1 Invariant mass

Consider a 4-vector

$$p_\mu = \{E, p_x, p_y, p_z\} = \{E,\ p\sin\theta\sin\phi,\ p\sin\theta\cos\phi,\ p\cos\theta\}$$

where $E^2 = m^2 + \bar{p}^2 = m^2 + p_x^2 + p_y^2 + p_z^2$. Despite this may look very familiar it is a fact that this cartesian and the spherical parametrization of p_μ do not make manifest the cylindrical symmetry around the beam axis. Therefore for many practical purposes in collider experiments p_μ is more conveniently written in terms of quantities that better show the cylindrical invariance of the problem. To this goal we define the z axis to be along the beam direction. The angular part of the 4-vector is described by two quantities, the azimuthal angle ϕ and the rapidity y. The angle ϕ is just denoting the direction in the $x - y$ plane. The angular information in y instead is interleaved with the energy information. In fact

$$y = \frac{1}{2}\log\frac{E + p_z}{E - p_z} = \frac{1}{2}\log\frac{E + p\cos\theta}{E - p\cos\theta}\,, \tag{1}$$

where θ is the polar angle of the cylindrical coordinates defined around the beam axis. Despite this parametrization obscured a bit the polar-angular information it has the great advantage to involve only quantities that have simple transformation properties under longitudinal boost. In fact the rapidity y transforms additively under longitudinal boosts

$$y \to y + y_{boost},\ \text{where}\ \cosh y_{boost} = \gamma,\ \sinh y_{boost} = \gamma\beta. \tag{2}$$

The energy information of the four-vector is encoded in the transverse momentum

$$p_T = \sqrt{p_x^2 + p_y^2}$$

and the mass of the particle, that we denote by m. All in all we can express a four-vector in cartesian coordinates as a function of y, ϕ, p_T, m as follows:

$$p_\mu(y, \phi, p_T, m) = \left\{\cosh(y)\sqrt{m^2 + p_T^2}, p_T\sin(\phi), p_T\cos(\phi), \sinh(y)\sqrt{m^2 + p_T^2}\right\}\,.$$

In these coordinates one can easily write the mass constraint for the decay of a particle a into the two decay products bc

$$a \to bc$$

that is

$$p_\mu^{(a)} = p_\mu^{(b)} + p_\mu^{(c)}\,.$$

From the conservation of four-momentum it follows that parametrizing the decay products four-momenta as

$$\begin{aligned} p_\mu^{(b)} &= \left\{\cosh\left(y_b\right)\sqrt{p_{T,b}^2+m_b^2}, \sin\left(\phi_b\right)p_{T,b}, \cos\left(\phi_b\right)p_{T,b}, \sinh\left(y_b\right)\sqrt{p_{T,b}^2+m_b^2}\right\} \qquad (3)\\ p_\mu^{(c)} &= \left\{\cosh\left(y_b+\Delta \mathrm{y}\right)\sqrt{p_{T,c}^2+m_c^2}, p_{T,c}\sin\left(\phi_b+\Delta\phi\right), p_{T,c}\cos\left(\phi_b+\Delta\phi\right), \sinh\left(y_b+\Delta \mathrm{y}\right)\sqrt{p_{T,c}^2+m_c^2}\right\}\end{aligned}$$

the invariant mass constraint is

$$\begin{aligned} m_a^2 &= p^{(a)}\cdot p^{(a)} = \left(p_\mu^{(b)}+p_\mu^{(c)}\right)^2 \\ &= 2\cosh\Delta \mathrm{y}\sqrt{p_{T,b}^2+m_b^2}\sqrt{p_{T,c}^2+m_c^2} - 2\cos\Delta\phi\, p_{T,b}p_{T,c} + m_b^2 + m_c^2\,. \qquad (4)\end{aligned}$$

2 Transverse Mass

In absence of information about the longitudinal components of the four-vectors useful information on the decay can still be extracted by restricting our attention to the accessible quantities, that are in general energies and momenta measured in the transverse plane.

For some quantities it is obvious how to neglect the p_z information, for instance the three-momentum $\bar{p}$ can be simply projected to the transverse plane by doing

$$\{p_x, p_y, p_z\} \to \{p_x, p_y\}\,.$$

However for quantities such as the energy of a four-vector the projection is less obvious. In fact the energy is by definition sensitive to all the components of the four-vector. One can choose to project the energy in a way that carries to the transverse sub-space different properties of the full four-vector. A commonly adopted projection, but not the only one existing in the literature (see [1] for an overview), is the projection that maintains the norm of the vector, *i.e.* mass of the particle. To this end one defines a projected (1+2)-vector $\tilde{p}_\alpha$

$$\tilde{p}_\alpha(\phi, p_T, m) = \left\{\sqrt{m^2+p_T^2}, p_T\sin(\phi), p_T\cos(\phi)\right\}\,,$$

which, by the choice we made for the temporal component

$$E = \sqrt{m^2+p_x^2+p_y^2+p_z^2} \to E_T = \sqrt{m^2+p_x^2+p_y^2}\,,$$

has the property to have norm square m^2 when dotted with the (1+2)-metric $\eta_T = \{1,-1,-1\}$

$$\tilde{p}\cdot\tilde{p} = m^2\,.$$

Operating this "mass preserving" projection on the four vectors $p_\mu^{(b)}$ and $p_\mu^{(c)}$ in eq.(3) we get two projected (1+2)-vectors $\tilde{p}_\alpha^{(b)}, \tilde{p}_\alpha^{(c)}$ which conserve the mass of the particle b and c, respectively.

These two (1+2)-vectors can be used to construct a quantity analogous to the one in eq.(4) [2]

$$\begin{aligned}\tilde{m}_{bc}^2 &= \left(\tilde{p}_\alpha^{(b)} + \tilde{p}_\alpha^{(c)}\right)^2 \\ &= 2\sqrt{p_{T,b}^2 + m_b^2}\sqrt{p_{T,c}^2 + m_c^2} - 2\cos(\Delta\phi) p_{T,b} p_{T,c} + m_b^2 + m_c^2 \leq m_a^2\end{aligned} \tag{5}$$

where the inequality follows from $\cosh \Delta y \geq 1$ and a direct comparison with eq.(4). From this it follows that

$$\tilde{m}_{bc} = m_{bc} \quad \text{for } y_b = y_c \,.$$

Since we are considering quantities projected on the transverse plane it is expected that in general the mass of the parent particle a cannot be entirely seen because it is partially employed to give b and c some motion along the z direction. In other words, by restricting our attention to the transverse plane, in general we are throwing away information on the global energy balance of the process. This is particularly clear if one looks at a concrete example.

For simplicity we take the case of the production of a single parent particle produced in a hadronic collision, followed by the decay into massless particles. This was in fact the first situation for the application [3] of the quantity $\tilde{m}_{bc}$ defined in eq.(5) to the case of the process

$$p\bar{p} \to W \to \ell\nu \,.$$

For the case of massless decay products b and c eq.(5) reduces to

$$m_{T,\ell\nu}^2 = 2 p_{T,\nu} p_{T,\ell} (1 - \cos\Delta\phi) \leq m_W^2 \,. \tag{6}$$

Assuming that the W is moving along the z direction we can boost the system back to the center of mass of the W. In the rest frame of the W boson we can define θ^*, the angle of the charged lepton with respect to the z direction. Therefore we have

$$p_{T,\ell} = E_\ell \sin\theta^* = \frac{m_W}{2} \sin\theta^* \,,$$

which is valid both in the rest frame of the W and in the center of mass of the $p\bar{p}$ laboratory. To saturate the inequality eq.(6) we need to take $\sin\theta^* = 1$ such as

$$p_{T,\nu} = p_{T,\ell} = \frac{1}{2} m_W \text{ and } \Delta\phi = \pi \,,$$

which corresponds to the emission of a lepton and a neutrino each with zero momentum in the longitudinal direction. This is precisely a situation where we expect the projection to the transverse plane to not change the global energy balance. In fact the decay has no longitudinal momentum from the very start, therefore is natural to saturate eq.(6).

More in general, regardless of the mass of the decay products and of the direction of motion of the parent particle, the events where the transverse mass is close to its maximum are events where the daughter particles can be boosted to a frame where they both are on the transverse plane, *i.e.*

$$p_z = 0$$

for both particles. This is guaranteed by the fact that they have same rapidity in the laboratory frame and the transformation of the rapidity of a four-momentum under a boost as given in eq.(2). For massless particles this is the same of saying that the decay products have some degree of collinearity, as implied by the fact that $\Delta y = 0$ implies that two massless particles have the same polar direction θ.

3 The "stransverse" mass m_{T2}

In presence of more than one invisible particle it no longer possible to define a transverse mass by simply projecting the momenta on the transverse plane and then computing a restricted Minkowski norm.

In fact in a process with at least two invisible final states there are new qualitative features of the kinematics that undermine the construction of m_T in a unambiguous way. A first limitation that we encounter is that in presence of two or more invisible particles we have experimental access only to one combination of their transverse momenta, that is vectorial sum of all the invisible particles

$$\bar{i}_T = \sum_{i=invisibles} \bar{p}_{T,i} = - \sum_{v=visibles} \bar{p}_{T,v} \, .$$

All the other combinations of the transverse momenta of the invisible particles are inaccessible at the experiment. On top of the increased number of non-measurable kinematical quantities we have to cope with another new feature. A system of invisible particles is different from a single invisible particle in that it does not posses a fixed mass, but rather its mass is different even-by-event. Hence we cannot define an analogous of the transverse mass $\tilde{m}$ of eq.(5) because the "constant" mass of the invisible particle would not be constant over the events collected by the experiment.

To overcome both these difficulties in certain situations we can make an ansatz. A common case at colliders is that of the pair production of states that decay in visible particle plus one invisible. A familiar example in the SM is the production of a pair of W boson

$$pp \to WW \to \ell\nu\ell\nu$$

but there are many other examples, especially among the signatures of SUSY and in processes involving new particles that belong to the sector that encompasses the Dark Matter particle

$$pp \to XX \to Y\chi Y\chi. \tag{7}$$

In a fully general model there can be any number of invisible particles. However in most models there are just two invisible particles per event, which adds up to the simplicity of this kind of final state as a motivation to start attacking the case of two invisibles. We can imagine that the invisible particles are two particles $\chi_{1,2}$ with identical mass m_χ, carrying transverse momentum p_{T,χ_1} and p_{T,χ_2} such that

$$\bar{p}_{T,\chi_1} + \bar{p}_{T,\chi_2} = \bar{\not{i}}_T \,.$$

Armed with this ansatz we can attack the case where Y is a single massless particle. We compute two transverse masses $\tilde{m}_{Y_1\chi_1}$ and $\tilde{m}_{Y_2\chi_2}$ using the formula eq.(5). For the properties of the transverse mass, if the ansatz was correct, each of these transverse masses would be a lower bound for m_X

$$\max\left(\tilde{m}_{Y_1\chi_1},\, \tilde{m}_{Y_2\chi_2}\right) < m_X \,.$$

Computational power permitting, we are free to vary our ansatz and determine the minimum of the "would-be" lower-bounds, which is

$$m_{T2} \equiv \min_{ansatz\, on\, p_{T,\chi}} \left(\max\left(\tilde{m}_{Y_1\chi_1},\, \tilde{m}_{Y_2\chi_2}\right)\right), \tag{8}$$

sometimes referred to as "stransverse" mass, for its origin in the context of SUSY spectrum mass measurements and SUSY discovery.

In general the transverse mass, and therefore m_{T2}, are functions of the mass of the invisible particle m_χ. This quantity is typically not known and in some situations it is crucial to know it to make use of m_{T2}. This is typically the case when one tries to use m_{T2} to measure the mass of the particles involved in the decay $X \to Y\chi$.

However, some information on the kinematics of the event can be reconstructed even when the mass m_χ relevant for the computation of m_{T2} is not know. In this respect the events with large m_{T2} enjoy particular properties, hence it makes sense to define

$$m_{T2}^{(max)} = \max_{events} m_{T2} \,. \tag{9}$$

In fact for some simple cases the extremization of eq.(9) has a solution in closed form as a function of the masses involved in the process. For the case of the process in eq.(7), the fact that we have two identical decay chains that produce each one visible and one invisible particle can be exploited to find that

$$m_{T2}^{(max)}(m_{\chi,trial}) = C + \sqrt{C^2 + m^2_{\chi,trial}} \,. \tag{10}$$

The events that populate the region of $m_{T2} \simeq m_{T2}^{(max)}$ are of course events where the transverse masses $\tilde{m}_{Y_1\chi_1}$, $\tilde{m}_{Y_2\chi_2}$ are large, that is to say events where the inequality eq.(5) is close to be saturated. As discussed right below eq.(5), this is happening because the invisible particles have the same rapidity of the visible particles with which the transverse mass is computed.

In cases where the process is more complicated than eq.(7) one can still rely on the fact that the visible particles can be clustered into a single object with four-momentum

$$P_{\mu,vis} = \sum_{v=visibles} p_{\mu,v}$$

and similarly for the invisible particles

$$P_{\mu,inv} = \sum_{i=invisibles} p_{\mu,i} \,.$$

In general these four-vectors do not respect any mass-shell relation because they are made of many particles physical four-momenta, that is to say $P_{inv}^2 = M_{cluster,inv}^2$ varies in every event. However, $M_{cluster,inv}$ is a sort of "effective" mass of the compound system of invisible particles, therefore can vary only within a range. In fact it cannot exceed the mass of the parent particle and has to be greater than the sum of the masses of the constituents invisible particles that we have clustered

$$M_{cluster,inv} \geq \sum_{i=invisibles} m_i \qquad \text{and} \qquad M_{cluster,inv} < m_X \,.$$

The situation where $M_{cluster,inv}$ is minimal is realized when there is a frame where the involved invisible particles are produced at rest, which is the same of saying that in all other frames they will have the same rapidity. This is a condition similar to the one that we encountered when discussing the maximization of m_T, however notice that here we are requiring it on two invisible particles, whereas for m_T one visible and one invisible were involved.

It is intuitive to see that for small effective mass of the invisible particles there is more momentum available in the visible system and therefore it is possible to attain larger m_T and m_{T2}. Motivated by the above reasoning one can compute m_{T2} under the naïve assumption [4] that only one invisible per decay chain has been produced. At this point it should not be surprising that the maximal m_{T2} that can be attained has the same form eq.(10) that is valid for a truly two body decay [4].

The fall-off of the probability to reach the maximum, of course, will reflect the fact for a compound system to give the maximal m_{T2} a number of alignments (in rapidity space, more in general, or in actual space for massless particles) among the invisible is required. The sensitivity of the slope of the m_{T2} distribution to the number of invisible particles makes it an interesting "counter" for the number of invisible particles produced [5].

The fact that the maximal m_{T2} is attained when the invisible particles line-up in rapidity implies a certain alignment in polar angle θ as well. This alignment is exact when the invisible particles are massless. As well known, the sum of two collinear massless four-vector is still a massless four-vector, therefore in such aligned situation it looks natural to have a convergence of the kinematics of many invisibles particles towards a single-invisible-like kinematics. In light of this fact the fact that the end-point of the m_{T2} has the same functional form eq.(10) independently of how many invisibles are present is somewhat expected.

4 Additional Material

For a general introduction to collider physics see [6] and also [7] for related fundamental QCD topics. For a shorter introduction [8] is a good starting point. A concise introduction about the use of transverse variables at colliders can be found in [9]. More in depth reviews are [10] and [1]. The original paper about m_{T2} is also a good insightful reading [11]. Examples of what can be done in cases where invisible particles are produced in association with more than one visible particle per decay are nicely described in [12] and [13].

References

[1] A. J. Barr, T. J. Khoo, P. Konar, K. Kong, C. G. Lester, K. T. Matchev, and M. Park, "Guide to transverse projections and mass-constraining variables," *Phys. Rev. D* **84** no. 9, (Nov., 2011) 095031, `arXiv:1105.2977 [hep-ph]`.

[2] A. Barr, C. Lester, and P. Stephens, "A variable for measuring masses at hadron colliders when missing energy is expected m_{T2}: the truth behind the glamour," *Journal of Physics G Nuclear Physics* **29** (Oct., 2003) 2343–2363, `arXiv:hep-ph/0304226`.

[3] J. Smith, W. van Neerven, and J. Vermaseren, "The transverse mass and width of the W boson," *Phys.Rev.Lett.* **50** (1983) 1738.

[4] K. Agashe, D. Kim, D. G. E. Walker, and L. Zhu, "Using M_{T2} to distinguish dark matter stabilization symmetries," *Phys. Rev. D* **84** no. 5, (Sept., 2011) 055020, `arXiv:1012.4460 [hep-ph]`.

[5] G. F. Giudice, B. Gripaios, and R. Mahbubani, "Counting dark matter particles in LHC events," *Phys. Rev. D* **85** no. 7, (Apr., 2012) 075019, `arXiv:1108.1800 [hep-ph]`.

[6] V. D. Barger and R. Phillips, *Collider Physics*. Addison-Wesley (Frontiers in Physics, 71), 1987.

[7] R. K. Ellis, W. J. Stirling, and B. Webber, "QCD and collider physics," *Camb.Monogr.Part.Phys.Nucl.Phys.Cosmol.* **8** (1996) 1–435.

[8] T. Han, "Collider Phenomenology:. Basic Knowledge and Techniques," in *Physics in $D \geq 4$ TASI 2004*, J. Terning, C. E. M. Wagner, and D. Zeppenfeld, eds., pp. 407–454. July, 2006. `arXiv:hep-ph/0508097`.

[9] B. Gripaios, "Tools for Extracting New Physics in Events with Missing Transverse Momentum," *International Journal of Modern Physics A* **26** (2011) 4881–4900, `arXiv:1110.4502 [hep-ph]`.

[10] A. J. Barr and C. G. Lester, "A review of the mass measurement techniques proposed for the Large Hadron Collider," *Journal of Physics G Nuclear Physics* **37** no. 12, (Dec., 2010) 123001, `arXiv:1004.2732 [hep-ph]`.

[11] C. G. Lester and D. J. Summers, "Measuring masses of semi-invisibly decaying particle pairs produced at hadron colliders," *Physics Letters B* **463** (Sept., 1999) 99–103, `arXiv:hep-ph/9906349`.

[12] W. S. Cho, D. Kim, K. T. Matchev, and M. Park, "Cracking the dark matter code at the LHC," *ArXiv e-prints* (June, 2012) , `arXiv:1206.1546 [hep-ph]`.

[13] L. Edelhauser, W. Porod, and R. K. Singh, "Spin Discrimination in Three-Body Decays," *JHEP* **1008** (2010) 053, `arXiv:1005.3720 [hep-ph]`.