Exploiting jet binning to identify the initial state of high-mass resonances

Markus A. Ebert,1 Stefan Liebler,1 Ian Moult,2 Iain W. Stewart,2 Frank J. Tackmann,1 Kerstin Tackmann,1 and Lisa Zeune3

1Deutsches Elektronen-Synchrotron (DESY), D-22607 Hamburg, Germany
2Center for Theoretical Physics, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, USA
3Nikhef, Theory Group, Science Park 105, 1098 XG Amsterdam, The Netherlands

(Received 7 June 2016; published 28 September 2016)

If a new high-mass resonance is discovered at the Large Hadron Collider, model-independent techniques to identify the production mechanism will be crucial to understand its nature and effective couplings to Standard Model particles. We present a powerful and model-independent method to infer the initial state in the production of any high-mass color-singlet system by using a tight veto on accompanying hadronic jets to divide the data into two mutually exclusive event samples (jet bins). For a resonance of several hundred GeV, the jet binning cut needed to discriminate quark and gluon initial states is in the experimentally accessible range of several tens of GeV. It also yields comparable cross sections for both bins, making this method viable already with the small event samples available shortly after a discovery. Theoretically, the method is made feasible by utilizing an effective field theory setup to compute the jet cut dependence precisely and model independently and to systematically control all sources of theoretical uncertainties in the jet binning, as well as their correlations. We use a 750 GeV scalar resonance as an example to demonstrate the viability of our method.

DOI: 10.1103/PhysRevD.94.051901

1. INTRODUCTION

The increased center-of-mass energy of the Large Hadron Collider (LHC) significantly enhances the sensitivity for the discovery of new heavy particles. Should a new high-mass state be found, a key goal will be to identify its production mechanism.

It is well known that the different patterns of initial-state radiation (ISR) for gluon- and quark-induced processes provide in principle a way to discriminate between these initial states. Typically, methods to exploit this fact require a substantial amount of data for the precise measurement of shapes of differential distributions. In this paper, we show that for any high-mass color-singlet system, the measurement of just two cross sections, namely dividing the data into events with and without additional hadronic jets in the final state, provides a strong discrimination between production mechanisms, which is furthermore experimentally accessible with event samples of limited size. The method is also theoretically clean, as it is both model independent and has well-controlled theory uncertainties.

As a concrete example, we investigate a color-singlet resonance with a mass of 750 GeV. The ATLAS and CMS experiments have recently reported some deviation from the background expectation in the diphoton invariant mass spectrum around 750 GeV [1,2]. Assuming the deviation to be a first sign of a new particle, a large number of proposals on its interpretation and possible property studies have been made [3]. Exploratory studies of the initial state have utilized the luminosity ratio between 8 and 13 TeV (which is limited by the available 8 TeV data), the transverse momentum and rapidity distribution of the new state [4], multiplicity and kinematic distributions of hadronic jets [5], and b-tagging [4,6]. Different techniques for tagging the initial state have been studied earlier, see e.g. Refs. [7–12]. Beyond its viability for small data sets, our method offers several additional advantages. Compared to considering additional jets at high \( p_T \) [5,6,13], the low \( p_T^{\text{jet}} \)-range we exploit has more discrimination power and is more model independent. Compared to the diphoton \( p_T \) spectrum, the \( p_T^{\text{jet}} \) of hadronic jets provides a more direct measure of ISR, making it insensitive to the possibility of more complicated decays of the resonance, for example, three-body [14,15] or cascade decays [16–21]. Our method is also unaffected by limited experimental acceptance for photons, which for example hinders fully exploiting the diphoton rapidity distribution to discriminate valence quarks by their different parton distribution function (PDF) shapes.

The 0-jet cross section is defined by requiring that all accompanying jets have \( p_T^{\text{jet}} \geq p_T^{\text{cut}} \). The QCD dynamics of low-\( p_T \) radiation produced in association with a hard scattering process into a final state \( F \) with total invariant mass \( m(F) = m_X \) can be described using the soft collinear effective theory (SCET) [22–25]. At the scale \( \mu = p_T^{\text{cut}} \ll m_X \), the leading effective field theory (EFT) Lagrangian has the form (see e.g. [26–28])

\[
\mathcal{L}_{\text{eff}}(p_T^{\text{cut}}) = \mathcal{L}_{\text{SCET}} + c_{yqF} \lambda_{yqF}^b \bar{b}^b F + \sum_q c_{qF} \lambda_{qF}^b \bar{q}^b X^{\lambda_{qF}}. \tag{1}
\]
Here, $\mathcal{L}_{\text{SCET}}$ is the universal Lagrangian encoding the interactions of soft and collinear quarks and gluons. The gauge-invariant operators $B_{\mu}B_{\nu}$ and $\bar{X}_{\mu}\gamma_{\nu}\bar{q}g$ describe the annihilation of energetic gluons or quarks $q = u, d, s, c, b$ along the beam directions, $n = (1, \hat{z})$ and $\bar{n} = (1, -\hat{z})$, with helicities $\lambda_1$ and $\lambda_2$ (implicitly summed over), and $\mathcal{F}$ collects all fields required to produce $F$. All hard degrees of freedom are integrated out, including quarks and gluons of virtuality $\sim m_X$ as well as any intermediate new heavy degrees of freedom leading to $F$.

We stress that $\mathcal{L}_{\text{eff}}(m_X)$ provides a completely model-independent description of the small-$p_T^\text{cut}$ region, up to power corrections suppressed by $(p_T^\text{cut}/m_X)^2$. It is valid for any produced color-singlet system $X$ leading to $F$, for example the decay of a finite-width resonance, the Standard Model background $pp \to F$, and even the signal-background interference [29], as all of the dynamics of the production and decay of $X$ are contained in the hard Wilson coefficients $c_{ij}^{\text{H,F}}$ and because the leading perturbative SCET dynamics are insensitive to the helicity structure of the operators and the details of $\mathcal{F}$. The 0-jet cross section thus only depends on the hard coefficients

$$|c_{ij}|^2 = \int d\Phi F \sum_{\lambda_1\lambda_2}|c_{ij}^{\lambda_1\lambda_2}(\Phi_F)|^2,$$

where the integral is over the final-state phase space for $F$ including any kinematic selection cuts. As a result, the $p_T^\text{cut}$ dependence is independent of any details of $F$ and in particular also the spin of $X$. (See Ref. [30] for a detailed analysis in a specific new-physics context.)

To predict the inclusive cross section $pp \to X \to F$ we need to know the Lagrangian at the scale $\mu \sim m_X$, which contains the full QCD Lagrangian plus the (effective) interactions of $X$ with quarks and gluons. This becomes somewhat more model dependent, and requires, for example, specifying the spin of $X$. For our concrete study we take $X$ to be a scalar, coupling to gluons and quarks via the effective Wilson coefficients $C_q$ and $C_g$ as

$$\mathcal{L}_{\text{eff}}(m_X) \supset -\frac{C_g}{1\text{ TeV}} \alpha_s G^{\mu\nu}G_{\mu\nu}X - \sum_q C_q \bar{q}qX,$$

where $G^{\mu\nu}$ is the gluon field strength and $\alpha_s$ is the strong coupling. (We assume that $X$ does not couple directly to top quarks, as this would have shown up in $t\bar{t}$ production.) Comparing quark and gluon luminosities as is often done is equivalent to using Eq. (3) at leading order (LO). With Eq. (3) specified, we can now match it onto Eq. (1) and compute $|c_{qgF}|^2$ and $|c_{qFF}|^2$.

For our purposes, any model can be represented by Eq. (3) at leading order in $\alpha_s(m_X)$. Treated as an EFT, Eq. (3) is a priori only correct to $\mathcal{O}(m_X/\Lambda)$, where $\Lambda$ is the mass scale of additional heavy degrees of freedom that induce the effective interactions of $X$. For example, the possibility of real QCD radiation from internal heavy states is not captured by Eq. (3). However, even for $\Lambda \sim m_X$ hard emissions only affect the inclusive cross section by $\mathcal{O}(\alpha_s(m_X))$, while emissions below the scale $p_T^\text{cut}$ are power suppressed. Similarly, a different choice of $\mathcal{L}_{\text{eff}}(m_X)$ in Eq. (3) (e.g. for a spin-2 resonance) changes the inclusive cross section and the matching in Eq. (4) only by terms of $\mathcal{O}(\alpha_s(m_X))$, i.e., at the 10%–20% level. The crucial point is that the $p_T^\text{cut}$ dependence for $p_T^\text{cut} \ll m_X$ is described by Eq. (1). Hence, for more complicated scenarios than the one considered here, our main conclusions regarding the initial-state discrimination are unaffected as they rest on the dynamics at the scale $\mu \sim p_T^\text{cut}$, which is described model independently.

II. CALCULATIONAL SETUP

Considering for simplicity the narrow-width approximation, we have

$$|c_{qgF}(\mu_H)|^2 = B(X \to F)|C_q(\mu_H)(1 + \cdots)|^2,$$

$$|c_{qFF}(\mu_H)|^2 = B(X \to F)|\alpha_s(\mu_H)C_g(\mu_H)(1 + \cdots)|^2,$$

where $\mu_H \sim m_X$ is the hard matching scale, and the ellipses indicate the $\sigma_i(\mu_H)$ corrections from hard virtual QCD emissions. The branching ratio $B \equiv B(X \to F)$ also depends on all Wilson coefficients $C_i$, but will drop out in our final analysis.

The jet cross sections we consider are given by

$$\sigma_{\geq 0} = |C_g|^2 \sigma_{\geq 0}^g + \sum_q |C_q|^2 \sigma_{\geq 0}^q,$$

$$\sigma_0(p_T^\text{cut}) = |C_g|^2 \sigma_0^g(p_T^\text{cut}) + \sum_q |C_q|^2 \sigma_0^q(p_T^\text{cut}),$$

$$\sigma_{\geq 1}(p_T^\text{cut}) = |C_g|^2 \sigma_{\geq 1}^g(p_T^\text{cut}) + \sum_q |C_q|^2 \sigma_{\geq 1}^q(p_T^\text{cut}),$$

where $\sigma_{\geq 0} = \sigma_0(p_T^\text{cut}) + \sigma_{\geq 1}(p_T^\text{cut})$. We take $C_i \equiv C_i(m_X)$ as the unknown parameters to be determined from the data. Their evolution from the fixed input scale $m_X$ to the hard matching scale $\mu_H$ is included in the $\sigma_m$ in Eq. (5), so they are defined to be scale independent to all orders. The $\sigma_{\geq 0}$ and $\sigma_{\geq 0}$ are the inclusive cross sections that follow from the $ggX$ and $q\bar{q}X$ operators in $\mathcal{L}_{\text{eff}}(m_X)$ in Eq. (3). The inclusive 1-jet cross sections are computed as $\sigma_{\geq 0}(p_T^\text{cut}) = \sigma_{\geq 0}^q(p_T^\text{cut})$. The 0-jet cross sections contain large Sudakov logarithms of $p_T^\text{cut}/m_X$, which are resummed utilizing the $p_T^\text{cut}$ resummation framework of Refs. [31,32] based on SCET (see also Refs. [33–36]). They are given by
EXPLOITING JET BINNING TO IDENTIFY THE ...

\[ |C_p|^2 \sigma_0^p(p_T^\text{cut}) = \frac{\pi}{6E_{\text{cm}}^2} \left[ \frac{m_X^2}{p_T^\text{cut}} \int dY B_p(p_T^\text{cut}, \mu) \times B_q(p_T^\text{cut}, \mu) S_{qq}(p_T^\text{cut}, \mu) + \sigma_0^\text{nons}(p_T^\text{cut}) \right] \]

The \( B_i \) are quark and gluon beam functions, which describe the dynamics of collinear radiation along the beam directions. In general, the type of the incoming parton is changed by both collinear PDF evolution and fixed-order corrections, so beyond LO an operator in the Lagrangian receives contributions from all PDFs. However, in the 0-jet cross section, both of these effects only occur up to the scale \( p_T^\text{cut} \) and are contained in the beam functions. Above the scale \( p_T^\text{cut} \), the parton type of the initial state is uniquely defined and matches between the beam function and the operator in the Lagrangian [26]. Similarly, the dynamics of wide-angle soft radiation, described by the soft functions \( S_{gg/qq^*} \), is unique to the parton type and does not change it. The fact that the jet veto freezes the initial-state parton type at the scale \( p_T^\text{cut} \) is what lends our method its strong sensitivity to \( \alpha_q \) and \( \alpha_g \). For all our results we use a jet radius of 0.4. The analogous results for \( d \) and \( s \) quarks are mostly indistinguishable from \( u \) quarks and are not shown. The split of the cross section into the 0-jet and \( \geq 1 \)-jet bins is clearly different for the different initial states, allowing one to distinguish light quarks (\( u,d,s \)), \( c \), \( b \) quarks, and gluons. The discrimination between \( b \) quarks and gluons is less good, in part due to the sizeable uncertainty in the gluon cross sections. The optimal \( p_T^\text{cut} \) value for discriminating between different possible initial states depends on the true initial states, and \( C_{\text{resum}} \), is induced by the binning cut and is correspondingly treated as a migration uncertainty that is fully anti-correlated between \( \sigma_0 \) and \( \sigma_1 \) and drops out of \( \sigma_2 \). The individual uncertainty contributions are estimated using profile scale variations [45,46] for the relevant resummation scales, as discussed in detail in Ref. [32]. Finally, we use a complex hard scale \( \mu_H = -im_X \) to resum large virtual QCD corrections. The corresponding resummation uncertainty, \( C_{\text{resum}} \), is estimated by varying the phase of \( \mu_H \) and corresponds to a yield uncertainty. The perturbative uncertainties are treated as fully correlated among all quark flavors and uncorrelated between quarks and gluons. We use the MMHT2014nnlo68cl [47] PDFs with the corresponding \( \alpha_s(m_Z) = 0.118 \) and 3-loop running. The parametric PDF uncertainties, \( C_{\text{PDF}} \), are constructed from the 25 independent eigenvectors of MMHT2014nnlo68cl. They are subdominant compared to the perturbative uncertainties.

III. INITIAL-STATE DISCRIMINATION

As a demonstration of the technique, we consider a hypothetical scalar resonance of mass \( m_X = 750 \text{ GeV} \) produced in 13 TeV \( pp \) collisions. In Fig. 1 we show the ratio \( \sigma_0(p_T^\text{cut})/\sigma_1(p_T^\text{cut}) \) as a function of \( p_T^\text{cut} \) for different initial states \( i = u, c, b, g \) including all theoretical uncertainties, see Eq. (7). For all our results we use a jet radius of 0.4. The analogous results for \( d \) and \( s \) quarks are mostly indistinguishable from \( u \) quarks and are not shown. The split of the cross section into the 0-jet and \( \geq 1 \)-jet bins is clearly different for the different initial states, allowing one to distinguish light quarks (\( u,d,s \)), \( c \) quarks, \( b \) quarks, and gluons. The discrimination between \( b \) quarks and gluons is less good, in part due to the sizeable uncertainty in the gluon cross sections. The optimal \( p_T^\text{cut} \) value for discriminating between different possible initial states depends on the true initial states.
state. Numerically, we observe only a mild sensitivity of the discrimination between initial states on the $p_T^{\text{cut}}$ value within the range $p_T^{\text{cut}} \in [25, 65]$ GeV. The $p_T^{\text{cut}}$ value can thus be chosen to optimize the experimental sensitivity with limited statistics. A roughly equal split of the cross section is achieved at $p_T^{\text{cut}} \approx 25$ GeV (65 GeV) for a light-quark (gluon) induced signal. In our subsequent analysis we use $p_T^{\text{cut}} = 40$ GeV, for which the cross section ratio is between 0.5 and 2 for any initial state. Note that effects from hadronization and multiparton interactions, which are not included in our calculations, can affect the leading jet $p_T$ spectrum at small $p_T$. However, their effects partially compensate each other. We checked that the net effect in the cross section ratios we consider becomes negligible above $p_T^{\text{cut}} \gtrsim 20$ GeV.

To study the constraints on the Wilson coefficients $C_i$ from measuring $\sigma_0$ and $\sigma_{\geq 1}$, we minimize the $\chi^2$ function

$$\chi^2(C_i) = \sum_{m,n \in \{0, \geq 1\}} (\sigma_{m,\text{meas}} - \sigma_m) ((C^{-1})_{mn}(\sigma_{n,\text{meas}} - \sigma_n),$$

where $\sigma_{m,\text{meas}}$ is the measured $pp \to X \to F$ cross section in bin $m$, $\mathcal{C}$ is the sum of the experimental and theory covariance matrices, and $\sigma_m$ is the predicted cross section in bin $m$ in Eq. (5). In the narrow-width approximation and considering a single decay channel (e.g. $F = \gamma\gamma$), we only constrain $C_i\sqrt{B}$ [see Eq. (4)]. To render our results independent of the details of $F$, we define

$$C_i^{\text{incl}} = \sqrt{\frac{\sigma_{\geq 0}}{\sigma_{0}}},$$

which is the value of $C_i\sqrt{B}$ for which the measured inclusive cross section is completely attributed to initial state $i$. By considering the ratios $|C_i/C_i^{\text{incl}}|$, our analysis only depends on the ratio $\sigma_0/\sigma_{\geq 1}$ but not the absolute cross sections or $B$. With more than one final state, the sum in Eq. (8) runs over the respective bins for all final states, and the dependence of the different branching ratios on the Wilson coefficients $C_i$ and the associated uncertainties need to be taken into account.

FIG. 2. $\Delta\chi^2 = 1$-contours for various scenarios. (a) gluon signal, (b) mixed gluon/$u$-quark signal, (c) $u$-quark signal, (d) $b$-quark signal, (e) $c$-quark signal, (f) gluon signal. The constraints from $\sigma_0$ and $\sigma_{\geq 1}$ are shown by the blue and green bands, respectively. The combined constraint from both are shown by the orange/yellow regions. The inner darker regions correspond to theory uncertainties only, while the full lighter bands include both theory and assumed experimental uncertainties.
IV. CONCLUSIONS

Should the deviation in the diphoton spectrum at 750 GeV manifest into a discovery, the method proposed here can be readily applied to identify its initial state. It is then preferable for the measurements to be fiducial in the kinematics of the X decay products to minimize the model dependence introduced by acceptance corrections.

We restricted our attention to discriminate quark- and gluon-initiated production. Using our method, it will also be possible to identify photoproduction, which has been considered in several recent studies [48–55]. In this case, the 0-jet cross section is given in terms of photon beam functions calculated in terms of photon PDFs at the scale \( \mu \sim p_T^\text{cut} \) and without QCD evolution above \( p_T^\text{cut} \). This implies that the \( \sigma_0/\sigma_{\geq 1} \) ratio will be substantially larger than for light quarks, thus providing a good discrimination against the production via quarks and gluons. The utility of a jet-veto to distinguish photon production from vector-boson fusion or gluon initial states was discussed in [50].

Our method allows for an early, model-independent, and theoretically clean identification of the production mechanism of any new high-mass color-singlet state. Since the ratio \( \sigma_0/(p_T^\text{cut})/\sigma_{\geq 1}/(p_T^\text{cut}) \) depends on \( p_T^\text{cut}/m_X \), and \( p_T^\text{cut} \geq 25 \text{ GeV} \) is experimentally feasible at the LHC, we expect it to work well for masses \( m_X \gtrsim 300 \text{ GeV} \).

ACKNOWLEDGMENTS

This work was partially supported by the DFG Emmy-Noether Grant No. TA 867/1-1, the Office of Nuclear Physics of the U.S. Department of Energy under the Grant No. DE-SC0011090, the Collaborative Research Center SFB 676 of the DFG “Particles, Strings and the Early Universe,” the Simons Foundation through the Investigator Grant No. 327942, the Netherlands Organization for Scientific Research (NWO) through a VENI Grant, the PIER Helmholtz Graduate School, and by a Global MISTI Collaboration Grant from MIT.