Mass Reconstruction for High Multiplicity Final States Using the Boundary of Phase Space

Matthew D. Klimek[*] **and Yuan-Pao Yang**

Theory Group, Department of Physics and Texas Cosmology Center, The University of Texas at Austin, Austin, TX 78712 U.S.A.

E-mail: klimek@physics.utexas.edu, yjp1986@utexas.edu

The lack of conclusive evidence for new physics in Run I of the LHC suggests that future discoveries may manifest themselves with small numbers of signal events. In this case, it will be crucial to use analysis techniques that extract as much information as possible from a limited number of events. Previously, a technique exploiting correlations in the full multi-particle phase space to efficiently reconstruct intermediate and invisible masses in decay chains has been demonstrated for the case of four final state particles. We review this technique and discuss its generalization to five or more final state particles. We also demonstrate a new technique for implementing a mass finding algorithm based on the same principle in the presence of realistic backgrounds with Voronoi tessellations.

UTTG-26-16

---

[*]Speaker.

1. Introduction

Given the lack of any discovery beyond the Higgs boson at the LHC, any future discovery, even in the high luminosity phase, is likely to come with a small number of signal events. Therefore, extracting all available information from a limited sample of events will be a top priority.

Many promising discovery channels include the production of colored partners. If the sector containing these particles is stabilized by some symmetry, the resulting cascade decays will contain a stable, invisible particle in the final stage. Traditional mass determination methods for decays with invisible particles rely on the measurement of the endpoints of 1-dimensional kinematic distributions. These methods work because the shape of the boundary of phase space for decays is a function only of the masses involved in the decay.[1] However, there are many events that lie close to the boundary but that would *not* be near the endpoint of any of a set of 1-dimensional distributions. Additionally, the location of many endpoints is set by the differences between masses in the decay, and have very weak dependence on the absolute mass scale. This translates into poor sensitivity of endpoint-based methods to the mass scale.

A simple illustration of this for the Dalitz plot for 3-body decays is shown in figure 1. There are many events outside the boundary specified by some incorrect mass hypothesis (dashed line), but the endpoints given by the projection of the incorrect hypothesis are consistent with the data. Moreover, the locations of the endpoints depend only on the differences between the parent mass M and daughter masses m_i. Any spectrum with the same mass differences would be indistinguishable to an endpoint study.
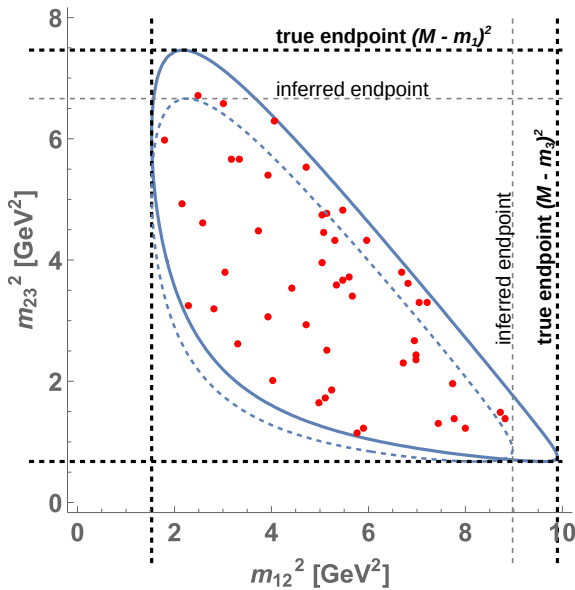
Figure 1: Dalitz plot with simulated data from a 3-body decay. The true boundary of phase space is shown in solid blue. All data points are contained within the true boundary. A boundary corresponding to an incorrect mass hypothesis is shown in dashed blue.

We will describe how to use the full boundary of phase space to determine masses with much better efficiency and sensitivity to absolute mass scale than is possible with 1-d distributions when there are at least 4 final state particles, summarizing work presented in [1, 2]. We will apply our technique to events where the cascade decay takes place on one side of the event, while the other

---

[1]Although the distribution of the decay events in phase space depends on the matrix element and spins of the particles, these factors do not change the shape of the phase space boundary.

side includes only an invisible partner, such as gluino-LSP associated production in the MSSM. The same technique can be used in a symmetric event at the cost of combinatoric uncertainties. To demonstrate the principle of our method, we also ignore uncertainties associated with the presence of backgrounds. In section 4, we will briefly discuss other methods that are being developed to deal with these complications.

2. Phase Space Boundary Enhancement

The standard form of the phase space volume element is expressed as a function of individual components of 4-momenta, which are not manifestly Lorentz invariant. A less well-known formulation expressed purely in terms of Lorentz scalars can be constructed as follows [3].

For an n-body decay, we define M_n as the $n \times n$ matrix with elements $p_i \cdot p_j$ and Δ_i as the coefficients of the characteristic polynomial of M_n, that is, $\text{Det}\left[\lambda 1_{n \times n} - M_n\right] \equiv \lambda^n - \left(\sum_{i=1}^{n} \Delta_i \lambda^{n-i}\right)$. The kinematically accessible region of phase space corresponds to $\Delta_{1,2,3} > 0$, $\Delta_4 \geq 0$ and $\Delta_{5,\dots,n} = 0$ [3]. For $n \geq 4$, the phase space volume element locally has the form [3, 2]

$$dPS_n = (\text{const.}) \times M_X^{-2} \left(\prod_{i<j}^{n} dm_{ij}^2\right) \Delta_4^{-(n-3)/2} \Theta(\Delta_4) \left(\prod_{\alpha \leq \beta}^{n-4} \delta(e_{\alpha\beta})\right) \delta\left(\sum_{i<j}^{n} m_{ij}^2 - \text{const.}\right). \quad (2.1)$$

Here M_X is the parent mass, the final δ-function enforces energy conservation, and the $e_{\alpha\beta}$ are a set of constraints that are linear in all m_{ij}^2 to first order. From this expression, we see that the phase space density scales as $\Delta_4^{-(n-3)/2}$, which results in an enhancement near the boundary $\Delta_4 = 0$ which becomes stronger with increasing n. The forms of the $e_{\alpha\beta}$ are complicated, which makes this expression difficult to use in practice. This expression can be rearranged into a symmetrical form, at the expense of introducing non-linear arguments into the δ-functions, as

$$dPS_n = (\text{const.}) \times M_X^{-2} \left(\prod_{i<j}^{n} dm_{ij}^2\right) \Delta_4^{(n-5)/2} \Theta(\Delta_4) \delta(\Delta_5) \cdots \delta(\Delta_n) \delta\left(\sum_{i<j}^{n} m_{ij}^2 - \text{const.}\right). \quad (2.2)$$

We have seen that when expressed in terms of Lorentz scalars m_{ij}^2, the most useful events for determining the masses (i.e., for finding the boundary) are the most plentiful. Fitting a boundary to the data, even with a limited number of events, provides an efficient way to measure the masses.

3. Boundary Fitting with Likelihood Functions

In order to demonstrate the efficacy of our method versus traditional endpoint or similar 1-d distribution methods, we will introduce quality-of-fit functions for each that can be used to find the masses that provide the best fit to a data sample. We find the best-fit masses for many 100-event samples of Monte Carlo data with fixed true masses, and plot the corresponding histograms to illustrate the performance of the two methods.

We define the measured position of a kinematic endpoint as the highest value obtained within the data sample for the observable in question. The best-fit mass hypothesis for the endpoint method will minimize the quality-of-fit function

$$Q = \Xi \sum_{i=\text{endpoints}} \left(\frac{\mathscr{O}_i^{\text{predicted}} - \mathscr{O}_i^{\text{measured}}}{\mathscr{O}_i^{\text{measured}}}\right)^2, \quad (3.1)$$

2

where $\Xi \to \infty$ if any one of the measured endpoints exceeds the predicted value calculated according to the hypothesis, so that it is rejected, and otherwise $\Xi = 1$. The expressions for all endpoints used in our analysis are listed in [1, 2].

Turning to the full boundary method, since the phase space density becomes large near the boundary of the accessible region, the likelihood function should favor hypotheses for which as many events as possible lie near the boundary, with no events outside. To quantify the quality-of-fit, we introduce an event-wise likelihood function $\mathscr{L}_{\text{event}}(\{m_{ij}^2\}|\{M_i\})$ which is equal to the normalized probability density for the event with data $\{m_{ij}^2\}$ calculated as if the mass hypothesis $\{M_i\}$ were correct. The total likelihood for a given sample is

$$\mathscr{L}(\text{data}|\{M_i\}) = \prod_{\text{events}} \mathscr{L}_{\text{event}}(\{m_{ij}^2\}|\{M_i\}). \tag{3.2}$$

The details of the calculation of the $\mathscr{L}_{\text{event}}$ are described in [1, 2]. As for the endpoint method, hypotheses for which there exist events in the sample that fall outside the allowed region calculated according to the hypothesis are excluded. The best-fit hypothesis is the one that maximizes $\mathscr{L}$.

As mentioned in the introduction, one of the primary disadvantages of the endpoint method is its poor sensitivity to the absolute mass scale. For this reason, we present our results in a basis $\{\alpha, \beta, \gamma, \ldots\}$ where α parametrizes a shift of all masses equally, and the others are orthogonal linear combinations of the masses that encode the mass differences.

As an example, our results for spectrum 1 of the 2+2+3 topology as listed in [2] are shown in figure 2. The results for other spectra and topologies are presented in detail in [1, 2]. It is easy to see that the multidimensional phase space method yields both more precise and more accurate results for the overall mass scale as well as for the mass differences.
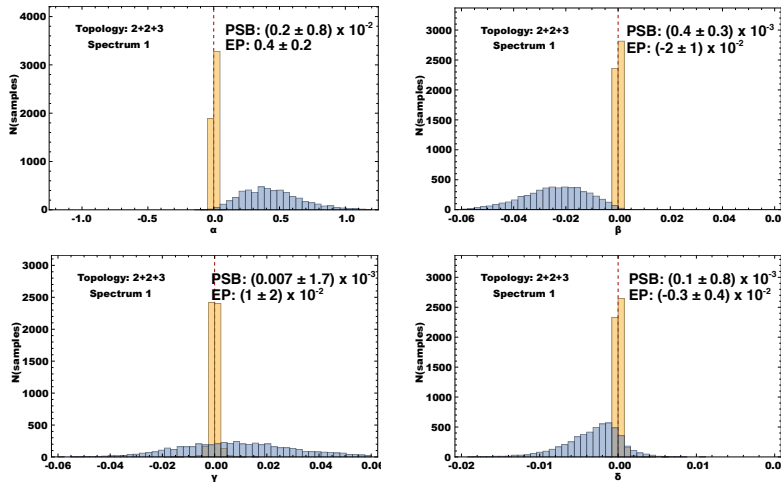
Figure 2: Distribution of the best-fit values of α, β, γ and δ (see discussion in text) for the endpoint method (EP; blue) and the phase space boundary method (PSB; yellow), using the first benchmark spectrum for the 2+2+3 topology of [2] and data samples of 100 events. The true masses correspond to α, β, γ and δ all being zero. The mean and standard deviations are shown.

4. Boundary Fitting with Voronoi Tessellation

It is worth emphasizing that all of the mass information is encoded by the shape of the boundary regardless of the distribution of events in the interior. As such, the likelihood method described above, while easy to implement, is not ideal, since it is also fitting the interior distribution by making an assumption about the decay matrix element and final state spins. We have also ignored

the complications associated with combinatoric uncertainties in 2-sided events and the presence of backgrounds. Methods are being developed to address these issues by fitting the shape of the boundary directly, even in the presence of backgrounds and making no reference to the distribution of events within, with Voronoi tessellation.

Given a space populated with a set of events, Voronoi tessellation divides the space into cells such that each cell contains the points which are nearest to a given event than to any other. The sharp concentration of signal events on the phase space boundary ensures that cells on the boundary will tend to have smaller volumes than many of their neighbors that do not straddle the boundary. Therefore, boundary cells may be selected by the large standard deviation of their neighbors' volumes.

For squark decays $\tilde{q} \to jl_n l_f \tilde{\chi}^0$ proceeding through a cascade of 2-body decays it has already been demonstrated [4] that mass measurement is possible even in the presence combinatoric and simple Standard Model backgrounds by using Voronoi tessellation, as illustrated in fig. 3.
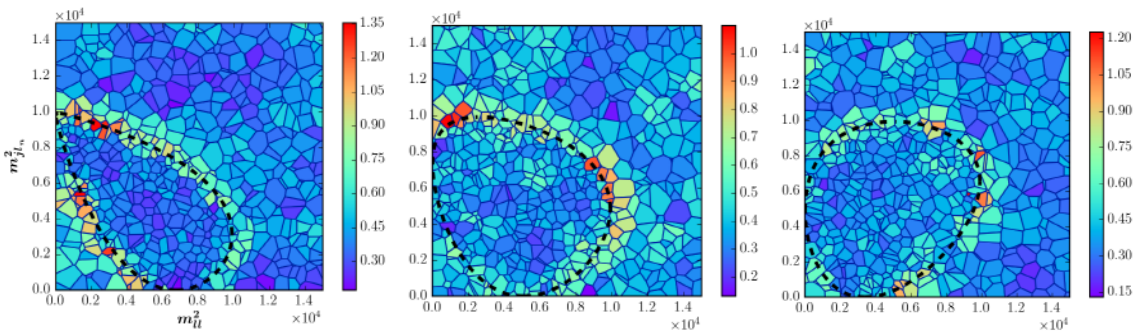
Figure 3: Two-dimensional slices of the phase space of a squark cascade decay plane at a fixed value of $m_{jl_f}^2 = (8000, 6000, 4000)$ GeV^2 (left to right). Cells are colored according to the standard deviation of the volumes of their neighbors as defined in [4]. The true location of the boundary is shown in dashed lines.

Acknowledgments

References

[1] P. Agrawal, C. Kilic, C. White, and J.-H. Yu, "Improved mass measurement using the boundary of many-body phase space," *Phys. Rev.* **D89** no. 1, (2014) 015021, arXiv:1308.6560 [hep-ph].

[2] B. Altunkaynak, C. Kilic, and M. D. Klimek, "Multi-Dimensional Phase Space Methods for Mass Measurements and Decay Topology Determination," arXiv:1611.09764 [hep-ph].

[3] N. Byers and C. N. Yang, "Physical regions in invariant variables for *n* particles and the phase-space volume element," *Rev. Mod. Phys.* **36** (Apr, 1964) 595–609. http://link.aps.org/doi/10.1103/RevModPhys.36.595.

[4] D. Debnath, J. S. Gainer, C. Kilic, D. Kim, K. T. Matchev, and Y.-P. Yang, "Identifying Phase Space Boundaries with Voronoi Tessellations," *Eur. Phys. J.* **C76** no. 11, (2016) 645, arXiv:1606.02721 [hep-ph].