Top Jets at the LHC

Leandro G. Almeida, Seung J. Lee, Gilad Perez, Ilmo Sung & Joseph Virzi

Top Jets at the LHC

Leandro G. Almeida[a], **Seung J. Lee**[a,b], **Gilad Perez**[a,b],
Ilmo Sung[a] **and Joseph Virzi**[c]

[a] *C. N.Yang Institute for Theoretical Physics, Stony Brook University,*
Stony Brook, NY 11794-3840, USA

[b] *Department of Particle Physics, Weizmann Institute of Science, Rehovot 76100, Israel*

[c] *University of California; LBNL, Physics Division, 1 Cyclotron Rd., Berkeley, CA 94720, USA*

Abstract

We investigate the reconstruction of high p_T hadronically-decaying top quarks at the Large Hadron Collider. One of the main challenges in identifying energetic top quarks is that the decay products become increasingly collimated. This reduces the efficacy of conventional reconstruction methods that exploit the topology of the top quark decay chain. We focus on the cases where the decay products of the top quark are reconstructed as a single jet, a "top-jet". The most basic "top-tag" method based on jet mass measurement is considered in detail. To analyze the feasibility of the top-tagging method, both theoretical and experimental aspects of the large QCD jet background contribution are examined. Based on a factorization approach, we derive a simple analytic approximation for the shape of the QCD jet mass spectrum. We observe very good agreement with the Monte Carlo simulation. We consider high-p_T $t\bar{t}$ production in the Standard Model as an example, and show that our theoretical QCD jet mass distributions can efficiently characterize the background via sideband analyses. We show that with 25 fb^{-1} of data, our approach allows us to resolve top-jets with $p_T \geq 1$ TeV, from the QCD background, and about 1.5 TeV top-jets with 100 fb^{-1}, without relying on b-tagging. To further improve the significance we consider jet shapes (recently analyzed in 0807.0234 [hep-ph]), which resolve the substructure of energy flow inside cone jets. A method of measuring the top quark polarization by using the transverse momentum of the bottom quark is also presented. The main advantages of our approach are: (i) the mass distributions are driven by first principle calculations, instead of relying solely on Monte Carlo simulation; (ii) for high p_T jets ($p_T \geq 1$ TeV), IR-safe jet shape variables are robust against detector resolution effects. Our analysis can be applied to other boosted massive particles such as the electroweak gauge bosons and the Higgs.

1   Introduction

The Large Hadron Collider (LHC) is expected to uncover some of the most interesting mysteries of nature. We expect to probe the underlying principles of electroweak symmetry breaking (EWSB) and what stabilizes the weak scale against radiative corrections from unknown microscopic dynamics. Due to its large mass, the top quark induces the most severe contributions to the Higgs quadratic divergence. Furthermore, in almost every known natural model of EWSB, the top sector plays a crucial role in breaking the EW symmetry. Thus, the top sector might hold a key to a new physics (NP) discovery. Many interesting models of EWSB predict new particles with mass $\sim$ TeV scale. In several known examples, the new particles decay into highly boosted top quark pairs ($pp \to X \to t\bar{t}$), or other decay chains containing a single top quark ($pp \to X \to t\,Y$). In addition, the Standard Model (SM) predicts that the LHC will produce more than 10^5 top quarks with $p_T \geq 1\,\text{TeV}$, significantly enhancing our ability to study high p_T tops and resolving beyond the SM dynamics.

Top quarks decay dominantly into hadronic final states ($t \to bW \to bq\bar{q}$) with a branching ratio $\sim 2/3$, providing potentially enhanced statistics. In the present work, we focus on highly boosted top quarks (decaying through the hadronic channel), and on the dominant QCD jet background. We refer to a top quark that decays hadronically as a *hadronic top*. For moderately boosted top quarks ($p_T \sim 500$ GeV), conventional top quark reconstruction methods, which exploit the decay chain topology, remain adequately efficient (see *e.g.* [1]). As the top quark p_T approaches $1\,\text{TeV}$, the situation significantly changes [2, 3, 4, 5].§ The average separation of the top quark decay products approaches the limits of reliable jet reconstruction (cone size $R \sim 0.4$), and starts to encroach upon the detector resolution ($R \sim 0.1$). As a result, the efficiency of conventional reconstruction methods drops quickly. The performance of b-tagging and light jet rejection is expected to drop substantially in this kinematic regime. At present, there is very little published data on b-tagging at $p_T \sim 1$ TeV [7]. We perform our analysis without accounting for the possible benefits of b-jet identification.¶

We turn our focus away from this family of "conventional" reconstruction methods. We examine the situation where the decay products of at least one top quark are reconstructed as a single jet, or *top-jet*. In semileptonic $t\bar{t}$ events, for example, the leptonic top may still be reconstructed via semi-conventional reconstruction methods, giving up on lepton isolation cuts [2], see also [4, 8]. These methods call for further extensive study due to expected reducible backgrounds and uncertainties related to the ability to measure the collimated semi-leptonic top mass (dileptonic $t\bar{t}$ events are also analyzed in [9]). Hadronic top, on the other hand, will give rise to a top-jet. There will still be some small, but non-negligible, number of $t\bar{t}$ events where one of the top quarks reconstructs as a top-jet, but the other top

---

§For earlier works in the case of boosted EW bosons see also [6].

¶The possibility of b-tagging jets, when the top quark reconstructs to 2 (or more) jets, one of which has a mass $\sim M_W$ and the sum of the two jets has a mass $\sim m_t$, is outside the main focus of this paper.

quark can be reconstructed via conventional methods (or semi-conventional methods where one of the tops is manifested as a two-jet object). In this paper, we focus on the top-jet itself as a means of identifying $t\bar{t}$ events. The main reasoning behind that is as follows:

(i) We find that for $p_T > 1\,\mathrm{TeV}$ the majority of hadronic-tops are manifested as top-jets, even for cone size as small as $R = 0.4$. Thus, it is clear that our tools will be applicable for a wide range of top momenta.

(ii) The distributions and shapes of both background and signal can be understood via first principle calculations as shown in this study and in Ref. [10]. It may allow for a cleaner analysis, in the sense that a more direct contact between actual data (expected to arrive soon) and the microscopical theory can be made.

Apart from substructure, to leading order, top jets provide four pieces of information, namely its energy, two angles and mass (just as any QCD jet, ignoring the possibility of b-tagging). Without a mass cut, the QCD jet background swamps the hadronic top signal by orders of magnitude. The most basic tagging method after giving up conventional methods is to use the jet mass as a discriminator between the QCD background and the hadronic top signal; the high-p_T top-jet mass distribution should peak around the top mass while the QCD jet mass distribution peaks near zero. However, using a jet mass as a discriminator is more complicated for several reasons. Due to radiation, QCD jets acquire a large tail in the mass distribution. The cross section for acquiring large jet mass, for example near the top mass, increases substantially with p_T and cone size. Top-jets also broaden due to radiation, hardening their jet mass distribution.$^{\|}$ Furthermore, a finite jet reconstruction cone size will not always capture all the daughters of the top quark decay chain, thus softening its mass distribution. The net effect is a smearing of the expected naive, broadened, Breit-Wigner distribution for the top jet mass distribution. Detector effects further smear the distribution, making the above idealized description unrealistic.

Nevertheless, jet mass cuts should retain some rejection power against the QCD background [12, 13, 14]. Our study addresses this issue in both quantitative and qualitative manner, by considering the experimental and theoretical aspects of the analysis. On the theoretical front, based on a factorization approach, we derive a simple approximation for the shape of the QCD jet mass spectrum. We demonstrate that there is good agreement between our simple analytic predictions and Monte Carlo (MC) results. We are able to compute from first principle various features related to a jet mass cut. We evaluate its significance in the form of a semi-analytical expression for the rejection power and show that it is independent of pseudorapidity. We provide a quantitative study of the distribution of the signal and background, via MadGraph/MadEvent [15, 16, 17](MG/ME) and Sherpa [18]. We consider the detector resolution by using transfer functions [19], smearing jets according to a profile obtained from full Geant4 Atlas simulation. Transfer functions

---

$^{\|}$For a detailed recent study see [11] and references therein.

provide a versatile mechanism to explore such effects as shifts in jet energy scale (JES), etc.

We apply the results of our studies to analyze boosted SM top quark pair production, an important discovery channel for NP [2, 3, 4, 8, 20, 21, 22, 23, 24]. To put results into perspective, we use both $25\,\mathrm{fb}^{-1}$ and $100\,\mathrm{fb}^{-1}$ of integrated luminosity as reference luminosities. At this time, these correspond to many years of data taking. We show that using single- and double-tagging methods with our jet functions (defined below) to analyze jet mass distributions, we can significantly separate the Standard Model $t\bar{t}$ signal from the QCD background. Our theoretical QCD jet mass distributions can efficiently characterize the background via sideband analyses. With 25 fb^{-1} of data, our approach allows us to resolve 1 TeV top-jets from the QCD background, and about 1.5 TeV top-jets with 100 fb^{-1}, if we exploit the kinematics of the so-called "away" side of the event, without relying on b-tagging. The essence of the away side mass cut is that it preferentially keeps the $t\bar{t}$ signal over the background. We analyze the mass distribution in more detailed manner, as simple counting methods are inadequate. As described above, the $t\bar{t}$ signal is expected to exhibit pronounced structure near the top quark mass. In order to resolve this "peak" against the QCD background, we need to understand the shape of both the $t\bar{t}$ signal and the QCD background. To characterize the background we perform a sideband analysis to reduce contamination by the signal. Our theory-driven ansatz for the QCD background is an admixture of quark- and gluon-jet functions, the coefficients of which we analyze by fitting in the sidebands (outside the top mass window). We interpolate the results of the fit into the top mass window ($140\,\mathrm{GeV} \leq m_J \leq 210\,\mathrm{GeV}$). Armed with *shapes* for the signal and background, we fit them into the data to obtain the normalization constants. These normalization constants are the magnitude of the signal and background. The errors associated with the normalization provides a measure of the significance of the measurement.

To further improve the significance we consider jet shapes [10], which resolve substructure of energy flow inside cone jets. In a companion paper [10], we explore the possibility that, requiring a large jet mass, perturbative predictions for jet shapes differ between jets that originate from the decay of heavy particles, and those which result from the showering of light quarks and gluons. With such additional handles, we might have a chance to distinguish boosted $t\bar{t}$ signal from the QCD background even at a smaller integrated luminosities.** We discuss jet substructure later in the text.

We turn our attention to the use of b-jets as spin analyzers for the top quarks. For highly boosted top quarks, chirality is approximately equal to helicity and is conserved to a good approximation. Information about the top chirality is encoded in the angular distribution of the decay products [2, 28, 29]. Naively, one would argue that for hadronic tops this information is inaccessible due to collimation and the absence of leptons which are known to be good spin analysers [28, 29]. We explore the possibility of using p_T of the

---

**There are other approaches dealing with a similar situation in a different perspective in recent literature [5, 25, 26, 27].

b-quark for measuring the top quark polarization, which is important for exploring NP. For this, we explore the case when at least one of the boosted top quark can be resolved into more than two jets. We also consider the possibility of using p_T of the lepton for measuring the top quark polarization for semi-leptonically decaying tops.

This work has two main focal points, namely QCD jet mass distributions and hadronic $t\bar{t}$ signal, and is structured as follows. In the next section, we discuss the MC generation and detector simulation. In section 3 we focus on highly boosted QCD jets. The jet mass distribution is examined numerically, via MC methods, and analytically, via jet functions. The salient points of the jet functions are introduced, leaving detailed derivations for the Appendix. Section 4 discusses the top-jet signal. In section 5, we compare high p_T hadronic $t\bar{t}$ events with QCD jets. In section 6 we discuss jet shapes [10], which can be used as additional discriminants against the background. Section 7 discusses the hadronic top quark polarization by using the transverse momentum of the bottom quark. We conclude in section 8.

2   Event Simulation

2.1   Monte Carlo Generation

The Sherpa [18] and MG/ME (version 4) [15, 17] MC generators were used to produce $t\bar{t}$ and QCD jet events, with parameters appropriate to the LHC. To effect partonic level cuts during the generation of QCD jets (p_T (≥ 1 parton) ≥ 800 GeV), we used customized code provided by the Sherpa authors applicable to Sherpa V1.1.0. For technical reasons, $t\bar{t}$ events were generated using Sherpa version 1.0.9, whereas QCD jet events were generated with Sherpa version 1.1.0. MG/ME interfaces to Pythia V6.4 (for parton shower and fragmentation) [30]. For jet reconstruction, we used SISCone V1.3 [31] for both Sherpa and MG/ME. Cross sections are calculated to leading order. Jets are defined via the cone algorithm [32] with $R = 0.4$ and $R = 0.7$, referred to as C4 and C7, respectively. Jets have $p_T > 50$ GeV and $|\eta| \leq 2$. At the hard scatter level, final state partons are required to have $p_T \geq 20$ GeV. For MG/ME events, the final state partons have $|\eta| \leq 4.5$.

We do not account for pile-up effects nor characterize the underlying event. Efficiencies for triggering and reconstruction of jets at these energies are very close to unity; the corrections are negligible and are not considered. The strong coupling constant was allowed to run. Throughout the analysis, we used Sherpa V1.0.9 with CTEQ6M parton distribution functions (PDF) [33]. Comparisons to MG/ME were made whenever appropriate, and also occasionally to Pythia (version 8.1) [34] for $2 \rightarrow 2$ process without matching. In such cases, the distinct curves are marked accordingly. The events used in the analysis were inclusive, i.e. $pp \rightarrow t\bar{t}(j)$ and $pp \rightarrow jj(j)$, with matching (see [35] for a detailed discussion): modified MLM [36] for MG/ME and CKKW [37] for Sherpa.

2.2 Cross Sections

In table 1 we present cross sections for producing final state (hadronic level) jets with $p_T \geq 1\,\text{TeV}$ for the different MC simulations. There are large uncertainties in the cross sections, due to differences between the MLM and CKKW matching, between MC generators, and between PDFs. It is outside the scope of this paper to explore the reasons behind these differences. We estimate a 100% systematic uncertainty associated with the $t\bar{t}$ cross section, and a 20% systematic uncertainty in the QCD jet cross section.

Process	Generator	PDF	Matching	Cross Section
$pp \to t\bar{t}(j)$	SHERPA 1.0.9	CTEQ6M	CKKW	135 fb
$pp \to t\bar{t}(j)$	SHERPA 1.1.2	CTEQ6M	CKKW	149 fb
$pp \to t\bar{t}(j)$	MG/ME 4	CTEQ6M	MLM	68 fb
$pp \to t\bar{t}(j)$	MG/ME 4	CTEQ6L	MLM	56 fb
$pp \to t\bar{t}$	Pythia 6.4	CTEQ6L	-	157 fb
$pp \to t\bar{t}$	Pythia 8.1	CTEQ6M	-	174 fb
$pp \to jj(j)$	SHERPA 1.1.0	CTEQ6M	CKKW	10.2 pb
$pp \to jj(j)$	MG/ME 4	CTEQ6L	MLM	8.54 pb
$pp \to jj(j)$	MG/ME 4	CTEQ6M	MLM	9.93 pb
$pp \to jj$	Pythia 6.4	CTEQ6L	-	13.7 pb
$pp \to jj$	Pythia 8.1	CTEQ6M	-	13.3 pb

Table 1: Cross sections for producing final state $R = 0.4$ leading cone jets with $p_T \geq 1\,\text{TeV}$ and $|\eta| \leq 2$. Generation level cuts were imposed as follows. Final state partons from the hard scatter were required to have $p_T \geq 20\,\text{GeV}$. For MG/ME, final state partons have $|\eta| \leq 4.5$. Processes with a trailing (j) suffix indicate that $2 \to 2$ and $2 \to 3$ processes are represented.

2.3 Modelling Detector Effects

A transfer function, trained with full ATLAS detector simulation on high p_T jet and high p_T $t\bar{t}$ samples, was used to map particle level jets (Atlas truth jet reconstruction) onto a full simulation model [19]. Transfer functions work by feeding back the differences between the target collection (Full Simulation) and the source collection (Truth Jets). The differences and efficiencies are stored as distributions, in the form of histograms, and binned in p_T and η. We refer to the collection of the smearing distributions as a *transfer function*. It is important to note that transfer functions are applicable on events with similar jet multiplicity and topology. We applied the transfer function (trained on Atlas truth jets) to SISCone truth jets, which preserve the salient characteristics of the Atlas truth jets. We used the transfer function to effect p_T and mass smearing, but not reconstruction efficiency. At the energies considered in this paper, reconstruction efficiency is very close to unity. In

5

summary, the results of the transfer function should be viewed simply as realistic detector smearing.

In this paper, a jet is transferred as follows. The transverse momentum and mass of truth-level jets are smeared according to the appropriate distribution. For the purposes of modeling the effects of the JES, the means of the p_T distributions are shifted accordingly, without cross correlation to the mass smearing. This is a subtle point. Depending on the reconstruction mechanism, reported jet masses may depend proportionally on the JES; a JES shift results in a jet mass shift. In our study of the effects of the JES, we do not make a correlation between the p_T and mass distributions. This effect is much smaller, and such precision is not warranted in these studies.
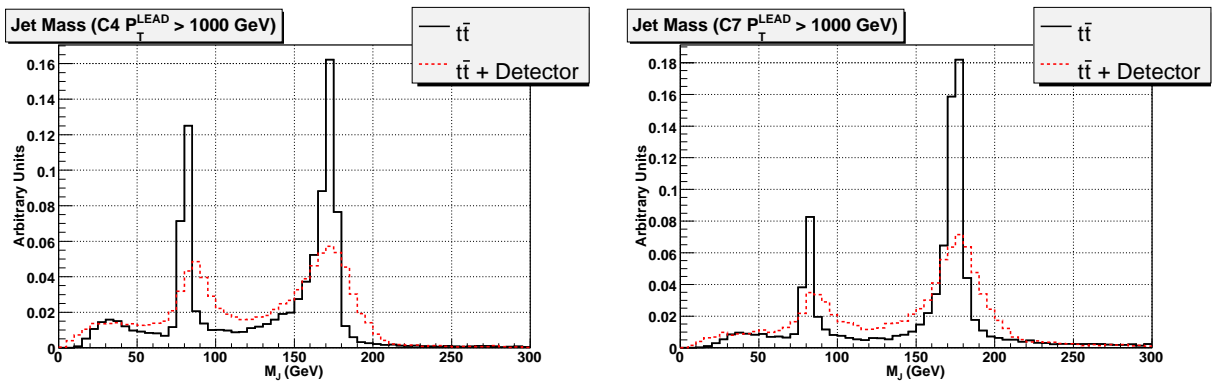

Figure 1: We compare the mass distribution of the leading jet $\left(p_T^{lead} \geq 1000\,\text{GeV}\right)$ for the $t\bar{t}$ signal with (the red dotted curve) and without (the black solid curve) leading detector effects. The plot on the left corresponds to C4 jets; the plot on the right corresponds to C7 jets.

In Fig. 1, we compare the $t\bar{t}$ jet mass distributions for C4 and C7 jets, with and without detector smearing, for $p_T^{lead} \geq 1000\,\text{GeV}$. We see, as expected, that due to the finite cone size even the top jet mass distribution is far from the naive Breit-Wigner shape. In cases where the outgoing b quark is outside the cone, we expect that the top jet mass to be peaked around the W mass. In cases where one of the quarks from the W decay is outside the cone we expect a smooth distribution with a typical invariant mass of roughly $m_t/\sqrt{2}$, etc. These effects are present even at the truth level, without detector effects. The black curve shows a smooth distribution with a spurious peak around the W mass. The red curve demonstrates how the detector effects further smear the top jet mass distribution.

3   QCD Jet Background

If jet mass methods are to be viable, we must be able to characterize the dominant QCD jet background [38]. One of the primary points in this work is that we are able to understand

the QCD jet background *analytically* as well as through MC simulations. In this section, we present the summary of our analytic calculations of the QCD jet mass distribution based on the factorization formalism [39, 40], which is presented in the Appendix. We compare our theoretical prediction with simulated MC data. Note that the final states, which induce the jet masses, simulated by MC event generators are much more complicated (due to radiation, showering etc.) than our simple two body final states. Yet, as we shall see, we can consistently describe the simulated MC data.

3.1 Analytic Prediction

We are interested in looking at the following processes:

$$H_a(p_a) + H_b(p_b) \to J_1(m_{J_1}^2, p_{1,T}, R) + X$$

$$H_a(p_a) + H_b(p_b) \to J_1(m_{J_1}^2, p_{1,T}, R) + J_2(m_{J_2}^2, p_{2,T}, R) + X$$

where, H_i are the initial hadrons, p_i being the corresponding momenta, and the final states include jets in the direction of the outgoing partons of the underlying process, with a fixed jet mass, m_{J_i}, "cone size" $R^2 = \Delta\eta^2 + \Delta\phi^2$ and tranverse momenta, $p_{i,T}$.

We begin with the factorized hadronic cross section for single inclusive jet processes,

$$\frac{d\sigma_{H_A H_B \to J_1 X}(R)}{dp_T dm_J d\eta} = \sum_{abc} \int dx_a\, dx_b\, \phi_a(x_a)\, \phi_b(x_b) \frac{d\hat{\sigma}_{ab \to cX}}{dp_T dm_J d\eta}(x_a, x_b, p_T, \eta, m_J, R),$$
(3.1)

which in the limit of small R, we can further factorize into (see Appendix B),

$$\frac{d\sigma_{H_A H_B \to J_1 X}(R)}{dp_T dm_J d\eta} = \sum_{abc} \int dx_a\, dx_b\, \phi_a(x_a)\, \phi_b(x_b) H_{ab \to cX}(x_a, x_b, p_T, \eta, R)$$
$$\times J_1^c(m_J, p_T, R).$$
(3.2)

The factorization and renormalization scales are chosen to be p_T, ϕ_i is the PDF for the initial hadrons, $H_{ab \to cX}$ denotes the perturbative cross section, and J^c denotes jet functions, whose matrix elements are defined in Appendix A (see *e.g.* [41] for recent reviews and references therein). Furthermore the J^cs are, by definition, normalized as

$$\int dm_J\, J^c = 1.$$
(3.3)

We have used the fact that the jet functions do not depend on η in the leading expansion (see Appendix A). Therefore, we can write Eq. (3.2) for the hadronic cross section as

$$\frac{d\sigma(R)}{dp_T dm_J} = \sum_c J^c(m_J, p_T, R) \frac{d\hat{\sigma}^c(R)}{dp_T},$$
(3.4)

7

where c represents the flavour of the jet, and where

$$\frac{d\hat{\sigma}^c(R)}{dp_T} = \sum_{ab} \int dx_a\, dx_b\, \phi_a\, \phi_b \int d\eta \int dm_J\, \frac{d\hat{\sigma}_{ab \to cX}(R)}{dp_T dm_J d\eta}\,. \tag{3.5}$$

We employ the jet functions given in the Appendix by Eqs. (A.14) and (A.16), for fixed jet mass and R at the next-to-leading order (NLO) with running coupling effects. As we will see below, these results are consistent with the MC data for sufficiently large $(m_J \geq \mathcal{O}(100\,\mathrm{GeV}))$ jet masses.

At the lower end of the jet mass spectrum, where $m_J \ll p_T R$, the jet mass distribution is dominated by higher order corrections and non-perturbative physics [41], which are beyond the scope of our work, as our interest lies in the region of high jet mass. We note this causes complications when trying to predict the moments of the mass distributions, such as the mean and RMS, unless we introduce a lower cutoff on the mass.

In the Appendix, we provide the full NLO result for the jet function in term of θ_S, the angle of the softer particle with respect to the jet axis. These exact results can be approximated by the eikonal approximation introduced in Appendix B as

$$\begin{aligned}
J^{(eik),c}(m_J, p_T, R) &= \alpha_{\mathrm{S}}(p_T)\frac{4\,C_c}{\pi m_J} \log\left(\frac{1}{z}\tan\left(\frac{R}{2}\right)\sqrt{4-z^2}\right) \tag{3.6}\\
&\simeq \alpha_{\mathrm{S}}(p_T)\frac{4\,C_c}{\pi m_J} \log\left(\frac{R\,p_T}{m_J}\right)\,,
\end{aligned}$$

where $\alpha_{\mathrm{S}}(p_T)$ is the strong coupling constant at the appropriate scale, $z = \frac{m_J}{p_T}$, c represents the flavour of the parton which initiated the jet and C_c equals $C_F = 4/3$ for quarks, and $C_A = 3$ for gluons. These expressions agree with the full NLO jet functions to the level of about 1% and 10% for quark and gluon initiated jets in the region of the top mass window, respectively (checked for $R = 0.4$ and 0.7 and $p_T \gtrsim 1\,\mathrm{TeV}$).

We can interpret the jet function as a probability density functions for a jet with a given p_T to acquire a mass between m_J and $m_J + \delta m_J$. Our rather simple treatment is valid for the higher end of the jet mass spectrum (above $m_J \sim \mathcal{O}(100\,\mathrm{GeV})$), where NLO perturbative calculation captures the dominant physics. In Fig. 2 we show the gluon jet mass distribution from (A.16) with running (red, dashed), and fixed (blue, dotted) coupling, along with the eikonal jet function (green, dashed-dotted) with fixed coupling. The fixed scales are chosen to be p_T. For reference we also superimpose in the Fig. a $1/m_J$ curve which has the same dimension as that of our jet functions and is roughly of the form of the soft function (*cf* Appendix B). It is remarkable that our theory curves are significantly different from simple $1/m_J$ curve whose normalization is chosen such that this curve overlaps with our theory curves around the top mass. This indicates that logarithmic factor is very important in our theory prediction. Note that at lower masses the running is much harder than the fixed cases since the configurations associated with this mass region have lower k_T (the radiated gluon momenta), leading to a larger α_{S}. Also, the eikonal
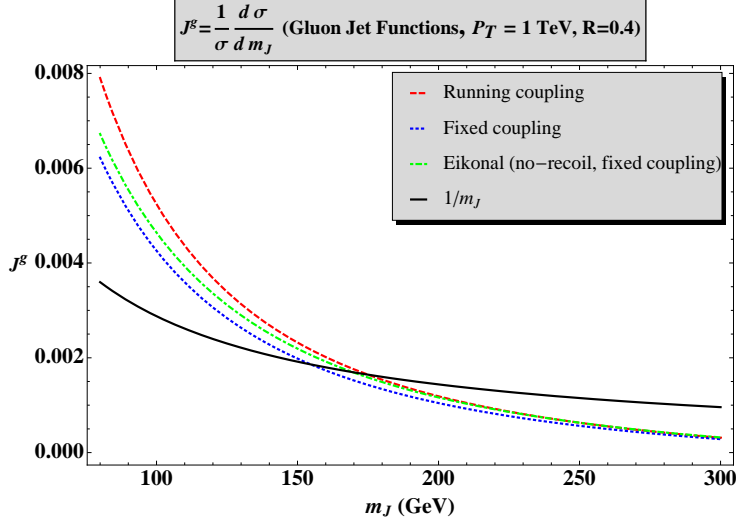
8

Figure 2: Various theoretical gluon-jet mass distributions, along with a $1/m_J$ curve, are plotted for $p_T = 1\,\text{TeV}$ and $R = 0.4$. Plotted are the jet mass distribution from (A.16) with running (red, dashed), and fixed (blue, dotted) coupling, along with the eikonal jet function (green, dashed-dotted) with fixed coupling. For the jet functions with no running the scales were chosen be p_T.

approximation is equivalent to a no recoil approximation, thus resulting overall in a harder process than the result in Eq. (A.16) at fixed scales.

For the purpose of comparing the mass distributions obtained from jet functions and the MC simulations, Eq. (3.5) can be matched to $(d\sigma^c(R)/dp_T)_{\text{MC}}$ obtained from MC, leading to the following relation,

$$\frac{d\sigma^c_{pred}(R)}{dp_T dm_J} = J^c\left(m_J, p_T, R\right)\left(\frac{d\sigma^c\left(R\right)}{dp_T}\right)_{\text{MC}}, \tag{3.7}$$

for the prediction of quark and gluon jet mass distribution based on perturbative calculated jet functions, Eqs. (A.14) and (A.16). Note, however, that this would require us to split the MC output in terms of the parton flavours c, which for realistic simulation leads to ambiguities especially when matching is used. Therefore, for our analysis, instead, we use the analytic result to suggest bounds for the "data" distribution from the MC. There is, however, no a posteriori way to determine the flavour which initiated the jet (as with real data). Thus, we write

$$\frac{d\sigma_{pred}(R)}{dp_T dm_J}_{\ upper\ bound} = J^g\left(m_J, p_T, R\right)\sum_c\left(\frac{d\sigma^c\left(R\right)}{dp_T}\right)_{\text{MC}}, \tag{3.8}$$

$$\frac{d\sigma_{pred}(R)}{dp_T dm_J}_{\ lower\ bound} = J^q\left(m_J, p_T, R\right)\sum_c\left(\frac{d\sigma^c\left(R\right)}{dp_T}\right)_{\text{MC}}, \tag{3.9}$$
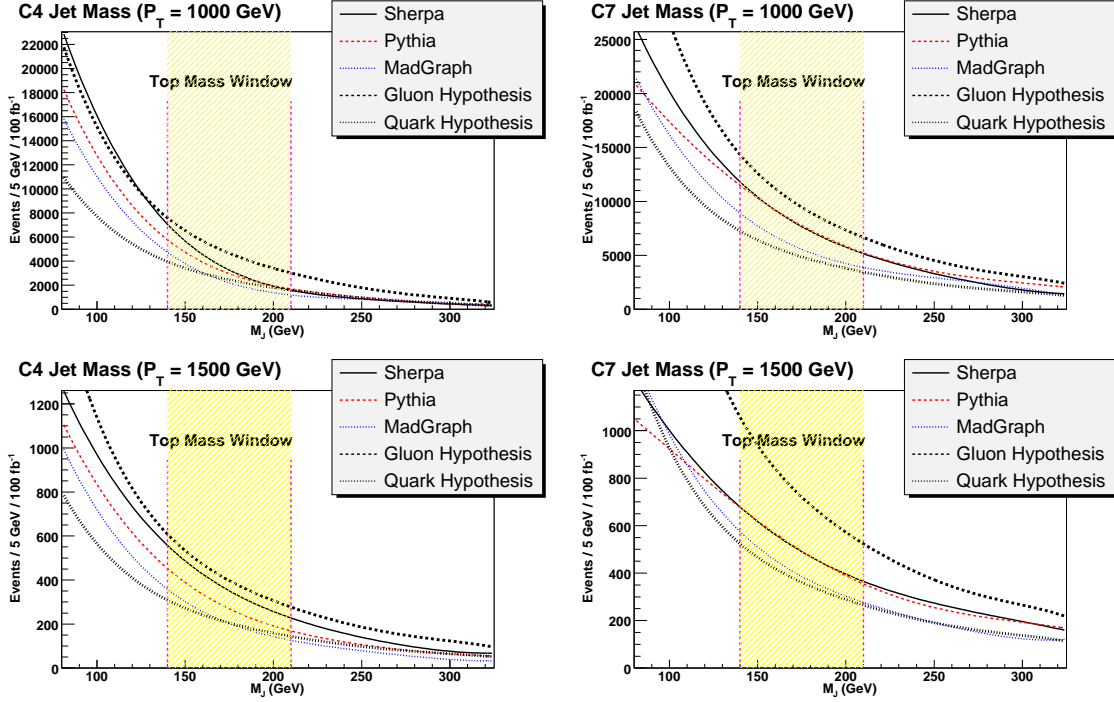
9

Figure 3: The jet mass distributions for Sherpa, Pythia and MG/ME are plotted for different p_T and jet cone sizes. The quark and gluon mass distributions from the jet functions are overlaid, using Eqs. (3.8) and (3.9). The upper left plot corresponds to $950\,\text{GeV} \le p_T \le 1050\,\text{GeV}$ and $R = 0.4$. The upper right plot corresponds to $950\,\text{GeV} \le p_T \le 1050\,\text{GeV}$ and $R = 0.7$. The lower left plot corresponds to $1450\,\text{GeV} \le p_T \le 1550\,\text{GeV}$ and $R = 0.4$. The lower right plot corresponds to $1450\,\text{GeV} \le p_T \le 1550\,\text{GeV}$ and $R = 0.7$.

exploiting the fact that $J^g > J^q$ in the region of high jet mass, as can be seen in Eq. (3.7).

3.2   Jet Function, Theory vs. MC Data

In this section, we compare a set of theory-based bounds for the jet mass distribution to the mass distribution obtained via MC event generators. This part contains one of our main results, where we demonstrate that our theoretical predictions are in agreement with the MC data. In Fig. 3, we compare the quark and gluon jet mass distributions from Eqs. (3.8) and (3.9) to the distributions from different MC generators (MG/ME, Sherpa and Pythia). We perform this comparison at fixed p_T, since we are interested in the relative shapes of these distributions around the top mass window. Note that above $m_J \sim \mathcal{O}(100\,\text{GeV})$, the *shapes* of three MC distributions are very similar. Sherpa and MG/ME distributions interpolate between the quark jet function (lower bound) and the gluon jet function (upper bound) as expected. For the purposes of comparing shapes, Pythia and MG/ME are rescaled so that their total cross sections agree with Sherpa. This cross section scaling does

not affect the predictive quality of the theory curves, since it affects both sides of Eqs. (3.8) and (3.9). The scaling allows us to present the results from the different event generators on a single plot. Note, as mentioned before, that for $m_J \ll p_T R$, higher order corrections will contribute, pushing the distribution down, with a Sudakov-like suppression, which can be seen in the lower mass region for $p_T = 1.5$ TeV and $R = 0.7$.

In a typical experimental setup, a lower cut over p_T will be assumed and the distributions will be integrated above that p_T^{min} cut. Thus we can integrate over the appropriate region on Eq. (3.7), which leads to the analog of Eqs. (3.8) and (3.9) for the p_T-integrated jet mass cross section,

$$\frac{d\sigma_{pred}^c(R)}{dm_J} \quad = \quad \int_{p_T^{min}}^{\infty} dp_T \, J^c\left(m_J, p_T, R\right) \sum_{c'} \left(\frac{d\sigma^{c'}(R)}{dp_T}\right)_{\text{MC}}, \qquad (3.10)$$

where J^c is defined as before. The MC differential cross section is obtained by summing over the contributions from both quark and gluon jets. Therefore, the cross section's shape is characterized by an admixture of quark and gluon jets and should interpolate between the two curves, $c = q$ and g. In Fig. 4, we compare leading jet mass distribution for
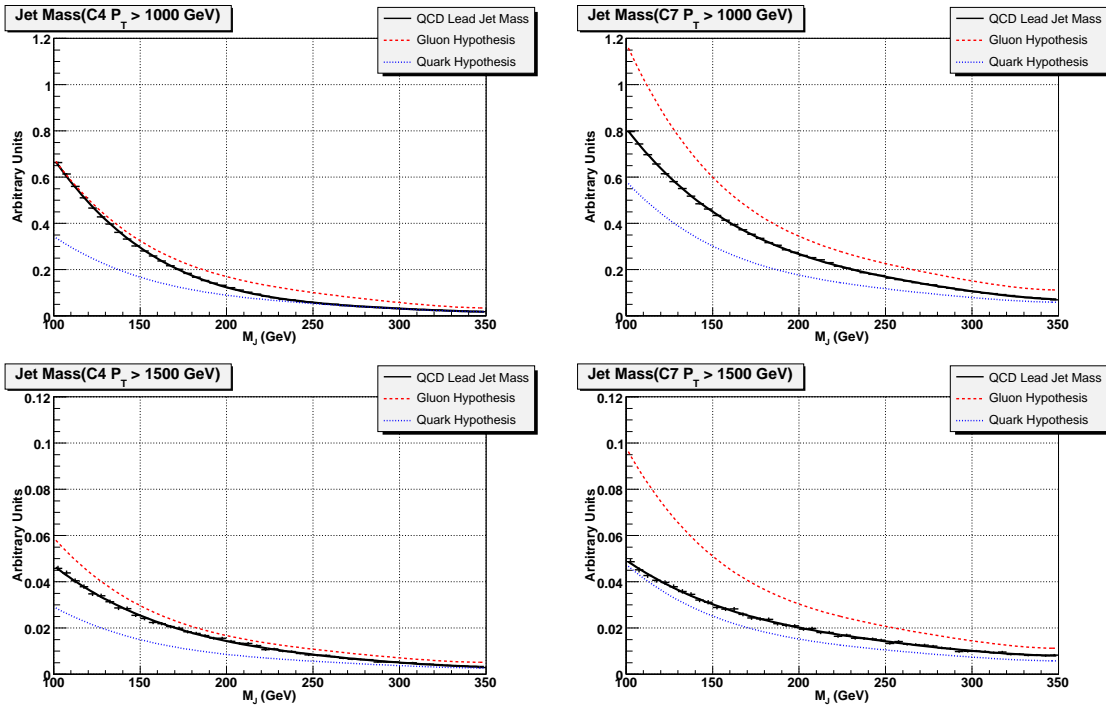

Figure 4: Comparison between the theoretical jet mass distributions and MC leading jet mass distribution from Sherpa. The minimum p_T and cone size are indicated on the plots. A gluon (quark) hypothesis is the prediction made if the entire contribution were from gluon (quark) jets (cf Eq. (3.10)).

events where the leading jet has $p_T \geq 1$ TeV obtained from Sherpa. The quark and gluon

curves, obtained from Eq. (3.10), with use of the jet functions in Eqs. (A.14) and (A.16), correspond to the cases where the lead jets are all quark or gluons jets, respectively.

As before, we find the bounds for the total cross section

$$\sigma(R)_{upper\ bound} = \int_{p_T^{min}}^{\infty} dp_T \sum_c \left(\frac{d\sigma^c(R)}{dp_T}\right)_{MC} \int_{140\,GeV}^{210\,GeV} J^g(m_J, p_T, R)\, dm_J\,, \quad (3.11)$$

$$\sigma(R)_{lower\ bound} = \int_{p_T^{min}}^{\infty} dp_T \sum_c \left(\frac{d\sigma^c(R)}{dp_T}\right)_{MC} \int_{140\,GeV}^{210\,GeV} J^q(m_J, p_T, R)\, dm_J\,. \quad (3.12)$$

In table 2, we refer to the gluon and quark jets from the results in Eqs. (3.11) and (3.12), respectively. The numbers in the table were calculated as follows. From a MC sample corresponding to $100\ \text{fb}^{-1}$ of data, we extracted the number of events with C4 lead jet $p_T \geq 1000(1500)\,\text{GeV}$ and $140\,\text{GeV} < m_J < 210\,\text{GeV}$, the top mass window. We repeated this exercise for C7 jets. The *data* column contains these results.

p_T^{lead} cut	Cone Size	Data	Quark hypothesis	Gluon hypothesis
1000 GeV	C4	113749	70701	135682
1000 GeV	C7	197981	131955	260045
1500 GeV	C4	10985	6513	12785
1500 GeV	C7	13993	11164	22469

Table 2: Comparison of Sherpa MC data to predictions of pure-quark and pure-gluon hypothesis, for the number of events with leading jet with mass between 140 GeV and 210 GeV. The data is compared to the bounds given in Eqs. (3.11) and (3.12). The statistics reflect $100\ \text{fb}^{-1}$ of integrated luminosity.

3.2.1 Fractional Fake Rate

With the theoretical machinery discussed in the previous section, we are able to make a prediction of the rate at which QCD jets will fake the mass signature of top-jets. We define the fractional fake rate as the fraction of jets with $140\,\text{GeV} \leq m_J \leq 210\,\text{GeV}$, for given p_T and R. We estimate the upper and lower bounds of the fractional fake rate as

$$\int_{140\,GeV}^{210\,GeV} dm_J\, J^q(m_J, p_T, R) \leq \text{Fractional fake rate} \leq \int_{140\,GeV}^{210\,GeV} dm_J\, J^g(m_J, p_T, R)\,.$$
$$(3.13)$$

In Fig. 5, we plot the fractional fake rate as a function of jet transverse momentum. To predict the number of fakes in our sample, we fold the differential cross section for QCD jet production (Fig. 6) with the fractional fake rate (Fig. 5). Again we expect a Sudakov-like suppression when $m_J \ll p_T R$, thus flatting the theoretical fractional fake rate as p_T increases. This can be seen more predominately for $R = 0.7$ in Fig. 5.

12
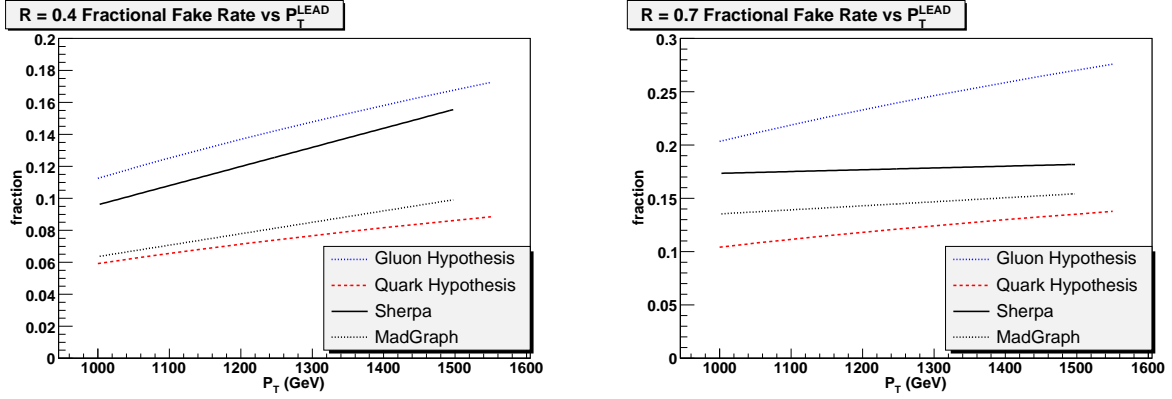
Figure 5: The fraction of jets which acquire $140\,\mathrm{GeV} \leq m_J \leq 210\,\mathrm{GeV}$ as a function of the transverse momentum of the leading jet. Quark- and gluon-hypothesis curves yield the prediction for the fractional fake rate, if all jets were either quark- or gluon-jets, respectively. The plot on the left corresponds to C4 jets; the plot on the right corresponds to C7 jets (cf Eq. (3.13)).
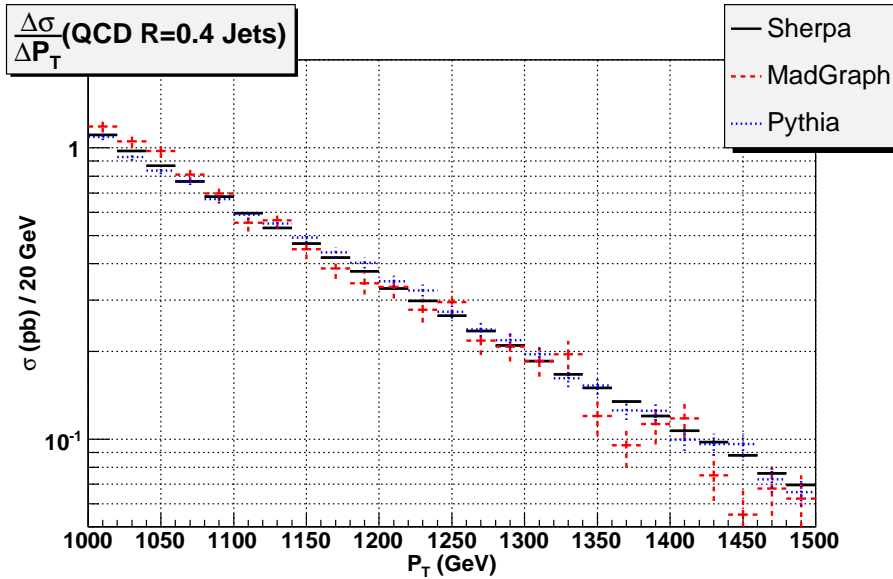
Figure 6: The differential cross section for QCD ($R = 0.4$) jet production with respect to the p_T of the leading jet. Sherpa, MG/ME and Pythia are represented.

3.2.2 Pseudorapidity Independence of the Jet Mass Distribution

In general, we expect that NP signals will have a pseudorapidity dependence. Therefore, the study of pseudorapidity dependence may provide a tool for NP searches (for an interesting discussion see [42]). In Fig. 7, we plot the jet mass distributions for central and outer jets. We observe consistency with the approximation that the distributions are to leading order, independent of pseudorapidity.
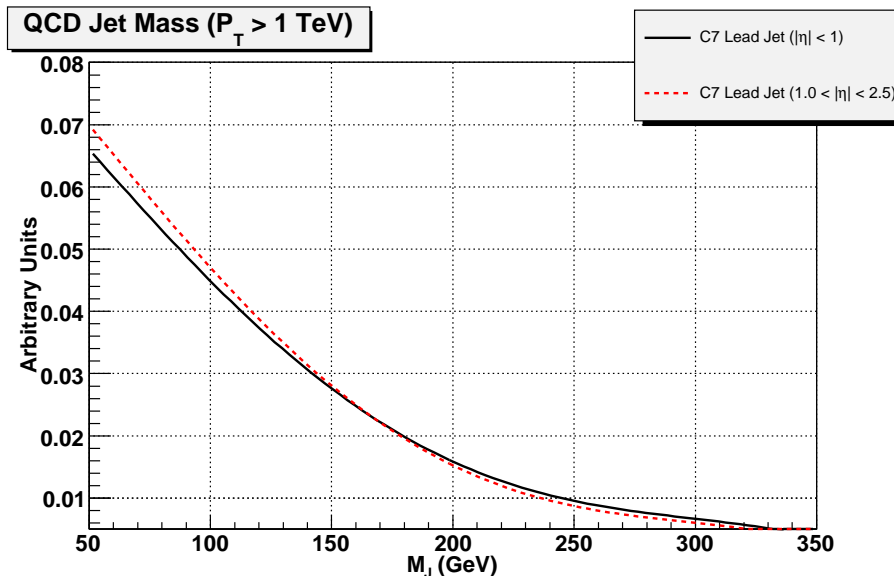

Figure 7: The $R = 0.7$ jet mass distribution for central jets ($|\eta| < 1$) and for jets with $1 \leq |\eta| \leq 2.5$. Jets have $p_T \geq 1 \, \text{TeV}$. This plot is produced with the Sherpa MC.

4 High p_T Hadronic Top Quarks

In this section, we discuss the collimation of the top quark decay products. In Fig. 8, we plot the rate of collimation as a function of the top p_T (for a related discussion and analysis see [4, 5, 8]). We define collimation rate as the fraction of top quarks which reconstruct to a jet having $140 \, \text{GeV} \leq m_J \leq 210 \, \text{GeV}$.

To examine the efficiency of the jet mass methods, it is instructive to look at mass distributions for the signal and background. We examine the distributions for events where the leading jet p_T exceeds $1000 \, \text{GeV}$ and $1500 \, \text{GeV}$ with C4 and C7 jets. In Fig. 1, we plot the jet mass distribution for the $t\bar{t}$ signal for $p_T{}^{lead} \geq 1000 \, \text{GeV}$. The efficiency of C7 jets for capturing the hadronic top is greater than that for C4 jets. For C4 jets, we still observe pronounced structure around the W-mass (M_W), which diminishes for C7 jets. We also note that the peak for the C7 jets moves closer to the top mass, indicating a higher

14

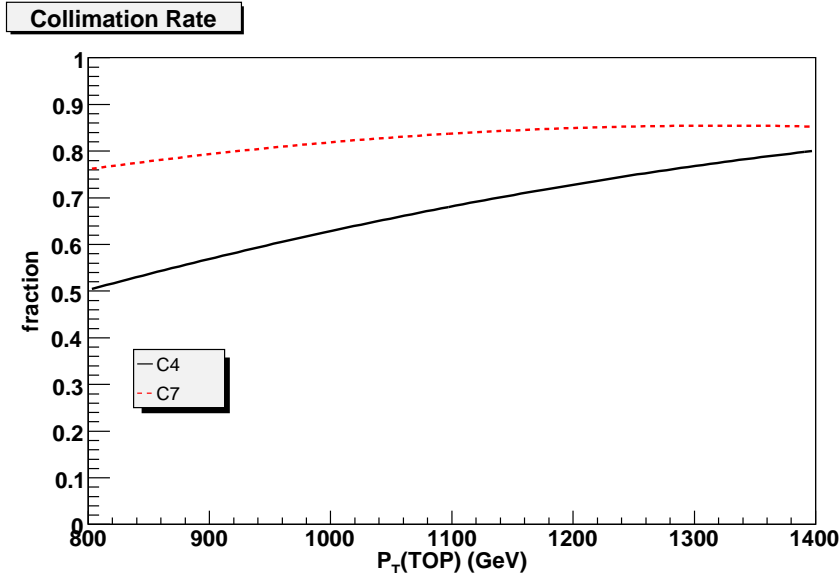Figure 8: The collimation rate for top quarks as a function of their transverse momentum, for C4 (black solid curve) and C7 (red dashed curve) jets. Collimation rate is defined as the fraction of top quarks with $140\,\mathrm{GeV} \leq m_J \leq 210\,\mathrm{GeV}$.

efficiency for capturing the hadronic top. We expect that detector effects will further smear the signal. Fig. 1 also shows the mass distributions including leading detector effects (using transfer functions).

We note that the analysis has an inherent tension with regard to choosing the cone size for the jet. The reconstruction cone should be sufficiently wide to capture all the daughter products of the hadronic top. On the other hand, we need to keep the cone appropriately small to keep out the QCD jet background and other soft contamination [5].

We describe the gross features of the top mass distribution, without providing a detailed analytic expression for the top jet. [†] At next to leading order, we expect the top-jet to be broken into two contributions, J^t_{QCD} and J^t_{EW} (at leading order it is just given by the top bare mass). The first contribution, J^t_{QCD}, is similar to that of the QCD jets. It is characterized by a very short time scale of $\mathcal{O}(10)\,\mathrm{GeV}$ and makes the top-jet mass harder. Using factorization, this process can be calculated by methods similar to the one discussed in the Appendix, fixing the mass of the final parton to m_t and assuming it is stable. For our purposes, the resulting broadening is subdominant for a top mass window of $\pm 35\,\mathrm{GeV}$. [‡] At leading order, the second contribution, J^t_{EW}, is expected to be kinematical in nature, due to the weak decay of the top quark. The time scale is longer (of order $\Gamma_t/\gamma_t = \mathcal{O}(0.2)\,\mathrm{GeV}$,

---

[†]For a more precise analysis in the case of e^+e^- collider and without providing a finite cone size see [11].

[‡]It is crucial to understand this behavior if one aims to improve the top mass measurement at the LHC. At the moment this has been studied only for lepton colliders [11].

15

where Γ_t is the top quark width and γ_t is the Lorentz boost). The main effect here is top mass softening, because the jet cone may not capture all the particles from the top quark decay chain. This kinematic effect depends solely on $m_t/(p_T R)$. It should reduce the mass of the top-jet and is expected to exhibit structure near the W mass. Since the top jet mass softening is a kinematic process, it should be well described by simple phase space generators. We schematically express the top jet mass as a sum of three contributions

$$m_J^t \sim m_t + \delta m_{QCD} + \delta m_{EW} \,, \tag{4.14}$$

where the jet mass function can be schematically written as a convolution of three different sources

$$J^t(m_J, m_t, R, p_T) \sim \int dm_{QCD} \, dm_{EW} \, dm_0 \, \delta(m_0 - m_t) \, \delta(m_J - m_0 - m_{QCD} - m_{EW}) \times$$
$$J_{QCD}^t(m_{QCD}, R, p_T) \times J_{EW}^t(m_{EW}, m_t/(p_T R)) \,. \tag{4.15}$$

The top mass is large, so we are not concerned about uncertainties in the lower jet mass spectrum. We conclude that existing MC tools should well describe this part of our studies.

5 $t\bar{t}$ Jets vs. QCD Jets at the LHC

In this section, we combine the results of the previous discussions, and apply them to analyze energetic SM $t\bar{t}$ events vis-a-vis QCD jet production at the LHC. The main purpose of this section is to understand how well we can discriminate our signal from the overwhelming QCD background. We illustrate an example analysis using the jet functions, and evaluate their performance on MC data. Unfortunately, it is very difficult to outline a one-size-fits-all analysis. Therefore, we perform a broad-strokes analysis that contains sufficient detail to provide general guidance. We do not attempt to invoke advanced, but analysis-specific, procedures that could provide further refinement. It is also important to bear in mind that the final evaluation of the jet functions, as precision analysis tools, can really only be done on real data. The primary reason is that we expect the jet functions to describe physics data. The MC distributions are, at this point, an approximation to what we believe will be LHC data. A precision analysis will show the strains between the jet function-based shape predictions and the effective distribution that MC uses to generate its mass distribution.

We examine two cases in detail, both at truth-level (no detector effects) and accounting for detector effects. The first case, *single tagging*, consists of "top-tagging" (requiring $140 \, \text{GeV} \leq m_J \leq 210 \, \text{GeV}$) the leading jets satisfying a p_T cut. The second case, *double tagging*, consists of top-tagging the leading and subleading jets, with a p_T cut only on the leading jet.

5.1 Peak Resolution

In this analysis, one objective is to resolve the excess of events where the mass of the leading jet lies in the top mass window ($140\,\mathrm{GeV} \leq m_J \leq 210\,\mathrm{GeV}$). It is important to note that we are not hunting for a peak; we already know its location. The issue is that of resolving its magnitude and estimate the probability that the background would fluctuate to yield the observed data. To estimate the significance of such a measurement, we perform a rudimentary analysis for resolving peaks. We emphasize that it is misleading to estimate the significance as $S/\sqrt{B}$, where the signal and background are separate MC samples. These numbers represent an unrealizable scenario, and tend to be optimistic. In real data, there is no way to separate the signal from background with certainty. Furthermore, at the present time, we cannot trust MC to provide the precise shape of QCD jet mass distributions. Therefore, we derive our approximations to the background *shape* directly from the "data", via sideband analysis (outside the top mass window). We use our previous knowledge of the shape of the background in the sideband region, to minimize the number of degree of freedoms involved in the sideband fit. We will discuss this further in the next section.

After approximating the shape of the background in the sideband region, we interpolate the shape of the background into the top mass window. The primary challenges are that our background is large and also has large uncertainties, which induce large uncertainties in the signal. We discuss this in more detail at the conclusion of this section. For the shape of the $t\bar{t}$ signal inside the top mass window, we use MC. In principle, the shape of the top mass distribution can be also derived semi-analytically, as discussed in section 4 (see also [11]). However, to leading order we expect the MC data to provide us with a reliable shape (it should capture the radiation at the leading log approximation, also the, phase space, population of the top decay products is purely kinematical). For simplicity we use the simulation data for this step in our analysis. These shapes, after normalization to unit area, are referred to as *probability density functions*. Unfortunately, the standard acronym for probability density functions conflicts with existing usage for parton distribution functions in this paper. To avoid confusion, we simply refer to them as *shapes*. We use the approximate shapes for the signal and background to perform an extended maximum likelihood fit to the sample, with jet mass distribution $F(m_J)$, thereby obtaining the background and signal normalizations. We define a jet mass distribution $F(m_J)$ as

$$F(m_J) = N_B \times b(m_J) + N_S \times s(m_J), \tag{5.16}$$

where N_B is the predicted background, and N_S is the predicted signal in the top mass window. $b(m_J)$ and $s(m_J)$ are used to denote the background and signal shapes, respectively. Both N_B and N_S are allowed to float independently.

5.1.1 Sideband Background Analysis

We perform a sideband analysis in order to avoid the $t\bar{t}$ signal-rich region. The basic goal is to understand the shape of the background by examining a region where there is no signal.

In the sidebands, in particular the low side, the signal contaminates the background. In Fig. 1, we see that the $t\bar{t}$ signal does not vanish outside the top mass window. Although it is small compared to the QCD background as can be seen in Fig. 10, this contamination substantially impacts resolution of the peak. We attempt to purify the background in this region, by rejecting energetic jets consistent with originating from a top quark decay, i.e. - signal, as follows. For a candidate event where the leading jet passes preselection criteria, all jets within a cone $R = 1$ are (vectorially) added into a single combined jet. We call this a group jet, although this definition differs slightly from that in J. Conway, *et al.*, in [5]. If the group mass, m_G, of the combined jet falls within the top mass window, the candidate event is rejected. This discriminant tends to reject events where the decay products of the top quarks are not fully collimated, i.e. reconstructed as a single jet. We must understand any biases introduced by this discriminant. Fig. 9 shows the effect of the m_G cuts on the background and signal. The background shape is left relatively intact, but the signal is substantially diminished.
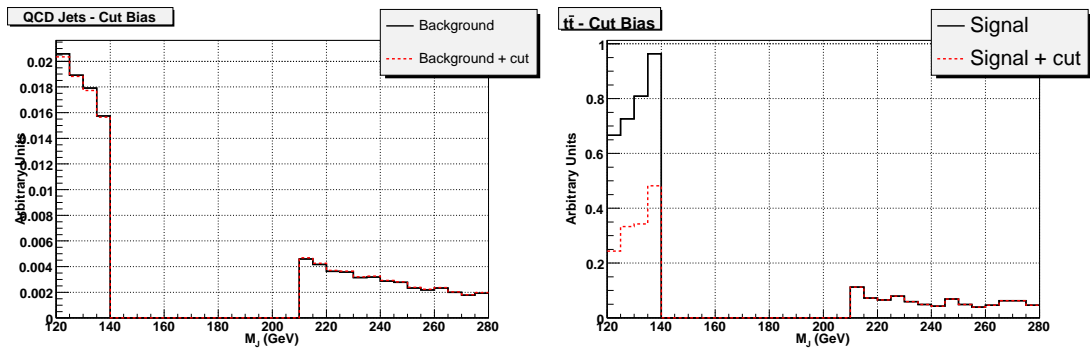
Figure 9: The jet mass distributions for the $t\bar{t}$ and QCD jet samples in the sidebands. The plot on the left depicts the shape of the QCD jet sample before and after making a combined jet mass cut on m_G, as described in Sec. 5.1. Both curves are normalized to unit area, to show the similarity of the shapes before and after the cut. The plot on the right depicts the effect of the combined jet mass cut on the $t\bar{t}$ signal. The red (dashed) curve shows the effect of the cut relative to the original jet mass distribution (black solid curve). Note: Unlike the left plot, these curves are not renormalized.

Advanced use of this m_G discriminant is outside the scope of this analysis, possibly leading to more sophisticated analyses (see *e.g.* J. Conway, *et al.*, in [5]). We simply use it to efficiently reject signal events in the sidebands, while keeping the majority of background events.

We analyze the shape of the background in the sidebands using the jet functions derived in section 3. We expect real QCD jets to be an admixture of quark and gluon jets. Therefore, our Ansatz posits the admixture of quarks and gluons as a fraction. We expect small corrections (deviations from a constant admixture) to arise from different sources. For example, we do not consider events with a leading jet of fixed p_T, but rather impose a

lower p_T cut. Our Ansatz for the jet mass distribution assumes the following form

$$b(m_J) \propto \beta\left(m_J\right) \times J^Q\left(m_J; p_T^{min}, R\right) + \left(1 - \beta\left(m_J\right)\right) \times J^G\left(m_J; p_T^{min}, R\right), \qquad (5.17)$$

where $\beta\left(m_J\right)$ is a linear polynomial $\left(\beta_0 + \beta_1 \frac{m_J}{p_T^{min} R}\right)$. Note that with $b(m_J)$ defined above, along with Eq. (5.16), the total number of degree of freedom involved in the sideband fit is four: β_0, β_1, N_B and N_S.

5.1.2 Significance

After resolving the magnitude of the signal $(t\bar{t})$ peak against that of the QCD jet background, via the methods outlined in the previous sections, we now discuss how to interpret those results. Our analysis is based on *log-likelihood ratio* method.[§] A background+signal hypothesis to describe a data sample is only meaningful if a background-only hypothesis is unlikely to describe that sample. We estimate the *statistical* significance, n_σ, of the peak as

$$n_\sigma = \sqrt{2\left(\log\mathcal{L} - \log\mathcal{L}_0\right)}, \qquad (5.18)$$

where $\mathcal{L}_0$ is the value of the maximized likelihood function obtained from fitting the data to the background shape alone (equivalent to setting N_S to zero in Eq. (5.16)), and $\mathcal{L}$ is the value of the maximized likelihood function obtained from fitting the data to the background shape and signal shape.[¶] The functional form of the likelihood function is given by

$$\mathcal{L} = \prod_{k=1}^{N_{\text{BINS}}} \frac{\exp\left(-F\left(m_k\right)\right) \times \left[F\left(m_k\right)\right]^{N_k}}{N_k!}, \qquad (5.19)$$

where we are fitting for the functional form of $F\left(m_J\right)$ as given by Eq. (5.16). Here, m_k and N_k refer to the value of the mass at the center, and the occupancy, of the k-th bin, respectively.

5.2 Single Top-Tagging

For each of the signal $(t\bar{t})$ and background (QCD jets) samples, we preselect events with a p_T cut on the leading jet. In Fig. 10 we plot the jet mass distribution including detector effects for the signal and background, including the theoretical upper and lower bound for the background. We show the number of events with jet mass in the range $140\,\mathrm{GeV} \leq m_J \leq 210\,\mathrm{GeV}$. For reference, the number of events for the signal and background, at the truth-level, are presented in table 3. It is clear that the background is roughly two orders

---

[§]An excellent discussion may be found in the The Review of Particle Physics [43].

[¶]Except in pathological cases, the significance is well approximated by $\frac{S}{\Delta S}$, where S is the fitted signal, and ΔS is the error on S.

of magnitude larger than the signal. Once we add detector effects the significance of the signal is further deteriorated. We conclude that a simple counting method would not be effective here.
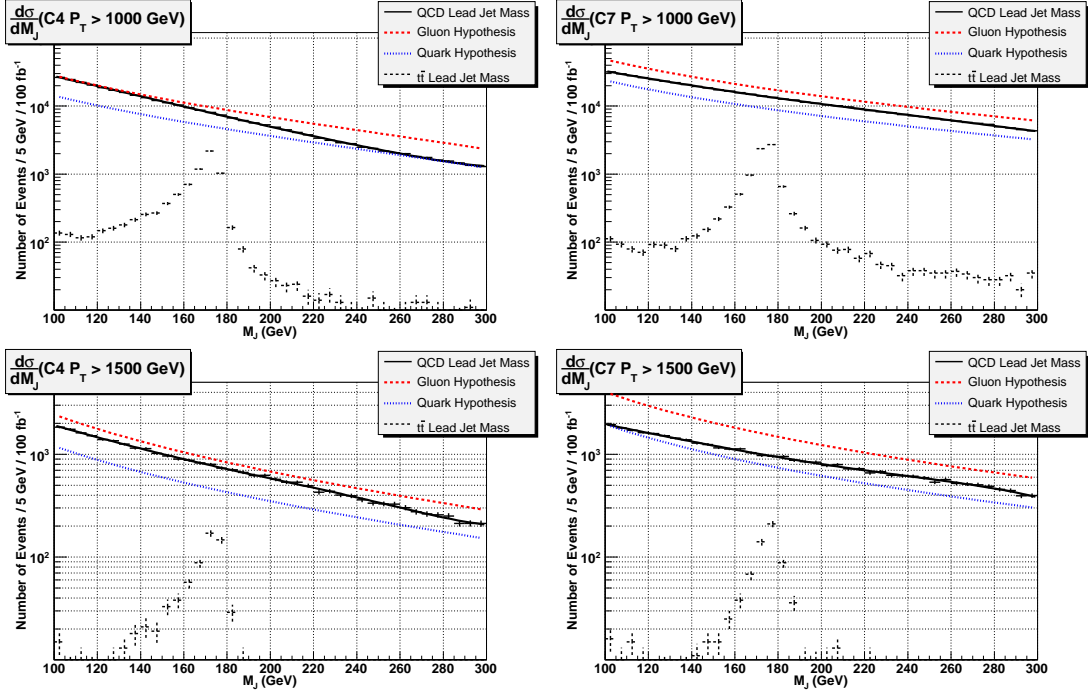
Figure 10: The jet mass distributions for the $t\bar{t}$ and QCD jet samples. The plots on the top row correspond to a $p_T{}^{lead} \geq 1000\,\mathrm{GeV}$. The plots on the bottom row correspond to a $p_T{}^{lead} \geq 1500\,\mathrm{GeV}$. The plots on the left correspond to $R = 0.4$; the plots on the right correspond to $R = 0.7$. The theoretical bounds, Eq. (3.10), are also plotted. These numbers are tabulated in table 3.

5.2.1   Detector Effects

Here, we repeat the truth-level procedure from above, accounting for the leading effects of detector resolution and $\pm 5\%$ jet energy scale. We also tabulate the relative change in acceptance of the signal and background, due to detector resolution and energy scale, which we define as

$$\Delta_{\mathrm{JES}} = \frac{N_{\mathrm{JES}} - N_{\mathrm{TRUTH}}}{N_{\mathrm{TRUTH}}}, \tag{5.20}$$

where N_{JES} is the number of events passing the selection criteria after detector smearing and JES effects have been applied. These results are tabulated in table 4, which shows how the signal and background are affected differently by smearing effects. We see that the

net effects of the detector smearing plus the uncertainties in the JES lead to substantial uncertainties $\mathcal{O}(10\% - 30\%)$ in the signal and background. As anticipated, this leads to a clear failure of simple counting type analyses and calls for a different approach, which will be introduced in the following in the form of sideband analyses and jet shapes.

p_T^{lead} cut	Cone Size	$t\bar{t}$ (S)	Background (B)	S/B
1000 GeV	C4	6860	113749	0.060
1000 GeV	C7	8725	197981	0.044
1500 GeV	C4	630	10985	0.057
1500 GeV	C7	689	13993	0.049

Table 3: Truth-level (no detector effects) results for single-tag jet mass method using, reflecting $100\,\mathrm{fb}^{-1}$ of integrated luminosity.

p_T^{lead} cut	Cone	S (0% JES)	Δ_0	+5% JES	Δ_5	-5% JES	Δ_{-5}
1000 GeV	C4	5778	-15.8%	6562	-4.3%	4798	-30.1%
1000 GeV	C7	7367	-15.6%	8543	-2.1%	6037	-30.8%
1500 GeV	C4	741	17.6%	934	48.3%	536	-14.9%
1500 GeV	C7	789	14.5%	1119	62.4%	601	-12.8%
p_T^{lead} cut	Cone	B (0% JES)	Δ_0	+5% JES	Δ_5	-5% JES	Δ_{-5}
1000 GeV	C4	107661	-5.4%	122291	7.5%	90232	-20.7%
1000 GeV	C7	192710	-2.7%	224666	13.5%	154733	-21.8%
1500 GeV	C4	13615	23.9%	18144	65.2%	10108	-8.0%
1500 GeV	C7	18712	33.7%	25361	81.2%	13407	-4.2%

Table 4: Acceptance of signal and background for the single tag method, relative to truth-level analysis, accounting for the leading effects of detector resolution and jet energy scale (JES). The $t\bar{t}$ signal is represented in the top half; the QCD jet background is represented in the bottom half. The statistics reflect $100\,\mathrm{fb}^{-1}$ of integrated luminosity. Δ_{JES} is the relative change in background for the indicated JES, relative to truth-level analysis in table 3 (cf Eq. (5.20)).

5.2.2 Results for single tagging

We now apply the analysis described in the previous sections to resolve the peak related to the top quark in the signal region, the top mass window. First we perform a sideband background analysis, to resolve the shape of the background. After applying the cuts described in Sec. 5.1, we fit the background to our Ansatz. Fig. 11 shows an example of such background fit to our Ansatz. The results of this fit described by Eq. (5.16) and below are shown in Fig. 12, which demonstrates how the detector affects the signal resolution.
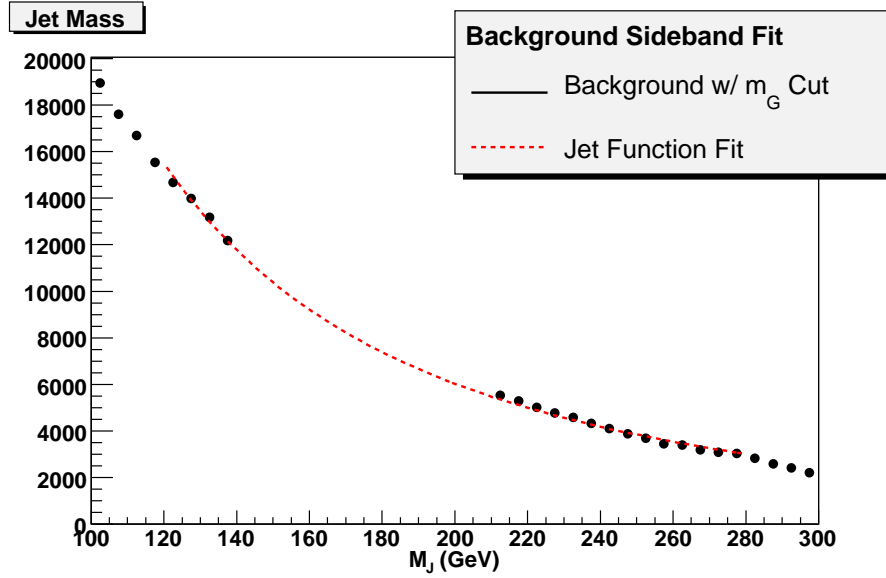
Figure 11: A typical example of fitting jet functions to the jet mass distribution in the sideband regions $(120 \, \text{GeV} \leq m_J \leq 140 \, \text{GeV}) \cup (210 \, \text{GeV} < m_J < 280 \, \text{GeV})$. This plot corresponds to a single-tag analysis with C7 jets with $p_T \geq 1000 \, \text{GeV}$.
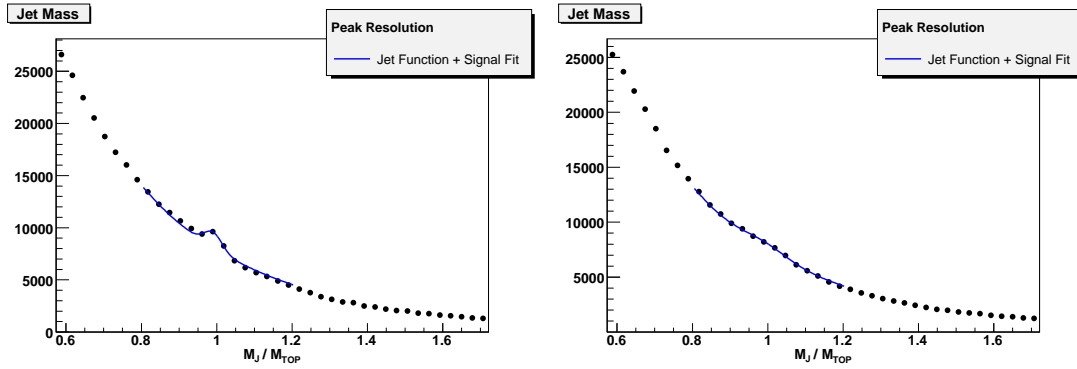
Figure 12: The results of fitting jet functions + signal shape to the jet mass distribution in the top mass window. The plot on the left corresponds to a truth-jet analysis. The plot on the right depicts the effects of detector smearing. The statistics reflect $100 \, \text{fb}^{-1}$ of integrated luminosity.

Our main results have been summarized in the tables below. The results of the fitting procedures for the different p_T cuts and cone sizes are shown in tables 5 and 6 for integrated luminosities of 100 fb^{-1} and 25 fb^{-1} respectively, and subsequently in tables 9 and 10 for the double top tagging case which is discussed in the following subsection. Our model for the background in these analyses was already introduced in subsection 5.1.2. Apart from the cone size and p_T^{min}, for each JES, we show the result of the fit regarding the number of background (B_{FIT}) and signal (S_{FIT}) events in the mass window and their ratio. ΔS is the error on S_{FIT} and p-value and χ^2/ndf are given to describe the quality of the fit in each case [43]. For our analysis, the total number of degree of freedom is 14 (18 bins $-$ 4 fit parameters: β_0, β_1, N_B and N_S).

Most importantly, we give the statistical significance, n_σ (defined in Eq. (5.18)), which is a measure of the probability that fluctuations of the *proposed* background yield in the observed data. The significance value is only as good as the p-value which indicates the goodness of fit. We point out that for entries in which the p-value is lower than, say 5%, the significance figure is probably not reliable. The fitting procedure on that data sample requires further examination, for residuals and bias analyses, for example, but this falls outside the scope of this work. We find two such instances of failed fits, both in table 5. This also suggests how we are to interpret the results of the tables. The relatively large background to signal ratio means that small errors in the background induce relatively large errors in the signal. Furthermore, we have not quantified correlations between the background and signal shapes. Similarities in the shapes can lead to small ambiguities, which are reflected in the fitting parameter errors. The combination of these two difficulties gives rise to an effect, which, strictly speaking, is a defect in the analysis. We remind the reader that we have a large uncertainty in the $t\bar{t}$ signal cross section (see table 1), which we have not accounted for in the analysis. We have singled out Sherpa MC data for use in our analyses, and the reader should bear this in mind when interpreting the results in tables 5, 6, 9, 10. Small errors in the background shape can yield good fits with high significance figures, and still have relatively large errors in the signal. We are led to interpret the results in the tables as the significance of the peak, relative to the indicated background shape hypothesis (the jet functions in our case). We find that our single tagging method allows us to resolve the $t\bar{t}$ signal from the QCD background with $p_T^{min} \sim 1$ TeV and 25 fb^{-1} of data. This jet mass analysis does not include any b-tagging or jet-shapes (to be discussed in the following section).

5.3 Double Top-Tagging

The above analyses related to single top-tagging are useful not only for $t\bar{t}$ production, but rather for general cases in which we expect to have at least one very energetic top jet. However, for the cases where there is more than one heavy high-p_T particle, we certainly have more information which can be used to distinguish signal from the QCD background. Clearly, $t\bar{t}$ events contain more information than what is encoded in a single top jet mass.

$p_T{}^{lead} \geq 1000 \, \text{GeV}$ Cone $R = 0.4$

JES	B_{FIT}	S_{FIT}	ΔS	n_σ	p-value	χ^2/ndf	$(S/B)_{\text{FIT}}$
0%	106571	6868	671	10.3	0.73	0.74	0.064
5%	120717	8137	715	11.4	0.01	2.01	0.067
-5%	89136	5895	615	9.6	0.95	0.46	0.066

$p_T{}^{lead} \geq 1000 \, \text{GeV}$ Cone $R = 0.7$

JES	B_{FIT}	S_{FIT}	ΔS	n_σ	p-value	χ^2/ndf	$(S/B)_{\text{FIT}}$
0%	189185	10892	800	13.7	0.09	1.52	0.058
5%	219189	14020	859	16.4	0.02	1.87	0.064
-5%	151556	9214	720	12.9	0.63	0.83	0.061

$p_T{}^{lead} \geq 1500 \, \text{GeV}$ Cone $R = 0.4$

JES	B_{FIT}	S_{FIT}	ΔS	n_σ	p-value	χ^2/ndf	$(S/B)_{\text{FIT}}$
0%	13562	794	224	3.6	1.00	0.26	0.059
5%	17803	1275	256	5.0	0.89	0.58	0.072
-5%	10155	489	193	2.5	0.94	0.49	0.048

$p_T{}^{lead} \geq 1500 \, \text{GeV}$ Cone $R = 0.7$

JES	B_{FIT}	S_{FIT}	ΔS	n_σ	p-value	χ^2/ndf	$(S/B)_{\text{FIT}}$
0%	18456	1045	252	4.2	0.75	0.72	0.057
5%	24921	1559	284	5.4	0.96	0.45	0.063
-5%	13315	693	213	3.3	1.00	0.20	0.052

Table 5: Estimate of upper limit on significance of peak resolution via single tag method, accounting for detector smearing. S_{FIT} and B_{FIT} are the results of an extended maximum likelihood fit. ΔS is the error on S_{FIT}. Significance n_σ is defined in Eq. (5.18). These results are derived with 100 fb^{-1} of integrated luminosity.

$p_T{}^{lead} \geq 1000\,\text{GeV}$ Cone $R = 0.4$

JES	B_{FIT}	S_{FIT}	ΔS	n_σ	p-value	χ^2/ndf	$(S/B)_{\text{FIT}}$
0%	26642	1712	335	5.1	1.00	0.19	0.064
5%	30206	1995	346	5.8	0.96	0.45	0.066
-5%	22371	1379	288	4.8	1.00	0.11	0.062

$p_T{}^{lead} \geq 1000\,\text{GeV}$ Cone $R = 0.7$

JES	B_{FIT}	S_{FIT}	ΔS	n_σ	p-value	χ^2/ndf	$(S/B)_{\text{FIT}}$
0%	47277	2730	399	6.8	0.98	0.38	0.058
5%	54870	3419	424	8.1	0.87	0.60	0.062
-5%	37910	2274	354	6.4	1.00	0.21	0.060

$p_T{}^{lead} \geq 1500\,\text{GeV}$ Cone $R = 0.4$

JES	B_{FIT}	S_{FIT}	ΔS	n_σ	p-value	χ^2/ndf	$(S/B)_{\text{FIT}}$
0%	3381	201	112	1.8	1.00	0.06	0.059
5%	4418	346	130	2.7	1.00	0.07	0.078
-5%	2519	136	96	1.4	1.00	0.09	0.054

$p_T{}^{lead} \geq 1500\,\text{GeV}$ Cone $R = 0.7$

JES	B_{FIT}	S_{FIT}	ΔS	n_σ	p-value	χ^2/ndf	$(S/B)_{\text{FIT}}$
0%	4609	259	125	2.1	1.00	0.18	0.056
5%	6231	382	144	2.6	1.00	0.12	0.061
-5%	3320	174	99	1.6	1.00	0.06	0.052

Table 6: Estimate of upper limit on significance of peak resolution via single tag method, accounting for detector smearing. S_{FIT} and B_{FIT} are the results of an extended maximum likelihood fit. ΔS is the error on S_{FIT}. Significance n_σ is defined in Eq. (5.18). These results are derived with 25 fb^{-1} of integrated luminosity.

We augment the single-tag analysis for the $t\bar{t}$ signal, by simply requiring that the subleading jet mass be in the top mass window, without imposing a p_T cut. This cut preferentially removes more background events than the signal events, without biasing the distributions. The sideband analysis, applied to the leading jet, remains the same as for the single top-tagging case. As we shall see even this simple treatment yields a sizable improvement in the significance. Roughly half of the events have smaller p_T than the minimum p_T for the leading jet as shown in Fig. 13. Although, by definition, a subleading jet has smaller p_T than the leading one, its p_T distribution is peaked at the $p_T{}^{min}$, and only small portion of events are in the smaller p_T tail region. The number of events for the signal and background, at the truth-level, are presented in table 7. To get an idea on how the subleading mass cut affects our signal and background samples, one can compare the numbers given in table 3 with the ones in 7. For example, we see that at truth level for $R = 0.4$ and $p_T{}^{min} = 1\,\mathrm{TeV}$ the size of the signal sample is decreased by 50% while the background sample by roughly 12%. This is consistent with the results shown in Figs. 8 and 5 in which the analysis is done for a fixed p_T.

In principle, one could apply a sideband analysis to the subleading jet. However, due to the fact that the p_T is allowed to float, the required analysis would necessarily be more complicated. The double-tagging method increases the signal-to-background ratio, and the significance of the measurements increases. The leading effects of detector resolution and jet energy scale on the signal and background acceptance can be seen in Tables 9 and 10. We find that our double tagging method yields a reach of up to $p_T{}^{min} \sim 1.5\,\mathrm{TeV}$ with $100\,\mathrm{fb}^{-1}$, without relying on b-tagging or jet-shapes (to be discussed in the following section).
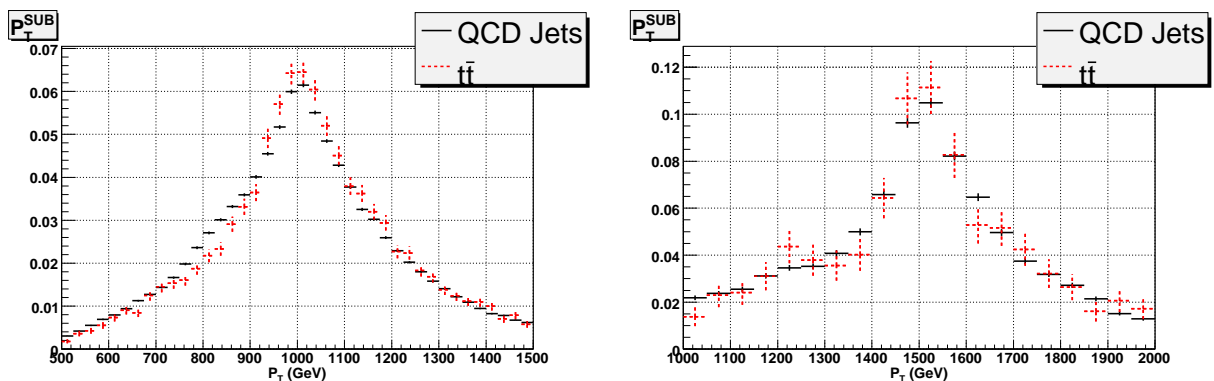

Figure 13: We compare the p_T distribution of the subleading jet for the $t\bar{t}$ signal with (the red dotted curve) and without (the black solid curve) leading detector effects. The plot on the left, right corresponds to C4 jets with $\left(p_T{}^{lead} \geq 1000, 1500\,\mathrm{GeV}\right)$ respectively.

p_T^{lead} cut	Cone Size	$t\bar{t}$ (S)	Background (B)	S/B
1000 GeV	C4	3430	13505	0.254
1000 GeV	C7	6302	36765	0.171
1500 GeV	C4	403	1874	0.215
1500 GeV	C7	458	2724	0.168

Table 7: Truth-level (no detector effects) results for double-tag jet mass method using, reflecting $100\,\mathrm{fb}^{-1}$ of integrated luminosity.

p_T^{lead} cut	Cone	S (0% JES)	Δ_0	+5% JES	Δ_5	-5% JES	Δ_{-5}
1000 GeV	C4	2601	-24.2%	2868	-16.4%	2228	-35.0%
1000 GeV	C7	4563	-27.6%	5351	-15.1%	3765	-40.3%
1500 GeV	C4	403	0.0%	489	21.3%	292	-27.5%
1500 GeV	C7	487	6.3%	688	50.2%	352	-23.1%
p_T^{lead} cut	Cone	B (0% JES)	Δ_0	+5% JES	Δ_5	-5% JES	Δ_{-5}
1000 GeV	C4	13680	1.3%	15187	12.5%	12054	-10.7%
1000 GeV	C7	39361	7.1%	45596	24.0%	32192	-12.4%
1500 GeV	C4	2373	26.6%	3109	65.9%	1746	-6.8%
1500 GeV	C7	4195	54.0%	5651	107.5%	3014	10.6%

Table 8: Acceptance of signal and background for the double tag method, relative to truth-level analysis, accounting for the leading effects of detector resolution and jet energy scale (JES). The $t\bar{t}$ signal is represented in the top half; the QCD jet background is represented in the bottom half. The statistics reflect $100\,\mathrm{fb}^{-1}$ of integrated luminosity. Δ_{JES} is the relative change in background for the indicated JES, relative to truth-level analysis in table 7 (cf Eq. (5.20)).

$p_T{}^{lead} \geq 1000\,\mathrm{GeV}$ Cone $R = 0.4$

JES	B_{FIT}	S_{FIT}	ΔS	n_σ	p-value	χ^2/ndf	$(S/B)_{\mathrm{FIT}}$
0%	13488	2789	237	11.8	0.99	0.33	0.207
5%	14653	3395	255	13.3	0.94	0.50	0.232
-5%	11762	2516	212	11.9	0.99	0.31	0.214

$p_T{}^{lead} \geq 1000\,\mathrm{GeV}$ Cone $R = 0.7$

JES	B_{FIT}	S_{FIT}	ΔS	n_σ	p-value	χ^2/ndf	$(S/B)_{\mathrm{FIT}}$
0%	38101	5813	358	16.2	0.72	0.76	0.153
5%	43993	6943	386	18.0	0.66	0.81	0.158
-5%	31290	4655	320	14.6	0.57	0.89	0.149

$p_T{}^{lead} \geq 1500\,\mathrm{GeV}$ Cone $R = 0.4$

JES	B_{FIT}	S_{FIT}	ΔS	n_σ	p-value	χ^2/ndf	$(S/B)_{\mathrm{FIT}}$
0%	2341	430	94	4.6	0.99	0.35	0.184
5%	2968	624	110	5.7	0.96	0.45	0.210
-5%	1593	436	79	5.5	0.82	0.66	0.274

$p_T{}^{lead} \geq 1500\,\mathrm{GeV}$ Cone $R = 0.7$

JES	B_{FIT}	S_{FIT}	ΔS	n_σ	p-value	χ^2/ndf	$(S/B)_{\mathrm{FIT}}$
0%	4053	625	129	5.2	1.00	0.28	0.154
5%	5532	801	128	6.3	0.93	0.50	0.145
-5%	2965	399	100	4.0	1.00	0.14	0.135

Table 9: Estimate of upper limit on significance of peak resolution via double tag method, accounting for detector smearing, and jet energy scale (JES). S_{FIT} and B_{FIT} are the results of an extended maximum likelihood fit. ΔS is the error on S_{FIT}. Significance n_σ is defined in Eq. (5.18). These results are derived with $100\,\mathrm{fb}^{-1}$ of integrated luminosity.

$p_T^{lead} \geq 1000\,\text{GeV}$ Cone $R = 0.4$

JES	B_{FIT}	S_{FIT}	ΔS	n_σ	p-value	χ^2/ndf	$(S/B)_{\text{FIT}}$
0%	3367	696	119	5.9	1.00	0.08	0.207
5%	3658	848	128	6.7	1.00	0.12	0.232
-5%	2931	631	106	6.0	1.00	0.07	0.215

$p_T^{lead} \geq 1000\,\text{GeV}$ Cone $R = 0.7$

JES	B_{FIT}	S_{FIT}	ΔS	n_σ	p-value	χ^2/ndf	$(S/B)_{\text{FIT}}$
0%	9521	1452	181	8.1	1.00	0.19	0.152
5%	10997	1732	193	9.0	1.00	0.20	0.158
-5%	7817	1162	160	7.3	1.00	0.22	0.149

$p_T^{lead} \geq 1500\,\text{GeV}$ Cone $R = 0.4$

JES	B_{FIT}	S_{FIT}	ΔS	n_σ	p-value	χ^2/ndf	$(S/B)_{\text{FIT}}$
0%	577	111	47	2.4	1.00	0.08	0.192
5%	737	155	55	2.8	1.00	0.11	0.210
-5%	393	109	40	2.8	1.00	0.16	0.277

$p_T^{lead} \geq 1500\,\text{GeV}$ Cone $R = 0.7$

JES	B_{FIT}	S_{FIT}	ΔS	n_σ	p-value	χ^2/ndf	$(S/B)_{\text{FIT}}$
0%	1005	159	70	2.7	1.00	0.06	0.158
5%	1376	200	64	3.1	1.00	0.12	0.145
-5%	739	96	50	1.9	1.00	0.04	0.130

Table 10: Estimate of upper limit on significance of peak resolution via double tag method, accounting for detector smearing, and jet energy scale (JES). S_{FIT} and B_{FIT} are the results of an extended maximum likelihood fit. ΔS is the error on S_{FIT}. Significance n_σ is defined in Eq. (5.18). These results are derived with 25 fb^{-1} of integrated luminosity.
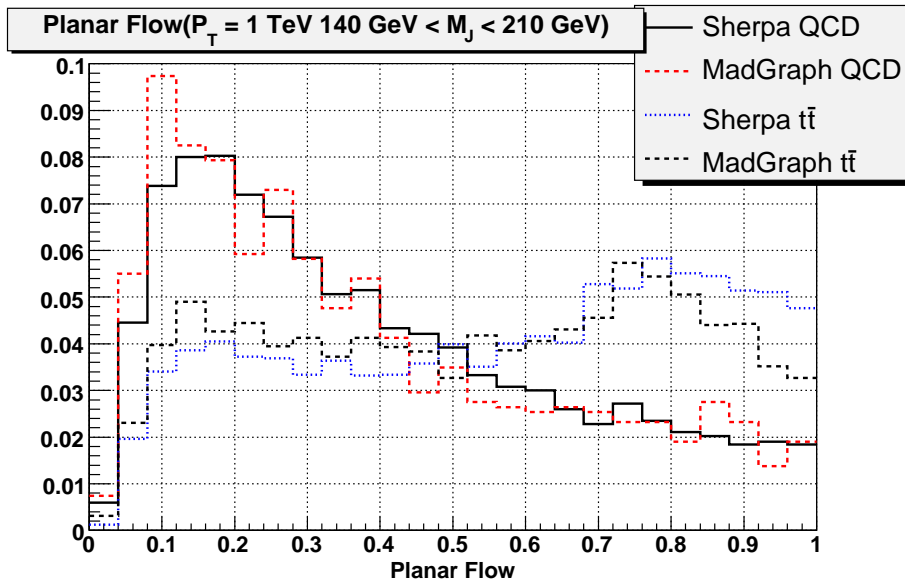
Figure 14: The planar flow distribution is plotted for $t\bar{t}$ and QCD jets with mass in the top mass window, $140\,\text{GeV} \leq m_J \leq 210\,\text{GeV}$. Sherpa and MG/ME distributions are represented.

6 Jet Substructure

We discussed simple single- and double-mass tagging methods, and we found that we may need additional handles in order to resolve SM $t\bar{t}$ signals for smaller integrated luminosities or higher p_T. We discuss briefly the possibility of using substructure to further analyze energetic jets in the top mass window. We defer the details to our recent work in [10] (see also [27]), where we derive simple analytic expressions to approximate the jet shape variable distributions. For developing additional tools to reslove $t\bar{t}$ signals, there are approaches which exploit information outside of hadronic calorimeter [44] such as tracker or electromagnetic calorimeter. But we limit ourselves to the information encoded only within the hadronic calorimeter to develop significance for resolving $t\bar{t}$ signals. We do not also discuss the possibility of b-tagging for high p_T top-jet [7], which is still under speculation for the range of p_T relevant for our analysis.

Jet shapes are the extensions of well-known event shapes, used at lepton colliders, applied to the analysis of energy flow inside single jets. The fact that we consider only jets with high mass is crucial since it allows us to control the shape of various distributions related to energy flow in a perturbative manner. As an example, we examine the *planar flow* variable, which measures the extent to which the energy flow inside the jet is linear or planar. Planar flow (Pf) is defined as follows. We first define an (unnormalized) event
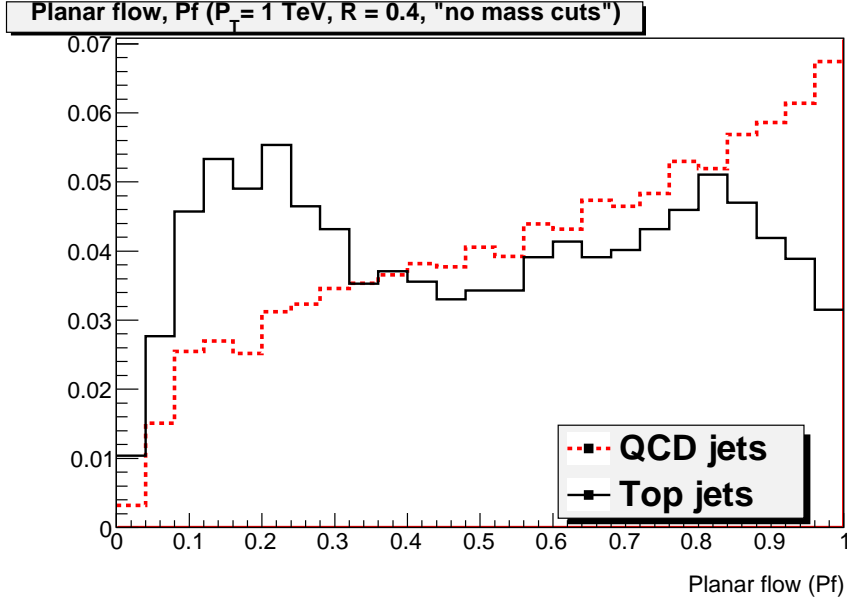
Figure 15: The planar flow distribution is plotted for $t\bar{t}$ and QCD jets without fixing jet mass. MG/ME distributions are represented.

shape tensor I_w as$^{\parallel}$

$$I_w^{kl} = \sum_i w_i \frac{p_{i,k}}{w_i} \frac{p_{i,l}}{w_i} \, , \tag{6.21}$$

where w_i is the energy of particle i in the jet, and $p_{i,k}$ is the kth component of its transverse momentum relative to the thrust axis, which typically coincides with the jet axis. Given I_w, we define Pf as

$$Pf = \frac{4 \det(I_w)}{\mathrm{tr}(I_w)^2} = \frac{4\lambda_1\lambda_2}{(\lambda_1 + \lambda_2)^2} \, , \tag{6.22}$$

where $\lambda_{1,2}$ are the eigenvalues of I_w. Pf approaches zero for linear shapes and approaches unity for isotropic depositions of energy. In Fig. 14, we plot the planar flow distributions for QCD jets and $t\bar{t}$. As can be seen by comparing Figs. 14 and 15, it is crucial to consider only events in the top mass window. Without a jet mass cut, the jet shape analysis loses its rejection power.

31
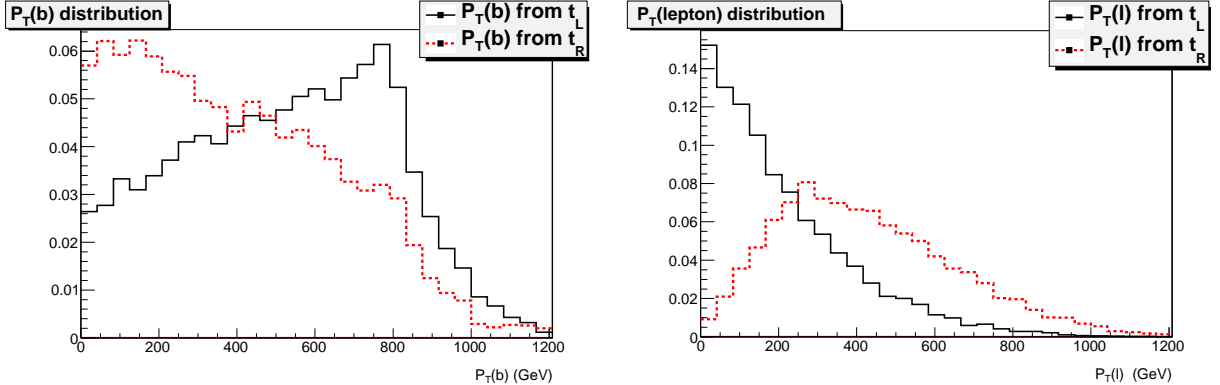
Figure 16: In the plot on the left, we show a comparison of the p_T distribution of b quark from t_L (the solid curve) vs. t_R (the dashed curve). In the plot on the right we show the p_T distribution of the charged lepton from t_L (the solid curve) vs. t_R (the dashed curve). We have imposed a lower cut, $p_T^{min} = 1000$ GeV.

7 Top Quark Polarization Measurement

In this section, we consider how to exploit b-quarks to measure the polarization of highly boosted hadronic tops. Various new physics models have particle spectra which couple preferentially to one chirality, giving rise to parity violation. Since chirality equals helicity for ultra-relativistic fermions, highly boosted top quarks can help us probe parity violation in the bottom/top quark sector. As is well-known, the top quark decays before the hadronization process occurs, and measurement of the top quark polarization from its leptonic daughters has been studied [2, 3]. We propose using the transverse momentum of the b-quark, inferred from the b-tagged jet, to perform similar measurements. The p_T distribution for the b-quark depends on the chirality of the top-quark. The b-tagged jets should, therefore, also act as good spin analyzers. In Fig. 16, we compare the p_T distributions for leptons and b-quarks, for both left- and right-handed top quarks.

As mentioned earlier, the issue of b-tagging at high p_T is quite challenging at this time (for recent studies see [7]), and a fully quantitative study is not yet available. The main idea is to examine the p_T distribution of b-tagged jets, in events where we believe these jets originate from $t \to bW$. In order to measure the p_T of the b quark, we need to require at least one of the top-jets should be resolved into more than two jets, since we cannot measure the p_T of the b quark inside a single top-jet. As shown in Fig. 8, even for high p_T ($p_T \geq 1000$ GeV) top jet, with cone size $R = 0.4$, $\sim 30\%$ of top-jets can be resolved into more than two jets. By fixing the cone size for jet reconstruction, it is important to understand any biases towards right-handed or left-handed top quarks. Bottom quarks from t_L have a harder p_T distribution than those from t_R, while the opposite is true for leptons from leptonic top quark decays. If one uses a small reconstruction cone, the efficiencies for

---

$\parallel$The overall normalization is not important to this discussion.

jet mass reconstruction between the t_L and t_R may differ. We found a negligible bias using cone jets with $R = 0.4$.

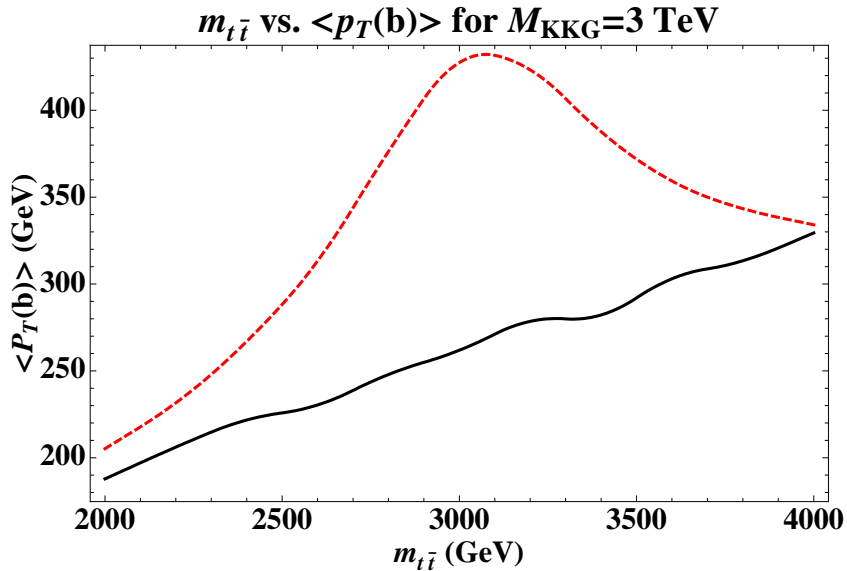**$m_{t\bar{t}}$ vs. $<p_T(b)>$ for M_{KKG}=3 TeV**

Figure 17: We compare the $\langle P_T \rangle$ distribution of the b quark, as predicted by the Standard Model (black solid curve) and by RS1 model with SM fields propagating in the bulk (red dashed curve).

We can develop this discussion further with an example, namely the Randall Sundrum (RS) [45] model with the SM fields propagating in the bulk. We consider the case where the first Kaluza-Klein (KK) excitation of the gluon has a mass $M_{KKG} = 3$ TeV. We perform this analysis at partonic level. In the model we are considering for using b-quark $\langle p_T \rangle$, KK excitations of the gluon couple to left-handed top quarks $\sim 5\times$ stronger than to right-handed top quarks. Typical cross sections for KK gluon production are relatively small. The (background) SM top quarks are produced dominantly via parity-invariant QCD processes, and tend to wash out the signal. In order to resolve the signature, we are compelled to correlate deviations in the b-quark $\langle p_T \rangle$ spectrum to an excess in KK gluon production. In Fig. 17, we compare the mean value of the b-quark p_T spectrum, for the Standard Model and RS1 scenarios with SM fields propagating in the bulk. When correlated to the invariant mass of the KK gluon, we see a substantial deviation in the distribution of the b-quark $\langle p_T \rangle$. In Fig. 18, we compare the mean value of the lepton p_T spectrum, for the Standard Model and RS1 scenarios with SM fields propagating in the bulk, where KK excitations of the gluon couple to right-handed top quarks $\sim 5\times$ stronger than to left-handed top quarks.
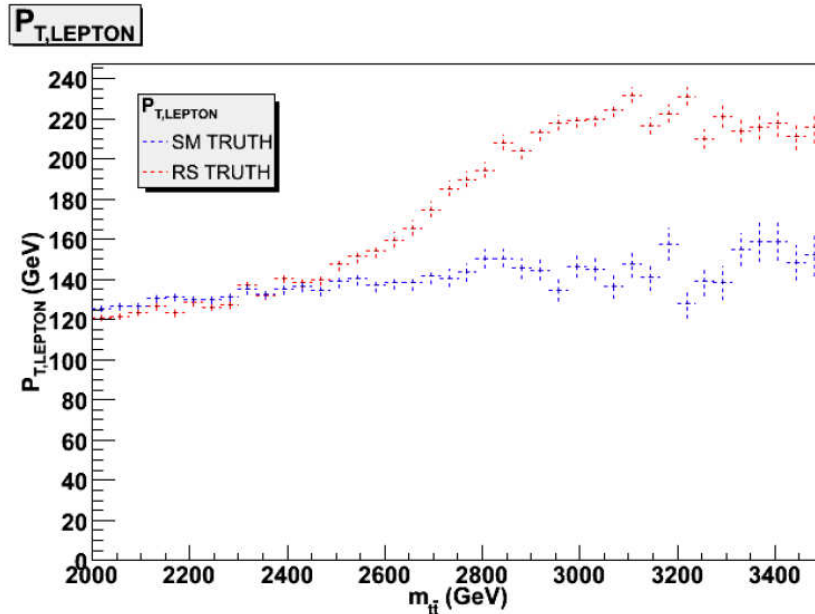
Figure 18: We compare the $\langle p_T \rangle$ distributions of the lepton, as predicted by the Standard Model (blue curve) and by RS1 model with SM fields propagating in the bulk (red curve).

8   Conclusions

In this study we have mainly focused on high p_T, hadronically decaying, tops in cases where they are fully collimated. Above p_T of 1 TeV the majority of the top daughters particles will be found inside a single cone jet even when the cone size is as small as $R = 0.4$. Therefore, they are simply denoted as top-jets. The leading background for top-jets comes from high p_T QCD jets. We provided analytic expressions for the QCD jet functions which approximate the background and show consistency with MC data. As an example we consider the case of SM $t\bar{t}$ production, and demonstrate how these jet functions, via side band analysis, allow us to efficiently resolve 1 TeV top-jet from the QCD background with $25\,\mathrm{fb}^{-1}$, and $\sim 1.5\,\mathrm{TeV}$ top-jets with $100\,\mathrm{fb}^{-1}$.

A wide class of new physics models posits $t\bar{t}$ production mechanisms which would significantly contribute to the mass distributions, possibly allowing resolution of excess production with less data. To further improve the significance we consider jet shapes (recently analyzed in [10] and also in [27]), which resolve substructure of energy flow inside cone jets. Augmentation of the analysis, such as the use of jet substructure in combination with a jet mass cut and b-tagging, may improve the signal resolution, allowing us to discover NP signal through top quark channel even with lower luminosity or higher p_T cut. We provided such an example using the *planar flow* jet shape variable, and a detailed analysis is presented in our recent work [10]. We also proposed using the transverse momentum of

34

the bottom quarks to measure top quark polarization as a probe of parity violation.

In this paper we mostly focused on fairly extreme (but not uncommon at the LHC) kinematical configurations where the tops are fully collimated. This has several advantages such as having direct contact with theoretical based calculation of the jet functions and also the ability to consider arbitrarily high top momenta (at least in principle). However, it is clear that some fraction of the hadronic tops will be reconstructed in 2-jet (intermediate) or $\geq$ 3-jet (conventional) topologies. The fraction of events related to the different topologies is a function of the cone size and p_T. Solid reconstruction algorithms and analyses must be flexible enough to interpolate between these different regions. We note that our approach is complimentary to others that have been proposed recently [5, 6, 26, 27]. In most cases, the difference is related to the fact that the tops considered are not fully collimated and a two-jet topology is exploited to increase the signal to background ratio. It would be very interesting and important to derive theoretically based techniques to control the corresponding distribution of the background relevant to the intermediate region. It is likely that there are overlaps between the different regions. Such issues are important to examine in detail. Mastering these complimentary methods may help to make potential new physics observations more robust, if verified via multiple and independent techniques.

Finally, we want to emphasize that the analysis proposed herein is also applicable to other processes involving, highly boosted, heavy particles such as electroweak gauge bosons, the Higgs and other new physics particles, to which QCD is a leading background as well.

Acknowledgments

Appendix

A Jets at Fixed Invariant Mass

Here we give details of the definitions and calculations for the jet functions that we employ in section 3. Single inclusive Jet cross sections have been studied intensively [38, 46, 47, 48]. Here, we are interested in computing the QCD background to jets of measured mass. The main background to the production of $t\bar{t}$ pairs is from dijet production from hadronic collisions,

$$H_a(p_a) + H_b(p_b) \rightarrow J_1(m_{J_1}^2, p_{1,T}, \eta_1, R) + J_2(m_{J_2}^2, p_{2,T}, \eta_2, R) + X, \tag{A.1}$$

where the final states are jets in the directions of the outgoing partons, each with a fixed jet mass m_J^2, a "cone size" $R^2 = \Delta\eta^2 + \Delta\phi^2$, and transverse momenta, $p_{i,T}$. For simplicity we choose the cone sizes equal for the two jets, although they can be different. For $R < 1$, we can isolate the leading (R^0) dependence of such cross-sections in factorized "jet" functions,

$$
\begin{aligned}
\frac{d\sigma_{H_A H_B \rightarrow J_1 J_2}}{dp_T dm_{J_1}^2 dm_{J_2}^2 d\eta_1 d\eta_2} &= \sum_{abcd} \int dx_a\, dx_b\, \phi_a(x_a)\, \phi_b(x_b) H_{ab\rightarrow cd}\left(x_a, x_b, p_T, \eta_1, \eta_2, \alpha_S(p_T)\right) \\
&\times J_1^c(m_{J_1}^2, p_T \cosh\eta_1, R, \alpha_S(p_T))\, J_2^d(m_{J_2}^2, p_T \cosh\eta_2, R, \alpha_S(p_T)),
\end{aligned}
\tag{A.2}
$$

with corrections that vanish as powers of R. Here the ϕ's are parton distribution functions for the initial hadrons, $H_{ab\rightarrow cd}$ is a perturbative $2 \rightarrow 2$ QCD hard-scattering function, equal to the dijet Born cross section at lowest order, and the J_i are jet functions, which are defined below. Jet function J_i summarizes the formation of a set of final state particles with fixed invariant mass and momenta collinear to the ith outgoing parton. Corrections to the cross section of order R^0 can only occur through collinear enhancements which factorize into these functions [49].

Following Ref. [40] we define jet function for quarks at fixed jet mass by

$$
\begin{aligned}
J_i^q(m_J^2, p_{0,J_i}, R) &= \frac{(2\pi)^3}{2\sqrt{2}\,(p_{0,J_i})^2} \frac{\xi_\mu}{N_c} \sum_{N_{J_i}} \text{Tr}\left\{\gamma^\mu \langle 0|q(0)\Phi_\xi^{(\bar{q})\dagger}(\infty, 0)|N_{J_i}\rangle\langle N_{J_i}|\Phi_\xi^{(\bar{q})}(\infty, 0)\bar{q}(0)|0\rangle\right\} \\
&\times \delta\left(m_J^2 - \tilde{m}_J^2(N_{J_i}, R)\right) \delta^{(2)}(\hat{n} - \tilde{n}(N_{J_i}))\delta(p_{0,J_i} - \omega(N_{J_c})),
\end{aligned}
\tag{A.3}
$$

where $\tilde{m}_J^2(N_{J_i}, R)$ is the invariant mass of all particles within the cone centered on direction $\hat{n}$ in state N_{J_i}. Correspondingly, gluon jet functions are defined by

$$
\begin{aligned}
J_i^g(m_J^2, p_{0,J_i}, R) &= \frac{(2\pi)^3}{2(p_{0,J_i})^3} \sum_{N_{J_i}} \langle 0|\xi_\sigma F^{\sigma\nu}(0)\Phi_\xi^{(g)\dagger}(0, \infty)|N_{J_i}\rangle\langle N_{J_i}|\Phi_\xi^{(g)}(0, \infty) F_\nu^{\,\rho}(0)\xi_\rho|0\rangle \\
&\times \delta\left(m_J^2 - \tilde{m}_J^2(N_{J_i}, R)\right) \delta^{(2)}(\hat{n} - \tilde{n}(N_{J_i}))\delta(p_{0,J_i} - \omega(N_{J_c})).
\end{aligned}
\tag{A.4}
$$

These functions absorb collinear enhancements to the outgoing particles that emerge from the underlying hard perturbative process and fragment into the observed jets. The Φ's are path ordered exponentials (Wilson lines) defined by

$$\Phi_\xi^{(f)}(\infty, 0; 0) = \mathcal{P} \left\{ e^{-ig \int_0^\infty d\eta\, \xi \cdot A^{(f)}(\eta\, \xi^\mu)} \right\}, \tag{A.5}$$

where $\mathcal{P}$ indicates ordering along the integral and where ξ is a direction with at least one component in the direction opposite to the jet. The full hadronic cross-section is independent of the choice for ξ. As indicated, the gauge field $A^{(f)}$ is a matrix in the representation of the generators for parton f. In general the jet function depends on $\vec{\xi} \cdot \hat{n}$, but for simplicity we suppress this dependence below. Finally the jet functions in Eqs. (A.3) and (A.4) are normalized such that at lowest order we have

$$J_i^{(0)}(m_{J_i}^2, p_{0,J_i}, R) = \delta(m_{J_i}^2). \tag{A.6}$$

A.1 Jet Functions at Next-to-Leading Order

At next-to-leading order, contributions to the jet mass distributions for light quark or gluon jets have only two particles in their final states. For the quark jet we have the following matrix element which has to be calculated to $\mathcal{O}(g^2)$,

$$\begin{aligned}
J_i^q(m_J^2, p_{0,J_i}, R) &= \frac{(2\pi)^3}{2(p_{0,J_i})^2} \frac{\xi_\mu}{N_c\sqrt{2}} \sum_{\sigma,\lambda} \int \frac{d^3p}{(2\pi)^3 2\omega_p} \frac{d^3k}{(2\pi)^3 2\omega_k} \mathrm{Tr} \left\{ \gamma^\mu \langle 0|q(0)\Phi_\xi^{(\bar{q})\dagger}(\infty,0)|p,\sigma;k,\lambda\rangle \right. \\
&\quad \left. \times \langle p,\sigma;k,\lambda|\Phi_\xi^{(\bar{q})}(\infty,0)\bar{q}(0)|0\rangle \right\} \delta\left(m_J^2 - (p+k)^2\right) \delta^{(2)}(\hat{n} - \hat{n}_{\vec{p}+\vec{k}}) \delta(p_{0,J_i} - p^0 - k^0),
\end{aligned} \tag{A.7}$$

where σ and λ denote the polarizations, and p and k the momenta of the final-state quark and gluon respectively with $\hat{n}_{\vec{p}+\vec{k}} \equiv \frac{\vec{p}+\vec{k}}{|\vec{p}+\vec{k}|}$. Similarly, for the gluon jet we have

$$\begin{aligned}
J_i^g(m_J^2, p_{0,J_i}, R) &= \frac{(2\pi)^3}{4(p_{0,J_i})^3} \sum_{N_{J_i}} \int \frac{d^3p}{(2\pi)^3 2\omega_p} \frac{d^3k}{(2\pi)^3 2\omega_k} \langle 0|\xi_\sigma F^{\sigma\nu}(0)\Phi_\xi^{(g)\dagger}(0,\infty)|p,\sigma;k,\lambda\rangle \\
&\quad \times \langle p,\sigma;k,\lambda|\Phi_\xi^{(g)}(0,\infty) F_\nu{}^\rho(0)\xi_\rho|0\rangle \delta\left(m_J^2 - (p+k)^2\right) \delta^{(2)}(\hat{n} - \hat{n}_{\vec{p}+\vec{k}}) \delta(p_{0,J_i} - p^0 - k^0),
\end{aligned} \tag{A.8}$$

where p and k are the final state momenta within the cone size, R. To evaluate these matrix elements, we need the rules for vertices shown in Fig. 19 for the field strengths. The double lines represent the perturbative expansion of the Wilson lines (A.5) in the ξ-direction (see Eq. (A.11)), whose vertices and propagators are shown in Fig. 20. The resulting diagrammatic contributions to the quark and gluon jet functions at next-to-leading order are shown in Fig. 21 and Fig. 22 respectively.
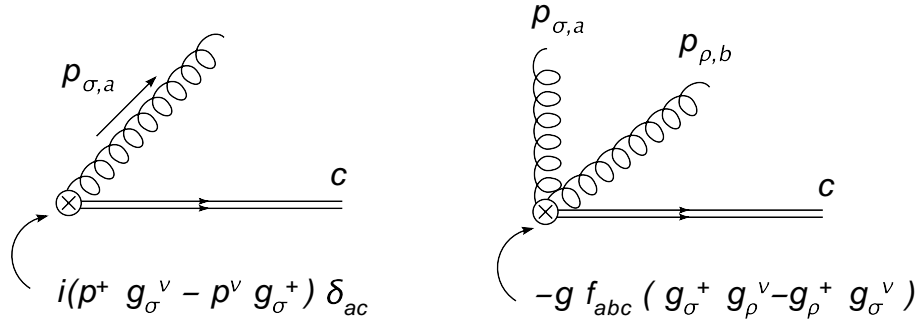
37

$p_{\sigma,a}$

c

$i(p^+\, g_\sigma{}^\nu - p^\nu\, g_\sigma{}^+)\, \delta_{ac}$

$p_{\sigma,a}$

$p_{\rho,b}$

c

$-g\, f_{abc}\, (\, g_\sigma{}^+\, g_\rho{}^\nu - g_\rho{}^+\, g_\sigma{}^\nu\,)$

Figure 19: Feynman rules associated with the $F^{+\nu}$ operator at the end of a Wilson line.

j i $-i\, g\, t_{a,ij}\, \xi_\mu$

$k_{\mu,\, a}$

$i\, (\, \xi \cdot k + i\, \epsilon\,)^{-1}$
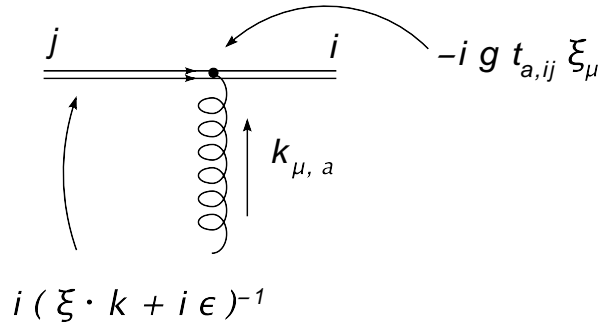
Figure 20: Feynman rules associated with eikonal lines, from the expansion of the Wilson lines.

(a) (b) (c)
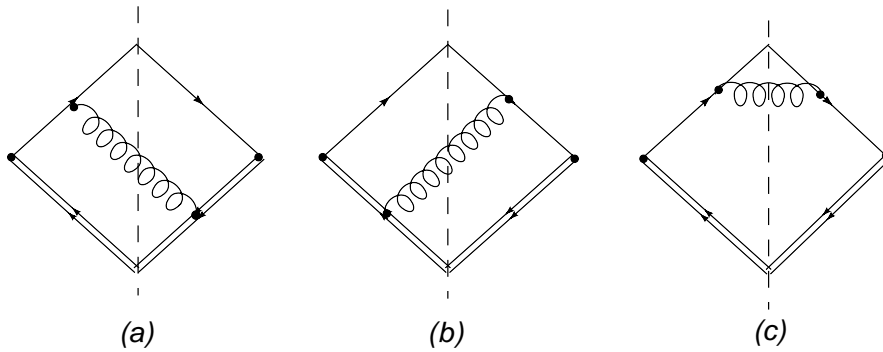
Figure 21: Real contributions to the quark jet function at order $\alpha_{\rm S}$.
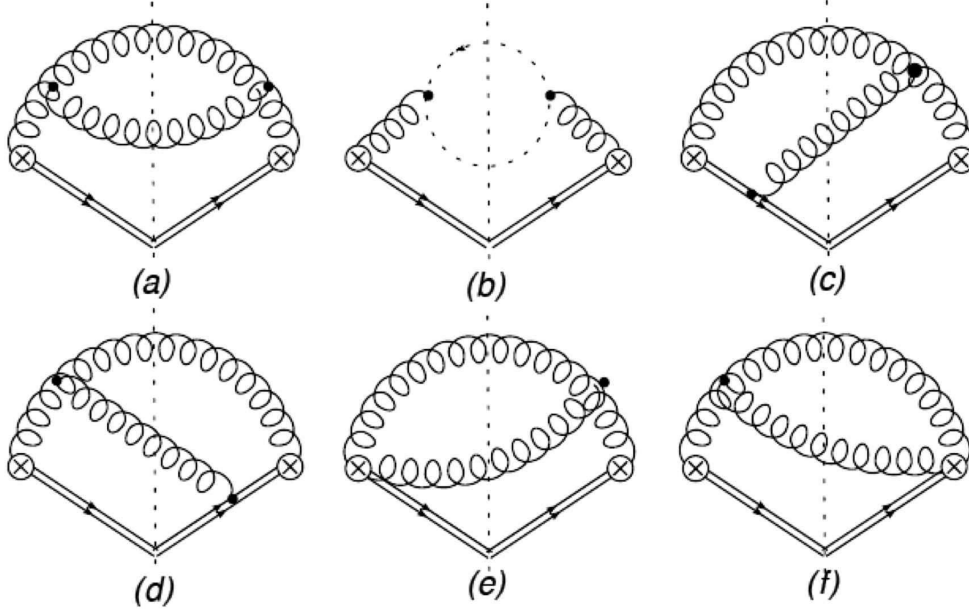
Figure 22: Real non-vanishing contributions to the gluon jet function in Feynman gauge at NLO.

We choose a frame where the jet is in the $\eta_J = \phi_J = 0$ direction and the vector ξ is light-like and in a direction opposite to the jet,

$$p_{J_i} = p_{0,J_i}(1, \beta_i, 0, 0) \quad \xi = \frac{1}{\sqrt{2}}(1, -1, 0, 0), \tag{A.9}$$

where $\beta_i = \sqrt{1 - m_{J_i}^2/p_{0,J_i}^2}$ is the velocity of the jet. In this frame we parametrize the momenta p and k above by

$$p = p^0(1, \cos\theta_p, \sin\theta_p, 0) \quad k = k^0(1, \cos\theta_k, -\sin\theta_k, 0), \tag{A.10}$$

where $\theta_{p,k}$ represents the angle of each particle to the jet axis $\hat{n}$. The path ordered exponentials are expanded order-by-order in g_s, related to the rules in Fig. 20 by the expansion,

$$\begin{aligned}
\Phi_\xi(\infty, 0; 0) &= \mathcal{P}\left\{ e^{-ig\int_0^\infty d\eta\,\xi\cdot A(\eta\,\xi^\mu)} \right\} \\
&= 1 - ig\int \frac{d^4k}{(2\pi)^4} \frac{i}{\xi\cdot k + i\epsilon}\xi\cdot A(k) + \dots .
\end{aligned} \tag{A.11}$$

We begin with the calculation of the quark jet function, which readily reduces to an integral over the energy and angle of one of the particles,

$$\begin{aligned}
J_i^{q(1)}(m_J^2, p_{0,J_i}, R) &= \frac{\beta_i}{8\sqrt{2}} \int \frac{d\cos\theta_k dk_0 k_0}{\pi(p_{0,J} - k_0)}|\mathcal{M}(p, k)|^2 \\
&\quad \times \delta(m_j^2 - 2k^0 p_{0,J}(1 - \beta_i\cos\theta_k))\Theta(R - \theta_k),
\end{aligned} \tag{A.12}$$

where we choose k to represent the gluon and p the quark. For k the softer momentum, we easily see that $\theta_k \geq \theta_p$. Therefore, $p_0 = k_0$ fixes the minimum angle for the softest particle, and we find $\cos(\theta_{S,min}) = \beta_i$. The region $\omega_p < \omega_k$ is found by simply interchanging p and k in $|\mathcal{M}(p,k)|^2$ so that

$$J_i^{q(1)}(m_J^2, p_{0,J_i}, R) = \frac{\beta_i}{16\sqrt{2}} \int_{\cos(R)}^{\beta_i} \frac{d\cos\theta_S}{(2\pi)^2} \frac{m_{J_i}^2/p_{0,J}^2}{\left(2(1-\beta_i\cos\theta_S) - \frac{m_{J_i}^2}{p_{0,J}^2}\right)} \frac{1}{p_{0,J}(1-\beta_i\cos\theta_S)}$$
$$\times \left(|\mathcal{M}_{q_i}(p,k)|^2 + |\mathcal{M}_{q_i}(k,p)|^2\right) . \tag{A.13}$$

The evaluation of $|\mathcal{M}_{q_i}(p,k)|^2$ is straightforward from the diagrams of Fig. 21, and we find

$$J_i^{q(1)}(m_J^2, p_{0,J_i}, R) = \frac{C_F\beta_i}{4m_{J_i}^2} \int_{\cos(R)}^{\beta_i} \frac{d\cos\theta_S}{\pi} \frac{\alpha_S(k_0) z^4}{(2(1-\beta_i\cos\theta_S) - z^2)(1-\beta_i\cos\theta_S)} \times$$
$$\left\{ z^2 \frac{(1+\cos\theta_S)^2}{(1-\beta_i\cos\theta_S)} \frac{1}{(2(1+\beta_i)(1-\beta_i\cos\theta_S) - z^2(1+\cos\theta_S))} + \right.$$
$$\left. \frac{3(1+\beta_i)}{z^2} + \frac{1}{z^4} \frac{(2(1+\beta_i)(1-\beta_i\cos\theta_S) - z^2(1+\cos\theta_S))^2}{(1+\cos\theta_S)(1-\beta_i\cos\theta_S)} \right\},$$
$$\tag{A.14}$$

where $z = \frac{m_{J_i}}{p_{0,J_i}}$, $p_{0,J_i} = \sqrt{m_{J_i}^2 + p_T^2}$, and $k_0 = \frac{p_{0,J_i}}{2} \frac{z^2}{1-\beta_i\cos\theta_S}$.

The calculation of the gluon jet function proceeds along the same lines, with the exception that both particles in the final states are now identical, and the presence of the field strengths, which appear at the end of each Wilson line. The rules for these vertices, as mentioned before, are shown in Fig. 19. Once again, we can write the gluon jet function as an integral over the angle of the softer particle,

$$J_i^{g(1)}(m_J^2, p_{0,J_i}, R) = \frac{\beta_i}{16m_{J_i}^2} \int_{\cos(R)}^{\beta_i} \frac{d\cos\theta_S}{(2\pi)^2 p_{0,J_i}^2} \frac{z^2}{(2(1-\beta_i\cos\theta_S) - z^2)(1-\beta_i\cos\theta_S)} |\mathcal{M}_{g_i}(p,k)|^2,$$
$$\tag{A.15}$$

where $|\mathcal{M}_{g_i}(p,k)|^2$ is symmetric under the interchange of p and k. We find from the diagrams shown in Fig. 22, the result

$$J_i^{g(1)}(m_J^2, p_{0,J_i}, R) = \frac{C_A\beta_i}{16m_{J_i}^2} \int_{\cos(R)}^{\beta_i} \frac{d\cos\theta_S}{\pi} \frac{\alpha_S(k_0)}{(1-\beta\cos\theta_S)^2(1-\cos^2\theta_S)(2(1+\beta) - z^2)}$$
$$\times \left(z^4(1+\cos\theta_S)^2 + z^2(1-\cos^2\theta_S)(2(1+\beta_i) - z^2) + (1-\cos\theta_S)^2(2(1+\beta_i) - z^2)^2 \right)^2.$$
$$\tag{A.16}$$

These one-loop expressions have been used to generate the comparisons to event generator output given in Section 3.

B   R-dependence

It is of interest to isolate the leading logarithmic contributions in both gluon and quark jets, which can be found from eikonal graphs in the adjoint and fundamental representations respectively,

$$J^{(eik),c}(m^2_{J_1}, p_T, R) = \frac{2\,C_c}{\sqrt{2}p_T} g^2 \int \frac{d^3k}{(2\pi)^3 2\omega_k} \frac{\xi \cdot p_J}{\xi \cdot k} \frac{\xi \cdot p_J}{2p_J \cdot k}$$
$$\times \delta\left(m^2_{J_1} - (p_1 + k)^2\right) \Theta\left(p_T - k_T\right) . \tag{A.17}$$

Parametrizing k as
$$k = k_T \left(\cosh \eta_k, \cos \phi_k, \sin \phi_k, \sinh \eta_k\right) , \tag{A.18}$$

this leads to

$$J^{(eik),c}(m^2_{J_1}, p_T, R) = g^2 \frac{C_c}{(2\pi)^3} \int dk_T\, k_T\, d\phi_k d\eta_k \frac{1}{k_T^2 \left(\cosh^2 \eta_k - \cos^2 \phi_k\right)}$$
$$\times \delta\left(2p_T\, k_T\left(\cosh \eta_k - \cos \phi_k\right) - m^2_{J_1}\right)\Theta\left(p_T - k_T\right).$$
$$= g^2 \frac{C_c}{(2\pi)^3} \int d\phi_k d\eta_k \frac{1}{m^2_{J_1}} \frac{1}{k_T^2\left(\cosh^2 \eta_k - \cos^2 \phi_k\right)}$$
$$\times \delta\left(2p_T\, k_T\, r - m^2_{J_1}\right)\Theta\left(\cosh \eta_k - \cos \phi_k - \frac{m^2_{J_1}}{p_T^2}\right). \tag{A.19}$$

In this expression we can change the variables to

$$\eta_k = r \cos \theta, \quad \phi_k = r \sin \theta . \tag{A.20}$$

Since we are dealing with highly collimated jets we can expand the integrand in r and integrate over θ, finding

$$J^{(eik),c}(m^2_{J_1}, p_T, R) \simeq g^2 \frac{2C_c}{(2\pi)^2} \int_{m_J/p_T}^{R} dr \frac{1}{m^2_{J_1}} \left\{\frac{1}{r} + \mathcal{O}(r^3)\right\}$$
$$\simeq \alpha_S(p_T) \frac{C_c}{m^2_{J_1}\pi} \left\{\log\left(\frac{R^2\, p_T^2}{m_J^2}\right) + \mathcal{O}(R^4)\right\}, \tag{A.21}$$

which shows explicitly the logarithmic behavior in R. Leading logarithmic contributions can be exponentiated, giving us a qualitative description of lower jet masses,

$$J^{(eik),c}(m^2_{J_1}, p_T, R) \simeq \frac{\alpha_S}{\pi} C_c \frac{1}{m_J^2} \log\left(\frac{R^2 p_T^2}{m_J^2}\right) \exp\left\{-\frac{\alpha_S}{2\pi} C_c \log^2\left(\frac{R^2 p_T^2}{m_J^2}\right)\right\} . \tag{A.22}$$

Without the above approximations, the eikonal jet function is given by

$$J^{(eik),c}(m_J, p_T, R) = \alpha_S(p_T) \frac{4\,C_c}{\pi m_J} \log\left(\frac{p_T}{m_J} \tan\left(\frac{R}{2}\right) \sqrt{4 - \left(\frac{m_J}{p_T}\right)^2}\right) . \tag{A.23}$$
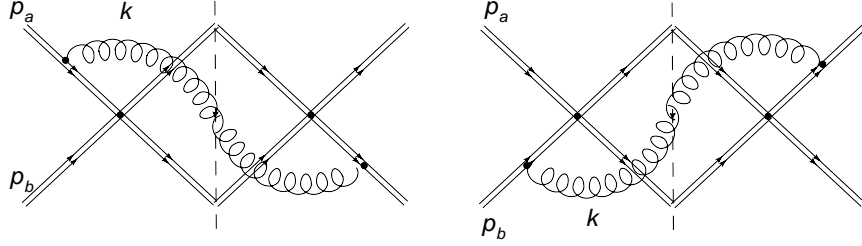
41

Figure 23: Contributions to the jet mass from the soft function.

As we have observed above, all R^0 behavior in the cross section can be found from the jet functions. We can also estimate the contribution of soft initial-state radiation on the cone-jet masses. Here we verify that such radiation is sub-leading in powers of R^2. Contributions due to wide angle gluons come from a "soft function" [49], which is defined in terms of an eikonal cross section,

$$S(m_{J_i}^2) \sim \sum_{N_s} \sigma^{(eik)}(N_s)\delta(m_{J_1}^2 - \tilde{m}_J^2(N_s, R)). \tag{A.24}$$

Diagrams that can contribute to the jet mass are illustrated in Fig. 23. The initial state radiation shown behaves as

$$S \sim \int d^4k\delta(k^2)\frac{p_a \cdot p_b}{(p_a \cdot k)(p_b \cdot k)}\delta(m_{J_1}^2 - 2p_1 \cdot k)\Theta(R - \tilde{R}(\eta_k, \phi_k)), \tag{A.25}$$

with p_a and p_b the momenta of incoming partons, neither of which is in the direction of the observed jets. Choosing a frame where the initial momenta are given by

$$p_a = \frac{\sqrt{s}}{2}(1, 0, 0, 1), \quad p_b = \frac{\sqrt{s}}{2}(1, 0, 0, -1), \tag{A.26}$$

and parametrizing the radiated gluon's momentum k as in Eq. (A.18) above, we find

$$\begin{aligned} S &\sim \int dk_T d\phi_k d\eta_k \frac{1}{k_T} \frac{1}{2p_T(\cosh\eta_k - \cos\phi_k)}\delta\left(k_T - \frac{m_{J_1}^2}{2p_T(\cosh\eta_k - \cos\phi_k)}\right) \\ &\sim \frac{2\pi}{m_{J_1}^2}\int_0^R dr\, r = \frac{\pi R^2}{m_{J_1}^2}, \end{aligned} \tag{A.27}$$

which is, as expected, power-suppresed in R compared to the logarithmic dependence we get from the jet function.

References

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

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

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

[4] U. Baur and L. H. Orr, Phys. Rev. D **76**, 094012 (2007) [arXiv:0707.2066 [hep-ph]].

[5] G. Brooijmans, ATLAS note, ATL-PHYS-CONF-2008-008; J. Conway, *et al.*, LPC Workshop on Early Physics at CMS, UC Davis (2007); G. Brooijmans *et al.*, arXiv:0802.3715 [hep-ph]; M. Vos, talk given in the *ATLAS Flavour Tagging Meeting*, CERN (2008).

[6] 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]; J. M. Butterworth, A. R. Davison, M. Rubin and G. P. Salam, arXiv:0809.2530 [hep-ph].

[7] See, for example, L. March, E. Ros and B. Salvachúa, ATL-PHYS-PUB-2006-002, ATL-COM-PHYS-2005-032; L. March, E. Ros and S.G.d.l. Hoz, ATL-COM-PHYS-2006-031, ATL-PHYS-CONF-2006-007; L. March, E. Ros, B. Salvacha, ATL-Phys-PUB-2006-002; M. Vos, ATL-PHYS-PUB-2008-000, ATL-PHYS-CONF-2008-016;

[8] U. Baur and L. H. Orr, arXiv:0803.1160 [hep-ph].

[9] Y. Bai and Z. Han, arXiv:0809.4487 [hep-ph].

[10] L. G. Almeida, S. J. Lee, G. Perez, G. Sterman, I. Sung and J. Virzi, arXiv:0807.0234 [hep-ph].

[11] 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].

[12] W. Skiba and D. Tucker-Smith, Phys. Rev. D **75**, 115010 (2007) [arXiv:hep-ph/0701247].

[13] B. Holdom, JHEP **0708**, 069 (2007) [arXiv:0705.1736 [hep-ph]].

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

[15] F. Maltoni and T. Stelzer, JHEP **0302**, 027 (2003) [arXiv:hep-ph/0208156].

[16] T. Stelzer and W. F. Long, Comput. Phys. Commun. **81**, 357 (1994) [arXiv:hep-ph/9401258].

[17] J. Alwall *et al.*, JHEP **0709**, 028 (2007) [arXiv:0706.2334 [hep-ph]].

[18] T. Gleisberg et al., JHEP **0402** (2004) 056 [arXiv:hep-ph/0311263].

[19] J. Virzi, "ATLAS Transfer Function,
https://twiki.cern.ch/twiki/bin/view/Atlas/TransferFunction."

[20] R. Frederix and F. Maltoni, arXiv:0712.2355 [hep-ph].

[21] V. Barger, T. Han and D. G. E. Walker, Phys. Rev. Lett. **100**, 031801 (2008)
[arXiv:hep-ph/0612016].

[22] T. Han, R. Mahbubani, D. G. E. Walker and L. T. E. Wang, arXiv:0803.3820 [hep-ph].

[23] K. Agashe, H. Davoudiasl, G. Perez and A. Soni, Phys. Rev. D **76**, 036006 (2007)
[arXiv:hep-ph/0701186]; B. Lillie, L. Randall and L. T. Wang, JHEP **0709**, 074 (2007)
[arXiv:hep-ph/0701166]; B. Lillie, J. Shu and T. M. P. Tait, Phys. Rev. D **76**, 115016
(2007) [arXiv:0706.3960 [hep-ph]]; H. Davoudiasl, G. Perez and A. Soni, Phys. Lett.
B **665**, 67 (2008) [arXiv:0802.0203 [hep-ph]].

[24] For a recent review see: T. Han, arXiv:0804.3178 [hep-ph].

[25] S. D. Ellis, Z. Kunszt and D. E. Soper, Phys. Rev. Lett. **69**, 3615 (1992) [arXiv:hep-ph/9208249].

[26] D. E. Kaplan, K. Rehermann, M. D. Schwartz and B. Tweedie, arXiv:0806.0848 [hep-ph].

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

[28] F. Hubaut, E. Monnier, P. Pralavorio, K. Smolek and V. Simak, Eur. Phys. J. C **44S2**,
13 (2005) [arXiv:hep-ex/0508061].

[29] W. Bernreuther, J. P. Ma and T. Schroder, Phys. Lett. B **297**, 318 (1992); S. Choi,
FERMILAB-THESIS-1999-07; K. A. Johns [D0 Collaboration], Int. J. Mod. Phys. A
16S1A (2001) 366; W. Bernreuther, A. Brandenburg, Z. G. Si and P. Uwer, Phys.
Lett. B **509**, 53 (2001) [arXiv:hep-ph/0104096]; B. Abbott *et al.* [D0 Collaboration],
Phys. Rev. Lett. **85**, 256 (2000) [arXiv:hep-ex/0002058]; T. Torma, JHEP **0111**, 055
(2001) [arXiv:hep-ph/9912281]; V. D. Barger, J. Ohnemus and R. J. N. Phillips, Int.
J. Mod. Phys. A **4**, 617 (1989); G. L. Kane, G. A. Ladinsky and C. P. Yuan, Phys.
Rev. D **45**, 124 (1992); G. Mahlon and S. J. Parke, Phys. Rev. D **53**, 4886 (1996)
[arXiv:hep-ph/9512264]; D. Atwood, A. Aeppli and A. Soni, Phys. Rev. Lett. **69**, 2754
(1992); D. Atwood, S. Bar-Shalom, G. Eilam and A. Soni, Phys. Rept. **347**, 1 (2001)
[arXiv:hep-ph/0006032]; M. Arai, N. Okada, K. Smolek and V. Simak, Phys. Rev. D
70, 115015 (2004) [arXiv:hep-ph/0409273]; C. Csaki, J. Heinonen and M. Perelstein,
JHEP **0710**, 107 (2007) [arXiv:0707.0014 [hep-ph]]; M. M. Nojiri and M. Takeuchi,
arXiv:0802.4142 [hep-ph].

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

[31] G. P. Salam and G. Soyez, JHEP **0705**, 086 (2007) [arXiv:0704.0292 [hep-ph]].

[32] J.E. Huth *et al.*, *Proceedings of Research Directions For The Decade: Snowmass Accord 1990.*

[33] J. Pumplin, A. Belyaev, J. Huston, D. Stump and W. K. Tung, JHEP **0602**, 032 (2006) [arXiv:hep-ph/0512167]; J. Campbell and R. K. Ellis, "MCFM, A Monte Carlo for FeMtobarn processes at Hadron Colliders, http://mcfm.fnal.gov."

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

[35] J. Alwall *et al.*, Eur. Phys. J. C **53**, 473 (2008) [arXiv:0706.2569 [hep-ph]].

[36] M. L. Mangano, M. Moretti, F. Piccinini, R. Pittau and A. D. Polosa, JHEP **0307**, 001 (2003) [arXiv:hep-ph/0206293].

[37] S. Catani, F. Krauss, R. Kuhn and B. R. Webber, JHEP **0111**, 063 (2001) [arXiv:hep-ph/0109231].

[38] S. D. Ellis, J. Huston, K. Hatakeyama, P. Loch and M. Tonnesmann, Prog. Part. Nucl. Phys. **60**, 484 (2008) [arXiv:0712.2447 [hep-ph]]; A. Banfi, G. P. Salam and G. Zanderighi, JHEP **0707**, 026 (2007) [arXiv:0704.2999 [hep-ph]].

[39] J. C. Collins, D. E. Soper and G. Sterman, Adv. Ser. Direct. High Energy Phys. **5**, 1 (1988) [arXiv:hep-ph/0409313].

[40] C. F. Berger, T. Kucs and G. Sterman, Phys. Rev. D **68**, 014012 (2003) [arXiv:hep-ph/0303051].

[41] G. Sterman, arXiv:hep-ph/0412013; I. M. Dremin, AIP Conf. Proc. **828**, 30 (2006) [arXiv:hep-ph/0510250]; T. Han, arXiv:hep-ph/0508097; G. Sterman, arXiv:hep-ph/9606312.

[42] P. Meade and L. Randall, arXiv:0708.3017 [hep-ph].

[43] C. Amsler *et al.* [Particle Data Group], Phys. Lett. B **667** (2008) 1, http://www-pdg.lbl.gov/2008/reviews/statrpp.pdf.

[44] M. Strassler, "Unusual Physics Signatures at the LHC," talk presented at the 2007 Phenomenology Symposium - Pheno 07, University of Wisconsin, Madison, May 7-9, 2007.

[45] L. Randall and R. Sundrum, Phys. Rev. Lett. **83**, 3370 (1999) [arXiv:hep-ph/9905221].

[46] M. Furman, Nucl. Phys. B **197**, 413 (1982).

[47] F. Aversa, P. Chiappetta, M. Greco and J. P. Guillet, Nucl. Phys. B **327**, 105 (1989), Z. Phys. C **46**, 253 (1990).

[48] N. Kidonakis and J. F. Eq.wens, Phys. Rev. D **63**, 054019 (2001) [arXiv:hep-ph/0007268].

[49] N. Kidonakis, G. Oderda and G. Sterman, arXiv:hep-ph/9805279; N. Kidonakis, G. Oderda and G. Sterman, Nucl. Phys. B **531**, 365 (1998) [arXiv:hep-ph/9803241].