A variable for measuring masses at hadron colliders when missing energy is expected; m_{T2}: the truth behind the glamour

Alan Barr, Christopher Lester and Phil Stephens

Cavendish Laboratory, University of Cambridge, Madingley Road, Cambridge, CB3 0HE, UK

E-mail: barr@hep.phy.cam.ac.uk, lester@hep.phy.cam.ac.uk and stephens@hep.phy.cam.ac.uk

Abstract
We present the kinematic variable, m_{T2}, which is in some ways similar to the more familiar 'transverse mass', but which can be used in events where two or more particles have escaped detection. We define this variable and describe the event topologies to which it applies, then present some of its mathematical properties. We then briefly discuss two case studies which show how m_{T2} is vital when reconstructing the masses of supersymmetric particles in mSUGRA-like and AMSB-like scenarios at the large hadron collider.

(Some figures in this article are in colour only in the electronic version)

1. Introduction

Reconstructing *R*-parity conserving supersymmetric events will be difficult at the large hadronic collider (LHC) because of the following factors which limit our knowledge of the event:

- two massive particles have escaped undetected;
- the masses of these particles are unknown;
- the masses of their 'parent' particles are unknown;
- the centre-of-mass energy of the collision is not known, and
- the boost along the beam axis of the collision centre-of-mass is not known either.

An example of such an event is shown schematically in figure 1, where a pair of supersymmetric particles have been produced, each of which has decayed to some visible and some invisible daughters.

An important question to ask is 'What model-independent information about sparticle masses can be deduced from events of this type?'. The question can be seen to be a harder version of a number of older problems with which we are more familiar.
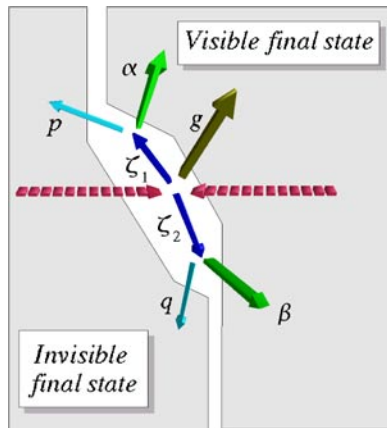

Figure 1. Schematic representation of a simple R-parity conserving event at the LHC in which supersymmetric particles were pair produced. The colliding protons are shown coming in from the left and right. The collision has pair produced two massive susy particles, ζ_1 and ζ_2 (dark blue). Each of these has been shown decaying to something visible (α or β) and to an undetected neutralino (p or q). The typical event will also contain some initial- or final-state radiation, or other debris, represented here by g. In this figure it has been assumed that g consists entirely of visible particles.

Searches at the large electron–positron collider (LEP) for pair-produced sparticles have much in common with this problem, although to first order they did not suffer from the latter two of the above problems. In some ways, the problem at the LHC is more akin to that faced by the other hadronic collider experiments, such as UA1, UA2, CDF or DO [1–4] where the W-mass has been measured from its decay to a lepton and a neutrino. This has been achieved using the 'transverse mass' event variable, m_T [5], a variable which on an event-by-event basis generates a lower bound for the W-mass, and the endpoint of whose distribution is the W-mass.

A generalization of m_T, known as the 'cluster transverse mass' [6], or 'minimum invariant mass' [7] was proposed for cases where more than one visible particle was observed. The interest at that time was motivated by the search for the leptonic decay of a t-quark with mass around 35 GeV.

It was noted in [7] that extensions to the standard model (SM), such as supersymmetry, might be investigated using similar variables. It is just such a generalized variable that is presented here. The generalization is required for supersymmetry searches at the LHC, since there will be at least *two* partially invisible decays per event, and in addition the masses of the daughter particles are not necessarily known.

The approach to the problem first proposed in [8] and subsequently developed in [9, 10] proposes the creation of a new kinematic variable, m_{T2},[1] analogous to the transverse mass, whose kinematic endpoint carries model-independent information about (to first order) the mass difference between the primary and the secondary supersymmetric particles. Mention is made in [10] of generalizations to this variable (m_{T3}, m_{T4}, ...) which may be used when events contain extra missing particles (e.g. neutrinos) as well as the two neutralinos.

The purpose of this paper is not to discuss new physics results which might be obtained with m_{T2} (for these the reader is encouraged to read [9, 10]) but rather it aims to take a closer look at more technical issues concerning the use and interpretation of m_{T2}, and its related

---

[1] Because of its use in supersymmetric events, m_{T2} has acquired the nickname of the 'stransverse' mass.

variables. It is hoped that by concentrating information on m_{T2} in this way, this paper can act as a repository of m_{T2} 'know-how' for future investigations.

2. A concrete example

It is perhaps easiest to introduce and motivate the definition of the Cambridge m_{T2} variable using a concrete example. This allows the ingredients that make up m_{T2} to be introduced, one at a time in an almost 'natural' way. Readers who would prefer a 'top down' description of m_{T2}, i.e. a description which starts with a definition and then works towards its consequences, are directed to skip to section 3 where this approach is taken.

The concrete example which will be used here is taken from [10]. This paper considers an (anomaly-mediated) R-parity conserving supersymmetric model whose key property is that it predicts a lightest chargino nearly mass degenerate with the lightest neutralino. With particular choices of model parameters, the only chargino decay mode available is

$$\chi_1^+ \to \chi_1^0 \pi^+, \tag{1}$$

which produces isolated pions which were selected in events with large missing momentum. Of course one could equally well consider pairs of squark decays such as $\tilde{q}_R \to q\chi_1^0$ or slepton pair production and decay $\tilde{\ell} \to \ell\chi_1^0$. One would then substitute 'quark-jet' or 'lepton' for 'pion' in the following discussion.

Events containing two such decays, i.e. events containing two simultaneous decays of an unseen particle of unknown mass into another invisible particle of unknown mass and visible particle, are exactly the sort of events that we hope to analyse with m_{T2}. This we shall now begin to do.

Considering for the moment just one of the decays of the form (1), one can write the Lorentz-invariant statement,

$$m_{\chi_1^+}^2 = m_\pi^2 + m_{\chi_1^0}^2 + 2\left[E_T^\pi E_T^{\chi_1^0} \cosh(\Delta\eta) - \mathbf{p}_T^\pi \cdot \mathbf{p}_T^{\chi_1^0} \right], \tag{2}$$

where $\mathbf{p}_T^\pi$ and $\mathbf{p}_T^{\chi_1^0}$ indicate pion and neutralino 2-vectors in the transverse plane, and the transverse energies are defined by

$$E_T^\pi = \sqrt{\left(\mathbf{p}_T^\pi\right)^2 + m_\pi^2} \qquad \text{and} \qquad E_T^{\chi_1^0} = \sqrt{\left(\mathbf{p}_T^{\chi_1^0}\right)^2 + m_{\chi_1^0}^2}. \tag{3}$$

Also

$$\eta = \frac{1}{2} \log\left[\frac{E + p_z}{E - p_z} \right], \tag{4}$$

is the true rapidity, so that

$$\tanh\eta = p_z/E, \qquad \sinh\eta = p_z/E_T, \qquad \cosh\eta = E/E_T. \tag{5}$$

In a hadron collider, only the transverse components of a missing particle's momentum can be inferred, so it is useful to define the transverse mass,

$$m_T^2\left(\mathbf{p}_T^\pi, \mathbf{p}_T^{\chi_1^0}; m_{\chi_1^0}\right) \equiv m_{\pi^+}^2 + m_{\chi_1^0}^2 + 2\left(E_T^\pi E_T^{\chi_1^0} - \mathbf{p}_T^\pi \cdot \mathbf{p}_T^{\chi_1^0} \right), \tag{6}$$

which, because $\cosh(x) \geqslant 1$, is less than or equal to the mass of the lightest chargino, with equality only when the rapidity difference between the neutralino and the pion, $\Delta\eta_{\chi_1^0\pi}$ is zero. All other $\Delta\eta$ lead to $m_T < m_{\chi_1^+}$, so *if* we knew the neutralino momentum, we could use m_T to give an event-by-event lower bound on the lightest chargino mass. m_T has been used in this way in the measurement of the $W^\pm$-mass.

In events considered in this example, however, there are expected to be two unseen lightest supersymmetric particles (LSPs)[2]. Since only the *sum* of the missing transverse momentum of the two neutralinos is known, the best that can be done is to evaluate the quantity

$$\min_{\mathbf{q}_T^{(1)}+\mathbf{q}_T^{(2)}=\, \not{\mathbf{p}}_T}\left[\max\left\{m_T^2\left(\mathbf{p}_T^{\pi^{(1)}},\mathbf{q}_T^{(1)};m_{\chi_1^0}\right),m_T^2\left(\mathbf{p}_T^{\pi^{(2)}},\mathbf{q}_T^{(2)};m_{\chi_1^0}\right)\right\}\right], \tag{7}$$

which is thus a *lower bound* on the square of the transverse mass, m_T, for events where two decays of the type (1) occur. Note that this minimization has forced us to introduce a pair of dummy 2-vectors $\mathbf{q}_T^{(1)}$ and $\mathbf{q}_T^{(2)}$ which, constrained by the minimization condition, parametrize our lack of knowledge about the two *true* neutralino momenta. Finally, we must recognize that under most circumstances, the value of $m_{\chi_1^0}$ is unlikely to be known, or may only be known with limited precision. In order to make our ignorance of $m_{\chi_1^0}$ explicit, we thus define a new free parameter, χ, calling it the 'neutralino mass parameter', intending it to denote any guess we might have as to the true neutralino mass $m_{\chi_1^0}$. Using it in place of m_χ, we convert (7) into the following definition of a new kinematic variable:

$$m_{T2}^2(\chi)\equiv\min_{\mathbf{q}_T^{(1)}+\mathbf{q}_T^{(2)}=\, \not{\mathbf{p}}_T}\left[\max\left\{m_T^2\left(\mathbf{p}_T^{\pi^{(1)}},\mathbf{q}_T^{(1)};\chi\right),m_T^2\left(\mathbf{p}_T^{\pi^{(2)}},\mathbf{q}_T^{(2)};\chi\right)\right\}\right]. \tag{8}$$

The quantity defined in (8) is the Cambridge m_{T2} variable which is the subject of this paper.

Staying within the framework of this example, we can now go on to describe some of the desirable model-independent properties which m_{T2} possesses.

2.1. Properties of $m_{T2}(\chi)$

Firstly, it is worth noting that the m_{T2} variable is not strictly a 'variable', and would more correctly be termed a 'function', as it retains a dependence on the unknown parameter χ. Ideally, χ would be set equal to the mass of the missing heavy particle, but in most of the situations in which the variable is likely to be used, the mass of the invisible object is unlikely to be known, or may only be known with a large uncertainty. The χ dependence remains, therefore. A more detailed discussion of how this can affect the use of m_{T2} is given in section 2.2.2.

Secondly, *from its method of construction*, it is clear that for any given event

$$m_\pi+m_{\chi_1^0}\leqslant m_{T2}\left(m_{\chi_1^0}\right)\leqslant m_{\chi_1^+}, \tag{9}$$

and

$$m_\pi+\chi\leqslant m_{T2}(\chi). \tag{10}$$

It is certainly not immediately clear, however, that events can always exist for which m_{T2} is capable of reaching all of these endpoints. In fact it turns out that such events do always exist, and proof of this is given in section 3.3. So, having defined the quantity $m_{T2}^{\max}(\chi)$ by

$$m_{T2}^{\max}(\chi)=\max_{\text{many events}}\left[m_{T2}(\chi)\right], \tag{11}$$

the important result to draw from all of this is that the upper kinematic limit of m_{T2} satisfies

$$m_{T2}^{\max}\left(m_{\chi_1^0}\right)=m_{\chi_1^+}. \tag{12}$$

This is the main model-independent statement that m_{T2} is able to offer.

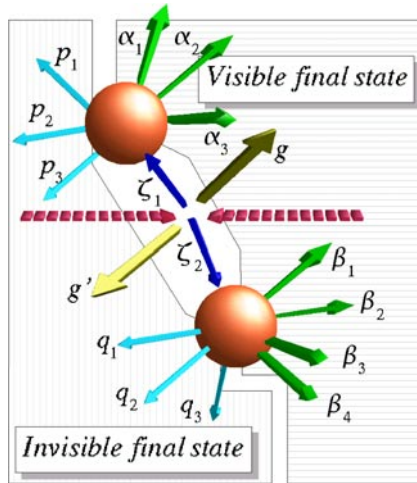[2] Though there may also be other unseen particles (see section 2.2.1).

Figure 2. Schematic representation of an *R*-parity conserving event at the LHC in which supersymmetric particles are pair produced. The colliding protons are shown coming in from the left and right. The collision has pair produced two massive susy particles, ζ_1 and ζ_2 (dark blue). Each of these has been shown decaying to a collection of visible particles (α_i or β_i) and to a set of undetected particles (p_i or q_i). The purpose of the large spherical blobs is to hide the details of the decay process(es) involved; in principle they may contain anything, from one large *n*-body decay, to $n - 1$ successive two-body decays. The typical event will also contain some initial- or final-state radiation, or other debris, represented here by g (the visible component) and g' (the invisible component). Comments in the text apply only principally to events in which g' is small enough to be neglected.

2.2. Going beyond pairs of two-body decays

The scenario in which m_{T2} has been introduced, thus far, is relatively simple; each event contains a pair of charginos, and each of these decays via a two-body decay into a charged pion and an unseen neutralino. We will now consider in more detail what happens when:

- the neutralinos are not the only missing particles;
- the initial (e.g. chargino) decays are not both two-body decays, and
- $m_{T2}(\chi)$ is evaluated at values of $\chi \neq m_{\chi_1^0}$.

2.2.1. Extra missing particles and multi-particle decays.
The need for m_{T2} to be adaptable to situations in which the neutralinos are not the only unobserved final-state particles may again be demonstrated using as an example the model of [10]. In this model, there were found to be some regions of parameter space in which the three-body chargino decay

$$\chi_1^{\pm} \rightarrow l^{\pm} \nu \chi_1^0, \tag{13}$$

had a rate comparable to that of the two-body decay (1) which we have already seen. Here, the presence of the neutrino (or antineutrino) in the final state means that we have even less information about the event than before. Nevertheless, one would like to benefit, if possible, from events in which one or two of these decays occur in place of the usual two-body decays. This type of event is just one of the general class of events depicted in figure 2.

It is clear that one can immediately generalize the m_{T2} of (8) to suit events such as those in figure 2 in the following way. Define the new variable m_{TX} by

$$m_{TX}^2 = \min_{\text{consistent splittings}} \left[\max \left\{ \left(\sum_i \alpha_i + \sum_j p_j \right)^2, \left(\sum_i \beta_i + \sum_j q_j \right)^2 \right\} \right]. \qquad (14)$$

The phrase 'consistent splittings', describing the constraint on the overall minimization, needs a little explanation. There are two sets of unknown momenta. The first of these is $\mathcal{F} = \{p_i\} \cup \{q_j\}$, containing the unknown momenta of all the unobserved *final*-state particles. The other set, $\mathcal{H} = \{h_i\}$, contains the momenta of any on-mass-shell particles which were present at an intermediate stage during the decays of the initial pair of sparticles (ζ_1 or ζ_2) to their final states. In other words, $\mathcal{H}$ contains the momenta of any intermediate particles *hidden* within the large blobs in figure 2. Minimization over 'consistent splittings', then, means minimization over all $p_i, q_j \in \mathcal{F}$ and all $h_i \in \mathcal{H}$ subject to:

- all p_i, q_i and h_i being on their respective mass shells;
- momenta being conserved at all 'hidden' vertices in which a short-lived intermediate particle with momentum $h_i \in \mathcal{H}$ decays, and
- the transverse components of $B = \sum_i p_i + \sum_j q_j$ being consistent with the measured missing momentum $\not{p}_T$.

It is because the last of these requirements that we need events in which g', the momentum carried by any invisible particles which are not descendants of a supersymmetric particle (see figure 2) is negligible. Were there to be a large tail in the distribution of g', this would degrade the performance of m_{TX} and m_{T2}.

Example. We illustrate the remarks of the previous section by returning to the example of [10] in the case where charginos could decay either by the two-body decay of (1) or the three-body decay (13). We can categorize events in this scenario by the number of missing particles in the event. When both charginos decay via (1) we only have *two* missing particles (the neutralinos). For each three-body decay which takes the place of one of these two-body decays we gain an *extra* missing particle in the form of a neutrino (or antineutrino). In short, the three categories of events could be summarized as those containing one of the following:

$$\chi_1^\pm \chi_1^\pm \rightarrow \left\{ \pi^\pm \chi_1^0 \pi^\pm \chi_1^0, \text{ or } e^\pm \nu \chi_1^0 \pi^\pm \chi_1^0, \text{ or } e^\pm \nu \chi_1^0 e^\pm \nu \chi_1^0 \right\}.$$

The events had been produced by a phase-space-only Monte Carlo generator. Three distributions of the quantity m_{TX}, defined in (14), were then generated from each of these sets of events. Using the number of missing particles to categorize these events, the values of m_{TX} measured in each case are referred to as m_{T2}, m_{T3} and m_{T4}. The resulting distributions for $m_{TX}(m_{\chi_1^0}) - m_{\chi_1^0}$ are shown in figure 3.

It has already been mentioned that a key property of m_{T2} is that the kinematic endpoint of its distribution occurs at $m_{T2}^{\max}(m_{\chi_1^0}) = m_{\chi_1^+}$ and so it is reassuring to see in figure 3 that a large number of events reach this endpoint. In the vicinity of the endpoint, the edge is seen to be sharp and near vertical. This shows that at the partonic level a measurement of $m_{T2}^{\max}$ would provide an excellent constraint on the masses of the sparticles involved. In section 4 plots from [9, 10], which include realistic detector effects, will show that the subsequent smearing of the m_{T2} edge, while significant, is still small.

Looking next at the m_{T3} and m_{T4} distributions, it is clear that the event fall-off in the vicinity of their kinematic endpoints is much less steep than in the case of m_{T2}. This is hardly surprising, given the reduced amount of information available in these events. Later,
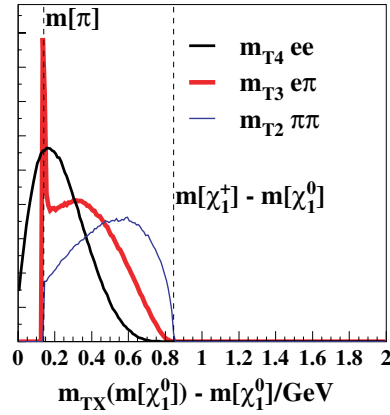
Figure 3. Simulations of $m_{TX}(m_{\chi_1^0}) - m_{\chi_1^0}$ for $X = 2, 3, 4$ using a simple phase-space Monte Carlo generator program for a pair of decays $\tilde{q} \to \chi_1^+ q$ followed by $\chi_1^+ \to \chi_1^0 \pi$ or $\chi_1^+ \to \chi_1^0 e \nu_e$. As the number of invisible particles increases, the proportion of events near the upper limit decreases. Within the figure, subscripts are indicated by square brackets.

in section 3.4, the relative fraction of events in the vicinity of the edge will be seen, more quantitatively, to be due to the larger number of simultaneous conditions that events near the edge must satisfy. Although the endpoint, itself, becomes increasingly harder to detect as the number of missing particles increases, the m_{TX} distributions are *all* capable of inferring the mass scale associated with (in this case) $m_{\chi_1^+} - m_{\chi_1^0}$ from the overall widths of their distributions, which each scale with the endpoint position, albeit with some dependence on the decays themselves, and on factors such as the detector acceptance over the width of the distribution.

Finally, one notes that the m_{T3} distribution has sharp peak at $m_{T3} = m_{\chi_1^0} + m_\pi$, not seen in the m_{T2} and m_{T4} distributions. It will be shown later, in section 3.5, that this is an effect which can occur whenever the hypothesized decays on each side of the event are different[3].

2.2.2. Other values of χ. Now we return to a brief look at the effect of evaluating m_{TX} distributions at values of χ different to the true neutralino mass.

Figure 4 shows the same data as in figure 3, but in addition it shows the distributions that would be obtained by evaluating $m_{TX}(\chi)$ at values of $\chi = m_{\chi_1^0} \pm 10\%$. In this particular example, where $m_{\chi_1^0} = 162$ GeV, 10% (16 GeV) errors in χ result in similar *fractional* errors in $\Delta M_{\tilde{\chi}_1}$, i.e. of a few tens of MeV. This shows that m_{T2} can be sensitive to small mass differences. In this example, too, we see a *positive* correlation between the change in χ and the change in the position of the endpoint. These examples are not always typical, however. For example, in [8] the authors considered m_{T2} in the context of the double decay of a pair of sleptons, each to a lepton and a neutralino at SUGRA point 5, one of the five supergravity points proposed at [11] and described in [12]. In this model, the difference in mass between the decaying and final sparticles ($157.1 - 121.5 = 35.6$ GeV) is approximately 40 times larger than in the AMSB case, and so at SUGRA point 5 we see a *negative* correlation between changes in

---

[3] 'Different', in this context, means 'being such that the minimum total invariant mass attainable by the particles on one side of the event is not equal to the minimum total invariant mass attainable by the particles on the other side of the event'. This happens principally when the particle content of each decay differs. In the case of m_{T3} in, the AMSB example scenario, the two dissimilar minima are $m_\pi + m_{\chi_1^0}$ and $m_e + m_\nu + m_{\chi_1^0}$.
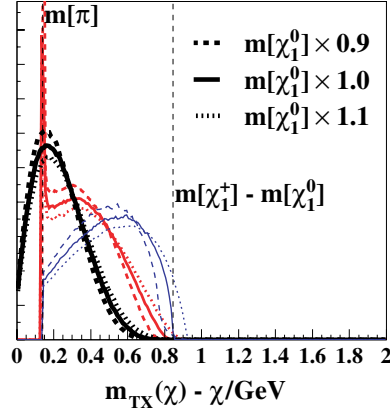
Figure 4. The distortion of $m_{TX}(\chi) - \chi$ when the LSP mass parameter, χ, is varied by $\pm 10\%$ about the 'ideal' value of $m_{\chi_1^0}$. These curves show that $m_{TX}(\chi) - \chi$ remains sensitive to the mass difference $\Delta M_{\tilde{\chi}_1} = m_{\chi_1^+} - m_{\chi_1^0}$. In this simulation $\Delta M_{\tilde{\chi}_1} = 0.845$ GeV, $m_{\chi_1^0} = 161.6$ GeV, and the electron and neutrino mass were neglected. The normalization is arbitrary. Within the figure, subscripts are indicated by square brackets.
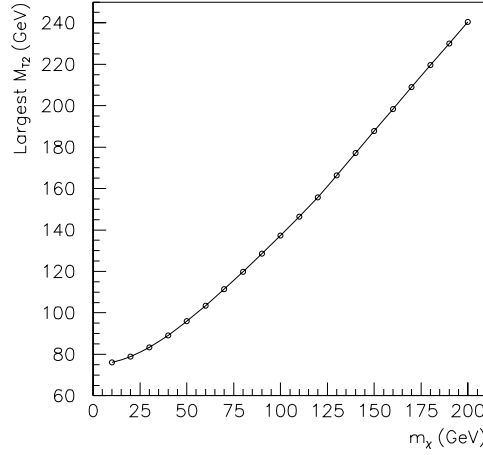

Figure 5. Variation of $m_{T2}^{\max}(\chi)$ with χ for a set of $\tilde{l}^+\tilde{l}^- \rightarrow l^+\chi_1^0 l^-\chi_1^0$ events generated by a phase-space Monte Carlo using $m_{\tilde{l}} = 157.1$ GeV and $m_{\chi_1^0} = 121.5$ GeV. Note that $m_{T2}^{\max}(\chi) - \chi$ decreases as χ increases. The asymptote has unit gradient.

χ and $m_{T2}(\chi) - \chi$. This is illustrated in figure 5. Differing kinds of behaviour, such as these, are typical of a variable such as m_{T2} which has input scales (e.g. m_π, $m_{\chi_1^+}$, χ and $m_{\chi_1^+} - \chi$) which can have a large number of relative hierarchies associated with them. For example, the AMSB scenario has $m_\pi \approx m_{\chi_1^+} - \chi \ll \chi \leqslant m_{\chi_1^+}$, while SUGRA point 5 has $m_l \ll m_{\tilde{l}} - \chi \approx \chi < m_{\tilde{l}}$.

We now take a final look at how m_{T2} depends upon χ by looking not just at events near the kinematic endpoint, but at events in general. To help, we define $y(\chi)$, a rescaling of $m_{T2}(\chi)$, as follows:

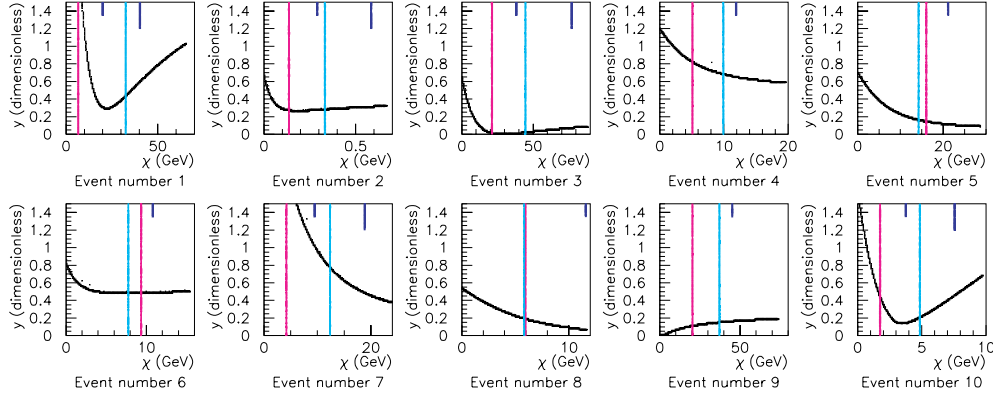$$y(\chi) \equiv \frac{m_{T2}(\chi) - \chi - m_\pi}{m_{\chi_1^+} - m_{\chi_1^0} - m_\pi}. \tag{15}$$

Figure 6. These plots show examples of how $y(\chi)$, defined in (15), can depend on χ. The plots were generated using the following procedure. Ten sets of masses satisfying $m_\pi + m_{\chi_1^0} < m_{\chi_1^+}$ were randomly generated. According to each set of masses, a phase-space Monte Carlo generated a single event of the type shown in figure 1 containing two $\chi_1^\pm \to \pi^\pm \chi_1^0$ decays. The plots above show how, in each event, the value of $y(\chi)$ (a dimensionless rescaling of $m_{T2}(\chi)$) depended upon χ over the range $0 < \chi < 2m_{\chi_1^0}$. The true value of the neutralino mass (i.e. that used in the Monte Carlo for the decay) is marked by the vertical line at the centre of each plot (cyan), while the other vertical line marks the value of the pion mass (magenta). The short and long vertical ticks (dark blue) mark $m_{\chi_1^+}/2$ and $m_{\chi_1^+}$, respectively.

By looking back at (9) and (12), this variable can be seen to map $m_{T2}\big(m_{\chi_1^0}\big)$ into the range $[0, 1]$.[4] This makes it easier to compare values of m_{T2} coming from events with different sparticle masses. A value of $y\big(m_{\chi_1^0}\big)$ close to 0 (or 1) indicates an event close to the lower (or upper) kinematic endpoint of the $m_{T2}\big(m_{\chi_1^0}\big)$ distribution. Figure 6 shows how $y(\chi)$ varies with χ for ten random events, each generated using a random set of masses satisfying $m_\pi + m_{\chi_1^0} < m_{\chi_1^+}$ as described in the figure caption. The main conclusion to draw from these plots is that there is no easy way to say, in advance, how $m_{T2}(\chi)$ will vary with χ in a given event, even in the vicinity of $m_{\chi_1^0}$. In general $m_{T2}(\chi)$ can rise, fall or even be stationary with respect to to χ near $m_{\chi_1^0}$, depending on the masses of the particles involved in the decays.

3. Some mathematical results concerning the variable m_{TX}

In section 2, m_{T2} and its analogues were defined using laboratory-frame momenta and with minimization conditions (such as $\not{q}_T^{(1)} + \not{q}_T^{(2)} = \not{p}_T$) not specified in Lorentz-invariant forms. The definition of m_{T2} from section 2 may be summarized as follows:

$$m_{T2}^2\big(\mathbf{p}_T^{l_1}, \mathbf{p}_T^{l_2}, \not{\mathbf{p}}_T; \chi\big) \equiv \min_{\not{\mathbf{p}}_1 + \not{\mathbf{p}}_2 = \not{\mathbf{p}}_T} \big[\max\{m_T^2\big(\mathbf{p}_T^{l_1}, \not{\mathbf{p}}_1; \chi\big), m_T^2\big(\mathbf{p}_T^{l_2}, \not{\mathbf{p}}_2; \chi\big)\}\big], \tag{16}$$

where

$$m_T^2\big(\mathbf{p}_T^l, \mathbf{p}_T^{\tilde{\chi}}; \chi\big) \equiv m_l^2 + \chi^2 + 2\big(E_T^l E_T^{\tilde{\chi}} - \mathbf{p}_T^l \cdot \mathbf{p}_T^{\tilde{\chi}}\big), \tag{17}$$

in which $E_T^l = \sqrt{|\mathbf{p}_T^l|^2 + m_l^2}$, $E_T^{\tilde{\chi}} = \sqrt{|\mathbf{p}_T^{\tilde{\chi}}|^2 + \chi^2}$. In the following sections we will replace these definitions by equivalent, but explicitly Lorentz-invariant ones, which are easier to manipulate mathematically.

---

[4] Note that $m_{T2}(\chi)$ is in general *not* constrained to lie in this range.

3.1. Definitions

The natural way to write m_{T2} in a manifestly (1, 2)-Lorentz-invariant form is as follows:

$$m_{T2}^2(\alpha, \beta, \Sigma, \Lambda; \chi) \equiv \min_{\left\{\substack{p + q = \sqrt{s}\Lambda - \Sigma \\ p^2 = q^2 = \chi^2}\right\}} [\max\{(\alpha + p)^2, (\beta + q)^2\}]. \qquad (18)$$

Here, m_{T2} has been written as a function of the four (1 + 2)-dimensional Lorentz vectors which describe each event (α, β, Σ and Λ) and one real parameter χ. The transverse Lorentz vectors of the two visible particles coming from each of the hidden decays are represented by α and β, while Σ represents the total transverse energy–momentum seen in the event. This is consistent with the notation used in figure 1. The only new vector, Λ, defines the laboratory frame by being the (1, 2)-energy–momentum of a particle of unit mass ($\Lambda^2 = 1$) at rest in the laboratory. The total transverse momentum of the event (visible and invisible) can only be assumed to be zero in the laboratory frame, and so knowledge of how to boost to the laboratory frame is essential. This is why Λ is needed.

Note that (18) includes a minimization over $\sqrt{s}$, a parameter which accounts for our lack of knowledge of the centre-of-mass energy of the whole event. The requirement that the hypothesized neutralino momenta are real, i.e. $(\sqrt{s}\Lambda - \Sigma)^2 \geqslant (2\chi)^2$, constrains $\sqrt{s}$ to be chosen from the region in which

$$\sqrt{s} \geqslant \Lambda \cdot \Sigma + \sqrt{4\chi^2 + ((\Lambda \cdot \Sigma)^2 - \Sigma^2)}. \qquad (19)$$

Similarly, one can also define m_{TX} in a manifestly Lorentz-invariant form:

$$m_{TX}^2(\hat{\alpha}, \hat{\beta}, \Sigma, \Lambda; \chi) \equiv \min [\max\{(\hat{\alpha} + \hat{p})^2, (\hat{\beta} + \hat{q})^2\}] \qquad \text{over}$$

$$\left\{\begin{array}{c} \hat{p} + \hat{q} = \sqrt{s}\Lambda - \Sigma, \\ p_i, q_i \text{ and } h_i \text{ all on their mass shells, and} \\ \text{momenta conserved at all internal } h \text{ decays} \end{array}\right\}, \qquad (20)$$

in which the same notation has been used as in figure 2 and equation (14), and in which the 'hats' indicate summation over all vectors of a set (e.g. $\hat{\alpha} = \sum_i \alpha_i$).

If desired, one may remove the 'max' at the expense of moving to Lorentz 4-vectors and remembering to minimize over all possible longitudinal boost of the centre of momentum (here denoted by the Lorentz boost L_z):

$$m_{T2}^2(\alpha, \beta, \Sigma, \Lambda; \chi) \equiv \min_{\left\{\substack{D^2 = D_1^2 = D_2^2 \\ p + q = \sqrt{s}L_z\Lambda - \Sigma \\ p^2 = q^2 = \chi^2}\right\}} [D^2], \qquad (21)$$

where

$$D_1^2 = (\alpha + p)^2, \qquad (22)$$

and

$$D_2^2 = (\beta + q)^2. \qquad (23)$$

This way of representing m_{T2} most clearly captures the spirit in which it provides an event-by-event lower bound on the initial sparticle mass.

3.2. Results concerning m_{T2}

In the case where both visible particles have the same mass, i.e. in the case where $\alpha^2 = \beta^2 = m_l^2$, (18) may be rewritten in the form

$$m_{T2}^2{}'(\alpha, \beta, \Sigma, \Lambda; \chi) \equiv m_l^2 + \chi^2 + \min_{\left\{\substack{p + q = \sqrt{s}\Lambda - \Sigma \\ p^2 = q^2 = \chi^2}\right\}} [2\max\{\alpha \cdot p, \beta \cdot q\}]. \qquad (24)$$

It was shown in [10] that the solution of (24) must select vectors p and q for which $\alpha \cdot p = \beta \cdot q$. Using this information, one may perform half of the minimization in (24) analytically. This allows (24) to be rewritten as a minimization over a single real variable, $\sqrt{s}$, as follows:

$$m_{T2}^2{'}(\alpha, \beta, \Sigma, \Lambda; \chi) \equiv m_l^2 + \chi^2 + \tfrac{1}{2} \min_{\sqrt{s}} \left[(\sigma \cdot B)Q - \sqrt{\sigma^2 Q - 4m_l^2} \sqrt{B^2 Q - 4\chi^2} \right], \quad (25)$$

where

$$Q = 1 - \frac{(\Delta \cdot B)^2}{(\sigma \cdot B)^2 - \sigma^2 B^2}, \qquad (\Rightarrow 0 \leqslant Q \leqslant 1) \quad (26)$$

$$\sigma = \alpha + \beta, \quad (27)$$

$$\Delta = \alpha - \beta, \quad (28)$$

and

$$B = \sqrt{s}\Lambda - \Sigma. \quad (29)$$

We should note that the constraint which has just been imposed, namely $\alpha \cdot p = \beta \cdot q$, is more stringent than the $\sqrt{s}$ constraint (19) which was only there to ensure that the hypothesized neutralinos were not tachyonic. As a consequence, the range over which $\sqrt{s}$ may be varied when performing the minimization in (25) must be replaced by the stronger condition that each of the quantities under radicals in (25) be positive.

It is interesting to note that if we define two new transverse Lorentz vectors ($\overline{\sigma}$ and $\overline{B}$) via a rescaling of existing transverse Lorentz vectors according to

$$\overline{\sigma} = \sigma \sqrt{Q}, \quad (30)$$

and

$$\overline{B} = B \sqrt{Q}, \quad (31)$$

then we can rewrite (25) in the form

$$m_{T2}^2{'}(\alpha, \beta, \Sigma, \Lambda; \chi) \equiv m_l^2 + \chi^2 + \tfrac{1}{2} \min_{\sqrt{s}} \left[(\overline{\sigma} \cdot \overline{B}) - \sqrt{\overline{\sigma}^2 - 4m_l^2} \sqrt{\overline{B}^2 - 4\chi^2} \right]. \quad (32)$$

This is not much of an improvement in itself, but it motivates the definition of two new Lorentz 4-vectors;

$$\tilde{\sigma} = \left(\overline{\sigma}, \sqrt{\overline{\sigma}^2 - 4m_l^2} \right), \quad (33)$$

and

$$\tilde{B} = \left(\overline{B}, \sqrt{\overline{B}^2 - 4\chi^2} \right), \quad (34)$$

which we see, by construction, satisfy the following fixed-mass relations:

$$m_{\tilde{\sigma}} = \sqrt{\tilde{\sigma}^2} = 2m_l, \quad (35)$$

and

$$m_{\tilde{B}} = \sqrt{\tilde{B}^2} = 2\chi. \quad (36)$$

In terms of these new Lorentz 4-vectors, then, we can finally rewrite (32) as

$$m_{T2}^2{'}(\alpha, \beta, \Sigma, \Lambda; \chi) \equiv m_l^2 + \chi^2 + \tfrac{1}{2} \min_{\sqrt{s}} (\tilde{\sigma} \cdot \tilde{B}), \quad (37)$$

or

$$m_{T2}{'}(\alpha, \beta, \Sigma, \Lambda; \chi) \equiv \tfrac{1}{2} \min_{\sqrt{s}} |\tilde{\sigma} + \tilde{B}|. \quad (38)$$

It is interesting to note that the constant mass relations (35) and (36), taken together with the definition of $m_{T2}{'}$ shown in (38), make it self-evident that the value of $m_{T2}{'}$ obtained in a given event is bounded below by $m_l + \chi$, as expected.

Approximations. To get a better idea of the way in which m_{T2} depends on its inputs, one might hope to find a concise closed-form analytic definition of the variable. Thus far, however, m_{T2} and m_{T2}' have resisted all attempts to write them in forms simpler than (16), (18) and (25), except in a few special cases. For example, in the special case of events in which the spatial part of the total visible transverse momentum is seen to be zero in the laboratory frame (i.e. events for which $(\Sigma \cdot \Lambda)^2 = \Sigma^2$) one can show that (25) is equivalent to

$$m_{T2}^2{}''(\alpha, \beta, \Lambda; \chi) \equiv m_l^2 + \chi^2 + \chi \sqrt{4(\alpha \cdot \Lambda)(\beta \cdot \Lambda) - (-\Delta^2)}$$

$$\left(= m_l^2 + \chi^2 + \chi \sqrt{2\left(E_\alpha E_\beta + m_l^2 + \mathbf{p}_T^\alpha \cdot \mathbf{p}_T^\beta\right)} \text{ in the laboratory frame}\right). \qquad (39)$$

The limit of validity of (39) can be explored as follows. The laboratory-frame energy that this special case solution assigns to p and q is given by

$$p \cdot \Lambda = q \cdot \Lambda = \frac{(\sigma \cdot \Lambda)\chi}{\sqrt{4(\alpha \cdot \Lambda)(\beta \cdot \Lambda) - (-\Delta^2)}}, \qquad (40)$$

and so the velocity of the boost needed to take the laboratory frame to the one in which the invisible particles are back to back could be written, in this special case, as

$$(0 =) \quad |\mathbf{v}|^2 = \mathbf{p}_\Sigma^2 \big/ (p \cdot \Lambda + q \cdot \Lambda)^2 \qquad (41)$$

$$= \frac{\left\{(\Sigma \cdot \Lambda)^2 - \Sigma^2\right\} \left\{4(\alpha \cdot \Lambda)(\beta \cdot \Lambda) - (-\Delta^2)\right\}}{4(\sigma \cdot \Lambda)^2 \chi^2}. \qquad (42)$$

$$\left(= \frac{\mathbf{p}_\Sigma^2 \left(E_\alpha E_\beta + m_l^2 + \mathbf{p}_T^\alpha \cdot \mathbf{p}_T^\beta\right)}{2(E_\alpha + E_\beta)^2 \chi^2} \text{ in the laboratory frame}\right). \qquad (43)$$

In the light of the above, we can interpret (39) as the leading term in an expansion of $m_{T2}^2{}'$ in powers of $|\mathbf{v}|^2$ as defined in (42). Given a particular event, all the quantities in (42) may be evaluated, so one can safely use (39) to evaluate $m_{T2}^2{}'$ for events in which $|\mathbf{v}|^2$ is observed to satisfy $|\mathbf{v}|^2 \ll 1$.[5]

3.3. Extremal values of m_{T2}

In this section, we show that the maximum value which $m_{T2}(m_{\chi_1^0})$ can attain, for a given set of particle masses, is indeed the mass of the initial sparticle[6]. We start from definition (8). We also describe the region of decay phase space which contains events which occur close to this kinematic endpoint.

To find the range of values m_{T2} may take we first let $f_1 = m_T^2\left(\mathbf{p}_T^{\pi^{(1)}}, \mathbf{q}_T^{(1)}; m_{\chi_1^0}\right)$, and $f_2 = m_T^2\left(\mathbf{p}_T^{\pi^{(2)}}, \mathbf{q}_T^{(2)}; m_{\chi_1^0}\right)$. We then note that the minimum over a parameter x of the maximum of $f_1(x)$ and $f_2(x)$ can occur at a local minimum, $f_{1(2)}'(x^*) = 0$, provided $f_{1(2)}(x^*) > f_{2(1)}(x^*)$, as shown in figure 7(a). Alternatively the minimum can occur when the functions cross one another when $f_1 = f_2$ (figure 7(b)) or at a boundary (figure 7(c)). The parameter x corresponds to the fraction of the missing momentum (in one of the transverse

---

[5] The reader is warned not to mistake $|\mathbf{v}|$ for the speed associated with an actual boost (real or conjectured) connected with the neutralino pair; $|\mathbf{v}|$ could, for example, even exceed the speed of light if χ were made sufficiently small! It should only be assumed that as $|\mathbf{v}| \to 0$, $|\mathbf{v}|$ will tend to the speed, in the laboratory frame, associated with the energy–momentum vector B. (This is the B which was originally defined in (29) and whose $\sqrt{s}$ value was selected by the minimization process in (25).)

[6] Up to this point, within the context of the AMSB example, it has only been shown that $m_{T2}(m_{\chi_1^0})$ is bounded above by $m_{\chi_1^+}$. It has not yet been shown that m_{T2} can attain this bound. The purpose of this section is to show that it can.
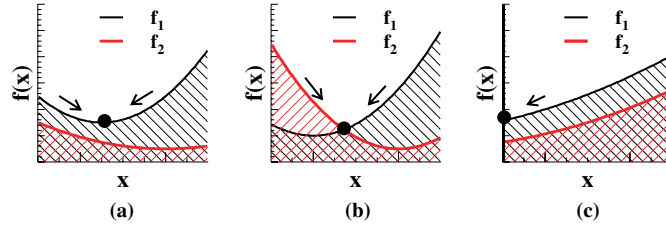
Figure 7. A diagram demonstrating that the minimization over some parameter of the maximum of two well-behaved functions may occur either at (a) a minimum value of one of them, or (b) when they are equal, or (c) at the boundary of the domain.

directions) which is assigned to each half of the event. Since $f_1, f_2 \to \infty$ as $x \to \pm\infty$ figure 7(c) is not relevant to our minimization problem.

To find which of (a) or (b) is pertinent, consider an *unconstrained* minimization over $\not{q}_T$, of $m_T^2 (\mathbf{p}_T^\pi, \not{q}_T; m_{\chi_1^0})$. Using the relationship

$$\frac{\partial \not{E}_T}{\partial \not{q}_k} = \frac{\not{q}_k}{\not{E}_T}, \tag{44}$$

where $\not{E}_T^2 = \not{q}_T^2 + m_{\chi_1^0}^2$, it is straightforward to show that

$$\frac{\partial m_T^2}{\partial \not{q}_k} = 2 \left(E_T^\pi \frac{\not{q}_k}{\not{E}_T} - p_k^\pi \right), \qquad k = 1, 2. \tag{45}$$

This means that at an unconstrained minimum of m_T^2 we have

$$\mathbf{v}_T^\pi = \not{\mathbf{u}}_T, \tag{46}$$

where we introduce the notation $\mathbf{v}_T \equiv \mathbf{p}_T / E_T$, $\not{\mathbf{u}}_T \equiv \not{q}_T / \not{E}_T$. Note that although $\mathbf{p}_T$ is the transverse momentum of a visible particle, $\mathbf{v}_T$ is not quite its transverse velocity, as it has E_T, rather than E in the denominator.

Using the basis (t, x, y) with the metric diag$(1, -1, -1)$, one can write

$$m_T^2 = \left(E_T^{\text{tot}}, \mathbf{p}_T^{\text{tot}} \right) \cdot \left(E_T^{\text{tot}}, \mathbf{p}_T^{\text{tot}} \right), \tag{47}$$

where $E_T^{\text{tot}} = E_T^\pi + \not{E}_T$ and $\mathbf{p}_T^{\text{tot}} = \mathbf{p}_T^\pi + \not{q}_T$. This (1+2)-dimensional Lorentz invariant can be evaluated in any frame boosted from the laboratory in the transverse plane. Equation (46) reveals that at the unconstrained minimum the transverse velocities $\mathbf{v}_T^\pi$ and $\not{\mathbf{u}}_T$ are equal; a statement necessarily true in all transverse frames, including the special one in which both the transverse velocities (and associated momenta) are zero. Evaluating (47) in this frame, we find that the unconstrained minimum of (47) then becomes $\left(m_\pi + m_{\chi_1^0}, 0, 0 \right) \cdot \left(m_\pi + m_{\chi_1^0}, 0, 0 \right)$, and we recover the expected result

$$m_T^{\text{min}} = m_\pi + m_{\chi_1^0}. \tag{48}$$

We therefore conclude that the function m_T^2 has only one stationary value and it is the global minimum, and is common to both sides of the event provided the same type of particles are emitted $\left(m_\pi^{(1)} = m_\pi^{(2)} \right)$. Thus when f_1 is minimum it cannot be greater than f_2, and so the minimization in (8) forces $f_1 = f_2$. This could of course occur when *both* f_1 and f_2 are at their global minima, in which case m_{T2} takes its minimum value:

$$m_{T2}^{\text{min}} = m_\pi + m_{\chi_1^0}. \tag{49}$$

To summarize, when the same particles are emitted from both sides of the event, m_{T2} may be defined as the minimum of $m_T^{(1)}$ subject to the two constraints $m_T^{(1)} = m_T^{(2)}$, and

$\mathbf{p}_T^{(1)} + \mathbf{p}_T^{(2)} = \mathbf{p}_T$. The condition for the minimization can be calculated by Lagrange-multiplier methods, the result of which is that the vectors $\mathbf{u}_T^{(1,2)}$ of the *assigned* neutralino momenta $\mathbf{q}_T^{(1,2)}$ must satisfy [7]

$$\left(\mathbf{u}_T^{(1)} - \mathbf{v}_T^{\pi^{(1)}}\right) \propto \left(\mathbf{u}_T^{(2)} - \mathbf{v}_T^{\pi^{(2)}}\right), \tag{50}$$

where again, $\mathbf{v}_T^\pi = \mathbf{p}_T^\pi / E_T^\pi$.

   To find the maximum of m_{T2} over many events we note that for each event the minimization will select hypothesized momenta satisfying (50). We now note events can occur in which the *true* transverse velocities of the neutralinos were exactly those which were assigned by the minimization, i.e. they can satisfy

$$\mathbf{v}_T^{\chi_1^0(1)} = \mathbf{u}_T^{(1)}, \qquad \mathbf{v}_T^{\chi_1^0(2)} = \mathbf{u}_T^{(2)}. \tag{51}$$

These events will have both hypothesized transverse masses equal not only to each other but also to true transverse masses which would have been calculated if the neutralino momenta had been known:

$$m_T^{(i)}\left(\mathbf{p}_T^{\pi^{(i)}}, \mathbf{p}_T^{\chi_1^0(i)}\right) = m_T^{(i)}\left(\mathbf{p}_T^{\pi^{(i)}}, \mathbf{q}_T^{(i)}\right). \tag{52}$$

If events occur where, in addition to the transverse components of the neutralino momenta satisfying (51), the rapidity differences satisfy $\eta_{\chi_1^0(1)} = \eta_{\pi(1)}$ and $\eta_{\chi_1^0(2)} = \eta_{\pi(2)}$, then by (2) m_{T2} will equal the true mass of the chargino. Combining this with (49) and recalling that m_{T2} cannot be greater than the chargino mass by construction, we can see that the event-by event distribution of m_{T2} can span the range

$$m_{\chi_1^0} + m_\pi \leqslant m_{T2} \leqslant m_{\chi_1^+}. \tag{53}$$

3.4. Extremal values of m_{TX}

In the last section we looked at the conditions under which events can generate m_{T2} values near the kinematic endpoint. Here we will look at some of the ways these conditions become modified for m_{T3} and m_{T4} events.

   Consider once again events from the AMSB scenario in which a chargino is produced and then decays to $\chi_1^0 e \nu_e$. If we expand the Lorentz invariant

$$\left(m_{\chi_1^+}\right)^2 = \left(p_{\chi_1^0} + p_e + p_\nu\right)^2, \tag{54}$$

we obtain three mass-squared terms for each of the decay particles and three cross terms. The cross terms can each be written in the form

$$2 p_a \cdot p_b = 2\left[E_T^{(a)} E_T^{(b)} \cosh(\Delta \eta_{ab}) - \mathbf{p}_T^{(a)} \cdot \mathbf{p}_T^{(b)}\right], \tag{55}$$

like the cross term in (2). If the neutralino and neutrino transverse momenta were individually known we could evaluate the transverse mass,

$$m_T^2 = m_{\chi_1^0}^2 + m_e^2 + 2\left[\left(E_T^e E_T^\chi - \mathbf{p}_T^e \cdot \mathbf{p}_T^\chi\right) + \left(E_T^\nu E_T^\chi - \mathbf{p}_T^\nu \cdot \mathbf{p}_T^\chi\right) + \left(E_T^e E_T^\nu - \mathbf{p}_T^e \cdot \mathbf{p}_T^\nu\right)\right], \tag{56}$$

where the neutrino mass is assumed to be negligible. m_T will be equal to the χ_1^+ mass in events where $\Delta \eta_{ab} = 0$ for all pairs of e, ν_e and χ_1^0.

---

[7] We comment that this minimization-enforcing relationship is also obtained in more general events in which the masses of the visible particles (or groups of particles) on each side of the event are unequal. It should however be noted that this condition only enforces the minimum contained within m_{T2}'s definition *in the case that the minimum at a point where $m_T^{(1)} = m_T^{(2)}$*. Cases where this does not occur are presented in section 3.5.

Using, in (8), the three-particle definition of m_T from (56) instead of the two-particle definition (6), one defines m_{T4}, the analogue of m_{T2} for the case of four missing particles. The constraint on the unobserved momenta will, of course, have to be modified to read

$$\mathbf{q}_T^{\nu(1)} + \mathbf{q}_T^{\chi(1)} + \mathbf{q}_T^{\nu(2)} + \mathbf{q}_T^{\chi(2)} = \not{\mathbf{p}}_T, \tag{57}$$

where the labels (1) and (2) indicate which chargino the particles were emitted from.

The conditions for the minimization required to calculate m_{T4} can be calculated just as for m_{T2}. The Euler–Lagrange (EL) equations involving

$$\frac{\partial \left(m_T^{(i)}\right)^2}{\partial \mathbf{q}_T^{\nu(i)}} \qquad \text{and} \qquad \frac{\partial \left(m_T^{(i)}\right)^2}{\partial \mathbf{q}_T^{\chi_1^0(i)}}, \tag{58}$$

show that the minimization will select the invisible particles' momenta such that $\mathbf{u}_T^{\chi_1^0(i)} = \mathbf{u}_T^{\nu(i)}$. The other EL equations reproduce (50) but with electrons replacing pions.

This means that when calculating m_{T4} one can replace the missing particles from each chargino decay with a pseudo-particle with mass equal to the sum of the masses of those invisible particles and proceed as for m_{T2}. In the case of leptonic chargino decay the mass of the neutrino can be safely neglected in comparison with that of the χ_1^0, and the constraint $\mathbf{u}_T^{\chi(i)} = \mathbf{u}_T^{\nu(i)}$ is equivalent to $\mathbf{q}_T^{\nu(i)} = (0, 0)$.

The distribution over events of m_{T4} will have fewer entries near the upper kinematic limit $\left(m_{T4} = m_{\chi_1^+}\right)$ because when more particles go undetected an event at that limit must satisfy a larger number of constraints. For fully leptonic chargino-pair decay, there are six constraints of the type $\Delta\eta = 0$, two $\mathbf{p}_T^{\nu(i)} = 0$ and finally the modified constraint from (50). This effect can be seen in figure 3 for events where a total of two, three and four invisible particles are produced.

3.5. Asymmetric decays

In the preceding two sections we have seen that when the decays on each side of the event are the same (i.e. both initial sparticles decay to the same set of daughter particles) then the resulting kinematic variables, m_{T2} and m_{T4}, have very similar properties. The only significant difference we have seen is the reduced density of events near the upper kinematic endpoint of m_{T4} relative to m_{T2}. Why, then, is the m_{T3} distribution, shown in figure 3, seen to have a shape significantly different from the m_{T2} and m_{T4} distributions? Specifically, why does it have the strong peak at low values not shared by the other two?

The difference occurs because the visible particles on each side of an m_{T3} event are different (on one side χ_1^0, e, ν and on the other to χ_1^0, π^+) and so the unconstrained minima of the values of m_T on each side of the event are not equal as they are in the case of m_{T2} and m_{T4}:

$$\min_{\not{\mathbf{q}}_T^{(1)}} \left(m_T^{(1)}\left(\mathbf{p}_T^{\pi}, \not{\mathbf{q}}_T^{(1)}\right)\right) = m_\pi + m_{\chi_1^0} \neq m_e + m_{\chi_1^0} = \min_{\not{\mathbf{q}}_T^{(2)}} \left(m_T^{(2)}\left(\mathbf{p}_T^{e}, \not{\mathbf{q}}_T^{(2)}\right)\right). \tag{59}$$

It is thus possible for some of the events can then fall into the category shown in figure 7(a), producing a peak of events with $m_{TX} = m_{\chi_1^0} + m_\pi$.

4. LHC case studies using m_{T2}

In this section we highlight some physics studies for the LHC which demonstrate that the background and the detector effects do not prevent m_{T2} from being a useful experimental variable. We investigate points from three different models, under two different classes of mass hierarchy.

4.1. Case 1—mSUGRA-like points

The first two points discussed are the mSUGRA point 5 (S5) and a point from the optimized string model (O1) discussed in [9]. The relevant parameters of these models are

$$m_{3/2} = 300\,\text{GeV}, \qquad m_0 = 100\,\text{GeV}, \qquad A_0 = 300\,\text{GeV}, \qquad \tan\beta = 2.1,$$
$$m_{3/2} = 250\,\text{GeV}, \qquad \tan\beta = 10, \qquad \theta = \pi/4,$$

respectively and $\mu > 0$ in both cases. For these points we are looking at dislepton production from a hard process which decays as $\tilde{l}^{\pm} \rightarrow \tilde{\chi}_1^0 l^{\pm}$, and so the mass hierarchy is $m_l \ll m_{\tilde{l}} - \chi \approx \chi < m_{\tilde{l}}$.

For S5 and O1, all events, except the $qq \rightarrow W^+W^-$ background processes, were simulated by HERWIG−6.0 [13]. The W-pair events were generated by ISAJET−7.42 [14]. The events for these two points were generated at 100 fb^{-1}. This is expected to correspond to running at high luminosity for one year.

Since there are two different processes being analysed, there are different cuts to apply. As this is not intended to introduce new physics, here we present only the major cuts used. For more details about the cuts and the techniques used, see [9, 10].

The events used for S5 and O1 are required to have one opposite sign same family (OSSF) pair of isolated leptons with $p_T^{l_1} > 50\,\text{GeV}$ and $p_T^{l_2} > 30\,\text{GeV}$. These events cannot contain any other isolated leptons. Also, events containing one or more jets with $p_T^j > 40\,\text{GeV}$ are vetoed. This helps reduce the standard model backgrounds.

The variable ΔM is defined as

$$(\Delta M)^2 \equiv \tfrac{1}{4}\big(M_{T2}^2(m_l)\big) - m_l^2. \tag{60}$$

This variable is what is studied for the points S5 and O1, for reasons given in [9]. The desired dislepton events have very little jet activity and the dislepton production cross sections are typically two orders of magnitude smaller than the squark/gluino production cross sections. There are also irreducible SM backgrounds (primarily $W^+W^- \rightarrow l^+l^-\nu\bar{\nu}$ and $t\bar{t} \rightarrow b\bar{b}W^+W^- \rightarrow jjl^+l^-\nu\bar{\nu}$ in cases where jets are below the reconstruction threshold or are outside detector acceptance) which have signatures identical to dislepton events. The smallness of the signal and the presence of these backgrounds would cause problems for naive straight-line fitting technique. Instead, the technique described in [9] is used for the estimation of the edge precision.

Figure 8 shows the ΔM distributions obtained at S5 and O1 after applying the cuts to a 100 fb^{-1} sample of signal dislepton events. Events from lighter sleptons ($\tilde{e}_R$ and $\tilde{\mu}_R$) occupy the unhatched region in each plot, while the events from heavier left-sleptons are cross-hatched. It will be noted that events from both light and heavy sleptons succeed in passing the cuts in both models. In principle, then, there are two edges to be observed in each of the models: one for the lighter slepton and one for the heavier. We note, however, that as the slepton masses increase, their production is strongly suppressed, and so there are very few heavy slepton events at O1 where there is in fact none within 10 GeV of the kinematic limit. It is readily observed that at the three remaining edges, where statistics are higher, there is good agreement between the theoretical prediction and the observed endpoint of each distribution.

Significant numbers of SM background events also pass the cuts. These can be well modelled by their opposite sign different family (OSDF) counterpart. As the signals from dislepton pair production are expected to be purely OSSF, we can use OSDF background subtraction. Supersymmetric backgrounds also have to be considered. Again, in this case the OSSF distributions are well modelled by OSDF events passing the same cuts. So again, different-family background subtraction is used.
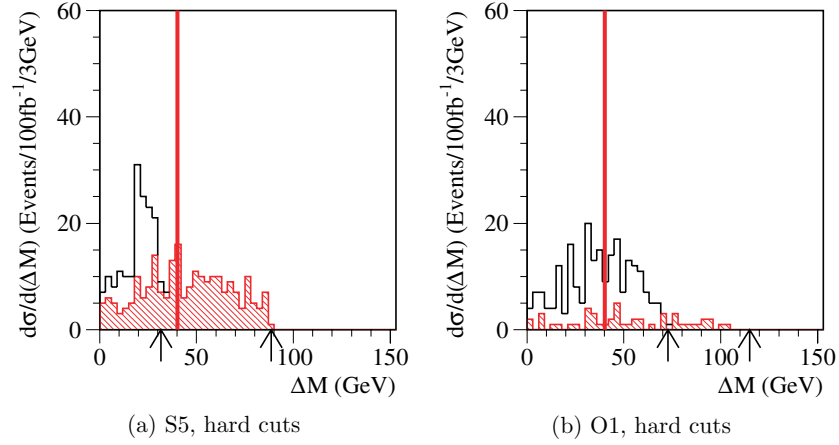
(a) S5, hard cuts (b) O1, hard cuts

Figure 8. ΔM distributions obtained at S5(a) and O1(b) after applying the cuts to a $100\,\mathrm{fb}^{-1}$ sample of $\tilde{e}_R\bar{\tilde{e}}_R$, $\tilde{\mu}_R\bar{\tilde{\mu}}_R$, $\tilde{e}_L\bar{\tilde{e}}_L$ and $\tilde{\mu}_L\bar{\tilde{\mu}}_L$ dislepton events. Events from light right-sleptons (unhatched) are stacked on top of those from heavier left-sleptons (hatched). Only events from OSSF lepton combinations are shown. The plots are generated without OSDF background subtraction, but were it to be performed, no significant differences would be apparent as only 4 (12) signal events are able to pass OSDF soft cuts at S5 (O1). Arrows indicate the values of ΔM^{max} predicted by theory for the two types of slepton in each model. A red vertical line is drawn through each plot at half the W-mass.
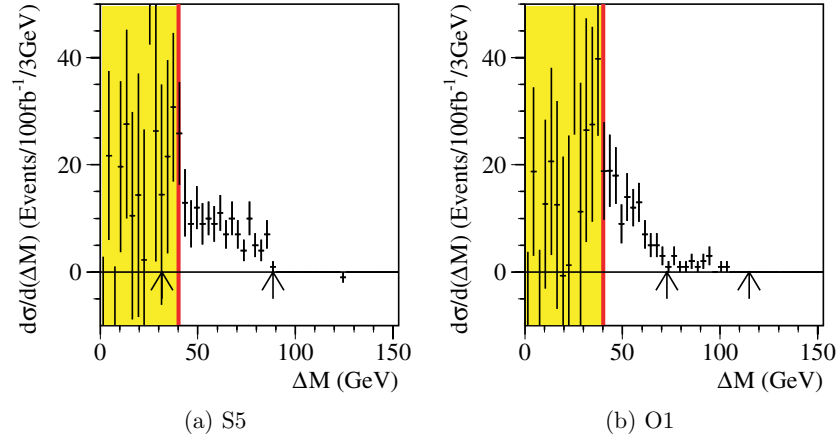

(a) S5 (b) O1

Figure 9. Flavour-subtracted ΔM distributions for combined signal and background at S5(a) and O1(b) after applying the cuts. A red vertical line is drawn through each plot at half the W-mass. Plots are shaded to the left of this line in order to draw attention to the events which reconstruct *above* this point. Compare these plots with those of figure 8 which contain only signal events. The arrows of figure 8 are provided for comparison.

All events for S5 and O1 (signals and backgrounds) are combined in figure 9 after different-family background subtraction. The reader is encouraged to compare these plots with those from figure 8 showing the desired signal shapes. As expected, all signal shape information is lost to the left of $m_W/2$ due to obliteration by the SM backgrounds. To the right of this point, at least one clear edge is observable in both models (the left-slepton edge at S5, and the right-slepton edge at O1) and in both sets of cuts. The hard cuts are able to suppress

Table 1. Parameters of the four AMSB-like points studied. The μ parameter was adjusted so that different values of $\Delta M_{\tilde{\chi}1}$ could be investigated.

Name	m_0 (GeV)	$m_{3/2}$ (TeV)	$\tan\beta$	μ
SPS-300	450	60	10	300
SPS-250	450	60	10	250
A-250	500	36	10	250
A-200	500	36	10	200

Table 2. The lightest chargino mass, the mass difference, $\Delta M_{\tilde{\chi}1} = m_{\chi_1^+} - m_{\chi_1^0}$, and two chargino branching ratios for the four AMSB-like points discussed in section 4.2. The hadronic branching ratios can be found in [10].

Point	$m_{\chi_1^+}$ (GeV)	$\Delta M_{\tilde{\chi}1}$ (MeV)	$\chi_1^+ \to \chi_1^0 e^+ \nu_e$	$\chi_1^+ \to \chi_1^0 \mu^+ \nu_\mu$
SPS-300	165	886	17.0%	15.9%
SPS-250	159	1798	21.9%	21.5%
A-250	101	766	15.4%	13.9%
A-200	97	1603	22.5%	22.2%

the supersymmetric backgrounds to such a degree that there is even compelling evidence at O1 for the existence of *two* edges, although the lack of statistics in the higher edge limits the precision with which the endpoint may be located.

4.2. Case 2—AMSB-like scenarios

The characteristic signature for anomaly-mediated supersymmetry breaking is the near mass-degeneracy of the lightest chargino and the lightest neutralino. The χ_1^+ therefore decays to a neutralino plus (relatively) light standard model particles. For a small mass difference, $\Delta M_{\tilde{\chi}1} = m_{\chi_1^+} - m_{\chi_1^0}$, the largest χ_1^+ branching ratios are to $\chi_1^0 \pi^+$ and to $\chi_1^0 l^+ \nu_l$, where $l \in e, \mu$. The mass hierarchy,

$$m_\pi \quad \text{or} \quad (m_l + m_\nu) \approx m_{\chi_1^+} - \chi \ll \chi \leqslant m_{\chi_1^+},$$

is therefore very different to the previous case study.

HERWIG−6.3 was used to generate $30\,\text{fb}^{-1}$ of unweighted inclusive supersymmetry events. HERWIG was also used to generate the background. For all the points, the results were passed through the ATLAS fast detector simulator, ATLFAST [15]. The signal-enhancing cuts require missing transverse energy, $\not\!\!E_T^{\min} = 500\,\text{GeV}$, leading jet transverse momentum, $p_{T(J_1)}^{\min} = 400\,\text{GeV}$ and transverse sphericity, $S_T^{\min} = 0.05$. There are also cuts on the tracks, these are described in more detail in [10].

We consider four AMSB-like points, which have the parameters given in table 1. Two of the points (SPS-300 and SPS-250) are based on the point suggested for study at Snowmass [16]. In all four cases the μ parameter has been adjusted at the electroweak scale in order to investigate different values of $\Delta M_{\tilde{\chi}1}$, as discussed in [10]. Some masses and branching ratios can be found in table 2.

Two of the points (SPS-300 and A-250) have large branching ratios for the decay $\tilde{\chi}_1^\pm \to \pi^\pm \tilde{\chi}_1^0$. This means that chargino-pair decay can easily generate the topology shown in section 1. We therefore plot distributions of $m_{T2} - m_{\tilde{\chi}_1^0}$, for which signal events in a perfect
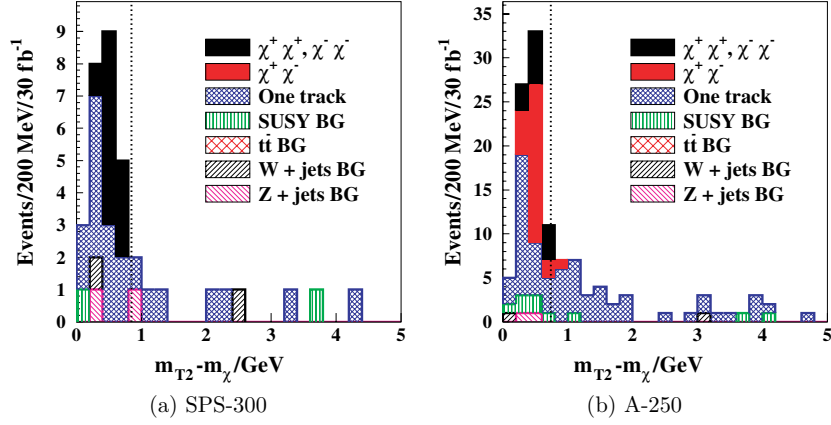
Figure 10. The $m_{T2} - m_{\tilde{\chi}_1^0}$ distribution for (a) the point SPS-300, and (b) the point A-250. The signal consists of the two solid regions labelled $\chi^x + \chi^x$ in the legend. The upper kinematic limit of $m_{T2} - m_{\chi_1^0}$ for signal events is marked with a dotted line. Note the sharp fall-off in the distribution near the kinematic edge at $m_{T2} - m_{\chi_1^0} = \Delta M_{\tilde{\chi}_1}$.
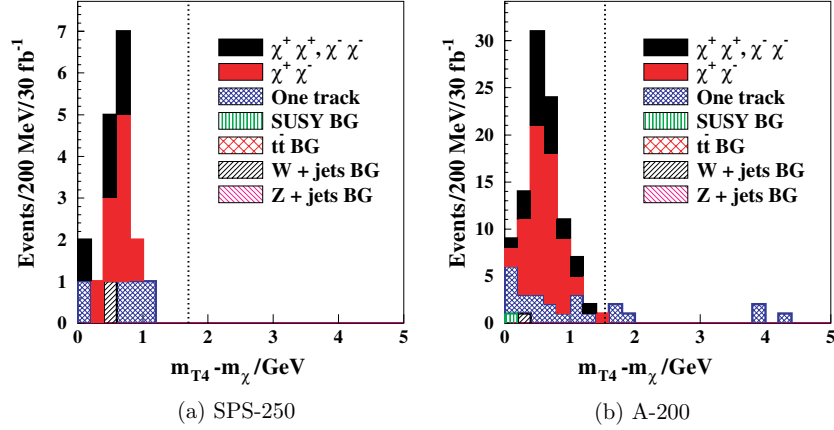

Figure 11. The $m_{T2} - m_{\tilde{\chi}_1^0}$ distribution for (a) the point SPS-250, and (b) the point A-200. The signal consists of the two solid regions labelled $\chi^x + \chi^x$ in the legend. The upper kinematic limit of $m_{T2} - m_{\chi_1^0}$ for signal events is marked with a dotted line. Note the fall-off in the distribution near the kinematic edge at $m_{T2} - m_{\chi_1^0} = \Delta M_{\tilde{\chi}_1}$.

detector would lie in the range $[m_\pi, \Delta M_{\tilde{\chi}_1}]$. The results (see figure figure 10) show that m_{T2} could be used to measure the small mass difference between the $\tilde{\chi}_1^+$ and the $\tilde{\chi}_1^0$ in this model, provided the signal cross section is sufficiently large.

The other two points each have a larger leptonic branching ratio, and so for these points the fully leptonic channel was investigated. Since there are now four missing particles in the final state (two neutralinos, and two neutrinos), distributions of $m_{T4} - m_{\chi_1^0}$ were plotted. For a perfect detector, these are restricted to lie in the range $[m_{e/\mu}, \Delta M_{\tilde{\chi}_1}]$.

The signal events are again indicated by the solid shades in the histograms in figure 11. Again, it can be observed that the distribution lies within the expected range. The distribution

is skewed to lower values because more particles are missing, and so more constraints must be satisfied for an event to approach the upper limit (as was seen in figure 3).

The sensitivity of m_{TX} to the estimated mass of the neutralino was shown in figure 4. It has been found that m_{TX} shows similar insensitivity to measurement uncertainties in the missing transverse momentum vector. This behaviour can be (at least partially) understood from the non-relativistic limit of m_{T2}, when the proportionality in (50) becomes an equality and

$$m_{T2}^2 - \left(m_\pi + m_{\chi_1^0}\right)^2 = \frac{1}{4m_\pi m_{\chi_1^0}}\left(m_\pi \not{p}_T - m_{\chi_1^0}\mathbf{p}_T^{\pi_1} - m_{\chi_1^0}\mathbf{p}_T^{\pi_2}\right)^2 + \mathcal{O}((\mathbf{v}_T \cdot \mathbf{v}_T)^2). \qquad (61)$$

The low sensitivity to the (possibly poorly measured) quantities $m_{\chi_1^0}$ and $\not{p}_T$ comes from the fact that in (61) they are multiplied by the quantities $\mathbf{p}_t^\pi$ and m_π respectively, which are both small in this mass regime.

5. Conclusion

This paper has attempted to achieve three objectives. Firstly, it seeks to introduce a new set of kinematic variables $\{m_{T2}, m_{T3}, \ldots\}$, which are specially designed to extract information from a particular class of troublesome events that we are likely to see at next generation hadron and lepton colliders. These events are those containing a pair of particles of identical (but unknown) mass which subsequently decay into groups of particles, each containing one or more invisible (possibly massive) particles. An example of this kind of event might be pair production of sleptons at the LHC, followed by subsequent sleptonic decay to leptons and neutralinos. Secondly, this paper attempts to get to the bottom of these new variables; it describes the regimes in which they can or cannot be trusted, develops useful approximations to them, and shows generally how one could go about calculating this variable for real. The approximations to the variables are not only useful in their own right, but are even more useful as guides which illustrate the dependence of the variables upon their inputs. Finally, this paper seeks to show with a couple of examples, real use of these variables in physics analyses. These hopefully show that m_{T2} and its chums are able to provide vital and new information about particle masses from events that would at first glance appear to contain so many unknown quantities as to be useless.

In conclusion, we believe that m_{T2} is an invaluable tool for physicists working at the LHC, and other future colliders, and we hope that this paper will encourage its use.

References

[1] Arnison G *et al* (UA1 Collaboration) 1983 Experimental observation of isolated large transverse energy electrons with associated missing energy at $\sqrt{s} = 540$ GeV *Phys. Lett.* B **122** 103–16
[2] Banner M *et al* (UA2 Collaboration) 1983 Observation of single isolated electrons of high transverse momentum in events with missing transverse energy at the CERN $p\bar{p}$ collider *Phys. Lett.* B **122** 476–85
[3] Affolder T *et al* (CDF Collaboration) 2001 Measurement of the *W* boson mass with the Collider Detector at Fermilab *Phys. Rev.* D **64** 052001 (*Preprint* hep-ex/0007044)
[4] Abazov V M *et al* (D0 Collaboration) 2001 Improved *D*0 *W* boson mass determination *Preprint* hep-ex/0106018 (FERMILAB-CONF-01-284-E)
[5] Barger V D, Martin A D and Phillips R J N 1983 Perpendicular *eν* mass from W decay *Z. Phys.* C **21** 99
[6] Barger V D, Martin A D and Phillips R J N 1983 Evidence for the *t* quark in $p\bar{p}$ collider data *Phys. Lett.* B **125** 339
[7] Berger E L, DiBitonto D, Jacob M and Stirling W J 1984 The minimum invariant mass: a technique for heavy quark searches at collider energy *Phys. Lett.* B **140** 259
[8] Lester C G and Summers D J 1999 Measuring masses of semi-invisibly decaying particle pairs produced at hadron colliders *Phys. Lett.* B **463** (*Preprint* hep-ph/9906349)

[9] Lester C G 2001 Model independant sparticle mass measurements at ATLAS *PhD Thesis* Cambridge University (*Preprint* CAV-HEP 02/13)

[10] Barr A J, Allanach B C, Lester C G, Parker M A and Richardson P 2002 Discovering anomaly-mediated supersymmetry at the LHC *J. High Energy Phys.* JHEP03(2002)45 (*Preprint* hep-ph/0208214)

[11] LHCC Supersymmetry Workshop (Oct. 1996)

[12] Polesello G, Poggioli L, Richter-Was E and Söderqvist J 1997 Precision SUSY measurements with ATLAS for SUGRA point 5, ATLAS Internal Note, (ATL-PHYS-97-111)

[13] Corcella G *et al* 1999 *HERWIG 6.1 release note* (*Preprint* hep-ph/9912396)

[14] Paige F E, Protopopescu S D, Baer H and Tata X 1998 ISAJET 7.40: a Monte-Carlo event generator for pp, $p\bar{p}$, and e^+e^- reactions (*Preprint* hep-ph/9810440)

[15] Richter-Was E, Froidevaux D and Poggioli L 1998 ATLFAST 2.0: a fast simulation package for ATLAS *Tech. Rep.* ATL-PHYS-98-131

[16] Allanach B C *et al* 2002 The snowmass points and slopes: benchmarks for SUSY searches *Preprint* hep-ph/0202233