. The two differences are that this time (1) $m_{llq} = \max(m_{llq_1}, m_{llq_2})$ (because a threshold[2] is expected rather than an endpoint) and (2) events must satisfy an additional constraint that m_{ll} must exceed $m_{ll}^{\rm max}/\sqrt{2}$. The resulting m_{llq} distribution may be seen in figure 3. This plot is not flavour-subtracted.

A threshold is clearly observed a little above $100\,{\rm GeV}$, though it is difficult to state its precise position due to uncertainty in the expected shape of the edge and the manner in which it is modified by detector effects. This is discussed further below.

2.3.5 $m_{lq}^{\rm max}$ and $m_{lq}^{\rm min}$ plots

In creating the $m_{lq}^{\rm max}$ and $m_{lq}^{\rm min}$ plots the following steps are taken. First, one of the two hardest jets in the event is selected by the same method used in section 2.3.3, i.e. by looking for the combination yielding the lower value of m_{llq}. Having identified this jet (call it q), the quantities m_{l_1q} and m_{l_2q} are formed. The larger of these two combinations $m_{lq}^{\rm high} = \max(m_{l_1q}, ml_2q)$ and the lower of them $m_{lq}^{\rm low} = \min(m_{l_1q}, ml_2q)$ are identified. The distribution of $m_{lq}^{\rm high}$ is plotted in figure 4 and the endpoint located therein is identified as being $m_{lq}^{\rm max}$. The distribution of $m_{lq}^{\rm low}$ is plotted in figure 5 and the endpoint located therein is identified as being $m_{lq}^{\rm min}$.
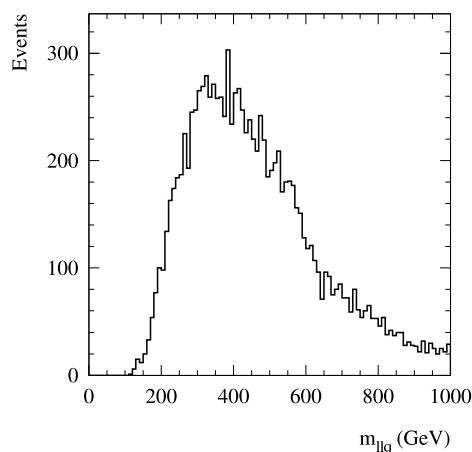

Figure 3: The llq invariant mass threshold plot.

---

[2] The terms "endpoint" and "threshold" are used to refer the the extremal values of a random variable or observable at respectively high and low mass values. The term "edge" describes the shape of the distribution of that variable near its endpoint or threshold.
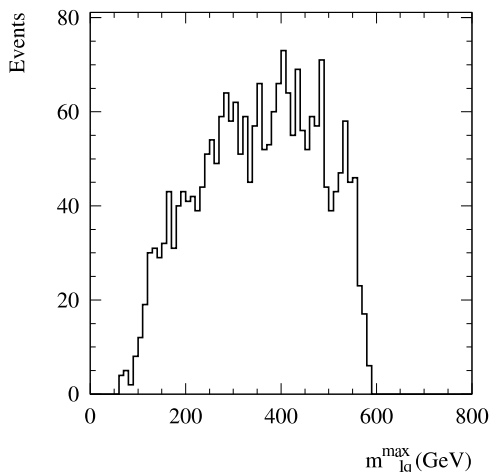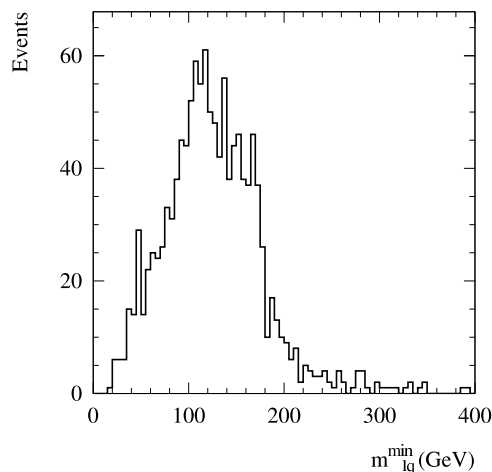
Figure 4: The lq max invariant mass plot.

Figure 5: The lq min invariant mass threshold plot.

Edge	Predicted (GeV)	Observed (GeV)
ll edge	57.64	57.5±2.5
llq edge	600.1	600±10
llq threshold	134.0	150±30
lq max edge	592.1	590±10
lq min edge	181.7	180±10

Table 3: Predicted and observed edge positions for the mSUGRA mass point described in the text. Error estimates have been obtained 'by eye', and reflect lack of information regarding the precise shapes of the endpoints.

For the $m_{lq}^{\max}$ plot (figure 4) events were subject to the additional constraint that one of the llq invariant masses formed with the two hardest jets must be above the llq endpoint, and the other must be below.

The $m_{lq}^{\min}$ plot (figure 5) has one additional cut: the dilepton invariant mass must be less than the value of $m_{ll}^{\max}$ observed in figure 1.

Both plots exhibit endpoints, and the edge is particularly abrupt in the $m_{lq}^{\max}$ histogram. Although there are events beyond the endpoint in the $m_{lq}^{\min}$ plot (due to SUSY background processes), there is nevertheless a convincing edge at $\approx 180\,\mathrm{GeV}$.

2.4 Comparison of observed and predicted edges

The edges predicted by equations (2.1) to (2.5) are summarised in table 3, where the spread of the squark masses has been ignored, approximating them at a common value of 750 GeV, and all other masses are taken from table 1. The observed positions of the endpoints are also given.

It is common when extracting the observed edges from plots such as those above to fit a function to the endpoint in order to determine both the precision and the accurate position. For the purposes of this analysis, the edges, and their estimated errors, have been determined 'by eye' for several reasons. Firstly, not all edges can be fitted with

functions (in the case of the llq threshold, for example, the correct shape is not known). Indeed, recent work in [8] highlights the need for caution in applying these functions too readily without first investigating the theoretical shape of the distribution, as endpoints can often exhibit tails or 'feet' that will be confused with SUSY background and hence may lead to inaccurate measurements. The shapes of the endpoints for distributions involving quarks vary significantly over the parameter space, introducing a model dependence into the precision with which one may realistically measure endpoint positions and hence masses. Given that the purpose of this note is primarily to use an arbitrary example in order to demonstrate our use of Markov chain sampling techniques, a full investigation of how to resolve this model dependence is considered to be beyond the scope of this paper, and we will use the conservative errors given in table 3. For those interested, the fitting of endpoint functions has been done in work leading to [13] which contains estimates of the precision expected if one were to take a more optimistic view.

3. Mass reconstruction

3.1 Background

Having obtained measurements of kinematic edges, the next step is to attempt to reconstruct the masses involved in the squark cascade decay. This has been done using a Markov Chain Monte Carlo method, details of which may be found in appendix A. The technique is an excellent way of efficiently exploring high dimensional parameter spaces, and it is in section 4 that the full advantages of the technique become apparent.

3.2 Application of Metropolis algorithm

We now apply the sampling techniques described in the appendix to our mass reconstruction problem. The five endpoints observed in the previous section essentially provide an (over-constrained) set of simultaneous equations in the four unknowns $m_{\tilde{q}}$, $m_{\tilde{e}_L}$, $m_{\tilde{\chi}_2^0}$ and $m_{\tilde{\chi}_1^0}$, and these can be solved to determine the masses. Given a set of observed edges $\mathbf{e}^{obs}$, and a set of postulated masses $\mathbf{m}$, the ultimate goal is to evaluate $p(\mathbf{m}|\mathbf{e}^{obs})$ and thus to find the regions of parameter space favoured by the data. The best way of doing this is to sample masses $\mathbf{m}$ from $p(\mathbf{m}|\mathbf{e}^{obs})$, subsequently histogramming the samples to reveal directly the shape of the probability distribution.

Using Bayes' Theorem we know that

$$p(\mathbf{m}|\mathbf{e}^{obs}) \propto p(\mathbf{e}^{obs}|\mathbf{m})p(\mathbf{m}). \tag{3.1}$$

We choose the prior $p(\mathbf{m})$ to be uniform[3] over the mass space considered. This choice seems a good as any other, and has the added benefit that plots of our posterior distribution

---

[3]Some points $\mathbf{m}$ in mass space do not satisfy the hierarchy $m_{\tilde{q}} > m_{\tilde{\chi}_2^0} > m_{\tilde{e}_L} > m_{\tilde{\chi}_1^0} > 0$ required by our decay chain. Under our model, then, these points yield $p(\mathbf{e}^{obs}|\mathbf{m}) = 0$ and veto the selection of such points. While this veto is technically part of the likelihood (given our model) it simplifies later discussion of the likelihood in more complicated scenarios if we pull the veto out of the likelihood and move it into the prior $p(\mathbf{m})$. In practise then, our effective prior is uniform over all of the region of mass space in which the required hierarchy is present, and zero elsewhere. The effect is the same as if we had left the veto in the likelihood, but the likelihoods will be simpler to describe and define.

$p(\mathbf{m}|\mathbf{e}^{obs})$ are also just plots of the likelihood $p(\mathbf{e}^{obs}|\mathbf{m})$, permitting the effects of other priors $p(\mathbf{m})$ to be easily imagined.

One can sample from $p(\mathbf{m}|\mathbf{e}^{obs})$ using the Metropolis Method as follows. First a mass point $\mathbf{m}$ is chosen, and $p(\mathbf{m}|\mathbf{e}^{obs})$ is evaluated using equation (3.1). For the edges e_1, e_2, e_3, e_4, and e_5, the likelihood $p(\mathbf{e}^{obs}|\mathbf{m})$ is given by the product

$$p(\mathbf{e}^{obs}|\mathbf{m}) = \prod_{i=1}^{5} p(e_i^{obs}|\mathbf{m}), \tag{3.2}$$

where

$$p(e_i^{obs}|\mathbf{m}) \approx \frac{1}{\sqrt{2\pi\sigma_i^2}} \exp\left(-\frac{(e_i^{obs} - e_i^{pred}(\mathbf{m}))^2}{2\sigma_i^2}\right) \tag{3.3}$$

in which σ_i is the statistical and fit error associated with the edge measurement of edge e_i, and where e_i^{obs} and $e_i^{pred}(\mathbf{m})$ are respectively the observed and predicted positions of the edge. This probability distribution assigns a weight $p(\mathbf{m}|\mathbf{e}^{obs})$ to each point $\mathbf{m}$ in mass space, including the errors associated with the endpoint measurements. Note that $p(\mathbf{m}|\mathbf{e}^{obs})$ is the equivalent of the $P^*(\mathbf{x})$ defined later on in equation (A.2), as it is defined only up to an unknown normalisation constant.

So, in order to plot the probability distribution, one follows the following steps of the Metropolis Algorithm:[4]

1. A new mass point $\mathbf{m}^{proposal}$ is suggested on the basis of the current point $\mathbf{m}^{current}$. The mass-space proposal distribution for the Metropolis Algorithm was chosen to be a 4-dimensional gaussian whose width in each dimension was $5\,\mathrm{GeV}$ and whose centre was the position of the current point $\mathbf{m}^{current}$. The widths were chosen for the efficiency reasons outlined in section A.4 and will not effect the results once convergence has occurred.

2. $p(\mathbf{m}^{proposal}|\mathbf{e}^{obs})$ is evaluated at the proposed point.

3. A decision is made on whether to jump to the new point, or remain at the current point on the basis (see equation (A.3)) of the ratio of $p(\mathbf{m}^{proposal}|\mathbf{e}^{obs})$ to $p(\mathbf{m}^{current}|\mathbf{e}^{obs})$.

4. If a decision to *not jump* is made, then the next point in the chain $\mathbf{m}^{next}$ is again set equal to $\mathbf{m}^{current}$, otherwise it is set equal to $\mathbf{m}^{proposal}$. When proposals are rejected, therefore, successive points in the chain are duplicates of each other.

5. All steps are repeated until the sampler has sufficiently explored the interesting regions of parameter space.

It is noted that in the real ATLAS detector, one might have a systematic shift of the endpoints due to the jet energy scale error, and this is considered in section 5.

---

[4]See appendix A.3 for discussion of the motivations behind each of these steps, and for definitions of "proposal functions" and the decision mechanism.

3.3 Mass space plots

The Metropolis sampler ensures that points which are more likely are sampled more often. One can observe the shape of the probability distribution by simply histogramming the sampled points. This is a 4 dimensional shape in mass space, which can be viewed as a projection onto each pair of axes. This is done in figure 6, revealing that a lengthy region of parameter space is compatible with the edge data, and extra information is required to constrain this further. Note that the endpoint equations discussed previously are sensitive principally to mass differences, and hence one observes lines in each plane of the mass space, constraining each mass only in terms of the others. Given that the endpoint data does not set the overall mass scale, the system is able to wander to high masses without affecting the position of the endpoints provided that the mass differences remain consistent. In the next section, we show that one can use other measurements from the LHC that are sensitive to the mass scale to constrain these regions further.

Finally, it is noted that the lines are broader in the plots featuring squark masses, and this is due to the fact that the end points were calculated using an average squark mass, whilst the Monte Carlo events feature a range of squark masses. Hence the resolution is smeared somewhat relative to the other masses.

4. Cross-section information in mSUGRA space

4.1 Background

In principle, any piece of relevant information may be used to further constrain the regions consistent with the kinematic edge analysis presented in the previous section. This may be in the form of further kinematic edges, which will provide a direct constraint on the weak scale sparticle masses, or in the form of constraints at the SUSY scale. The greater the number of relevant pieces of information one is able to obtain, the better the precision of the mass measurements.

One example is given here, and developed further in this section. It should be possible to measure the cross-section of events with missing p_T greater than $500\,\text{GeV}$ in the ATLAS detector to a reasonable precision. As the masses of sparticles increase, the missing p_T will increase, but the total production cross-section will decrease and hence the high mass solutions encountered in the previous section will lead to missing p_T cross-sections that are lower than the value obtained at the coannihilation point. Thus, the cross-section information can be added to the definition of the probability function for the Markov Chain to give a tighter constraint on the SUSY masses.

It should be noted that up to now we have performed a model independent analysis but, from here on in, some model will have to be assumed in order to draw conclusions from our measurements. This is because endpoint data can be analysed purely in the mass space S_{mass} (hereafter "M") defined by the weak scale masses, but inclusive measurements must be compared to a given scenario (through the use of a suitable Monte Carlo generator) before any conclusions can be drawn, and therefore must be analysed in the space of parameters, S_{model} of that model. In section 4, we investigate the constraints imposed
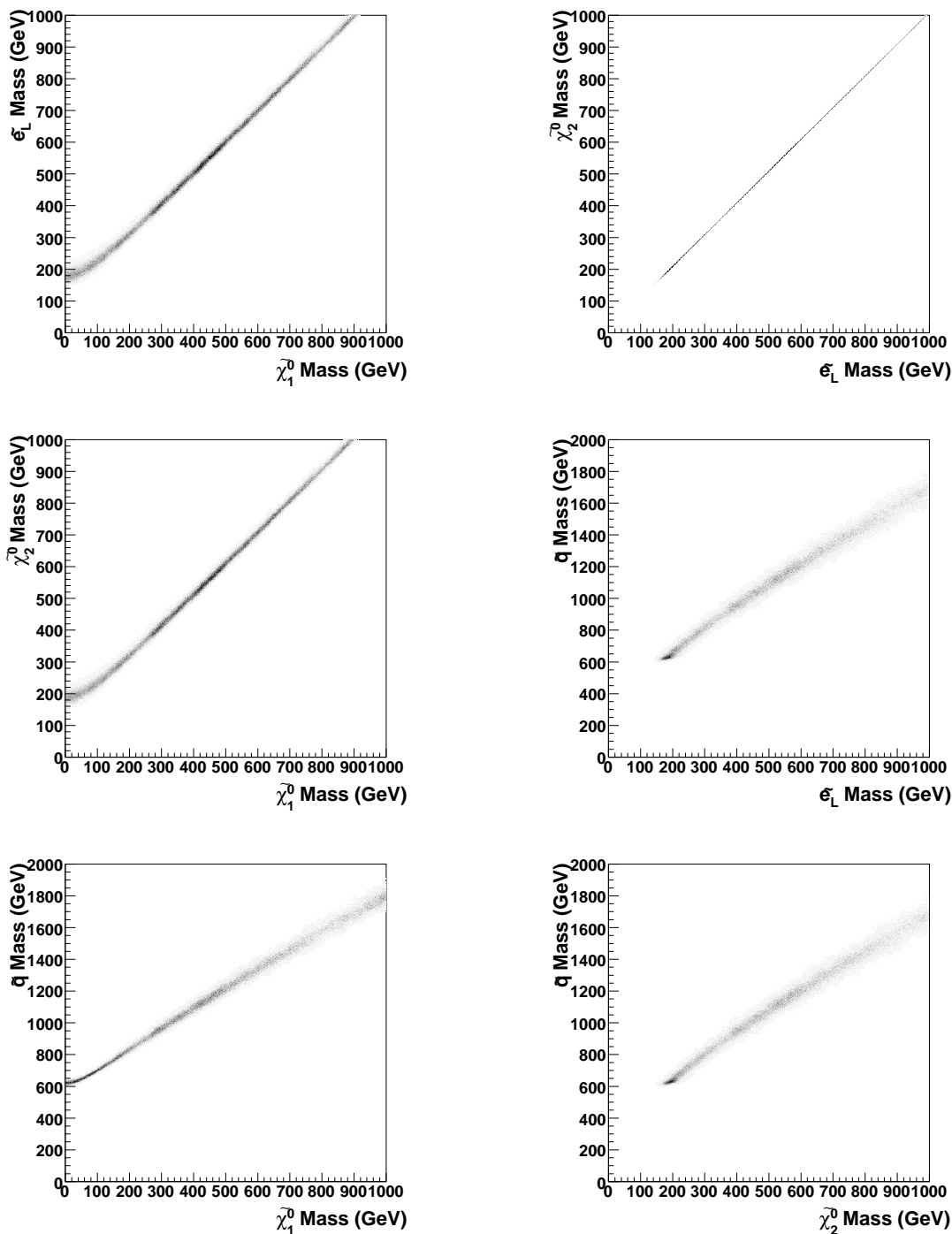
Figure 6: The region of mass space consistent with the kinematic edge measurements described in the text, obtained using a Markov chain sampler.

by a cross-section measurement on the parameter space S_{mSUGRA} (hereafter "P") of a particular model, mSUGRA, in order to introduce the technique in a familiar context. The limitations of this approach will become apparent by the end of section 4 and will be tackled in section 5.

In view of this change of the constrained-space, (from the space of weak-scale masses $\mathbf{m} \in M$ to the space of mSUGRA models $\mathbf{p} \in P$) the description of the Metropolis algorithm in section 3.2 must, in section 4, be considered re-written in terms of $p(\mathbf{p}|\mathbf{e}^{obs})$ rather than $p(\mathbf{m}|\mathbf{e}^{obs})$. This is made more explicit in section 5.1 when a *further* enlargement of the constrained-space is made to accommodate uncertainty in the absolute jet energy scale.

4.2 Cross-section measurement

4.2.1 Implementation

It is assumed in this study that the cross-section of events with missing p_T greater than 500 GeV can be measured at ATLAS. One can then pick points in the mSUGRA parameter space S_{mSUGRA}, work out the mass spectrum, generate Monte Carlo events and work out the cross-section of events passing this cut. Only certain points in the parameter space are consistent with this measurement, and these will give a range of masses that are consistent. Naively, the overlap of this region of the mass space with the region consistent with the edge data will give the new region of mass space that is compatible with the ATLAS data. In fact, since the end points are not entirely independent of the cross-section measurement, one needs to include both sets of information in the final analysis in order to obtain the required region. The 'overlap' picture is approximately true, however, given that the measurement of the cross-section is not strongly related to the measurements of the edge positions, and is a useful guide for preliminary investigation before the final analysis is carried out.

A plot of the missing p_T cross-section in the m_0, $m_{1/2}$ plane for fixed $\tan\beta$ and positive μ is shown in figure 7. As can be seen, there is a lot of variation over the parameter space and a measurement of even modest precision will be able to rule out large areas.

The full process of picking mSUGRA points and obtaining the cross-section that pass the missing p_T cut has been accomplished by successively running `ISAJET`, `HERWIG` and `ATLFAST`, with 1000 events being generated at each point. This is rather time consuming, however, and a simple scan of the mSUGRA parameter space is unfeasible if any more than two of the mSUGRA parameters are varied. For this reason, we again use the Metropolis sampling technique introduced in the previous section and, indeed, it is here that the power of the method becomes apparent. The algorithm has been used to map the interesting region of the parameter space with fewer points than would be required in a scan in order to obtain similar performance.

To demonstrate this, consider the following. There are four and a half parameters in the mSUGRA parameter space, though we have held A_0 constant for simplicity.[5] Of the remaining parameters, one is simply a sign (the sign of μ), and hence one sampling run was performed with this sign positive, and another with it negative. In any one application of the software, then, three parameters are varied — m_0, $m_{1/2}$ and $\tan\beta$ — and even a coarse scan requiring 100 points along each axis would require one million points for each sign of

---

[5]In retrospect there was no compelling reason to hold A_0 fixed, and in a later study we expect to look at the effect of allowing A_0 to vary and be measured by the data along with all the other model parameters.
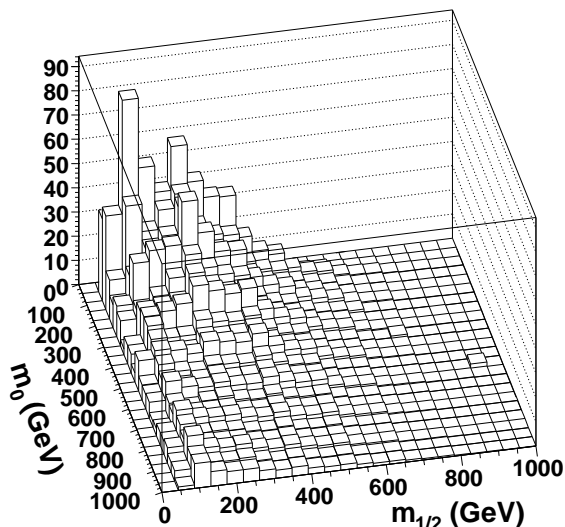
Figure 7: The cross-section in picobarns for events passing a missing p_T cut of $500\,\text{GeV}$, for $\tan\beta$ = 10 and positive μ, obtained using `HERWIG`. The value at the coannihilation point is 2.03 pb. The irregularity in the plot comes from the statistical error from having only simulated 1000 events at each point, in keeping with the method used in the sampler.

μ. The Metropolis algorithm maps the interesting regions of the space in approximately 15,000 points per sign of μ, a dramatic improvement without which the analysis would have taken many months, if not years.

Even with this improvement, it was still necessary to reduce the run time of `HERWIG` significantly through the use of a parallel computer. Although the Metropolis algorithm itself cannot be parallelised, we have adapted `HERWIG` to run on a parallel machine with the use of `MPI` code, thereby substantially reducing the run time per point.

4.2.2 Definition of Metropolis quantities for cross-section

We now define the Metropolis algorithm for use with (only) the cross-section data. As in the previous section, we require the definition of the probability distribution $p(\mathbf{p}|\sigma^{obs})$ from which samples are to be taken, in which σ^{obs} represents the cross-section supposedly "observed" or measured by the experiment. Lacking real data, we take σ^{obs} to be 2.04 pb, the value predicted by a `HERWIG` and `ATLFAST` simulatuion of the coannihiliation point of section 1.2. The evaluation of $p(\mathbf{p}|\sigma^{obs})$ necessitates the definition of a suitable prior $p(\mathbf{p})$ on the model space P which again we take to be flat (but equal to zero for invalid values of any of the model parameters $p_i \in \mathbf{p}$). Finally the Metropolis sampler's proposal distribution must be modified to act on the model space P rather than on the mass space M. The proposal distribution was again chosen to be a multi-dimensional gaussian centred on the current point $\mathbf{p} \in P$. The widths of the proposal distribution in m_0, $m_{1/2}$ and $\tan\beta$ were respectively usually $25\,\text{GeV}$, $25\,\text{GeV}$ and 2, except when both cross-section *and* edge constraints were in use simultaneously (only in sections 4.3 and beyond) in which case a smaller set of widths was used ($5\,\text{GeV}$, $5\,\text{GeV}$ and 2). The widths were chosen for the

efficiency reasons outlined in section A.4 and will not affect the results once convergence has occurred.

The sampled probability distribution $p(\mathbf{p}|\sigma^{obs})$ follows a similar definition to that encountered previously for $p(\mathbf{m}|\mathbf{e}^{obs})$. The analogue of equation (3.2) is then just the single term $p(\sigma^{obs}|\mathbf{p})$ quantifying the cross-section likelihood according to:

$$p(\sigma^{obs}|\mathbf{p}) \approx \frac{1}{\sqrt{2\pi\sigma_{err}^2}} \exp\left(-\frac{(\sigma^{obs} - \sigma^{pred}(\mathbf{p}))^2}{2\sigma_{err}^2}\right), \tag{4.1}$$

where σ_{err} is the error associated with the cross-section measurement σ^{obs}, and $\sigma^{pred}(\mathbf{p})$ is the value of the cross-section expected at the point $\mathbf{p}$ in mSUGRA parameter space P as again predicted by a `HERWIG` and `ATLFAST` simulation.

The error σ_{err} on the observed cross-section σ_{obs} was taken to be ten per cent. This figure was chosen somewhat arbitrarily, for similar reasons to those given when explaining the sizes of the errors assumed for the endpoint measurements (see section 2.4): this paper is designed to illustrate a *method*, not to claim that a particular measurement can be made with a certain precision. In contrast, if we had access to *real* data, it would be of vital importance to make the estimation of the cross-section error as accurate as possible. The eventual precision of the final answer will be strongly correlated with the error attributed to the cross-section. In retrospect, the chosen value of ten per cent probably underestimates the combination of (1) the statistical error, (2) luminosity error, (3) the theoretical uncertainty on the signal cross-section, and (4) the combined experimental and theoretical uncertainty on the prediction for the number of standard model events likely to pass the signal cuts. If we were to be granted further time on the supercomputer and were able to start the analysis again from scratch, we would probably re-simulate with a larger and more realistic error of thirty percent.[6] Further work (beyond the scope of this paper) should be done to investigate the expected size of this error, and to confirm that the effect of increasing this error estimate is just to enlarge the size of the final regions. Within this article, however, the cross section error will be taken to be the stated ten per cent — and this will be sufficient for the purpose of demonstrating how the proposed method can be used in practice.

Certain regions of mSUGRA parameter space P are known to be unphysical — for example there may be no electroweak symmetry breaking or there may be a charged LSP. In both cases, `ISAJET` will detect this and fail to run. Furthermore there are points $\mathbf{p}$ for which `HERWIG` will not run. When any of these problems occur we take the point $\mathbf{p}$ to be unphysical and multiply the likelihood by zero (as unphysical points *cannot* have generated the observed data!).

---

[6]One might hope to estimate what would have happened with a choice of $\sigma_{err} = 30\%$ by reweighting our exiting "10%" samples. Unfortunatley we find it is only possible to re-weight our existing samples *robustly* to reflect values of σ_{err} satisfying $\sigma_{err} \leq 12\%$. Attempting to reweight more strongly leads to a very small number of very strongly reweighted samples dominating the whole reweighted distribution. There are insufficient statistics in the tails of our existing sampling to allow us to identify the parts of space that are relevant where a more generous error to have been assumed.
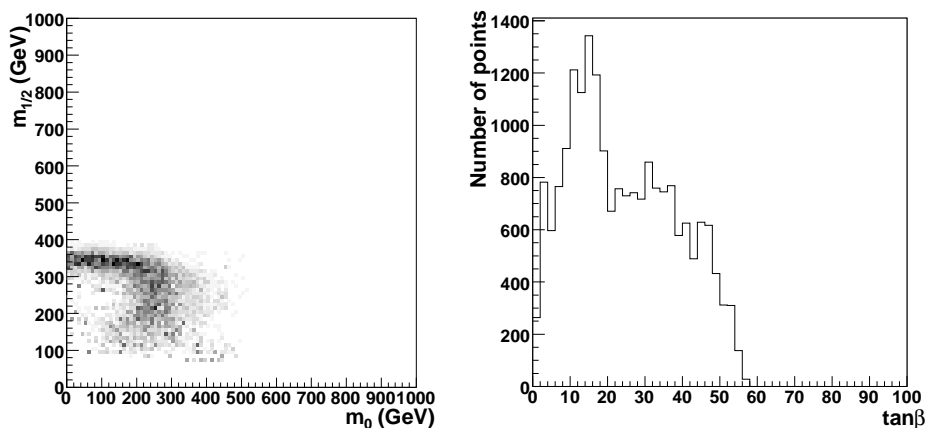
Figure 8: The region of mSUGRA parameter space consistent with the measurement of the cross-section of events with missing p_T greater than $500\,\mathrm{GeV}$, for positive μ.
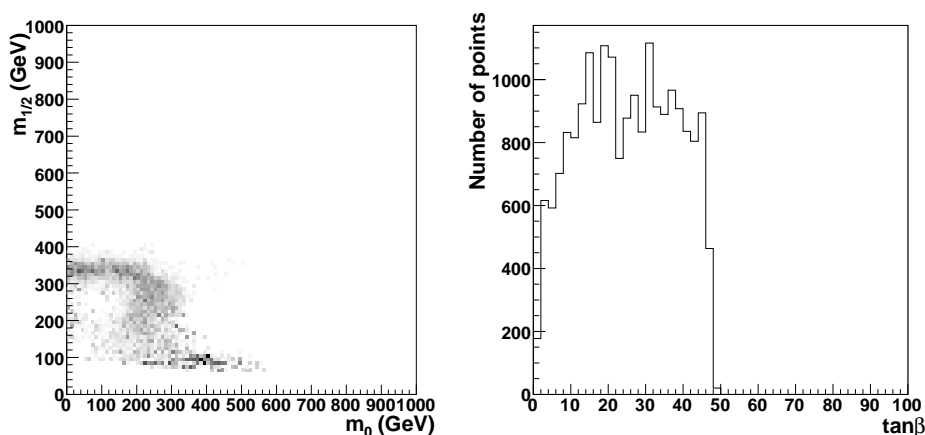

Figure 9: The region of mSUGRA parameter space consistent with the measurement of the cross-section of events with missing p_T greater than $500\,\mathrm{GeV}$, for negative mu.

4.2.3 Results in mSUGRA space (for cross-section information alone)

The results of the Markov Chain in mass space for positive μ can be seen in figure 8, with those for negative μ presented in figure 9. The distributions look very similar in the $m_0, m_{1/2}$ plane, reflecting a lack of sensitivity to the sign of μ. The $\tan\beta$ distribution is approximately flat for negative μ, whilst there is some insignificant preference for the 'correct' value of $\tan\beta = 10$ in the positive μ case.

4.2.4 Results in mass space (for cross-section information alone)

We now relate the results in figures 8 and 9 to the weak scale mass space in which we have already observed the regions consistent with the kinematic edge analysis. The positive μ and negative μ data sets presented previously have been evolved to the weak scale using ISAJETand combined into a single data set by weighting each of the two sets by the average likelihood of that set. The region obtained in mass space is shown in figure 10, and is

dramatically different from that obtained using the edge analysis. The overlap between the regions found by the two methods (figures 6 and 10) is shown in figure 11, and was obtained by multiplying the previous data sets.

The overlap of the two regions has produced much tighter constraints on the particle masses, even with a relatively conservative estimate of the precision of the endpoint measurements. It is worth noting that the projections of the region of overlap on each pair of axes give different size regions in each plane, with the smallest being that in the plane of the neutralino masses. This could be used to remove some of the area shown in the other planes, although the strictly correct procedure (followed in section 4.3) is to run a Markov Chain with the edge and cross-section information implemented at the same time.

4.3 Further analysis

The overlap plots presented in the previous subsection give a rough idea of what to expect from the combination of edge and cross-section information, but the approach is only approximately valid given that the cross-section measurement is not independent of the kinematic edge positions. In order to be fully rigorous, one must run a Markov Chain whose probability density function combines both the cross-section and the edge information at the same time — in other words one must sample this time from $p(\mathbf{p}|\mathbf{e}^{obs}, \sigma^{obs})$.

Accordingly, a Metropolis sampler of $p(\mathbf{p}|\mathbf{e}^{obs}, \sigma^{obs})$ was set to explore the mSUGRA parameter space P.

At each point $\mathbf{p} \in P$ the number of events passing the missing p_T cut was obtained from the `ATLFAST` output whilst the `ISAJET` mass spectrum was used to find the expected position of the endpoints. This information was then compared to the 'measured' information (in this case, the endpoints shown earlier, and the cross-section obtained through Monte Carlo simulation of the coannihilation point) in the definition of the probability weight for each point $\mathbf{p} \in P$. The likelihood $p(\mathbf{e}^{obs}, \sigma^{obs}|\mathbf{p})$, the analogue of equations (3.2) and (4.1), is this time the product of the pair of them:

$$p(\mathbf{e}^{obs}, \sigma^{obs}|\mathbf{p}) = p(\sigma^{obs}|\mathbf{p}) \prod_{i=1}^{5} p(e_i^{obs}|\mathbf{m}(\mathbf{p})). \tag{4.2}$$

The same flat prior $p(\mathbf{p})$ on mSUGRA space P was used as in section 4.2.2. The likelihood was multiplied by zero if the sparticle masses $\mathbf{m}(\mathbf{p})$ obtained at a point $\mathbf{p}$ were not consistent with the mass hierarchy required for the squark decay chain to exist. The Metropolis algorithm's proposal distribution was the same as that used previously in section 4.2.2. Chains were run separately for positive and negative μ.

4.3.1 Results for cross-section and edge measurements *together*

The mSUGRA space results for cross-section *and* edge measurements are shown in figures 12 and 13, with the results in mass space shown in figure 14. Note that inclusion of the cross-section information greatly improves the precision in the $m_0, m_{1/2}$ plane.

We would like to emphasise at this stage that the majority of the apparent improvement is *not* the result of the inclusion of the cross-section measurement — but is rather a well
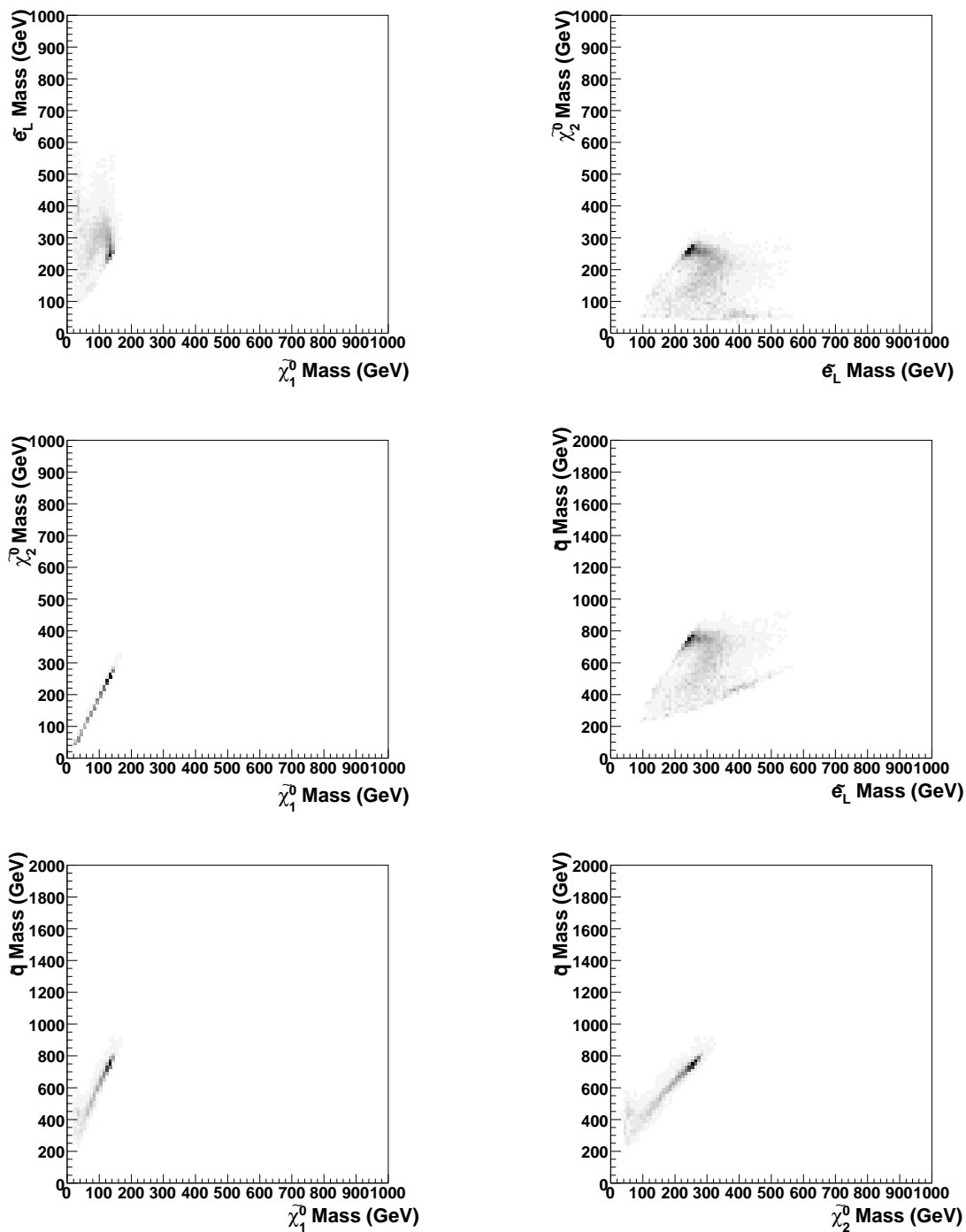
Figure 10: The region of mass space consistent with a measurement at 10% precision of the cross-section of events with missing p_T greater than $500 \, \text{GeV}$, obtained using a Markov chain sampler.

known side-effect of the fit taking place in a model space which is more tightly constrained (masses depend primarily on just the two parameters m_0 and $m_{1/2}$) than the original mass space (four free masses). Many points in mSUGRA space are now rejected as they
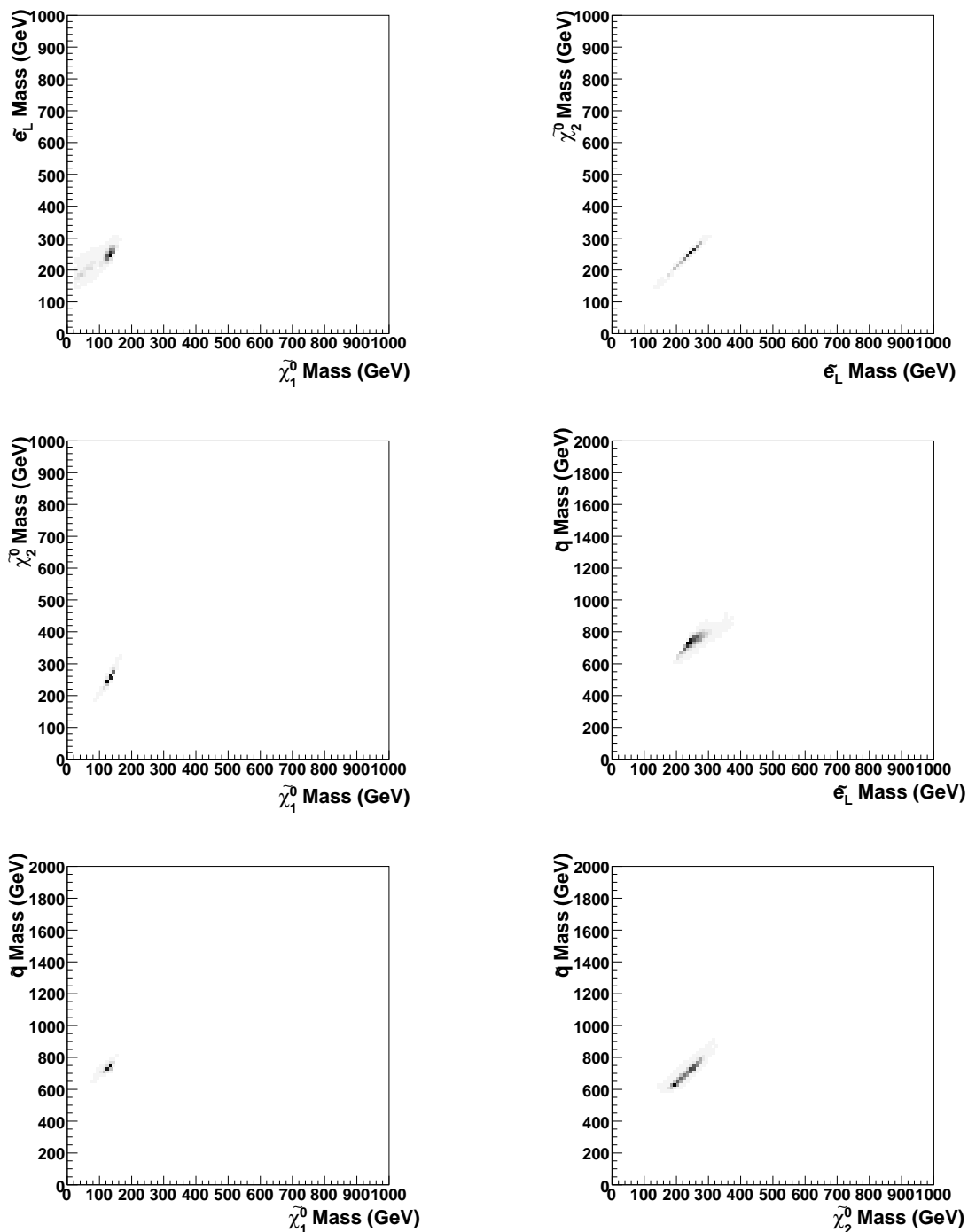
Figure 11: The region of mass space consistent with a measurement at 10% precision of the cross-section of events with missing p_T greater than $500\,\mathrm{GeV}$, overlapped with a measurement of the squark decay kinematic endpoints obtained in section 2.

give the wrong mass hierarchy for the decay chain. This leads to a jump in mass-space precision, at the expense of incorporating some model dependence. If we are prepared to accept the model dependence introduced by moving to the space of a particular model (in this case mSUGRA) we are forced to accept "uncomfortably" tight constraints on the
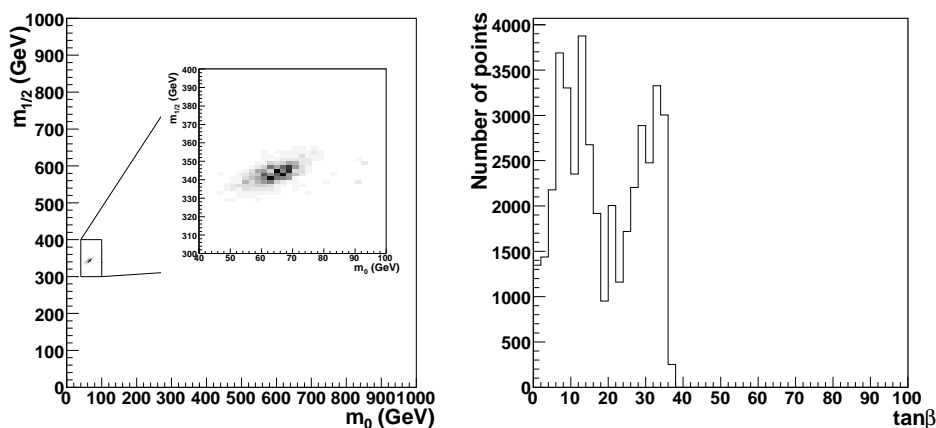
Figure 12: The region of mSUGRA parameter space consistent with the measurement of the cross-section of events with missing p_T greater than $500\,\mathrm{GeV}$ and with the endpoint measurements obtained in section 2, for positive μ.
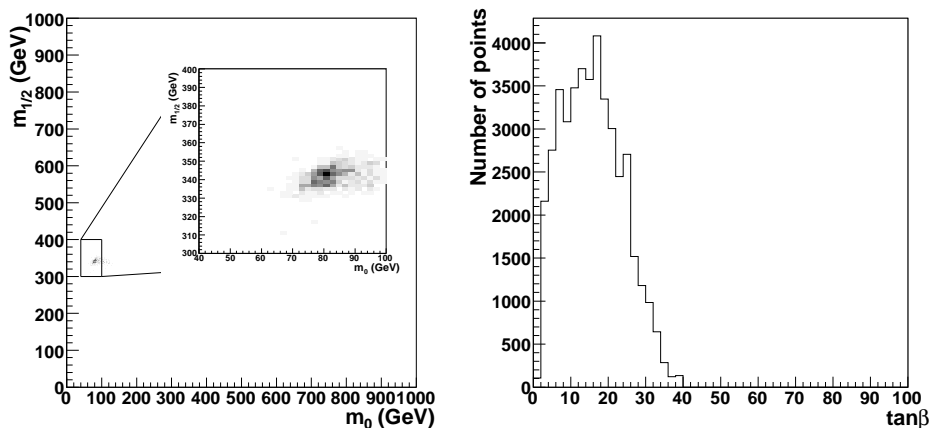

Figure 13: The region of mSUGRA parameter space consistent with the measurement of the cross-section of events with missing p_T greater than $500\,\mathrm{GeV}$ and with the endpoint measurements obtained in section 2, for negative μ.

compatible regions of parameter space. Why uncomfortable? Uncomfortable because the choice of mSUGRA was somewhat arbitrary, and made without a strong degree of belief that mSUGRA is an effective theory of Nature. Given this lack of confidence in mSUGRA itself, there seems little use in being able to quote tiny errors on the parts of it which are compatible with the data — especially when even small departures from the mSUGRA model might lead to significant changes in the sparticle spectra or properties.

However, this very distaste is now the motivation for recognising that we are no longer restricted to looking at overly constrained models like mSUGRA, and suggests that we can now look at a wider class of models in which we hope to have a higher degree of faith. In this way we can lose some of the unpleasant model dependence just introduced, and can for the first time actually put the cross-section measurement in a position in which it can play

an *active* role in constraining competing theories. We thus hope to illustrate the power of our technique.

In section 5 we go on to increase the dimensionality of the parameter space in exactly this way (by relaxing the conditions that impose, for example, universal gaugino masses at the GUT scale, etc) and still maintain good precision by using the endpoint data together with the cross-section measurement. Inclusive and exclusive data is combined to explore more general SUSY models in order to learn yet more about the SUSY lagrangian.

Finally, it is crucial to note that there is another limitation in the analysis so far in that it has been assumed that one has established that the particles involved in the decay chain are the two lightest neutralinos and the left-handed slepton. In practise, one could just as easily fit endpoints using, for example, the heaviest two neutralinos and the right-handed slepton. This ambiguity ought to be reflected in the analysis, and has only rarely been considered before (see for example [7, 8]). This is also considered in section 5.

5. Going beyond mSUGRA

We have seen thus far that one can sample from the mSUGRA parameter space using both kinematic endpoint data and a simple cross-section measurement. Endpoint data alone gives more than adequate precision within the framework of mSUGRA, provided one samples the mSUGRA parameter space and assumes that one has identified the particles in the decay chain correctly. The aim of this section is to generalise this analysis to include both ambiguity in the decay chain and more general assumptions about the mechanism of SUSY breaking. We will also consider the effect of the jet energy scale error on the endpoint positions, thus demonstrating how one would include correlated experimental effects in our analysis.

5.1 Effect of a jet energy scale error

Any detector such as ATLAS does not measure the energy of jets perfectly, but instead has some energy scale error. Given that most of the endpoints feature a quark jet, it is worth investigating the effect of the energy scale error on the positions of the endpoints, and the subsequent effect on our precision in the mSUGRA parameter space.

Firstly, it is noted that for jets whose energy exceeds $70\,\mathrm{GeV}$ (the likely energy of the jet in our endpoints given the relatively large mass difference between the squarks and the neutralinos), the energy scale error is expected to be of the order of 1 per cent [9]. This is much lower than the errors we have already attributed to the endpoints that arise from mismeasurement, and hence the effect will not cause a discernible difference to our results. We have nevertheless included the effect in our analysis as an example of how one can incorporate experimental effects in our analysis.

To accommodate the effect of an unknown shift s in the absolute jet energy scale, we add s to the parameter set explored by the sampler. In other words, the sampler now wanders around the extended space $Q = P \otimes S$ defined as the product of the mSUGRA parameter space $P = \{\mathbf{p}\}$ with the set S of possible values of s. At each point $\mathbf{q} = (\mathbf{p}, s) \in Q$ we work out the masses $\mathbf{m}(\mathbf{p})$ of the particles in the decay chain. We then calculate
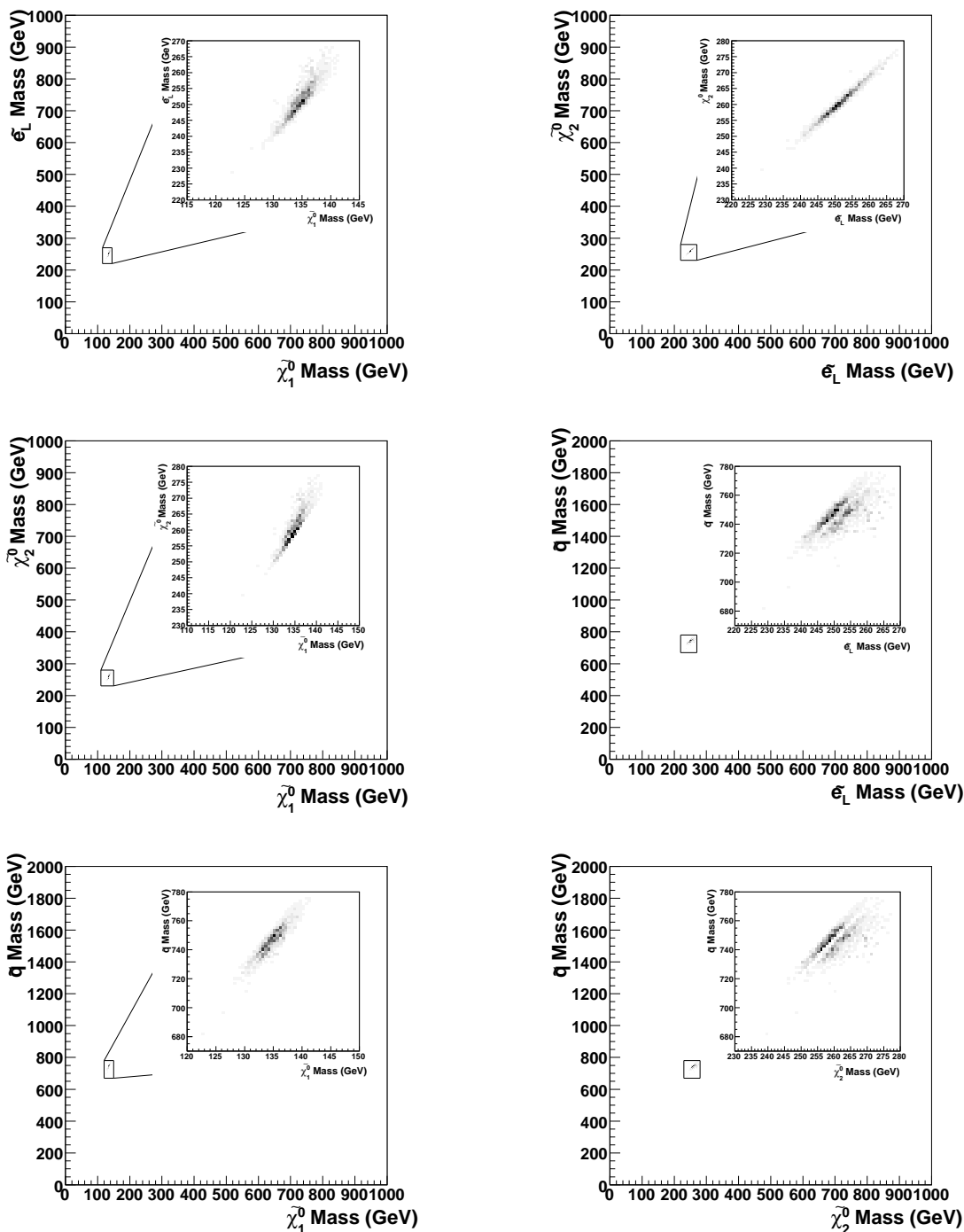
Figure 14: The region of mass space consistent with a measurement at 10% precision of the cross-section of events with missing p_T greater than $500\,\text{GeV}$ combined with the endpoints measured in section 2, obtained using a Markov chain sampler in mSUGRA space.

the "idealised" positions of the edges corresponding to these masses (as before) but we then *move* the positions of these edges by the amount predicted by the current hypothesis s for the the absolute jet energy scale correction. The resulting modified edge positions
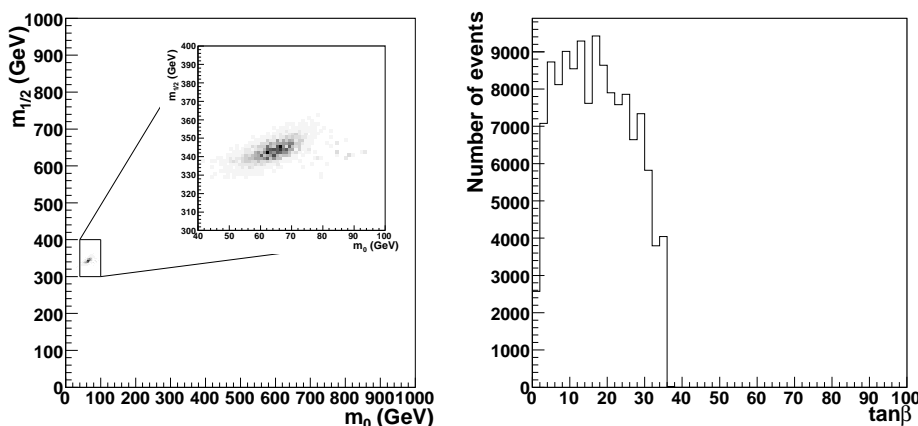
Figure 15: The region of mSUGRA parameter space consistent with the endpoint measurements obtained in section 2, for positive μ, with a 1 per cent jet energy scale error included.

$\mathbf{e}^{pred} = \mathbf{e}^{pred}(\mathbf{q}) = \mathbf{e}^{pred}(\mathbf{m}(\mathbf{p}), s)$, which now depend on s, are the values which are used in the new version of equation (3.3).

Having extended P to the larger space Q, our goal is now to sample not from $p(\mathbf{p}|\mathbf{e}^{obs})$ but from $p(\mathbf{q}|\mathbf{e}^{obs})$. The latter is proportional to $p(\mathbf{e}^{obs}|\mathbf{q})p(\mathbf{q})$. The first term $p(\mathbf{e}^{obs}|\mathbf{q})$ may be calculated almost exactly as before in equation (3.3) but with the new modified edge positions $\mathbf{e}^{pred}(\mathbf{m}(\mathbf{p}), s)$ described above. The last term $p(\mathbf{q})$ may be decomposed by independence into two parts: $p(\mathbf{p})p(s)$. The first of these, $p(\mathbf{p})$, is the mSUGRA-space prior which we have seen before,[7] while the other, $p(s)$, is the expected distribution of the final uncertainty in the absolute jet energy scale. Following [9] we take $p(s)$ to be a gaussian of width 1%.

In order to determine the particular amounts δ_i by which the ith endpoint should be shifted for a given jet energy scale correction factor s, we run a toy Monte Carlo simulation at that point and for that edge.[8] This is done once with and once without the correction factor s multiplying the jet energies. The positions of the endpoints are compared in the two cases. Different endpoints are thus shifted by different fractions of the energy scale error s.

The results including uncertainty in the jet energy scale are shown in figures 15 and 16 for positive and negative μ respectively and are comparable to those obtained previously (figures 12 and 13) when uncertainty in the jet energy scale was not considered.

5.2 Chain ambiguity in mSUGRA

In order to investigate the effect of chain ambiguity on the mSUGRA parameter space, the edge data from section 2 are here used in an mSUGRA fit *without* the assumption that the particles in the decay chain have been identified correctly. It is still true that there are few processes that can give the characteristic endpoints associated with the squark cascade decay already described, and it should be sufficient merely to include the possibility that any of the neutralinos may be produced in the decay (provided of course that the one further down the chain is lighter than that above it) and that one has ambiguity over the

---

[7]We must remember that, as in earlier sections, the likelihood $p(\mathbf{e}^{obs}|\mathbf{q})$ will be zero (given our model) at points where the masses of the particles in the chain do not obey the necessary mass hierarchy. It was computationally easier for us to place this veto into the prior $p(\mathbf{p})$ as before.

[8]Strictly speaking the toy Monte Carlo simulation is only needed for the *llq* edge and the *llq* threshold as the shifts in the edge positions for the *other* edges are linear in $\sqrt{s}$ and may be calculated analytically.
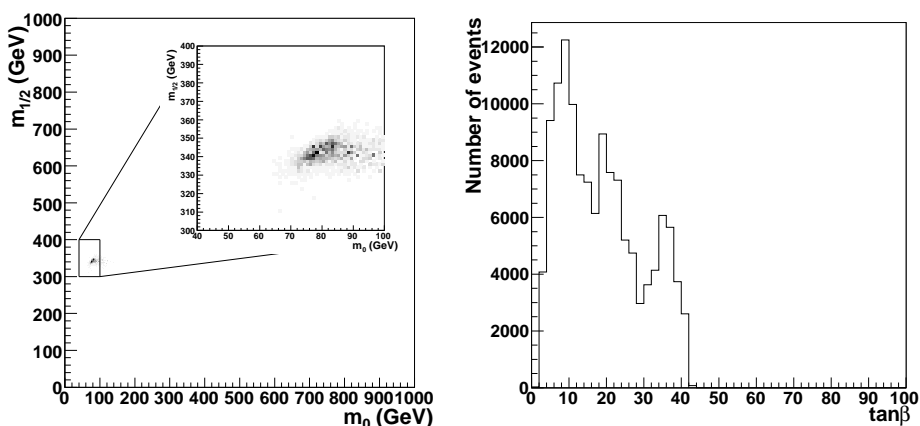
Figure 16: The region of mSUGRA parameter space consistent with the endpoint measurements obtained in section 2, for negative μ, with a 1 per cent jet energy scale error included.

slepton chirality. This gives twelve possible mass hierarchies (see table 4) and each of these gives a series of possible endpoints in the mass spectra. The issue of how to deal with parts of parameter space able to generate the same final state through three- rather than two-body decays (for example when the sleptons are too massive to produce directly) is beyond the scope of this document but is ideal for further study.

There can easily be points in parameter space at which almost all sparticle production goes through one particular hierarchy (say H_1), but in which a different hierarchy (say H_2) has endpoint locations which are a much better fit to the *positions* of the "observed" edges. This could be true even if the cross section for H_2 was much less than for H_1. Events from H_2 might not even be observable. It is very costly to accurately determine the observability (after realistic detector cuts and consideration of backgrounds) of each of the hierarchies in table 4 at every point in parameter space visited by the Markov Chain.

Name	Hieracrchy						
H_1	$m_{\tilde{q}}$	$>$	$m_{\tilde{\chi}_2^0}$	$>$	$m_{\tilde{e}_L}$	$>$	$m_{\tilde{\chi}_1^0}$
H_2	$m_{\tilde{q}}$	$>$	$m_{\tilde{\chi}_3^0}$	$>$	$m_{\tilde{e}_L}$	$>$	$m_{\tilde{\chi}_1^0}$
H_3	$m_{\tilde{q}}$	$>$	$m_{\tilde{\chi}_3^0}$	$>$	$m_{\tilde{e}_L}$	$>$	$m_{\tilde{\chi}_2^0}$
H_4	$m_{\tilde{q}}$	$>$	$m_{\tilde{\chi}_4^0}$	$>$	$m_{\tilde{e}_L}$	$>$	$m_{\tilde{\chi}_1^0}$
H_5	$m_{\tilde{q}}$	$>$	$m_{\tilde{\chi}_4^0}$	$>$	$m_{\tilde{e}_L}$	$>$	$m_{\tilde{\chi}_2^0}$
H_6	$m_{\tilde{q}}$	$>$	$m_{\tilde{\chi}_4^0}$	$>$	$m_{\tilde{e}_L}$	$>$	$m_{\tilde{\chi}_3^0}$
H_7	$m_{\tilde{q}}$	$>$	$m_{\tilde{\chi}_2^0}$	$>$	$m_{\tilde{e}_R}$	$>$	$m_{\tilde{\chi}_1^0}$
H_8	$m_{\tilde{q}}$	$>$	$m_{\tilde{\chi}_3^0}$	$>$	$m_{\tilde{e}_R}$	$>$	$m_{\tilde{\chi}_1^0}$
H_9	$m_{\tilde{q}}$	$>$	$m_{\tilde{\chi}_3^0}$	$>$	$m_{\tilde{e}_R}$	$>$	$m_{\tilde{\chi}_2^0}$
H_{10}	$m_{\tilde{q}}$	$>$	$m_{\tilde{\chi}_4^0}$	$>$	$m_{\tilde{e}_R}$	$>$	$m_{\tilde{\chi}_1^0}$
H_{11}	$m_{\tilde{q}}$	$>$	$m_{\tilde{\chi}_4^0}$	$>$	$m_{\tilde{e}_R}$	$>$	$m_{\tilde{\chi}_2^0}$
H_{12}	$m_{\tilde{q}}$	$>$	$m_{\tilde{\chi}_4^0}$	$>$	$m_{\tilde{e}_R}$	$>$	$m_{\tilde{\chi}_3^0}$

Table 4: The twelve mass hierarchies considered in section 5.2.

For this reason, in this article we adopt the following conservative position. We choose not to consider the (un)observability of end points at different points in parameter space. Instead we assume that *every* hierarchy consistent with the masses of a given point is potentially visible. This assumption is conservative because in reality only a few hierarchies will be visible. The consequence of our assumption is that we will not reject points of pa-

rameter space that a more in-depth analysis might be able to reject. It would be interesting for further work to pursue the possibility of making stronger statements at each point in parameter space based not only on the positions of the observed edges, but also based on the number of event in them, and the number of events in distributions which were *not* observed to have edges etc. How to cope with points at which heavy sleptons force three-rather than two-body neutralino decays should also be investigated. In depth analyses of this kind are beyond the scope of this paper, however.

If we label the N_a different mass assignments with a tag a_i, the likelihood for the i-th observed edge at each point $\mathbf{p}$ in the mSUGRA parameter space P now becomes:

$$
\begin{aligned}
p(e_i^{obs}|\mathbf{p}) &= \sum_{j=1}^{N_a} p(e_i^{obs}|\mathbf{p}, a_j)p(a_j) \\
&= \sum_{j=1}^{N_a} p(e_i^{obs}|\mathbf{m}_{a_j}(\mathbf{p}))p(a_j)
\end{aligned}
\tag{5.1}
$$

where $p(a_i)$ is the prior for the mass assignments, and N_a gives the number of assignments open at that point in parameter space. If we assume that each of the assignments is equally likely, the prior $p(a_i)$ is simply $1/N_a$. The term $p(e_i^{obs}|\mathbf{m}_{a_i}(\mathbf{p}))$ is calculated using equation (3.3) with the masses corresponding to the assignment a_i.

Equation (5.1) makes the conservative assumption that any observed edge could have come from any observed chain (i.e. not necessarily from the same chain as that generating a different observed edge). Furthermore (but less realistically) it assumes that there is no correlation between the chains generating each of the edges, whereas in many parts of parameter space it is highly likely that there is only one dominant chain. It is thus arguable that equation (5.1) should be replaced by the stronger statement

$$
\begin{aligned}
p(\mathbf{e}^{obs}|\mathbf{p}) &= \sum_{j=1}^{N_a} p(\mathbf{e}^{obs}|\mathbf{p}, a_j)p(a_j) \\
&= \sum_{j=1}^{N_a} p(\mathbf{e}^{obs}|\mathbf{m}_{a_j}(\mathbf{p}))p(a_j)
\end{aligned}
\tag{5.2}
$$

which says that *all* the observed edges were the result of the same (albeit unknown and unidentified) chain of sparticles. We choose to present results using (5.1) rather than (5.2).

The results for positive μ are seen in figure 17, whilst those for negative μ are in figure 18. The precision is worse than that encountered previously, but not by much. It may be seen that there are two favoured regions in each plot, rather than the single region encountered previously. The region at larger m_0 is one in which hierarchy H_1 dominates the sum (5.1). The lower m_0 region has (5.1) dominated by hierarchy H_7 in which the right-slepton is substituted for the left-slepton.

The next course of action is to view the regions in the weak scale mass space that correspond to the chosen mSUGRA points, and here we have a problem. Since we are now assuming that we do not know exactly which particles are in the decay chain, we can no
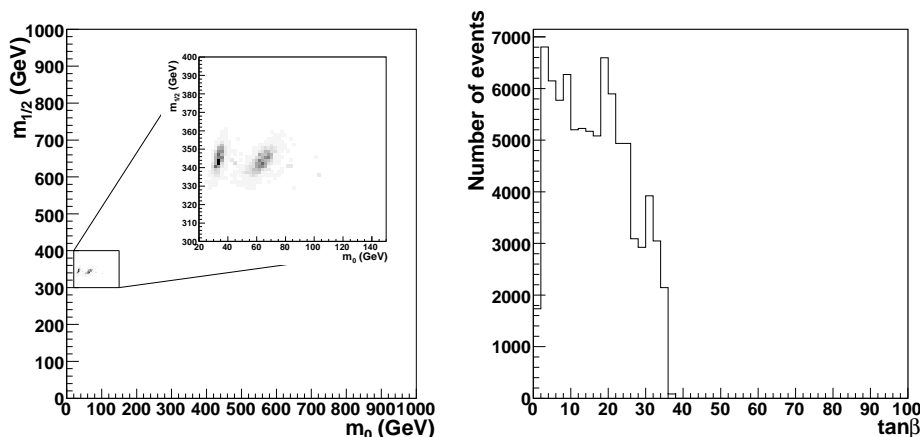
Figure 17: The region of mSUGRA parameter space consistent with the endpoint measurements of section 2, without the assumption that the neutralinos and slepton in the squark decay chain have been correctly identified. For full details, see text. Results are shown for positive μ.
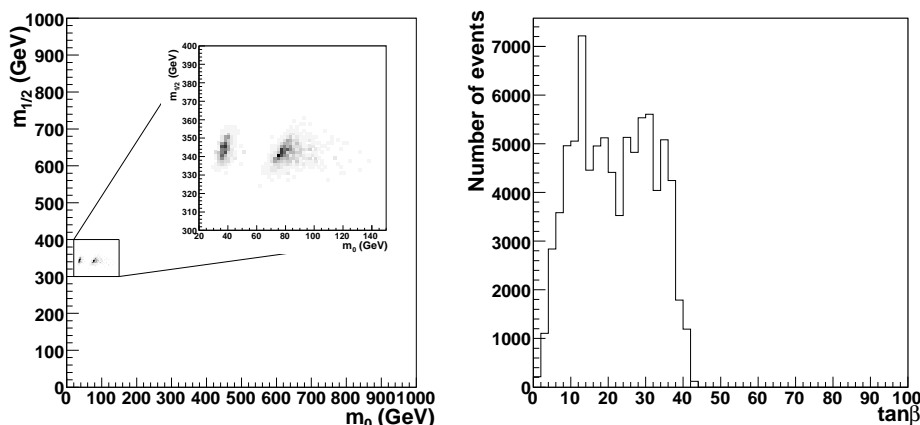

Figure 18: The region of mSUGRA parameter space consistent with the endpoint measurements of section 2, without the assumption that the neutralinos and slepton in the squark decay chain have been correctly identified. For full details, see text. Results are shown for negative μ.

longer take the points in the mSUGRA plane and claim that they give us the masses of the lightest two neutralinos and the left handed slepton. Instead, we can merely say that we have measured a neutralino-like object and a slepton-like object, but that we need some more facts before we can say anything more.

We can, however, use some other information to tell us more about the particles in the decay chain. For a start, we can look at the width of the distribution for each mass (neutralino 1, neutralino 2, etc) that results from the mSUGRA points and use these widths as a qualitative guide. If the endpoints are really caused by a single mass hierarchy, the masses in this chain should generally fit the data better than other hierarchies, and this will manifest itself in a smaller spread of masses for the masses involved in the correct hierarchy. In our case, the endpoints should all be caused by a decay chain featuring the

lightest two neutralinos and the left handed slepton, so we expect these masses to have narrower distributions. This is indeed the case for the neutralinos, as seen in figure 19, though the selectron results are less different.

Note that figure 19 does not yet show mass measurements. The plots could only be interpreted as mass measurements if further work were able to establish the identities of the particles involved and confirm that they came predominantly from just one chain. Here we show them only to help get a hold on which decay chains appear to be consistent with the results.

There are other things that can be done. Having had our attention drawn to a small region of the mSUGRA parameter space, we can look within that region at the branching ratios for the different possible mass hierarchies, after which we might find that there are not enough events of a given type to be consistent with the observed endpoints. Therefore, although a decay chain featuring a neutralino 3 and neutralino 2 may fit a given endpoint slightly better than the correct chain, it might be impossible for that chain to produce an endpoint with the same number of events present as has been observed. This, in conjunction with the width of the mass distributions, might be enough to confirm the nature of the true decay chain, but it would be foolish to assume that the true chain will always be easy to identify.

Given that the region in mSUGRA space has not substantially increased in size, we will not add the cross-section information at this stage. Instead, we will investigate the effect of relaxing some of the assumptions of the mSUGRA model.

5.3 A non-universal SUGRA model

The mSUGRA model assumes universality of the scalar and gaugino masses at the GUT scale, and also unifies the trilinear couplings at the GUT scale. Although this helps in reducing the SUSY breaking parameter set to a manageable level, reality may present a more complicated case. Hence, there is a very strong motivation for developing techniques that are either model independent or are at least able to tackle some more general SUSY models.

In this subsection, we investigate the effect of relaxing the assumption of universal GUT scale gaugino masses, whilst still retaining the chain ambiguity and jet energy scale effects encountered in the sections 5.1 and 5.2. It is important to realise that this is merely a first example of the use of the techniques developed here; one could just as easily relax more of the mSUGRA assumptions provided that one has made enough measurements to provide suitable constraints on the resulting model.

5.3.1 Kinematic edge constraints on non-universal SUGRA

The parameter set for the SUGRA model now becomes m_0, $\tan\beta$, A_0, $\text{sgn}(\mu)$, M_1, M_2 and M_3. A Metropolis sampler was used to sample from this parameter space (along with the jet energy scale error s), with the mass spectrum of each point found using `ISAJET 7.69`. Chain ambiguity was incorporated in the same way as described in section 5.2. The results are seen in figures 20 and 21: it should be noted that the previous m_0 vs $m_{1/2}$ plot has been superseded by three plots against the various GUT scale gaugino masses. The
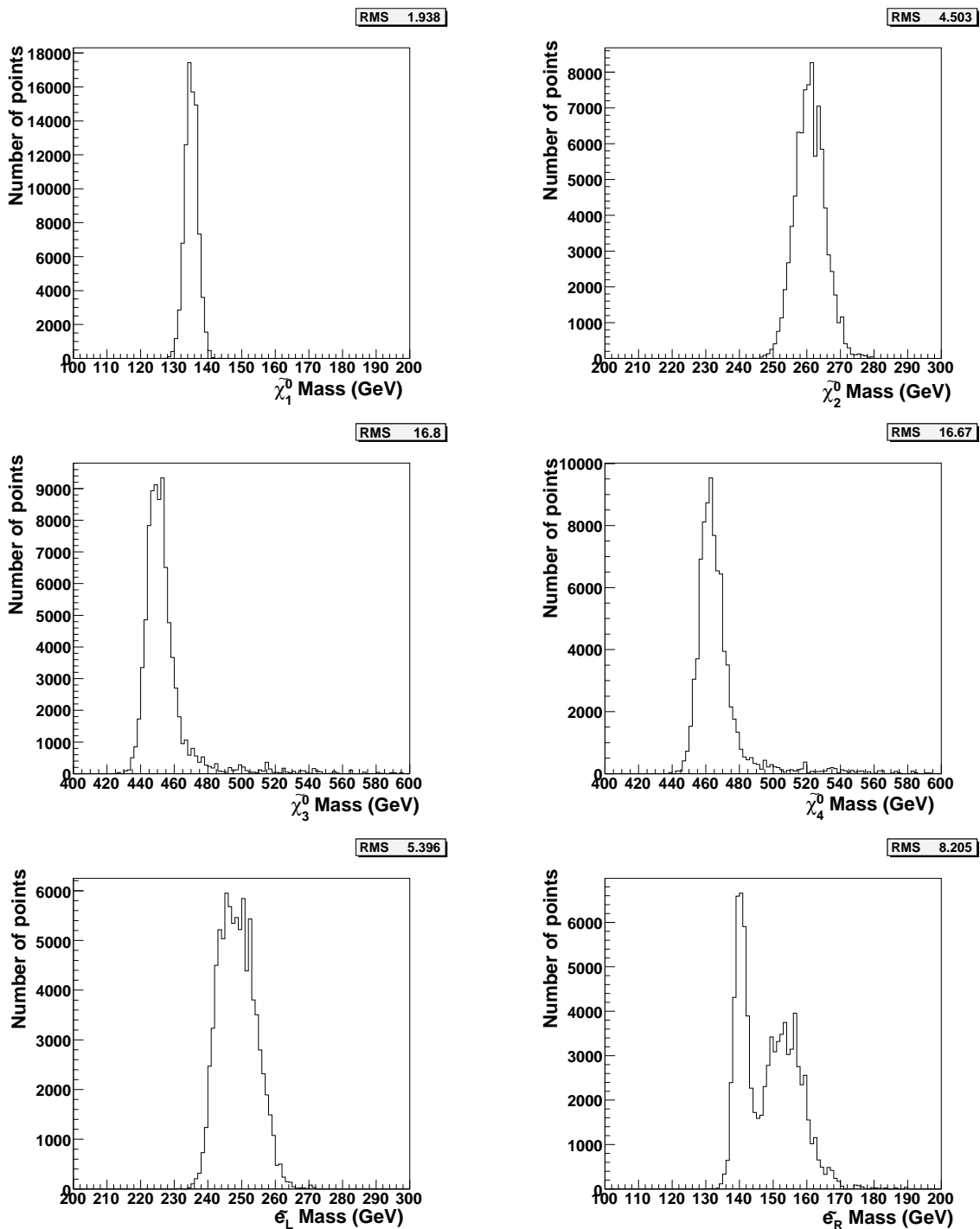
Figure 19: The mass distributions obtained from the mSUGRA chain, in which one has not assumed that the particles in the decay chain have been identified. These are not to be confused with mass measurements! The width of each plot (RMS about the mean) is recorded for each plot in GeV.

plots shown contain 800,000 points, after which the sampler was still clearly exploring new areas of the parameter space. In these plots, the Markov Chain has not yet converged, and
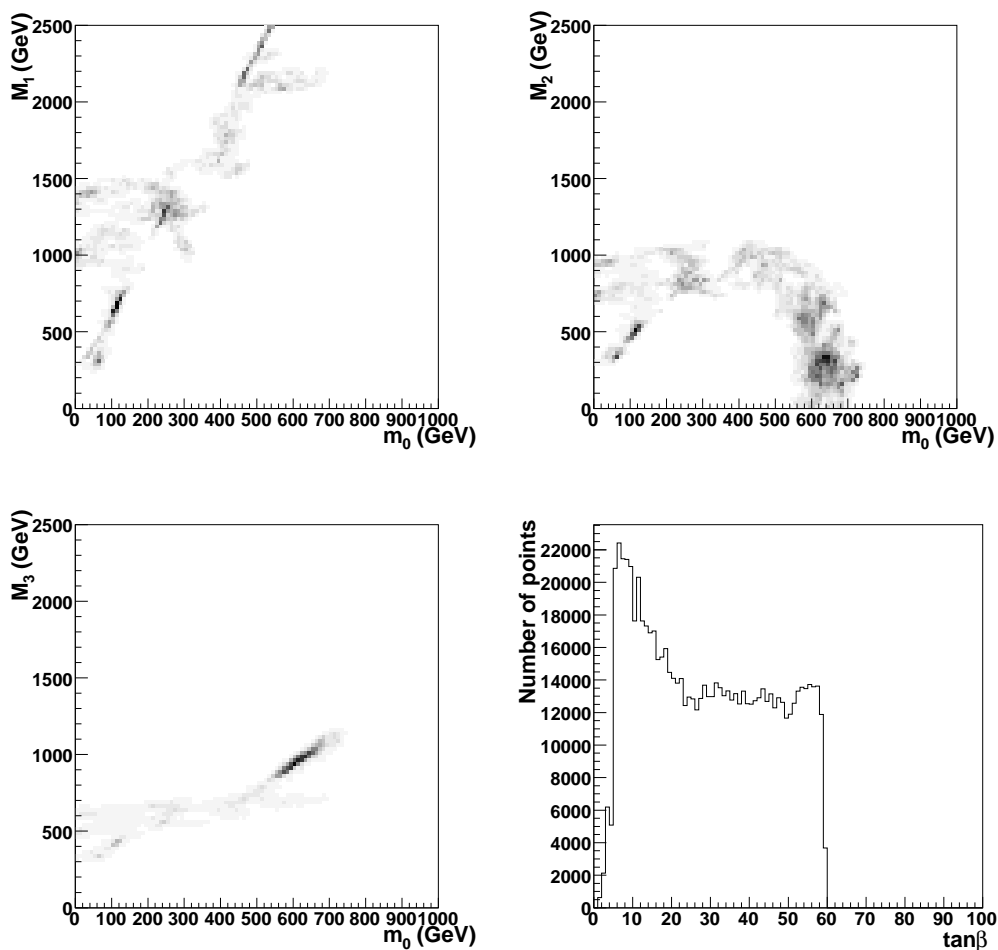
Figure 20: The region of our non-universal SUGRA parameter space consistent with the endpoint measurements of section 2, with chain ambiguity included. Results are shown for positive μ.

this lack of convergence is sufficient to show that the endpoint data *alone* do not provide sufficient information to adequately constrain the non-universal SUGRA model, and so we have indeed reached a point where we need to consider additional measurements — such as the cross-section.

5.3.2 Kinematic edge data *and* cross-section constraints on non-universal SUGRA

A further Metropolis sampler was used to explore the parameter space of our non-universal SUGRA model using both the cross-section information *and* the edge data in the definition of the probability weight for each point. The results for positive μ are seen in figure 22, whilst those for negative μ are seen in figure 23, and the difference from the plots described above is immediately apparent. The system is much more tightly constrained, and it has not wandered too far from the region corresponding to an mSUGRA model in which M_1,
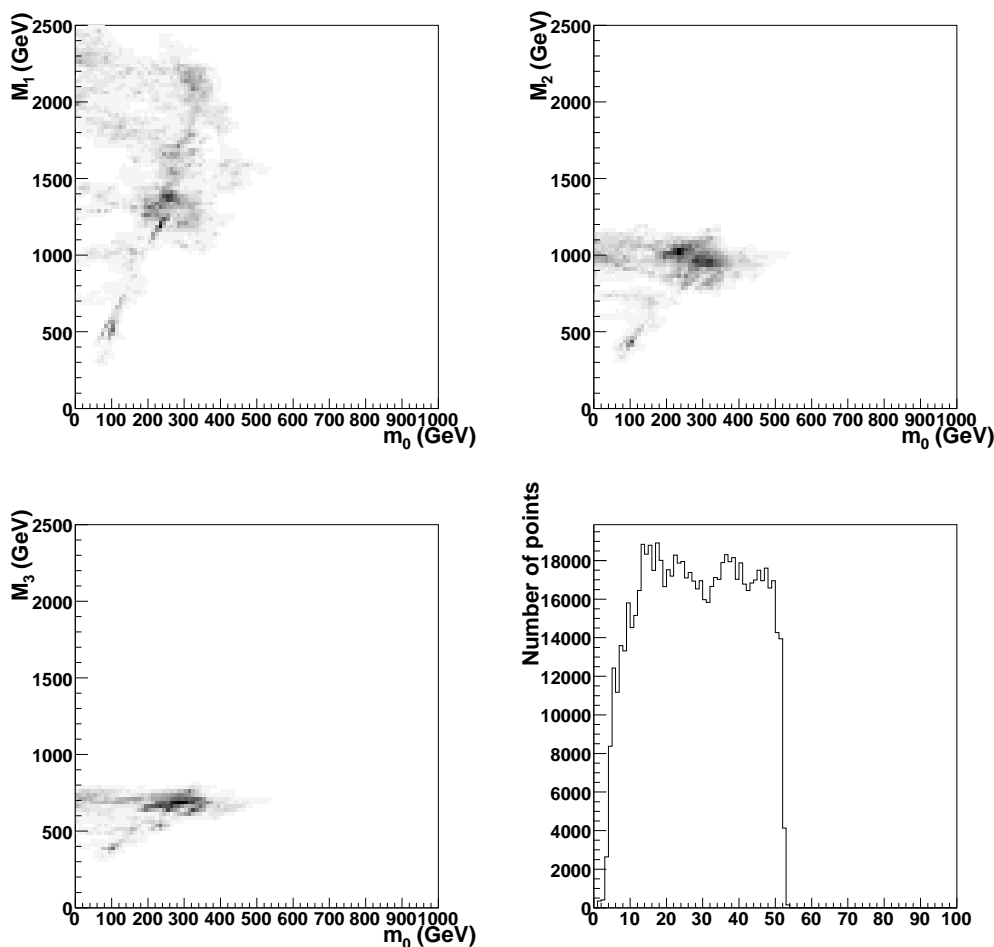
Figure 21: The region of our non-universal SUGRA parameter space consistent with the endpoint measurements of section 2, with chain ambiguity included. Results are shown for negative μ.

M_2 and M_3 are degenerate. One can convert this GUT scale region to a region in mass space as before (see figure 24), though with the previous disclaimer that we have not yet identified which of the particles are involved in the decay chain but merely the range on the various masses that might be involved. Further work in the form of Monte Carlo studies targeted in the selected region at the GUT scale might possibly identify which masses are involved and hence improve the precision further, a study that is perfectly feasible given the relatively small extent of the region allowed by our data.

The results presented here are very encouraging, however, showing that even with only one extra observable we can afford to be more honest about our lack of information regarding decay processes whilst still obtaining adequate precision within the framework of mSUGRA, and reasonable precision in a more general model.
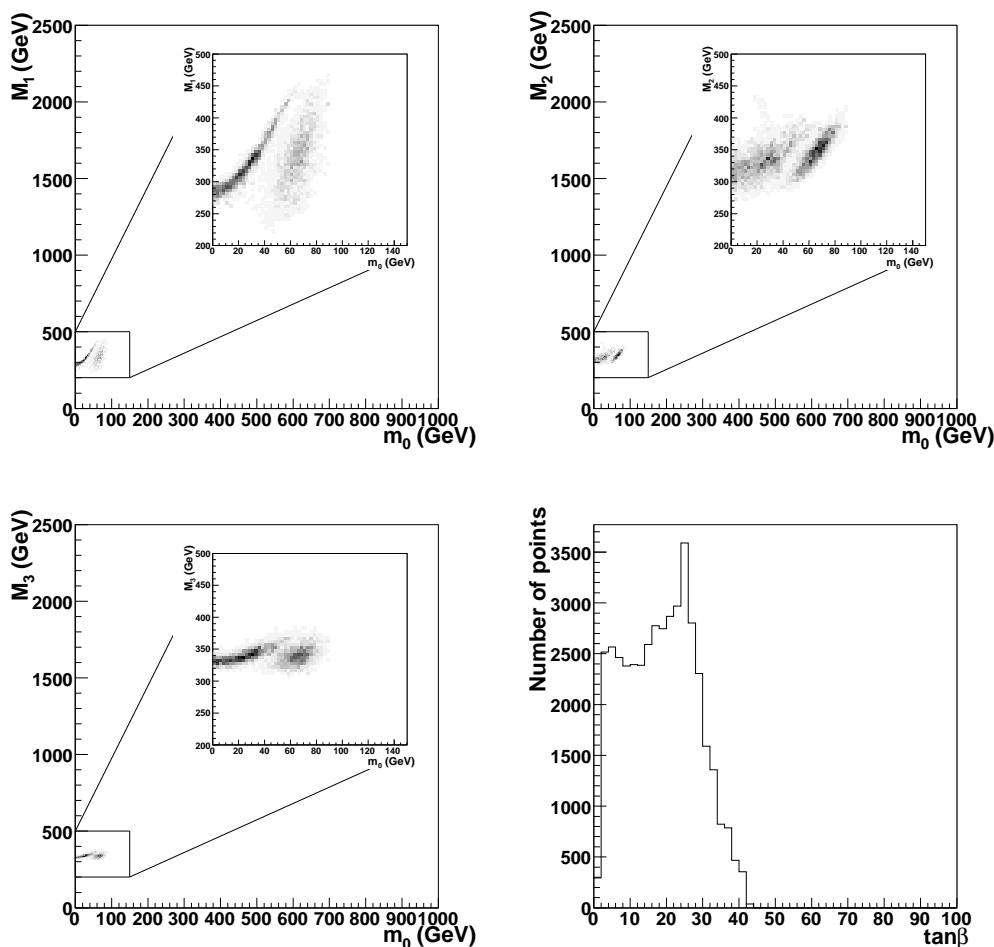
Figure 22: The region of our non-universal SUGRA parameter space consistent with the endpoint measurements of section 2 and the cross-section measurement, with chain ambiguity included. Results are shown for positive μ.

6. Conclusions

We have used Markov Chain sampling techniques to combine kinematic endpoint measurements with a cross-section measurement in order to obtain precision SUSY mass measurements in simulated ATLAS data. Previous analyses have been extended to include ambiguity in the decay chain responsible for the endpoints, and a preliminary study has been made of a non-universal SUGRA model. Throughout it has been shown that the precision of mass measurements is greatly improved through the use of inclusive data, and the technique described offers a rigorous and general approach to the problem of constraining SUSY at the LHC. Reasonable precision has been obtained even with a fairly conservative estimate of the errors on the endpoints themselves.

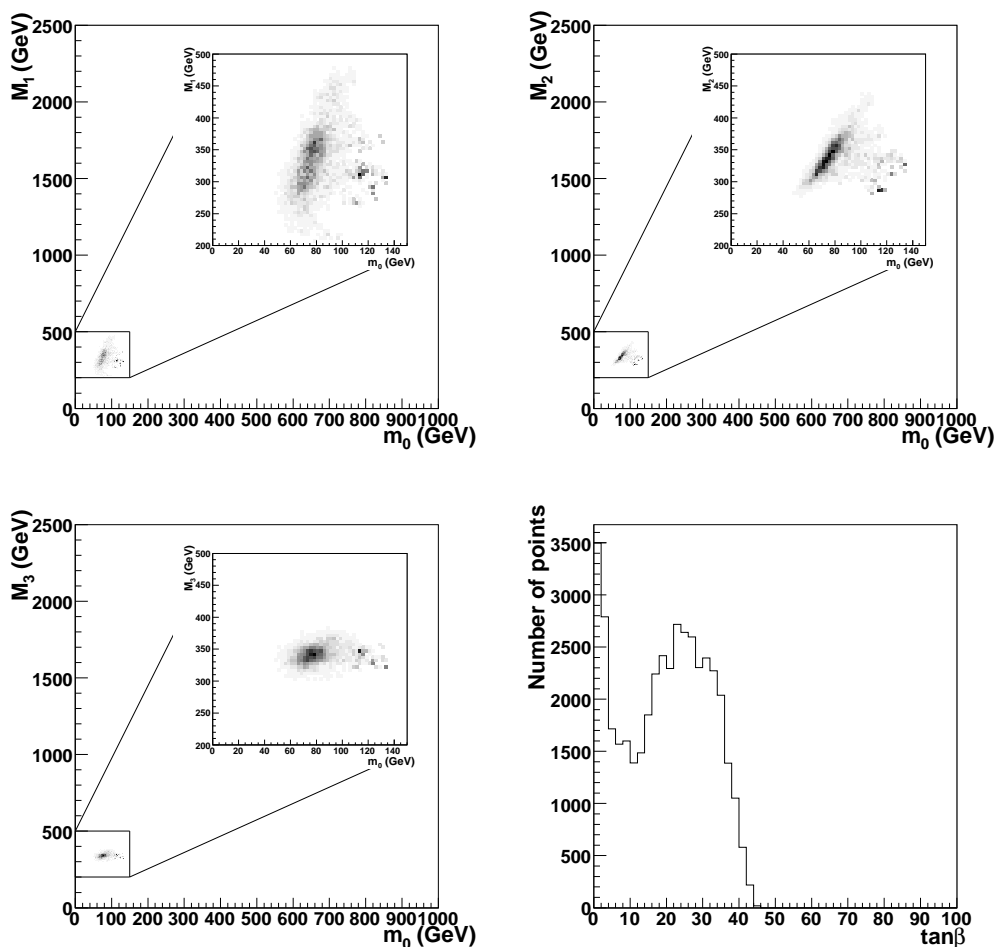The work described here is the first step toward what is hoped will be a powerful

Figure 23: The region of our non-universal SUGRA parameter space consistent with the endpoint measurements of section 2 and the cross-section measurement, with chain ambiguity included. Results are shown for negative μ.

technique for future analysis. By collecting inclusive observables, one can start to look at more and more general models, with the final result limited only by the ability of physicists to come up with new pieces of information. At the very least, the Markov Chain approach is a powerful framework for combining information and exploring multi-dimensional parameter spaces in an efficient manner.

As a final note, it is worth remarking that the technique is not limited solely to data obtained at the LHC. Any piece of relevant data is potentially useful, with obvious examples being cross-section limits for rare decay processes, and dark matter measurements that are currently already being used to set limits on theories (see, for example, [14, 15] and [16]). As we start to explore models with greater numbers of parameters, this extra knowledge could prove invaluable in providing a sufficient number of constraints, and this will be the subject of future papers.
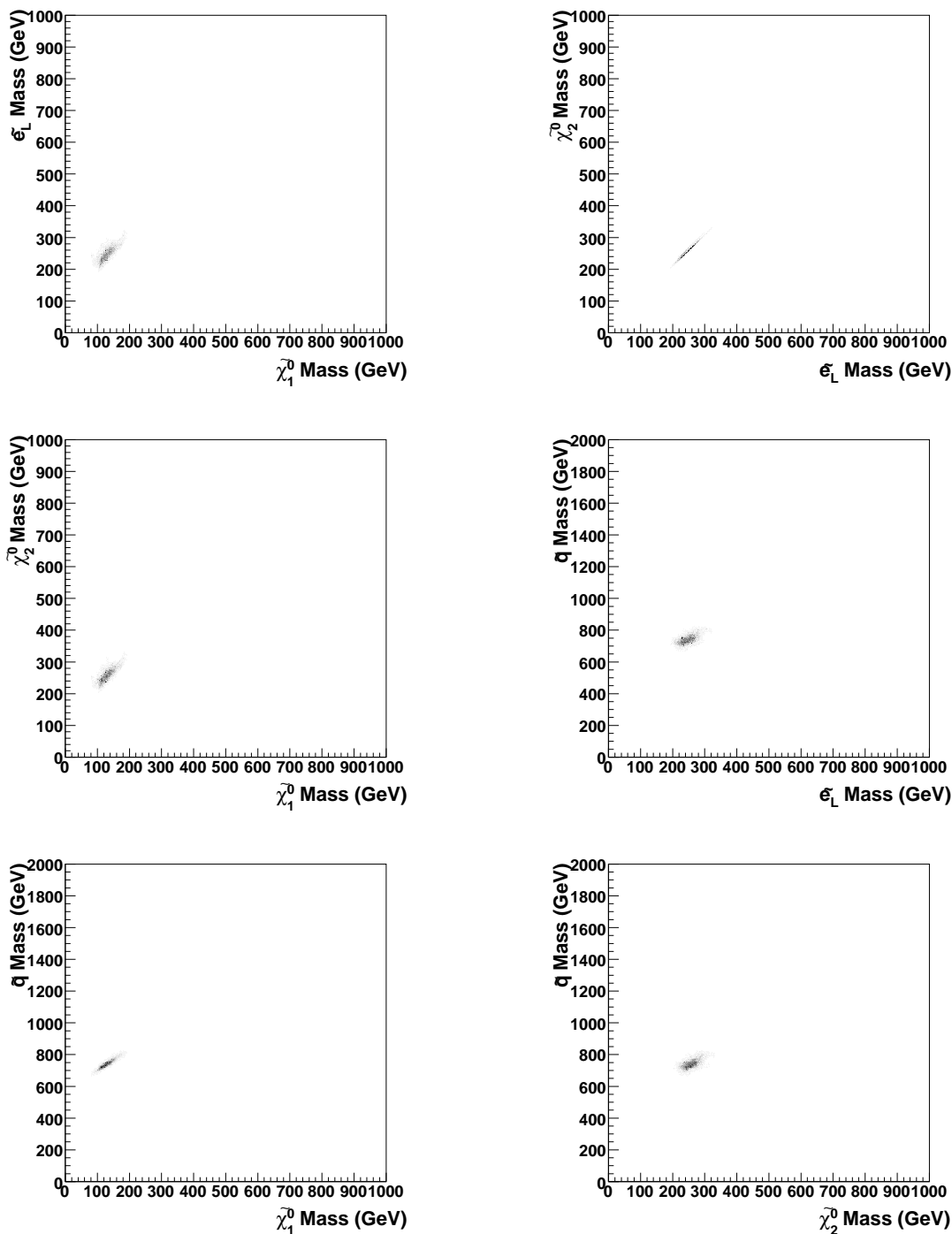
Figure 24: The region of mass space corresponding to the non-universal SUGRA parameter space region obtained in the text.

Acknowledgments

A. Markov chain sampling

There follows a brief review of the relevant techniques involved in the Markov chain methods used in our analysis. For a more comprehensive explanation, see [17]).

A.1 Markov chains

Let X_i be a (possibly infinite) discrete sequence of random variables. $X_1, X_2, \ldots$ is said to have the *Markov property* if:

$$P(X_{i+1} = x_{i+1}|X_i = x_i, X_{i-1} = x_{i-1}, \ldots, X_1 = x_1) = P(X_{i+1} = x_{i+1}|X_i = x_i) \quad \text{(A.1)}$$

for every sequence $x_1, \ldots, x_i, x_{i+1}$ and for every $i \geq 1$. A sequence of random variables with the Markov property is called a *Markov chain*.

Suppose i is a discrete step in a time variable. The Markov property is then equivalent to stating that, given a present element of the sequence X_i, the conditional probability of the next element in the sequence is dependent only on the present. Thus, at each time i the future of the process is conditionally independent of the past given the present.

A.2 Sampling and probability distributions

Suppose we wish to determine a probability distribution $P(\mathbf{x})$; for example, the posterior probability of a model's parameters given some data. It is assumed in general that $\mathbf{x}$ is an N-dimensional vector and that $P(\mathbf{x})$ can be evaluated only to within a normalisation constant Z; i.e. we can evaluate the function $P^*(\mathbf{x})$ where:

$$P(\mathbf{x}) = \frac{P^*(\mathbf{x})}{Z} \quad \text{(A.2)}$$

Although $P(\mathbf{x})$ cannot be obtained analytically, we can in theory solve the problem by sampling from $P(\mathbf{x})$ and plotting the results. Two immediate problems present themselves; the first is that Z is in general unknown. The second, which holds true even if we know Z, is that it is not obvious how to sample from $P(\mathbf{x})$ efficiently without visiting every position $\mathbf{x}$. We would like a way to visit places in $\mathbf{x}$-space where $P(\mathbf{x})$ is large in preference to places where $P(\mathbf{x})$ is small, thus giving a description of the probability distribution with a minimum of computational effort.

A.3 The Metropolis-Hastings algorithm

The above problem can be solved through the use of Markov Chain Monte Carlo methods, one example of which is the Metropolis-Hastings algorithm. This makes use of a proposal

density Q which depends on the current state of a system, which we label $\mathbf{x}^{(\mathbf{t})}$. (This state is really a point in a Markov Chain, and may be, for example, a particular choice of the parameters in the model whose probability distribution we are trying to sample). The density $Q(\mathbf{x}'; \mathbf{x}^{(\mathbf{t})})$ (where $\mathbf{x}'$ is a tentative new state, or the next point in the Markov chain) can be any fixed density from which it is possible to draw samples; it is not necessary for $Q(\mathbf{x}'; \mathbf{x}^{(\mathbf{t})})$ to resemble $P(\mathbf{x})$ for the algorithm to be useful, and it is common to choose a simple distribution such as a gaussian with a width chosen for the reasons outlined in section A.4.

Assuming that it is possible to evaluate $P^*(\mathbf{x})$ for any $\mathbf{x}$ as above, the first step in the Metropolis-Hastings algorithm is to generate a new state $\mathbf{x}'$ from the proposal density $Q(\mathbf{x}'; \mathbf{x}^{(\mathbf{t})})$. The decision on whether to accept the new state is made by computing the quantity:

$$a = \frac{P^*(\mathbf{x}')Q(\mathbf{x}^{(\mathbf{t})}; \mathbf{x}')}{P^*(\mathbf{x}^{(\mathbf{t})})Q(\mathbf{x}'; \mathbf{x}^{(\mathbf{t})})} \tag{A.3}$$

Equation (A.3) exists to ensure that the sampled distribution *does not depend on the choice of Q*.

If $a \geq 1$ the new state is accepted, otherwise the new state is accepted with probability a. It is noted that if Q is a simple symmetric density, the ratio of the Q functions in equation (A.3) is unity, in which case the Metropolis-Hastings algorithm reduces to the Metropolis method, involving a simple comparison of the target density at the two points in the Markov Chain.

If Q is chosen such that $Q(\mathbf{x}'; \mathbf{x}) > 0$ for all $\mathbf{x}, \mathbf{x}'$, the probability distribution of $\mathbf{x}^{(\mathbf{t})}$ tends to $P(\mathbf{x}) = P^*(\mathbf{x})/Z$ as $t \to \infty$. Thus, by choosing points via the Metropolis algorithm and then plotting them, we have achieved our goal of obtaining a description of $P(\mathbf{x})$ in an efficient manner.

A.4 Efficiency of the Metropolis-Hastings algorithm

Note that the presence of the caveat $t \to \infty$ implies that there is an issue of convergence in the application of the Metropolis-Hastings algorithm. Each element in the sequence $\mathbf{x}^{(\mathbf{t})}$ has a probability distribution that is dependent on the previous value $\mathbf{x}^{(\mathbf{t-1})}$ and hence, since successive samples are correlated with each other, the Markov Chain must be run for a certain length of time in order to generate samples that are effectively independent — at which point we say the chain has "converged". The time it takes for the chain to converge depends on the particular $P(\mathbf{x})$ being sampled, and on the details of Q. You cannot modify $P(\mathbf{x})$, but you are free to choose the form of Q so as to reduce the number of points which must be sampled before convergence is reached. Remember that equation (A.3) exists to ensure that the sampled distribution *does not depend on your choice of Q*.

Finding a sensible Q is a balance between choosing distributions that are wide (and thus lead to successive samples being relatively **un**-correlated) and choosing distributions which are too wide (and which then take a long time to random walk from one end of the sample space to the other). The widths of the proposal functions Q used in this paper were chosen to be as large as possible, subject to the Markov Chain's efficiency (the fraction of proposal points being accepted) not falling much below one in twenty. This choice

only affects the sampler's time to convergence and not the shape of the resultant sampled distributions once convergence has been reached.

B. Convergence of Markov chains

It is a good idea to check on the convergence of processes involving Markov chains. Results (e.g. means of quantities evaluated under the posterior distribution) could be very wrong if the markov chain has not yet fully explored the bulk of the parameter space. Even if all the important parts of the space *have* been explored, numerical results could still be adversely affected by poor equilibriation between distant parts of the distribution.

There is no 100% reliable way of assessing whether a Markov Chain has fully converged — for example there is always the possibility of there being an isolated or almost isolated region of *very large probability* which has not (yet) been found by the chain. This might happen if there were a large "barrier of poor probability" separating the explored from the unexplored regions. We cation readers therefore to take quantitative convergence measures (such as those given below) with a pinch of salt — they provide a *guide* to the convergence properties of the chain in question, but by themselves they do not tell you all you might wish to know. In the opinion of the authors, experience gained from using Markov Chain samplers combined with knowledge of the sorts of regions that the chain *ought* to be exploring (given the physical constraints of the problem) often provides a much clearer estimate of whether a chain has converged than many statistical tests provide by themselves.

In spite of the above caution, convergence measures are nonetheless informative. We have chosen to illustrate the convergence properties of the chain in this paper (for the mSUGRA data presented in section 4) using the relatively simple technique of observing the variation with time of the mean of the parameters which we are interested in.

Figures 8 and 9 display the region of mSUGRA parameter space consistent with the cross-section of events with missing p_T greater than $500\,\mathrm{GeV}$. Three parameters were varied in the study (m_0, $m_{1/2}$ and $\tan\beta$), and the mean of each of these parameters can be plotted against 'time'[9]. Convergence can be said to occur when the mean tends towards a constant value, though one has to assume that there is no hidden region of high probability that the sampler failed to find.

Results are shown in figures 25, 26 and 27. The sampler was started at a point far from the true point, and in each case one can observe a rapid move towards a region of higher probability, followed by a gradual convergence to at or better than the 10% level. The final mean value need not correspond to the value in the straw model; in the case of positive μ this is because the endpoints are mismeasured whilst in the negative μ case there is in fact no correct mean parameter value as the model being investigated has positive μ. The same method was used to assess the convergence of all other data sets in the main body of the text.

It is noted that the case for the convergence of $\tan\beta$ is less compelling than that for the other parameters, and it is suggested that this is a reflection of the fact that the

---

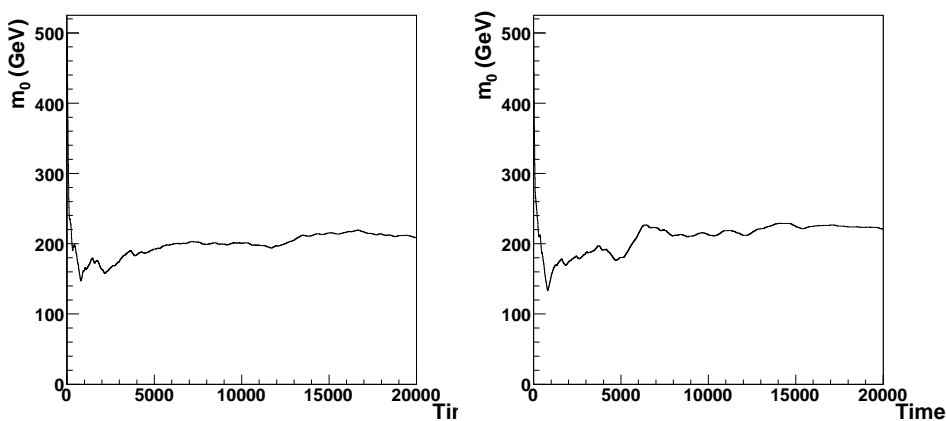[9]the time variable actually labels successive points in the Markov Chain.
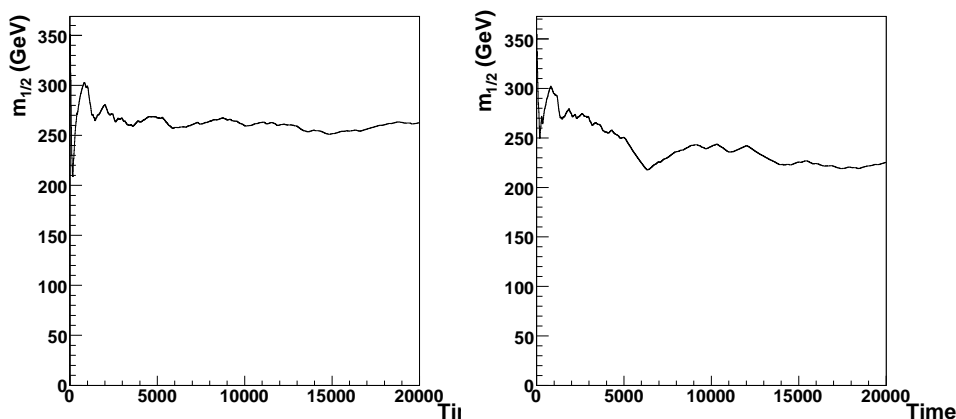
Figure 25: The variation of the mean m_0 value with time for the mSUGRA data presented in figures 8 and 9. Positive μ data is shown on the left, whilst negative μ data is on the right.

Figure 26: The variation of the mean $m_{1/2}$ value with time for the mSUGRA data presented in figures 8 and 9. Positive μ data is shown on the left, whilst negative μ data is on the right.

constraints used in the analysis have little dependence on $\tan\beta$. A flat distribution in $\tan\beta$ should ultimately arise from the sampling, and it takes a relatively long time to establish given the poor level of constraint. Essentially, this illustrates that a knowledge of the underlying physics is often as useful as quantitative tests in establishing the behaviour of a system near convergence.

Later on, when one has access to real LHC data and wishes to make statements about constants of nature, it will be much more important to test convergence rigorously, for example using methods like those in [16] based on [18].

References

[1] WMAP collaboration, D.N. Spergel et al., *First year Wilkinson microwave anisotropy probe (WMAP) observations: determination of cosmological parameters*, Astrophys. J. Suppl. **148** (2003) 175 [`astro-ph/0302209`].
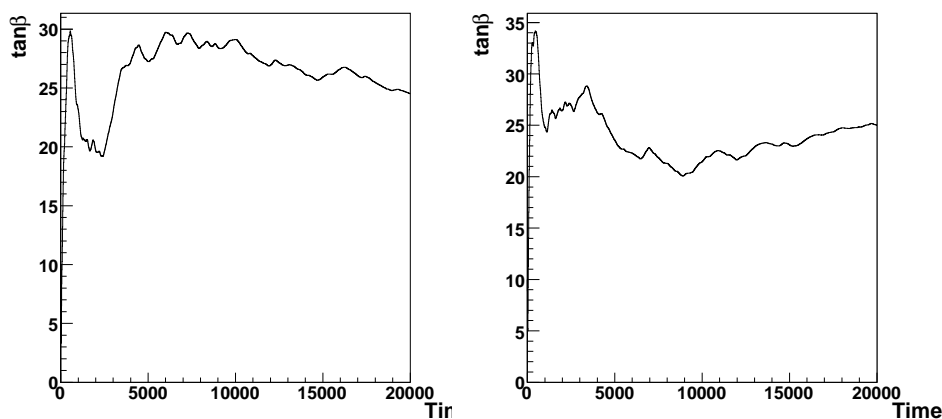
Figure 27: The variation of the mean tanβ value with time for the mSUGRA data presented in figures 8 and 9. Positive μ data is shown on the left, whilst negative μ data is on the right.

[2] M. Battaglia et al., *Proposed post-lep Benchmarks for supersymmetry*, *Eur. Phys. J.* **C 22** (2001) 535 [hep-ph/0106204].

[3] M. Battaglia et al., *Updated post-WMAP Benchmarks for supersymmetry*, *Eur. Phys. J.* **C 33** (2004) 273 [hep-ph/0306219].

[4] G. Corcella et al., *Herwig 6: an event generator for hadron emission reactions with interfering gluons (including supersymmetric processes)*, *JHEP* **01** (2001) 010 [hep-ph/0011363].

[5] G. Corcella et al., *Herwig 6.5 release note*, hep-ph/0210213.

[6] S. Moretti, K. Odagiri, P. Richardson, M.H. Seymour and B.R. Webber, *Implementation of supersymmetric processes in the herwig event generator*, *JHEP* **04** (2002) 028 [hep-ph/0204123].

[7] C.G. Lester, *Model independant sparticle mass measurements at ATLAS*, PhD thesis, University of Cambridge, Cambridge 2001.

[8] B.K. Gjelsten, D.J. Miller and P. Osland, *Measurement of SUSY masses via cascade decays for SPS 1a*, *JHEP* **12** (2004) 003 [hep-ph/0410303].

[9] ATLAS collaboration, *ATLAS detector and physics performance Technical Design Report*, CERN/LHCC 99-15, 1999.

[10] B.C. Allanach, C.G. Lester, M.A. Parker and B.R. Webber, *Measuring sparticle masses in non-universal string inspired models at the LHC*, *JHEP* **09** (2000) 004 [hep-ph/0007009].

[11] E. Richter-Was, D. Froidevaux, and L. Poggioli, *ATLAS Internal Note*, ATL-PHYS-98-131, 1999.

[12] F.E. Paige, S.D. Protopescu, H. Baer and X. Tata, *Isajet 7.69: a Monte Carlo event generator for pp, $\bar{p}p$ and e^+e^- reactions*, hep-ph/0312045.

[13] G. Comune, *Co-annihilation model parameter measurements*, talk given to the CERN-ATLAS Supersymmetry Working Group, 16 February 2005. http://agenda.cern.ch/fullAgenda.php?ida=a05518.

[14] H. Baer and C. Balazs, χ^2 *analysis of the minimal supergravity model including WMAP,* $g(\mu) - 2$ *and* $b \rightarrow s\gamma$ *constraints, JCAP* **05** (2003) 006 [`hep-ph/0303114`].

[15] E.A. Baltz and P. Gondolo, *Markov chain Monte Carlo exploration of minimal supergravity with implications for dark matter, JHEP* **10** (2004) 052 [`hep-ph/0407039`].

[16] B.C. Allanach and C.G. Lester, *Multi-dimensional mSUGRA likelihood maps,* `hep-ph/0507283`.

[17] D.J.C. MacKay, *Information theory, inference and learning algorithms.* Cambridge University Press, Cambridge 2003.

[18] A. Gelman and D.B. Rubin, *Inference from iterative simulation using multiple sequences, Statistical Science* **4** (1992) 457.