Three-particle templates for a boosted Higgs boson

Leandro G. Almeida, Ozan Erdoğan, José Juknevich, Seung J. Lee, Gilad Perez, and George Sterman

Three-particle templates for boosted Higgs

Leandro G. Almeida[a], Ozan Erdoğan[b], José Juknevich[c], Seung J. Lee[d], Gilad Perez[c,e],
George Sterman[b]

[a] *Institut de Physique Theorique, CEA-Saclay, F-91191, Gif-sur-Yvette cedex, France*

[b] *C.N. Yang Institute for Theoretical Physics*
Stony Brook University, Stony Brook, New York 11794-3840, USA

[c] *Department of Particle Physics and Astrophysics*
Weizmann Institute of Science, Rehovot 76100, Israel

[d] *Department of Physics, Korea Advanced Institute of Science and Technology*
335 Gwahak-ro, Yuseong-gu, Daejeon 305-701, Korea

[e] *CERN, Theory Division, CH1211 Geneva 23, Switzerland*

Abstract

We explore the ability of three-particle templates to distinguish color neutral objects from QCD background. This method is particularly useful to identify the standard model Higgs, as well as other massive neutral particles. Simple cut-based analysis in the overlap distributions of the signal and background is shown to provide a significant rejection power. By combining with other discriminating variables, such as planar flow, and several variables that depend on the partonic template, three-particle templates are used to characterize the influence of gluon emission and color flow in collider events. The performance of the method is discussed for the case of a highly boosted Higgs in association with a leptonically-decaying W boson.

1   Introduction

Processes with top quarks, standard model Higgs and W/Z jets play a key role in high energy collisions. They allow for tests of perturbative QCD and are important backgrounds for more exotic phenomena. Over the past few years, scenarios have been proposed in which these heavy particles are produced at large transverse momentum [1–11]. At high enough p_T, their decay products will appear as heavy, collimated jets [12, 13]. Even such exotic final states,

however, will coexist with a substantial tail of the mass distribution of light-parton QCD jets [14, 15]. Thus it will generally be necessary to study jet substructure systematically to distinguish such a signal.

Precise predictions for the substructure of such boosted jets are severely limited by the complexity of the final state: in general it will be difficult to find an experimental observable perfectly correlated with a parton shower history, including color flow and hadronization. Nevertheless, a number of methods to analyze high-p_T jets have been proposed. Generally, these methods depend on differences in the substructure of light-parton QCD jets compared to those from particle decays. Diagnostics to detect this difference include infrared safe jet shapes [16–20], direct analyses of jet substructure [3, 21–34], grooming methods to improve jet mass resolution by reducing jet contamination [35–37], and matrix element methods [38]. Jet substructure techniques have also been studied in the context of specific particle searches, where they have been shown to extend substantially the reach of traditional search techniques in a wide variety of scenarios [39–48]. In addition, experiments have started to use boosted hadronic objects as a probe of new physics in data [49–51](for a summary of the experimental and theoretical progress in jet substructure, see, for instance, [52, 53]).

In Ref. [54] a template overlap method was developed for the quantitative comparison of the energy flow of observed jets at high-p_T with the flow from selected sets (the templates) of partonic states. The template overlap method can be summarized as follows, adopting the notation of Ref. [54]. We denote a set of particles or calorimeter towers in a physical (or event-generated) jet by $|j\rangle$. The jet, of course, is identified by some algorithm. Any such jet is to be compared with a large set of configurations $|f\rangle$, which represent sets of partonic momenta $p_1, \ldots, p_N$ that would be identified as a jet by the same algorithm we apply to the physical jets. For a given j, we determine the template state for which some measure, $\mathcal{F}(j, f) \equiv \langle f|j\rangle$, is maximized. Although this measure is in principle completely free, we choose below a Gaussian in energy differences within angular regions surrounding each of the partons of the template. When the energy flow of the state perfectly tracks that of a template, the measure is maximized at unity. Any region of partonic phase space $\{f\}$, can define a template, although in this analysis we will be interested in templates that can describe boosted Higgs decays, in terms of the energies and invariant masses of its partons.

Although applied to boosted Higgs in this paper, the template overlap approach is very flexible and can be applied to a wide variety of particle processes for which theoretical models have been established. For each set of templates with a definite number of particles, N, the overlaps provide us with a tool to match unequivocally arbitrary final states j to partonic partners $f[j]$. Once a "peak template" $f[j]$ is found, we can use it to characterize the energy flow of the state, which gives additional information on the likelihood that it is signal or background. Similarly, it is possible to analyze the template functions found in this way to further discriminate events. We will give examples below.

As was shown in [54], the application of these ideas is particularly straightforward for top jets. Much of the QCD background is characterized by two sub-jets, with very different

energy flow from the three-parton templates in general. Indeed, for a lowest order partonic QCD jet consisting of the original parton plus one soft gluon, there is no template state from top decay that matches the energy flow. This gives a fundamental discrimination, to which we can add additional information from event shapes.

In practice the procedure outlined in [54] is not yet optimal for disentangling a Higgs, as opposed to top, signal from a QCD background, because at lowest order (corresponding to templates with $N = 2$) both boosted Higgs and QCD jets consist of two particles. A successful identification strategy should also make use of observables that do not map exclusively onto the minimum number of partons. We go a further step in this direction by employing templates with three partons, with a cut to ensure that they are well-separated in phase space. To the extent that physical final states reflect partonic energy flow in both signal and background, the population of three-particle templates found from a given sample reflects the short-distance dynamics that produced the corresponding physical states.

In the following section we review the definition of template overlap used in this paper. Next, in Sec. 3, we introduce an efficient procedure to construct the templates and maximize the overlap. The discriminating power of several observables, including planar flow, and several new partonic template variables is discussed in Sec. 4. We then study Higgs tagging performance of template overlap in Sec. 5. Finally, we conclude in Sec. 6, leaving a few explicit calculations to the appendices.

2 The Template Overlap Method

The method developed in [54] makes use of template overlaps that are functionals of energy flow of jet events. The templates are sets of partonic momenta $f = \{p_1, \ldots, p_N\}$, with

$$\sum_{i=1}^{N} p_i = P, \quad P^2 = M^2, \tag{1}$$

which represent the decay products of a signal of mass M. Here we have adopted the expectation that a good, if not the best, rejection power is obtained when we use the signal distribution itself to construct our templates (see [55]). For example, the lowest order templates for Higgs have $N = 2$ and for the top decay, $N = 3$, with phase space in the latter case restricted by the W mass: $t \to b + W \to b + (q\bar{q})$. The number of particles in the templates is not necessarily fixed, and templates with more than the minimum number of particles are possible. We will find, however, that combining templates in the full phase space for $N = 3$ and $N = 2$ already delivers encouraging results for the Higgs.

In order to compare the energy flow of any given (physical or event-generated) jet to that of the signal on the unit sphere, denoted by Ω, following the notation of Ref. [54], we represent the template energy flow as $dE(f)/d\Omega$ while that of the jet is denoted by $dE(j)/d\Omega$.

We may then represent a general overlap functional $Ov(j, f)$ between a jet and a template as

$$Ov(j, f) = \langle j | f \rangle = \mathcal{F} \left[\frac{dE(j)}{d\Omega}, \frac{dE(f)}{d\Omega} \right] . \tag{2}$$

Here, the functional $\mathcal{F}$ is in principle completely free, and by construction it is maximized by the corresponding template for a given j. The measure for the matching between j and the template f can be taken as the weighted difference of their energy flows integrated over some specific region in the phase space. In our analysis, we will take the functional $\mathcal{F}$ as a Gaussian in energy differences within angular regions surrounding each parton in the template, i.e. around directions of the N template momenta p_i,

$$Ov_N(j, p_1, \ldots, p_N) = \max_{\tau_N^{(R)}} \exp \left[- \sum_{a=1}^{N} \frac{1}{2\sigma_a^2} \left(\int d^2\hat{n} \, \frac{dE(j)}{d^2\hat{n}} \theta_N(\hat{n}, \hat{n}_a^{(f)}) - E_a^{(f)} \right)^2 \right], \tag{3}$$

where the direction of parton a in the template is $\hat{n}_a$ and its energy is $E_a^{(f)}$. These energies set the widths σ_a of the Gaussians; while the functions $\theta_N(\hat{n}, \hat{n}_a^{(f)})$ restrict the angular integrals to (nonintersecting) regions with a definite width surrounding each of the partons of the template. The corresponding state will be referred as the "peak template" $f[j]$ for state j.

A new element of our analysis is to use templates with more than the minimum number of particles. This gives us the ability to resolve details of jet substructure, facilitating the capture of possible gluon radiation in the heavy particle decay, while still eliminating contamination from pile-up and the underlying event.

A key step for defining template overlap is to choose appropriately the set of template states over which the maximization is performed. Ideally, one would determine Ov by maximizing over all possible template states. In practice, however, one needs to introduce a discretization in the space of template states. In order to make sure that the maximum overlap found by this reduced set is very close to the true maximum, a large number of template states is needed, roughly of order 100^{3N-4}. In general, such a maximization procedure can be computationally intensive when the number of template particles, N, is large, but in section 3 we present an efficient algorithm to perform this task. We emphasize that, once generated, the same set of template states is used for all the data.

In summary, the output of the peak template method for any physical state j is the value of the overlap, $Ov(j, f)$, and also the identity of the template state $f[j]$ to which the best match is found. As we shall see, this will be of particular value when we apply our method to boosted Higgs.

4

3 Analysis overview

Motivated by Ref. [3], we now describe the application of the template method to the production of a Higgs boson in association with a W boson, $p + p \rightarrow W + H$, followed by the dominant light Higgs boson decay, to two b-tagged jets, including schemes for generating templates and for discretizing the data. We wish to investigate the tagging efficiency for this process, and the fake rates from the background process $p + p \rightarrow W + jets$. We emphasize that the method we propose is quite general, and it can be used for other massive object searches as well.

In this section, we describe the construction and use of templates with $N = 3$ for Higgs decay. Three particle templates will allow us to test the influence of gluon emission and color flow, through their effect on energy flow. This will provide an additional tool to discriminate between QCD and Higgs jets. For example, we expect soft radiation from the boosted color singlet Higgs to be concentrated between the b and the $\bar{b}$ decay products. This is to be contrasted to a jet initiated by a light parton, whose color is correlated with particles in other parts of phase space, producing radiation in the gaps between those particles and the jet system. In this way, the template for a given state can provide evidence on its origin. The template method has the advantage of not requiring any special algorithmic technique and allows us to consider more elaborate jet substructure observables, and we will look at a few specific kinds in the following sections.

3.1 Matrix Element Optimization

In the preliminary study of Ref. [54], the template states, $f = p_1, \ldots, p_N$, were generated evenly over all the phase space by a brute force method. This approach was very effective for the top at LO, $t \rightarrow b + W \rightarrow b + (q\bar{q})$, where the pair invariant mass was constrained to be m_W. However, the full, unconstrained 3-particle phase space translates to an increase of computational runtime. If we have N-body templates described by $3N - 4$ phase-space variables and we divide each of these variables into, say, ~ 100 bins, then we have $\sim 100^{3N-4}$ total bins in the absence of additional kinematic constraints. A coarser discretization might generate errors in the estimation of $Ov_N(j, f)$, when the matrix element is changing rapidly. We here present a new scheme which significantly reduces the number of the template states needed for the computation of $Ov_N(j, f)$.

A more efficient approach is to generate templates in phase space according to the size of squared matrix elements of signal events. This will ensure that most signal events will match templates very closely. In regions where matrix elements are small, they are changing slowly, so that rare signal events will be well matched even where the density of templates is lower. Such matrix element weighting methods are simple to implement and allow one

to deemphasize the uninteresting majority of phase space*. The idea is to associate to each template f a weight defined as the probability to produce and observe a template in a given model, labelled by the letter α. We can readily evaluate such a probability in the ideal situation where the resolution of the detector is perfect. The probability is then given by the squared matrix element, summed over final state colors and polarizations and averaged over initial ones,

$$\mathcal{P}(\mathbf{p}|\alpha) = \frac{1}{\sigma_\alpha} d\Phi \, |M_\alpha(\mathbf{p})|^2. \tag{4}$$

Here M_α is the matrix element, $d\Phi$ is the phase-space measure and σ_α is the total cross section.

The generation of templates according to the cross section in Eq. (4) involves three steps:

- A sampling Monte Carlo routine is used to sample the phase space, $\{\mathbf{x}^{(\mathbf{r})}\}_{r=1}^{3N-4}$ with r running over independent kinematical variables, according to the probability distribution $\mathcal{P}(\mathbf{p}|\alpha)$.

- The same algorithm computes the templates, i.e. the parton momenta $\{p_i\}_{i=1}^{N}$, from the set of $\{\mathbf{x}^{(\mathbf{r})}\}_{r=1}^{3N-4}$.

- The matrix elements contain singularities in certain kinematic configurations. For example, those with final states of several jets are singular if the jets are nearly collinear or if the energy of a jet approaches zero. Hence, after the momenta are generated, a call is made to a routine that is used to apply cuts to the generated templates. A phase space point that fails the cut is then rejected and the template is not evaluated, and the state reverts to a template with fewer particles.

3.2   Selection and Discretization of the Data

We generate events for $W^+ + H \to l^+ \nu_l \, b\bar{b}$ and $W^+ + jets \to l^+ \nu_l + jets$ in a configuration with large transverse momentum, using PYTHIA 8.150 [60], SHERPA 1.3.0 [61] (with CKKW matching [62]), and MADGRAPH [63] interfaced to PYTHIA 6 [64] (with MLM matching [65]). Jets are reconstructed using FASTJET [66], and the anti-k_T algorithm [67] with large effective cone size $R = 0.4, 0.7$. In this paper, we have chosen plausible values for R, based on a combination of physics input and a trial-and-error, but have not attempted to optimize them systematically.† For each event, we find the jet with the highest transverse momentum and impose a jet mass window for the Higgs. We choose the jet mass window to be 110 GeV

---

*In different contexts, such an approach has been applied to Tevatron data [59].

†An optimal jet radius would be a compromise between taking it large enough to include perturbative final-state radiation, and small enough to avoid too much contamination from the underlying event and initial-state radiation.

$\leq m_J \leq 130$ GeV, with our reference Higgs boson mass chosen to be $m_H = 120$ GeV, and jet energy 950 GeV $\leq P_0 \leq 1050$ GeV. This gives us a set of final states j.

For any state j, we determine the measured (or Monte Carlo (MC) generated) energy distribution, $dE(j)/d\Omega$, in the physical θ-ϕ plane with respect to the jet axis for each reconstructed jet, and we discretize this data into a jet-energy configuration. In our demonstration for the Higgs, we discretize the θ-ϕ plane into cells of size $\Delta\theta = 0.04$ and $\Delta\phi = 0.1$ [‡]. Next, we again assemble a table of energies $E(\text{row}_m, \text{column}_n)$, where row_m and column_n are the row and column number corresponding to the discretized values of θ and ϕ.

Before proceeding further, let us mention a few words concerning our choice of event selection criteria and the interpretation of our results. Our focus in this work is to demonstrate that the combination of two- and three-particle templates can give a qualitative improvement in the separation of Higgs signal events from QCD background. In order to simplify the analysis we have restricted our study to very large transverse momenta, not realistic at this time of the LHC running. However, we regard the findings described below as a proof of principle, and we expect that more realistic parameters [69] (for example, a lower p_T cut) and $b-$tagging will provide even higher discriminating power.

3.3 Construction of template functions

We define the leading order (next to leading order) templates in terms of the lowest-order (next-to-lowest-order) decays of the Higgs, schematically,

$$|f\rangle = |h\rangle^{(\text{LO})} = |p_1, p_2\rangle, \tag{5}$$

and

$$|f\rangle = |h\rangle^{(\text{NLO})} = |p_1, p_2, p_3\rangle. \tag{6}$$

Our templates will be a set of discretized partonic states corresponding to given points in phase space. We wish to generate an ample number of template states to cover both two- and three-particle phase space for Higgs decay. To this end, we use a sampling Monte Carlo routine to generate points in a region of phase-space of final-state momenta, with a density proportional to the differential cross section, as described in Sec. 3.1 above.

We assume that the lifetime of the Higgs particle is long enough so that its decay is incoherent with the evolution of the hard scattering. Thus, the rest frame of the particle is identical to the CM frame of its decay products. Starting from the hard parton-level matrix elements, we can readily find the distribution of template momenta in the Higgs jet rest frame, and then, by a straightforward Lorentz transformation, boost it to the lab frame.

[‡] In a typical experimental setup the energy is discretized according to the detector resolution, and each pair (row,col) corresponds to a specific cell in the calorimeter. At the LHC experiments [68], for instance, electromagnetic calorimeter cell size (in η and ϕ) is of $\mathcal{O}(0.025 \times 0.025)$ and of $\mathcal{O}(0.1 \times 0.1)$ for hadronic calorimeter cells.

For the two-body Higgs decay, two angles define the two-body state of the daughter particles. We choose these as the polar and azimuthal angles in the Higgs rest frame, relative to the boost axis that links the Higgs rest frame with the lab frame. In these coordinates, the Lorentz invariant differential cross section for the two-body Higgs decay has the simple form

$$\frac{d\Gamma}{\Gamma_0} = \frac{1}{4\pi} d\phi \, d\cos\theta, \tag{7}$$

where Γ_0 is the tree-level decay width for a light Higgs,

$$\Gamma_0 = \frac{N_c G_F m_H m_b^2}{4\sqrt{2}\pi} \left(1 - \frac{4m_b^2}{m_H^2}\right)^{3/2}, \tag{8}$$

with m_b the b-quark mass, m_H the Higgs mass, G_F the Fermi coupling constant and $N_c = 3$ the color factor.

By straightforward Lorentz transformations of particle momenta, the two angles identified above determine the energies and directions of the two decay products of the Higgs at LO. The two angular parameters θ and ϕ are distributed according to (7). We generate a large set of template states so that we are confident of identifying the peak value of overlap. We can now encode the two physical angles in terms of row and column numbers, corresponding to the data discretization scheme. Each template consists of the information (row$_a$, column$_a$, E_a) for each of the two daughter particles. We exclude those templates having polar angles (in the lab frame) larger than the cone size R.

Next, we consider the templates motivated by the on-shell next-to-leading order decays of the Higgs. In this case, more variables enter the problem. They can be chosen as $\{x_1, x_2, x_3\}$, x_i denoting the fractional energies of the (massless) partons in the jet CM frame (normalized so that $0 < x_i < 1, \sum x_i = 2$), and three Euler angles ψ, θ, ϕ to parameterize the orientation of the decay products in the Higgs rest frame (see Appendix A). Since the Higgs boson is a scalar, its decay products are spherically symmetric and therefore distributed uniformly in ψ, ϕ and $\cos\theta$. Specifically, in the rest frame of the scalar Higgs, the b, $\bar{b}$ quarks and the gluon, g have energy fractions given by x_1, x_2 and x_3, respectively, distributed according to the differential cross section,

$$\frac{d\Gamma}{\Gamma_0} = \frac{1}{8\pi^2} C_F \alpha_s \frac{(1 - x_1 - x_2)^2 + 1}{(1 - x_1)(1 - x_2)} dx_1 dx_2 \, d\psi \, d\cos\theta \, d\phi. \tag{9}$$

Here α_s is the coupling constant and $C_F = (N^2 - 1)/2N$. There exists a certain region of phase space, corresponding to the soft or collinear emission of a gluon, in which the three-parton templates simulates a two-parton template. This is precisely the region in which the differential distribution (9) contains divergences, which cancel in the total transition probability after the inclusion of virtual gluon exchange diagrams. We regulate these divergences in terms of the invariant mass of the pair of nearly-collinear partons, which has a simple relation to the variables x_i. For a partonic configuration to correspond to a three-particle

template, we require

$$y_{ij} = \frac{(p_i + p_j)^2}{m_H^2} = (x_i + x_j - 1) > y, \tag{10}$$

where y is a cut-off parameter. In the following, we take $y = 0.05$.

As for the two-parton templates, we generate a large number of three-parton templates, of order one million, distributed according to (9), while satisfying the constraint (10). By Lorentz transformations of particle momenta, analogous to the two-particle case, the two energy fractions and the three angles identified above determine the energies and directions of the three decay products of the Higgs at NLO. As for the discretization of the data, we again encode two physical angles in terms of row and column number corresponding to the data discretization scheme. A given three-particle template consists of a list $(\text{row}_a, \text{column}_a, E_a, a=1,2,3)$ for each of three daughter particles of NLO Higgs decay (b, $\bar{b}$ and g). We also exclude those templates having particles whose polar angles in the boosted frame, relative to the jet axis, are larger than the cone size R.

Clearly, templates cannot be meaningful for zero momentum partons. However, the invariant-mass cut Eq. (10) and the requirement that the templates are within a cone of size R imposes a lower cutoff on the energies of the template partons,

$$\frac{2\,m_H^2\,y}{P_0\,R^2} \lesssim E_a\,. \tag{11}$$

For the benchmark parameters used here, $P_0 = 1000$ GeV, $m_H = 120$ GeV, $y = 0.05$, the corresponding cutoffs are $E_a \gtrsim 3(9)$ GeV for $R = 0.7(0.4)$. Although the cutoff corresponding to $R = 0.7$ happens to be particularly close to the hadronization scale, we checked that imposing a more stringent, explicit cut on our templates, namely $E_a \geq 20$ GeV, did not have a sensible effect on the results.

3.4 Two- and three-particle template overlap

We are now ready to implement Eq. (3) for the Higgs, by defining an overlap between templates, $|f\rangle$, and jet states, $|j\rangle$, $Ov = \langle j|f\rangle$. Defined as above, our two-parton templates each have two cells corresponding to two daughter partons (q and $\bar{q}$) with their row and column numbers determined by the data discretization scheme. In addition, we have three-parton templates each having three cells corresponding to three daughter partons (q, $\bar{q}$ and g) with their row and column numbers determined by the same discretization scheme.

We compute the overlap between data state j and two- or three-body template f from the unweighted sum of the energy in the nine cells of state j surrounding and including the

occupied cells of template state f,

$$Ov_N(j,f) = \max_{\tau_N^{(R)}} \exp\left[-\sum_{a=1}^{N} \frac{1}{2\sigma_a^2} \left(\sum_{k=i_a-1}^{i_a+1} \sum_{l=j_a-1}^{j_a+1} E(k,l) - E(i_a,j_a)^{(f)} \right)^2 \right], \qquad (12)$$

where $N = 2$ or 3. Here, $E(i_a, j_a)^{(f)}$ is the energy in the template state for particle a whose direction is labelled by indices i_a and j_a, according to the discretization table described in Sec. 3.2. If one of the sums extends outside the jet cone, we set the corresponding energies $E(k,l)$ to zero. We fix σ_a (for the ath parton) by that parton's energy, $\sigma_a = E(i_a, j_a)^{(f)}/2$.

The use of $N = 2$ and $N = 3$ templates follows three steps. Starting from a hard massive jet on angular scale R, one identifies the two subjets within it using two-parton templates. Within the same angular region, one then further takes the three hardest subjets that appear using three-parton templates, if the energies of three subjets satisfy Eq. (11). If this condition is not satisfied, only the two-body templates are used. The whole procedure is illustrated in Fig. 1.
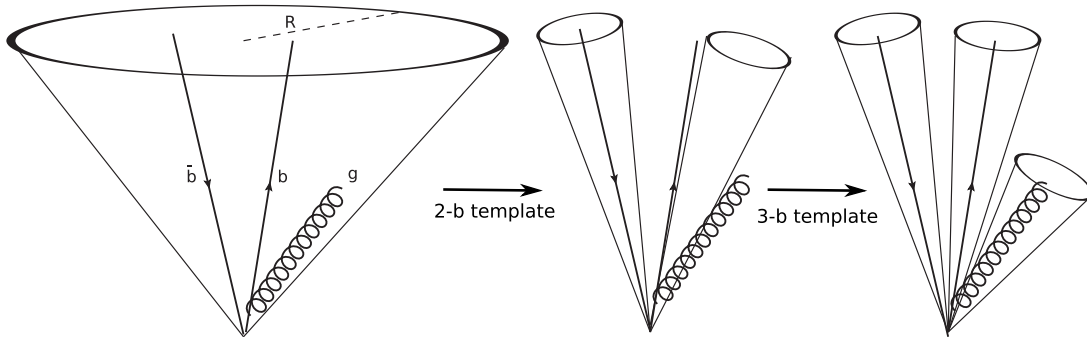

Figure 1: Stages of the template method described in text.

3.5   Parton-level studies

To check the consistency of our method, we first test our ideas using parton-level events, which are simulated by a Monte Carlo sampling of Eq. (9). The Monte Carlo samples used for this validation include final state radiation up to $\mathcal{O}(\alpha_s)$. The events are conveniently regularized as in Eq. (10) via an invariant-mass cut algorithm with $y = 0.05$. The signal samples thus contain only q and $\bar{q}$ quarks and up to one gluon g in the final state. These events are not run through full parton showering and hadronization simulations. Consequently, the generated partons in each event are distributed according to the same squared matrix elements used in the template generation, Eqs. (7) and (9).

In Fig. 2, the results are shown for the two- and three-body overlap given by Eq. (12). The data shows that, as required for consistency, each peak value Ov_2 (Ov_3) is close to unity

for all the events in our Higgs decaying into a $b\bar{b}$ ($b\bar{b}g$) sample, around 80% (20%) of the total (see Fig. 13 in Appendix B). Of the 22% with $Ov_3 > 0.8$, 14% have small Ov_2 while 8% have substantial two-body overlap, $0.2 < Ov_2 < 1$. As we will see below, the effect of showering is to spread out Higgs decays over the full range of Ov_2 and Ov_3.
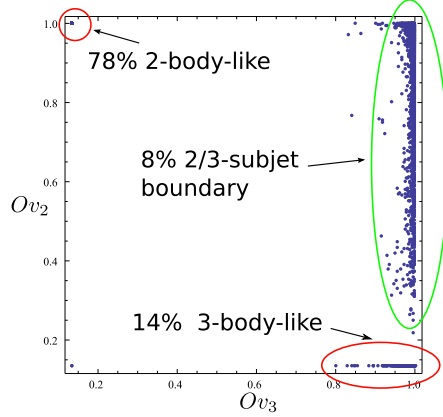

Figure 2: A scatter plot of two-parton template overlap *vs.* three-parton template overlap for LO parton-level MC output for Higgs decay, with jet energy, $P_0 = 1$ TeV, $m_H = 120$ GeV.

3.6 Template overlaps for Higgs and QCD jets

We now proceed to test the ability of two- and three-parton templates to distinguish between energetic Higgs jets from QCD jets. We apply templates to events generated by PYTHIA [60], MG/ME [63], and SHERPA [61]. The event selection was described in Sec. 3.2.

In Fig. 3 we exhibit typical overlap distributions for showered Higgs and QCD jets (for the same jet mass and energy) for event generators PYTHIA (version 8) for $2 \rightarrow 2$ process without matching, MADGRAPH/MADEVENT (MG/ME) (with MLM matching interfaced into PYTHIA V6.4), and SHERPA 1.2.1 (with CKKW matching). In the plots on the left panels of Fig. 3, we compare the template overlap distribution from Eq. (12) with $N = 2$ for Higgs and QCD jets. The corresponding plots on the right panels of Fig. 3 show similar distributions when three-particle templates ($N = 3$) are used. Clearly, showering smears the Higgs distributions significantly, although Higgs events are concentrated at larger peak overlaps than QCD events for both $N = 2$ and $N = 3$. These differences, although not adequate to isolate a small signal, can serve as a foundation for further enrichment by use of additional analysis of template and jet states, as discussed below.[§]

[§]Note also the moderate variations between the different Monte Carlo generators. Since the template overlap method depends on the precise radiation pattern within a jet, these variations are expected.
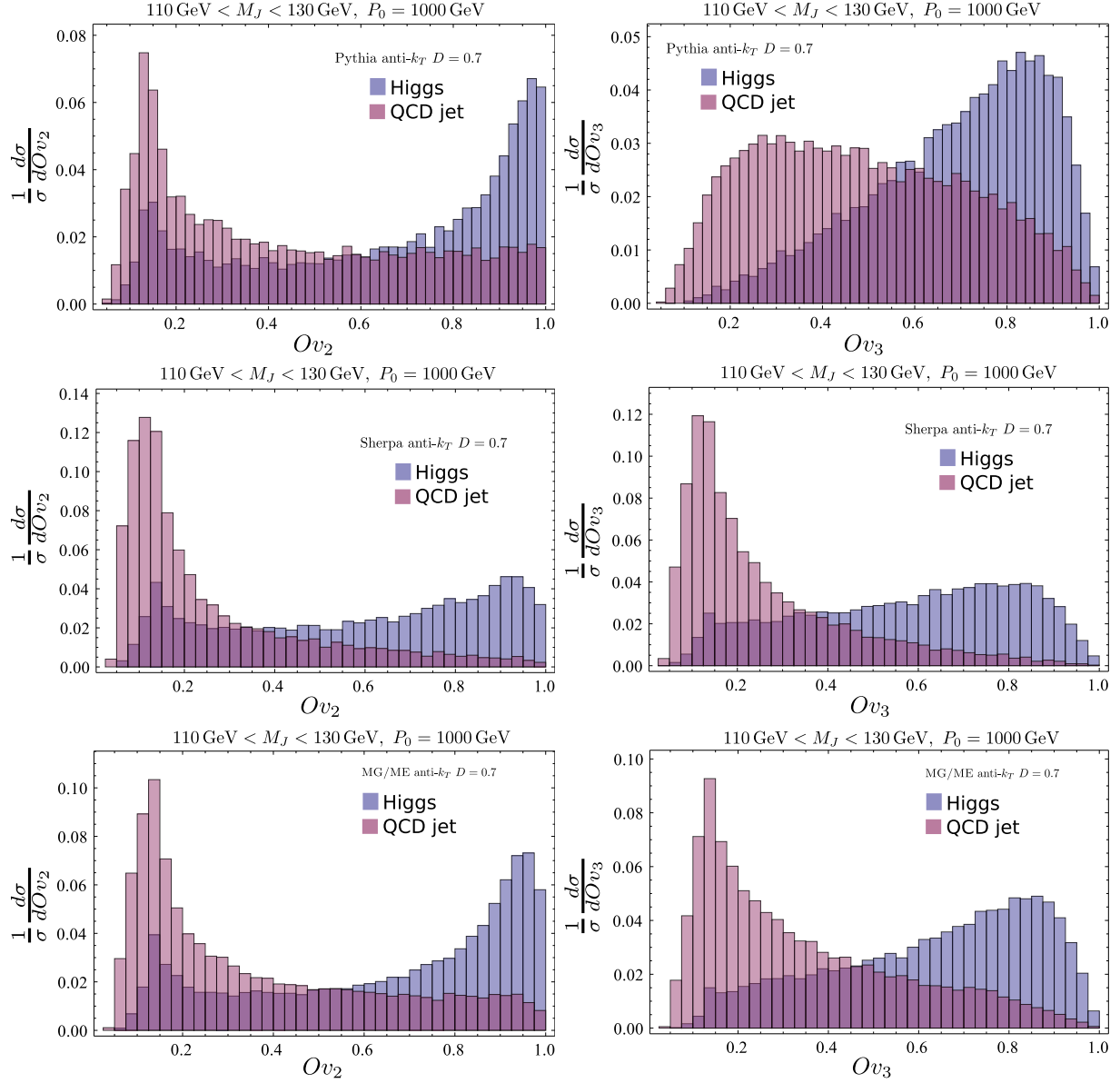
Figure 3: Comparison of histograms of template overlap Ov with Higgs jets and QCD jets from different MCs [from top to bottom, PYTHIA, SHERPA and MADGRAPH], for $R = 0.7$, 950 GeV$\leq P_0 \leq$1050 GeV, 110 GeV$\leq m_J \leq$130 GeV and $m_H = 120$ GeV using 2-body templates (left) and 3-body templates (right).

4 Single-variable discriminants

While overlap distributions help enrich samples of Higgs events by capturing the constrast in the radiation pattern for the Higgs signal and QCD background, there is still a large tail of QCD background events with relatively large Ov. We proceed to develop and study several single-variable discriminants, which provide us with additional tools to maximize the separation between QCD and Higgs jets. Many of these variables, although intimately related, are worth studying individually, since they can be used to highlight different characteristics of jet substructure.

Infrared-safe observables, such as jet shapes, guarantee that we can make meaningful comparisons between theoretical computations and their analogous experimental measurements. The method identifies the template that a physical state most closely resembles. We can also analyze, for example, angular distributions of template partons, considered as subjets, without the energy weights in event shapes. The template can thus be used to identify jet substructure [3, 21–38]. The overlap method enables us to make subjet identification by providing a mapping between energy-unweighted variables and the template that defines the energy flow distributions. In what follows, we describe these two categories of tools that will allow us to further reject background events with a large Ov.

4.1 Planar flow

Jet shape variables are an especially interesting class of observables for jet studies, and have received considerable attention in the past several years in the context of boosted jet identification [16, 17, 20]. The common feature of all jet shapes is that they involve moments of the energy of observed particles and are thus smooth functionals of energy flow within a jet. In this manner, they are complementary to the information provided by template overlaps, which is associated with jumps and spikes in energy flow.

Following Ref. [54], we will make use of the jet shape planar flow in the form,

$$Pf = \frac{4 \det(I_\omega)}{\operatorname{tr}(I_\omega)^2},$$
(13)

where I_ω is defined by,

$$I_\omega^{kl} = \frac{1}{m_J} \sum_i \omega_i \frac{p_{i,k}}{\omega_i} \frac{p_{i,l}}{\omega_i},$$
(14)

with m_J the jet mass, ω_i the energy of particle i in the jet, and $p_{i,k}$ the kth component of its transverse momentum relative to the axis of the jet's momentum. Jets attributed to two-body final states have a differential jet function fixed at zero planar flow,

$$\frac{1}{J} \left(\frac{dJ}{dPf} \right)_{2\,\text{body}} = \delta(Pf).$$
(15)

This would apply at leading order for events with highly boosted Higgs and QCD jets. On the other hand realistic QCD and Higgs jets have nonzero Pf, because of QCD radiation effects that smear the distribution.

We expect soft radiation from the boosted color singlet Higgs to be concentrated between the b and $\bar{b}$ decay products. This is to be contrasted to a jet initiated by a light parton, whose color is correlated with particles in other parts of the event, producing radiation in the gaps between those particles and the jet system. Therefore, we expect that planar flow for Higgs jets will be peaked toward a lower value than that of QCD jets. In the studies we show below, the combination of Ov and Pf gives a strong background (QCD) suppression with quite substantial signal (Higgs decay) efficiency.
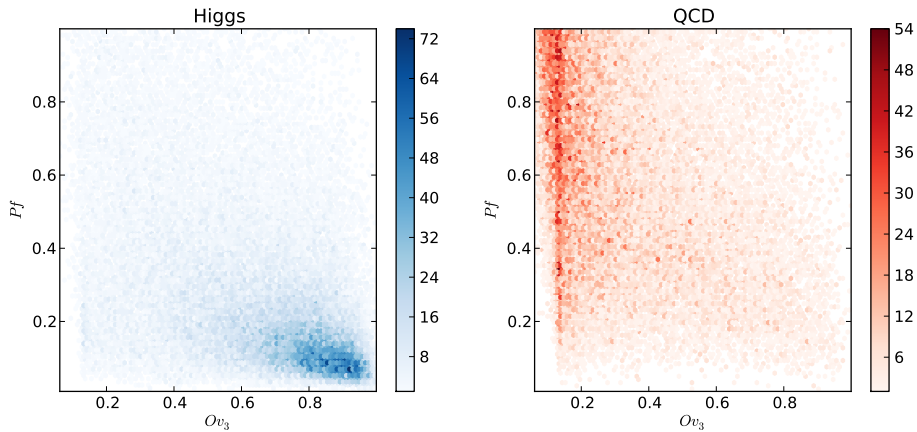

Figure 4: Density plots of planar flow Pf vs. template overlap Ov for Higgs jets and QCD jets from PYTHIA [60], for $R = 0.7$, 950 GeV$\leq P_0 \leq$1050 GeV, 110 GeV$\leq m_J \leq$130 GeV using three-body templates.

Figure 4 shows the two-dimensional distributions of MC events (obtained via PYTHIA [60]) in the Pf vs. Ov_3 plane for the signal and background. The scatter plot shows that signal events cluster around large overlap while, at the same time, Pf is essentially below 0.2. By contrast, QCD events tend to be spread over the entire area. These plots also confirm our expectation that Higgs jets tend to have smaller Pf values than QCD jet events (for the same ratio m_J/P_0). Clearly, any set of events chosen from the bottom right of these plots, with $Pf < Ov_3$, is highly enriched in three-body Higgs events compared with background. The clear difference in these scatter plots shows the potential of the template overlap method.

14

4.2 Template variables

We will refer to the second set of variables for discriminating jets as template variables. We focus on variables which are sensitive to differences between jet radiation for signal and background. For the particular case of the Higgs boson signal, we aim to show the potential discriminating power of this category of variables. In future work it should be possible to improve upon our naïve analysis below, which is based on simple rectangular cuts.

We can test these ideas on planar flow. The Pf computed from the template of any event is a template variable, and can be compared to the corresponding Pf of the event. For comparison purposes, in Fig. 5 we show the two-dimensional distributions for the actual planar flow and its template variable version for Higgs jets. The data corresponds to Higgs jets from parton-level MC output based on Eq. (9) (left) and PYTHIA (right). The parton level plot validates the use of templates, since points are concentrated on the diagonal. The distribution from PYTHIA is slightly different, but the correlation between the actual planar flow and its template version is evident. One can see that the effect of showering is to smear the Pf distribution to larger values.
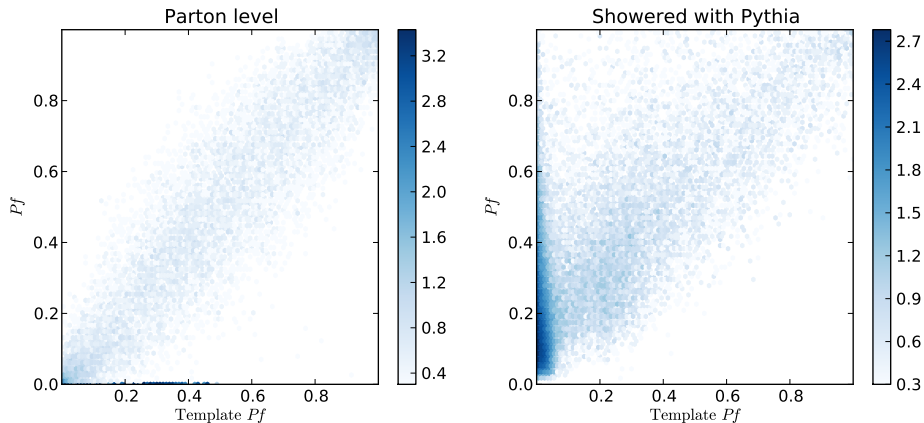
Figure 5: Density plots of Pf *vs.* partonic template Pf for Higgs decay events from LO parton-level MC output (left) and PYTHIA (right), with $P_0 = 1$ TeV, $m_H = 120$ GeV. The intensity of the shading is proportional to the density of points.

We first consider variables constructed out of templates with the minimum number of particles. This has been discussed in detail in [54], so our discussion will be brief. As described in [17], at lowest order the signal phase space for the Higgs decay is characterized by simple kinematic parameters. For example, in Higgs decays, there are only two variables that characterize the decay $h \rightarrow b\bar{b}$. One of them is the jet energy P_0, and a convenient choice for the other variable is the angle, θ_s, between the jet axis and the softer of the two particles. At fixed m_J/P_0, Higgs events tend to be peaked towards smaller values of θ_s than

15

QCD jets [17]. Therefore, even within the two-body description, θ_s already provides useful information of jet energy flow. An analysis of this template variable and its application to two-body Higgs decay was presented in [54].

Once we include the contribution from leading order radiation to the Higgs decay, the phase space is now characterized by five variables, $(x_1, x_2, \psi, \theta, \phi)$ as defined in Sec. 3.3. This gives us more freedom to contrive template variables. In the following, we discuss distributions in these variables, beginning with the energy fractions, x_1 and x_2. Here and below, x_1 is the energy fraction for the b quark while x_2 is that of $\bar{b}$ in the template.

We may anticipate that x_1 and x_2 are not ideal discriminants between Higgs decay and QCD events, because both distributions have singularities at $x_1 = 1$ and $x_2 = 1$, which reflect collinear and soft gluon emission. This is seen in Fig. 6, where we show two-dimensional distributions for $\{x_1, x_2\}$ for Higgs and QCD peak templates. The top panels show the $\{x_1, x_2\}$ distributions at the MADGRAPH level before showering, and the bottom panels after events are showered and jets reconstructed with PYTHIA and FASTJET. We have also imposed a quality cut on the three-body template overlap $Ov_3 > 0.6$. We can see that the discriminating power of the energy fractions is reduced by requiring an overlap cut in the showered events. While one might still be able to improve signal to background by drawing a contour around the collinear regions, it is clear that the $\{x_1, x_2\}$ variables alone will not be very effective in discriminating boosted Higgs jets from QCD background, and in any case, in these regions we expect resummation effects to be important.

Template variables based on angular distributions are more promising than energy fractions. To be specific, we will consider the following (lab frame) variables, and look at their distributions.

- The angles between the jet axis and the template momenta θ_{iJ},

$$1 - \cos\theta_{iJ} = \frac{z\,x_i\,m_J}{2E_i}\,, \tag{16}$$

  with $z = m_J/P_0$.

- The angular separations: θ_{12}, θ_{13}, θ_{23},

$$1 - \cos\theta_{ij} = \frac{(x_i + x_j - 1)\,m_J^2}{2E_iE_j}\,. \tag{17}$$

- The angle between the jet axis and the softest of the partons: $\tilde{\theta}_s$,

$$1 - \cos\tilde{\theta}_s = \frac{z\,x_s\,m_J}{2E_s}\,, \tag{18}$$
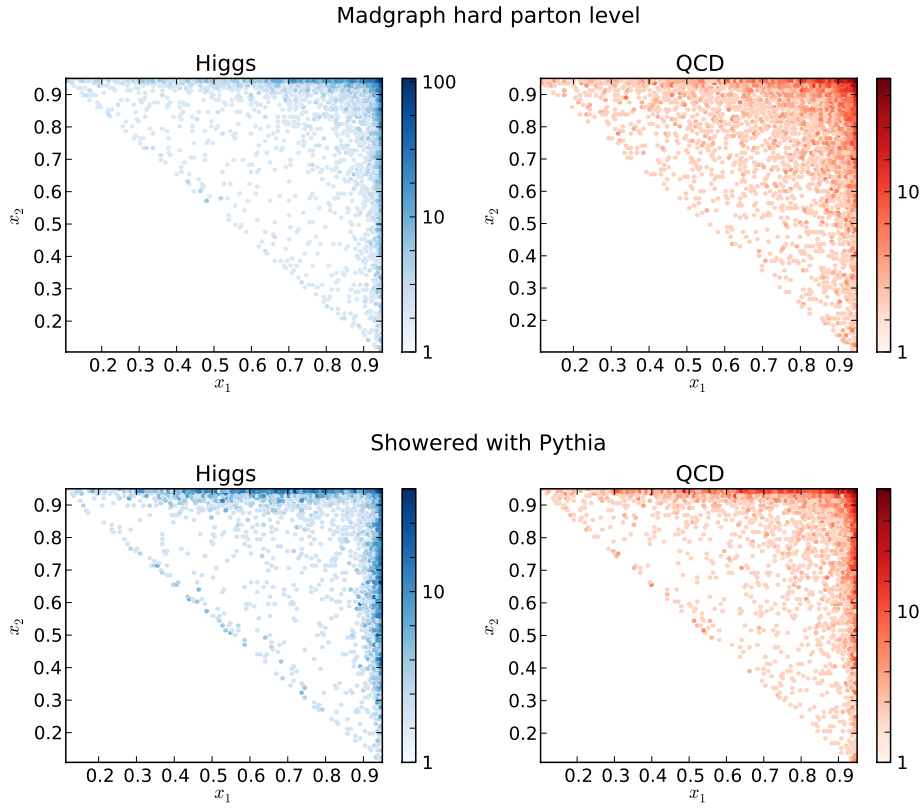
  where $E_s = \min\{E_i\}$.

16

Figure 6: Density plots of x_1 vs. x_2 for Higgs signal (blue) and QCD background (red), for the LHC. MADGRAPH hard parton-level (top) and showered jet level with PYTHIA (bottom). Both are shown only in the Higgs mass window 110 GeV $< m_J <$ 130 GeV, with $P_0 = 1$ TeV, $m_H = 120$ GeV.

- $r_\theta = \min\{\theta_{13}/\theta_{12}, \theta_{23}/\theta_{12}\}$, found by finding the minimum of $(1 - \cos\theta_{i3})$, $i = 1, 2$, given by

$$\min\left\{ \frac{(1 - x_2)E_2}{(1 - x_3)E_3}, \frac{(1 - x_1)E_1}{(1 - x_3)E_3} \right\} . \tag{19}$$

- The three-body angular variable $\bar{\theta}$,

$$\bar{\theta} = \sum_i \sin\theta_{iJ} , \tag{20}$$

with θ_{iJ} given by Eq. (16).

The expression for the energy E_i of particle i is fairly simple, and is given in Appendix A, Eqs. (26-28).

The distributions of the variables Eqs. (17)-(20) are shown in Fig. 7. All of these variables are shown for anti-k_T $R = 0.7$ jets. Since our focus is on the difference in the shapes of various observables, all of the kinematic distributions are normalized to unity. The angular variables r_θ and $\bar{\theta}$ offer the promise of reasonable discriminating power, because they are directly tied to physical features of the signal, as follows.

In the Higgs decay to a quark-antiquark pair and a gluon, $h \to q\bar{q}g$, we expect events where the gluon is soft to be predominant. In the boosted frame, this radiation appears dominantly within an angular region spanned by the dipole formed by the quark and the antiquark [70]. In contrast, in the perturbative expansion, jets initiated by quark or gluon radiation would have a color connection with the rest of the event resulting in a bias for large angle soft gluon emission towards other jets in the event or the beam. One can take advantage of this in order to focus on regions in the phase space where Higgs decay events are more likely to dominate over QCD events (*e.g.* with ggg final state).

The variables r_θ and $\bar{\theta}$ are designed to measure the difference in angular ordering in our peak templates between the Higgs boson signal and the QCD background. We note that both $\bar{\theta}$ and θ_{12} for three-body templates are somewhat analogous to θ_s for two-body templates.

5   Higgs tagging performance

In this section, we investigate the tagging efficiencies for Higgs jets and the mistag rates for QCD jets using template overlap. We apply it to events generated by PYTHIA 8.150 [60], SHERPA 1.3.0 [61] (with CKKW matching [62]), and MADGRAPH [63] interfaced to PYTHIA 6 [64] (with MLM matching [65]) also. The event selection was described in Sec. 3.2.

As seen in Sec. 3.6, both two- and thee-body template overlap have substantial discriminating power. This can also be seen in the scatter plots, shown in Fig. 8, of Ov_2 and Ov_3 for
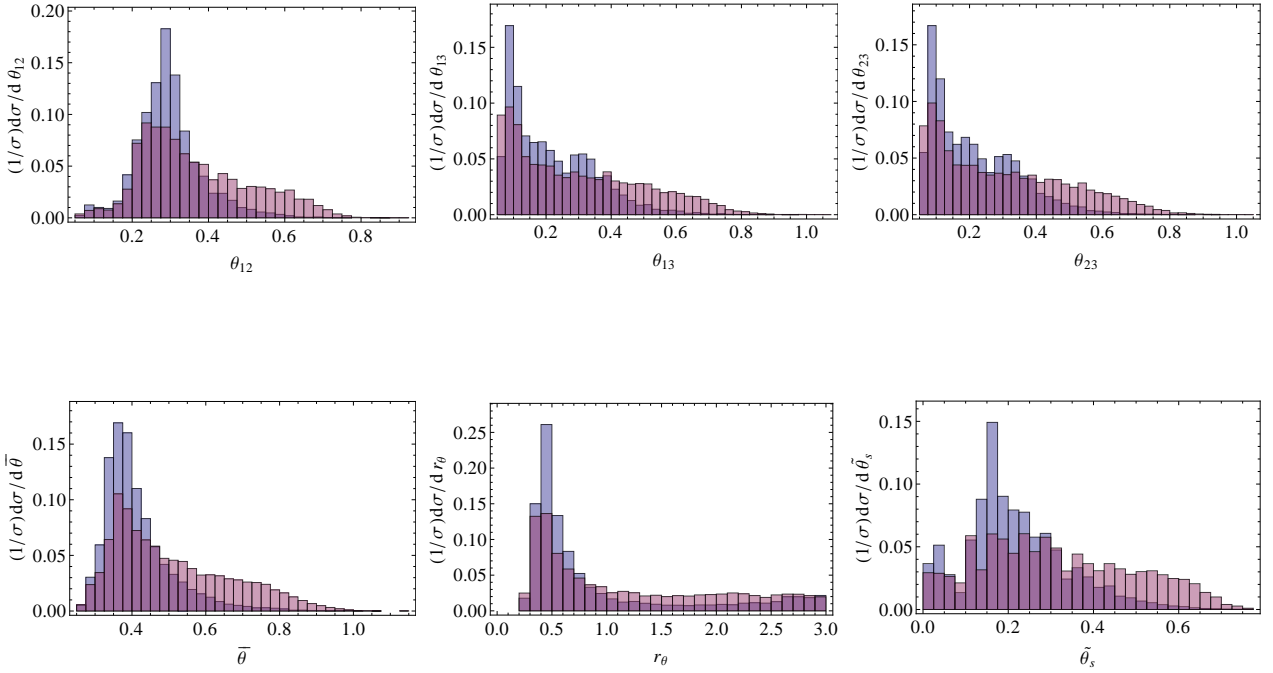
Figure 7: A selection of various template variable distributions for Higgs jets (blue) and QCD background (purple) at the LHC. Events satisfy selection cuts and the Higgs mass window cut, 110 GeV $< m_J <$ 130 GeV. Horizontal axes are in radians or dimensionless units as appropriate, and vertical axes are in arbitrary units with signal and background normalized to the same area.

19

Higgs signal (left panel) and dijet production (right panel). While the signal events cluster around the upper right corner of the plot, most QCD jet events are localized diagonally opposite in the lower left. It follows immediately that making tight cuts on each observable, by drawing a rectangular window in the upper right corner of the scatter plot, makes a good discriminator to separate signal from background.
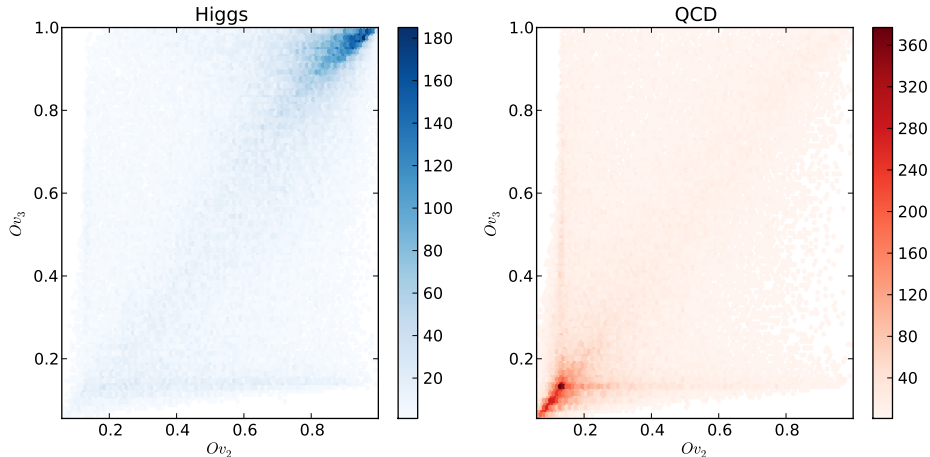

Figure 8: Density plots of 2-body overlap *vs.* 3-body overlap for boosted Higgs and QCD jets with $R = 0.7$ and same number of events (20000).

We now assess the additional discriminating power offered by the template variables discussed in section 4.2. In the simple analysis of this paper, we only look at the effect that a simple cut or window has on the number of signal and background events, leaving the use of more sophisticated methods for future work. We did not find that any of the template variables were qualitatively better than the planar flow cut. On the other hand, we found that planar flow and the $\bar{\theta}$ variable were somewhat complimentary. The two dimensional distribution Pf vs $\bar{\theta}$ for both Higgs jets and QCD jets is shown in Fig. 9. One can see that Pf and $\bar{\theta}$ are approximately independent and, thus, drawing a contour to separate signal from background will do better than any single straight line.

In Fig. 10, we summarize the boosted Higgs tagging efficiency versus background rejection using the two overlaps observables Ov_2 and Ov_3 and the two variables $\bar{\theta}$ and Pf. For a given lower cut on Ov_3 (denoted by the same color line in each frame) the efficiency is controlled by the lower cut on two-parton overlap Ov_2. Each point on one of these curves corresponds to a specific choice of Ov_2 at fixed Ov_3, and hence to the set of points within a rectangle that includes the upper right corners of the corresponding scatter plots in Fig. 8. The results depend on the choice of Ov_3 cut, but it is clear that any cut above 0.8 leads to a substantial increase in efficiency. Without using properties of the data itself, such as jet shapes, it appears that purely template overlap variables can be used to remove significant amount of background. Planar flow cuts remove some of the background contributions, but
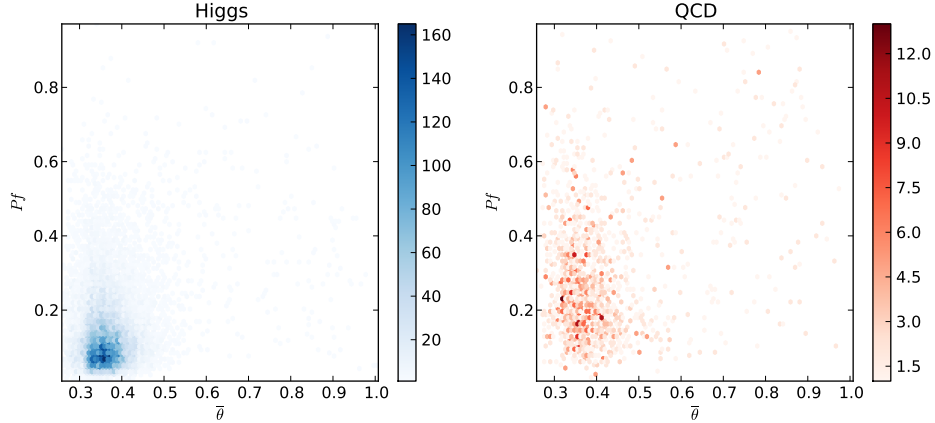
20

Figure 9: Density plots of Pf vs $\bar{\theta}$ for anti-k_T $R = 0.7$ jets for Higgs signal (left) and QCD background (right) at the LHC with selection cuts and quality cut $Ov_2 > 0.8$ and $Ov_3 > 0.8$.

$\bar{\theta}$ distributions also show some remaining discriminating power. Once we combine the fake rate and efficiency from a jet mass cut (fake rate: $\sim 10\%$, efficiency: $\sim 70\%$) with template overlap, $\bar{\theta}$ and planar flow, we find, for example, at efficiency of 10%, a fake rate of 0.05% (with $Ov_3 > 0.8$, $Pf < 0.2$ and $\bar{\theta} < 0.4$).

Our final results for the Higgs jet case are summarized in Tables 1 and 2 for the three event generators and $R = 0.4, 0.7$, respectively, that result from including these simple, naive one-dimensional cuts in Ov_2, Ov_3, $\bar{\theta}$ and Pf at fixed signal efficiency of $S = 10\%$. It is evident from the numbers presented that the template overlap method works well for events generated by any of the MC generators. In each case, we find a large enhancement of signal compared to background, typically of the order of fifteen or more. Taking into account the rejection of QCD jets by imposing a mass window, these numbers (for a single massive jet) are multiplied by factors of ten to twenty. The efficiency for finding a jet within the Higgs mass window is small for $R = 0.4$, since a small cone is unlikely to capture all of the QCD radiation in Higgs decay. For the bigger cone size $R = 0.7$, the efficiency is increased roughly by a factor of 2.

For comparison, we have applied the BDRS Higgs tagger [3] as implemented by FASTJET 3 [66] to the PYTHIA samples with the same jet radius and invariant mass window for QCD jets and Higgs jets. The BDRS Higgs tagger has two distinct parts: First, a mass-drop criteria to determine whether a jet contains substructure consistent with a Higgs boson decaying into bottom-antibottom quark pairs. If it passes the criteria, a jet is then filtered by reclustering it with a smaller jet radius R_{filt} and only keeping the 3 hardest of these subjets as the pieces of the final filtered jet. One can then place cuts on the mass of the filtered jets, m_{filt}, to further reject background from QCD jets. Our results can be seen in Fig. 11, which compares the Template Overlap method to the BDRS Higgs tagger in a
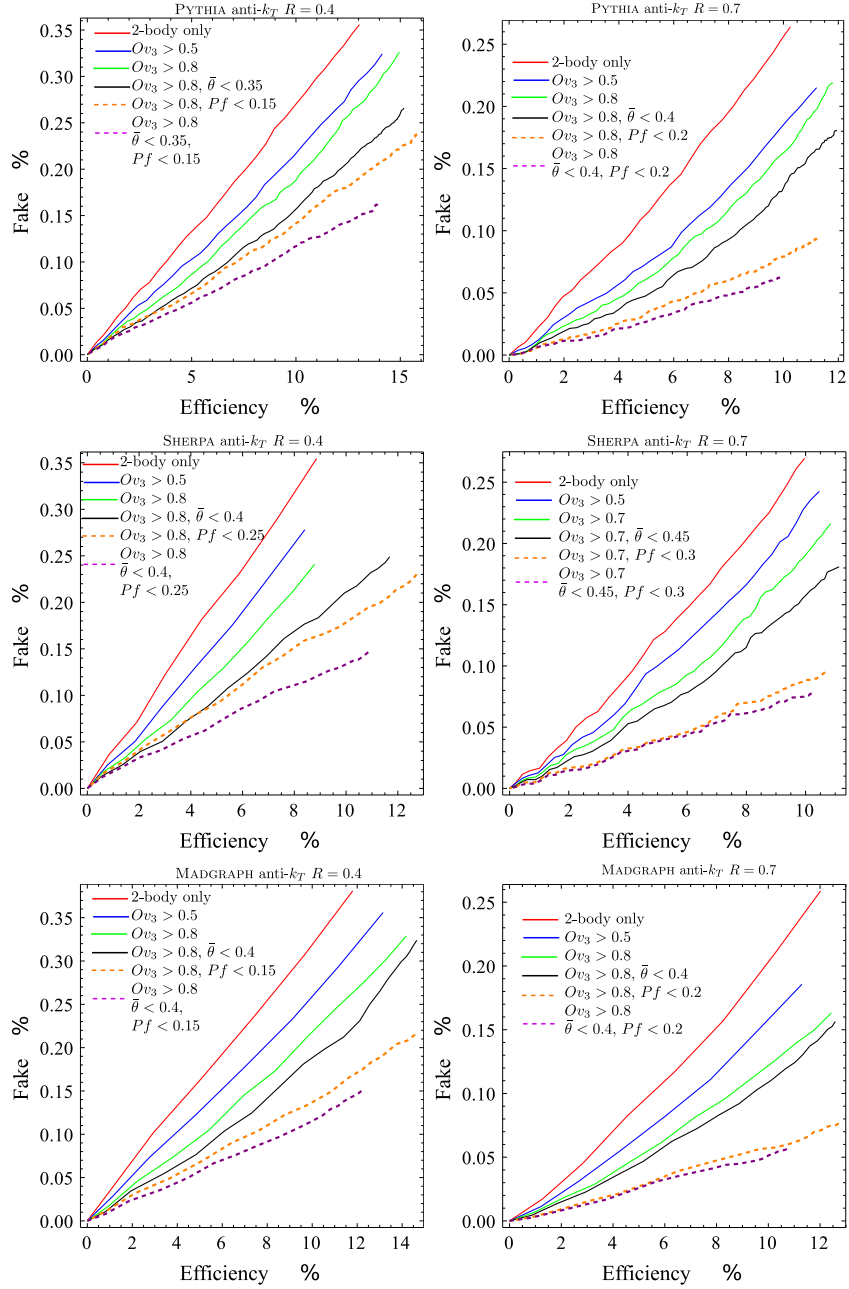
Figure 10: Comparison of fake rate *vs.* signal efficiency with various cuts on three-body template overlap Ov_3 with Higgs jets and QCD jets. The curves represent variations on the minimal value of two-body template overlap Ov_2. The frames show results from different MC [top PYTHIA, middle SHERPA, and bottom MG/ME with $R = 0.4$ (left) and $R = 0.7$ (right)], for a Higgs mass window selection criteria 110 GeV $< m_J <$ 130 GeV and 950 GeV $< P_0 <$ 1050 GeV. Both efficiency and fake rates decrease as we increase the overlap cut. The dashed curves denote the case when Pf cut is implemented, while the solid curves have no Pf cut.

simple, non-optimized test. We see that the template-based approach and the BDRS Higgs tagger perform very similarly. While not an optimal comparison, this should serve as a useful point of reference to establish the effectiveness of Template Overlap.
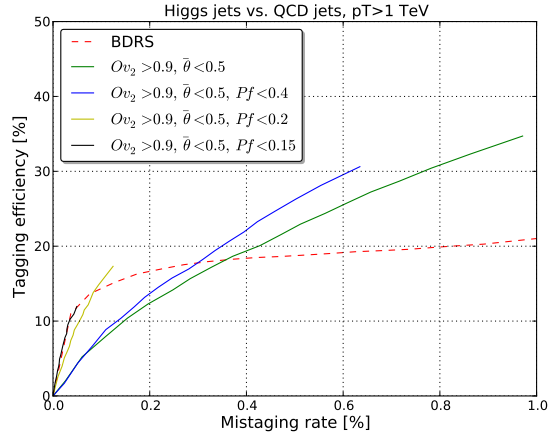

Figure 11: Efficiency versus mistag rate for the Template Overlap method and the BDRS Higgs tagger. The filled curves represent variations on the minimal value of three-body template overlap Ov_3, while the dashed curve corresponds to a varying mass window for the filtered jet mass, m_{filt}. The results are from PYTHIA 8 with $R = 0.7$ and same selection criteria as in the template case above.

We present these results for demonstration purposes only, and have not carried out a systematic study of how to maximize rejection power. Because the template method naturally provides scatter plots like those in Figs. 8 and 9, we can imagine optimizing cuts on the data. We may also investigate improvements in the overlap functional Eq. (12).

6 Conclusions

The template overlap method has the ability to maximize the physics reach for massive jets by matching jet energy flow with that of a boosted partonic decay. Template overlap is a particularly interesting observable, since it directly measures the N-prong nature of a jet. In this paper, we have proposed and tested several modifications to the template overlap method as implemented in [54], combining two- and three-body template states to the analysis of boosted Higgs. Using several MC simulated samples, we have shown that template overlap offers the promise of a successful boosted Higgs tagger, validating the preliminary study in [54]. We have demonstrated how the inclusion of three particle templates allows us to test the influence of gluon emission and color flow, through their effect on energy flow, and have illustrated its use through the construction of several partonic template observables. Different event generators give moderately different averages for our template overlaps. We

23

MC	Jet mass cut only		Mass cut + Ov +$\bar{\theta}$+ Pf	
	Higgs-jet efficiency [%]	fake rate [%]	Higgs-jet efficiency [%]	fake rate [%]
PYTHIA 8	60	5	10	0.10
MG/ME	60	5	10	0.10
SHERPA	50	5	10	0.10

Table 1: Efficiencies and fake rates for jets with $R = 0.4$ (using anti-k_T: $D = 0.4$), 950 GeV$\leq P_0 \leq$1050 GeV, 110 GeV$\leq m_J \leq$130 GeV and $m_H = 120$ GeV. The left pair of columns shows efficiencies and fake rates found by imposing the jet mass window only. The right pair takes into account the effects of cuts in both $Ov's$, $\bar{\theta}$ and Pf in addition to the mass window. For the different MC simulations, we have imposed various cuts on Ov, $\bar{\theta}$ and Pf variables: for PYTHIA v8 [60] $Ov_2 \geq 0.8$, $Ov_3 \geq 0.8$, $\bar{\theta} < 0.4$ and $Pf < 0.2$, for MG/ME [63] interfaced to PYTHIA v6 [64](with MLM matching [65]) $Ov_2 \geq 0.8$, $Ov_3 > 0.8$, $\bar{\theta} < 0.4$ and $Pf < 0.2$ and for SHERPA [61] (with CKKW matching [62]) $Ov_2 \geq 0.7$, $Ov_3 > 0.7$, $\bar{\theta} < 0.45$ and $Pf < 0.3$.

nevertheless find in each case excellent and similar rejection power. For the Higgs case studied in this work we get a rejection power of order 1:200 when combined with a jet mass cut, with sizable efficiencies.

We should emphasize that our event selection was chosen in a kinematical regime that at present is unrealistic for the LHC. However, our findings should serve as a proof of concept for many of the ideas, and, based on ongoing research [69,71], we expect an extended phenomenological analysis to deliver similar qualitative behaviour in terms of rejection power. The fact that the rejection powers are strong is encouraging and should help motivate possible modifications and improvements of the template method with different analyses in mind. We have not yet made use of b-tagging, which is a natural extension to the method. Since template overlap cuts are independent of b-tagging, we expect the use of b-tagging to result in significant improvements in background rejection [69]. Additional discrimination power would be possible using multivariate techniques, and we leave them for future study. Another point that we have not investigated in detail and that might be interesting in a future study is how it is possible to fix some of the parameters in an optimal way given the properties of the events we want to reconstruct. As we have seen in Section 5, the choice of the jet cone radius R has a strong impact on the performance of the method; and it might be interesting to see how R could be adapted dynamically when varying the jet selection criteria [72]. It is also worth noting that the template method is quite general and could be employed in a broader set of applications to both Standard Model and beyond the Standard Model physics. As the LHC continues to explore the energy frontier of particle physics, template overlap provides us with an interesting tool for further development of jet substructure techniques.

MC	Jet mass cut only		Mass cut + Ov + $\bar{\theta}$ + Pf	
	Higgs-jet efficiency [%]	fake rate [%]	Higgs-jet efficiency [%]	fake rate [%]
PYTHIA 8	70	10	10	0.05
MG/ME	70	10	10	0.05
SHERPA	60	10	10	0.05

Table 2: Efficiencies and fake rates for jets with $R = 0.7$ (using anti-k_T: $D = 0.7$), 950 GeV$\leq$ $P_0 \leq$1050 GeV, 110 GeV$\leq m_J \leq$130 GeV and $m_H = 120$ GeV. The left pair of columns shows efficiencies and fake rates found by imposing the jet mass window only. The right pair takes into account the effects of cuts in both $Ov's$, $\bar{\theta}$ and Pf in addition to the mass window. For the different MC simulations, we have imposed various cuts on Ov, $\bar{\theta}$ and Pf variables: for PYTHIA v8 [60] $Ov_2 \geq 0.8$, $Ov_3 \geq 0.8$, $\bar{\theta} < 0.4$ and $Pf < 0.2$, for MG/ME [63] interfaced to PYTHIA v6 [64](with MLM matching [65]) $Ov_2 \geq 0.8$, $Ov_3 > 0.8$, $\bar{\theta} < 0.4$ and $Pf < 0.2$ and for SHERPA [61] (with CKKW matching [62]) $Ov_2 \geq 0.7$, $Ov_3 > 0.7$, $\bar{\theta} < 0.45$ and $Pf < 0.3$.

Acknowledgments

Appendices

A   Kinematic of three-body decays

We consider the process $pp \rightarrow HX \rightarrow q\bar{q}gX$ of Fig. 12, where q and $\bar{q}$ are quark and antiquark and g is a gluon. In the approximation where the Higgs boson is exactly on-shell, the cross section for this process can be written in a factorized form as

$$d\sigma_{pp \rightarrow Xq\bar{q}g} = d\sigma_{pp \rightarrow XH} \frac{d\Gamma_{H \rightarrow q\bar{q}g}}{\Gamma_0}. \tag{21}$$

It follows that the rest frame of the Higgs is identical to the center-of-mass frame of its decay products. In that frame the three final state particles will lie in a plane. The phase space can therefore be specified by giving three Euler angles (ψ, θ, ϕ) that specify the orientation
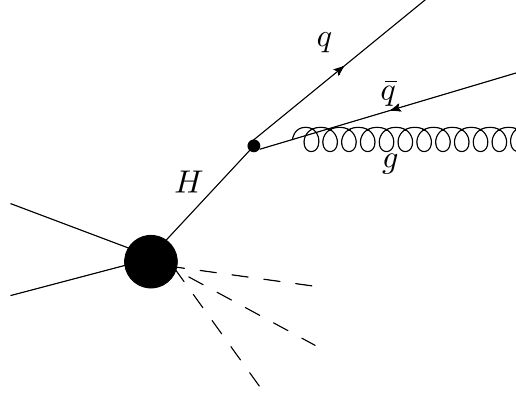
25

Figure 12: A schematic diagram of a Higgs production process and decay. For our study, the Higgs boson is assumed to be exactly on-shell.

of the final system relative to the Higgs boson and fractional energy variables

$$x_i = \frac{E_i^*}{m_H/2} = \frac{2p_i \cdot q}{m_H^2}, \tag{22}$$

with E_i^* the CM energy, and

$$0 < x_i < 1 . \tag{23}$$

Here, q^μ is the Higgs momentum with $q_\mu q^\mu = m_H^2$, p_i^μ are the momenta of the outgoing partons $(q, \bar{q}, g)$. Energy conservation gives $\sum x_i = 2$, which implies that only two of the x_i are independent.

Let θ_{ij} be the angles between the momenta of the partons i and j. One can relate the angles between the momenta to the energy fractions defined by Eq. (22). Momentum conservation gives the relation between the energy fractions of the massless decay products and the angles between their momenta in the rest frame of the Higgs,

$$(1 - \cos \theta_{ij}) = \frac{2(1 - x_k)}{x_i x_j} , \quad i \neq j \neq k . \tag{24}$$

We can also obtain the angles between the boosted momenta in any frame,

$$(1 - \cos \theta_{ij}) = \frac{m_H^2}{2 E_i E_j}(1 - x_k) . \tag{25}$$

The energies E_i after a Lorentz boost $\vec{\gamma} = \gamma \hat{n}$ are given by

$$E_1 = \frac{1}{2}(1 + \beta n_1)\gamma m_H x_1 , \tag{26}$$

$$E_2 = \frac{1}{2}(1 + \beta n_1 \cos \theta_{12} - \beta n_2 \frac{2\sqrt{S}}{x_1 x_2})\gamma m_H x_2 , \tag{27}$$

$$E_3 = \frac{1}{2}(1 + \beta n_1 \cos \theta_{13} + \beta n_2 \frac{2\sqrt{S}}{x_1 x_3})\gamma m_H x_3 \tag{28}$$

26

and we have defined,

$$S \equiv (1 - x_1)(1 - x_2)(1 - x_3). \tag{29}$$

Here, the boost direction is given by a unit vector in terms of two Euler angles,

$$\hat{n} = (n_1, n_2, n_3) = (\cos\phi\sin\theta, \sin\phi\sin\theta, \cos\theta). \tag{30}$$

In our conventions, the initial z-axis is perpendicular to the plane of the decay in the rest frame of the decaying particle, and we chose the x-axis to point in the direction of the quark.

B   Templates at NLO

As discussed in Sec. 3.1, templates were generated in three-particle phase space with a density that reflects the NLO differential decay rate for Higgs decay, Eq. (9). To be counted as a three-particle template, each pair of partons in the final state should have a fractional invariant mass, $y_{ij} > y$, in Eq. (10), where in this study we have taken $y = 0.05$. As shown in Fig. 2, the output of partonic level Monte Carlos suggests that about 20% of events generated in this way will pass this cut. In fact, we can confirm this estimate and calculate NLO y dependence analytically, by integrating the differential rate, Eq. (9) subject to $y_{ij} > y$ for all pairs i, j. Following familiar notation [70], this result divided by the total NLO decay rate is labelled f_3, the three-jet fraction. The explicit expression is

$$\begin{aligned}
f_3 = \frac{C_F\alpha_s}{4\pi} \Bigg[&-4\mathrm{Li}_2\left(2 + \frac{1}{y-1}\right) + 6y\log\left(\frac{1}{y^2} - \frac{3}{y}2\right) \\
&+ \log(y)\left(-4\log\left(2y^2 - 3y + 1\right) + 4\log(y) + 3\right) + (2y - 1)(5y - 7) \\
&+ \log(1 - 2y)(4\log(1 - y) - 3) - 6\tanh^{-1}(1 - 2y)\Bigg] ,
\end{aligned} \tag{31}$$

We can regard the remaining fraction of events as a two-jet fraction, labelled

$$f_2 = 1 - f_3. \tag{32}$$

The fractions f_2 and f_3 are shown in Fig. 13 as a function of y. The points correspond to the reference value, $y = 0.05$ used in our study.
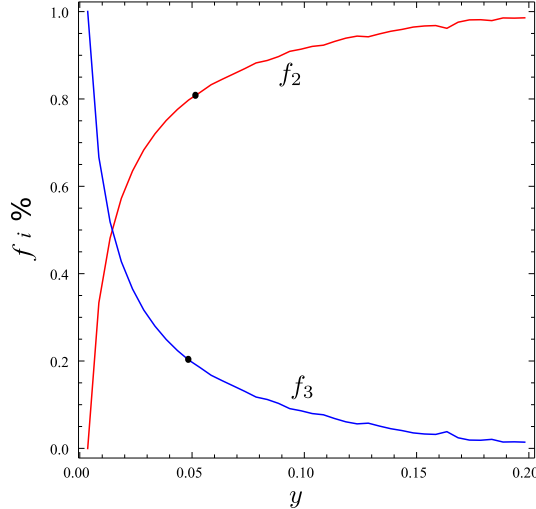
Figure 13: The values of f_3 and f_2 from Eq. (31)-(32). Notice that the choice $y = 0.05$ corresponds to roughly $f_2 = 0.8$ and $f_3 = 0.2$.

References

[1] J. M. Butterworth, J. R. Ellis and A. R. Raklev, JHEP **0705**, 033 (2007) [arXiv:hep-ph/0702150].

[2] D. Benchekroun, C. Driouichi, A. Hoummada, SN-ATLAS-2001-001, ATL-COM-PHYS-2000-020, EPJ Direct 3, 1 (2001); J. M. Butterworth, B. E. Cox and J. R. Forshaw, Phys. Rev. D **65**, 096014 (2002) [arXiv:hep-ph/0201098].

[3] J. M. Butterworth, A. R. Davison, M. Rubin and G. P. Salam, Phys. Rev. Lett. **100**, 242001 (2008) [arXiv:0802.2470 [hep-ph]].

[4] J. M. Butterworth, J. R. Ellis, A. R. Raklev and G. P. Salam, Phys. Rev. Lett. **103**, 241803 (2009) [arXiv:0906.0728 [hep-ph]].

[5] K. Agashe, A. Belyaev, T. Krupovnickas, G. Perez and J. Virzi, Phys. Rev. D **77**, 015003 (2008) [arXiv:hep-ph/0612015].

[6] B. Lillie, L. Randall and L. T. Wang, JHEP **0709**, 074 (2007) [arXiv:hep-ph/0701166].

[7] A. L. Fitzpatrick, J. Kaplan, L. Randall and L. T. Wang, JHEP **0709**, 013 (2007) [arXiv:hep-ph/0701150].

[8] K. Agashe, H. Davoudiasl, G. Perez and A. Soni, Phys. Rev. D **76**, 036006 (2007) [arXiv:hep-ph/0701186].

[9] B. Lillie, J. Shu and T. M. P. Tait, Phys. Rev. D **76**, 115016 (2007) [arXiv:0706.3960 [hep-ph]].

[10] K. Agashe *et al.*, Phys. Rev. D **76**, 115015 (2007) [arXiv:0709.0007 [hep-ph]].

[11] T. Han, S.J. Lee, F. Maltoni, G. Perez, Z. Sullivan, T.M.P. Tait and L. T. Wang, in P. Nath *et al.*, Nucl. Phys. Proc. Suppl. **200-202**, 185 (2010) [arXiv:1001.2693 [hep-ph]].

[12] S. Fleming, A. H. Hoang, S. Mantry and I. W. Stewart, arXiv:0711.2079 [hep-ph]; S. Fleming, A. H. Hoang, S. Mantry and I. W. Stewart, Phys. Rev. D **77**, 074010 (2008) [arXiv:hep-ph/0703207]; A. H. Hoang and I. W. Stewart, arXiv:0808.0222 [hep-ph].

[13] A. Banfi, G. P. Salam and G. Zanderighi, JHEP **0707**, 026 (2007) [arXiv:0704.2999 [hep-ph]].

[14] S. D. Ellis, J. Huston, K. Hatakeyama, P. Loch and M. Tonnesmann, Prog. Part. Nucl. Phys. **60**, 484 (2008) [arXiv:0712.2447 [hep-ph]].

[15] L. G. Almeida, S. J. Lee, G. Perez, I. Sung and J. Virzi, Phys. Rev. D **79**, 074012 (2009) [arXiv:0810.0934 [hep-ph]].

[16] J. Thaler and L. T. Wang, JHEP **0807**, 092 (2008) [arXiv:0806.0023 [hep-ph]].

[17] L. G. Almeida, S. J. Lee, G. Perez, G. Sterman, I. Sung and J. Virzi, Phys. Rev. D **79**, 074017 (2009) [arXiv:0807.0234 [hep-ph]].

[18] J. Thaler and K. Van Tilburg, JHEP **1103**, 015 (2011) [arXiv:1011.2268 [hep-ph]].

[19] J. Thaler and K. Van Tilburg, arXiv:1108.2701 [hep-ph].

[20] G. Gur-Ari, M. Papucci and G. Perez, arXiv:1101.2905 [hep-ph].

[21] D. E. Kaplan, K. Rehermann, M. D. Schwartz and B. Tweedie, Phys. Rev. Lett. **101**, 142001 (2008) [arXiv:0806.0848 [hep-ph]].

[22] G. D. Kribs, A. Martin, T. S. Roy and M. Spannowsky, arXiv:0912.4731 [hep-ph].

[23] S. Chekanov and J. Proudfoot, arXiv:1002.3982 [hep-ph].

[24] A. Katz, M. Son and B. Tweedie, JHEP **1103**, 011 (2011) [arXiv:1010.5253 [hep-ph]].

[25] A. Katz, M. Son and B. Tweedie, Phys. Rev. D **83**, 114033 (2011) [arXiv:1011.4523 [hep-ph]].

[26] B. Bhattacherjee, M. Guchait, S. Raychaudhuri and K. Sridhar, Phys. Rev. D **82**, 055006 (2010) [arXiv:1006.3213 [hep-ph]].

[27] C. Englert, C. Hackstein and M. Spannowsky, Phys. Rev. D **82**, 114024 (2010) [arXiv:1010.0676 [hep-ph]].

[28] C. Hackstein and M. Spannowsky, Phys. Rev. D **82**, 113012 (2010) [arXiv:1008.2202 [hep-ph]].

[29] Y. Cui, Z. Han and M. D. Schwartz, Phys. Rev. D **83**, 074023 (2011) [arXiv:1012.2077 [hep-ph]].

[30] J. Gallicchio, J. Huth, M. Kagan, M. D. Schwartz, K. Black and B. Tweedie, JHEP **1104**, 069 (2011) [arXiv:1010.3698 [hep-ph]].

[31] T. Plehn, M. Spannowsky and M. Takeuchi, JHEP **1105**, 135 (2011) [arXiv:1102.0557 [hep-ph]].

[32] K. Rehermann and B. Tweedie, JHEP **1103**, 059 (2011) [arXiv:1007.2221 [hep-ph]].

[33] M. Jankowiak and A. J. Larkoski, JHEP **1106**, 057 (2011) [arXiv:1104.1646 [hep-ph]].

[34] J. -H. Kim, Phys. Rev. D **83**, 011502 (2011) [arXiv:1011.1493 [hep-ph]].

[35] S. D. Ellis, A. Hornig, C. Lee, C. K. Vermilion and J. R. Walsh, arXiv:1001.0014 [hep-ph]; S. D. Ellis, C. K. Vermilion and J. R. Walsh, arXiv:0912.0033 [hep-ph]; S. D. Ellis, C. K. Vermilion and J. R. Walsh, Phys. Rev. D **80**, 051501 (2009) [arXiv:0903.5081 [hep-ph]].

[36] D. Krohn, J. Thaler and L. T. Wang, JHEP **1002**, 084 (2010) [arXiv:0912.1342 [hep-ph]].

[37] D. E. Soper and M. Spannowsky, arXiv:1005.0417 [hep-ph].

[38] D. E. Soper and M. Spannowsky, Phys. Rev. D **84**, 074002 (2011) [arXiv:1102.3480 [hep-ph]].

[39] A. Falkowski, C. Grojean, A. Kaminska, S. Pokorski and A. Weiler, JHEP **1111**, 028 (2011) [arXiv:1108.1183 [hep-ph]].

[40] Y. Bai and J. Shelton, arXiv:1107.3563 [hep-ph].

[41] A. Hook, M. Jankowiak and J. G. Wacker, arXiv:1102.1012 [hep-ph].

[42] G. D. Kribs, A. Martin and T. S. Roy, arXiv:1012.2866 [hep-ph].

[43] S. Yang and Q. -S. Yan, arXiv:1111.4530 [hep-ph].

[44] C. Englert, J. Jaeckel, E. Re and M. Spannowsky, arXiv:1111.1719 [hep-ph].

[45] J. Gallicchio and M. D. Schwartz, Phys. Rev. Lett. **107**, 172001 (2011) [arXiv:1106.3076 [hep-ph]].

[46] B. Bellazzini, C. Csaki, J. Hubisz and J. Shao, Phys. Rev. D **83**, 095018 (2011) [arXiv:1012.1316 [hep-ph]].

[47] I. Sung, Phys. Rev. D **80**, 094020 (2009) [arXiv:0908.3688 [hep-ph]].

[48] J. Gallicchio and M. D. Schwartz, arXiv:1001.5027 [hep-ph].

[49] G. Aad *et al.* [Atlas Collaboration], Phys. Rev. D **83**, 052003 (2011) [arXiv:1101.0070 [hep-ex]].

[50] T. Aaltonen *et al.* [CDF Collaboration], arXiv:1106.5952 [hep-ex].

[51] R. Alon, E. Duchovni, G. Perez, A. P. Pranko and P. K. Sinervo, arXiv:1101.3002 [hep-ph].

[52] A. Abdesselam *et al.*, Eur. Phys. J. C **71**, 1661 (2011) [arXiv:1012.5412 [hep-ph]].

[53] L. G. Almeida, R. Alon and M. Spannowsky, arXiv:1110.3684 [hep-ph].

[54] L. G. Almeida, S. J. Lee, G. Perez, G. Sterman, I. Sung, Phys. Rev. **D82**, 054034 (2010). [arXiv:1006.2035 [hep-ph]].

[55] J. M. Maldacena, JHEP **0305**, 013 (2003) [arXiv:astro-ph/0210603].

[56] G. Sterman, Phys. Rev. D **19**, 3135 (1979).

[57] Y. S. Lai and B. A. Cole, arXiv:0806.1499 [nucl-ex].

[58] G. Sterman, Phys. Rev. D **17**, 2789 (1978).

[59] T. Aaltonen *et al.* [The CDF Collaboration], Phys. Rev. D **81**, 052011 (2010) [arXiv:1002.0365 [hep-ex]].

[60] T. Sjostrand, S. Mrenna and P. Skands, Comput. Phys. Commun. **178**, 852 (2008) [arXiv:0710.3820 [hep-ph]].

[61] T. Gleisberg, S. Hoche, F. Krauss, M. Schonherr, S. Schumann, F. Siegert and J. Winter, JHEP **0902**, 007 (2009) [arXiv:0811.4622 [hep-ph]].

[62] S. Hoeche, F. Krauss, S. Schumann and F. Siegert, JHEP **0905**, 053 (2009) [arXiv:0903.1219 [hep-ph]].

[63] F. Maltoni and T. Stelzer, JHEP **0302**, 027 (2003) [arXiv:hep-ph/0208156]; T. Stelzer and W. F. Long, Comput. Phys. Commun. **81**, 357 (1994) [arXiv:hep-ph/9401258]. JHEP **0709**, 028 (2007) [arXiv:0706.2334 [hep-ph]].

[64] T. Sjostrand, S. Mrenna and P. Skands, JHEP **0605**, 026 (2006) [arXiv:hep-ph/0603175].

[65] M. L. Mangano, M. Moretti, F. Piccinini and M. Treccani, JHEP **0701**, 013 (2007) [arXiv:hep-ph/0611129].

[66] M. Cacciari and G.P. Salam, Phys. Lett. B 641 (2006) 57 [hep-ph/0512210]; M. Cacciari, G.P. Salam and G. Soyez, http://fastjet.fr/

[67] M. Cacciari, G. P. Salam and G. Soyez, JHEP **0804**, 063 (2008) [arXiv:0802.1189 [hep-ph]].

[68] ATLAS Detector and physics performance TDR, CERN-LHCC-99-14; CMS Physics and performance TDR, Volume II: CERN-LHCC-2006-021.

[69] M. Backovic, J. Juknevich, G. Perez and J. Winter, In preparation.

[70] R.K. Ellis, W.J. Stirling and B.R. Webber, *QCD and Collider Physics*, Cambridge University Press (1996)

[71] G. Gur-Ari, M. Field, D.A. Kosower, L. Manelli and G. Perez, In preparation.

[72] D. Krohn, J. Thaler and L. -T. Wang, JHEP **0906**, 059 (2009) [arXiv:0903.0392 [hep-ph]].