Guide to transverse projections and mass-constraining variables

A. J. Barr, T. J. Khoo, P. Konar, K. Kong, C. G. Lester, K. T. Matchev, and M. Park

Phys. Rev. D **84**, 095031 — Published 22 November 2011

DOI: 10.1103/PhysRevD.84.095031

A storm in a "T" cup: the connoisseur's guide to transverse projections and mass-constraining variables

A. J. Barr,[1] T. J. Khoo,[2] P. Konar,[3] K. Kong,[4] C. G. Lester,[2] K. T. Matchev,[5] and M. Park[5]

[1]*Department of Physics, Denys Wilkinson Building, Keble Road, Oxford OX1 3RH, UK*
[2]*Department of Physics, Cavendish Laboratory, JJ Thomson Avenue, Cambridge, CB3 0HE, UK*
[3]*Theoretical Physics Group, Physical Research Laboratory, Ahmedabad, Gujarat - 380 009, India*
[4]*Department of Physics and Astronomy, University of Kansas, Lawrence, KA 66045, USA*
[5]*Department of Physics, University of Florida, Gainesville, FL 32611, USA*

This paper seeks to demonstrate that many of the existing mass-measurement variables proposed for hadron colliders (m_T, $m_{\rm eff}$, m_{T2}, missing $\vec{p}_T$, h_T, $\sqrt{\hat{s}}_{\rm min}$, etc.) are far more closely related to each other than is widely appreciated, and indeed can all be viewed as a common mass bound specialized for a variety of purposes. A consequence of this is that one may understand better the strengths and weaknesses of each variable, and the circumstances in which each can be used to best effect. In order to achieve this, we find it necessary first to revisit the seemingly empty and infertile wilderness populated by the subscript "T" (as in "p_T") in order to remind ourselves what this process of transversification actually means. We note that, far from being simple, transversification can mean quite different things to different people. Those readers who manage to battle through the barrage of transverse notation distinguishing "$\top$" from "$\vee$" or from "$\circ$", and "early projection" from "late projection", will find their efforts rewarded towards the end of the paper with (i) a better understanding of how collider mass variables fit together, (ii) an appreciation of how these variables could be generalized to search for things more complicated than supersymmetry, (iii) will depart with an aversion to thoughtless or naïve use of the so-called "transverse" methods of any of the popular computer Lorentz-vector libraries, and (iv) will take care in their subsequent papers to be explicit about which of the 61 identified variants of the "transverse mass" they are employing.

I. INTRODUCTION

Almost every analysis of data from hadron colliders uses at some point a variable which represents a "projection" of an energy or momentum into the plane transverse to the beams. The typical reason for performing these projections is that one does not wish the analysis to be sensitive to the unknown momentum – along the direction of the beams – of the quarks or gluons which collide in the 'hard' interaction. Given the widespread use of such variables it is perhaps surprising that many collider physicists are probably unaware that there exist at least two commonly-used ways of projecting of a Lorentz energy–momentum vector into the transverse plane, and that these two different methods have very different properties when the mass is non-zero (see Section III below). Furthermore, as explained later in Section V, for each of those transverse projections, there are at least two inequivalent ways that transverse vectors can be "added together", each of which has benefits and weaknesses. A careful definition of what we mean by a transverse projection forms the first part of this paper.

The later part of the paper (Sections VI–XI) deals with mass-scale (or energy-scale) variables, a variety of which have been proposed in the run-up to the LHC data-taking[1]. Though some of these variables have been constructed from careful consideration of the Lorentz symmetries of space-time, others have been created in a somewhat *ad-hoc* process, after simulations demonstrate that they provide good signal-to-background discrimination, or that they are highly correlated with the mass of some particle or particles. The main aim of this part of the paper is to demonstrate that many of these seemingly *ad-hoc* definitions are in fact not only well-motivated from the kinematical perspective, but also that the associated variables are more closely related than one might have thought.

Figure 1 illustrates some of the variables that are found to be connected in ways that are not widely appreciated. One might argue that we add little to the sum total of human knowledge by merely showing the relationships between existing variables which are already known to work well in particular roles. However, careful study of their similarities and differences not only gives insights into why (and under what circumstances) these choices are appropriate, it also fits them into a common framework – from which it is straightforward to make generalizations to more complex decay topologies.

The paper is organized as follows; first we carefully define our notation for Lorentz 1+3 vectors and their transverse projections in Section II. Then in Section III we describe the two common but inequivalent transverse projections, which we shall denote by subscripts $\top$ or $\vee$. We also introduce the special case of a "massless" transverse projection, denoted by $\circ$. In Section IV we compare the results from the three different types of projections: $\top$, $\vee$ and $\circ$. In Section V we highlight the differences between projecting into the transverse plane *before* or *after* forming composite objects. Section VI describes the general event topology targeted by new physics searches

[1] For a recent review see [1].

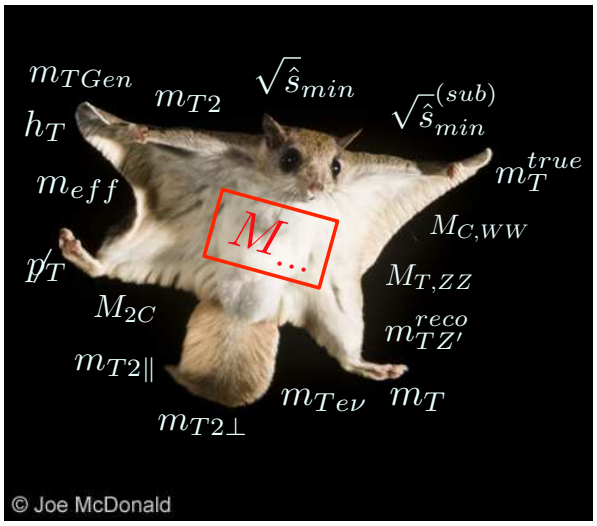

FIG. 1: The stretched, webbed limbs of the *Glaucomys volans* have been adapted by generations of natural selection to provide an ideal visual illustration of the various different, yet related, transverse mass variables (and incidentally provide an appropriate aerodynamic shape for gliding flight).

in channels with missing momentum.

All of those ingredients are put to work in Sections VII to X, which contain the main results of this paper. In Section VII we introduce the general class of mass-constraining variables which can be usefully applied for studying events containing invisible particles. The set of possible transverse mass variables is extended in Section VIII, where we consider additionally projected one-dimensional objects. Some mathematical properties of these mass-constraining variables are discussed in Section IX. Some of the variables have previously appeared elsewhere in the literature and we clarify the corresponding connections in Section X. In Section XI we illustrate the use of these variables with two simple examples: an s-channel resonant production process, for which we take inclusive Higgs boson production $pp \to h \to W^+W^- \to \ell^+\ell^- + \not{E}_T$, and a pair-production process represented by top quark production $pp \to t\bar{t} \to b\bar{b}\ell^+\ell^- + \not{E}_T$. Section XII contains a short summary and conclusions.

Appendix A contains a short guide to the currently existing computer libraries and codes which can be used for computing some of the variables described in the main body of the text. Appendix B provides derivations of extremal mass-bound results and other general mathematical proofs which are used elsewhere in the paper.

II. NOTATION AND CONVENTIONS

A. Labelling momenta and their components

In general, capital letters (P, Q, M, E, etc.) will refer to genuine 1+3 dimensional vectors, while lowercase letters (p, q, m, e, etc.) will refer to "less than 1+3" dimensional constructs. Lower indices $i, j, \ldots$ label *individual* final state particles, while lower indices $a, b, \ldots$ are used for parent particles and the corresponding *collections* of final state particles defined below in Sec. VI. We also use upper indices $\mu, \nu, \ldots$ to label the components of 1+3 vectors, and upper indices $\alpha, \beta, \ldots$ to label the components of the projected 1+2 dimensional transverse "vectors" of the types defined in Section III. The 1+3 metric $g_{\mu\nu}$ is diag$(1, -1, -1, -1)$ and the 1+2 dimensional metric $g_{\alpha\beta}$ is diag$(1, -1, -1)$. Thus the 1+3 energy-momentum vector for some particle is written $P^\mu = \left(E, \vec{P}\right)$ and the corresponding mass denoted by a capital M:

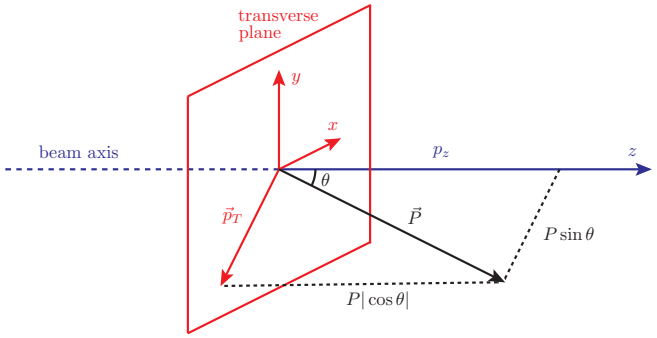

FIG. 2: The standard geometry of a collider experiment. The z axis (in blue) is oriented along the beam, while the x and y axes (in red) define the transverse plane. Any 3-dimensional vector $\vec{P}$ can be uniquely decomposed into a longitudinal component p_z and a transverse component $\vec{p}_T$.

$$M^2 = P^\mu P_\mu = E^2 - \vec{P}^{\,2}. \tag{1}$$

As illustrated in Fig. 2, any 3-dimensional vector $\vec{P}$ can be trivially decomposed into a transverse and a longitudinal component:

$$\vec{P} \equiv (\vec{p}_T, p_z)\,. \tag{2}$$

The transverse momentum $\vec{p}_T = (p_x, p_y)$ of the particle is, of course, 2-dimensional, so it has a lowercase "p". Similarly, the longitudinal momentum p_z is 1-dimensional, and is also lowercase. By contrast, the energy E measured in the detector is a component of a "1+3 dimensional thing", since it is given in terms of the 1+3 dimensional mass M and the 3-dimensional momentum $\vec{P}$:

$$E = \sqrt{M^2 + \vec{P}^{\,2}} = \sqrt{M^2 + \vec{p}_T^{\,2} + p_z^2}. \tag{3}$$

When it comes to projecting geometric 3-vectors like $\vec{P}$, the decomposition shown in eq. (2) and Fig. 2 is unambiguous. One has no other choice – the very definition of the transverse plane requires one simply to dispose of the z-component to arrive at $\vec{p}_T = (p_x, p_y)$. All the transverse projections considered in this paper (and any others

that one might invent) must share this property, or else they cannot justify being so named.

However, "projecting" the time-like component E is not, in itself, a well defined operation. What does it mean? There is not a single correct answer, but rather a number of different answers, each with different properties and motivations. How one should (and even *whether* one should) project time-like components of 1+3 Lorentz vectors is dependent on what one is trying to achieve.

B. Labelling transverse projections

In the particle physics literature, one can find evidence of at least three different types of "transverse projection" being applied to (1+3)-Lorentz vectors – although this diversity is not obvious at first glance, as the majority of papers do not explicitly state which projection they are using.[2] Even those papers which define the projection explicitly, usually neither comment on *why* the particular choice was made, nor comment on what would happen were another projection to have been used.

One of the main objectives of this paper is to place these three main types of projection side-by-side so that their differences, the things that they share, and their respective uses can be directly compared. Before we describe them in more detail, we make some remarks about notation.

In the literature, **all** of the types of projection are labelled by the **same** symbol: the letter "T". Since in this document we need to clearly distinguish the three types of projection, it is necessary for us to create our own notation for each - and we use the three symbols "$\top$", "$\vee$" and "$\circ$" for that purpose.

We will continue to use the letter "T" to indicate "generic" transverse quantities, *i.e.* quantities which are either common to all projections (e.g. the transverse momentum (2)-vector $\vec{p}_T$ already commented upon, the missing tranverse momentum vector $\vec{\not{p}}_T$, or the transverse upstream visible momentum vector $\vec{u}_T$ defined below in Section VI A) or for quantities which for historical reasons carry a transverse subscript, but which may not be tied to one type of projection to the exclusion of others (e.g. h_T).

Note that certain quantities, such as the so-called "transverse energy" and "transverse mass", are different in each of the projections. For this reason the symbol "e_T" is effectively meaningless, and should appear nowhere in this document (outside this sentence) unlike $e_\top$, $e_\vee$ and $e_\circ$ (which are all different and all well-defined). Similarly, m_T is also ambiguous, and should be specified as being $m_\top$, $m_\vee$ or $m_\circ$. In contrast, "$\vec{p}_T$" is perfectly legitimate, and indeed (as we have already noted) is equivalent to $\vec{p}_\top$, $\vec{p}_\vee$ and $\vec{p}_\circ$:

$$\vec{p}_T \equiv \vec{p}_\top \equiv \vec{p}_\vee \equiv \vec{p}_\circ \,. \tag{4}$$

III. TRANSVERSE PROJECTIONS

In this Section we describe the three different types of projections "$\top$", "$\vee$" and "$\circ$". While reading this and the following sections, the reader may find it helpful to refer to Table I for notational reference, and also to see how the results for each projection compare to those of the others.

A. The mass-preserving "$\top$" projection

The first approach we will describe, which will be denoted by a "$\top$" subscript, is the most common in the mass measurement literature. For example it is found in the early literature on the transverse mass when it was used to measure the W mass [2–6] and in the generalization of the transverse mass to pair production, namely M_{T2} (the stransverse mass) [7–29] as well as in literature relating to M_{CT} [30–35] and in reviews of the field [1].

In the $\top$ projection one defines the 1+2 dimensional transverse energy[3] $e_\top$ and transverse momentum $\vec{p}_\top$ in terms of the 1+3 dimensional mass M and 1+3 dimensional components according to

$$e_\top \equiv \sqrt{M^2 + \vec{p}_T^{\,2}} \equiv \sqrt{E^2 - p_z^2}, \tag{5}$$

$$\vec{p}_\top \equiv \vec{p}_T, \tag{6}$$

$$m_\top \equiv M. \tag{7}$$

In this case, the components of the 1+2 dimensional quantity

$$p_\top^\alpha \equiv (e_\top, \vec{p}_\top) \tag{8}$$

satisfy the mass shell condition

$$e_\top^2 - \vec{p}_\top^{\,2} = m_\top^2 = M^2 \tag{9}$$

with the 1+3 dimensional mass M.

The equivalence class for this projection function – the set of 1+3 vectors which map to the *same* 1+2 projected

[2] This may be because all forms turn out to be equivalent for massless particles.

[3] Note that in equation (5), it is the middle expression $\sqrt{M^2 + \vec{p}_T^{\,2}}$ that we use to justify our "calling" the LHS a (transverse) "energy" – since it is square root of a "mass squared plus a transverse momentum squared". Someone who saw the right hand expression first, $\sqrt{E^2 - p_z^2}$, could argue differently, and might reasonably expect us to call the whole quantity a "longitudinal mass" – since it is a square root of an "energy squared minus a longitudinal momentum squared". All this really goes to show is that the "name" of the quantity is to some extent a matter of convention rather than physics.

Quantity	Transverse projection method		
	Mass-preserving '$\top$'	**Speed-preserving '$\vee$'**	**Massless '$\circ$'**
Original (4)-momentum (1+3)-mass invariant Transverse momentum	$P^\mu = (E, \vec{p}_T, p_z)$ $M = \sqrt{E^2 - \vec{p}_T^{\,2} - p_z^2}$ $\vec{p}_T \equiv (p_x, p_y)$		
(1+2)-vectors	$p^\alpha_\top \equiv (e_\top, \vec{p}_\top)$	$p^\alpha_\vee \equiv (e_\vee, \vec{p}_\vee)$	$p^\alpha_\circ \equiv (e_\circ, \vec{p}_\circ)$
Transverse momentum under the projection	$\vec{p}_\top \equiv \vec{p}_T$	$\vec{p}_\vee \equiv \vec{p}_T$	$\vec{p}_\circ \equiv \vec{p}_T$
Transverse energy under the projection	$e_\top \equiv \sqrt{M^2 + \vec{p}_T^{\,2}}$	$e_\vee \equiv E\,\lvert\sin\theta\rvert = \lvert\vec{p}_T\rvert / V$	$e_\circ \equiv \lvert\vec{p}_T\rvert$
Transverse mass under the projection	$m^2_\top = e^2_\top - \vec{p}^{\,2}_\top$	$m^2_\vee \equiv e^2_\vee - \vec{p}^{\,2}_\vee$	$m^2_\circ \equiv e^2_\circ - \vec{p}^{\,2}_\circ = 0$
Relationship between transverse quantity and its (1+3) analogue	$m_\top = M$	$m_\vee = M\,\lvert\sin\theta\rvert$	$m_\circ = 0$
	$\frac{1}{v_\top} = \frac{1}{V}\sqrt{1 + (1 - V^2)\frac{p_z^2}{p_T^2}}$	$v_\vee = V$	$v_\circ = 1$
Equivalence classes under $(1+3) \overset{\text{proj}}{\longmapsto} (1+2)$	All P^μ with the same p_x, p_y and M	All P^μ with the same p_x, p_y and V	All P^μ with the same p_x and p_y

TABLE I: A comparison of the three transversification methods introduced in Section III.

vector under $\top$ – consists of the set of 1+3 vectors with the same $\vec{p}_T$ and M:

$$\left(\sqrt{M^2 + p_T^2 + p_z^2}, \vec{p}_T, p_z\right) \overset{\top}{\longmapsto} \left(\sqrt{M^2 + p_T^2}, \vec{p}_T\right). \tag{10}$$

The fact that all members of the equivalence class share the same mass is what motivates us to call this the "mass preserving" $\top$ projection.

Given its dominant use in the literature, it is something of a surprise that the nomenclature of the $\top$ projection is *not* adopted in the commonly used high-energy physics computer libraries such as `CLHEP` [36] or `ROOT` [37] which instead implement the alternative $\vee$ projection introduced below in Section III B. The $\top$ projection is, however, used in the "Oxbridge stransverse mass library" [38] and the U.C. Davis M_{T2} library [39]. See Appendix A and Table VII in it for a summary of library conventions.

B. The speed-preserving "$\vee$" projection

Alternatively one can follow the method of the `CLHEP` [36] and `ROOT` [37] libraries and "project" the energy on the transverse plane, using the same angle θ as for the momentum vector. As alteady seen in Fig. 2, the magnitude p_T of the transverse momentum $\vec{p}_T$ is related to the magnitude P of the 3-dimensional momentum $\vec{P}$ by

$$p_T = P\,\sin\theta, \tag{11}$$

with

$$\tan\theta \equiv \frac{p_T}{p_z}. \tag{12}$$

Thus by analogy with (11) one can define the transverse energy in terms of its 1+3 dimensional counterpart E as

$$e_\vee \equiv E\sin\theta. \tag{13}$$

Then for any *individual* 1+3 momentum vector we have the $\vee$ version of the "transverse" components

$$e_\vee \equiv E\sin\theta = \frac{p_T}{\sqrt{p_T^2 + p_z^2}}\,E, \tag{14}$$

$$\vec{p}_\vee \equiv \vec{p}_T, \tag{15}$$

$$m_\vee \equiv M\sin\theta = \frac{p_T}{\sqrt{p_T^2 + p_z^2}}\,M. \tag{16}$$

We can take the angle θ to be defined in $(0, \pi)$, so that $e_\vee$ and $m_\vee$ are always nonnegative.

In this $\vee$ method of projection we can also introduce 1+2 "vectors" which now have components

$$p^\alpha_\vee \equiv (e_\vee, \vec{p}_\vee). \tag{17}$$

The $\vee$ projected components obey a different mass shell relation than the $\top$ projected components in (9):

$$e^2_\vee - p^2_\vee = m^2_\vee \le M^2, \tag{18}$$

with the 1+2 dimensional $\vee$ projected mass $m_\vee$.

Just as an aside, one could also define the "longitudinal" components in complete analogy to (14)-(16)

$$e_z \equiv E\,|\cos\theta| = \frac{|p_z|}{\sqrt{p_T^2+p_z^2}}\,E, \tag{19}$$

$$p_z \equiv p_z, \tag{20}$$

$$m_z \equiv M\,|\cos\theta| = \frac{|p_z|}{\sqrt{p_T^2+p_z^2}}\,M, \tag{21}$$

although in what follows we shall not be making any use of those. The connection between the 1+3 dimensional quantities and the $\vee$ 1+2 dimensional components is

$$E^2 = e_\vee^2 + e_z^2, \tag{22}$$

$$M^2 = m_\vee^2 + m_z^2\ . \tag{23}$$

For massive vectors[4] the equivalence classes of the $\vee$ projection are different from those of the $\top$ projection. The mass-shell relation (18) implies that all the 1+3 vectors which map to the same 1+2 vector under the $\vee$ projection share the same value of $m_\vee = M\sin\theta$ and thus generally do *not* preserve the usual invariant mass M, since $m_\vee \neq M$ for any $\theta \neq \frac{\pi}{2}$.

A more physical picture of the equivalence class of vectors for the $\vee$ projection can be found by considering the 3-speed of the particle

$$V \equiv \frac{P}{E}. \tag{24}$$

After the $\vee$ projection, the corresponding 2-speed is given by

$$v_\vee \equiv \frac{p_\vee}{e_\vee} = \frac{p_T}{e_\vee} = \frac{P\sin\theta}{E\sin\theta} = \frac{P}{E}. \tag{25}$$

Eqs. (24) and (25) reveal that the $\vee$ projection is "speed preserving", i.e.

$$v_\vee = V, \tag{26}$$

which justifies our choice of subscript notation for this kind of transverse projection. The equivalence class for the $\vee$ projection therefore consists of all 1+3 vectors with the same $\vec{p}_T$ and speed V:

$$\left(\frac{\sqrt{p_T^2+p_z^2}}{V}, \vec{p}_T, p_z\right) \overset{\vee}{\longmapsto} \left(\frac{p_T}{V}, \vec{p}_T\right). \tag{27}$$

Note that members belonging to the same equivalence class under the $\vee$ projection (27) have the same speed, but different masses, while members of the same equivalence class under the $\top$ projection (10) have the same mass, but different speeds.

[4] See section IV B for comments concerning the massless case.

C. The massless "$\circ$" projection

The massless "$\circ$" projection defines components

$$e_\circ \equiv |\vec{p}_T|\,, \tag{28}$$

$$\vec{p}_\circ \equiv \vec{p}_T \tag{29}$$

and thereby defines a massless 1+2 vector of the form

$$p_\circ^\alpha = (|\vec{p}_T|, \vec{p}_T)\,. \tag{30}$$

The main feature of this projection is that the 1+2 vector $p_\circ^\alpha$ *always* has a null invariant

$$g_{\alpha\beta}\,p_\circ^\alpha\,p_\circ^\beta \equiv m_\circ^2 = 0. \tag{31}$$

It should be noted that $p_\top^\alpha$ and $p_\vee^\alpha$ have three degrees of freedom ($\{e_\top, p_x, p_y\}$ and $\{e_\vee, p_x, p_y\}$, correspondingly). Therefore their equivalence classes are one-dimensional, and can be parameterized by the coordinate p_z, as indicated in (10) and (27). In contrast, our '$\circ$' projected vector $p_\circ^\alpha$ has only *two* degrees of freedom, p_x and p_y — the time-like component being fully specified from p_x and p_y through $e_\circ = |\vec{p}_T|$. The equivalence class of any $p_\circ^\alpha$ vector is therefore also a $4-2=2$-dimensional object, parameterized by, say, p_z and E:

$$(E, \vec{p}_T, p_z) \overset{\circ}{\longmapsto} (|\vec{p}_T|\,, \vec{p}_T)\,. \tag{32}$$

IV. COMPARISON OF THE DIFFERENT TRANSVERSE PROJECTIONS

The three different projections discussed in Section III are pictorially represented in Fig. 3. For a given fixed value of p_T, the white region in the figure depicts all possible allowed values of the energy E and the longitudinal momentum p_z. (The yellow-shaded region $E^2 < p_T^2 + p_z^2$ is forbidden because it corresponds to a tachyonic particle with $M^2 < 0$, travelling with superluminal speed.) In this figure, we consider the plane of energy *squared* versus momentum *squared*, and in order to retain the information about the sign of the longitudinal momentum component, we plot $\mathrm{sign}(p_z)p_z^2$, so that the mapping from the (E, p_z)-plane to the $(E^2, \mathrm{sign}(p_z)\,p_z^2)$-plane is one-to-one.

Each of the three transverse projections maps a point with some given[5] values of E and p_z onto the $p_z = 0$ axis as shown. In the case of $\top$, the projection is along a line of constant mass M and results in transverse energy squared $e_\top^2 = M^2 + p_T^2$. In the $(E^2, \mathrm{sign}(p_z)\,p_z^2)$-plane, lines of constant M are straight lines, which explains our choice of quadratic power scale on the axes. Fig. 3 illustrates that the equivalence class of vectors under the

[5] For definiteness, in Fig. 3 we have chosen an illustration point with $p_z < 0$.

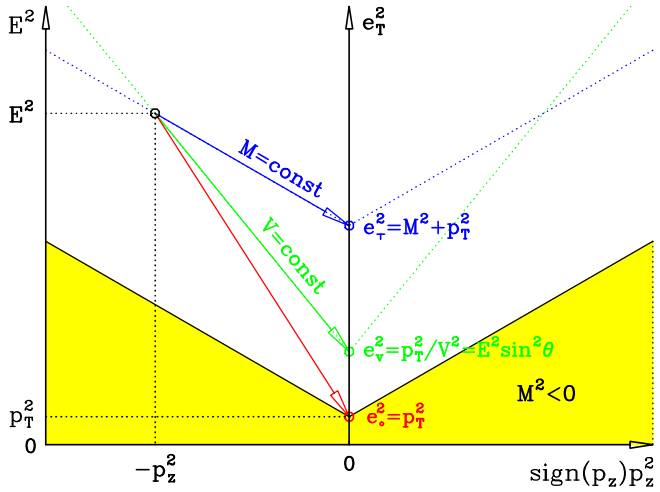

FIG. 3: A pictorial representation of the three transverse projections discussed in Section III. The colored arrows represent the mappings under the $\top$, $\vee$ and $\circ$ projections. The blue and green dotted lines represent the equivalence classes of the projected points under the $\top$ and $\vee$ projections, respectively.

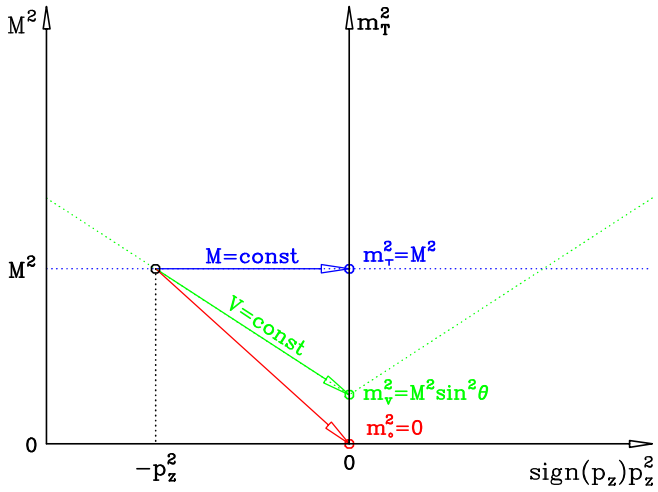

FIG. 4: The same as Fig. 3, but plotted in the $(M^2, \mathrm{sign}(p_z)\, p_z^2)$-plane.

$\top$ projection is one-dimensional: it is represented by the two blue dotted straight lines, which can be simply parameterized by the value of p_z.

The $\vee$ projection, on the other hand, projects along a line of constant speed V, as indicated in Fig. 3. In the $(E^2, \mathrm{sign}(p_z)\, p_z^2)$-plane, lines of constant V are also straight lines, albeit with a different slope. The resulting value of the transverse energy is $e_\vee = p_T/V$. The corresponding equivalence class of vectors is given by the two green dotted lines, and can also be parameterized in terms of a single parameter, say p_z.

Finally, the massless "$\circ$" projection maps *any* allowed point in the (E, p_z)-plane to the massless 1+2 vector with transverse energy $e_\circ = p_T$. The equivalence class of vectors in this case is two-dimensional, and is represented by the whole white shaded region in Fig. 3.

All of the previous discussion can be recast in the language of the $(M^2, \mathrm{sign}(p_z)\, p_z^2)$-plane, as shown in Fig. 4. In this case, the whole $M^2 \geq 0$ half-plane is allowed, and the $\top$ projection projects horizontally onto the $p_z = 0$ axis, following the blue arrow. The $\vee$ projection is also done along a straight line, following the green arrow. As before, the equivalence classes for the $\top$ and $\vee$ operations are straight lines, while the equivalence class for the "$\circ$" case is given by the whole $M^2 \geq 0$ half-plane.

A. A hierarchy among projections

As illustrated in Fig. 4, the definition of each projection imposes a hierarchy among the projected masses of the form:

$$M = m_\top \geq m_\vee \geq m_\circ = 0. \tag{33}$$

We draw attention to this hierarchy here as it will have very close analogues in the more complicated experimental mass bounds derived from each type of projection in the later sections.

Given the mass hierarchy (33), eqs. (9), (18) and (28) imply that a similar hierarchy exists for the projected energies:

$$E \geq e_\top \geq e_\vee \geq e_\circ = p_T, \tag{34}$$

which is illustrated in Fig. 3.

B. Equivalence in the massless limit

We note that in the special (but common) case in which the original four-vector is massless ($M = 0$) all projections are equivalent since

$$\lim_{M\to 0} e_\top = \lim_{M\to 0} e_\vee = e_\circ = |\vec{p}_T| \tag{35}$$

and thus

$$\lim_{M\to 0} p_\top^\alpha = \lim_{M\to 0} p_\vee^\alpha = p_\circ^\alpha = (|\vec{p}_T|, \vec{p}_T)\,. \tag{36}$$

Clearly the projections are not equivalent for massive particles, nor for collections of massless particles (unless they be collinear) since collections of massless particles can have large total invariant mass – the equivalence extends only to application to *individual* massless particles.

In practice, the statement above may also be taken as saying that all the projections are equivalent in the *high-energy limit* – i.e. the limit in which the momentum of a particle is much greater than its mass – again only at the level of *individual* high-energy particles.

Since all the projections are equivalent in the above limits, and since most individual reconstructed particles

in high-energy physics experiments satisfy one of those limits due to the small masses of the leptons and light quarks, one might wonder what all the fuss is about. However, the importance of the distinctions will be seen to arise and become very large when we consider *composite* particles, i.e. collections of massless "daughter" particles[6]. Composite particles are expected to have non-negligible masses, even when they consist of sums of (approximately) massless particles. As we already learned from the simple example considered in Figs. 3 and 4, not only do these composite particles generate very different projected 1+2 vectors, but the classes of equivalent four-vectors associated with those projections are very different as well.

V. SUMMING AND PROJECTING: EARLY VERSUS LATE PROJECTIONS

In forming transverse kinematic variables for composite particles, one needs to perform two separate operations: summation of the momentum vectors of the daughter particles, and projecting into the transverse plane. The *order* of these operations does not matter for the two *space-like* vector components:

$$\sum_i \vec{p}_{i\top} = \left(\sum_i \vec{P}_i\right)_\top , \tag{37}$$

$$\sum_i \vec{p}_{i\vee} = \left(\sum_i \vec{P}_i\right)_\vee , \tag{38}$$

$$\sum_i \vec{p}_{i\circ} = \left(\sum_i \vec{P}_i\right)_\circ , \tag{39}$$

where we use an index i to label the momenta of the individual daughter particles and the sums run over all such daughter particles[7].

However, projecting before or after the sum can make a very significant difference to the value of the time-like ($e_\top$, $e_\vee$ or $e_\circ$) component of the final 1+2 vector – and therefore the operations of projecting and summing do *not* generally commute:

$$\sum_i e_{i\top} \neq \left(\sum_i E_i\right)_\top , \tag{40}$$

$$\sum_i e_{i\vee} \neq \left(\sum_i E_i\right)_\vee , \tag{41}$$

$$\sum_i e_{i\circ} \neq \left(\sum_i E_i\right)_\circ . \tag{42}$$

One can see clearly how the order makes a difference if one considers an extreme case consisting of a pair of massless daughter particles travelling in opposite directions along the beam pipe, i.e. with 1+3 momenta

$$P_1^\mu = (E,0,0,+E) , \tag{43}$$

$$P_2^\mu = (E,0,0,-E) . \tag{44}$$

If one were to project these 1+3 momenta into the transverse plane before summing (a combined operation hereafter called *early projection*), one would find that the resulting 1+2 dimensional vector

$$\sum_i p_{i\top}^\alpha = p_{1\top}^\alpha + p_{2\top}^\alpha \tag{45a}$$

$$= (E,0,0,E)_\top + (E,0,0,-E)_\top \tag{45b}$$

$$= (0,0,0) + (0,0,0) \tag{45c}$$

$$= (0,0,0) \tag{45d}$$

is null. A null sum would also be obtained if we had used the $\vee$ or $\circ$ projections.[8] However if one were first to sum the Lorentz 1+3 vectors P_i^μ and then later project into the transverse plane (hereafter denoted *late projection*) one would find that

$$\left(\sum_i P_i^\mu\right)_\top = (P_1^\mu + P_2^\mu)_\top \tag{46a}$$

$$= ((E,0,0,E) + (E,0,0,-E))_\top \tag{46b}$$

$$= (2E,0,0,0)_\top \tag{46c}$$

$$= (2E,0,0) , \tag{46d}$$

which is clearly not the same as was found in (45d). This extreme case shows that while projecting early has the effect of reducing dependence on longitudinal momenta, projecting late means that the resultant projected composite retains much more sensitivity to the original relative momenta along the beam directions.

This concludes this section, whose main purpose was simply to highlight the difference between the "early" and the "late" transverse projection. It also underscores

[6] The need for considering composite particles arises when dealing with short-lived heavy resonances, which decay promptly to a certain collection of daughter particles, which in turn are seen in the detector. The energy and momentum of the parent resonance are correspondingly obtained by summing the measured energies and momenta of the daughter particles.

[7] Recall our convention that lowercase letters refer to 1+2 dimensional quantities and capital letters refer to 1+3 dimensional quantities. Thus in the left-hand-sides of eqs. (37)-(39) we are adding 2-dimensional transverse vectors, while in the right-hand-sides we are first adding the corresponding 3-vectors, then projecting their sum onto the transverse plane.

[8] In fact for this example we have chosen massless vectors for which the '$\top$', '$\vee$', and '$\circ$' projections are identical.

the need to develop the proper notation to distinguish between these two types of transverse projections, which we shall do below in Section VI C. The differences between the two projections will be further illustrated with the physics examples considered in the later sections. One may reasonably wonder which one of the two projections is more appropriate and should be used. In principle, the answer to this question will depend on the analysis being performed. If one is initially building a composite particle from two leptons, e.g. from a Z-boson decay $Z \to e^+e^-$, then the relative longitudinal momentum of the positron and the electron is probably a safe quantity to retain full sensitivity to in one's calculations. However, in cases where jets at large rapidity $|\eta|$ are concerned, the probability of QCD radiation grows rapidly as one gets closer and closer to the beam direction. One will often prefer not to have the high-energy end of the composite-particle spectrum dominated by combinations of low $|p_T|$, high-energy forward-going jets with other low $|p_T|$, high-energy backward-going jets, so in this latter case, early projection would probably be appropriate. Nevertheless, giving a universal prescription for selecting the "correct" transverse projection for *collections* of particles is beyond the scope of this paper. The best method will depend on non-kinematic factors, such as the size of any backgrounds, the detector resolution, and other factors that will vary from case to case.

VI. INTERPRETING EVENTS

A. Characterizing an event

Analysis of an event is a game. The aim of the game is to *interpret* the available information within a particular framework or hypothesis. In this paper we wish to employ a very general framework that will be useful for searches and mass measurements at hadron colliders (pp, $p\bar{p}$ or even $\bar{p}\bar{p}$ for that matter). Specializations of this framework will then be useful in a wide variety of different contexts. The general layout of an event is represented in Fig. 5. The figure comprises: two incoming objects, denoted by the proton lines on the left hand side; an interaction, represented by an oval 'blob'; and some final state objects, contained within the rectangles on the right hand side. Since it is the final state objects that provide the kinematic information about the event, we now take some time to explain rather carefully what we mean by them.

We define final state objects of two types. A *visible* final state object is one that leaves a signal in the detector that betrays its presence. Those signals may then be reconstructed and interpreted as an individual particle – for example as photon, electron or muon – or the signals may be indicative of a composite object, such as a QCD or tau jet. The "visible object" category is deliberately allowed to be sufficiently broad as to permit the inclusion of very heavy, visibly decaying, composite objects such as Z, W or top quarks. The classification of a final state object as 'visible' here implies not only that a signal consistent with the presence of some particle has been observed, but also that the full Lorentz energy-momentum vector of that particle can be reconstructed from the observed signal (to within some experimental precision). For most heavy visible objects (jets, W, Z, H bosons, ...), the four-momentum of the visible object must be calculated from the vector sum of its constituents.

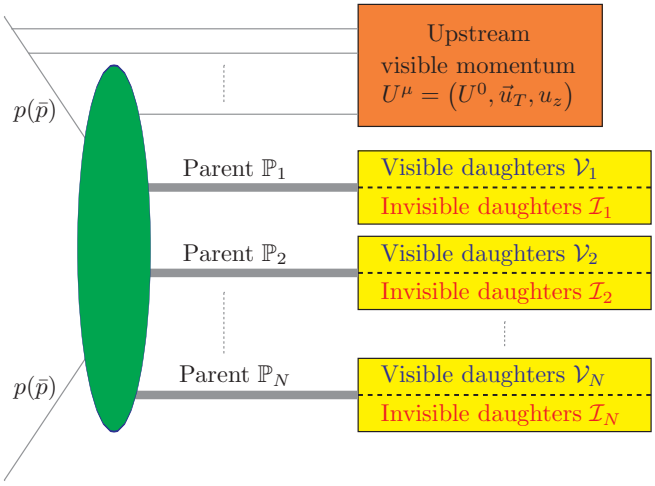

FIG. 5: The event topology for new physics searches and measurements used in this paper.

By contrast, an *invisible* final state object is one that leaves no direct signal, but the existence of which is demanded by the *interpretation* of the event being imposed by the analyst. The numbers, types and masses of any invisible final state particles form part of the interpretation of the event. The 3-momentum vectors of all invisible particles are *a priori* unknown, and are constrained only by conservation of the total momentum of the event in the plane transverse to the beam. The general framework can accommodate a final state hypothesis in which invisible particles have particular known (or rather assumed) masses, but it can also be applied when some or indeed all of those invisible particles have unknown masses.

As illustrated in Fig. 5, the next step in interpreting the event is to partition the combined set of all final state objects (visible and invisible) into subsets, which are represented by rectangles in the figure. Each final state object must be found in one and only one such subset. There is one subset per *parent* plus one further subset, the latter being labelled "upstream visible momentum" in the figure.

In our interpretation a *parent* is any short-lived object that is believed to have decayed to produce the visible and invisible final state objects in its associated set. The term 'parent' is usually associated with a short-lived heavy state, most often a reasonably narrow resonance (whether produced directly in the "hard scatter" or from decays of even heavier objects).

The general framework presented permits a variety of different interpretations for any given event. For any particular interpretation there is a corresponding partitioning of the final state into subsets. For example an event which contains evidence of an electron and a positron, and which is hypothesized to also contain a neutrino and an anti-neutrino, could be partitioned into parent/daughter combinations: $W^+ \to \{e^+, \nu_e\}$ and $W^- \to \{e^-, \bar{\nu}_e\}$ for one analysis; however another analysis might find it more appropriate to partition those objects according to the interpretation $Z^0 \to \{e^+, e^-\}$ and $Z^0 \to \{\nu_e, \bar{\nu}_e\}$.

The subset corresponding to any parent may contain any number (including zero) of visible particles and any number (including zero) of invisible particles — though it is not meaningful to have a totally empty set of daughters. The framework is very general, in that the number of parents can be arbitrary, and the nature, mass and decay mode of any parent need not be related to those of any other. There is therefore a great deal of freedom in performing the partition into subsets. We shall later be constraining the masses of the parents so the subsets should be chosen to correspond to the descendants of the parents whose invariant masses we are interested in.

Figure 5 also shows the one special (non-'parent') subset into which visible final state particles may be allocated. That set is labelled "*upstream visible momentum* (UVM)", and is designed to be a 'catch-all' that will accommodate any visible particle not allocated to any of the parent sets. This is a special set in the following senses: firstly it is permitted to contain (by assertion) only visible objects; and secondly, and crucially, final state objects allocated to this set are *not* used directly to constrain the mass of any parent. Objects in this set are only used to keep track of overall energy-momentum conservation. We do not specify the elements found in this UVM set, but in practical applications it almost always contains some contribution from "soft" particles that are unallocated to any parent. Such soft components usually include calorimeter energy found outside of jets, and low energy jets from multiple parton interactions, and perhaps from initial state radiation (ISR). As well as these soft components, one must include any other visible objects not associated with any parent. The UVM set will often contain more than just 'soft' activity — since any type of visible particle can end up therein — possibly including decay products (of heavy progenitor particles) that the analyst chose not to allocate to any parent. In practice, every hadron collider event has *some* amount of UVM. Furthermore, as discussed in [10, 12, 21, 24–26, 34], the presence of a *significant* amount of UVM can in fact be beneficial in mass reconstruction studies.

Apart from the reconstructed physics objects, another important experimental quantity is the *missing transverse momentum* in the event. This quantity is the experimental collaboration's best estimate of the amount (and direction) of momentum in any particular event that has been carried away in the plane transverse to the beam by invisible particles. It is an important quantity insofar as we will wish to apply the constraint that the missing momentum in an event is entirely due to the invisible final state objects.

B. Notation used to characterize events

We require considerable amount of notation to describe events and the hypotheses and interpretations that we layer on top of them. We have summarized the notation we have adopted in Table II — and we recommend that readers immediately compare the first section of that table with any of the three small concrete examples provided in Figures 6, 7 and 8 in order to follow later sections. For the simplest pieces of notation, Table II serves as the primary definition. Notation that requires more explanation will be described in more detail either below or at first point of use.

The N parents are labelled $\mathbb{P}_a$, $(a = 1, 2, \ldots, N)$. The set of observed visible (hypothesized invisible) daughters associated with $\mathbb{P}_a$ is labelled $\mathcal{V}_a$ ($\mathcal{I}_a$). Since no visible or invisible particle has more than one parent, we have $\mathcal{V}_a \bigcap \mathcal{V}_b = 0$ and $\mathcal{I}_a \bigcap \mathcal{I}_b = 0$ when $a \neq b$, and so the number of visible (invisible) particles may either be written as the sum of the number of visible (invisible) daughters of each parent $N_\mathcal{V} = \sum_{a=1}^{N} |\mathcal{V}_a|$, $(N_\mathcal{I} = \sum_{a=1}^{N} |\mathcal{I}_a|)$ or as the number of elements from the set of all visible (invisible) daughters $N_\mathcal{V} = |\mathcal{V}|$ $(N_\mathcal{I} = |\mathcal{I}|)$ where $\mathcal{V} = \bigcup_{a=1}^{N} \mathcal{V}_a$ $(\mathcal{I} = \bigcup_{a=1}^{N} \mathcal{I}_a)$.

As seen in Table II, in our conventions the letter "P" ("p") will be used to denote measured momenta, and the letter "Q" ("q") will be used for the momenta of any invisible or hypothesized particles. Correspondingly, the individual 4-momenta P_i^μ, $(i \in \mathcal{V})$, of the visible daughters are measured and known, while the individual 4-momenta Q_i^μ, $(i \in \mathcal{I})$, of the invisible daughters are not measured and remain unknown. We denote the masses of the visible final state particles by M_i and those of the hypothesized invisible final state particles by $\tilde{M}_i$. Similarly, we will find it convenient to denote the 3-speeds of the visible final state particles as V_i and the 3-speeds of the hypothesized invisible final state particles by $\tilde{V}_i$. In some places we will need to refer to sets of these masses or speeds, and so we define: (i) the set consisting of the hypothesized masses of *all* invisible particles:

$$\tilde{\mu} = \left\{ \tilde{M}_i \mid i \in \mathcal{I} \right\}, \tag{47}$$

(ii) the set containing only the hypothesized masses of the invisible particles assocated with parent $\mathbb{P}_a$:

$$\tilde{\mu}_a = \left\{ \tilde{M}_i \mid i \in \mathcal{I}_a \right\}, \tag{48}$$

(iii) the set consisting of the hypothesized 3-speeds of *all* invisible particles:

$$\tilde{v} = \left\{ \tilde{V}_i \mid i \in \mathcal{I} \right\}, \tag{49}$$

	Symbol	Meaning	See also
objects and sets	$\lvert\mathcal{A}\rvert$	Cardinal number (number of elements) of any finite set $\mathcal{A}$.	Figures 6, 7 and 8
	$\mathbb{P}_a$	a^{th} parent ($a \in \{1, 2, \ldots, N\}$)	
	$\mathcal{P}$	Set of all parents $\mathcal{P} \equiv \{\mathbb{P}_1, \mathbb{P}_2, \ldots, \mathbb{P}_N\}$	
	$\mathcal{V}_a$	Set of visible final state objects associated with the a^{th} parent	
	$\mathcal{I}_a$	Set of invisible final state objects associated with the a^{th} parent	
	$\mathcal{V} \equiv \bigcup_a \mathcal{V}_a$	Set of all visible final state objects ($\mathcal{V} \equiv \{\mathbb{V}_1, \mathbb{V}_2, \ldots, \mathbb{V}_{N_\mathcal{V}}\}$)	
	$\mathcal{I} \equiv \bigcup_a \mathcal{I}_a$.	Set of all invisible final state objects ($\mathcal{I} \equiv \{\mathbb{I}_1, \mathbb{I}_2, \ldots, \mathbb{I}_{N_\mathcal{I}}\}$)	
	$N \equiv \lvert\mathcal{P}\rvert$	Number of parents assumed for the interpretation being applied	
	$N_\mathcal{V} \equiv \lvert\mathcal{V}\rvert$	Total number of visible final state objects	
	$N_\mathcal{I} \equiv \lvert\mathcal{I}\rvert$	Total number of invisible final state objects	
	indices	For notational purposes, indices are used interchangably with the names of the particles they identify. For example: "$\mathcal{V}_a$" and "$\mathcal{V}_{\mathbb{P}_a}$" are equivalent; "$i \in \mathcal{V}$" and "$i \in \{1, 2, \ldots, N_\mathcal{V}\}$" are equivalent; "$a \in \mathcal{P}$" and "$a \in \{1, 2, \ldots, N\}$" are equivalent, etc.	
1+3 momenta	$P_i^\mu = (E_i, \vec{p}_{iT}, p_{iz})^\mu$	1+3 momentum components of the i^{th} final state visible object ($i \in \mathcal{V}$)	
	$Q_i^\mu = (\tilde{E}_i, \vec{q}_{iT}, q_{iz})^\mu$	Hypothesized 1+3 momentum components of the i^{th} final state invisible ($i \in \mathcal{I}$)	
	$\mathbf{P}_a^\mu \equiv \sum_{i\in\mathcal{V}_a} P_i^\mu$	Sum of 1+3 momentum components of visible objects belonging to parent $\mathbb{P}_a$	(53)
	$\mathbf{Q}_a^\mu \equiv \sum_{i\in\mathcal{I}_a} Q_i^\mu$	Sum of 1+3 momentum components of invisible objects belonging to parent $\mathbb{P}_a$	(57)
	$U^\mu \equiv \left(U^0, \vec{u}_T, u_z\right)^\mu$	Total 1+3 momentum components of the 'UVM' set	(51)
derived quantities	$\vec{\not{p}}_T$	Missing transverse momentum two vector (magnitude $\lvert\vec{\not{p}}_T\rvert = \not{p}_T$)	(51)
	$M_a \equiv M_{\mathbb{P}_a}$	Mass of the a^{th} parent ($a \in \mathcal{P}$)	
	$M_i \equiv M_{\mathbb{V}_i}$	Mass of the i^{th} visible final state object ($i \in \mathcal{V}$)	
	$\tilde{M}_i \equiv M_{\mathbb{I}_i}$	Hypothesized mass of the i^{th} invisible ($i \in \mathcal{I}$)	
	$\tilde{\mu}_a \equiv \left\{\tilde{M}_i \mid i \in \mathcal{I}_a\right\}$	Set of hypothesised masses of the invisibles associated with parent $\mathbb{P}_a$	
	$\tilde{\mu} \equiv \bigcup_a \tilde{\mu}_a$	Set of the hypothesised masses of all invisibles	
	$\mathcal{M}_a$	Hypothesized 1+3 dim. invariant mass of the composite parent particle $\mathbb{P}_a$	(62)
	$\mathbf{M}_a$	1+3 dim. invariant mass of the visibles in $\mathcal{V}_a$	(63)
	$\tilde{\mathbf{M}}_a$	1+3 dim. invariant mass of the invisibles in $\mathcal{I}_a$	(64)
	V_i	3-speed of the i^{th} visible ($i \in \mathcal{V}$)	
	$\tilde{V}_i$	Hypothesized 3-speed of the i^{th} invisible ($i \in \mathcal{I}$)	
	$\tilde{v}_a \equiv \left\{\tilde{V}_i \mid i \in \mathcal{I}_a\right\}$	Hypothesised 3-speeds of the invisibles associated with parent $\mathbb{P}_a$	(50)
	$\tilde{v} \equiv \bigcup_a \tilde{v}_a$	Set of hypothesised 3-speeds of all the invisibles	(49)
	$\not{\mathbf{M}}_a \equiv \sum_{i\in\mathcal{I}_a} \left[\tilde{M}_i\right]$	Sum of the masses of those invisibles associated with parent $\mathbb{P}_a$	(93)
	$\not{\mathbf{M}} \equiv \left\{\not{\mathbf{M}}_a \mid a \in \mathcal{P}\right\}$	Set of all 'invisible particle mass sum parameters'	(94)
	$\not{\mathbf{V}}_a \equiv \max_{i\in\mathcal{I}_a} \left[\tilde{V}_i\right]$	Largest hypothesised 3-speed of any invisible associated with parent $\mathbb{P}_a$	(117)
	$\not{\mathbf{V}} \equiv \left\{\not{\mathbf{V}}_a \mid a \in \mathcal{P}\right\}$	Set of all 'maximum invisible 3-speed parameters'	(118)
1+2 d	$p_{iT}^\alpha = (e_{iT},\, \vec{p}_{iT})^\alpha$	1+2 dim. projected energy-momentum vector for the i^{th} visible	Sec. III
	$q_{iT}^\alpha = (\tilde{e}_{iT},\, \vec{a}_{iT})^\alpha$	Hypothesized 1+2 dim. projected energy-momentum vector for the i^{th} invisible	

TABLE II: Notation used in the description of events.

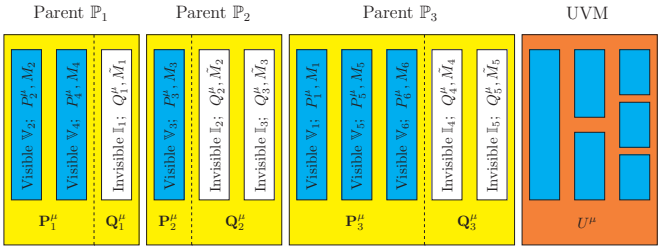

FIG. 6: This figure illustrates the notation used to label physics objects and their assignments to parent hypotheses. The figure shows a hypothesis in which six ($N_\mathcal{V} = 6$) visible physics objects $\mathcal{V} = \{\mathbb{V}_1, \mathbb{V}_2, \mathbb{V}_3, \mathbb{V}_4, \mathbb{V}_5, \mathbb{V}_6\}$ and five ($N_\mathcal{I} = 5$) invisible physics objects $\mathcal{I} = \{\mathbb{I}_1, \mathbb{I}_2, \mathbb{I}_3, \mathbb{I}_4, \mathbb{I}_5\}$ have been assigned to three ($N = 3$) parents $\mathcal{P} = \{\mathbb{P}_1, \mathbb{P}_2, \mathbb{P}_3\}$ according to the assignments $\mathcal{V}_1 = \{\mathbb{V}_2, \mathbb{V}_4\}$, $\mathcal{V}_2 = \{\mathbb{V}_3\}$, $\mathcal{V}_3 = \{\mathbb{V}_1, \mathbb{V}_5, \mathbb{V}_6\}$, $\mathcal{I}_1 = \{\mathbb{I}_1\}$, $\mathcal{I}_2 = \{\mathbb{I}_2, \mathbb{I}_3\}$ and $\mathcal{I}_3 = \{\mathbb{I}_4, \mathbb{I}_5\}$. The number of visible physics objects assigned to each parent in turn are therefore $|\mathcal{V}_1| = 2$, $|\mathcal{V}_2| = 1$ and $|\mathcal{V}_3| = 3$ and the number of invisible physics objects assigned to each parent in turn are $|\mathcal{I}_1| = 1$, $|\mathcal{I}_2| = 2$ and $|\mathcal{I}_3| = 2$.

and (iv) the set containing only the hypothesized 3-speeds of the invisible particles assocated with parent $\mathbb{P}_a$:

$$\tilde{v}_a = \left\{ \tilde{V}_i \mid i \in \mathcal{I}_a \right\}. \tag{50}$$

We denote the missing transverse *momentum* two-vector by the symbol[9] $\vec{\not{p}}_T$ and its magnitude thus $\not{p}_T$. Note that some authors use variants of the symbol "$\not{\vec{E}}_T$" to denote the missing transverse momentum,[10] but the distinction is necessary in this paper as we shall (as others should) make important distinctions between energy and momentum.

We wish to apply the constraint that the missing momentum in an event is entirely due to the $N_\mathcal{I}$ invisible particles with momenta Q_i^μ, rather than to jet mismeasurement, for example. In other words, we use the relationships expressed in:

$$\sum_{i=1}^{N_\mathcal{I}} \vec{q}_{iT} = \vec{\not{p}}_T \equiv -\vec{u}_T - \sum_{i=1}^{N_\mathcal{V}} \vec{p}_{iT}. \tag{51}$$

in which the first equality represents our desire to constrain the momenta of the invisible particles (and only those particles) using $\vec{\not{p}}_T$, while the second equality reminds us of our assumptions of how $\vec{\not{p}}_T$ is constructed as an experimentally measurable quantity. These relationships also remind us that we have assumed (i) that there are no sources of invisible momentum other than those coming from the parent decays, and (ii) that we have defined the "Upstream visible momentum" to contain all visible momentum deposits which did not originate from the decay of any parent.

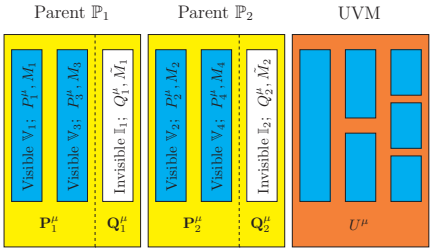

FIG. 7: This figure is provided for the benefit of readers unable to imagine a simpler version of Figure 6. (Readers finding this figure helpful need not admit this to close friends, relatives or colleagues.) The figure shows a hypothesis in which four ($N_\mathcal{V} = 4$) visible physics objects $\mathcal{V} = \{\mathbb{V}_1, \mathbb{V}_2, \mathbb{V}_3, \mathbb{V}_4\}$ and two ($N_\mathcal{I} = 2$) invisible physics objects $\mathcal{I} = \{\mathbb{I}_1, \mathbb{I}_2\}$ have been assigned to two ($N = 2$) parents $\mathcal{P} = \{\mathbb{P}_1, \mathbb{P}_2\}$ according to the assignments $\mathcal{V}_1 = \{\mathbb{V}_1, \mathbb{V}_3\}$, $\mathcal{V}_2 = \{\mathbb{V}_2, \mathbb{V}_4\}$, $\mathcal{I}_1 = \{\mathbb{I}_1\}$ and $\mathcal{I}_2 = \{\mathbb{I}_2\}$. The number of visible physics objects assigned to each parent in turn are therefore $|\mathcal{V}_1| = 2$, $|\mathcal{V}_2| = 2$ and the number of invisible physics objects assigned to each parent in turn are $|\mathcal{I}_1| = 1$, $|\mathcal{I}_2| = 1$. An explicit physics example corresponding to this figure is discussed in Section XI B.

When considering the decay of a single parent $\mathbb{P}_a$

$$\mathbb{P}_a \longrightarrow \mathcal{V}_a \cup \mathcal{I}_a. \tag{52}$$

it is useful to have notation that can refer to composite quantities, e.g. the total four momentum posessed by the visible daughters of $\mathbb{P}_a$, or the total invariant mass of that collection of visible daughters. Accordingly, as illustrated in Figures 6, 7 and 8, we denote by $\mathbf{P}_a^\mu$ the total (1+3)-momentum of the visible daughters of parent $\mathbb{P}_a$:

$$\mathbf{P}_a^\mu \equiv \left(\mathbf{E}_a, \vec{\mathbf{p}}_{aT}, \mathbf{p}_{az} \right) \equiv \sum_{i \in \mathcal{V}_a} P_i^\mu, \tag{53}$$

or in components

$$\vec{\mathbf{p}}_{aT} \equiv \sum_{i \in \mathcal{V}_a} \vec{p}_{iT}, \tag{54}$$

$$\mathbf{p}_{az} \equiv \sum_{i \in \mathcal{V}_a} p_{iz}, \tag{55}$$

[9] Note that due to its status as an experimentally measurable quantity, for the missing transverse momentum $\vec{\not{p}}_T$ we use the letter "p" as opposed to "q", even though at high values $\vec{\not{p}}_T$ is interpreted as the total transverse momentum of *invisible* particles.

[10] By right, since its meaning is derived from conservation of *momentum* in the transverse plane, the missing transverse momentum ought universally to be known as $\vec{\not{p}}_T$. Alas, much of the hadron-collider literature, especially that from the experimental collaborations, calls the missing transverse momentum the "missing *energy*" or "missing transverse *energy*" and denotes its magnitude "$\not{E}_T$" and its two vector by some variant of "$\not{\vec{E}}_T$". This is perhaps a result of history (a hang over from e^+e^- or LEP terminology where the collision of point-particles from mono-energetic beams meant that one really *could* talk about missing energy) and the fact that $\vec{\not{p}}_T$ is often reconstructed, at least in part, from calorimetric *energy* deposits under the assumption they were produced by massless physics objects.

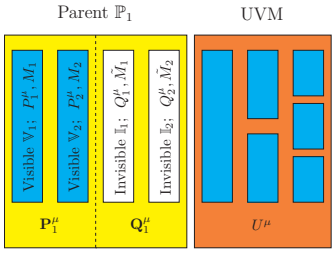

FIG. 8: This figure is provided for the benefit of readers unable to imagine an *even simpler* version of Figure 6 than was shown in Figure 7. (Readers finding this figure helpful are advised to seek gainful employment in some other field.) The figure shows a hypothesis in which two ($N_{\mathcal{V}} = 2$) visible physics objects $\mathcal{V} = \{\mathbb{V}_1, \mathbb{V}_2\}$ and two ($N_{\mathcal{I}} = 2$) invisible physics objects $\mathcal{I} = \{\mathbb{I}_1, \mathbb{I}_2\}$ have been assigned to one ($N = 1$) parent $\mathcal{P} = \{\mathbb{P}_1\}$ according to the assignments $\mathcal{V}_1 = \{\mathbb{V}_1, \mathbb{V}_2\}$ and $\mathcal{I}_1 = \{\mathbb{I}_1, \mathbb{I}_2\}$. For completeness we note $|\mathcal{V}_1| = |\mathcal{I}_1| = 2$. An explicit physics example corresponding to this figure is discussed in Section XI A.

$$\mathbf{E}_a = \sum_{i\in\mathcal{V}_a} \sqrt{M_i^2 + \vec{p}_{iT}^{\,2} + p_{iz}^2} \tag{56a}$$

$$= \sum_{i\in\mathcal{V}_a} \frac{|\vec{P}_i|}{V_i} = \sum_{i\in\mathcal{V}_a} \frac{\sqrt{\vec{p}_{iT}^{\,2} + p_{iz}^2}}{V_i}, \tag{56b}$$

where the former (latter) expression for $\mathbf{E}_a$ will be relevant later on for $\top$ ($\vee$) transverse projections since it is written in a form which depends explicitly on the masses (speeds) of the visible particles.

Similarly, we denote the total hypothesized (1+3)-momentum of the invisible daughters of parent $\mathbb{P}_a$ by

$$\mathbf{Q}_a^\mu \equiv \left(\tilde{\mathbf{E}}_a, \vec{\mathbf{q}}_{aT}, \mathbf{q}_{az}\right) \equiv \sum_{i\in\mathcal{I}_a} Q_i^\mu, \tag{57}$$

or in components

$$\vec{\mathbf{q}}_{aT} \equiv \sum_{i\in\mathcal{I}_a} \vec{q}_{iT}, \tag{58}$$

$$\mathbf{q}_{az} \equiv \sum_{i\in\mathcal{I}_a} q_{iz}, \tag{59}$$

$$\tilde{\mathbf{E}}_a = \sum_{i\in\mathcal{I}_a} \sqrt{\tilde{M}_i^2 + \vec{q}_{iT}^{\,2} + q_{iz}^2}\,, \tag{60a}$$

$$= \sum_{i\in\mathcal{I}_a} \frac{|\vec{Q}_i|}{\tilde{V}_i} = \sum_{i\in\mathcal{I}_a} \frac{\sqrt{\vec{q}_{iT}^{\,2} + q_{iz}^2}}{\tilde{V}_i}, \tag{60b}$$

where again the former (latter) expression for $\tilde{\mathbf{E}}_a$ will be relevant for $\top$ ($\vee$) transverse projections since it is written in a form which depends explicitly on the masses (speeds) of the invisible particles.

As already indicated in eqs. (53-60), we shall use **boldface** script to label "composite" momenta. Each parent is thus also treated as a composite particle, which has (1+3) momentum

$$\mathbf{P}_a^\mu + \mathbf{Q}_a^\mu \tag{61}$$

with (*a priori* unknown) (1+3) dim. invariant mass

$$\mathcal{M}_a \equiv \sqrt{g_{\mu\nu}\,(\mathbf{P}_a + \mathbf{Q}_a)^\mu\,(\mathbf{P}_a + \mathbf{Q}_a)^\nu}. \tag{62}$$

The important distinction between the **bold-face** notation for composite momenta and the ordinary notation for the momenta of individual particles is pictorially illustrated in Figures 6, 7 and 8.

Note that $\tilde{\mathbf{E}}_a$ in (57) (whose tilde is necessary to distinguish it from the energy $\mathbf{E}_a$ of the *visible* composite daughter of parent a) might legitimately be termed the missing *energy*[11] of the parent $\mathbb{P}_a$. We also introduce masses for the respective composite daughter objects as follows

$$\mathbf{M}_a \equiv \sqrt{\mathbf{E}_a^2 - \vec{\mathbf{p}}_{aT}^{\,2} - \mathbf{p}_{az}^2}, \tag{63}$$

$$\tilde{\mathbf{M}}_a(Q_i^\mu) \equiv \sqrt{\tilde{\mathbf{E}}_a^2 - \vec{\mathbf{q}}_{aT}^{\,2} - \mathbf{q}_{az}^2}, \tag{64}$$

where again a tilde refers to the invisible object. Note that the invariant mass (64) of any composite invisible daughter is not "constant" or a fixed function of measured momenta. It depends on the hypothesized invisible momenta Q_i^μ and so is part of the event hypothesis.

C. Notation for "early" and "late" transverse projections

When forming transverse kinematic variables corresponding to composite parent or daughter objects, one needs to construct the transverse 1+2 dim. analogues of (53) and (57). In doing so, one inevitably has to face the issue discussed in Section V — whether the agglomeration of individual particles into a composite object is done before or after projecting into the transverse plane. As we already saw in Section V, the two outcomes are generally quite different, since the composite object is usually massive. This is why we shall need to develop some additional notation to help us keep track of the order in which those operations are performed. Correspondingly, for the remainder of this paper we shall adopt the following principle: *in forming transverse quantities for composite objects, the order in which the various operations of agglomeration and projection are taken will be specified by the order (from left to right) of the corresponding subscript indices.*

Let us illustrate this principle with a few relevant examples. The "late-projected" (or "early-partitioned")

[11] Really the missing *energy* rather than the missing *momentum*!

version of the composite visible momentum (53) is denoted by $\mathbf{p}^{\alpha}_{aT}$

$$\mathbf{p}^{\alpha}_{aT} \equiv (\mathbf{e}_{aT}, \vec{\mathbf{p}}_{aT}) \tag{65}$$

while the alternative "early-projected" (or "late-partitioned") version is denoted by $\mathbf{p}^{\alpha}_{Ta}$:

$$\mathbf{p}^{\alpha}_{Ta} \equiv (\mathbf{e}_{Ta}, \vec{\mathbf{p}}_{Ta})\,. \tag{66}$$

We remind the reader that the generic index "T" in (65) and (66) stands for either "$\top$", "$\vee$" or "$\circ$", as discussed in Section III.

We already saw in Section V (eqs. (37-39)) that the space-like components of (65) and (66) are equivalent for any choice of "T":

$$\vec{\mathbf{p}}_{aT} \equiv \vec{\mathbf{p}}_{Ta} = \sum_{i\in\mathcal{V}_a} \vec{p}_{iT}, \tag{67}$$

but the time-like components $\mathbf{e}_{aT}$ and $\mathbf{e}_{Ta}$ are generally different. For example, in the case of $T = \top$, the late-projected (early-partitioned) transverse energy $\mathbf{e}_{a\top}$ is given by

$$\mathbf{e}_{a\top} \;=\; \sqrt{\mathbf{M}_a^2 + \vec{\mathbf{p}}_{aT}^{\,2}} = \sqrt{\mathbf{E}_a^2 - \mathbf{p}_{az}^2} \tag{68a}$$

$$= \sqrt{\left(\sum_{i\in\mathcal{V}_a} \sqrt{M_i^2 + \vec{p}_{iT}^{\,2} + p_{iz}^2}\right)^2 - \left(\sum_{i\in\mathcal{V}_a} p_{iz}\right)^2}, \tag{68b}$$

while the early-projected (late-partitioned) transverse energy $\mathbf{e}_{\top a}$ is given by

$$\mathbf{e}_{\top a} \;=\; \sum_{i\in\mathcal{V}_a} e_{i\top} = \sum_{i\in\mathcal{V}_a} \sqrt{M_i^2 + \vec{p}_{iT}^{\,2}}\,. \tag{69}$$

In the case of $T = \vee$ projections, the corresponding transverse energies are given by

$$\mathbf{e}_{a\vee} \;=\; \frac{\mathbf{p}_{aT}}{\sqrt{\mathbf{p}_{aT}^2 + \mathbf{p}_{az}^2}}\,\mathbf{E}_a\,, \tag{70}$$

$$\mathbf{e}_{\vee a} \;=\; \sum_{i\in\mathcal{V}_a} e_{i\vee} = \sum_{i\in\mathcal{V}_a} \frac{p_{iT}}{V_i}\,. \tag{71}$$

Finally, for $T = \circ$, the two transverse energies are

$$\mathbf{e}_{a\circ} \;=\; \left|\sum_{i\in\mathcal{V}_a} \vec{p}_{iT}\right|, \tag{72}$$

$$\mathbf{e}_{\circ a} \;=\; \sum_{i\in\mathcal{V}_a} p_{iT}. \tag{73}$$

The same conventions apply to the transverse projections of the composite momentum of a collection of invisible daughter particles: the "late-projected" (or "early-partitioned") version of the composite invisible momentum (57) is denoted by $\mathbf{q}^{\alpha}_{aT}$

$$\mathbf{q}^{\alpha}_{aT} \equiv (\tilde{\mathbf{e}}_{aT}, \vec{\mathbf{q}}_{aT})\,, \tag{74}$$

while the alternative "early-projected" (or "late-partitioned") version is denoted by $\mathbf{q}^{\alpha}_{Ta}$:

$$\mathbf{q}^{\alpha}_{Ta} \equiv (\tilde{\mathbf{e}}_{Ta}, \vec{\mathbf{q}}_{Ta})\,. \tag{75}$$

Again, the space-like components of (74) and (75) are the same:

$$\vec{\mathbf{q}}_{aT} \equiv \vec{\mathbf{q}}_{Ta} = \sum_{i\in\mathcal{I}_a} \vec{q}_{iT}, \tag{76}$$

but the time-like components are not. Altogether, there are 6 different possibilities:

$$\tilde{\mathbf{e}}_{a\top} \;=\; \sqrt{\tilde{\mathbf{M}}_a^2 + \vec{\mathbf{q}}_{aT}^{\,2}} = \sqrt{\tilde{\mathbf{E}}_a^2 - \mathbf{q}_{az}^2} \tag{77a}$$

$$= \sqrt{\left(\sum_{i\in\mathcal{I}_a} \sqrt{\tilde{M}_i^2 + \vec{q}_{iT}^{\,2} + q_{iz}^2}\right)^2 - \left(\sum_{i\in\mathcal{I}_a} q_{iz}\right)^2}, \tag{77b}$$

$$\tilde{\mathbf{e}}_{\top a} \;=\; \sum_{i\in\mathcal{I}_a} \tilde{e}_{i\top} = \sum_{i\in\mathcal{I}_a} \sqrt{\tilde{M}_i^2 + \vec{q}_{iT}^{\,2}}\,, \tag{78}$$

$$\tilde{\mathbf{e}}_{a\vee} \;=\; \frac{\mathbf{q}_{aT}}{\sqrt{\mathbf{q}_{aT}^2 + \mathbf{q}_{az}^2}}\,\tilde{\mathbf{E}}_a \tag{79a}$$

$$= \frac{\left|\sum_{i\in\mathcal{I}_a} \vec{q}_{iT}\right| \sum_{i\in\mathcal{I}_a} \frac{\sqrt{q_{iT}^2 + q_{iz}^2}}{V_i}}{\sqrt{\left(\sum_{i\in\mathcal{I}_a} \vec{q}_{iT}\right)^2 + \left(\sum_{i\in\mathcal{I}_a} q_{iz}\right)^2}}, \tag{79b}$$

$$\tilde{\mathbf{e}}_{\vee a} \;=\; \sum_{i\in\mathcal{I}_a} \tilde{e}_{i\vee} = \sum_{i\in\mathcal{I}_a} \frac{q_{iT}}{\tilde{V}_i}\,, \tag{80}$$

$$\tilde{\mathbf{e}}_{a\circ} \;=\; \left|\sum_{i\in\mathcal{I}_a} \vec{q}_{iT}\right|, \tag{81}$$

$$\tilde{\mathbf{e}}_{\circ a} \;=\; \sum_{i\in\mathcal{I}_a} q_{iT}. \tag{82}$$

In general, our principle of index ordering will extend to *any* transverse invariant mass or transverse energy variable. For example, in analogy to (68-73) and (77-82) there will be six different versions of the transverse masses of the composite parent particles and they will be denoted by $\mathcal{M}_{aT}$ or $\mathcal{M}_{Ta}$, with $T \in \{\top, \vee, \circ\}$.

D. Comments on the characterization framework

Note that we do not place any a priori restrictions on the values of N, $N_{\mathcal{I}}$ or on the way invisible particles are partitioned into the subsets $\mathcal{I}_a$. In contrast, many studies on supersymmetry (SUSY) or Universal Extra Dimensions (UED) in the hadron collider literature are predicated on the following assumptions:

- $N = 2$. This assumption is motivated if the new particles are charged under a conserved Z_2 parity, like R-parity in supersymmetry or KK-parity in UED. However, other discrete symmetries are also possible, e.g. Z_3 [40, 41] and higher [42, 43], which could in principle allow for $N > 2$. Even in models with a Z_2 parity one could still consider the production of any *even* number of parents, e.g. $N = 4$, $N = 6$, etc.

- $|\mathcal{I}_a| = 1$ for all a. In the conventional models with conserved Z_2 parity, this assumption implies that the decay of each parent generates one and only one massive invisible particle, excluding the possibility of any neutrinos appearing among the invisible particles. However, this assumption is not guaranteed — even in the conventional SUSY models with conserved R-parity, SM neutrinos can easily appear among the decay products of charginos, sleptons, W-bosons, heavy flavor quarks (especially top), taus, etc. Furthermore, R-parity conservation only guarantees that a given SUSY parent must decay into an *odd* (not necessarily 1) number of SUSY particles. Finally, a Z_3 symmetry could allow *two* massive invisible particles per parent, see e.g. [44, 45].

Because of all these caveats, we prefer to keep our discussion as general as possible, and first define our invariant mass variables below in Sec. VII for any N and $N_{\mathcal{I}}$, before specializing to $N = 1$ and $N = 2$ for illustration purposes only.

One might ask whether the methods proposed here can be usefully applied to events with "the wrong" value of N. The answer to this question is "yes", and we shall demonstrate this explicitly below in Section XI B (see in particular Figure 12) where we shall apply $N = 1$ variables in an example where not one, but two parents were produced in the hard scatter. That study will show that one can sometimes obtain useful information from variables with "the wrong" value of the parent number N.

E. Choosing the partitioning

In conclusion of this section, one more comment regarding the partitioning is in order. One may wonder how one should decide whether a given visible particle should be counted among the set of visible daughters or whether it should be included in the "Upstream visible momentum" category. The answer to this question depends on the particular case at hand. There are simple cases of final states where the outgoing particles can be unambiguously associated with the particle sets $\mathcal{V}_a$ that match the expected decay products of an assumed parent. For example, a high p_T, isolated reconstructed lepton is unlikely to have come from the typical sources of UVM such as initial state radiation (ISR), multiple parton-parton interactions (MPI), multiple hadron-hadron interactions (pileup) etc., and can probably be safely counted as a visible daughter. On the other hand, there are also cases (typically involving jets of hadrons) where the correct partitioning is not obvious at all. In such cases, one possible approach is to consider all possible partitions, see e.g. [9]. Another possible approach would be to devise a certain set of cuts, using the generic differences between the kinematics of ISR jets and jets from heavy parent decays [28, 46–48]. Examples of choices for particular physical examples can be found in Section X.

VII. THE MASS-BOUND VARIABLES

A. Guiding principles

The guiding principle we employ for creating useful hadron-collider event variables, is that: *we should place the best possible bounds on any Lorentz invariants of interest, such as parent masses or the center-of-mass energy $\hat{s}^{1/2}$, in any cases where it is not possible to determine the actual values of those Lorentz invariants due to incomplete event information.* Such incomplete information could take the form of lack of knowledge of the longitudinal momentum of the primary collision, or lack of knowledge of the 4-momenta of individual invisible particles, or lack of knowledge of the number of invisible particles which were present, etc.

We contrast this principle with the alternative approach that is used to motivate event variables without any explicit regard to whether they have an interpretation as an optimal bound of a Lorentz invariant. This alternative approach tends to recommend the use of variables that are somewhat ad-hoc, but by construction possess useful invariances (such as invariance under longitudinal boosts) which are designed to remove sensitivity to quantities that are unknown. One example of this latter class of variables, which are usually considered to be simply "made up" without reference to our guiding principle, would include the missing transverse momentum $\not{p}_T$ (aready seen in (51)) obtained by adding all transverse visible momenta vectorially. Another would be the so called h_T variable[12] which is defined as the *scalar* sum

[12] Note that the definition of h_T in the literature is not well standardized. Indeed even one LHC experiment has managed to define it in three different and inequivalent ways in the space of just a few years, and sometimes even inconsistently in a single document (see Section 2 of [1] for further details). The definition we adopt in equation (83) is the definition which appears, at present, to be the most widely used in the literature. We note that a conceivable consequence of this paper might be that purists will in the future settle on a definition in which h_T is defined as a sum of transverse energies e_T instead of transverse momenta, whereby three different variants would be possible: $h_\top$, $h_\vee$ and $h_\circ$ (though these three definitions will be almost equivalent under most practical experimental conditions, where

of the transverse momenta of some class of visible objects (typically jets) in the event:

$$h_T \equiv \sum_{i=1}^{N_{\mathcal{V}}} p_{iT}. \tag{83}$$

Another example is the sum of these two variables:

$$m_{\text{eff}} \equiv h_T + \not{p}_T, \tag{84}$$

a quantity which can be traced back to the original literature [49] and has become known as an "effective mass", even though it is not a mass.[13]

The main disadvantage of variables like h_T and m_{eff}, is that they do not utilize all the information available; for example they are completely insensitive to all angles in the transverse plane. This is why here we would like to construct a more optimal class of variables, to wit, those which bound the invariants of interest. These too must be invariant under global longitudinal boosts since a bound cannot depend on unknown quantities. However by explict construction we can ensure that they also make best use of any available kinematic information.

B. Construction of mass-bounding variables

We are now ready to define the general procedure that can be used to construct the mass-bound variables. In fact, we shall describe a broad class of such variables, where each individual variable $M_{\{indices\}}$ will be labelled by a certain set of indices $\{indices\}$ indicative of the way the particular variable was constructed, namely:

- Since we are targeting the general event topology of Fig. 5, where we imagine the inclusive production of N parents, each one of our variables will necessarily carry a corresponding index N. In the process of constructing such a variable, we will have to partition (and then agglomerate) the observed visible particles in the event into N groups $\mathcal{V}_a$, $(a = 1, 2, \ldots, N)$, as already explained in Section VI A. We will then form the 1+3 dimensional invariant mass of each parent $\mathbb{P}_a$

$$\mathcal{M}_a \equiv \sqrt{g_{\mu\nu}\,(\mathbf{P}_a^\mu + \mathbf{Q}_a^\mu)(\mathbf{P}_a^\nu + \mathbf{Q}_a^\nu)}, \tag{85}$$

  which is constructed out of the 1+3 momenta $\mathbf{P}_a^\mu$ and $\mathbf{Q}_a^\mu$ of the respective composite daughter particles (see Section VI B).

- Optionally, instead of the 1+3 dimensional parent mass (85), we may choose to consider the corresponding early-partitioned (late-projected) transverse mass

$$\mathcal{M}_{aT} \equiv \sqrt{g_{\alpha\beta}\,(\mathbf{p}_{aT}^\alpha + \mathbf{q}_{aT}^\alpha)(\mathbf{p}_{aT}^\beta + \mathbf{q}_{aT}^\beta)}, \tag{86}$$

  or the late-partitioned (early-projected) transverse mass

$$\mathcal{M}_{Ta} \equiv \sqrt{g_{\alpha\beta}\,(\mathbf{p}_{Ta}^\alpha + \mathbf{q}_{Ta}^\alpha)(\mathbf{p}_{Ta}^\beta + \mathbf{q}_{Ta}^\beta)}, \tag{87}$$

  where $\mathbf{p}_{aT}^\alpha$, $\mathbf{p}_{Ta}^\alpha$, $\mathbf{q}_{aT}^\alpha$ and $\mathbf{q}_{Ta}^\alpha$ are the 1+2 dimensional momentum vectors defined in (65), (66), (74) and (75), correspondingly, and the index T takes values in $\{\top, \vee, \circ\}$, as explained in Section III.[14]

- The last step is to consider the *largest* hypothesized parent mass ($\max[\mathcal{M}_a]$, $\max[\mathcal{M}_{aT}]$ or $\max[\mathcal{M}_{Ta}]$ as appropriate) and *minimize* it over all possible values of the unknown invisible momenta consistent with the constraints. This minimization is always a well-defined, unambiguous operation, which yields a unique numerical answer [50], which we shall denote as

$$M_N \;\equiv\; \min_{\sum \vec{q}_{iT} = \vec{\not{p}}_T} \left[\max_a [\mathcal{M}_a]\right], \tag{88}$$

$$M_{NT} \;\equiv\; \min_{\sum \vec{q}_{iT} = \vec{\not{p}}_T} \left[\max_a [\mathcal{M}_{aT}]\right], \tag{89}$$

$$M_{TN} \;\equiv\; \min_{\sum \vec{q}_{iT} = \vec{\not{p}}_T} \left[\max_a [\mathcal{M}_{Ta}]\right], \tag{90}$$

the visible particles are approximately massless).

[13] In keeping with our conventions from Section II, we use lowercase letters for both h_T and m_{eff}, since they are not 1+3 dimensional quantities.

[14] We should point out that the projection specification $T \in \{\top, \vee, \circ\}$ refers to operations on the *visible* particles. One should keep in mind that the visible and the invisible composite particles are a priori independent and so could, in principle, be treated differently, both in terms of the order of the operations, as well as regarding the type of transverse projections. For example, consider the M_{NT} class of variables, where one first forms composite visible particles and transversifies later. In principle, for the invisible particles, one could perform those operations in the opposite order and instead of (86) consider

$$\sqrt{g_{\alpha\beta}\,(\mathbf{p}_{aT}^\alpha + \mathbf{q}_{Ta}^\alpha)(\mathbf{p}_{aT}^\beta + \mathbf{q}_{Ta}^\beta)}$$

instead. Furthermore, one could choose a different type of transverse projection for the invisibles than for the visible sector, e.g.

$$\sqrt{g_{\alpha\beta}\,(\mathbf{p}_{a\top}^\alpha + \mathbf{q}_{a\vee}^\alpha)(\mathbf{p}_{a\top}^\beta + \mathbf{q}_{a\vee}^\beta)}$$

and so on. One might therefore wonder whether projected variables need to carry additional indices indicating how the *invisible* sector is being handled. In the following, for simplicity we shall assume that the invisible particles are always projected in exactly the same way as the corresponding visible particles, so that the transversification indices uniquely describe the transverse projections of both visible and invisible daughters. Those readers who are curious about the remaining cases (when the visibles and the invisibles are projected differently) can easily infer the corresponding results from the formulas given below.

Type of variables	Operations			Notation
	First	Second	Third	
Unprojected	Partitioning	Minimization	—	M_N
Early partitioned (late projected) M_{NT}	Partitioning	$T = \top$ projection	Minimization	$M_{N\top}$
	Partitioning	$T = \vee$ projection	Minimization	$M_{N\vee}$
	Partitioning	$T = \circ$ projection	Minimization	$M_{N\circ}$
Late partitioned (early projected) M_{TN}	$T = \top$ projection	Partitioning	Minimization	$M_{\top N}$
	$T = \vee$ projection	Partitioning	Minimization	$M_{\vee N}$
	$T = \circ$ projection	Partitioning	Minimization	$M_{\circ N}$

TABLE III: Method of constructing the mass-bound variables and corresponding notation. The Table lists the sequence of operations performed in the calculation of each variable. "Partitioning" refers to the operations discussed in Sec. VI A and VI B of partitioning the final state particles into daughter sets and then adding the momenta in each set to form corresponding composite daughter particles. "Minimization" implies minimizing the largest (suitably projected) parent mass with respect to (the relevant components of) the missing momenta of all invisible particles; while the remaining operations involve the different types of transverse projections defined and discussed in Section III.

as indicated in Table III. The minimization over the unknown parameter is performed in order to guarantee that the resultant variable cannot be larger than the mass of the heaviest parent, resulting in an event-by-event lower bound on the mass of the heaviest parent.

These are the basic steps, leading to the variables displayed in Table III. This basic set of variables will be further extended in Section VIII below, by considering a second level of projections *within* the transverse plane. For the remainder of this section, however, we shall stick to the basic procedures above and focus on the simplest classes of variables displayed in Table III, namely the "unprojected" M_N and the "singly projected" M_{NT} and M_{TN} variables.

C. The variables: M_N, M_{NT} and M_{TN}

In this subsection we provide analytic formulas (where available) for calculating each of the basic mass-bound variables from Table III on an event-by-event basis.

1. *The usual ("unprojected") invariant mass: M_N*

Here we work directly with the usual (1+3)-dimensional invariant masses $\mathcal{M}_a$ of the parent particles $\mathbb{P}_a$:

$$\mathcal{M}_a^2(\mathbf{P}_a, \mathbf{Q}_a, \tilde{\mu}_a) \equiv (\mathbf{P}_a + \mathbf{Q}_a)^2 \tag{91a}$$

$$= \left(\mathbf{E}_a + \tilde{\mathbf{E}}_a\right)^2 - (\vec{\mathbf{p}}_{aT} + \vec{\mathbf{q}}_{aT})^2 - (\mathbf{p}_{az} + \mathbf{q}_{az})^2 \,. \tag{91b}$$

The unprojected invariant mass variable M_N is defined by the right hand side of

$$M_N(\not{\mathbf{M}}) \equiv \min_{\sum \vec{q}_{iT} = \not{\vec{p}}_T} \left[\max_a \left[\mathcal{M}_a(\mathbf{P}_a, \mathbf{Q}_a, \tilde{\mu}_a)\right]\right], \tag{92}$$

where the minimization needs to be performed over $3N_{\mathcal{I}}$ degrees of freedom ($\vec{q}_{iT}$ and q_{iz} for $i = 1, 2, \ldots, N_{\mathcal{I}}$), subject to the two scalar constraints (51) supplied by transverse momentum conservation. The invisible particle momenta $\vec{q}_{iT}$ and q_{iz} are fixed by the minimization and M_N does not depend on them.

Note that we have emphasized in the left hand side of (92) that M_N turns out *not* to be a function of the $N_{\mathcal{I}}$ individual invisible mass hypotheses $\tilde{M}_i$ in $\tilde{\mu} = \bigcup_a \tilde{\mu}_a$, but instead turns out (see proof in Section IX A) to be a function of the set

$$\not{\mathbf{M}} = \left\{\not{\mathbf{M}}_a \mid a \in \mathcal{P}\right\}. \tag{93}$$

containing the N "invisible mass-sum parameters, $\not{\mathbf{M}}_a$" defined by

$$\not{\mathbf{M}}_a \equiv \sum_{i \in \mathcal{I}_a} \tilde{M}_i. \tag{94}$$

These mass parameters are simple arithmetic sums of the hypothesized masses of the individual invisible particles associated with any given parent $\mathbb{P}_a$.

Notice the simplification in going from the individual parent masses $\mathcal{M}_a$ to the variable M_N. The individual parent masses $\mathcal{M}_a$ collectively depend on *all* invisible particle masses $\tilde{M}_i$, (a total of $N_{\mathcal{I}}$ parameters), while the invariant mass variable M_N defined in (92) only depends on the N summed-invisible-mass parameters $\not{\mathbf{M}}_a$, ($a = 1, 2, \ldots, N$), which are simply related to the individual particle masses $\tilde{M}_i$ via (94). In the most common cases of $N = 1$ or $N = 2$, we will therefore have to deal with only one or two unknown invisible mass-sum

parameters. A similar reduction in complexity will be found when we consider the $\vee$ projected variables, but there the mass bound will end up depending on a speed-related parameter for each parent. We see that from now on the index N can be interpreted not only as the number of parents, but also as the number of relevant independent mass inputs characterizing the invisible sector.

The preceding discussion is best illustrated with a specific example. Let us consider the simplest case of $N = 1$. The minimization of the corresponding variable M_1 with respect to $\vec{q}_{iT}$ and q_{iz} is straightforward. One finds that the minimum is located at [50]

$$\vec{q}_{iT} \;=\; \not{\vec{p}}_T \frac{\tilde{M}_i}{\not{\mathbf{M}}_1}, \tag{95}$$

$$q_{iz} \;=\; \mathbf{p}_{1z} \frac{\tilde{M}_i}{\not{\mathbf{M}}_1} \sqrt{\frac{\not{\mathbf{M}}_1^2 + \not{p}_T^2}{\mathbf{M}_1^2 + \mathbf{p}_{1T}^2}}, \tag{96}$$

and its value (see [46]) is given by

$$M_1^2(\not{\mathbf{M}}_1) \equiv \left(\sqrt{\mathbf{M}_1^2 + \mathbf{p}_{1T}^2} + \sqrt{\not{\mathbf{M}}_1^2 + \not{p}_T^2} \right)^2 - u_T^2 \tag{97}$$

in which, to save space, we have slightly abused our notation by writing $M_1^2(\not{\mathbf{M}}_1)$ in place of $M_1^2(\{\not{\mathbf{M}}_1\})$ — a convention we will adopt throughout this document wherever $N = 1$. We remind the reader that $\mathbf{M}_1$ is the measured (1+3)-mass of the (single) visible composite daughter (see also eq. (63))

$$\mathbf{M}_1 \equiv \sqrt{\mathbf{E}_1^2 - \vec{\mathbf{p}}_{1T}^2 - \mathbf{p}_{1z}^2}, \tag{98}$$

while $\not{\mathbf{M}}_1$ is the only invisible mass parameter needed[15] defined in (94)

$$\not{\mathbf{M}}_1 \equiv \sum_{i=1}^{N_\mathcal{I}} \tilde{M}_i. \tag{99}$$

In Ref. [46], the quantity $M_1(\not{\mathbf{M}}_1)$ defined in (97) was labelled $\sqrt{\hat{s}}_{\min}^{(\mathrm{sub})}$:

$$M_1(\not{\mathbf{M}}_1) \equiv \sqrt{\hat{s}}_{\min}^{(\mathrm{sub})}(\not{\mathbf{M}}_1), \tag{100}$$

since it provides a lower bound on the parton-level center-of-mass energy of the parent subsystem $\mathcal{V}_1 \oplus \mathcal{I}_1$, not counting the uninteresting upstream visible momentum U^μ. In the special case of a vanishing upstream momentum ($u_T = 0$), $M_1(\not{\mathbf{M}}_1)$ reduces to the global variable $\sqrt{\hat{s}}_{\min}$ from [50]:

$$\lim_{u_T \to 0} M_1(\not{\mathbf{M}}_1) = \sqrt{\hat{s}}_{\min}(\not{\mathbf{M}}_1). \tag{101}$$

We will not consider the next simplest example (M_2) until Section X H, as simple analytic (as opposed to numerical or iterative) formulae for it are only known to exist in certain special cases [51], such as when $\not{\mathbf{M}}_1 = \not{\mathbf{M}}_2 = \mathbf{M}_1 = \mathbf{M}_2 = 0$, or when the upstream visible momentum $\vec{u}_T$ is either zero or (anti-)parallel to the missing transverse momentum $\not{\vec{p}}_T$.

2. *The early partitioned, $\top$-projected invariant mass: $M_{N\top}$*

Here the momenta $\mathbf{P}_a^\mu$ and $\mathbf{Q}_a^\mu$ of the composite particles are first formed in 1+3 dimensions, as in (53) and (57), then *afterwards* are projected on the transverse plane according to the mass-preserving $\top$ method defined in eq. (8) of Sec. III A. This results in *transverse* masses of the parents given by

$$\mathcal{M}_{a\top}^2(\mathbf{p}_{a\top}^\alpha, \mathbf{q}_{a\top}^\alpha, \tilde{\mu}_a) \;\equiv\; (\mathbf{p}_{a\top} + \mathbf{q}_{a\top})^2 \tag{102a}$$

$$\equiv (\mathbf{e}_{a\top} + \tilde{\mathbf{e}}_{a\top})^2 - (\vec{\mathbf{p}}_{aT} + \vec{\mathbf{q}}_{aT})^2, \tag{102b}$$

where the transverse momenta $\vec{\mathbf{p}}_{aT}$ and $\vec{\mathbf{q}}_{aT}$ are given by (54) and (58), while the transverse energies $\mathbf{e}_{a\top}$ and $\tilde{\mathbf{e}}_{a\top}$ are given by (68) and (77).

Then the "early partitioned, $\top$-projected" variable $M_{N\top}$ is defined in a manner very similar to (92)

$$M_{N\top}(\not{\mathbf{M}}) \equiv \min_{\sum \vec{q}_{iT} = \not{\vec{p}}_T} \left[\max_a \left[\mathcal{M}_{a\top}(\mathbf{p}_{a\top}^\alpha, \mathbf{q}_{a\top}^\alpha, \tilde{\mu}_a) \right] \right]. \tag{103}$$

Just like M_N, this variable also depends only[16] on the N summed-invisible-mass parameters $\not{\mathbf{M}}_a$ within $\not{\mathbf{M}}$ as opposed to the $N_\mathcal{I}$ individual invisible masses $\tilde{M}_i$ within $\tilde{\mu}$. Eq. (103) again represents a constrained minimization problem for the $3N_\mathcal{I}$ variables $\vec{q}_{iT}$ and q_{iz}. Note that in spite of its "transverse" index, $M_{N\top}$ still depends on the

[15] Note the analogy between $\not{\vec{p}}_T$ and $\not{\mathbf{M}}_1$. $\not{\vec{p}}_T$ measures the *total* transverse momentum of the whole collection of missing particles. Similarly, $\not{\mathbf{M}}_1$ measures the *total* mass of the whole collection of missing particles. Both $\not{\vec{p}}_T$ and $\not{\mathbf{M}}_1$ are given by simple sums of the corresponding quantities $\vec{q}_{iT}$ and $\tilde{M}_i$ of the individual missing particles, compare (51) and (99).

[16] At this point readers who are familiar with the Cambridge m_{T2} variable [7, 8] have probably recognized that for the special case of $N = 2$, the $M_{N\top}$ variable (103) recovers the Cambridge m_{T2}. Note that the original literature [8] on the Cambridge m_{T2} variable also defined more general variables m_{TX}, e.g. m_{T3}, m_{T4}, etc. However, we caution readers to make the distinction between the index "N" in $M_{N\top}$, which refers to the number of hypothesized *parents*, and the index "X" in the Cambridge m_{TX}, which stood for the total number of *invisible particles* (in this paper denoted by $N_\mathcal{I}$). For example, the index "2" in the Cambridge m_{T2} notation implies the presence of exactly two *invisible* particles, the number of parents already being implicitly assumed to be two. In contrast, the variable $M_{2\top}$ defined in (103) does not imply any particular number of invisible particles, and in this sense is equivalent to the whole class of m_{TX} for any X.

longitudinal momenta q_{iz} through the transverse energy $\tilde{\mathbf{e}}_{a\top}$, see (77b).

In order to gain some intuition, let us again consider the simplest case of $N = 1$. The minimization of (103) is once again straightforward and the minimum is found at

$$\vec{q}_{iT} \;=\; \not{\vec{p}}_T \,\frac{\tilde{M}_i}{\not{\mathbf{M}}_1}, \tag{104}$$

$$q_{iz} \;=\; \mathbf{q}_{1z}\,\frac{\tilde{M}_i}{\not{\mathbf{M}}_1}, \tag{105}$$

with an arbitrary choice of $\mathbf{q}_{1z}$. This leads to

$$M^2_{1\top}(\not{\mathbf{M}}_1) \equiv \left(\sqrt{\mathbf{M}_1^2+\mathbf{p}_{1T}^2}+\sqrt{\not{\mathbf{M}}_1^2+\not{p}_T^2}\right)^2 - u_T^2\,. \tag{106}$$

Comparing (106) to (97), we see that

$$M_{1\top} = M_1. \tag{107}$$

This is in fact a special case of the more general mathematical identity

$$M_{N\top} = M_N, \tag{108}$$

for which a proof is provided in the appendix — see equation (B49). This identity reveals that "transverse" quantities do not necessarily "forget" about relative longitudinal momenta. In particular, (108) teaches us that whenever the composite particles are formed *before* the transverse projection, the information about the relative longitudinal momenta is retained, and the result is the same as if everything was done in 1+3 dimensions throughout. As a result, $M_{N\top}$ automatically inherits all the advantages and disadvantages of its 1+3 cousin M_N.

3. *The late partitioned, $\top$-projected invariant mass: $M_{\top N}$*

This is the first example of an "early projected", "late partitioned" variable. We follow the procedure of the previous subsection VII C 2, only this time we switch the order of the operations, and we first $\top$-project the momentum of each individual particle on the transverse plane, before forming composite particles. The transverse invariant mass of each composite parent is then given by

$$\mathcal{M}^2_{\top a}(\mathbf{p}^\alpha_{\top a},\mathbf{q}^\alpha_{\top a},\tilde{\mu}_a) \;\equiv\; (\mathbf{p}_{\top a}+\mathbf{q}_{\top a})^2 \tag{109a}$$

$$\equiv (\mathbf{e}_{\top a}+\tilde{\mathbf{e}}_{\top a})^2 - (\vec{\mathbf{p}}_{aT}+\vec{\mathbf{q}}_{aT})^2, \tag{109b}$$

with $\vec{\mathbf{p}}_{aT}$ and $\vec{\mathbf{q}}_{aT}$ still given by (54) and (58), while the composite transverse energies $\mathbf{e}_{\top a}$ and $\tilde{\mathbf{e}}_{\top a}$ are given by (69) and (78), correspondingly. Notice that these expressions do not contain the longitudinal momenta p_{iz} and q_{iz}. This is in contrast to the "early partitioned" case represented by (68) and (77), where the longitudinal momenta appear explicitly. The comparison between (68) and (77) on the one hand, and (69) and (78) on the other, nicely illustrates the main point of Section V — that by adding the momenta *before* the projection, one retains sensitivity to the relative longitudinal momenta. Conversely, when the operations are performed in reverse order and the transverse projection is done first, the longitudinal momenta completely drop out of the game.

Now we are ready to apply the usual definition and obtain

$$M_{\top N}(\not{\mathbf{M}}) \equiv \min_{\sum \vec{q}_{iT}=\not{\vec{p}}_T}\left[\max_a\left[\mathcal{M}_{\top a}(\mathbf{p}^\alpha_{\top a},\mathbf{q}^\alpha_{\top a},\tilde{\mu}_a)\right]\right]. \tag{110}$$

Let us again investigate the simplest case of $N = 1$. With the help of the transverse momentum conservation constraint (51), eq. (110) reduces to

$$\begin{aligned} M^2_{\top 1} \;&=\; \min_{\sum \vec{q}_{iT}=\not{\vec{p}}_T}\left[\left(\sum_{i=1}^{N_\mathcal{V}} e_{i\top}+\sum_{i=1}^{N_\mathcal{I}}\tilde{e}_{i\top}\right)^2 - u_T^2\right] \\ &=\; \left(\sum_{i=1}^{N_\mathcal{V}} e_{i\top}+\min_{\sum \vec{q}_{iT}=\not{\vec{p}}_T}\left[\sum_{i=1}^{N_\mathcal{I}}\tilde{e}_{i\top}\right]\right)^2 - u_T^2. \end{aligned}$$

The minimum is once again found at (104) and we get

$$M^2_{\top 1}(\not{\mathbf{M}}_1) = \left(\sum_{i=1}^{N_\mathcal{V}}\sqrt{M_i^2+\vec{p}_{iT}^{\,2}}+\sqrt{\not{\mathbf{M}}_1^2+\not{p}_T^2}\right)^2 - u_T^2. \tag{111}$$

As expected, this result differs from (106), although the two formulas follow a similar pattern. The difference is only in the term corresponding to the visible sector, where the transverse energy of the composite visible particle is computed differently, compare (68a) and (69).

An interesting result emerges if we consider the further simplification that all visible particles are massless, i.e. $M_i = 0, \forall\, i$. This, in fact, is a very good approximation for the leptons and quarks/gluons of the SM, whose masses can be safely neglected. Setting $M_i = 0$ in (111) and using (83), we get

$$\lim_{M_i\to 0} M^2_{\top 1}(\not{\mathbf{M}}_1) = \left(h_T+\sqrt{\not{\mathbf{M}}_1^2+\not{p}_T^2}\right)^2 - u_T^2. \tag{112}$$

This result is quite interesting. It allows us to reinterpret the usual h_T variable in terms of a bona fide invariant mass variable like $M_{\top 1}$, properly accounting for the effects of upstream visible momentum u_T *and* the total mass $\not{\mathbf{M}}_1$ of the invisible particles present in the event. We shall return to this point in the next Section IX.

Another interesting result follows from eq. (112) in the special case when we set $\not{\mathbf{M}}_1 = 0$. Using (84), we get

$$\lim_{M_i\to 0} M^2_{\top 1}(\not{\mathbf{M}}_1 = 0) = \left(h_T+\not{p}_T\right)^2 - u_T^2 = m^2_{\text{eff}} - u_T^2, \tag{113}$$

providing a connection between the "effective mass" m_{eff} and $M_{\top 1}(0)$.

4. The late partitioned, $\vee$-projected mass: $M_{\vee N}$

This is the second example of an "early projected" variable, only this time we use the speed-preserving $\vee$ projection described in Section III B. Correspondingly, the individual visible (invisible) particles will be characterized by their 3-speeds V_i ($\tilde{V}_i$) instead of their masses M_i ($\tilde{M}_i$) and so we remind the reader of the notation introduced in (49) and (50).

The 1+2 momentum vectors of the individual particles after the $\vee$ projection are obtained from (17)

$$p^\alpha_{i\vee} \;\equiv\; (e_{i\vee}, \vec{p}_{i\vee}) = \left(\frac{p_{iT}}{V_i}, \vec{p}_{iT}\right), \tag{114}$$

$$q^\alpha_{i\vee} \;\equiv\; (\tilde{e}_{i\vee}, \vec{q}_{i\vee}) = \left(\frac{q_{iT}}{\tilde{V}_i}, \vec{q}_{iT}\right). \tag{115}$$

Then we form composite particles with $\vee$ projected 1+2 momenta $\mathbf{p}^\alpha_{\vee a}$ and $\mathbf{q}^\alpha_{\vee a}$ given by (66) and (75), respectively.

The transverse parent masses are now formed in terms of $\mathbf{p}^\alpha_{\vee a}$ and $\mathbf{q}^\alpha_{\vee a}$ as follows

$$\mathcal{M}^2_{\vee a}(\mathbf{p}^\alpha_{\vee a}, \mathbf{q}^\alpha_{\vee a}, \tilde{v}_a) \;\equiv\; (\mathbf{p}_{\vee a} + \mathbf{q}_{\vee a})^2 \tag{116a}$$

$$\equiv (\mathbf{e}_{\vee a} + \tilde{\mathbf{e}}_{\vee a})^2 - (\vec{\mathbf{p}}_{aT} + \vec{\mathbf{q}}_{aT})^2, \tag{116b}$$

where the transverse energies $\mathbf{e}_{\vee a}$ and $\tilde{\mathbf{e}}_{\vee a}$ are specified by (71) and (80) and the transverse momenta $\vec{\mathbf{p}}_{aT}$ and $\vec{\mathbf{q}}_{aT}$ are given by (54) and (58).

This is a convenient place to introduce another two small pieces of notation.[17] Firstly we will need to define a "maximum invisible velocity parameter" $\not{\mathbb{V}}_a$ for each parent $\mathbb{P}_a$ according to

$$\not{\mathbb{V}}_a \equiv \max_{i\in\mathcal{I}_a}\left[\tilde{V}_i\right]. \tag{117}$$

Then we would like to denote by $\not{\mathbb{V}}$ the set of all the above velocity parameters, i.e.

$$\not{\mathbb{V}} = \left\{\not{\mathbb{V}}_a \mid a \in \mathcal{P}\right\}. \tag{118}$$

Now we are in a position to state (see proof in Section IX A) that the only dependence of the "late partitioned", $\vee$-projected mass variable $M_{\vee N}$ on the velocity parameters of the invisible particles is through $\not{\mathbb{V}}$, i.e.:

$$M_{\vee N}(\not{\mathbb{V}}) \equiv \min_{\sum \vec{q}_{iT} = \vec{\not{p}}_T}\left[\max_a\left[\mathcal{M}_{\vee a}(\mathbf{p}^\alpha_{\vee a}, \mathbf{q}^\alpha_{\vee a}, \tilde{v}_a)\right]\right]. \tag{119}$$

Once again, it is instructive to consider the special case of $N = 1$. With the help of (51), eq. (119) becomes

$$\begin{aligned} M^2_{\vee 1}(\not{\mathbb{V}}_1) \;&=\; \min_{\sum \vec{q}_{iT} = \vec{\not{p}}_T}\left[\left(\sum_{i=1}^{N_\mathcal{V}} e_{i\vee} + \sum_{i=1}^{N_\mathcal{I}} \tilde{e}_{i\vee}\right)^2 - u_T^2\right] \\ &=\; \left(\sum_{i=1}^{N_\mathcal{V}} e_{i\vee} + \min_{\sum \vec{q}_{iT} = \vec{\not{p}}_T}\left[\sum_{i=1}^{N_\mathcal{I}} \tilde{e}_{i\vee}\right]\right)^2 - u_T^2 \\ &=\; \left(\sum_{i=1}^{N_\mathcal{V}} e_{i\vee} + \min_{\sum \vec{q}_{iT} = \vec{\not{p}}_T}\left[\sum_{i=1}^{N_\mathcal{I}} \frac{q_{iT}}{\tilde{V}_i}\right]\right)^2 - u_T^2. \end{aligned}$$

The minimization selects the invisible particle with the largest speed, whose transverse momentum becomes $\vec{\not{p}}_T$, while all other invisible particles have $q_{iT} = 0$. This configuration leads to the final answer

$$M^2_{\vee 1}(\not{\mathbb{V}}_1) = \left(\sum_{i=1}^{N_\mathcal{V}} \frac{p_{iT}}{V_i} + \frac{\not{p}_T}{\not{\mathbb{V}}_1}\right)^2 - u_T^2. \tag{120}$$

When we make the approximation that all visible particles are massless ($V_i = 1$), we again obtain a relation to h_T:

$$\lim_{V_i\to 1} M^2_{\vee 1}(\not{\mathbb{V}}_1) = \left(h_T + \frac{\not{p}_T}{\not{\mathbb{V}}_1}\right)^2 - u_T^2, \tag{121}$$

which is the analogue of (112) for the case of $\vee$ transverse projections. But note that unlike (112), here the unknown parameter characterizing the invisible sector is the maximum *speed* parameter $\not{\mathbb{V}}_1$ instead of the summed-invisible-mass parameter $\not{\mathbb{M}}_1$.

Finally, if in addition we also assume that all invisible particles are massless as well, then

$$\tilde{V}_i = 1,\ \forall i \quad\Longrightarrow\quad \not{\mathbb{V}}_1 = 1,$$

so that

$$\lim_{V_i\to 1} M^2_{\vee 1}(\not{\mathbb{V}}_1 = 1) = \left(h_T + \not{p}_T\right)^2 - u_T^2 = m^2_{\text{eff}} - u_T^2, \tag{122}$$

which is the analogue of (113). The fact that (113) and (122) are the same should not come as a surprise: recall from Sec. IV B that the two transverse projections $\top$ and $\vee$ are equivalent in the massless limit.

5. The early partitioned, $\vee$-projected mass: $M_{N\vee}$

Here we follow a procedure analogous to that of Sec. VII C 2, where the composite momenta $\mathbf{P}^\mu_a$ and $\mathbf{Q}^\mu_a$ are first formed in 1+3 dimensions, before being projected on the transverse plane, only this time we use the $\vee$ projection for this purpose:

$$\sum_{i\in\mathcal{V}_a} P^\mu_i \;\longrightarrow\; \mathbf{P}^\mu_a \;\xrightarrow{\vee}\; \mathbf{p}^\alpha_{a\vee} = (\mathbf{e}_{a\vee}, \vec{\mathbf{p}}_{a\vee}), \tag{123}$$

$$\sum_{i\in\mathcal{I}_a} Q^\mu_i \;\longrightarrow\; \mathbf{Q}^\mu_a \;\xrightarrow{\vee}\; \mathbf{q}^\alpha_{a\vee} = (\tilde{\mathbf{e}}_{a\vee}, \vec{\mathbf{q}}_{a\vee}). \tag{124}$$

[17] Contrast with the definition of $\not{\mathbb{M}}_a$ in equation (94) and the definition of $\not{\mathbb{M}}$ in equation (93).

The transverse parent masses are now formed in terms of $\mathbf{p}^\alpha_{a\vee}$ and $\mathbf{q}^\alpha_{a\vee}$ as usual

$$\mathcal{M}^2_{a\vee}(\mathbf{p}^\alpha_{a\vee},\mathbf{q}^\alpha_{a\vee},\tilde{v}_a) \equiv (\mathbf{p}_{a\vee}+\mathbf{q}_{a\vee})^2 \tag{125a}$$

$$\equiv (\mathbf{e}_{a\vee}+\tilde{\mathbf{e}}_{a\vee})^2-(\vec{\mathbf{p}}_{aT}+\vec{\mathbf{q}}_{aT})^2 \tag{125b}$$

Here the composite transverse momenta $\vec{\mathbf{p}}_{a\vee}$ and $\vec{\mathbf{q}}_{a\vee}$ are still given by (54) and (58), while the transverse energies $\mathbf{e}_{a\vee}$ and $\tilde{\mathbf{e}}_{a\vee}$ are given by (70) and (79), correspondingly.

Then the early-partitioned, $\vee$-projected variable is defined as usual:

$$M_{N\vee}(\not{\mathbf{V}}) \equiv \min_{\sum\vec{q}_{iT}=\vec{\not{p}}_T}\left[\max_a\left[\mathcal{M}_{a\vee}(\mathbf{p}^\alpha_{a\vee},\mathbf{q}^\alpha_{a\vee},\tilde{v}_a)\right]\right]. \tag{126}$$

Once again, let us specify this to the case of $N=1$. Using (51), we get

$$M_{1\vee} = \min_{\sum\vec{q}_{iT}=\vec{\not{p}}_T}\left[(\mathbf{e}_{1\vee}+\tilde{\mathbf{e}}_{1\vee})^2-u_T^2\right] \tag{127a}$$

$$= \left(\mathbf{e}_{1\vee}+\min_{\sum\vec{q}_{iT}=\vec{\not{p}}_T}[\tilde{\mathbf{e}}_{1\vee}]\right)^2-u_T^2. \tag{127b}$$

The minimization is performed over the $3N_\mathcal{I}$ variables $\vec{q}_{iT}$ and q_{iz}, $i=1,2,\ldots,N_\mathcal{I}$ and the result is

$$M_{1\vee}(\not{\mathbf{V}}_1) = \left(\mathbf{e}_{1\vee}+\frac{\not{p}_T}{\not{\mathbf{V}}_1}\right)^2-u_T^2 \tag{128a}$$

$$= \left(\frac{\mathbf{p}_{1T}}{\sqrt{\mathbf{p}^2_{1T}+\mathbf{p}^2_{1z}}}\,\mathbf{E}_1+\frac{\not{p}_T}{\not{\mathbf{V}}_1}\right)^2-u_T^2, \tag{128b}$$

which is similar, but not equivalent to (120).

6. *The late partitioned, $\circ$-projected mass: $M_{\circ N}$*

Here we follow the procedure of Secs. VII C 3 and VII C 4, only this time we use the $\circ$ transverse projection from Sec. III C. One first forms the 1+2 momenta of the individual particles

$$p^\alpha_{i\circ} \equiv (e_{i\circ},\vec{p}_{i\circ}) = (p_{iT},\vec{p}_{iT}), \tag{129}$$

$$q^\alpha_{i\circ} \equiv (\tilde{e}_{i\circ},\vec{q}_{i\circ}) = (q_{iT},\vec{q}_{iT}), \tag{130}$$

then the composite momenta

$$\mathbf{p}^\alpha_{\circ a} \equiv (\mathbf{e}_{\circ a},\vec{\mathbf{p}}_{\circ a}) = \left(\sum_{i\in\mathcal{V}_a}p_{iT},\sum_{i\in\mathcal{V}_a}\vec{p}_{iT}\right), \tag{131}$$

$$\mathbf{q}^\alpha_{\circ a} \equiv (\tilde{\mathbf{e}}_{\circ a},\vec{\mathbf{q}}_{\circ a}) = \left(\sum_{i\in\mathcal{V}_a}q_{iT},\sum_{i\in\mathcal{V}_a}\vec{q}_{iT}\right). \tag{132}$$

The transverse parent masses are now formed in terms of $\mathbf{p}^\alpha_{\circ a}$ and $\mathbf{q}^\alpha_{\circ a}$ as usual:

$$\mathcal{M}^2_{\circ a}(\mathbf{p}^\alpha_{\circ a},\mathbf{q}^\alpha_{\circ a}) \equiv (\mathbf{p}_{\circ a}+\mathbf{q}_{\circ a})^2 \tag{133a}$$

$$\equiv (\mathbf{e}_{\circ a}+\tilde{\mathbf{e}}_{\circ a})^2-(\vec{\mathbf{p}}_{aT}+\vec{\mathbf{q}}_{aT})^2, \tag{133b}$$

and the "late partitioned", $\circ$-projected mass variable $M_{\circ N}$ is defined as before:

$$M_{\circ N} \equiv \min_{\sum\vec{q}_{iT}=\vec{\not{p}}_T}\left[\max_a\left[\mathcal{M}_{\circ a}(\mathbf{p}^\alpha_{\circ a},\mathbf{q}^\alpha_{\circ a})\right]\right]. \tag{134}$$

Notice that the $M_{\circ N}$ variables do not depend on any unknown parameters related to the invisible sector (i.e. we need no "$\not{\mathbf{O}}$" where previously we needed an $\not{\mathbf{M}}$ or a $\not{\mathbf{V}}$) and so can be uniquely computed in terms of the measured momenta of the visible particles and the missing transverse momentum alone.

Specializing (134) to the simplest case of $N=1$, we get

$$M^2_{\circ 1} = \min_{\sum\vec{q}_{iT}=\vec{\not{p}}_T}\left[\left(\sum_{i=1}^{N_\mathcal{V}}e_{i\circ}+\sum_{i=1}^{N_\mathcal{I}}\tilde{e}_{i\circ}\right)^2-u_T^2\right]$$

$$= \left(\sum_{i=1}^{N_\mathcal{V}}e_{i\circ}+\min_{\sum\vec{q}_{iT}=\vec{\not{p}}_T}\left[\sum_{i=1}^{N_\mathcal{I}}\tilde{e}_{i\circ}\right]\right)^2-u_T^2$$

$$= \left(\sum_{i=1}^{N_\mathcal{V}}p_{iT}+\min_{\sum\vec{q}_{iT}=\vec{\not{p}}_T}\left[\sum_{i=1}^{N_\mathcal{I}}q_{iT}\right]\right)^2-u_T^2.$$

The minimization over the $2N_\mathcal{I}$ variables $\vec{q}_{iT}$ is straightforward and we obtain several equivalent expressions for the answer

$$M^2_{\circ 1} = \left(\sum_{i=1}^{N_\mathcal{V}}p_{iT}+\not{p}_T\right)^2-u_T^2 \tag{135a}$$

$$= \left(h_T+\not{p}_T\right)^2-u_T^2, \tag{135b}$$

$$= m^2_{\text{eff}}-u_T^2. \tag{135c}$$

showing the close connection between $M_{\circ 1}$ and the usual h_T and m_{eff} variables. We see that in the absence of any upstream visible momentum ($\vec{u}_T=0$), the variable $M_{\circ 1}$ itself is nothing but the effective mass m_{eff}. However, these two variables differ if (as is typically the case) the event also has some nonzero upstream momentum u_T. The importance of the result (135c) is that it teaches us how to properly account for the presence of UVM in such cases: u_T should be *subtracted in quadratures* from m_{eff} in order to obtain the proper invariant mass variable (in this case $M_{\circ 1}$). Furthermore, it also reveals the physical meaning of the widely used m_{eff} variable (see also Sec. X B below): it is the minimum allowed transverse mass constructed out of "$\circ$"-projected momenta, for a semi-invisibly decaying parent, whenever that parent is produced exclusively with $u_T=0$ (i.e. with no additional upstream momentum in the event).

7. *The early partitioned, $\circ$-projected mass: $M_{N\circ}$*

Finally, we discuss the early partitioned, $\circ$-projected version $M_{N\circ}$, where the composite momenta are first

formed in 1+3 dimensions, then transversified via the "$\circ$" projection:

$$\sum_{i\in\mathcal{V}_a} P_i^\mu \;\longrightarrow\; \mathbf{P}_a^\mu \;\overset{\circ}{\longrightarrow}\; \mathbf{p}_{a\circ}^\alpha = (\mathbf{e}_{a\circ}, \vec{\mathbf{p}}_{a\circ})\,, \quad (136)$$

$$\sum_{i\in\mathcal{I}_a} Q_i^\mu \;\longrightarrow\; \mathbf{Q}_a^\mu \;\overset{\circ}{\longrightarrow}\; \mathbf{q}_{a\circ}^\alpha = (\tilde{\mathbf{e}}_{a\circ}, \vec{\mathbf{q}}_{a\circ})\,, \quad (137)$$

where in light of (72) and (81)

$$\mathbf{p}_{a\circ}^\alpha \;=\; (\mathbf{e}_{a\circ}, \vec{\mathbf{p}}_{a\circ}) = \left(\left|\sum_{i\in\mathcal{V}_a} \vec{p}_{iT}\right|, \sum_{i\in\mathcal{V}_a} \vec{p}_{iT}\right)\,, \quad (138)$$

$$\mathbf{q}_{a\circ}^\alpha \;=\; (\tilde{\mathbf{e}}_{a\circ}, \vec{\mathbf{q}}_{a\circ}) = \left(\left|\sum_{i\in\mathcal{I}_a} \vec{q}_{iT}\right|, \sum_{i\in\mathcal{I}_a} \vec{q}_{iT}\right)\,. \quad (139)$$

These (1+2) composite momenta are now used to form the corresponding transverse parent masses

$$\mathcal{M}^2_{a\circ}(\mathbf{p}^\alpha_{a\circ}, \mathbf{q}^\alpha_{a\circ}) \;\equiv\; (\mathbf{p}_{a\circ} + \mathbf{q}_{a\circ})^2 \quad (140a)$$

$$\equiv (\mathbf{e}_{a\circ} + \tilde{\mathbf{e}}_{a\circ})^2 - (\vec{\mathbf{p}}_{aT} + \vec{\mathbf{q}}_{aT})^2. \quad (140b)$$

Now the "early partitioned", $\circ$-projected mass variable $M_{N\circ}$ is defined as before:

$$M_{N\circ} \equiv \min_{\sum \vec{q}_{iT} = \vec{\not{p}}_T} \left[\max_a \left[\mathcal{M}_{a\circ}(\mathbf{p}^\alpha_{a\circ}, \mathbf{q}^\alpha_{a\circ})\right]\right]. \quad (141)$$

Just like its cousin $M_{\circ N}$ defined in (134), $M_{N\circ}$ does not depend on any unknown parameters like $\mathbb{M}_a$ or $\mathbb{V}_a$.

Specifying (141) to the simplest case of $N = 1$, we get

$$\begin{aligned} M^2_{1\circ} \;&=\; \min_{\sum \vec{q}_{iT} = \vec{\not{p}}_T} \left[(\mathbf{e}_{1\circ} + \tilde{\mathbf{e}}_{1\circ})^2 - u_T^2\right] \\ &=\; \left(\mathbf{e}_{1\circ} + \min_{\sum \vec{q}_{iT} = \vec{\not{p}}_T} [\tilde{\mathbf{e}}_{1\circ}]\right)^2 - u_T^2 \\ &=\; \left(\mathbf{e}_{1\circ} + \min_{\sum \vec{q}_{iT} = \vec{\not{p}}_T} \left[\left|\sum_{i=1}^{N_\mathcal{I}} \vec{q}_{iT}\right|\right]\right)^2 - u_T^2. \end{aligned}$$

The minimization over the $2N_\mathcal{I}$ variables $\vec{q}_{iT}$ gives

$$M^2_{1\circ} \;=\; \left(\left|\sum_{i=1}^{N_\mathcal{V}} \vec{p}_{iT}\right| + \not{p}_T\right)^2 - u_T^2 \quad (142a)$$

$$=\; \left(|\vec{\not{p}}_T + \vec{u}_T| + \not{p}_T\right)^2 - u_T^2 \quad (142b)$$

$$=\; 2\left(\vec{\not{p}}_T \cdot (\vec{\not{p}}_T + \vec{u}_T) + \not{p}_T\,|\vec{\not{p}}_T + \vec{u}_T|\right), \quad (142c)$$

providing a connection between our $M_{1\circ}$ variable and the usual missing transverse momentum $\not{p}_T$. In order to see the physical meaning of $\not{p}_T$, let us take the "no upstream momentum" limit $u_T \to 0$ in (142b) or (142c), resulting in

$$\lim_{u_T\to 0} M^2_{1\circ} = 4\not{p}_T^2. \quad (143)$$

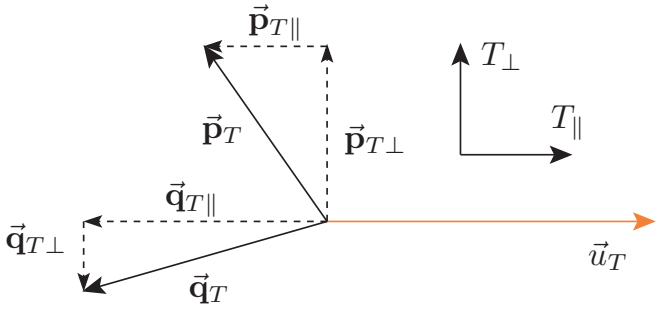

FIG. 9: Transverse vector decomposition onto the direction $T_\|$ specified by the UVM transverse momentum vector $\vec{u}_T$ and the direction $T_\perp$ orthogonal to it [25, 34]. All vectors shown are in the plane perpendicular to the beam axis.

One can thus interpret the variable $2\not{p}_T$ (and not just the $\not{p}_T$!) as the minimum allowed "$\circ$"-projected transverse mass of a semi-invisibly decaying parent, whenever the parent is produced exclusively with $u_T = 0$, i.e. with no additional upstream momentum in the event. However, in situations when the parent is produced inclusively, with $u_T \neq 0$, the relevant variable to consider would be $M_{1\circ}$ as given by (142b) or (142c), which properly accounts for the u_T effect (see also Sec. X A below).

VIII. ADDITIONALLY PROJECTED VARIABLES

A. Momentum decompositions with respect to $\vec{u}_T$

An additional level of projection *within* the plane transverse to the beam has been shown to be useful in certain circumstances [25, 34]. To orient such projections we note that the total transverse momentum $\vec{u}_T$ of the UVM category breaks the rotational symmetry of the transverse plane and selects two preferred directions $T_\|$ (along $\vec{u}_T$) and $T_\perp$ (transverse to $\vec{u}_T$), as shown in Fig. 9. Having projected the 1+3 momentum vectors onto the transverse plane as in Fig. 2, one may then additionally project the resulting 1+2 transverse momentum vectors onto these special directions, as illustrated in Fig. 9. The corresponding momentum components resulting from such "double transverse" projections will carry a "double transverse" index: "$T\perp$" for components along $T_\perp$ and "$T\|$" for components along $T_\|$ (see Fig. 9). For example, the $\vec{\not{p}}_T$ vector can be decomposed into a $T_\perp$ component $\vec{\not{p}}_{T\perp}$

$$\vec{\not{p}}_{T\perp} = \frac{1}{u_T^2}\vec{u}_T \times \left(\vec{p}_T \times \vec{u}_T\right), \quad (144)$$

and a $T_\|$ component $\vec{\not{p}}_{T\|}$

$$\vec{\not{p}}_{T\|} = \vec{\not{p}}_T - \vec{\not{p}}_{T\perp} = \frac{1}{u_T^2}\left(\vec{\not{p}}_T \cdot \vec{u}_T\right)\vec{u}_T\,. \quad (145)$$

By definition, the upstream transverse momentum vector $\vec{u}_T$ has only a $T_\parallel$ component, i.e.

$$\vec{u}_{T\parallel} \;=\; \vec{u}_T, \tag{146}$$

$$\vec{u}_{T\perp} \;=\; 0. \tag{147}$$

In view of (51) and (147), momentum conservation in the $T_\perp$ direction reads

$$\sum_{i=1}^{N_\mathcal{I}} \vec{q}_{iT\perp} = \vec{\not{p}}_{T\perp} = -\sum_{i=1}^{N_\mathcal{V}} \vec{p}_{iT\perp}\,. \tag{148}$$

It is precisely the absence of a $\vec{u}_{T\perp}$ term in this equation which allows one to derive exact analytical formulas for the $T_\perp$ doubly projected variables defined next in Section VIII B.

B. Doubly-projected mass bound variables

1. Homogeneously-doubly-projected mass bound variables

For our purposes, the additional projections in Fig. 9 allow us to extend the original set of mass-bound variables from Table III by considering the "doubly projected" variables shown in Table IV[18]. The benefit of such additionally projected varibles has been noted and discussed in [25, 34]. For example, the shapes and the kinematic endpoints of the distributions of $T_\perp$-projected variables can be independent of the value of u_T. Therefore, such distributions can be constructed from the whole event sample, without any loss in statistics due to a specific u_T selection. Furthermore, the relation (147) leads to significant simplifications in the analytical treatment of $T_\perp$ doubly projected variables. For example, for singly projected variables, the case of $N = 2$ is untractable by analytical means, and (apart from some special cases [51]) has to be treated numerically [38, 39]. In contrast, one can derive exact analytical formulas for calculating $N = 2$, $T_\perp$ doubly projected mass bound variables on an event-per-event basis, without any need for numerical minimizations [25, 34].

In general, the variables in Table IV are independent, with one exception:

$$M_{N\circ\perp} \equiv M_{\circ N\perp}. \tag{149}$$

Later on in Section XI B (see in particular Fig. 13(b)), we shall consider a specific example illustrating some of the homogeneously-doubly-projected variables from Table IV.

2. Heterogeneously-doubly-projected mass bound variables

Notice that in defining the mass bound variables in Table IV, we have chosen the second level of projection (along $T_\perp$) to be performed with *the same* type of transverse projection ("$\top$", "$\vee$" or "$\circ$") which was used to project into the transverse plane. Of course, this does not have to be the case — and by allowing for *different* types of transverse projections for T and for $T_\perp$, one would obtain 18 additional variables with "mixed" transverse projections. These heterogeneously-doubly-projected variables are listed in Table V, where the additional subindex on $\perp$ specifies the type of $T_\perp$ projection as being of the "$\top$", "$\vee$" or "$\circ$" type.[19] As usual, the sequence of indices in both Tables IV and V represents the order in which the operations are to be performed. For example, $M_{\circ N\perp_\top}$ means

- project all objects using the massless '$\circ$' projection, then
- partition and agglomerate into N parents, then
- project into the direction perpendendicular to $\vec{u}_T$ using the mass-preserving '$\top$' projection, then, as ever,
- minimize over all values of the unknown momenta that satisfy the constraints.

Interestingly, most of the "$\perp_\circ$" heterogeneously-doubly-projected variables turn out to be related to each other and to the corresponding homogeneously-doubly-projected variables from Table IV. For example:

$$M_{N\circ\perp} \;\equiv\; M_{N\circ\perp_\circ} = M_{N\top\perp_\circ} = M_{N\circ\perp_\top} = M_{\top N\perp_\circ}, \tag{150}$$

$$M_{\circ\perp N} \;\equiv\; M_{\circ\perp_\circ N} = M_{\circ\perp_\top N} = M_{\top\perp_\circ N}, \tag{151}$$

$$M_{\circ N\perp} \;\equiv\; M_{\circ N\perp_\circ} = M_{N\top\perp_\circ} = M_{N\circ\perp_\top} = M_{\top N\perp_\circ}, \tag{152}$$

where the last line (152) follows from (149) and (150). The one remaining variable $M_{\circ N\perp_\top}$ is rather similar to $M_{\top N\perp_\top} \equiv M_{\top N\perp}$, since the difference between them may arise only due to nonzero masses ($M_i \neq 0$) of the individual visible particles.

IX. PROPERTIES OF THE MASS-BOUND VARIABLES

We should stress that proliferating the number of kinematic variables in the literature is certainly not among the goals of this paper — on the contrary, we emphasise that these variables are different implementations of

[18] To save space, Table IV lists only $T_\perp$ projected variables. An analogous set of $T_\parallel$ projected variables is obtained by replacing the $T_\perp$ projection in Table IV with a $T_\parallel$ projection.

[19] Another set of 18 additional variables can be trivially obtained from Table V by considering a $T_\parallel$ type of projection at the second level instead.

Type of variables	Operations				Notation
	First	Second	Third	Fourth	
Early partitioned	Partitioning	$T = \top$ projection	$\perp = \top$ projection on $T_\perp$	Minimization	$M_{N\top\perp}$
doubly projected	Partitioning	$T = \vee$ projection	$\perp = \vee$ projection on $T_\perp$	Minimization	$M_{N\vee\perp}$
$M_{NT\perp}$	Partitioning	$T = \circ$ projection	$\perp = \circ$ projection on $T_\perp$	Minimization	$M_{N\circ\perp}$
Late partitioned,	$T = \top$ projection	$\perp = \top$ projection on $T_\perp$	Partitioning	Minimization	$M_{\top\perp N}$
doubly projected	$T = \vee$ projection	$\perp = \vee$ projection on $T_\perp$	Partitioning	Minimization	$M_{\vee\perp N}$
$M_{T\perp N}$	$T = \circ$ projection	$\perp = \circ$ projection on $T_\perp$	Partitioning	Minimization	$M_{\circ\perp N}$
In-between partitioned,	$T = \top$ projection	Partitioning	$\perp = \top$ projection on $T_\perp$	Minimization	$M_{\top N\perp}$
doubly projected	$T = \vee$ projection	Partitioning	$\perp = \vee$ projection on $T_\perp$	Minimization	$M_{\vee N\perp}$
$M_{TN\perp}$	$T = \circ$ projection	Partitioning	$\perp = \circ$ projection on $T_\perp$	Minimization	$M_{\circ N\perp}$

TABLE IV: An extended version of Table III, containing the additional variables found by including the option of a $T_\perp$ projection shown in Fig. 9. An analogous set of variables is obtained by considering a $T_\parallel$ projection instead.

Early partition	Hedged partition	Late partition
$M_{N\top\perp_\vee}$, $M_{N\top\perp_\circ}$	$M_{\top N\perp_\vee}$, $M_{\top N\perp_\circ}$	$M_{\top\perp_\vee N}$, $M_{\top\perp_\circ N}$
$M_{N\vee\perp_\top}$, $M_{N\vee\perp_\circ}$	$M_{\vee N\perp_\top}$, $M_{\vee N\perp_\circ}$	$M_{\vee\perp_\top N}$, $M_{\vee\perp_\circ N}$
$M_{N\circ\perp_\top}$, $M_{N\circ\perp_\vee}$	$M_{\circ N\perp_\top}$, $M_{\circ N\perp_\vee}$	$M_{\circ\perp_\top N}$, $M_{\circ\perp_\vee N}$

TABLE V: The 18 additional heterogeneously-doubly-projected transverse mass variables for each N, where the additional subindex on $\perp$ specifies the type of $T_\perp$ projection as being of the "$\top$", "$\vee$" or "$\circ$" type. As was the case in Tables III and IV, "partition" implies the combined operation of partitioning the objects and agglomerating them by summation into composite objects.

the general principle described in Section VII A. What's more we will soon begin to reveal further connections of these variables to each other (in Section IX C) and to existing proposals (in Section X). But before we proceed, perhaps now is a good time to summarize what we have accomplished so far.

The previous discussion has hopefully convinced the reader that, once the decision on the targeted event topology (Fig. 5) is made, the choice of relevant invariant mass variables is straightforward and rather unambiguous. Following the general recipe outlined in Section VII B, one is able to overcome the two main obstacles in any analysis involving missing momentum:

- *The fact that the momenta of the invisible particles are unknown.* To construct a bound, this problem is solved by performing a *minimization* over all possible values of the invisible momenta, consistent with the measured $\not{\vec{p}}_T$. The minimization fixes the values of the invisible momenta (e.g. as in (95) and (96)) and from that point on, one works with fully specified kinematics in the event. Of course, the momenta found in the process of minimization, are not equal to the *actual* momenta of the invisible particles in the event, although in some cases they can be close, see [19].

- *The fact that the total number and the masses of the invisible particles are unknown.* This problem is also resolved through the minimization — as we have seen in the explicit $N = 1$ examples discussed in Section VII C, the mass bound resulting from the minimization turned out to be a function which depends only on a set of N summed mass parameters (93) or a set of N 3-speeds (118), and is insensitive to the number of invisible particles or to the fine structure of the individual masses $\tilde{M}_a$ or 3-speeds $\tilde{V}_a$ connected to parent $\mathbb{P}_a$. We set out a general proof for general N in Section IX A below.

It should be recognized that for any practical application, there is no need to consider every one of the variables in Tables III and IV, since some will be better suited than others to the particular task at hand.

For example, we have seen that the "$\vee$"-projected quantities assume knowledge of the $\vec{p}_T$ and the speed of the particle, but leave the mass and p_z undetermined. This means that all of the "$\vee$"-projected variables in Tables III and IV should be considered appropriate only for experimental situations in which the p_T and speed of the particles are known, but nothing is known about their masses or longitudinal momenta. Such situations may exist – for example if p_T can be determined from the particles' bending radii in a solenoidal magnetic field and speed can be inferred from time-of-flight information or from the characteristic angle of any emitted Čerenkov

radiation. However such cases are the exception, rather than the norm in current experiments.[20] In what follows we shall therefore give greater attention to the remaining three classes of variables: "unprojected", "$\top$"-projected, and "$\circ$"-projected.

We shall denote a generic mass-bound variable as $M_{\mathcal{F}}$, where the composite index $\mathcal{F}$ is made up from (any number of) objects taken from the set $\{N, \top, \vee, \circ\}$. There are seven such possibilities[21]:

$$\mathcal{F} \in \{N, N\top, \top N, N\circ, \circ N, N\vee, \vee N\}. \tag{153}$$

For later convenience, we also introduce the generic notation $\mathcal{F}_\top$ for the $\top$-projected variables:

$$\mathcal{F}_\top \in \{N\top, \top N\} \tag{154}$$

$\mathcal{F}_\vee$ for their two "$\vee$"-projected counterparts:

$$\mathcal{F}_\vee \in \{N\vee, \vee N\}, \tag{155}$$

and $\mathcal{F}_\circ$ for the two "$\circ$"-projected equivalents:

$$\mathcal{F}_\circ \in \{N\circ, \circ N\}. \tag{156}$$

The large multiplicity is partially due to the different possible ways to transversify the energy-momenta of the composite daughter particles whose masses $\mathbf{M}_a$ and $\tilde{\mathbf{M}}_a$ are typically nonzero. First, one can choose whether or not to project, and then those projections can be of type $\top$, $\vee$ or $\circ$ (see Section III). In addition, as emphasized in Section V, the operations of partitioning into composite particles and transversifying do not commute, so that in general we obtain non-equivalent variables simply by switching the order of those operations. As illustrated in Tables III-V, we use the ordering of indices on each variable (from left to right) to indicate the order of the corresponding operations. For example $M_{2\top}$ means that we add the 1+3 vectors first to form two composite visible daughter particles and transversely project later, while $M_{\top 2}$ implies the opposite — make a $\top$ transverse projection before forming the composite daughter particles.

A. Dependence of mass bounds $M_{\mathcal{F}}$ on $\not{\mathbf{M}}$, $\not{\mathbf{V}}$, etc.

We have stated that the dependence of the mass bound variables, $M_{\mathcal{F}}$, on parameters of the hypothesis is always confined to a set of N parameters contained within $\not{\mathbf{M}}$ or $\not{\mathbf{V}}$ etc. We have not yet proved this statement for general values of N, or indicated whether we can demonstrate this to be true for other classes of projection not already discussed. All we have proved, so far, are the following statements, which are specific to $N = 1$ and consider at most one projection:

- That M_1 depends only on $\not{\mathbf{M}} = \{\not{\mathbf{M}}_1\}$ (see (97))
- That $M_{1\top}$ depends only on $\not{\mathbf{M}} = \{\not{\mathbf{M}}_1\}$ (see (106))
- That $M_{\top 1}$ depends only on $\not{\mathbf{M}} = \{\not{\mathbf{M}}_1\}$ (see (111))
- That $M_{\vee 1}$ depends only on $\not{\mathbf{V}} = \{\not{\mathbf{V}}_1\}$ (see (120))
- That $M_{1\vee}$ depends only on $\not{\mathbf{V}} = \{\not{\mathbf{V}}_1\}$ (see (128))
- That $M_{1\circ}$ depends on no hypothesis parameters[22],
- That $M_{\circ 1}$ depends on no hypothesis parameters.

We now seek to generalize the proofs of the above to all other values of N, in a manner that does not make specific requirements on $\mathcal{F}$. Specifically, we would like to prove that:

$$\text{"}M_{\mathcal{F}} \text{ depends only on } \not{\mathbf{S}}_{\mathcal{F}} = \left\{\not{\mathbf{S}}_1^{\mathcal{F}}, \ldots, \not{\mathbf{S}}_N^{\mathcal{F}}\right\}\text{"} \tag{157}$$

where $\not{\mathbf{S}}_{\mathcal{F}}$ is a set of N parameters, of which there is one ($\not{\mathbf{S}}_a^{\mathcal{F}}$) for each parent $\mathbb{P}_a$, and where the nature of $\not{\mathbf{S}}_a^{\mathcal{F}}$ depends on the type of projection in $\mathcal{F}$ (which may be arbitrary), and on a, but not on the number N of parents in total. In particular, we have already seen to expect $\not{\mathbf{S}}_{\mathcal{F}} = \{\not{\mathbf{M}}_1\}$ when $\mathcal{F} \in \{1, 1\top, \top 1\}$, and to expect $\not{\mathbf{S}}_{\mathcal{F}} = \{\not{\mathbf{V}}_1\}$ when $\mathcal{F} \in \{1\vee, \vee 1\}$, and we are now seeking to generalize these to results like "$\not{\mathbf{S}}_{\mathcal{F}} = \{\not{\mathbf{M}}_1, \not{\mathbf{M}}_2\}$ when $\mathcal{F} \in \{2, 2\top, \top 2\}$" or "$\not{\mathbf{S}}_{\mathcal{F}} = \{\not{\mathbf{V}}_1, \not{\mathbf{V}}_2, \not{\mathbf{V}}_3\}$ when $\mathcal{F} \in \{3\vee, \vee 3\}$"etc.

What we will actually succeed in proving is the marginally less general statement that:

$$\boxed{\begin{array}{l} M_{\mathcal{F}} \text{ depends only on } \not{\mathbf{S}}_{\mathcal{F}} = \left\{\not{\mathbf{S}}_1^{\mathcal{F}}, \not{\mathbf{S}}_2^{\mathcal{F}}, \ldots, \not{\mathbf{S}}_n^{\mathcal{F}}\right\}, \\ \text{when } N = n, \text{ provided that } M_{\mathcal{F}} \text{ depends only} \\ \text{on } \not{\mathbf{S}}_{\mathcal{F}} = \left\{\not{\mathbf{S}}_1^{\mathcal{F}}\right\} \text{ when } N = 1 \end{array}} \tag{158}$$

which reminds us that the generality of the desired (but unattainable) result (157) is constrained (for any particular projection $\mathcal{F}$) by the need to prove the result for the $N = 1$ case. In other words, though the proof of (158) found below will be valid for any projection, the desired result (157) will only be true for projections that experience simplification in the $N = 1$ case.

---

[20] Even in those cases, it is usual that $\vec{p}_T$ is determined from a measured track, so one would also expect to be able to reconstruct the polar angle θ, from that track, which would permit p_z and hence the full 1+3-vector of the particle to be determined. There would be no need to then restrict oneself to the subset of that information held by the corresponding "$\vee$"-projected quantities.

[21] The number of possibilities increases to 17 if one allows a second level of projection, as discussed in Section VIII.

[22] Note that depending on "no hypothesis parameters" is a special case of depending on a very dull set of parameters $\not{\mathbf{O}} = \{\not{\mathbf{O}}_1\}$ which contain no information.

The proof of (158) is astonishingly simple. Consider an arbitrary mass-bound variable

$$M_{\mathcal{F}} \equiv \min_{\sum \vec{q}_{iT} = \vec{p}_T} \left[\max_a \left[\mathcal{M}_a \left(\mathcal{S}_a^{\mathcal{F}} \right) \right] \right]. \tag{159}$$

where each parent mass $\mathcal{M}_a$ depends on a corresponding set of invisible parameters $S_i^{\mathcal{F}}$

$$\mathcal{S}_a^{\mathcal{F}} \equiv \left\{ S_i^{\mathcal{F}} \mid i \in \mathcal{I}_a \right\} \tag{160}$$

(compare this to the analogous relations (48) and (50)). Let us now perform the minimization in (159) in two steps. In the first step, for each parent, we hold the sum of the invisible daughters' momenta constant, and minimize over the internal partition of invisible momentum between those daughters:

$$M_{\mathcal{F}} = \min_{\sum \vec{q}_{aT} = \vec{p}_T} \left[\min_{\sum\limits_{i \in \mathcal{I}_a} \vec{q}_{iT} = \vec{q}_{aT}} \left[\max_a \left[\mathcal{M}_a \left(\mathcal{S}_a^{\mathcal{F}} \right) \right] \right] \right]. \tag{161}$$

Since the internal partitions over the invisible momenta are done independently for each parent, (161) can be equivalently rewritten as

$$M_{\mathcal{F}} = \min_{\sum \vec{q}_{aT} = \vec{p}_T} \left[\max_a \left[\min_{\sum\limits_{i \in \mathcal{I}_a} \vec{q}_{iT} = \vec{q}_{aT}} \left[\mathcal{M}_a \left(\mathcal{S}_a^{\mathcal{F}} \right) \right] \right] \right]. \tag{162}$$

Now we use the assertion that for $N = 1$ (i.e. for any individual parent) the minimization over internal partitions yields a function of a single parameter $\not{\mathbf{S}}_a^{\mathcal{F}}$ as opposed to the whole set of parameters $\mathcal{S}_a^{\mathcal{F}}$:

$$\min_{\sum\limits_{i \in \mathcal{I}_a} \vec{q}_{iT} = \vec{q}_{aT}} \left[\mathcal{M}_a \left(\mathcal{S}_a^{\mathcal{F}} \right) \right] \equiv \mathcal{M}_a \left(\not{\mathbf{S}}_a^{\mathcal{F}} \right). \tag{163}$$

Substituting (163) into (162), we obtain the desired result

$$M_{\mathcal{F}} \left(\not{\mathbf{S}}_{\mathcal{F}} \right) = \min_{\sum \vec{q}_{aT} = \vec{p}_T} \left[\max_a \left[\mathcal{M}_a \left(\not{\mathbf{S}}_a^{\mathcal{F}} \right) \right] \right], \tag{164}$$

which makes it obvious that $M_{\mathcal{F}}$ can only be a function of the set of parameters $\not{\mathbf{S}}_{\mathcal{F}} = \left\{ \not{\mathbf{S}}_1^{\mathcal{F}}, \not{\mathbf{S}}_2^{\mathcal{F}}, \ldots, \not{\mathbf{S}}_n^{\mathcal{F}} \right\}$.

B. Parental masses and upper kinematic endpoints

By construction, the mass-bound variables of Table III are designed to provide an event-by-event lower bound on the true invariant mass $M_{\mathcal{P}}^{\max}$ of the heaviest parent

$$M_{\mathcal{P}}^{\max} \equiv \max_{a \in \mathcal{P}} \left[M_a \right]. \tag{165}$$

Such bounding properties are contingent on us being able to make appropriate choices when analyzing the events. We initially restrict our discussion of the bound to the case where the set of momentum configurations $\mathcal{A}$ permitted under our assumptions is equal to the set of momenta $\mathcal{E}$ sampled by nature. We observe that to ensure $\mathcal{A} = \mathcal{E}$ we must (a) correctly reconstruct the event topology (the number of parents, the number and types of daughters, and the association of daughters to parents) and (b) employ the true values of the parameters $\not{\mathbf{S}}$ used in constructing of any $m_{\mathcal{F}}$ variable — i.e. $\not{\mathbf{V}}^{\text{true}}$ for $\mathcal{F}_\vee$, and $\not{\mathbf{M}}^{\text{true}}$ for $\mathcal{F}_\top$ or $\mathcal{F}_N$. Under $\mathcal{A} = \mathcal{E}$ conditions all $M_{\mathcal{F}}$ variables are designed to return values smaller than the mass of the heaviest parent

$$M_{\mathcal{F}} \leq M_{\mathcal{P}}^{\max} \quad (\mathcal{A} = \mathcal{E}). \tag{166}$$

From eq. (166), it follows directly that, if we were to consider the differential distribution of the same variable $M_{\mathcal{F}}$ over all events, the upper kinematic endpoint $M_{\mathcal{F}}^{\max}$ of this distribution also satisfies

$$M_{\mathcal{F}}^{\max} \equiv \max_{\text{all events}} \left[M_{\mathcal{F}} \right] \leq M_{\mathcal{P}}^{\max} \quad (\mathcal{A} = \mathcal{E}), \tag{167}$$

where we make explicit the requirement that the true values of the $\not{\mathbf{S}}$ parameters are used.

There remains the important question as to the circumstances under which the inequalities in (166) and (167) are saturated – i.e. the conditions for which a measurement of the $M_{\mathcal{F}}$ kinematic endpoint will provide a determination of (rather than simply a lower bound on) the largest parent mass $M_{\mathcal{P}}^{\max}$.

We observe that when $\mathcal{A} = \mathcal{E}$: (i) that for any selected event $\epsilon \in \mathcal{E}$ the minimization picks out some non-empty subset of momenta K_ϵ that satisfy the global minimum; (ii) that $K_\epsilon \subset \mathcal{A}$; (iii) we may define for convenience $\mathcal{M}_{\mathcal{F}}^{\max}(c) = \max_{a \in \mathcal{P}} \left[\mathcal{M}_{a\mathcal{F}}(c) \right]$ for any configuration $c \in \mathcal{A}$; (iv) that since we minimize over any unknown momentum components, the value of $\mathcal{M}_{\mathcal{F}}^{\max}$ evaluated for some minimum configuration $k \in K$ cannot exceed the value that would be obtained elsewhere in $\mathcal{A}$ (and therefore in $\mathcal{E}$); (v) that projections $T \in \{\top, \vee, \circ\}$ do not increase the invariant mass (33); (vi) that $M_{\mathcal{F}}$ can therefore not exceed the largest parent's invariant mass since

$$M_{\mathcal{F}} \equiv \mathcal{M}_{\mathcal{F}}^{\max}(K_\epsilon) \leq \mathcal{M}_{\mathcal{F}}^{\max}(\epsilon) \leq M_{\mathcal{P}}^{\max} \quad (\mathcal{A} = \mathcal{E}). \tag{168}$$

The necessary and sufficient condition for saturation of (167) is therefore that there exist some event ϵ for which two inequalities in (168) simultaneously become equalities.

Given that $M_{\mathcal{F}}^{\max} \equiv \max_{a \in \mathcal{P}} \left[\mathcal{M}_a(K_\epsilon) \right]$ it follows from (168) that a necessary condition for saturation is that

$$\exists (\epsilon \in \mathcal{E}, a \in \mathcal{P}) \left[\mathcal{M}_a(\epsilon) = M_{\mathcal{P}}^{\max} \right]. \tag{169}$$

There are cases for which (169) is *not* satisfied, and for which the inequality in (167) must therefore remain unsaturated. For example, if the decay of some $\mathbb{P}_a$ proceeds exclusively via an intermediate on-shell resonance, then the set of physically observed momenta $\mathcal{E}$ is further

restricted to a subset of $\mathcal{A}$ which need not contain any event satisfying (169).[23]

The "$\circ$" projection discards all previous information about the mass of the 1+3 vector being projected, and so calculation of $M_{\circ N}$ will return the same value that would be obtained if one were to set both $M_i = 0\,(\forall i \in \mathcal{V})$ and $\not{\mathbf{M}}_a = 0\,(\forall a \in \mathcal{P})$. If all events contain massive invisibles (or indeed massive visibles) in all daughter sets — as would be the case for models like R-parity conserving supersymmetry and UED — then (169) cannot be true for any $\mathcal{M}_{\circ a}$ and so $M_{\circ N}$ can only bound from below, rather than determine, the mass of the heaviest parent. In Appendix B 3 we prove the event-by-event inequality $M_{N\circ} \leq M_{\circ N}$, so any bound that is unsaturated for $M_{\circ N}$ must also be unsaturated for $M_{N\circ}$.

Compared to the $\mathcal{F}_\circ$ variables, the $\mathcal{F}_\top$ and $\mathcal{F} = N$ variables are subject to less stringent conditions for saturation, because they retain mass information during the process of (absence of) projection. Some of the necessary conditions can be inferred from the results of in App. B 1 and B 2. As an example of these less-stringent conditions, if $|\mathcal{I}_a| \neq 0\,\forall a$ then a necessary condition for saturation will include the existence of events $\epsilon \in \mathcal{E}$ with vanishing relative rapidity between $\mathbf{P}_a^\mu$ and $\mathbf{Q}_a^\mu$ for some $a \in \mathcal{P}$.

As discussed in Section VII C, some of the widely used collider variables like h_T, $m_{\rm eff}$ and $\not{p}_T$ belong to the $\mathcal{F}_\circ$ class of mass bound variables, and as such can generally only place a lower bound on the parent mass if nature produces heavy invisibles. In order to really measure the mass scale of the new particles when massive invisible particles are pervasive, one must work with variables which retain the dependence on the missing mass parameters and therefore belong to either the M_N or the $M_{\mathcal{F}_\top}$ class of mass-bound variables. Other than the full 1+3 dim. invariant mass, other common examples of such variables include $\sqrt{\hat{s}}_{\rm min}$ [50] (discussed below in Sec. X C) the transverse mass m_T in the form [53] that accounts for the mass of all daughters (shown in Sec. X D) and the 'stransverse mass' m_{T2} [7] (Sec. X H).

Conditions for saturation have been most thoroughly explored for the $N = 1$ and $N = 2$ cases. One might reasonably ask whether the bounds can still be saturated for larger N. Saturation is indeed possible for $N > 2$ topologies with similar kinematical configurations as for the $N = 1$ or $N = 2$ cases. With increasing numbers of invisible final state particles $N_{\mathcal{I}}$ (which may or may not be a result of increasing N) the density of states near the boundary is reduced. Saturation is still possible for any $N_{\mathcal{I}}$, but as the multiplicity of invisible particles increases, the number of events close to the endpoints becomes small, and experimentally determining the true end-point becomes increasingly challenging.

Before concluding this subsection, we note that the $M_{\mathcal{F}}$ variables are still useful even when the true values of the mass $\not{\mathbf{M}}$ (or speed $\not{\mathbf{V}}$) parameters are not known. The most conservative procedure in these situations of uncertainty is to minimize $M_{\mathcal{F}}$ over the complete physically relevant range of any unknown parameter. This leads to $\not{\mathbf{M}}_a \to 0$ for M_N, $M_{\mathcal{F}_\top}$, and to $\not{\mathbf{V}}_a \to 1$ for $M_{\mathcal{F}_\circ}$. The resulting, conservative, $M_{\mathcal{F}}$ variables still provide lower bounds on the mass of the heaviest parent – though those bounds will generally not be saturated.

A more sophisticated treatment is also possible. For example if the "true" value of the summed-mass parameter set $\not{\mathbf{M}}^{\rm true}$ for the calculation of $M_N(\not{\mathbf{M}})$ or $M_{\mathcal{F}_\top}(\not{\mathbf{M}})$ were not known — then one could still view the set of endpoint measurements for all possible values of $\not{\mathbf{M}}_a$ as one constraint among the $N+1$ unknowns $M_{\mathcal{P}}^{\rm max}$ and $\not{\mathbf{M}}_a$, $(a = 1, 2, \ldots, N)$. Not only is this valuable information on its own, the derived functional relationship $M_{\mathcal{P}}^{\rm max}(\not{\mathbf{M}})$ is, in addition, often sufficient for determining the *individual* mass parameters $\not{\mathbf{M}}_a$. The function $M_{\mathcal{P}}^{\rm max}(\not{\mathbf{M}})$, when viewed as an N-dimensional hyper-surface in the $(N+1)$-dimensional space spanned by $M_{\mathcal{P}}^{\rm max}$ and $\not{\mathbf{M}}$, exhibits certain ridge or crease features, which commonly originate from the point marking the set of true values of $\not{\mathbf{M}}^{\rm true}$ [24, 26]. (The one-dimensional version of this phenomenon was originally discussed in [10–13, 21] and is known as the m_{T2} "kink". Also see [54] for algebraic singularity in relation to the kink.)

C. Relations among the mass-bound variables

Some of the variables in Table III are related to each other, either in general[24]

$$M_{N\top}(\not{\mathbf{M}}) = M_N(\not{\mathbf{M}}), \tag{170}$$

$$M_{N\circ\perp} = M_{\circ N\perp} = M_{N\top\perp_\circ} = M_{N\circ\perp_\top} = M_{\top N\perp_\circ}, \tag{171}$$

$$M_{\circ\perp N} = M_{\circ\perp_\top N} = M_{\top\perp_\circ N}, \tag{172}$$

or under some special circumstances, e.g. massless particles:

$$M_{\top N}(\{\not{\mathbf{M}}_a = 0\}, \{M_i = 0\}) \;=\; M_{\circ N}, \tag{173}$$

$$M_{\top N\perp}(\{\not{\mathbf{M}}_a = 0\}, \{M_i = 0\}) \;=\; M_{\circ N\perp_\top}. \tag{174}$$

Given such exact identities like (170), the reader may wonder why we even bothered to introduce separately variables like $M_{N\top}$ and M_N. In our view, such redundancy is a virtue, since it offers deeper intuitive understanding of these kinematic variables, and allows one

[23] The conditions for saturation for an $N = 1$ example containing on-shell intermediate particles were explored in [52].

[24] Previously in (107) we already encountered the $N = 1$ version of eq. (170). The general proof for arbitrary N is provided in the appendix in equation (B49).

to think about the same fundamental quantity in different contexts, e.g. in (1+3)-dimensions or in (1+2)-dimensions.

We additionally find (see proof terminating in (B53) in Appendix B 3) that the mass-bounds from Table III obey a hierarchy:

$$M_N = M_{N\top} \geq M_{\top N} \geq M_{\circ N} \geq M_{N\circ}. \tag{175}$$

Similarly, the doubly-projected mass-bounds from Table IV obey the hierarchy

$$M_{N\top\perp} \geq M_{\top N\perp} \geq M_{\top\perp N} \geq M_{\circ\perp N} \geq M_{\circ N\perp} = M_{N\circ\perp}. \tag{176}$$

From these hierarchies, it becomes apparent that the $\top$-projected, late-partitioned variables bear a cost associated with the insensitivity to the longitudinal momenta. By dropping this information we necessarily weaken the bound relative to the early-partitioned versions. Interestingly enough, the order of projection and partition has the opposite effect with the $\circ$-projection, since both longitudinal and transverse information is contained in the masses of the agglomerates, and hence by throwing away the masses at a later stage, we in fact throw away maximal information and are forced to produce the worst possible bound!

X. CONNECTIONS TO OTHER VARIABLES IN THE LITERATURE

The existing literature is abundant with a number of (transverse) invariant mass variables which were suggested (at various times and for a variety of reasons) for the study of missing momentum event topologies (see [1] for a recent review). At the same time, the mass-bound variables which we defined earlier in Table III, were meant to be very general, since they target the rather generic event topology of Fig. 5, and are intended to have as few hidden assumptions as possible. It follows that we should be able to correlate the most useful mass-scale variables in the literature to one of our mass-bound variables from Table III.[25] The purpose of this subsection is to demonstrate that this is indeed the case.

A. Missing transverse momentum $\not{p}_T$

The defining feature of any "missing particle" event is the presence of missing momentum (more precisely, missing transverse momentum) $\not{p}_T$. This is due to the production and escape of a certain number of "invisible" particles, which are either sterile, or very weakly interacting, so that they are not seen in the detector. The $\not{p}_T$ distribution[26] is perhaps the most widely studied distribution in relation to new physics searches, especially in models with WIMP dark matter candidates like supersymmetry, UED and so on. Eq. (142) allows us to correlate the $\not{p}_T$ variable to our $M_{1\circ}$ variable as

$$M_{1\circ} \overset{u_T\to 0}{\longrightarrow} 2\not{p}_T. \tag{177}$$

We see that as $M_{1\circ}$ is defined more and more inclusively, it eventually becomes equal to twice the missing transverse momentum. Thus in the case of a singly produced parent, eq. (177) allows us to interpret the usual $\not{p}_T$ variable (more precisely, the variable $2\not{p}_T$) as a suitably constructed (in the $M_{1\circ}$ sense) transverse invariant mass of the parent (see also the discussion at the end of Sec. VII C 7). In accordance with (166), in the $u_T \to 0$ limit the upper kinematic endpoint of the $2\not{p}_T$ distribution gives a *lower bound* on the parent mass in events interpreted as single-parent ($N = 1$) production.

B. Effective mass $m_{\rm eff}$

The "effective mass" variable defined in (84) can be also directly related to one of our variables, namely the late-partitioned, "$\circ$"-projected variable $M_{\circ 1}$ discussed in Sec. VII C 6. The previously derived eq. (135c) reads

$$M_{\circ 1}^2 = m_{\rm eff}^2 - u_T^2. \tag{178}$$

Therefore, we obtain the correspondence

$$M_{\circ 1} \overset{u_T\to 0}{\longrightarrow} m_{\rm eff}, \tag{179}$$

allowing us to interpret $m_{\rm eff}$ as a suitably constructed (in the $M_{\circ 1}$ sense) transverse invariant mass of a singly produced, semi-invisibly decaying parent (see also the discussion at the end of Sec. VII C 6).

The comparison between eq. (177) and eq. (179) rather nicely illustrates the main point of Sec. V: that when it comes to transverse projections and forming composite particles, performing these operations in different order yields different results. In the case at hand, when forming composite particles *before* the "$\circ$" transverse projection, one obtains $2\not{p}_T$, while by forming composite particles *after* the "$\circ$" transverse projection, one obtains $m_{\rm eff}$.

[25] A corollary from this statement is that invariant mass variables which make similar sorts of assumptions but do not fit into the classification of Table III, are often both poorly motivated and sub-optimal.

[26] The missing transverse momentum is often labelled called "missing transverse energy" and labelled $\not{E}_T$ or $E_{\rm T}^{\rm miss}$ in experimental papers. As previously discussed we prefer to recognize the imporant distinction between energy and momentum, so use the symbol $\not{p}_T$.

C. Florida $\sqrt{\hat{s}}_{\rm min}$ and $\sqrt{\hat{s}}^{\rm (sub)}_{\rm min}$ variables

As already seen in eq. (100), in the special case of $N = 1$, the unprojected mass-bound variable $M_N = M_1$ is nothing but the subsystem $\sqrt{\hat{s}}^{\rm (sub)}_{\rm min}$ variable from [46]:

$$M_1(\not{\mathbf{M}}_1) \equiv \sqrt{\hat{s}}^{\rm (sub)}_{\rm min}(\not{\mathbf{M}}_1) = \left[\left(\sqrt{\mathbf{M}_1^2+\mathbf{p}_{1T}^2}+\sqrt{\not{\mathbf{M}}_1^2+\not{p}_T^2}\right)^2-u_T^2\right]^{1/2}. \quad (180)$$

Restricting to events with vanishing upstream momentum ($u_T = 0$), one gets the inclusive $\sqrt{\hat{s}}_{\rm min}$ variable from [50]:

$$\lim_{u_T\to 0} M_1(\not{\mathbf{M}}_1) \equiv \sqrt{\hat{s}}_{\rm min}(\not{\mathbf{M}}_1) = \sqrt{\mathbf{M}_1^2+\mathbf{p}_{1T}^2}+\sqrt{\not{\mathbf{M}}_1^2+\not{p}_T^2}. \quad (181)$$

As advocated in Refs. [46, 50], practical applications of $M_1(\not{\mathbf{M}}_1)$ need not be limited to events in which the *actual* number of parents was $N = 1$. The work of [46, 50] showed that in events with $N = 2$, the peak in the $M_1(\not{\mathbf{M}}_1)$ distribution is correlated with the parent mass threshold $\sum_{a=1,2} M_{\mathbb{P}_a}$, even if the two parent particles $\mathbb{P}_1$ and $\mathbb{P}_2$ are different.

Note that the mathematical identity (170) also allows us to write

$$\sqrt{\hat{s}}^{\rm (sub)}_{\rm min}(\not{\mathbf{M}}_1) = M_{1\top}(\not{\mathbf{M}}_1), \quad (182)$$

$$\sqrt{\hat{s}}_{\rm min}(\not{\mathbf{M}}_1) = \lim_{u_T\to 0} M_{1\top}(\not{\mathbf{M}}_1), \quad (183)$$

relating the $\sqrt{\hat{s}}^{\rm (sub)}_{\rm min}$ and $\sqrt{\hat{s}}_{\rm min}$ variables to the *transverse* invariant mass quantity $M_{1\top}$, which is simply the total transverse invariant mass in the event (after accounting for the potential presence of any transverse upstream momentum u_T).

D. Transverse mass

Perhaps the most popular variable which specifically targets a semi-invisibly decaying resonance, is the transverse mass $m_{Te\nu}$, which, as suggested by our notation, was first applied in searches for a leptonically decaying W-boson (see, e.g. [6]):

$$m_{Te\nu}^2 \equiv (e_{eT}+e_{\nu T})^2-(\vec{p}_{eT}+\vec{q}_{\nu T})^2 \quad (184a)$$

$$\approx 2(|\vec{p}_{eT}||\vec{q}_{\nu T}|-\vec{p}_{eT}\cdot\vec{q}_{\nu T}), \quad (184b)$$

where $\vec{p}_{eT}$ ($\vec{q}_{eT}$) is the transverse momentum of the lepton (neutrino), and in the second line one makes the approximation that the lepton and the neutrino are approximately massless. Assuming that the W boson is produced singly, with zero recoil (i.e. $u_T = 0$ in our language), the neutrino transverse momentum $\vec{q}_{eT}$ can be identified with the measured missing transverse momentum $\vec{\not{p}}_T$, and (184b) becomes

$$m_{Te\nu}^2 \approx 2p_{eT}\not{p}_T(1-\cos\phi_{e\nu}), \quad (185)$$

where $\phi_{e\nu}$ is the measured opening angle between the transverse vectors $\vec{p}_{eT}$ and $\vec{\not{p}}_T$.

In this simple example of a W-decay, the two daughter particles are massless, but the same idea can be easily generalized to the case of massive daughters as [53]

$$m_{Te\nu}^2(M_e,M_\nu) = M_e^2+M_\nu^2+2(e_{eT}e_{\nu T}-\vec{p}_{eT}\cdot\vec{q}_{\nu T}), \quad (186)$$

where M_e and M_ν are the electron and neutrino masses, respectively, and

$$e_{eT} \equiv \sqrt{M_e^2+\vec{p}_{eT}^{\,2}}, \quad (187a)$$

$$e_{\nu T} \equiv \sqrt{M_\nu^2+\vec{q}_{\nu T}^{\,2}}. \quad (187b)$$

Now let us obtain these results with our formalism. In general, we have a singly produced ($N = 1$) parent resonance, which decays to a single ($N_{\mathcal{V}} = 1$) visible daughter particle and a single ($N_{\mathcal{I}} = 1$) invisible daughter particle. Since there is only one particle in each daughter set, $\mathcal{V}_1 = \{e\}$ and $\mathcal{I}_1 = \{\nu\}$, there is no need to form composite particles, so the order of the operations becomes unimportant. However, if the daughter particles are massive, the two different types of transverse projections give two different versions of the transverse mass variable:

$$M_{1\top}^2(M_\nu) = M_{\top 1}^2(M_\nu) = \left(\sqrt{M_e^2+\vec{p}_{eT}^{\,2}}+\sqrt{M_\nu^2+\vec{q}_{\nu T}^{\,2}}\right)^2-\vec{u}_T^{\,2}, \quad (188)$$

$$M_{1\circ}^2 = M_{\circ 1}^2 = (p_{eT}+q_{\nu T})^2-\vec{u}_T^{\,2}. \quad (189)$$

Here eq. (188) follows simply from the general formulas (106) or (111) with the identifications $\mathbf{M}_1 = M_1 = M_e$, $\vec{\mathbf{p}}_{1T} = \vec{p}_{1T} = \vec{p}_{eT}$, $\not{\mathbf{M}}_1 = M_\nu$ and $\vec{\not{p}}_T = \vec{q}_{\nu T}$. Similarly, eq. (189) is obtained from either (135a) or (142a).

Now it is trivial to eliminate $\vec{u}_T$ using the transverse momentum relation (51) $\vec{u}_T = -\vec{p}_{eT}-\vec{q}_{\nu T}$, and show that eq. (188) is equivalent to (186):

$$M_{1\top}(M_\nu) = M_{\top 1}(M_\nu) = m_{Te\nu}(M_e,M_\nu) \quad (190)$$

while eq. (189) is equivalent to (184b):

$$M_{1\circ} = M_{\circ 1} = m_{Te\nu}(M_e = 0, M_\nu = 0). \quad (191)$$

E. Cluster transverse mass variables

Next we consider a couple of more complicated single resonance processes. The first example is $h \to ZZ \to$

$e^+e^-\nu\bar{\nu}$ where each Z-boson is assumed to be on-shell, one decaying invisibly, the other decaying visibly to a pair of leptons. For this particular scenario, Ref. [55] suggested the cluster transverse mass variable

$$M^2_{T,ZZ} = \left(E_{T,Z_1} + E_{T,Z_2}\right)^2 - \left(\vec{p}_{T,Z_1} + \vec{p}_{T,Z_2}\right)^2 \quad (192a)$$

$$= \left(\sqrt{M_Z^2 + p^2_{T,e^+e^-}} + \sqrt{M_Z^2 + \not{p}_T^2}\right)^2 - \left(\vec{p}_{T,e^+e^-} + \vec{\not{p}}_T\right)^2 . \quad (192b)$$

Note that the 1+3-dimenstional invariant mass of the visible and that of the invisible systems have each been constrained to be equal to M_Z^2. These two constraints reflect the on-shell hypothesis we have chosen to assume for each of the two Z bosons.[27]

Once again, we can obtain this variable from our $M_{1\top}$ or $M_{\top 1}$. In analogy to the case of $m_{T e\nu}$, we have a single parent resonance, the Higgs boson h, decaying to a single massive visible daughter, the first Z boson and a single massive invisible particle, the other Z-boson. This corresponds to $N = 1$, $N_{\mathcal{V}} = 1$, $N_{\mathcal{I}} = 1$, $\mathcal{V}_1 = \{Z \to e^+e^-\}$ and $\mathcal{I}_1 = \{Z \to \nu\bar{\nu}\}$. Correspondingly, we identify $\mathbf{M}_1 = M_1 = M_Z$, $\vec{\mathbf{p}}_{1T} = \vec{p}_{1T} = \vec{p}_{T,e^+e^-}$, $\not{\mathbf{M}}_1 = M_Z$ and $\vec{\not{p}}_T = \vec{q}_{T,\nu\bar{\nu}}$. Then (106) and (111) simply give

$$M^2_{1\top}(M_Z) = M^2_{\top 1}(M_Z) = \left(\sqrt{M_Z^2 + \vec{p}^{\,2}_{T,e^+e^-}} + \sqrt{M_Z^2 + \not{p}_T}\right)^2 - \vec{u}_T^{\,2}, \quad (193)$$

which is equivalent to (192b) in light of the momentum conservation relation $\vec{u}_T = -\vec{p}_{T,e^+e^-} - \vec{\not{p}}_T$. Thus we have proved

$$M_{1\top}(M_Z) = M_{\top 1}(M_Z) = M_{T,ZZ}. \quad (194)$$

Another interesting example is provided by the process $h \to W^+W^- \to e^+e^-\nu\bar{\nu}$, for which Ref. [56] proposed the cluster transverse mass variable

$$M^2_{C,WW} \equiv \left(\sqrt{M^2_{e^+e^-} + \vec{p}^{\,2}_{T,e^+e^-}} + \not{p}_T\right)^2 - \left(\vec{p}_{T,e^+e^-} + \vec{\not{p}}_T\right)^2 . \quad (195)$$

Here the two leptons are clustered together (even though they originate from different W-bosons, they have a common parent in h) and their total transverse momentum is $\vec{p}_{T,e^+e^-}$. The definition (195) is similar to (192b), the difference now being that the two leptons are not correlated, and their invariant mass does not have to be consistent with M_Z. In addition, the invisible mass parameter $\not{\mathbf{M}}_1$ is now set to zero (as opposed to M_Z), because the invisible particles (the two neutrinos) are massless.

The cluster variable (195) can be readily obtained from $M_{N\top}$ with the following interpretation: $N = 1$, $N_{\mathcal{V}} = 2$, $N_{\mathcal{I}} = 2$, $\mathcal{V}_1 = \{e^+, e^-\}$ and $\mathcal{I}_1 = \{\nu, \bar{\nu}\}$. Correspondingly, we identify $\mathbf{M}_1 = M_{e^+e^-}$, $\vec{\mathbf{p}}_{1T} = \vec{p}_{T,e^+e^-}$ and $\not{\mathbf{M}}_1 = 2M_\nu = 0$. Then the general formula (106) reduces to

$$M^2_{1\top}(0) = \left(\sqrt{M^2_{e^+e^-} + \vec{p}^{\,2}_{T,e^+e^-}} + \not{p}_T\right)^2 - \vec{u}_T^{\,2}, \quad (196)$$

which is the same as (195), so that

$$M_{1\top}(0) = M_{C,WW}. \quad (197)$$

Notice that in this example we are clustering two visible particles, and the order of operations becomes important. Therefore, here $M_{1\top}$ and $M_{\top 1}$ in general lead to distinct variables, unlike the case of (190) and (194).

F. The $m_T^{\rm true}$ transverse mass variable

Concerning the same $h \to W^+W^- \to e^+e^-\nu\bar{\nu}$ example, Ref. [57] advertized the variable (assuming massless neutrinos)

$$\left(m_T^{\rm true}\right)^2 \equiv M^2_{e^+e^-} + 2\Big(\not{p}_T\sqrt{M^2_{e^+e^-} + \vec{p}^{\,2}_{T,e^+e^-}} - \vec{\not{p}}_T \cdot \vec{p}_{T,e^+e^-}\Big), \quad (198)$$

which can be rewritten as

$$\begin{aligned}\left(m_T^{\rm true}\right)^2 &= M^2_{e^+e^-} + \vec{p}^{\,2}_{T,e^+e^-} + \vec{\not{p}}_T^{\,2} - \vec{p}^{\,2}_{T,e^+e^-} - \vec{\not{p}}_T^{\,2} \\ &\quad + 2\not{p}_T\sqrt{M^2_{e^+e^-} + \vec{p}^{\,2}_{T,e^+e^-}} - 2\vec{\not{p}}_T \cdot \vec{p}_{T,e^+e^-} \\ &= \left(\sqrt{M^2_{e^+e^-} + \vec{p}^{\,2}_{T,e^+e^-}} + \not{p}_T\right)^2 - \left(\vec{p}_{T,e^+e^-} + \vec{\not{p}}_T\right)^2 \\ &\equiv M^2_{C,WW}\,. \end{aligned} \quad (199)$$

From (197) and (199) it now follows that

$$M_{1\top}(0) = m_T^{\rm true}. \quad (200)$$

This connection in fact was the primary motivation for introducing the $m_T^{\rm true}$ variable in the first place [57].

G. The $m_{TZ'}^{reco}$ transverse mass variable

Our final single resonance example will be taken from a new physics scenario, namely a generic model with a new Z' gauge boson which decays to a SM Higgs boson h and a SM Z-boson as $Z' \to hZ \to b\bar{b}\nu\bar{\nu}$. For this particular topology, Ref. [58] considered the transverse mass variable

$$m_{TZ'}^{reco} \equiv \sqrt{M_h^2 + p_{Th}^2} + \sqrt{M_Z^2 + \not{p}_T^2}, \quad (201)$$

[27] We note that if one wishes to relax the assumption of an on-shell Z leading to the visible e^+e^- system one may do so by treating the electron and positron vectors as separate inputs to the visible system V. Similarly one may relax the assumption that the invisible system is the result of the decay of an on-shell Z by treating the neutrinos as independent invisible inputs.

where M_h (p_{Th}) is the measured invariant mass (transverse momentum) of the $b\bar{b}$ jet pair resulting from the decay $h \to b\bar{b}$.

In our language, the event topology $Z' \to hZ \to b\bar{b}\nu\bar{\nu}$ corresponds to a single parent, the Z' boson, thus $N = 1$. There is a single ($N_{\mathcal{V}} = 1$) visible daughter particle, which is the reconstructed Higgs boson: $\mathcal{V}_1 = \{h \to b\bar{b}\}$. There is also a single ($N_{\mathcal{I}} = 1$) invisible daughter particle, which is the invisibly decaying Z-boson: $\mathcal{I}_1 = \{Z \to \nu\bar{\nu}\}$. Thus we identify $\mathbf{M}_1 = M_1 = M_h$, $\vec{\mathbf{p}}_{1T} = \vec{p}_{1T} = \vec{p}_{Th}$ and $\not{\mathbf{M}}_1 = M_Z$.

Again, we have chosen to make assumptions about the 1+3-dimensional invariant masses of the visible and the invisible systems, requiring the former to be equal to M_h, and the latter to be equal to M_Z, reflecting our assumptions about the decay topology. As before it would be possible to independently relax either of both those assumptions by treating the $b\bar{b}$ and $\nu\bar{\nu}$ as independent inputs to the visible and invisible systems respectively.

If we retain the mass-shell constraint for both the h boson and the Z boson then (106) and (111) give

$$\begin{aligned} M^2_{1\top}(M_Z) &= M^2_{\top 1}(M_Z) \\ &= \left(\sqrt{M_h^2 + \vec{p}^{\,2}_{Th}} + \sqrt{M_Z^2 + \not{p}_T}\right)^2 - \vec{u}_T^{\,2}. \end{aligned} \tag{202}$$

Comparing to (201), we see that

$$\lim_{u_T \to 0} M_{1\top}(M_Z) = \lim_{u_T \to 0} M_{\top 1}(M_Z) = m^{reco}_{TZ'}. \tag{203}$$

Since $m^{reco}_{TZ'}$ was properly defined as a transverse mass variable, it is not surprising that it can be obtained as a special case of the mass-bounding variables shown in Table III. The importance of eq. (202) is that it shows the proper way to generalize $m^{reco}_{TZ'}$ to the case where the Z' is produced *inclusively*, with some non-vanishing UVM u_T in the event.

This concludes our discussion of singly produced resonances. We are hopeful that after all these examples, the reader is prepared to handle any assumed event topology, and will be able to construct the proper transverse invariant mass variable for the case at hand.

H. Cambridge m_{T2} variable

The variables considered in our previous examples referred either to the event as a whole (as in Secs. X A–X C) or to the production of a single resonance (as in Secs. X D–X G). We now move on to discussing variables intended to handle the production of more than one parent resonance ($N > 1$). Such cases are very common in new physics scenarios, especially if the new model contains a dark matter candidate, whose lifetime is protected by some discrete symmetry (typically a Z_2). In such models (e.g. supersymmetry, extra dimensions, little Higgs theories etc.) the main production mechanisms usually involve the *pair-production* of new particles, thus the case of $N = 2$ has received the most attention so far in the literature, although $N = 3, 4, \ldots$ cannot be ruled out, and in principle deserve attention as well.

A popular variable of this type is the Cambridge m_{T2} variable defined as [7]

$$m_{T2} \equiv \min_{\sum \vec{q}_{iT} = \vec{\not{p}}_T} \left[\max\left[\mathcal{M}_{1\top}, \mathcal{M}_{2\top}\right]\right] , \tag{204}$$

where $\mathcal{M}_{1\top}$ and $\mathcal{M}_{2\top}$ are the transverse masses of the two parent particles, and the minimization is done over all possible partitions of the transverse momenta of the invisible particles, consistent with the measured $\not{p}_T$.

We note that if there is only one visible particle belonging to each parent, we can immediately identify the Cambridge variable m_{T2} with both $M_{2\top}$ and $M_{\top 2}$ (and even with M_2 using (108)) since partitioning and projection commute for single particles. If, however, we intend to apply m_{T2} to events in which one or either parent has two or more physical daughters (e.g. when doing top quark mass measurements in the di-leptonic $t\bar{t} \to b\bar{b}l^+l^-\nu\bar{\nu}$ events) then $M_{2\top}$ will become inequivalent to $M_{\top 2}$ and we should decide which of these is the right thing to use. The answer to this question is subtle. The original m_{T2} paper [7] does not explicitly state how parent momenta should be constructed in the event that they have come from compound objects, so it is left up to users to decide which inputs to supply. It was certainly in the minds of the authors of [7] that users ought always to supply the maximal amount of trustworthy information to any analysis of any kind. In the context of m_{T2} this maxim would imply projecting only *after* combining the primary (1+3) momenta of any constituents of parents, provided that those constituents could be "trusted".[28] Only when defined in this manner (i.e. as $M_{2\top}$) can the maximum amount of information be squeezed from the variable in "clean" events. However, there can be benefits from using $M_{\top 2}$ (see for example Ref [9]) in high-multiplicity or inclusive situations in which the indiviual momenta making up each parent have dubiuos provenance or poorly measured longitudinal momenta. In such cases, one can benefit from using $M_{\top 2}$, even though its end point is less sharp for the signal, simply because it is less sensitive to longitudinal momenta and momenta at high rapidities.

One might ask which of the two m_{T2} choices — $M_{2\top}$ or $M_{\top 2}$ — is "better". Unfortunately, this question does not allow a "one size fits all" answer. Each of the two m_{T2} implementations has its unique advantages and disadvantages. The longitudinal correlations among the visible particles which are preserved by $M_{2\top}$ result in steeper, better defined endpoint structures — see Figs. 11(a) and

[28] This might include the case where an experiment that records dileptonic top-pairs with good (signed) b-tagging could allow the l^+ to be associated unambiguously with the b and the l^- to be associated with the $\bar{b}$.

12(a) below. On the other hand, $M_{\top 2}$ dampens the effects of any longitudinal momenta, which would be beneficial in circumstances where forward jet activity due to ISR may be a problem.

We hope that the current paper will serve as a reminder that studies using m_{T2} in which either parent is built from two or more reconstructed momenta should think carefully about the advantages and disadvantages of both approaches before choosing the option that is best for them. Both versions of m_{T2}, namely $M_{2\top}$ and $M_{\top 2}$, may prove to be useful, and it can be important to make the distinction between them.

In conclusion of this subsection, we highlight the analogy between m_{T2} and $\sqrt{\hat{s}}_{\rm min}$, two variables which are more closely related than one might think. We have shown that the (1+3) dimensional version of m_{T2}, together with the mathematical identity (170) implies

$$m_{T2}^{(1+3)}(\not{\mathbf{M}}) \equiv M_{2\top}(\not{\mathbf{M}}) = M_2(\not{\mathbf{M}}). \tag{205}$$

The second equality here emphasizes that, in spite of the transverse index "T", the m_{T2} variable is a bona fide 1+3 dimensional quantity. In other words, the apparent transverse projection in the definition (204) does not lead to any loss of useful information.[29] Of course, the same cannot be said about the (1+2)-dimensional version $m_{T2}^{(1+2)} \equiv M_{\top 2}$.

Now compare (205) to the analogous equation following from (180) and (182)

$$\sqrt{\hat{s}}_{\rm min}^{(\rm sub)}(\not{\mathbf{M}}_1) = M_{1\top}(\not{\mathbf{M}}_1) = M_1(\not{\mathbf{M}}_1). \tag{206}$$

We are reminded that $\sqrt{\hat{s}}_{\rm min}^{(\rm sub)}$ and $m_{T2}^{(1+3)}$ have essentially the same physical meaning: they both give a lower bound on a mass **in (1+3) dimensions** as a function of the corresponding invisible mass parameters. In the case of $\sqrt{\hat{s}}_{\rm min}^{(\rm sub)}$ that mass is the center of mass energy of the collision since it views the whole collision as a "single parent", while $m_{T2}^{(1+3)}$ hypothesizes that the collision was a $2 \to 2$ process, and therefore bounds the mass of the heavier of the two outgoing particles.

I. The doubly projected variables $m_{T2\perp}$ and $m_{T2\parallel}$

The doubly projected variables $m_{T2\perp}$ and $m_{T2\parallel}$ introduced in [25] are nothing but the one-dimensional analogues of the Cambridge variable m_{T2} (204), where one performs an additional projection on the directions $T_\perp$ and $T_\parallel$, correspondingly. As was the case for m_{T2} discussed above, the transverse projection should be interpreted in the $T = \top$ sense, and there are three possible versions of each variable, depending on the order of operations:

$$m_{T2\perp}(\not{\mathbf{M}}) = \begin{cases} M_{2\top\perp}(\not{\mathbf{M}}) & \text{in } 1+3 \text{ dims;} \\ M_{\top 2\perp}(\not{\mathbf{M}}) & \text{in } 1+2 \text{ dims;} \\ M_{\top\perp 2}(\not{\mathbf{M}}) & \text{in } 1+1 \text{ dims,} \end{cases} \tag{207}$$

and similarly for $m_{T2\parallel}(\not{\mathbf{M}})$. The example considered in [25] was inclusive chargino production, where each chargino parent decays to a visible lepton and an invisibly decaying sneutrino. In this case, each visible daughter partition $\mathcal{V}_a$ has a single massless visible particle (a lepton) and the distinction between the early partitioned version $m_{T2\perp}^{(1+3)} \equiv M_{2\top\perp}$, the in-between partitioned version $m_{T2\perp}^{(1+2)} \equiv M_{\top 2\perp}$ and the late partitioned version $m_{T2\perp}^{(1+1)} \equiv M_{\top\perp 2}$ does not become manifest. However, in more complicated scenarios with multiple visible daughter particles, one would in principle obtain different results from $m_{T2\perp}^{(1+3)}$, $m_{T2\perp}^{(1+2)}$ and $m_{T2\perp}^{(1+1)}$, which is why one is advised to carefully define which particular version of $m_{T2\perp}$ (and similarly for $m_{T2\parallel}$) is being used.

J. Additionally constrained variables

So far we have been discussing very general variables, which target the most general event topology of Fig. 5. Notice that we have made very few assumptions on how the decays (52) actually take place, and where such assumptions have been made (such as in Section X E), it is always possible to relax those constraints if desired. Also we did not use any additional information which may be available from the preliminary studies of other variables related to our events, for example the invariant mass distributions of the visible daughter collections $\mathcal{V}_a$. Armed with such additional information, one may in principle further constrain the minimization over the unknown momenta, and obtain new, more specialized versions of our variables. However, the downside is that such additional information typically comes at a cost: the need to make additional assumptions about the event topology.

As an example, consider the Oxford M_{2C} variable [14, 17], which is a variant of m_{T2}, subject to the following additional assumptions:

1. The two parents $\mathbb{P}_1$ and $\mathbb{P}_2$ are identical, with *a priori* unknown mass $M_{\mathbb{P}_1} = M_{\mathbb{P}_2}$, therefore the minimization over invisible momenta is performed subject to the additional constraint

$$(\mathbf{P}_1 + \mathbf{Q}_1)^2 = (\mathbf{P}_2 + \mathbf{Q}_2)^2 \tag{208}$$

with $\mathbf{P}_a$ and $\mathbf{Q}_a$ given by (53) and (57) correspondingly.

[29] This fact is found to be suprising to people who view m_{T2} and similar variables as acting on "projected" quantities. On the other hand it is no surprise to those who have always viewed m_{T2} as a variable insensitive to relative rapidity differences beteen the (total) invisible and (total) visible decay products of each parent — the line taken in [9].

2. There is only one and only one invisible particle in each invisible daughter set, i.e. that

$$|\mathcal{I}_1| = |\mathcal{I}_2| = 1.$$

3. There is more than one visible particle in each visible daughter set, i.e. that

$$|\mathcal{V}_1| = |\mathcal{V}_2| \geq 2.$$

4. There are no intermediate on-shell resonances, so that the decay (52) is effectively $(|\mathcal{V}_a| + |\mathcal{I}_a|)$-body for each $a = 1, 2$.

Given a large sample of events $\mathcal{S}$ that satisfy these conditions one can study the distribution of the invariant mass $M_{\mathcal{V}_a}$ of the visible particles in each set $\mathcal{V}_a$. If presented with some sufficiently large sample of events $\mathcal{S} = \{\epsilon \mid \epsilon \in \mathcal{A}\}$ one can measure the upper bound for the $M_{\mathcal{V}_a}$ distribution which will be found at the mass difference between the parent and the single invisible daughter:

$$\Delta^{\max} \equiv \max_{\text{events}} [M_{\mathcal{V}_a}] = M_{\mathbb{P}_a} - \not{\mathbf{M}}_a \tag{209}$$

One can then reuse this measurement for any event $\epsilon \in \mathcal{S}$ by asserting the constraint

$$\Delta^{\max} = \sqrt{(\mathbf{P}_a + \mathbf{Q}_a)^2} - \sqrt{\mathbf{Q}_a^2}, \tag{210}$$

during the process of minimization over $\mathbf{Q}_a$.

The advantage of such additionally constrained variables is that they are clearly better adapted for the study of the corresponding class of more restricted event topologies. And additional constraints can bring qualitatively new features, including otherwise unobtained *upper* bounds on parental masses [17, 22]. Their disadvantage is that they are better adapted (perhaps only suitable) for the study of those restricted topologies, and it is not clear how to interpret them once some of the assumptions hardwired in their definitions cease to be valid.

K. Other variables

In the literature one may sometimes encounter variables which have the appearance of a (transverse) invariant mass, but cannot be related to any of our variables in Table III. As an illustrative example, consider the transverse mass variable

$$M_{T_{WW}}^2 \equiv \Big(\sqrt{M_{e^+e^-}^2 + \vec{p}_{T,e^+e^-}^2} + \sqrt{M_{e^+e^-}^2 + \not{p}_T^2}\Big)^2 - \big(\vec{p}_{T,e^+e^-} + \vec{\not{p}}_T\big)^2, \tag{211}$$

proposed in Ref. [59] in relation to the $h \to W^+W^- \to e^+e^-\nu\bar{\nu}$ process discussed in Sec. X E. Comparing to the definition (195) of $M_{C,WW}$ and to the identity (196), we see that $M_{T_{WW}}$ can be formally obtained from $M_{C,WW} = M_{1\top}(0)$ with the rather ad hoc replacement

$$\not{\mathbf{M}}_1 = 0 \quad \to \quad \not{\mathbf{M}}_1 = M_{e^+e^-}. \tag{212}$$

However, there is no good physics justification for this conjecture and as a result, $M_{T_{WW}}$ cannot be related to any of the variables in Table III. Not surprisingly, subsequent studies [57] found that $M_{C,WW} = M_{1\top}(0)$ outperforms $M_{T_{WW}}$.

XI. SIMULATION: PHYSICS EXAMPLES

In this section we provide an illustration of our previous discussion with two specific physics examples from the Standard Model:

- *A case with* $N = 1$. Here we consider the inclusive (single) production of a SM Higgs boson (mostly from gluon fusion), followed by the decay of the Higgs to a leptonic W-pair:

$$\begin{aligned} pp &\to h + X \\ &\to W^+W^- + X \\ &\to \ell^+\ell^- + \not{p}_T + X, \end{aligned} \tag{213}$$

  where X plays the role of UVM and stands for jets from initial state radiation, unclustered hadronic energy, etc. In terms of our previous notation, this case involves one parent ($N = 1$), two visible particles ($N_{\mathcal{V}} = 2$) and two invisible particles ($N_{\mathcal{I}} = 2$).

- *A case with* $N = 2$. Here we consider dilepton events from inclusive $t\bar{t}$ pair production, where both W's decay leptonically:

$$\begin{aligned} pp &\to t\bar{t} + X \\ &\to b\bar{b}W^+W^- + X \\ &\to b\bar{b}\ell^+\ell^- + \not{p}_T + X. \end{aligned} \tag{214}$$

  This case corresponds to two parents ($N = 2$), four visible particles ($N_{\mathcal{V}} = 4$) and two invisible particles ($N_{\mathcal{I}} = 2$).

In both of those two cases, the events very closely resemble the typical SUSY-like events, in which there are two missing dark matter particles. Parton-level event simulation is performed with PYTHIA [60] at an LHC of 7 TeV, including the effects from the underlying event (using PYTHIA's default model for it).

A. An $N = 1$ example: a Higgs resonance

We start with the Higgs production process (213) for a Higgs boson mass $M_h = 200$ GeV. In the language of

Fig. 5, the Higgs resonance is treated as the only heavy parent particle ($N = 1$) and the event is partitioned as

$$\begin{aligned}\mathcal{V}_1 &= \{\ell^+, \ell^-\}, \\ \mathcal{I}_1 &= \{\nu_\ell, \bar{\nu}_\ell\}.\end{aligned}$$

This partitioning is pictorially represented in Fig. 8. We now concentrate on the five unprojected or singly projected variables which are of interest to us, namely $M_{\mathcal{F}}$ with $\mathcal{F} \in \{1, 1\top, \top 1, 1\circ, \circ 1\}$. Their distributions are shown in Fig. 11(a), and for proper comparison, we use the correct value of the missing mass parameter $\not{M}_1 = 0$ where necessary. In that case, according to the general property (167), all $M_{\mathcal{F}}$ variables are bounded from above by the parent mass, in this case M_h. For reference, Fig. 11(a) also shows the Breit-Wigner distribution of the Higgs resonance (yellow-shaded histogram). Fig. 11(a) confirms that the distributions obey the bound of eq. (167). Furthermore, it also shows that each of the five distributions appears to be saturated – i.e. that each has a kinematic endpoint at the value of the Higgs boson mass M_h (only a very tiny fraction of events is observed to exceed the bound, but this is due to the finite width of the Higgs parent).

We can confirm the endpoint is saturated for each of the variables $M_{\mathcal{F}} \in \{1, 1\top, \top 1, 1\circ, \circ 1\}$ by explicitly constructing an extremal event. We do so under the approximation that $M_\ell = M_\nu = 0$. Following the arguments of Sec. IX B we should construct an extremal event ϵ from the subset $\mathcal{E}$ sampled by nature of the total set of momentum configurations $\mathcal{A}$ that satisfy our general $N = 1$, $M_{\mathbb{P}_1} = M_H$ topology. In the case of the decay of interest (213) nature obliges us to impose an additional on-mass-shell condition for the intermediate $W^\pm$ bosons $\mathcal{E} = \{a \in \mathcal{A} \mid (P_i + Q_i)^2 = M_W^2 \ \ (i \in \{1, 2\})\}$.[30] An example of an extremal event $\epsilon \in \mathcal{E}$ that satisfies the constraints and that also saturates the two inequalities of (168) is (see also Fig. 10)

$$\begin{aligned}P_{\ell^+} &= (E_1, E_1, 0, 0) \\ P_{\ell^-} &= (E_2, E_2, 0, 0) \\ Q_\nu &= (E_2, -E_2, 0, 0) \\ Q_{\bar{\nu}} &= (E_1, -E_1, 0, 0),\end{aligned}$$

where $E_{1,2} = \frac{M_h}{4} \pm \frac{1}{2}\sqrt{M_h^2/4 - M_W^2}$.

[30] We assume the W to have narrow widths. In fact since we can construct an extremal event while imposing a strict on-shell requirement for the intermediate $W^\pm$ then we can certainly also do so when that requirement is relaxed by allowing the W to sample from its natural width distribution.

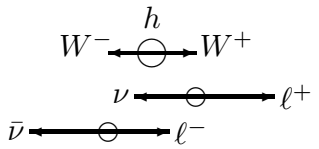

FIG. 10: Illustration of an example extremal configuration for the $N = 1$ variables when applied to the h example.

We note that in the cases of $M_{1\circ}$ and $M_{\circ 1}$, the kinematic endpoint coincides with the mass of the parent only because the final state objects in this example happened to be massless. In more general scenarios with massive particles the endpoints of $M_{1\circ}$ and $M_{\circ 1}$ will provide only an unsaturated lower bound on the parent mass, in line with (33).

Fig. 11(a) also allows us to compare the different $M_{\mathcal{F}}$ distributions to each other. As expected from the general property (170), the distributions of M_1 and $M_{1\top}$ (given in blue) are identical. As discussed in Section X C and shown in eqs. (180) and (182), they also coincide with the distribution of the $\sqrt{\hat{s}}_{\min}^{(\mathrm{sub})}$ variable from [46]. Similarly, in line with eq. (173), the distributions of $M_{\top 1}(\not{M}_1 = 0)$ and $M_{\circ 1}$ (shown in red) are practically indistinguishable, since the lepton masses are so tiny. Notice that this is only true when $M_{\top 1}$ is calculated with $\not{M}_1 = 0$, as was done here, otherwise the distributions of $M_{\top 1}$ and $M_{\circ 1}$ would generally be different. Finally, the distribution of $M_{1\circ}$ (shown in green) is distinct, as this variable is not related to any of the others.

Upon inspection of the shapes of different distributions in Fig. 11(a), one observes that M_1 and $M_{1\top}$ appear to peak closest to the parent mass M_h, and consequently, have the best defined endpoint structures. On the other hand, $M_{1\circ}$ peaks much farther from M_h, and has a rather low event population in the vicinity of its endpoint. Finally, the case of $M_{\top 1}(0) \simeq M_{\circ 1}$ represents an intermediate situation — the peak is found in between the peaks of $M_1 = M_{1\top}$ and $M_{1\circ}$; and the endpoint structure is more pronounced than the case of $M_{1\circ}$, but not as sharp as the case of $M_1 = M_{1\top}$. This is an inevitable consequence of the hierarchy (B53) among the mass bounds which is present in every event.

Next, in Fig. 11(b) we compare the distributions of the standard variables m_{eff} (red solid line) and $2\not{p}_T$ (green solid line) to their mass-bound counterparts $M_{\circ 1}$ (red dotted line) and $M_{1\circ}$ (green dotted line). Recall from the discussion in Sec. X B (and in particular eq. (179)) that $M_{\circ 1}$ is the analogue of m_{eff}, and in the limit of no upstream momentum the two variables become identical. This is confirmed in Fig. 11(b), which shows rather similar distributions for $M_{\circ 1}$ and m_{eff}. However, the analogy is not perfect and the m_{eff} distribution is slightly shifted

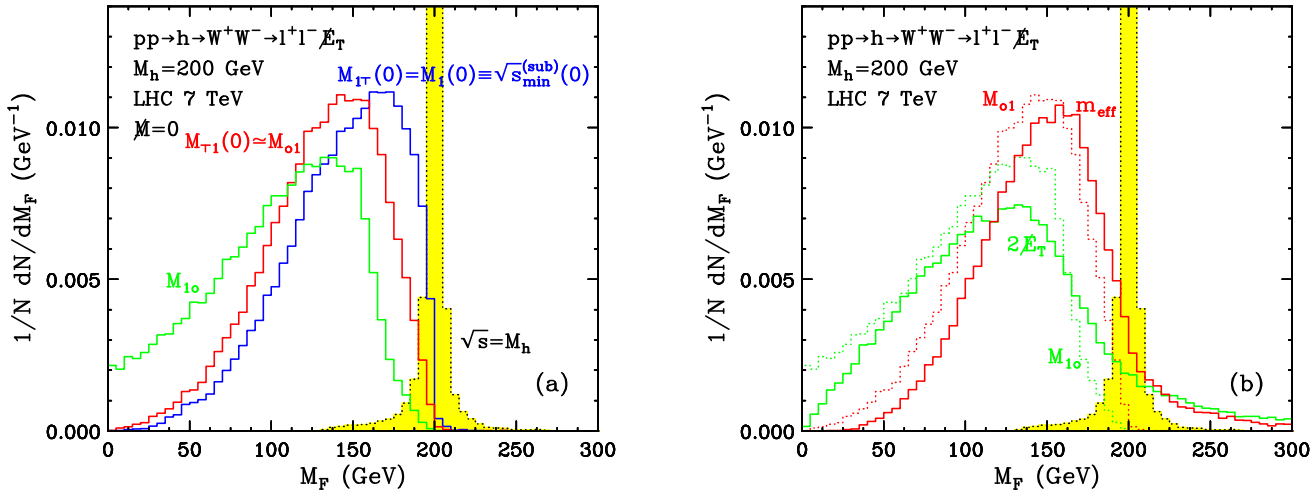

FIG. 11: (a) Unit-normalized distribution of the five $N = 1$ mass-bound variables $M_{\mathcal{F}}$, $\mathcal{F} \in \{1, 1\top, \top 1, 1\circ, \circ 1\}$ for the inclusive Higgs production process $h \to W^+W^- \to \ell^+\ell^- + \not{p}_T$ at a 7 TeV LHC, with $m_h = 200$ GeV and $\not{M} = 0$. The dotted (yellow-shaded) histogram gives the true $\sqrt{\hat{s}}$ distribution, which in this case is given by the Breit-Wigner h resonance. (b) Unit normalized distributions of the variables $m_{\rm eff}$ and $2\not{p}_T = 2\not{p}_T$ (solid lines), contrasted with $M_{\circ 1}$ and $M_{1\circ}$ (dotted lines).

to the right. The only[31] reason for this effect is the fact that we allow for initial state radiation in our sample, so that the Higgs parent is typically produced with some recoil and $u_T \neq 0$. This is why the the $m_{\rm eff}$ distribution does *not* terminate at M_h, but shows a long tail extending to $m_{\rm eff} > M_h$. In contrast, the $M_{\circ 1}$ distribution has an exact endpoint at M_h.

A similar analysis holds for the other pair of distributions (color coded in green) which are shown in Fig. 11(b). As explained in Sec. X A and seen from eq. (177), the variable $M_{1\circ}$ is the analogue of $2\not{p}_T$, since the two become identical in the limit of no upstream momentum ($u_T \to 0$). However, in the presence of upstream momentum, the proper behavior (an endpoint located at the parent mass) is retained only by the $M_{1\circ}$ distribution, while the $2\not{p}_T$ distribution picks up a long tail extending beyond the true value of M_h.

This concludes our discussion of the $N = 1$ unprojected and singly projected variables in relation to Higgs production (213). We note that one could also apply $N = 2$ variables to this example, this time considering the two W bosons as the two heavy parent particles, and partitioning as

$$\mathcal{V}_1 = \{\ell^+\}, \quad \mathcal{I}_1 = \{\nu_\ell\},$$
$$\mathcal{V}_2 = \{\ell^-\}, \quad \mathcal{I}_2 = \{\bar{\nu}_\ell\}.$$

The upper kinematic endpoints of the resulting distributions will be found at the corresponding parent mass, in this case the mass M_W of the W-bosons.

[31] We checked that when one restricts the plot only to events with $u_T = 0$, the distributions of $m_{\rm eff}$ and $M_{\circ 1}$ become identical, as required by eq. (179).

B. An $N = 2$ example: top quark pair production

As our next example, we consider dilepton events from the top quark pair production process (214). We assume that the two b-jets from the top quark decays have been tagged, which distinguishes them from QCD jets from initial state radiation. Correspondingly, the tagged b-jets will be included among the set of visible particles, while any remaining QCD jets will contribute to the UVM category.

We first reconsider the $N = 1$ variables already studied in Sec. XI A, and show that they can be useful even when there are multiple parents in the event. For the purpose of constructing $N = 1$ variables, the event is partitioned simply as

$$\mathcal{V}_1 = \{b, \bar{b}, \ell^+, \ell^-\},$$
$$\mathcal{I}_1 = \{\nu_\ell, \bar{\nu}_\ell\}.$$

Fig. 12(a) displays the distributions of the resulting $N = 1$ variables. Those distributions should be contrasted with the true $\sqrt{\hat{s}}$ distribution of the $t\bar{t}$ *pair*, which is shown in the figure with the yellow-shaded histogram. Just like in Fig. 11(a), we find only three distinct distributions, since $M_1 = M_{1\top}$ from (170) and $M_{\top 1}(\not{M}_1 = 0) \simeq M_{\circ 1}$ from (173). The hierarchical ordering of the three distributions is the same as in Fig. 11(a), the distribution of $M_1 = M_{1\top}$ being the hardest, and the distribution of $M_{1\circ}$ being the softest. Since all of our $N = 1$ variables are defined through minimization, each variable provides a lower bound on the true center-of-mass energy

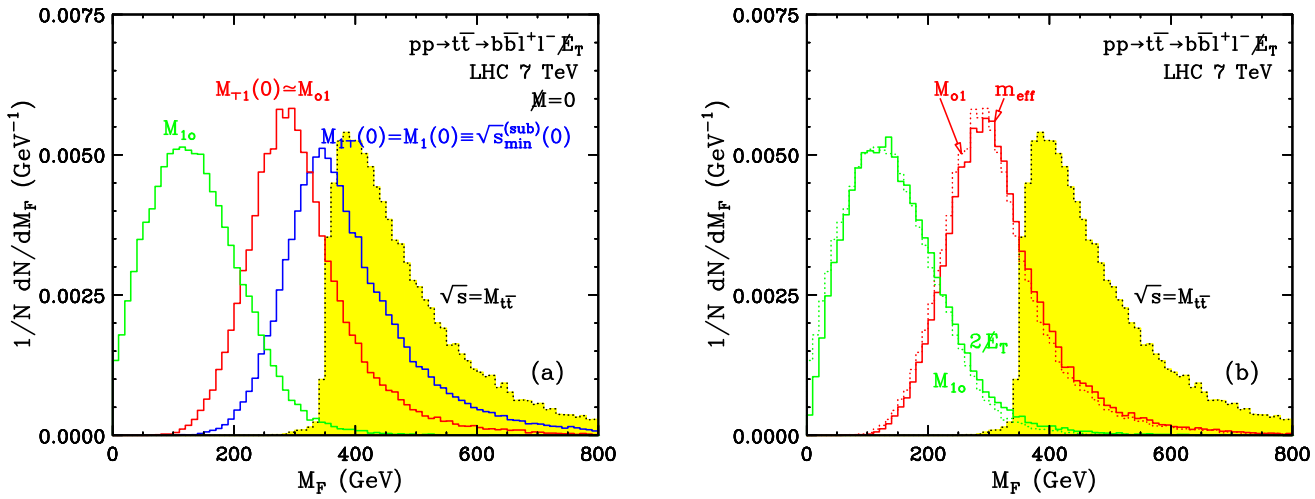

FIG. 12: The same as Fig. 11, but for the $t\bar{t}$ example. The yellow shaded distribution now gives the true invariant mass of the $t\bar{t}$ pair.

$\sqrt{\hat{s}}$ in the event. As one might have expected, it is the M_1 (or equivalently, the $M_{1\top}$) variable which offers the best (in the sense of being most stringent and meaningful) bound. Since M_1 and $M_{1\top}$ are identical to the $\sqrt{\hat{s}}^{(\rm sub)}_{\rm min}$ variable, their distribution exhibits the interesting property first noted in [50] in relation to $\sqrt{\hat{s}}_{\rm min}$: that the peak of the distribution is located very near the mass threshold for producing the two heavy parents, in this case the two top quarks. Indeed, notice how the peak in the (blue) $M_1 = M_{1\top} = \sqrt{\hat{s}}^{(\rm sub)}_{\rm min}$ histogram coincides with the onset of the (yellow-shaded) true $\sqrt{\hat{s}}$ distribution. When applied to searches for new physics, one can then use the peak in the $M_1 = M_{1\top} = \sqrt{\hat{s}}^{(\rm sub)}_{\rm min}$ distribution as a rough estimate of the new physics mass scale [46, 50].

In analogy to Fig. 11(b) here we can also perform a comparison of the usual variables $m_{\rm eff}$ and $2\not{p}_T$ to their mass-bound analogues $M_{\circ 1}$ and $M_{1\circ}$. In Fig. 12(b) we compare $m_{\rm eff}$ to $M_{\circ 1}$ (in red) and $2\not{p}_T$ to $M_{1\circ}$ (in green). This time the differences are much less pronounced that the single resonance case shown in Fig. 11(b). This suggests that for $N = 2$ processes, the variable $m_{\rm eff}$ $(2\not{p}_T)$ is on an equal footing with $M_{\circ 1}$ $(M_{1\circ})$.

We remind the reader that the $N = 1$ variables shown in Fig. 12(a) do not exhibit any upper kinematic endpoints, since they are being applied to $N = 2$ events, i.e. they have the "wrong" value of N and so the bounding relations (166) do not apply to any individual parent. Thus let us now discuss the $N = 2$ variables, which have the correct value of N and for which (166) holds. In the case of $N = 2$, a $t\bar{t}$ dilepton event is partitioned as

$$\begin{aligned} \mathcal{V}_1 &= \{b, \ell^+\}, \\ \mathcal{V}_2 &= \{\bar{b}, \ell^-\}, \\ \mathcal{I}_1 &= \{\nu_\ell\}, \\ \mathcal{I}_2 &= \{\bar{\nu}_\ell\}. \end{aligned}$$

This partitioning can be pictorially visualized in Fig. 7. Since we are primarily interested in the kinematical effects, for this illustrative example we make the simplifying assumption (unlikely to be realized in any real experiment) that each lepton can be associated with its sibling b-jet.

The distributions of the corresponding five $N = 2$ variables are shown in Fig. 13, where for illustrative purposes we use Monte Carlo truth information to properly assign the correct b-jet to each lepton. According to (166), these distributions are bounded from above by the individual parent mass, which in this case is the mass of the top quark. Correspondingly, in Fig. 13 the reference yellow-shaded distribution now shows the (average) top quark mass in the event, which follows the familiar Breit-Wigner shape (compare to the Higgs resonance shape in Fig. 11).

As before, we observe three distinct distributions, $M_2 = M_{2\top}$ (in blue), $M_{\top 2}(\{\not{\mathbf{M}}_a = 0\}) \simeq M_{\circ 2}$ (in red) and $M_{2\circ}$ (in green). All of them exhibit an upper endpoint less than or equal to the top quark mass M_t, in accordance with ((166), but the three shapes are considerably different. As before, and in agreement with the general hierarchy proven in the arguments leading up to (B53) for any event, the early partitioned versions M_2 and $M_{2\top}$ have the steepest endpoint, with the largest fraction of events near the endpoint.

Again, we can show that we expect the M_2, $M_{2\top}$, $M_{\top 2}$,

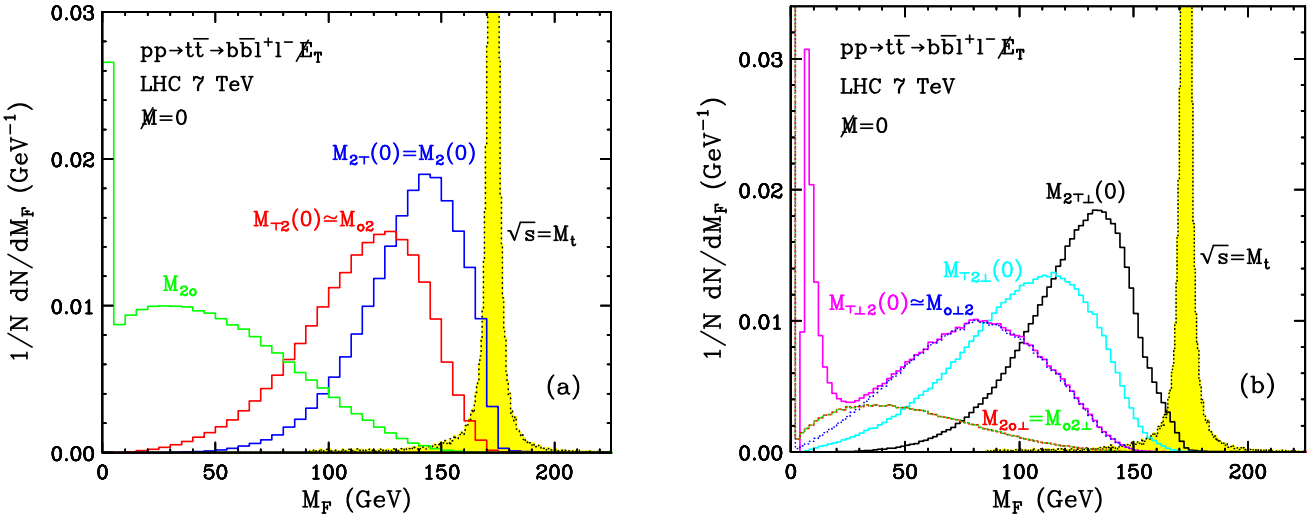

FIG. 13: The same as Fig. 12, but for $N = 2$ variables: (a) the unprojected M_2 and the singly projected variables $M_{2\top}$, $M_{\top 2}$, $M_{\circ 2}$ and $M_{2\circ}$; and (b) the doubly projected variables $M_{2\top\perp}$ (black), $M_{\top 2\perp}$ (cyan), $M_{\top\perp 2}$ (magenta), $M_{2\circ\perp}$ (red), $M_{\circ 2\perp}$ (green) and $M_{\circ\perp 2}$ (blue). The yellow shaded distribution now gives the average top quark mass in the event. In panel (b), "$\top$"-projected quantities are denoted with solid lines and are evaluated with $\not{\mathbf{M}}_1 = \not{\mathbf{M}}_2 = 0$, while "$\circ$"-projected quantities are denoted with dotted lines.

$M_{\circ 2}$ and $M_{2\circ}$ bounds to be saturated by explicity constructing an extremal event $\epsilon \in \mathcal{E}$ that satisfies the on-shell constraints of the t and $\bar{t}$ quarks and $W^{\pm}$ bosons. An example of such a configuration (see also Fig. 14) is

$$\begin{aligned} P_b = P_{\bar{b}} &= (E_b,\, p_b,\, 0,\, 0) \\ P_{\ell^+} = P_{\ell^-} &= (E_\ell,\, p_\ell,\, 0,\, 0) \\ Q_\nu = Q_{\bar{\nu}} &= (E_\nu,\, -E_\nu,\, 0,\, 0)\,, \end{aligned}$$

where,

$$\begin{aligned} E_b^2 &= p_b^2 + M_b^2 & E_\ell^2 &= p_\ell^2 + M_\ell^2 \\ p_b &= \lambda(M_t,\, M_b,\, M_W) & p_\ell &= M_\ell \sinh(\rho - \sigma) \\ E_\nu &= p^* e^{\sigma} & p^* &= \lambda(M_W,\, M_\ell,\, 0) \\ \sigma &= \sinh^{-1}\left(\frac{p_b}{M_W}\right) & \rho &= \sinh^{-1}\left(\frac{p^*}{M_\ell}\right), \end{aligned}$$

and the two-body momentum function is given by

$$\lambda(a,b,c) \equiv \frac{\sqrt{(a^2-(b+c)^2)\,(a^2-(b-c)^2)}}{2a}.$$

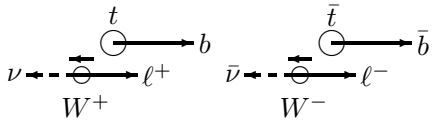

FIG. 14: Illustration of an example extremal configuration for the $N = 2$ variables when applied to the $t\bar{t}$ example.

Recall from eq. (205) that the early partitioned variables M_2 and $M_{2\top}$ are equivalent to the (1+3)-dimensional version $m_{T2}^{(1+3)}$ of the Cambridge m_{T2} variable. It appears therefore that for studies like the one presented here, where the UVM contributions can be safely identified and accounted for, $m_{T2}^{(1+3)}$ is preferable over $m_{T2}^{(1+2)}$.

We use the $t\bar{t}$ example to also illustrate the doubly projected variables from Table IV. Fig. 13(b) shows the doubly projected "$\top$"-projections (solid lines) and the "$\circ$"-projections (dotted lines). All "$\top$"-projected quantities are evaluated with $\not{\mathbf{M}}_1 = \not{\mathbf{M}}_2 = 0$. Each type of projection can be done in three different ways: early partitioning, $M_{2\top\perp}$ (black) and $M_{2\circ\perp}$ (red); late partitioning, $M_{\top\perp 2}$ (magenta) and $M_{\circ\perp 2}$ (blue); or in-between partitioning, $M_{\top 2\perp}$ (cyan) and $M_{\circ 2\perp}$ (green).

Similarly to the result from Fig. 13(a), Fig. 13(b) also reveals that the early partitioned, "$\top$"-projected variable $M_{2\top\perp}$ has the best defined endpoint structure, which clearly indicates the value of the parent mass M_t. As for the remaining variables, two are identically equal:

$$M_{2\circ\perp} \equiv M_{\circ 2\perp}, \tag{215}$$

which is a special case of the general identity (149), while two others are approximately equal:

$$M_{\top\perp 2}(\{\not{\mathbf{M}}_a = 0\}) \approx M_{\circ\perp 2}, \tag{216}$$

where a noticeable difference arises only at low values due to the finite mass of the b-quark — see eq. (174).

Existing variable	Mass-bound variable					
	$N=1$				$N=2$	
	$M_1(\not{\mathbf{M}}_1) = M_{1\top}(\not{\mathbf{M}}_1)$	$M_{\top 1}(\not{\mathbf{M}}_1)$	$M_{\circ 1}$	$M_{1\circ}$	$M_2(\not{\mathbf{M}}) = M_{2\top}(\not{\mathbf{M}})$	$M_{2\top\perp}(\not{\mathbf{M}})$
$2\not{p}_T = 2\not{E}_T$				$u_T \to 0$		
$m_{\rm eff}$		$\not{\mathbf{M}}_1 \to 0, u_T \to 0$	$u_T \to 0$			
$\sqrt{\hat{s}}^{(\rm sub)}_{\rm min}(\not{\mathbf{M}}_1)$	✓					
$\sqrt{\hat{s}}_{\rm min}(\not{\mathbf{M}}_1)$	$u_T \to 0$					
$m_{T e\nu}(M_e, M_\nu)$	✓	✓	$M_e, M_\nu \to 0$	$M_e, M_\nu \to 0$		
$M_{T,ZZ}(M_Z)$	✓	✓				
$M_{C,WW}$	$\not{\mathbf{M}}_1 \to 0$					
$m_T^{\rm true}$	$\not{\mathbf{M}}_1 \to 0$					
$m_{TZ'}^{reco}(M_Z)$	$u_T \to 0$	$u_T \to 0$				
$m_{T2}(\not{\mathbf{M}})$					✓	
$m_{T2\perp}(\not{\mathbf{M}})$						✓

TABLE VI: Correspondence between some of the existing variables in the literature, which were discussed in Section X, and the corresponding mass-bound variables. A checkmark (✓) implies an exact equivalence, otherwise the relevant limiting condition is listed. The last variable $m_{T2\perp}(\not{\mathbf{M}}_a)$ employs the doubly-projected $\perp$ construction described in Appendix VIII.

XII. CONCLUSIONS

The main "result" of this paper is the proposal made in Section VII of a general scheme for constructing and categorizing the basic invariant mass variables which are best suited for the study of missing energy events at hadron colliders. As a demonstration of the utility of this general scheme, in Section X we showed how a wide variety of widely used kinematic variables discussed in the literature can be properly accommodated in our framework. A short summary of this discussion is presented in Table VI, which exhibits the connections between the variables discussed in Section X and the corresponding mass-bound variables from Tables III and IV. The table reveals that one can give a new meaning to well-known variables like $\not{p}_T$ and $m_{\rm eff}$, which were originally introduced and defined in a way unrelated to any invariant mass considerations. Now we see that the same variables allow an alternative interpretation in terms of bounds on Lorentz invariants of interest as long as one is using the "massless" ($\circ$) type of projection for the transversification.

Another lesson from Table VI is that depending on the specific topology, the same bound may be constructed in different ways. A perfect illustration is provided by the variable $M_1 = M_{1\top}$. As discussed in detail in Sec. X E, even for the same final state (two leptons and missing energy), the variable $M_1 = M_{1\top}$ can emerge as differing bounds (either $M_{T,ZZ}$ or $M_{C,WW}$) depending on the choice of interpretation of the kinematical information.

But the value of the proposed scheme is not just in the accommodation of existing techniques and variables. The primary benefit from our approach is that, having understood the main principles behind the construction of a good invariant mass variable, the reader is now prepared to tackle almost any event topology, first by realizing what are the proper invariant mass variables for the case at hand, and second, knowing how to construct and calculate those variables. As discussed in Sections III–VI, there are a number of choices to be made along the way, related to the method of transversification, the partitioning of the event, and the exact order in which one takes all those operations. The main guiding principle through all this is that at the end of the day, one is always going to construct a bound on the mass of the heaviest parent. In that sense we are extending the principles and methods of construction put forth in [7] for m_{T2} and [50] for $\sqrt{\hat{s}}_{\rm min}$.

As we have seen, many of the generalized mass-bound variables are already in use at the LHC and elsewhere, but the majority have, for the moment, the status of solutions in search of problems.

Appendix A: Computer libraries offering "transverse" energy and mass variables

Though libraries should be a repository of human knowledge, any careful experimentalist will already have recognized that the computer libraries which support transverse projection methods for Lorentz vectors do not always produce the expected behavior. A selection of some of the most commonly used libraries and some of

Library	Object	Method/function name						
		$e_\top$	$e_\top^2$	$m_\top$	$m_\top^2$	m_{T2}	$e_\vee$	$e_\vee^2$
CLHEP[36]	LorentzVector	mt()	mt2()	–	–	–	et()	et2()
ROOT [37]	TLorentzVector	Mt()	Mt2()	–	–	–	Et()	Et2()
Fastjet [61]	Pseudojet	mperp()	mperp2()	–	–	–	Et()	Et2()
PGS [62]	–	–	–	–	–	–	v4et(p)	–
Oxbridge	LorentzVector	ET()	ET2()	LTV().mass()	LTV().masssq()	–	–	–
M_{T2} [38]	LorentzTransverseVector	Et()	Etsq()	mass()	masssq()	–	–	–
	Mt2_332_Calculator	–	–	–	–	mT2_332()	–	–
UCD M_{T2} [39]	mt2	Ea, Eb	Easq, Ebsq	–	–	get_mt2()	–	–
Defining equation in this paper		(5)		(7)		(205)	(13)	

TABLE VII: The versions of the transverse variables used in commonly used high-energy physics computer libraries and codes. A brief survey of experimental collaborations' software suggests that most follow the conventions of CLHEP. 'LTV' is a shorthand for the method getLorentzTransverseVector().

their methods for calculating transverse variables can be found in Table VII. In many cases the method of projection used (i.e. "$\top$" or "$\vee$") is undocumented and can only be determined by excavating the implementation. What is more, the names of the methods and functions in some cases produce output very different from what the user might expect. The result is that use of a plausible-sounding method can land the unwary user with a totally unexpected result – for example the CLHEP method called mt() returns the $\top$-projected transverse *energy* ($e_\top = \sqrt{M^2 + p_T^2}$), not the transverse mass they might have anticipated. Of course, because of the right-hand expression in eq. (5), one might fittingly call this quantity a "mass", but in that case the proper nomenclature should probably be a "longitudinal" mass and not a "transverse" mass.

To the extent that there is agreement on the conventions, one can see that the most commonly-used libraries (ROOT and CLHEP) use the $\vee$ convention when calculating "transverse energy" quantities. The Tevatron and LHC experimental collaborations tend to follow the "$\vee$" conventions when talking about "transverse energy" in calorimeters. For analyses where the transverse mass really matters, e.g. for $W \to \ell\nu$, the (ROOT and CLHEP) libraries have no function to return the 'usual' transverse mass of Refs [2–5, 55]: m_T must instead be calculated explicitly by the user.

Appendix B: Mass bounds on collections of momenta

In this section we present derivations of mass bounds on collections of arbitrary momenta, which may be represented by unprojected vectors and/or vectors transversified by any of the projections $\top$, $\vee$ and $\circ$. These cover the cases mentioned in VII, and justify the representation of multibody decays to visible and invisible particles in the form of a pair of composite momenta, where all visibles are projected identically (if at all) and all invisibles are likewise projected identically, though not necessarily by the same method as the visibles.

The question of *what* goes into the set of momenta from which we wish to generate the parental mass bound is not a mathematical question at all. However, once that set of momenta is formed, the question of *how to calculate the best bound making maximum use of the information contained in that set* is entirely mathematical. It is this mathematical question that we solve in the this section.

In essence, we try to answer the following question:

> Given a particular set of vectors, what is the greatest possible lower bound that we can place on the mass of any parent particle which could have have decayed to daughters characterized by that set? In particular, how does that bound depend on the dimensionalities and projection-types of the vectors characterizing the information about the daughters?

We shall denote the answer to that question as $\mathcal{M}\{\ldots\}$, where $\{\ldots\}$ is the set of vectors. We do not wish to restrict the set to contain only momenta of the same type (e.g. only four-momenta). Instead, we permit the set, if so desired, to be a heterogeneous mixture containing any number of four-momenta, $\top$-momenta, $\vee$-momenta, $\circ$-momenta or 2-momenta. For example, $\mathcal{M}\{A^\mu, B^\mu, c_\top^\alpha, d_\top^\alpha, e_\top^\alpha, f_\vee^\alpha, g_\circ^\alpha, \vec{h}_T\}$ would denote be the greatest possible lower bound on the mass of a particle assumed to have decayed to (at least) eight daughters, under the assumption that the only information from which we would wish that bound to be constructed were to comprise: the four-momenta of two daughters a and b; the masses and transverse two-momenta of three daughters c,

d and e; the three-speed and transverse two-momentum of daughter f; and the transverse two-momenta of particles g and h.[32]

1. Parental mass bounds from sets containing any two objects

Before considering parental bounds from arbitrary sets of momenta, we shall first consider the bound one obtains for each of the ten pair-wise combinations of the various types of vectors, i.e.:

$\mathcal{M}\{A^\mu, B^\mu\}$	$\mathcal{M}\{A^\mu, b^\alpha_\top\}$	$\mathcal{M}\{A^\mu, b^\alpha_\vee\}$	$\mathcal{M}\{A^\mu, \vec{b}_T\}$
–	$\mathcal{M}\{a^\alpha_\top, b^\alpha_\top\}$	$\mathcal{M}\{a^\alpha_\top, b^\alpha_\vee\}$	$\mathcal{M}\{a^\alpha_\top, \vec{b}_T\}$
–	–	$\mathcal{M}\{a^\alpha_\vee, b^\alpha_\vee\}$	$\mathcal{M}\{a^\alpha_\vee, \vec{b}_T\}$
–	–	–	$\mathcal{M}\{\vec{a}_T, \vec{b}_T\}$

To avoid imposing a physical interpretation on the vectors (other than that they are momenta), we generally work with A's and B's, as opposed to the P's and Q's used in the main text. The latter carry implications of visibility/invisibility that are irrelevant to the considerations of this section.

The list above appears to leave out the massless $\circ$-projection, but this is simply a special case of the $\top$ and $\vee$ projections, so the results for $\circ$ can be derived from the other two cases. In its place, we allow for combinations of vectors including transverse two-momenta $\vec{a}_T, \vec{b}_T$, in which the timelike component is simply unspecified. It will be seen that the bounds from combinations involving $\vec{a}_T, \vec{b}_T$ simply emerge to be the massless case.

a. *The $\mathcal{M}\{A^\mu, B^\mu\}$ parental mass bound*

We start with a straightforward case, taking care to be explicit about the sequence of operations that will also be required for the construction of the bound in the less trivial cases. The best parental mass bound[33] given a pair of daughter 1+3 momenta A^μ and B^μ is given by

$$\begin{aligned}\mathcal{M}^2\{A^\mu, B^\mu\} &= \min\left[M^2\right] \\ &= \min\left[P^\mu P_\mu\right] \\ &= \min\left[(A^\mu + B^\mu)(A_\mu + B_\mu)\right] \\ &= (A^\mu + B^\mu)(A_\mu + B_\mu), \\ &\equiv (A^\mu + B^\mu)^2 \end{aligned} \tag{B1}$$

where the first equality is simply a rephrasing of the meaning of $\mathcal{M}\{\}$ as the minimum mass consistent with the constraints. The second equality is from the definition of the inner product (or physically the definition of the mass), and the third equality is from the definition of a vector space (physically representing energy-momentum conservation). The fourth equality is a statement that the vectors A^μ and B^μ are fully specified, so the minimization is trivial (no parameters need be changed). The hopefully unsurprising outcome, then, is that the best lower bound on the parental mass is given by the invariant mass of the two daughter momenta A^μ and B^μ.

b. *The $\mathcal{M}\{A^\mu, b^\alpha_\top\}$ parental mass bound*

To calculate the bound

$$\mathcal{M}\{A^\mu, b^\alpha_\top\} \tag{B2}$$

we note that $b^\alpha_\top$ contains partial information about some 1+3 vector B^μ which projects to $b^\alpha_\top$ under the $\top$-projetion, about which the x and y components are known, but the z component, b_z is completely unspecified. The bound (B2) can therefore be rephrased,

$$\mathcal{M}^2\{A^\mu, b^\alpha_\top\} = \min_{b_z}\left[(A^\mu + B^\mu)^2\right]. \tag{B3}$$

For the minimization we recognise that provided either $M_A \neq 0$ or $|\vec{a}_T| \neq 0$, then M is unbounded above as $b_z \to \pm\infty$. Provided that we are dealing with particles produced with non-zero transverse momentum (which we shall assume hereafter), the solution must then be given by the local minimum

$$\begin{aligned} 0 &= \frac{\partial}{\partial b_z}(A^\mu + B^\mu)^2 \\ &= \frac{\partial}{\partial b_z}\left(M_A^2 + M_B^2 + 2\left(E_A E_B - \vec{a}_T \cdot \vec{b}_T - a_z b_z\right)\right). \end{aligned}$$

The minimization selects $b_z/E_B = a_z/E_A$ such that B^μ has equal rapidity to A^μ,

$$y_B = y_A. \tag{B4}$$

To calculate the value of the mass bound we recognise that, by the definition of the Lorentz transformation, the inner product of A^μ and B^μ, which we might denote by $g(A,B) \equiv A^\mu g_{\mu\nu} B^\nu$, is invariant under identical Lorentz transforms Λ of both vectors

$$g(A,B) = g(\Lambda A, \Lambda B). \tag{B5}$$

By letting Λ be a boost along the z-axis corresponding to rapidity change $-y_A$, which will then set both rapidities to zero, one finds that

$$g(A,B) = e^{(A)}_\top e^{(B)}_\top - \vec{a}_\top \cdot \vec{b}_\top. \tag{B6}$$

---

[32] Note that it makes no difference whether we use $\vec{h}_T$ instead of $h^\alpha_\circ$ as an input, as the information content of each is identical.

[33] Note that it is simplest to calculate the bound for the squared of the parental mass $\mathcal{M}^2\{\ldots\}$ rather than for the parental mass itself $\mathcal{M}\{\ldots\}$. This difference is of no consequence, and so for brevity we will talk only of "mass bounds" in the text, ignoring the square.

The best lower bound on the parent mass for daughters specified by a 1+3 momentum A^μ and a $\top$-projected 1+2 momentum $b^\alpha_\top$ is then given by

$$\begin{aligned}\mathcal{M}^2\{A^\mu, b^\alpha_\top\} &= M_A^2 + M_B^2 + 2\left(e^{(A)}_\top e^{(B)}_\top - \vec{a}_\top\cdot\vec{b}_\top\right)\\ &= (a_\top + b_\top)^\alpha (a_\top + b_\top)_\alpha\\ &\equiv (a^\alpha_\top + b^\alpha_\top)^2. \end{aligned} \tag{B7}$$

c. The $\mathcal{M}\{A^\mu, \vec{b}_T\}$ parental mass bound

The bound on a 1+3 Lorentz vector with a transverse two-vector can be found in a similar manner, but the 1+3 vector which projects to $\vec{b}_T$ is now given by some B which has both unknown z component *and* unknown mass. The bound is given by

$$\mathcal{M}\{A, \vec{b}_T\} = \mathcal{M}\{A, B\} = \min_{b_z, M_B}\left[A + B\right]. \tag{B8}$$

A similar argument to that which led to (B4) shows that the b_z component must be such that $y_B = y_A$. The M_B minimization selects $M_B = 0$, so that

$$\begin{aligned}\mathcal{M}^2\{A, \vec{b}_T\} &= M_A^2 + 2\left(e^{(A)}_T|\vec{b}_T| - \vec{a}_T\cdot\vec{b}_T\right)\\ &\equiv (a^\alpha_\top + b^\alpha_\circ)^2, \end{aligned} \tag{B9}$$

so the bound is formed by turning the transverse two-momentum into a $\circ$-projected 1+2 momentum. Comparing with (B7), we see that if $b^\alpha_\top$ is made massless $M_B = 0$, then $e^{(B)}_\top = |\vec{b}_T|$ and (B9) is reproduced.

d. The $\mathcal{M}\{a^\alpha_\top, b^\alpha_\top\}$ parental mass bound

For each of the 1+2 $\top$-projected vectors, the corresponding set of 1+3 dimensional objects shares the same transverse components and inner product (mass) as their $\top$-projected counterpart, but has arbitrary z momentum. The bound is then given by

$$\mathcal{M}^2\{a_T, b_T\} = \min_{a_z, b_z}\left[(A + B)^2\right] \tag{B10}$$

This time the minimizations force the rapidities of A and B to be equal, but leave the value of that rapidity $y_A = y_B$ free. Similarly to the previous cases,

$$\begin{aligned}\mathcal{M}^2\{a_T, b_T\} &= M_A^2 + M_B^2 + 2\left(e^{(A)}_T e^{(B)}_T - \vec{a}_T\cdot\vec{b}_T\right)\\ &= (a^\alpha_\top + b^\alpha_\top)^2. \end{aligned} \tag{B11}$$

e. The $\mathcal{M}\{a^\alpha_\top, \vec{b}_T\}$ parental mass bound

The limit on M is given by

$$\mathcal{M}^2\{a_T, \vec{b}_T\} = \min_{a_z, b_z, M_B}\left[(A + B)^2\right]. \tag{B12}$$

The minimizations set the rapidities to be equal $y_A = y_B$ (but undefined) and $M_B = 0$. The limit again appears in the form,

$$\begin{aligned}\mathcal{M}^2\{a_T, \vec{b}_T\} &= M_A^2 + 2\left(e^{(A)}_T|\vec{b}_T| - \vec{a}_T\cdot\vec{b}_T\right)\\ &\equiv (a^\alpha_\top + b^\alpha_\circ)^2. \end{aligned} \tag{B13}$$

f. The $\mathcal{M}\{\vec{a}_T, \vec{b}_T\}$ parental mass bound

The limit on M for a pair of transverse two-momenta is given by

$$\mathcal{M}^2\{\vec{a}_T, \vec{b}_T\} = \min_{a_z, b_z, M_A, M_B}\left[(A + B)^2\right]. \tag{B14}$$

The z minimizations again set the relative rapidities equal but arbitrary $y_A = y_B$, and the mass minimizations set $M_A = M_B = 0$.

$$\begin{aligned}\mathcal{M}^2\{\vec{a}_T, \vec{b}_T\} &= 2\left(|\vec{a}_T|\,|\vec{b}_T| - \vec{a}_T\cdot\vec{b}_T\right)\\ &\equiv (a^\alpha_\circ + b^\alpha_\circ)^2. \end{aligned} \tag{B15}$$

g. The $\mathcal{M}\{A^\mu, b^\alpha_\vee\}$ parental mass bound

The $\vee$ projection described in section III B maps all 1+3 vectors B^μ with the same transverse momentum $\vec{b}_T$ and velocity $V_B = |\vec{b}|/E_B$ to the same 1+2 vector $b^\alpha_\vee$. Therefore the longitudinal momentum component b_z is unspecified and the parental mass bound $\mathcal{M}\{A^\mu, b^\alpha_\vee\}$ is given by

$$\mathcal{M}^2\{A^\mu, b^\alpha_\vee\} = \min_{b_z}\left[(A + B)^2\right], \tag{B16}$$

From equations 22 and 23, we see that we can decompose the full (1+3)-dimensional energy and mass

$$E_B^2 = (e^B_\vee)^2 + (e^B_z)^2, \tag{B17}$$

$$M_B^2 = (m^B_\vee)^2 + (m^B_z)^2\,. \tag{B18}$$

each in terms of a transverse quantity ($e^B_\vee$, $m^B_\vee$) and a longitudinal quantity ($e^B_z = |\vec{b}_z|/V_B$, $m^B_z = |\vec{b}_z|/(V_B\gamma_B)$) with γ_B denoting the Lorentz factor $1/\sqrt{1 - V_B^2}$.

Using these relations, we can write the Lorentz-invariant quantity $(A + B)^2$ as

$$\begin{aligned}(A^\mu + B^\mu)^2 &= M_A^2 + (m^B_\vee)^2 + b_z^2/(V_B\gamma_B)^2\\ &\quad + 2\left(\frac{E_A}{V_B}\sqrt{b_T^2 + b_z^2} - \vec{a}_T\cdot\vec{b}_T - a_z b_z\right). \end{aligned} \tag{B19}$$

Leaving aside the trivial case of $b_T = 0$, we now attempt the minimization over b_z, requiring

$$\begin{aligned}0 &= \frac{\partial}{\partial b_z}(A^\mu + B^\mu)^2\\ &= 2\left(\frac{E_A}{V_B}\frac{b_z}{\sqrt{b_T^2 + b_z^2}} - a_z + \frac{b_z}{V_B^2\gamma_B^2}\right). \end{aligned} \tag{B20}$$

This gives rise to a quartic in b_z,

$$(b_T^2 + b_z^2)(b_z - \alpha)^2 - \epsilon^2\, b_z^2 = 0\ , \tag{B21}$$

where the constants

$$\alpha \ = \ a_z\, V_B^2\, \gamma_B^2, \tag{B22}$$
$$\epsilon \ = \ E_A\, V_B\, \gamma_B^2\, . \tag{B23}$$

The need to solve this quartic makes the $\mathcal{M}^2\{A^\mu, b_\vee^\alpha\}$ bound intractable in comparison with the similar $\mathcal{M}^2\{A^\mu, b_\top^\alpha\}$ bound. Similar difficulties are encountered in the following $\mathcal{M}^2\{a_\vee^\mu, b_\vee^\alpha\}$ case.

One might guess that the $y_B = y_A$ condition resulting from $\mathcal{M}^2\{A^\mu, b_\top^\alpha\}$ bound could represent the correct solution, since we are again working with a fully (1+3)-dimensional vector combined with a (1+2)-dimensional vector. But the solution this condition gives for b_z is not a root of the quartic in (B21). One can show that $b_z' = 0$ implies

$$b_z = \frac{b_T \alpha}{\sqrt{\epsilon^2 V_B^2 - \alpha^2}}\ , \tag{B24}$$

which when substituted into the LHS of (B21) gives

$$\frac{b_T^2 \alpha^2 \epsilon^2}{(\alpha^2 - \epsilon^2 V_B^2)^2}\Big(b_T^2 V_B^4 \ + \ (V_B^4 - 1)(\epsilon^2 V_B^2 - \alpha^2) \\ - \ 2\, b_T\, V_B^4 \sqrt{\epsilon^2 V_B^2 - \alpha^2}\Big),$$

which is in general non-zero, i.e. the equal rapidities condition only minimizes b_z under certain special conditions.

h. *The $\mathcal{M}\{a_\vee^\alpha, b_\vee^\alpha\}$ parental mass bound*

In the case of two $\vee$-projections, the mass bound is given by

$$\mathcal{M}^2\{a_\vee^\alpha, b_\vee^\alpha\} = \min_{a_z, b_z} \left[(A+B)^2\right]\ , \tag{B25}$$

the minimization of which involves finding a_z and b_z, with both V_A and V_B being held fixed, such that each of a_z and b_z is a root of a quartic like that in (B21).

In this situation, we can apply the same mass/energy decompositions (B17) and (B18) to a_z, to get a variant of (B19),

$$(A^\mu + B^\mu)^2 = \\ = (m_\vee^A)^2 + \frac{a_z^2}{V_A^2\gamma_A^2} + (m_\vee^B)^2 + \frac{b_z^2}{V_B^2\gamma_B^2} \\ + 2\left(\frac{\sqrt{a_T^2 + a_z^2}}{V_A}\frac{\sqrt{b_T^2 + b_z^2}}{V_B} - \vec{a}_T \cdot \vec{b}_T - a_z b_z\right). \tag{B26}$$

Differentiating by a_z and by b_z separately, the minimization imposes

$$0 \ = \ \frac{\partial}{\partial a_z}(A^\mu + B^\mu)^2 \tag{B27}$$
$$= \ 2\left(\frac{\sqrt{b_T^2 + b_z^2}}{\sqrt{a_T^2 + a_z^2}}\frac{a_z}{V_A V_B} - b_z + \ \frac{a_z}{V_A^2\gamma_A^2}\right)\ ,$$

and simultaneously

$$0 \ = \ \frac{\partial}{\partial b_z}(A^\mu + B^\mu)^2 \tag{B28}$$
$$= \ 2\left(\frac{\sqrt{a_T^2 + a_z^2}}{\sqrt{b_T^2 + b_z^2}}\frac{b_z}{V_A V_B} - a_z + \ \frac{b_z}{V_B^2\gamma_B^2}\right)\ .$$

Note that the fraction $|\vec{a}|/|\vec{b}|$ appears in both (B27) and (B28), albeit as a reciprocal in the latter. So, we can combine the two minimization constraints in the form of a quadratic in a_z and b_z:

$$\frac{a_z^2}{V_A^2\gamma_A^2} + \frac{b_z^2}{V_B^2\gamma_B^2} + a_z b_z\left(\frac{1}{V_A^2\gamma_A^2} + \frac{1}{V_B^2\gamma_B^2}\right) = 0\ . \tag{B29}$$

This can be solved to give $a_z = -c\, b_z$, with

$$c = 1 \quad \text{or} \quad \frac{V_A^2\gamma_A^2}{V_B^2\gamma_B^2}\ . \tag{B30}$$

Substituting this solution back into (B26) gives a pleasingly simple result

$$A^\mu + B^\mu)^2 = \\ = (m_\vee^A)^2 + (m_\vee^B)^2 + b_z^2\left(2\,c + \frac{c^2}{V_A^2\gamma_A^2} + \frac{1}{V_B^2\gamma_B^2}\right) \\ + 2\left(\frac{\sqrt{a_T^2 + c^2 b_z^2}}{V_A}\frac{\sqrt{b_T^2 + b_z^2}}{V_B} - \vec{a}_T \cdot \vec{b}_T\right). \tag{B31}$$

Since c was chosen to be positive, this expression is clearly minimised for $b_z = 0$, implying that $a_z = 0$ as well. If we then make the replacements $a_T/V_A = e_\vee^A$ and $b_T/V_B = e_\vee^B$, we find that the choice $a_z = b_z = 0$ gives, quite simply and in tune with our intuition and inductive sense,

$$\mathcal{M}^2\{a_\vee^\alpha, b_\vee^\alpha\} = (a_\vee^\alpha + b_\vee^\alpha)^2\, . \tag{B32}$$

i. *The $\mathcal{M}\{a_\top^\alpha, b_\vee^\alpha\}$ parental mass bound*

The bound $\mathcal{M}^2\{a_\top^\alpha, b_\vee^\alpha\}$ requires minimization over both a_z and b_z.

$$\mathcal{M}^2\{a_\top^\alpha, b_\vee^\alpha\} = \min_{a_z, b_z} \left[(A^\mu + B^\mu)^2\right]\ , \tag{B33}$$

with $(A^\mu + B^\mu)^2$ defined as before in (B19). First the minimization over a_z forces the rapidities of A and B to

be equal, $a_z/E_A = b_z/E_B$. Plugging this into (B19), we obtain

$$\begin{aligned}(A^\mu + B^\mu)^2 &= M_A^2 + (m_\vee^B)^2 + b_z^2/(V_B\gamma_B)^2 \qquad \text{(B34)}\\ &\quad + 2\left(E_A E_B - \vec{a}_T\cdot\vec{b}_T - b_z^2\frac{E_A}{E_B}\right).\\ &= M_A^2 + (m_\vee^B)^2 + b_z^2/(V_B\gamma_B)^2 \qquad \text{(B35)}\\ &\quad + 2\left(\frac{E_A}{E_B}(E_B^2 - b_z^2) - \vec{a}_T\cdot\vec{b}_T\right).\end{aligned}$$

The expression $(E_B^2 - b_z^2)$ can be written one of two ways – either as $(e_\top^B)^2$ or as $(e_\vee^B)^2 - b_z^2/(V_B\gamma_B)^2$. We choose the latter, since we have fixed $e_\vee^B$, but if we were to fix instead $e_\top^B$, we would rederive (B9).

A further simplification is implied by the equal rapidities condition, since $E_A = e_\top^A\, E_B/\sqrt{E_B^2 - b_z^2}$. Using this and (B17), we find

$$\begin{aligned}(A^\mu + B^\mu)^2 &= M_A^2 + (m_\vee^B)^2 + b_z^2/(V_B\gamma_B)^2 \qquad \text{(B36)}\\ &\quad + 2\left(e_\top^A\, e_\vee^B\sqrt{1 + \frac{b_z^2}{b_T^2\gamma_B^2}} - \vec{a}_T\cdot\vec{b}_T\right).\end{aligned}$$

Recognising that $b_T^2\gamma_B^2$ is positive, we see that the b_z minimization simply gives $a_z = b_z = 0$, and thus

$$\mathcal{M}^2\{a_\top^\alpha, b_\vee^\alpha\} = (a_\top^\alpha + b_\vee^\alpha)^2\,. \qquad \text{(B37)}$$

j. The $\mathcal{M}\{a_\vee^\alpha, \vec{b}_T\}$ parental mass bound

This mass bound is similar to $\mathcal{M}\{a_\vee^\alpha, b_\top^\alpha\}$ with the additional minimization of $M_B \mapsto 0$.

$$\begin{aligned}\mathcal{M}^2\{a_\vee^\alpha, \vec{b}_T\} &= \min_{a_z, b_z, M_B}\left[(A+B)^2\right] \qquad \text{(B38)}\\ &= (m_\vee^A)^2 + 2\left(e_\vee^A\, b_T - \vec{a}_T\cdot\vec{b}_T\right) \qquad \text{(B39)}\\ &= (a_\vee^\alpha + b_\circ^\alpha)^2\,. \qquad \text{(B40)}\end{aligned}$$

2. Arbitrarily large sets of (1+3)-, (1+2)$_\top$- and 2-vectors

The generalization of (B1) to an arbitrarily large set of fully specified 1+3 vectors $\mathcal{A} = \{A_i^\mu \mid 1 \le i \le |\mathcal{A}|\}$ is

$$\begin{aligned}\mathcal{M}^2\{\mathcal{A}\} &= \min\left[M^2\right] = \min\left[P^2\right]\\ &= \min\left[(\Sigma_i A_i^\mu)^2\right] = (\Sigma_i A_i)^2. \qquad \text{(B41)}\end{aligned}$$

Note that in the special case of fully specified 1+3 vectors, the mass bound for the set is the same as the mass bound of the single object formed of the sum of those vectors

$$\mathcal{M}\{\mathcal{A}\} = \mathcal{M}\{\Sigma_i A_i\}. \qquad \text{(B42)}$$

Let us further generalize our results to an arbitrary set of (1+3)-vectors $\mathcal{A}$ and $(1+2)_\top$-projected vectors $\mathcal{B}_\top = \left\{b_{j\top}^\alpha \mid 1 \le j \le |\mathcal{B}_\top|\right\}$. Each of the $b_{j\top}^\alpha$ has a (1+3)-vector equivalence class B_j^μ for which the z components can take any value. Writing $\mathcal{B} = \left\{B_j^\mu \mid 1 \le j \le |\mathcal{B}_\top|\right\}$ and $\mathcal{B}_z = \{b_{jz} \mid 1 \le j \le |\mathcal{B}_z|\}$, we can therefore write the mass bound as

$$\mathcal{M}\{\mathcal{A}, \mathcal{B}_\top\} = \mathcal{M}\{\mathcal{A}, \mathcal{B}\} = \min_{\mathcal{B}_z}\left[\left(\Sigma_i A_i^\mu + \Sigma_j B_j^\mu\right)^2\right]. \qquad \text{(B43)}$$

where each of the B_j^μ has a free z component. The result can be found by induction. The bound $\mathcal{M}\{K_1^\mu, \ldots, K_m^\mu\}$ for some set of fully specified (1+3)-vectors is given by the sum $\mathcal{M}\{\Sigma_{i=1,m} K_i^\mu\}$ by (B42). Adding a further 1+3 vector K_{m+1}^μ which has free z momentum to that set gives a bound $\mathcal{M}\{K_1^\mu, \ldots, K_{m+1}^\mu\}$. A similar argument to that which led to (B4) shows that the rapidity of K_{m+1}^μ must be equal to that of $\Sigma_{i=1,m} K_i^\mu$. With this constraint applied K_{m+1}^μ becomes a fully specified 1+3 vector, so we can treat it as one of the known 1+3 vectors and proceed with the next 1+2 ($\top$-projected) vector in the set.

Applying this argument sequentially to the B_j we find that

$$\mathcal{M}\{\mathcal{A}, b_{j\top}^\alpha\} = \mathcal{M}\{(\Sigma_i A_i^\mu), B_j^\mu\}, \qquad \text{(B44)}$$

where each of the B_j has the same rapidity as $\Sigma_i A_i^\mu$.

Since the set of 1+2 $\top$-projected vectors is isomorphic to the set of 1+3 vectors with fixed (but arbitrary) rapidity under the operations of addition and inner product, we can rewrite this bound as

$$\mathcal{M}\{\mathcal{A}, \mathcal{B}_\top\} = \mathcal{M}\{(\Sigma_i A_i^\mu), (\Sigma_j b_{j\top}^\alpha)\}, \qquad \text{(B45)}$$

the bound for the summed 1+3 vector $(\Sigma_i A_i^\mu)$ and the 1+2 $\top$-projected vector $(\Sigma_j b_{j\top}^\alpha)$, the explicit formula for which is given in (B7).

We can further extend the argument by allowing some other daughters parameterized only by their two-momentum to be added to the set,

$$\mathcal{M}\{\mathcal{A}, \mathcal{B}_\top, \mathcal{C}_T\}, \qquad \text{(B46)}$$

with $\mathcal{C}_T = \{\vec{c}_{kT} \mid 1 \le k \le |\mathcal{C}_T|\}$. Each 2-vector $\vec{c}_{kT}$ has a corresponding equivalence class which can be represented by a 1+2 $\top$-projected vector $c_{k\top}^\alpha$ with unknown mass. The arguments which led to (B44) apply equally to the $c_{k\top}$, so the corresponding C_k^μ 1+3 vector rapidities are set equal to $\Sigma_i A\mu_i$, but now we have the extra minimization over the masses which fixes $m_C = 0$ for each C_k^μ (or indeed $c_{k\top}^\alpha$) .

Therefore

$$\mathcal{M}\{\mathcal{A}, \mathcal{B}_\top, \mathcal{C}_T\} = \mathcal{M}\{(\Sigma_i A_i^\mu), (\Sigma_j b_{j\top}^\alpha + \Sigma_k c_{k\circ}^\alpha)\}\,. \qquad \text{(B47)}$$

Now (B47) has the same form as all the previous bounds, but in obtaining the result we have found out something non-trivial: one would *not* get the best bound on M if one were simply to replace the set of 2-vectors $\mathcal{C}_T$ by their sum $(\Sigma_k \vec{c}_{kT})$: One must instead add the corresponding massless 1+2 vectors $c_{k\circ}^\alpha$.

In principle we could now try to extend our bounds to include (arbitrarily large numbers of) $\vee$-projected 1+2 vectors. However we shall *not* do so for two reasons. The first reason is that in collider experiments such as the LHC, situations for which $\vee$ projection is appropriate are rare. It is only in very unusual cases where we might find ourselves knowing just the transverse momentum components and the size of the three-velocity, but not the azimuthal angle θ, the z-momentum, or the mass.

The second reason we do not pursue the $\vee$ vectors further is that one ends up with a real mess, as we have seen. The most basic pair-wise combination $\mathcal{M}\{A, b_\vee\}$ requires solution of a quartic equation in b_z. Only if one is solely interested in combining ($\top$, $\vee$, $\circ$)-projected vectors might the expressions be tractable, but the utility of such a combination is unclear.

3. Mass bound hierarchies

The similarity in the expressions for the mass bounds derived in the preceding sections allows for a further observation – that as progressively more information is neglected or unknown, the mass bound is lowered. Intuitively one would expect this, since the absence of hard information causes one to have to be progressively more conservative, but we can, with little additional work, show this explicitly to be the case.

We set out, therefore, to prove the hierarchy that was seen earlier in (175). Our proof proceeds in two stages. In the first stage we demonstrate the result for the case $N = 1$, in which the hierarchy becomes:

$$M_1 = M_{1\top} \geq M_{\top 1} \geq M_{\circ 1} \geq M_{1\circ}. \tag{B48}$$

In the second stage we extend this to general N.

Using the results of the previous section, we can treat each of the mass bound variables in terms of the composite visible and composite invisible objects described in section VI B. The equality in (B48) then results from the definition of M_1 as a concrete case of (B7), where A^μ represents the visible P^μ, and $b^\alpha_\top$ the invisible $q^\alpha_\top$. Similarly, $M_{1\top}$ is just (B11), where $a^\alpha_\top$ and $b^\alpha_\top$ stand in for $p^\alpha_\top$ and $q^\alpha_\top$. On comparing (B7) with (B11), we see that they are identical, and hence $M_1 = M_{1\top}$.

For the next statement, $M_{1\top} \geq M_{\top 1}$, we have to consider the difference between "early" and "late" partition, i.e. whether we retain information about the relative longitudinal momenta of the visibles. Let our visible composite $\mathbf{P}^\mu_a$ of parent $\mathbb{P}_a$ be composed of constituents P^μ_i, i.e.

$$\begin{aligned}\mathbf{P}^\mu_a &= \sum_{i\in\mathcal{V}_a} P^\mu_i \\ &= (\mathbf{E}_a, \vec{\mathbf{p}}_{aT}, \mathbf{p}_{az}) \,,\end{aligned}$$

with

$$\begin{aligned}\mathbf{E}_a &= \sum_{i\in\mathcal{V}_a} E_i \,, \\ \vec{\mathbf{p}}_{aT} &= \sum_{i\in\mathcal{V}_a} \vec{p}_{iT} \,, \\ \mathbf{p}_{az} &= \sum_{i\in\mathcal{V}_a} p_{iz} \,.\end{aligned}$$

We form the early-partitioned composite

$$\mathbf{p}^\alpha_{a\top} = \left(\sum_{i\in\mathcal{V}_a} P^\mu_i\right)_\top = (\mathbf{e}_{a\top}, \vec{\mathbf{p}}_{aT}) \,,$$

and the late-partitioned composite

$$\mathbf{p}^\alpha_{\top a} = \sum_{i\in\mathcal{V}_a} p^\alpha_{i\top} \,, = (\mathbf{e}_{\top a}, \vec{\mathbf{p}}_{Ta}) \,,$$

differing only in their energy components

$$\begin{aligned}\mathbf{e}_{a\top} &= \sqrt{\mathbf{E}^2_a - \mathbf{p}^2_{az}} \\ &= \sqrt{\mathbf{M}^2_a + \mathbf{p}^2_{aT}} \,, \\ \mathbf{e}_{\top a} &= \sum_{i\in\mathcal{V}_a} e_{i\top} \\ &= \sqrt{\mathbf{m}^2_{\top a} + \mathbf{p}^2_{Ta}} \,,\end{aligned}$$

where

$$\begin{aligned}\mathbf{M}^2_a &= \left(\sum_{i\in\mathcal{V}_a} P^\mu_i\right)^2 \,, \\ \mathbf{m}^2_{\top a} &= \left(\sum_{i\in\mathcal{V}_a} p^\alpha_{i\top}\right)^2 \,.\end{aligned}$$

Of course, it is established in the preceding sections B 1 a and B 1 b that $\mathbf{M}^2_a \geq \mathbf{m}^2_{\top a}$, since $\mathbf{m}_{\top a}$ could be constructed by repeated minimizations of $\mathbf{M}_a$ over the longitudinal momentum components $(p_z)_i$. Hence, $\mathbf{e}_{a\top} \geq \mathbf{e}_{\top a}$.

If we now define analogous quantities $\tilde{\mathbf{e}}_{a\top}, \tilde{\mathbf{e}}_{\top a}$ and $\vec{\mathbf{q}}_{aT}$ for the composite invisible particle, then all the same arguments apply. Forming the two mass variables as in (102) and (109),

$$\begin{aligned}M^2_{1\top} = (\mathbf{e}_{1\top} + \tilde{\mathbf{e}}_{1\top})^2 &- (\vec{\mathbf{p}}_{1T} + \vec{\mathbf{q}}_{1T})^2 \\ &\geq \\ (\mathbf{e}_{\top 1} + \tilde{\mathbf{e}}_{\top 1})^2 &- (\vec{\mathbf{p}}_{T1} + \vec{\mathbf{q}}_{T1})^2 = M^2_{\top 1} \,.\end{aligned}$$

Moving next to $M_{\circ 1}$, we note that this is simply the previous case, with an additional minimization over the masses M_i of the constituent particles, which must reduce the size of the bound, forcing $M_{\circ 1} \leq M_{\top 1}$.

For the final inequality, we recall the statement due to (B47), that says the bound is weakened (i.e. made

smaller) if we base the bound on the sum of the transverse two-vectors, rather than promoting them to $\circ$-projected (1+2)-vectors before summing. The difference is solely in the energy component – the late-partitioned $\mathbf{p}^{\alpha}_{\circ a}$ has energy component

$$\mathbf{e}_{\circ a} = \sum_{i\in\mathcal{V}_a} (p_T)_i \,,$$

whereas the early-partitioned $\mathbf{p}^{\alpha}_{a\circ}$ has energy component

$$\mathbf{e}_{a\circ} = \mathbf{p}_{aT} \,.$$

By the triangle inequality, $\mathbf{e}_{\circ a} \geq \mathbf{e}_{a\circ}$, yielding the final required result, that

$$\begin{aligned} M^2_{\circ 1} = (\mathbf{e}_{\circ 1} + \tilde{\mathbf{e}}_{\circ 1})^2 \;&-\; (\vec{\mathbf{p}}_{T1} + \vec{\mathbf{q}}_{T1})^2 \\ &\geq \\ (\mathbf{e}_{1\circ} + \tilde{\mathbf{e}}_{1\circ})^2 \;&-\; (\vec{\mathbf{p}}_{1T} + \vec{\mathbf{q}}_{1T})^2 = M^2_{1\circ} \,. \end{aligned}$$

Armed with this knowledge, we tackle the hierarchy when $N > 1$. We revisit the definitions of $M_N, M_{N\top}, M_{\top N}, M_{N\circ}$, and $M_{\circ N}$, from section VII C, as

$$\begin{aligned} M_N(\not{\mathbf{M}}) &\equiv \min_{\sum \vec{q}_{iT} = \vec{\not{p}}_T} \Big[\max_a \left[\mathcal{M}_a(\mathbf{P}_a, \mathbf{Q}_a, \tilde{\mu}_a)\right]\Big] \,, \\ M_{N\top}(\not{\mathbf{M}}) &\equiv \min_{\sum \vec{q}_{iT} = \vec{\not{p}}_T} \Big[\max_a \left[\mathcal{M}_{a\top}(\mathbf{p}_{a\top}, \mathbf{q}_{a\top}, \tilde{\mu}_a)\right]\Big] \,, \\ M_{\top N}(\not{\mathbf{M}}) &\equiv \min_{\sum \vec{q}_{iT} = \vec{\not{p}}_T} \Big[\max_a \left[\mathcal{M}_{\top a}(\mathbf{p}_{\top a}, \mathbf{q}_{a\top a}, \tilde{\mu}_a)\right]\Big] \,, \\ M_{N\circ}(\not{\mathbf{M}}) &\equiv \min_{\sum \vec{q}_{iT} = \vec{\not{p}}_T} \Big[\max_a \left[\mathcal{M}_{a\circ}(\mathbf{p}_{a\circ}, \mathbf{q}_{a\circ}, \tilde{\mu}_a)\right]\Big] \,, \\ M_{\circ N}(\not{\mathbf{M}}) &\equiv \min_{\sum \vec{q}_{iT} = \vec{\not{p}}_T} \Big[\max_a \left[\mathcal{M}_{\circ a}(\mathbf{p}_{\circ a}, \mathbf{q}_{\circ a}, \tilde{\mu}_a)\right]\Big] \,. \end{aligned}$$

At first glance, it might seem alarming that we assert $M_1 = M_{1\top}$, when M_1 seems to be built of a fully (1+3)-dimensional object $\mathcal{M}_1(\mathbf{P}_1, \mathbf{Q}_1, \tilde{\mu}_1)$. But in fact, with the components $(q_z)_i$ left free, the minimization will (for reasons identical to those in the discussion of early and late partitioning) be achieved when all the constituents of $\mathbf{Q}_1$ have equal rapidity to $\mathbf{P}_1$, meaning

$$\mathcal{M}_1(\mathbf{P}_1, \mathbf{Q}_1, \tilde{\mu}_1) \;=\; \mathcal{M}_{1\top}(\mathbf{p}_{1\top}, \mathbf{q}_{1\top}, \tilde{\mu}_1).$$

But this should apply to all N, since the only constraint on the invisibles of each parent $\mathbf{Q}_a$ is on their transverse momentum components. That is, for each of the N parents, given our inputs we will get

$$\mathcal{M}_a(\mathbf{P}_a, \mathbf{Q}_a, \tilde{\mu}_a) = \mathcal{M}_{1\top}(\mathbf{p}_{1\top}, \mathbf{q}_{1\top}, \tilde{\mu}_1),$$

and therefore we immediately see that

$$\begin{aligned} M_N(\not{\mathbf{M}}) &\equiv \min_{\sum \vec{q}_{iT} = \vec{\not{p}}_T} \Big[\max_a \left[\mathcal{M}_a(\mathbf{P}_a, \mathbf{Q}_a, \tilde{\mu}_a)\right]\Big] \\ &= \min_{\sum \vec{q}_{iT} = \vec{\not{p}}_T} \Big[\max_a \left[\mathcal{M}_{a\top}(\mathbf{p}_{a\top}, \mathbf{q}_{a\top}, \tilde{\mu}_a)\right]\Big] \\ &\equiv M_{N\top}(\not{\mathbf{M}}) \,. \end{aligned} \tag{B49}$$

Next one might ask whether the successive inequalities still hold. The very first one follows straightforwardly. Only in the input vectors to each of the N parental mass bounds $\mathcal{M}_{a\top}$ do $M_{N\top}$ and $M_{\top N}$ differ. Furthermore, since the late-partitioned input vectors $\mathbf{p}_{\top a}$ and $\mathbf{q}_{\top a}$ will have smaller energy components than their early-partitioned counterparts $\mathbf{p}_{a\top}$ and $\mathbf{q}_{a\top}$, each of the individual parental bounds follows the relation

$$\mathcal{M}_{a\top}(\mathbf{p}_{a\top}, \mathbf{q}_{a\top}, \tilde{\mu}_a) \geq \mathcal{M}_{\top a}(\mathbf{p}_{\top a}, \mathbf{q}_{\top a}, \tilde{\mu}_a) \,, \tag{B50}$$

for every possible choice of unprojected inputs $\mathbf{P}_a, \mathbf{Q}_a$.

To complete the argument, we need to establish that the global minimum considering all trial $\vec{q}_{iT}$ cannot increase if any or all of the parental bounds decrease.

The minimization probes the full space of $\{\vec{q}_{iT}\}$, subject to the constraint that their sum is the missing transverse momentum vector, with all other parameters having been specified. For the minimization to pick out a larger value for $M_{\top N}$ than for $M_{N\top}$, we must have

$$\max_a \left[\mathcal{M}_{\top a}(\mathbf{p}_{\top a}, \mathbf{q}_{\top a}, \tilde{\mu}_a)\right] > \max_a \left[\mathcal{M}_{a\top}(\mathbf{p}_{a\top}, \mathbf{q}_{a\top}, \tilde{\mu}_a)\right] \tag{B51}$$

for the same values of $\{\vec{q}_{iT}\}$ that give the value of $M_{N\top}$, if nowhere else. But we have already established (B50) for all a and all inputs. So we are led to the conclusion

$$M_{N\top}(\not{\mathbf{M}}) \geq M_{\top N}(\not{\mathbf{M}}) \,. \tag{B52}$$

Actually, we have achieved more than that. The same argument holds for the remaining levels of the hierarchy involving the $\circ$-projection. So we can boldly claim our final result and can retire to a well-deserved cuppa

$$\begin{aligned} M_N \;&=\; M_{N\top} \\ &\geq \\ &M_{\top N} \\ &\geq \\ &M_{\circ N} \\ &\geq \\ &M_{N\circ}. \end{aligned} \tag{B53}$$

Appendix C: Pronunciation guide

Following the release of the first version of this note to the arXiv, a pronunciation guide was requested. The authors do not wish to stifle innovation in this area, but tentatively suggest the formulations below.

Symbol	Pronunciation	IPA	Comment
T	tee	ti:	or 'generic tea' (yellow label)
$\top$	tee	ti:	or 'mass-preserving tea' (milky)
$\vee$	vee	vi:	
$\circ$	oh	əʊ	as in 'Oh my, cucumber sandwiches'
z	zed	zɛd	

Acknowledgments

This work is supported in part by a US Department of Energy grant DE-FG02-97ER41029, and by the Science and Technology Research Council of the United Kingdom. TJK is supported by a Dr. Herchel Smith Fellowship from Williams College. KCK is partially supported by the National Science Foundation under Award No. EPS-0903806 and matching funds from the State of Kansas through Kansas Technology Enterprise Corporation. We would like to thank W. Buttinger and B. Gripaios for useful discussions. We are grateful to Joe and Mary Ann McDonald for interrupting their Indian tiger safari to give permission for the use of the *Glaucomys volans* photograph.

[1] A. J. Barr and C. G. Lester, *A Review of the Mass Measurement Techniques proposed for the Large Hadron Collider*, *J. Phys.* **G37** (2010) 123001, [arXiv:1004.2732].
[2] W. van Neerven, J. Vermaseren, and K. Gaemers, *Lepton-jet events as a signature for W production in p anti-p collisions*, *NIKHEF-H/82-20* (Nov, 1982) 15.
[3] **UA1** Collaboration, G. Arnison *et. al.*, *Experimental observation of isolated large transverse energy electrons with associated missing energy at $s^{1/2}$ = 540 GeV*, *Phys. Lett.* **B122** (1983) 103–116.
[4] **UA2** Collaboration, M. Banner *et. al.*, *Observation of single isolated electrons of high transverse momentum in events with missing transverse energy at the CERN $\bar{p}p$ collider*, *Phys. Lett.* **B122** (1983) 476–485.
[5] J. Smith, W. L. van Neerven, and J. A. M. Vermaseren, *The transvere mass and width of the W boson*, *Phys. Rev. Lett.* **50** (1983) 1738.
[6] V. D. Barger, A. D. Martin, and R. Phillips, *Perpendicular electron neutrino mass from W decay*, *Z.Phys.* **C21** (1983) 99.
[7] C. G. Lester and D. J. Summers, *Measuring masses of semiinvisibly decaying particles pair produced at hadron colliders*, *Phys. Lett.* **B463** (1999) 99–103, [hep-ph/9906349].
[8] A. Barr, C. Lester, and P. Stephens, *m(T2) : The Truth behind the glamour*, *J. Phys.* **G29** (2003) 2343–2363, [hep-ph/0304226].
[9] C. Lester and A. Barr, *M_{TGen} : Mass scale measurements in pair-production at colliders*, *JHEP* **12** (2007) 102, [arXiv:0708.1028].
[10] B. Gripaios, *Transverse observables and mass determination at hadron colliders*, *JHEP* **02** (2008) 053, [arXiv:0709.2740].
[11] W. S. Cho, K. Choi, Y. G. Kim, and C. B. Park, *Transverse mass for pairs of gluinos*, *Phys. Rev. Lett.* **100** (2008) 171801, [arXiv:0709.0288]. © (2008) by the American Physical Society.
[12] A. J. Barr, B. Gripaios, and C. G. Lester, *Weighing WIMPs with kinks at colliders: Invisible particle mass measurements from endpoints*, *JHEP* **02** (2008) 014, [arXiv:0711.4008].
[13] W. S. Cho, K. Choi, Y. G. Kim, and C. B. Park, *Measuring superparticle masses at hadron collider using the transverse mass kink*, *JHEP* **02** (2008) 035, [arXiv:0711.4526].
[14] G. G. Ross and M. Serna, *Mass determination of new states at hadron colliders*, *Phys. Lett.* **B665** (2008) 212–218, [arXiv:0712.0943].
[15] M. M. Nojiri, Y. Shimizu, S. Okada, and K. Kawagoe, *Inclusive transverse mass analysis for squark and gluino mass determination*, *JHEP* **06** (2008) 035, [arXiv:0802.2412].
[16] W. S. Cho, K. Choi, Y. G. Kim, and C. B. Park, *Measuring the top quark mass with m_{T2} at the LHC*, *Phys. Rev.* **D78** (2008) 034019, [arXiv:0804.2185].
[17] A. J. Barr, G. G. Ross, and M. Serna, *The precision determination of invisible-particle masses at the LHC*, *Phys. Rev.* **D78** (2008) 056006, [arXiv:0806.3224].
[18] M. M. Nojiri, K. Sakurai, Y. Shimizu, and M. Takeuchi, *Handling jets + missing E_T channel using inclusive m_{T2}*, *JHEP* **10** (2008) 100, [arXiv:0808.1094].
[19] W. S. Cho, K. Choi, Y. G. Kim, and C. B. Park, *M_{T2}-assisted on-shell reconstruction of missing momenta and its application to spin measurement at the LHC*, *Phys. Rev.* **D79** (2009) 031701, [arXiv:0810.4853].
[20] H.-C. Cheng and Z. Han, *Minimal kinematic constraints and M_{T2}*, *JHEP* **12** (2008) 063, [arXiv:0810.5178].
[21] M. Burns, K. Kong, K. T. Matchev, and M. Park, *Using subsystem m_{T2} for complete mass determinations in decay chains with missing energy at hadron colliders*, *JHEP* **03** (2009) 143, [arXiv:0810.5576].
[22] A. J. Barr, A. Pinder, and M. Serna, *Precision Determination of Invisible-Particle Masses at the CERN LHC: II*, *Phys. Rev.* **D79** (2009) 074005, [arXiv:0811.2138].

[23] S.-G. Kim, N. Maekawa, K. I. Nagao, M. M. Nojiri, and K. Sakurai, *LHC signature of supersymmetric models with non-universal sfermion masses*, *JHEP* **10** (2009) 005, [arXiv:0907.4234].

[24] A. J. Barr, B. Gripaios, and C. G. Lester, *Transverse masses and kinematic constraints: from the boundary to the crease*, *JHEP* **11** (2009) 096, [arXiv:0908.3779].

[25] P. Konar, K. Kong, K. T. Matchev, and M. Park, *Superpartner Mass Measurement Technique using 1D Orthogonal Decompositions of the Cambridge Transverse Mass Variable M_{T2}*, *Phys.Rev.Lett.* **105** (2010) 051802, [arXiv:0910.3679].

[26] P. Konar, K. Kong, K. T. Matchev, and M. Park, *Dark Matter Particle Spectroscopy at the LHC: Generalizing M(T2) to Asymmetric Event Topologies*, *JHEP* **1004** (2010) 086, [arXiv:0911.4126].

[27] A. J. Barr and C. Gwenlan, *The race for supersymmetry: using M_{T2} for discovery*, *Phys. Rev.* **D80** (2009) 074007, [arXiv:0907.2713].

[28] J. Alwall, K. Hiramatsu, M. M. Nojiri, and Y. Shimizu, *Novel reconstruction technique for New Physics processes with initial state radiation*, *Phys. Rev. Lett.* **103** (2009) 151802, [arXiv:0905.1201].

[29] K. Choi, D. Guadagnoli, S. H. Im, and C. B. Park, *Sparticle masses from transverse mass kinks at the LHC: the case of Yukawa-unified SUSY GUTs*, *JHEP* **10** (2010) 025, [arXiv:1005.0618].

[30] D. R. Tovey, *On measuring the masses of pair-produced semi-invisibly decaying particles at hadron colliders*, *JHEP* **04** (2008) 034, [arXiv:0802.2879].

[31] M. Serna, *A short comparison between m_{T2} and m_{CT}*, *JHEP* **06** (2008) 004, [arXiv:0804.3344].

[32] G. Polesello and D. R. Tovey, *Supersymmetric particle mass measurement with the boost-corrected contransverse mass*, *JHEP* **03** (2010) 030, [arXiv:0910.0174].

[33] W. S. Cho, J. E. Kim, and J.-H. Kim, *Amplification of endpoint structure for new particle mass measurement at the LHC*, *Phys. Rev.* **D81** (2010) 095010, [arXiv:0912.2354].

[34] K. T. Matchev and M. Park, *A general method for determining the masses of semi-invisibly decaying particles at hadron colliders*, arXiv:0910.1584.

[35] A. J. Barr, C. Gwenlan, C. G. Lester, and C. J. S. Young, *A comment on 'Amplification of endpoint structure for new particle mass measurement at the LHC'*, arXiv:1006.2568.

[36] L. Lonnblad, *CLHEP: A project for designing a C++ class library for high-energy physics*, *Comput. Phys. Commun.* **84** (1994) 307–316.

[37] I. Antcheva *et. al.*, *ROOT: A C++ framework for petabyte data storage, statistical analysis and visualization*, *Comput. Phys. Commun.* **180** (2009) 2499–2512.

[38] A. J. Barr and C. G. Lester, "Oxbridge stransverse mass library." http://www.hep.phy.cam.ac.uk/~lester/mt2/index.html.

[39] H.-C. Cheng and Z. Han, "UCD stransverse mass library." http://particle.physics.ucdavis.edu/hefti/projects/doku.php?id=wimpmass.

[40] D. J. Castano and S. P. Martin, *Discrete symmetries and isosinglet quarks in low-energy supersymmetry*, *Phys. Lett.* **B340** (1994) 67–73, [hep-ph/9408230].

[41] H. K. Dreiner, C. Luhn, H. Murayama, and M. Thormeier, *Baryon Triality and Neutrino Masses from an Anomalous Flavor U(1)*, *Nucl. Phys.* **B774** (2007) 127–167, [hep-ph/0610026].

[42] H. K. Dreiner, C. Luhn, and M. Thormeier, *What is the discrete gauge symmetry of the MSSM?*, *Phys. Rev.* **D73** (2006) 075007, [hep-ph/0512163].

[43] H.-S. Lee, C. Luhn, and K. T. Matchev, *Discrete gauge symmetries and proton stability in the U(1)'-extended MSSM*, *JHEP* **07** (2008) 065, [arXiv:0712.3505].

[44] K. Agashe, D. Kim, M. Toharia, and D. G. E. Walker, *Distinguishing Dark Matter Stabilization Symmetries Using Multiple Kinematic Edges and Cusps*, arXiv:1003.0899.

[45] K. Agashe, D. Kim, D. G. E. Walker, and L. Zhu, *Using M_{T2} to Distinguish Dark Matter Stabilization Symmetries*, arXiv:1012.4460.

[46] P. Konar, K. Kong, K. T. Matchev, and M. Park, *RECO level $\sqrt{s}_{min}$ and subsystem $\sqrt{s}_{min}$: improved global inclusive variables for measuring the new physics mass scale in missing energy events at hadron colliders*, arXiv:1006.0653.

[47] M. M. Nojiri and K. Sakurai, *Controlling ISR in sparticle mass reconstruction*, *Phys. Rev.* **D82** (2010) 115026, [arXiv:1008.1813].

[48] D. Krohn, L. Randall, and L.-T. Wang, *On the Feasibility and Utility of ISR Tagging*, arXiv:1101.0810.

[49] I. Hinchliffe, F. E. Paige, M. D. Shapiro, J. Soderqvist, and W. Yao, *Precision SUSY measurements at CERN LHC*, *Phys. Rev.* **D55** (1997) 5520–5540, [hep-ph/9610544].

[50] P. Konar, K. Kong, and K. T. Matchev, *$\sqrt{\hat{s}}_{min}$: A global inclusive variable for determining the mass scale of new physics in events with missing energy at hadron colliders*, *JHEP* **03** (2009) 085, [arXiv:0812.1042].

[51] C. G. Lester, *The stransverse mass, MT2, in special cases*, arXiv:1103.5682.

[52] D. R. Tovey, *Transverse mass and invariant mass observables for measuring the mass of a semi-invisibly decaying heavy particle*, *JHEP* **11** (2010) 148, [arXiv:1008.3837].

[53] **Particle Data Group** Collaboration, K. Nakamura *et. al.*, *Review of particle physics*, *J. Phys.* **G37** (2010) 075021.

[54] I.-W. Kim, *Algebraic singularity method for mass measurement with missing energy*, *Phys. Rev. Lett.* **104** (2010) 081601, [arXiv:0910.1149].

[55] V. D. Barger, T. Han, and R. J. N. Phillips, *Improved transverse mass variable for detecting higgs boson decays into z pairs*, *Phys. Rev.* **D36** (1987) 295.

[56] V. D. Barger, T. Han, and J. Ohnemus, *Heavy leptons at hadron supercolliders*, *Phys. Rev.* **D37** (1988) 1174.

[57] A. J. Barr, B. Gripaios, and C. G. Lester, *Measuring the Higgs boson mass in dileptonic W-boson decays at hadron*

colliders, *JHEP* **07** (2009) 072, [arXiv:0902.4864].

[58] A. Katz, M. Son, and B. Tweedie, *Jet Substructure and the Search for Neutral Spin-One Resonances in Electroweak Boson Channels*, *JHEP* **03** (2011) 011, [arXiv:1010.5253].

[59] D. L. Rainwater and D. Zeppenfeld, *Observing $H \to W^{(*)}W^{(*)} \to e^{\pm}\mu^{\mp}\not{p}_T$ in weak boson fusion with dual forward jet tagging at the CERN LHC*, *Phys. Rev.* **D60** (1999) 113004, [hep-ph/9906218].

[60] T. Sjostrand, S. Mrenna, and P. Z. Skands, *PYTHIA 6.4 Physics and Manual*, *JHEP* **05** (2006) 026, [hep-ph/0603175].

[61] M. Cacciari and G. P. Salam, *Dispelling the N^3 myth for the k_t jet-finder*, *Phys. Lett.* **B641** (2006) 57–61, [hep-ph/0512210].

[62] J. Conway, "PGS - pretty good simulation." http://www.physics.ucdavis.edu/~conway/research/software/pgs/pgs4-general.htm.