The race for supersymmetry: using $m_{T2}$ for discovery

Alan J. Barr$^1$† and Claire Gwenlan$^1$‡

$^1$Denys Wilkinson Building, Keble Road, Oxford, OX1 3RH, United Kingdom

(Dated: July 15, 2009)

We describe how one may employ a very simple event selection, using only the kinematic variable $m_{T2}$, to search for new particles at the LHC. The method is useful when searching for evidence of models (such as $R$-parity conserving supersymmetry) which have a $Z_2$ parity and a weakly-interacting lightest parity-odd particle. We discuss the kinematic properties which make this variable an excellent discriminant against the great majority of Standard Model backgrounds. Monte Carlo simulations suggest that this approach could be used to discover supersymmetry with somewhat smaller integrated luminosities (or perhaps lower center-of-mass energies) than would be required for other comparable analyses.

I. INTRODUCTION

The search for new particles cannot be divorced from measurement of their masses. Indeed new particles have usually been discovered by measuring differential distributions in variables sensitive to their mass [1, 2]. A notable example was the successful $W$ boson discovery at the CERN Sp$\ddot{p}$S collider [1, 3] using the transverse mass, $m_T$. Using the right variables can improve the discrimination between signal and background, increase the statistical power and thus expedite discovery.

One of the primary goals of the Large Hadron Collider (LHC) is to search for physics beyond the Standard Model. Many well-motivated models of new physics [4] have a $Z_2$ parity (for example $R$-parity in supersymmetry) which predicts the presence of stable, presumed weakly-interacting lightest parity-odd particle. Such particles must be produced in pairs and would be invisible to the LHC detectors.

In this paper we explain why the $m_{T2}$ variable, which was originally proposed [4] for supersymmetric particle mass determination in hadron colliders, is also useful in searches for such models.

We start by recalling the definition of the transverse mass and its generalisation $m_{T2}$, and outline their defining properties (Section II). In Section III we show that experimentally-interesting kinematic configurations often lead to near-minimal values of $m_{T2}$. In Section IV we explain why these properties should cleanly separate new physics signals from the great majority of the Standard Model background events. A Monte Carlo simulation of a typical di-jet channel can be found in Section V. We discuss the results and conclude in Section VI.

II. DEFINITIONS

The familiar transverse mass, which was used for the $W$ boson discovery, is defined by

$$m^2_T \equiv m_v^2 + m_i^2 + 2(e_v e_i - v_T \cdot q_T), \quad (1)$$

where $m_v$ and $m_i$ are the masses of the visible and invisible systems respectively (e.g. the electron and neutrino for the $W$ case), $v_T$ and $q_T$ are their momenta in the transverse plane, and the transverse energies are defined by

$$e_v = m_v^2 + v_T^2 \quad \text{and} \quad e_i = m_i^2 + q_T^2.$$ 

The transverse momentum $q_T$ of the unobserved neutrino can be inferred from momentum conservation in the directions perpendicular to the beam: it is assumed to be equal to the missing transverse momentum, $p_T$ which is defined to be the the negative sum of the transverse momenta of all the particles observed in the detector (i.e. including hadrons in the $W$ case).

Why was $m_T$ useful for the $W$ search? The defining property of $m_T$ is as follows: for events in which a visible and an invisible system originate from the decay of a single parent particle of mass $m_0$, and when the correct daughter masses are used, $m_T$ satisfies

$$m_T \leq m_0 \quad (2)$$

by construction, with equality when the rapidity of all invisible daughters is equal to that of the visible system. The transverse mass gives the lowest event-by-event upper bound on $m_0$. The maximum value of $m_T$ over all events therefore delineates the boundary between allowed and forbidden values of the parent mass, $m_0$.

For the $W$ boson search the allowed values of $m_T$ could therefore span a range $m_\prec \lesssim m_T \lesssim m_W$ for signal events. The lower bound is given by $m_\prec = m_v + m_i$, which is zero if one assumes (as was done implicitly in

---

† Electronic address: a.barr@physics.ox.ac.uk
‡ Electronic address: c.gwenlan@physics.ox.ac.uk

1 Some of these observations were originally made in a talk at CERN by the authors in May 2007. Since then, the method has been employed, for example in [4], but the underlying properties which represented the motivation for the approach have not previously been published.

2 If the width of the particle and the experimental resolutions are assumed to be small.
the original $W$ search papers) that both $m_p$ and $m_l$ are zero. Other processes which could have led to lepton + \( p_T \) final states, for example leptonic decays of $\tau$ leptons or $B$ hadrons were also constrained by similar inequalities, but these backgrounds have smaller values of $m_0$. There was therefore a region bounded approximately by $m_B \lessapprox m_T \lessapprox m_W$ in which there were few background events and a preponderance of the signal.\footnote{In \cite{1} the lower end of this distribution was removed because of threshold cuts on the electron $p_T$ and $p$.}

A corresponding construction can be made for the types of LHC search described in the introduction. We are now interested in the case in which two new particles are produced, each of which decays to a set of (one or more) visible and (one or more) invisible daughters. We label the two branches of the decay with the superscripts (1) and (2) to distinguish them, and for the purposes of this paper assume that the visible particles can be unambiguously assigned to one or other parent, though this need not be done in general \cite{2}. According to our assumptions, each invisible system should include one dark matter candidate.

As was the case for single-particle decays, there will be some values of parent and (visible and invisible) daughter masses which are consistent with the observed momenta, and others which are not.

The $p_T$ constraint needs some modification compared to the single decay case. For a pair of invisible systems the transverse momentum of each the two invisible-particle systems may represent an individual lepton, or a jet or a di-jet pair. For example $v \chi$\cite{12}.\footnote{Alternative combinatorial procedures are available for the visible system \cite{13}.} Similarly each of the invisible-particle systems may represent a single massive invisible particle (e.g. $\nu$). For composite systems, the composite 2+1 vectors can be formed from the vector sum of the 2+1 vectors of their individual constituents.\footnote{If and when signals suggesting new invisible particles are observed, their masses can later be determined, for example by constructing variables sensitive to the kink in the maximum of $m_{T2}$ as a function of the input value of $m_i$ \cite{3,10}.}

### III. RELEVANT PROPERTIES

Later in Section \cite{IV} and \cite{V} we shall find that many background processes predict values of $m_{T2}$ near the global minimum, $m_\chi$. The value of this minimum is therefore of some interest. It follows from (\ref{eq:3}) that

$$m_\chi = \max(m_{\chi}^{(1)}, m_{\chi}^{(2)})$$

where $m_{\chi}^{(n)} = m_{i}^{(n)} + m_{i}^{(n)}$ is the global minimum of each of the two individual $m_{T2}^{(n)}$.

In order to construct $m_{T2}$ we require hypotheses for the invisible particle(s) masses $m_i^{(n)}$. In a search we usually don’t know a priori whether any invisible particles are massive (or if so, what those masses are) so we choose to set the invisible-particle mass hypotheses $m_i$ to zero. This is the only value guaranteed to preserve (\ref{eq:2}) if the invisible particles’ true masses are unknown.\footnote{\cite{5}} It is also the appropriate choice for background processes where the invisible particles are ($\sim$ massless) neutrinos.

In the rest of this section we show that the observable $m_{T2}(v_1, v_2, p_T, 0, 0)$ must adopt small values for a variety of kinematic configurations. In Section \cite{V} we use these results to show that such configurations are satisfied (at least approximately) by very many of the Standard Model processes which represent backgrounds to example searches.

**Lemma 1** When a pair of particles are produced, both with mass $m_0$, and each parent decays to a visible system $v$ and an invisible system $i$, then $m_{T2}(v_1, v_2, p_T, 0, 0) \leq m_0$.

**Proof** Consider the chained pair of inequalities

$$m_{T2}(v_1, v_2, p_T, 0, 0) \leq m_{T2}(v_1, v_2, p_T, m_i^{(1)}, m_i^{(2)}) \leq m_0.$$  

The first inequality follows since each $m_i^{(n)}$ is a monotonic function of the (non-negative) parameter $m_i^{(n)}$. The second inequality is satisfied by $m_{T2}$ by construction: one of the partitions of the missing momentum is the correct one, and for that partition each $m_T \leq m_0$ by \cite{2}.

**Lemma 2** When two particles are produced with different masses $m_1$ and $m_2$ and each parent decays to a visible system $v$ and an invisible system $i$ then $m_{T2}(v_1, v_2, p_T, 0, 0) \leq \max(m_1, m_2)$.
Proof As for Lemma 1, except that at the correct \( \mathbf{p}_T \) partition we can only be sure that \( m_T^{(1)} \leq m_1 \) and \( m_T^{(2)} \leq m_2 \) and so \( \max(m_T^{(1)}, m_T^{(2)}) \leq \max(m_1, m_2) \) for that partition.

Lemma 3 When \( \mathbf{v}_T^{(n)} = \mathbf{0} \) and \( m_v^{(n)} = m_i^{(n)} = 0 \) then \( m_T^2 = m_\perp \) for \( n \in \{1,2\} \).

Proof Without loss in generality let \( n = 2 \) (and so \( \pi = 1 \)). Now \( m_T^{(2)} = m_\perp \) \( \forall \mathbf{q}_T^{(2)} \). There exists a partition of \( \mathbf{p}_T \), with \( \mathbf{q}_T^{(1)} = \mathbf{0} \) for which \( m_T^{(1)} = m_\perp \). The result follows from (3) and (4).

An important example of Lemma 3 is the decay of a single parent of mass \( m_0 \). The second (non-existent) visible system is represented by \( \mathbf{v}_2 = (0,0) \). With \( m_i^{(1)} = 0 \), Lemma 3 shows that \( m_T^2 = m_\perp = \max(m_i^{(2)}, m_v^{(1)}) \) where the second equality is a result of (4).

Lemma 4 When \( \mathbf{p}_T = \mathbf{0} \) and \( m_i^{(1,2)} = 0 \) then \( m_T^2 = m_\perp \).

Proof For \( \mathbf{p}_T = \mathbf{0} \) there exists a trivial partition of the missing momentum with \( \mathbf{q}_T^{(1)} = \mathbf{p}_T = \mathbf{0} \). For that partition, \( m_T^{(1)} = m_\perp \) and \( m_T^{(2)} = m_\perp \); the result follows from (3) and (4).

Lemma 5 When \( m_i^{(1,2)} = 0 \), \( m_v^{(n)} = 0 \) and \( \mathbf{p}_T \parallel \mathbf{v}_T^{(n)} \) then \( m_T^2 = m_\perp \) for \( n \in \{1,2\} \).

Proof Without loss of generality let \( n = 1 \). There exists a partition of \( \mathbf{p}_T \), with \( \mathbf{q}_T^{(1)} = \mathbf{p}_T \) and \( \mathbf{q}_T^{(2)} = \mathbf{0} \). Each \( m_T \) is equal to its global minimum by (1), and the result follows from (4).

Lemma 6 When \( m_i^{(1,2)} = m_v^{(1,2)} = 0 \), and \( \mathbf{p}_T \) can be expressed as the sum \( \mathbf{p}_T = \sum x_k \mathbf{q}_k^1 \) for some real non-negative \( x_k \) then \( m_T^2 = m_\perp = 0 \).

Proof For the partition of \( \mathbf{p}_T \), given by these \( x_k \), \( \mathbf{q}_k^{(n)} \parallel \mathbf{v}_T^{(n)} \) simultaneously for both \( n \in \{1,2\} \). For that partition each \( m_T^{(n)} = m_\perp = 0 \). The result follows from (3) and (4).

Lemma 7 If either or both visible systems are composite, then each of the results in Lemmas 4–6 also hold when either or both the composite Lorentz 2+1 vectors \( \mathbf{v} \) are replaced by any subset of their (respective) constituents.

Proof Lemmas 3–6 follow exactly as before. For Lemmas 1 and 2 it is sufficient to show that \( m_T^2 \) cannot be increased when a constituent is removed from a composite (visible) system.

Now \( m_T^2 \) is the inner product of a sum of Lorentz 2+1 vectors

\[
 m_T^2 = (v + i)^2 = \left( \sum c_k \right)^2,
\]

where the sum runs over all the constituent \( c_k \), whether visible or invisible. Separating one of the visible particles, \( c_j \) and letting \( \sigma = \sum_{k \neq j} c_k \),

\[
 m_T^2 = \sigma^2 + v_j^2 + 2\sigma \cdot v_j.
\]

Since each of the constituent vectors is time-like (or null) each of the terms is non-negative and \( \sigma^2 \leq m_T^2 \). But \( \sqrt{\sigma^2} \) is precisely the transverse mass one obtains if \( v_j \) is omitted, so the transverse mass cannot be increased by omitting a particle. This result also holds at the partition chosen by the minimisation of (3) so \( m_T^2 \) is never increased by omitting one (or by induction more than one) of the visible particles.

Physically relevant configurations will rarely (if ever) conform to the precise configurations described in Lemmas 1–7; parent particles will be off-mass-shell, Standard Model particles will have small but non-zero masses, \( \mathbf{p}_T \) will never be exactly zero nor vectors exactly parallel. However, since \( m_T \) is a continuous function of its inputs, kinematic configurations close to those described above will have upper bounds close to these idealised cases.

IV. EXAMPLE

The properties described in Section III turn out to be particularly useful for various LHC searches. We demonstrate the reasons why by examining the concrete example of pair-production of heavy, strongly-interacting particles of mass \( m_0 \), each of which decays to a light-quark jet and an invisible particle. The characteristic final state is thus two (usually high-\( p_T \)) jets with significant \( \mathbf{p}_T \). Examples of models which could lead to such a final state can be found in Table I.

| Supersymmetry | \( q\bar{q} \rightarrow q\bar{q}^0 + q\bar{q}^0 \) |
|--------------|----------------------------------|
| UED          | \( q\bar{q} \rightarrow q\gamma + q\gamma \) |
| Leptoquarks  | \( LQ \rightarrow q\bar{q}^0 \) |

TABLE I: Examples of models of new physics producing dijets in association with \( \mathbf{p}_T \).

When considering how to search in such channels, a variety of levels of sophistication can be envisioned.

In a typical cut-based search one would select events with large \( \mathbf{p}_T \) and two high-\( p_T \) jets; require sufficiently large azimuthal angles \( \Delta \phi \) between each jet and the \( \mathbf{p}_T \) vector to reduce backgrounds from “fake” (from the mis-measurement of one of the jet energies) or “real” (from neutrinos in jets) \( \mathbf{p} \) from the jet; and apply further cuts to reduce background processes (such as leptonic \( t\bar{t} \) decays) which produce neutrinos and high-\( p_T \) jets with larger \( \Delta \phi \) [6]. The remaining events should be signal-enriched, and so, provided that the residual background contributions can be understood, one can check for an excess of events not explained by the Standard Model. The
most difficult task is to understand the contribution of the background to the selected events – for which a various techniques have been proposed. Applying such a set of many single-variable cuts is fairly straightforward but is somewhat wasteful of signal events (which one can assume will be in scarce supply at the time of any interesting discovery).

At the other extreme, the formally optimal search method (needing the fewest events) would require calculation of the likelihood for all events for every signal and background hypothesis. This is often prohibitively difficult in practice even if all such hypotheses can be identified and modelled.

A pragmatic approach, intermediate in complexity, is to find a single, easily-calculable observable which, while not optimal in the formal statistical sense, still gives very good discrimination between the majority of signal-like and background-like events, based on some rather general principles such as relativistic kinematics. In what follows we shall find that \( mT_2 \) is just such an observable.

For the example search channels, we identify the two visible systems \( v^{(m)} \) with the two highest-\( p_T \) jets. Our jets, though massive, should have a sufficiently small mass that \( m_j \approx 0 \) is a good approximation. For the reasons discussed in Section III we set both \( m_i^{(n)} \) to zero. For these choices, \( m< \) is small (equal to the larger of the \( m_{j}^{(n)} \)) and the mass conditions required for Lemmas 3–7 are satisfied in the \( m_j \to 0 \) limit.

| Process | \( m_{T_2}(v_1, v_2, p_T, 0, 0) \) | Comments |
|---------|----------------------------------|----------|
| QCD di-jet \( \to \) hadrons | \( = \max m_j \) by Lemmas 1–11 | fully hadronic decays |
| QCD multi jets \( \to \) hadrons | \( = \max m_j \) by Lemma 3 | any leptonic decays |
| \( t\bar{t} \) production | \( = \max m_j \) by Lemma 4 | fully hadronic decays |
| Single top / \( tW \) | \( \leq m_t \) by Lemmas 1–7 | any leptonic decays |
| Multi jets: “fake” \( \not{p}_T \) | \( = \max m_j \) by Lemma 4 | single mismeasured jet \( ^a \) |
| Multi jets: “real” \( \not{p}_T \) | \( = \max m_j \) by Lemma 6 | two mismeasured jets \( ^a \) |
| \( Z \to \nu\bar{\nu} \) | \( = 0 \) by Lemma 3 | single jet with leptonic \( b \) decay \( ^a \) |
| \( Z j \to \nu\bar{\nu} j \) | \( = m_j \) by Lemma 3 | two jets with leptonic \( b \) decays \( ^a \) |
| \( W \to \ell\nu \) \( ^b \) | \( = m_\ell \) by Lemma 3 | only ISR jet \( ^a \) |
| \( W j \to \ell\nu j \) \( ^b \) | \( \leq m_W \) by Lemma 2 | only ISR jet \( ^a \) |
| \( WW \to \ell\nu\ell\nu \) \( ^b \) | \( \leq m_W \) by Lemma 1 | also \( m_j \) for one ISR jet \( ^a \) |
| \( ZZ \to \nu\bar{\nu}\nu\bar{\nu} \) | \( \leq m_{LQ} \) | \( ^a \) assuming that the relevant jet(s) are identified with one of the two visible particle systems \( v^{(m)} \). |
| \( q\bar{q} \to q\bar{q} \chi_i \) | \( \leq m_q \) | \( ^a \) even if the lepton is mistaken for a jet. |
| \( q_1\bar{q}_1 \to q_1\bar{q}_1 \chi_i \) | \( \leq m_{q_1} \) | i.e. can take large values |

\(^a\) Assuming that the relevant jet(s) are identified with one of the two visible particle systems \( v^{(m)} \).
\(^b\) Even if the lepton is mistaken for a jet.

Table II lists a variety of Standard Model processes which form backgrounds to the search channel. From Lemmas 1–7 we expect to find restrictive upper bounds on \( mT_2 \) for many of these. Processes with small \( p_T \); with small jet \( p_T \); with small \( \Delta\phi_j \); or from production of one or two Standard Model particles; all bound \( mT_2 \) from above. Such processes constitute the majority of both the “physics” and “detector” backgrounds to this channel so one can remove the vast majority of the background simply by requiring large \( mT_2 \). No additional cuts are required – other than perhaps modest trigger requirements on jet \( p_T \) or on \( p_T \).

An example “physics” background is \( t\bar{t} \) production. This is pair-production of equal-mass particles, so \( mT_2 \) is bounded above by the mass of the top quark \( m_t \) by Lemma 1. One could argue that this is not a very strict bound, but it is much smaller than masses of the particles typically being searched for, and the top is the heaviest known Standard Model particle, so other similar background processes, e.g. hadronic \( \tau^+\tau^- \) decays should adopt smaller values of \( mT_2 \) again.

An example of a detector-induced background is “fake” \( p_T \) – events with no invisible particles, but where a substantial fraction of the energy of one of the leading jets is lost. This can happen when energy is deposited in inactive material the detector (e.g. cables, services or support structures). Such pathological events cannot necessarily be cleanly identified. In these cases the detector usu-
ally measures a jet with Lorentz 2+1 vector \( j' = \alpha j \) (\( 0 < \alpha < 1 \)), and one gains a contribution to \( p_T \) of \( (\alpha - 1)j_T \). In the absence of any other source of missing momentum, \( p_T \parallel j_T \), so \( m_{T2} \rightarrow m_\ell < \) by Lemma 4 or 6. Similar arguments apply to heavy-quark jets where leptonic decays lead to production of neutrinos close to the jet axis.

The other backgrounds in Table II are also forced to small values of \( m_{T2} \). The least restrictive is \( \leq m_\ell \) since the top is (assumed here to be) the heaviest Standard Model particle.

What values of \( m_{T2} \) are expected for any new particles? At the upper end it is clear from Lemma 1 that \( m_{T2} \leq m_0 \) for the processes in Table II. One needs a significant number of events with \( m_{T2} > m_\ell \) to have a signal region which is relatively free of background. Now if the correct value of \( m_\ell \) were to be used then the upper bound (\( m_0 \)) would be saturated since there is a significant density of states with \( m_{T2} \rightarrow m_\ell \) close to \( m_0 \). We have chosen to input the lowest conceivable value, and if the argument does not prove that the bound is saturated. However one can see from Table I that, provided \( m_\ell \ll |q_T| \), then events which are close to maximal when the true \( m_\ell \) is used will also remain close when we replace this by \( m_\ell = 0 \).

We therefore expect to find a large number of signal events, and very little background, in the region approximately bounded by \( m_\ell \leq m_{T2} \leq m_0 \), where \( m_0 \) is the mass of the new particle.

V. SIMULATION

To illustrate the results of Section III and Section IV, we generate Monte Carlo signal and background samples with Herwig++ 2.3.2 [13]. The background processes simulated are QCD, \( tt \), \( W \rightarrow \ell \nu+jets \), \( Z \rightarrow \ell^+\ell^-+jets \), and \( Z \rightarrow \ell\bar{\nu}+jets \). The contribution from diboson+jets is expected to be very small [8] so is not considered here. In order to generate sufficient QCD events in the high-\( p_T \) region, eight samples were generated in slices of the \( p_T \) of the hard scatter. For the SUSY signal, the SPS1a point [16] is used (\( m_0 = 100 \text{ GeV}, m_{1/2} = 250 \text{ GeV}, A_0 = -100 \text{ GeV}, \tan \beta = 10, \mu > 0 \)), as calculated by SPheno 2.2.3 [17]. Table III lists the leading order cross sections calculated by Herwig++, and the number of events generated for each of the processes considered.

We cluster hadrons (and \( \pi^0 \)'s) with fiducial pseudorapidity (\( |\eta| < 5 \)) and momentum \( (p_T > 0.5 \text{ GeV}) \) into jets using the fastjet implementation of the anti-\( k_T \) algorithm [18]. We use the \( E \) combination scheme and set \( R = 0.4 \) and \( p_T^{min} = 10 \text{ GeV} \). To simulate the detector effects we smear the majority (\( 1 - \epsilon \)) of the jet energies by a Gaussian probability density function of width

\[
\sigma(E)/E_j = \left( 0.6/\sqrt{E_j[\text{GeV}]} \right) \oplus 0.03
\]

where \( E_j \) is the unsmeared jet energy. This resolution is typical of one of the general-purpose LHC detectors [14]. Since the tails of the \( p_T \) distribution are often dominated by badly mismeasured jets, we simulate pathological energy-loss by applying a different smearing function to the remaining fraction (\( \epsilon = 0.1\% \)) of the jets with probability density:

\[
P(E) = \begin{cases} 2E/E_j^2 & \text{for } (0 < E < E_j) \\ 0 & \text{elsewhere} \end{cases}
\]

The missing transverse momentum is calculated from the negative vector sum of the visible fiducial hadrons (including \( \pi^0 \)) and is corrected for the jet smearing. We impose the simple requirement that each event contains at least two jets with \( p_T > 50 \text{ GeV} \). We then take the two highest \( p_T \) jets as \( j_{1,2} \) and calculate \( m_{T2}(j_1, j_2, p_T, 0, 0) \), for all events. We normalise to 100 pb\(^{-1} \) (using the leading order cross sections for both signal and background). The resulting distribution can be seen in Figure 1.

6 [8] suggests a larger value of \( \epsilon \sim 1\% \). We find a smaller value better matches the tails of the \( p_T \) and \( m_{T2} \) distributions found in full simulation. The detailed form of the transfer function clearly needs to be determined from the collision data, but our findings are not materially altered by changing epsilon from 0.1% to 1%.
As predicted, the region $m_t \lesssim m_{T_2} \lesssim m_{\tilde{q}}$ is dominated by signal events. The great majority of the Standard Model background events are found at small $m_{T_2}$, as required by the arguments of Section IV. The remaining backgrounds for which $m_{T_2}$ is not bounded above by the arguments in Section IV are: $Z \rightarrow \nu\bar{\nu}$ in association with two hard jets from initial state radiation (ISR); $W \rightarrow$ leptons plus two hard ISR jets; and three- (or more-) jet production where one of the jets looses a very large amount of energy, and that jet is not one of the two highest $p_T$ jets input to $m_{T_2}$. For the signal point examined, and with the level of sophistication used for our simulations, we find that the residual backgrounds for $m_{T_2} \gtrsim m_t$ are well below the supersymmetric signal predictions.

VI. DISCUSSION

The numerical simulations of Section V confirm the analytic results of Section III; $m_{T_2}$ adopts small values – equal to or less than the mass of Standard Model particles (e.g. $Z_{jj}$, $W_{jj}$, $t\bar{t}j$, $jjj$). The bounds of Section III do not apply to these since the variable $m_{T_2}$ was developed for the explicit two-parent case. To reduce the residual contribution from these $n > 2$-particle topologies one might consider using, rather than $m_{T_2}$, an $n$-parent generalisation of the transverse mass:

$$m_{T_n} \equiv \min_{\sum q_{T_i} = \mathbf{p}_T} \{\max m_T^{(i)}\} \text{ with } i \in \{1, \ldots, n\}. \tag{5}$$

Using (5) $n$-body background topologies (and also $m < n$-body topologies by a generalisation of Lemma 2) would be bounded from above by the mass of the heaviest parent. The problem would be that one would no longer expect pair-produced signal topologies (such as those in Table I) to obtain $m_{T_n}$ values close to the mass of the new heavy parent. Indeed a generalisation of Lemma 3 shows that for the simplest $Z_2$ signal case, di-jet + $p_T$, $m_{T_3}$ is forced towards the small value (max $m_j$) so one loses the discrimination power of $m_{T_2}$. Such $n > 2$ generalisations are therefore only likely to be appropriate when $n$ heavy signal particles are expected to be produced.

Our introduction focused on the simplest decay topologies of pair-produced heavy particles such as $\tilde{q}\tilde{q} \rightarrow q\chi^0_i\bar{q}\chi^0_1$ but our simulations also show excellent discrimination in more complicated cases using the same $m_{T_2}$ variable. Decay sequences with many steps, or indeed individual multi-daughter decays, such as $\tilde{g}\tilde{g} \rightarrow q\bar{q}\chi^0_i\tilde{g}\chi^0_1$ (via three-body decays), could have been input to $m_{T_2}$

| Sample | LO cross section (pb) | Number of events |
|--------|-----------------------|------------------|
| QCD ($17 < p_T < 35$ GeV) | $9.44 \times 10^4$ | $6 \times 10^7$ |
| QCD ($35 < p_T < 70$ GeV) | $5.99 \times 10^7$ | $5 \times 10^7$ |
| QCD ($70 < p_T < 140$ GeV) | $3.45 \times 10^6$ | $5 \times 10^7$ |
| QCD ($140 < p_T < 280$ GeV) | $1.57 \times 10^5$ | $5 \times 10^7$ |
| QCD ($280 < p_T < 560$ GeV) | 5280 | $4 \times 10^7$ |
| QCD ($560 < p_T < 1120$ GeV) | 116 | $3 \times 10^7$ |
| QCD ($1120 < p_T < 2240$ GeV) | 1.11 | $2 \times 10^7$ |
| QCD ($p_T > 2240$ GeV) | $1.15 \times 10^{-3}$ | $5 \times 10^7$ |
| $t\bar{t}$ | 231 | $2 \times 10^7$ |
| $W \rightarrow \ell\nu +$jets | 17,100 | $1.5 \times 10^8$ |
| $Z \rightarrow \ell^+\ell^- +$jets | 1780 | $9 \times 10^7$ |
| $Z \rightarrow \nu\bar{\nu} +$jets | 3440 | $5 \times 10^7$ |
| SUSY signal (SPS1a) | 13.5 | $2 \times 10^6$ |

TABLE III: Signal and background processes together with their leading order cross-sections, and number of events generated for each.
in a number of ways. There is no unique choice for constructing the two ‘visible particle systems’ from the \((k > 2)\) decay products. We could have chosen to form two composite systems (two di-jet systems for the gluino example), and use these as the visible inputs to \(m_{T2}\). This construction would have provided a large number of signal events close to the kinematic boundary and so would be appropriate for mass determination. However forming di-jet composite objects produces visible systems which no longer have \(m_v \approx 0\), so one would lose the desirable properties of Lemmas 3, 5 - 7. For our search, even though cascade or multi-body decays are expected, we have still chosen to form \(m_{T2}(j_1, j_2, p_T, 0, 0)\) using only the two highest \(p_T\) jets — precisely as for the simpler di-jet topology. This way we retain the desirable properties (Lemmas 3, 5 - 7) for the backgrounds. The \(m_{T2}\) distribution for the signal can still extend up to large values (close to \(m_\eta\) for the three-body decay), albeit with fewer near-maximal events than would be found by using composite two-jet visible systems. By inputting only the highest two \(p_T\) jets to \(m_{T2}\), not only do we take advantage of the background rejection properties, but we also combine many signal channels together, forming a larger sample of signal events (which therefore enjoys a larger statistical significance).

Of course even having found a good discriminating variable one is not absolved from the need to understand the residual background contributions. Previous studies [6, 14] show that the rates of a wide variety of Standard Model processes can be measured in control regions using observables which are used by the LHC experiments. Our analysis and simulations suggest that \(m_{T2}\), the natural kinematic observable for pair-produced particles, ought to be at the front of the queue.

We are grateful to C.G. Lester and to B. Gripaios for encouraging the wider dissemination of these results, and to Mihoko Nojiri and the IPMU, Tokyo, Japan, for hosting the workshop which prompted those discussions. Further thanks to C.G. Lester, B. Gripaios and to A. Pinder, and K. Matchev for helpful comments and discussions. AJB and CG are supported by Advanced Fellowships from the UK Science and Technology Facilities Council.

---

[1] G. Arnison et al. (UA1), Phys. Lett. B122, 103 (1983).
[2] J. J. Aubert et al. (E598), Phys. Rev. Lett. 33, 1404 (1974); G. Arnison et al. (UA1), Phys. Lett. B126, 398 (1983); S. Abachi et al. (D0), Phys. Rev. Lett. 74, 2632 (1995), hep-ex/9503003; F. Abe et al. (CDF), Phys. Rev. Lett. 74, 2626 (1995), hep-ex/9503002.
[3] G. Arnison et al. (UA1), Phys. Lett. B129, 273 (1983); M. Banner et al. (UA2), Phys. Lett. B122, 476 (1983).
[4] S. Dimopoulos and H. Georgi, Nucl. Phys. B193, 150 (1981); H.-C. Cheng, K. T. Matchev, and M. Schmaltz, Phys. Rev. D66, 056006 (2002), hep-ph/0205314; H.-C. Cheng and I. Low, JHEP 09, 051 (2003), hep-ph/0308199.
[5] C. G. Lester and D. J. Summers, Phys. Lett. B463, 99 (1999), hep-ph/9906349.
[6] G. Aad et al. (The ATLAS) (2009), 0901.0512.
[7] C. Lester and A. Barr, JHEP 12, 102 (2007), 0708.1028; M. M. Nojiri, Y. Shimizu, S. Okada, and K. Kawagoe, JHEP 06, 035 (2008), 0802.2412.
[8] H.-C. Cheng and Z. Han, JHEP 12, 063 (2008), 0810.5178.
[9] W. S. Cho, K. Choi, Y. G. Kim, and C. B. Park, Phys. Rev. Lett. 100, 171801 (2008), 0709.0288; A. J. Barr, B. Gripaios, and C. G. Lester, JHEP 02, 014 (2008), 0711.4008.
[10] B. Gripaios, JHEP 02, 053 (2008), 0709.2740; W. S. Cho, K. Choi, Y. G. Kim, and C. B. Park, JHEP 02, 035 (2008), 0711.4526.
[11] B. C. Allanach, C. G. Lester, M. A. Parker, and B. R. Webber, JHEP 09, 004 (2000), hep-ph/0007009; A. J. Barr, C. G. Lester, M. A. Parker, B. C. Allanach, P. Richardson, JHEP 03, 045 (2003), hep-ph/0208214; A. Barr, C. Lester, and P. Stephens, J. Phys. G29, 2343 (2003), hep-ph/0304226; G. Weiglein et al. (LHC/LC Study Group), Phys. Rept. 426, 47 (2006), hep-ph/0410364; P. Meade and M. Reece, Phys. Rev. D74, 015010 (2006), hep-ph/0601124; M. Baumgart, T. Hartman, C. Kilic, and L.-T. Wang, JHEP 11, 084 (2007), hep-ph/0608172; B. Fuks, AIP Conf. Proc. 903, 165 (2007), hep-ph/0610316; W. S. Cho, Y. G. Kim, K. Y. Lee, C. B. Park, and Y. Shimizu, JHEP 04, 054 (2007), hep-ph/0703163; J. P. Conlon, C. H. Kom, K. Suruliz, B. C. Allanach, and F. Quevedo, JHEP 08, 061 (2007), 0704.3403; W. S. Cho and K. Choi, AIP Conf. Proc. 939, 219 (2007), 0706.2871; H.-C. Cheng, J. F. Gunion, Z. Han, G. Marandella, and B. McElrath, JHEP 12, 076 (2007), 0707.0630; B. Fuks (2007), 0710.2002; G. G. Ross and M. Serna, Phys. Lett. B665, 212 (2008), 0712.0943; M. M. Nojiri, G. Polesello, and D. R. Tovey, JHEP 05, 014 (2008), 0712.2718; Y. No-
mura, M. Papucci, and D. Stolarski, JHEP 07, 055 (2008), 0802.2582; D. R. Tovey, JHEP 04, 034 (2008), 0802.2879; M. M. Nojiri et al. (2008), 0802.3672; T. Han, R. Mahbubani, D. G. E. Walker, and L.-T. Wang, JHEP 05, 117 (2009), 0803.3820; W. S. Cho, K. Choi, Y. G. Kim, and C. B. Park, Phys. Rev. D78, 034019 (2008), 0804.2185; M. Serna, JHEP 06, 004 (2008), 0804.3344; K. Hamaguchi, E. Nakamura, and S. Shirai, Phys. Lett. B666, 57 (2008), 0805.2502; R. Kitano, JHEP 11, 045 (2008), 0806.1057; A. J. Barr, G. G. Ross, and M. Serna, Phys. Rev. D78, 056006 (2008), 0806.3224; M. M. Nojiri, K. Sakurai, Y. Shimizu, and M. Takeuchi, JHEP 10, 100 (2008), 0808.1094; S. Y. Choi, M. Drees, A. Freitas, and P. M. Zerwas, Phys. Rev. D78, 095007 (2008), 0808.2410; O. Brandt (ATLAS) (2008), 0808.2934; P. Wienemann (ATLAS), AIP Conf. Proc. 1078, 286 (2009), 0809.2204; W. S. Cho, K. Choi, Y. G. Kim, and C. B. Park, Phys. Rev. D79, 031701 (2009), 0810.4853; M. Burns, K. Kong, K. T. Matchev, and M. Park, JHEP 03, 143 (2009), 0810.5576; A. J. Barr, A. Pinder, and M. Serna, Phys. Rev. D79, 074005 (2009), 0811.2138; A. Read et al., J. Phys. Conf. Ser. 119, 062041 (2008); P. Konar, K. Kong, and K. T. Matchev, JHEP 03, 085 (2009), 0812.1042; J. Hisano, M. M. Nojiri, and W. Sreethawong, JHEP 06, 044 (2009), 0812.4406; Y. Bai and Z. Han (2009), 0902.0006; M. Burns, K. T. Matchev, and M. Park, JHEP 05, 094 (2009), 0903.4371; K. T. Matchev, F. Moortgat, L. Pape, and M. Park (2009), 0906.2417.

[10] H. S. Lee (CDF) (2009), 0905.4659; Simultaneous template-based top quark mass measurement in the lepton+jets and dilepton channels using 3.2 fb-1 of CDF data (the dilepton channel uses a new variable mt2.), CDF note 9679.

[13] A. Barr et al., in preparation.

[14] CMS physics technical design report, volume II (2006), CERN-LHCC-2006-021, CMS-TDR-008-2.

[15] M. Bähr et al. (2008), 0812.0529; M. Bähr et al. (2008), 0803.0883.

[16] B. Allanach et al., Eur. Phys. J. C25, 113 (2002), hep-ph/0202233.

[17] W. Porod, Comput. Phys. Commun. 153, 275 (2003), hep-ph/0301101.

[18] M. Cacciari and G. P. Salam, Phys. Lett. B641, 57 (2006), hep-ph/0512210.

[19] M. Cacciari, G. P. Salam, and G. Soyez, JHEP 04, 063 (2008), 0802.1189.

[20] L. Randall and D. Tucker-Smith, Phys. Rev. Lett. 101, 221803 (2008), 0806.1049.

[21] (2008), cMS-PAS-SUS-08-005.

[22] A. Papaefstathiou and B. Webber (2009), 0903.2013.