QCD Resummation for Jet Substructures

Hsiang-nan Li, Zhao Li, and C.-P. Yuan

Phys. Rev. Lett. **107**, 152001 — Published 3 October 2011

DOI: 10.1103/PhysRevLett.107.152001

QCD resummation for jet substructures

Hsiang-nan Li,[1,*] Zhao Li,[2,†] and C.-P. Yuan[2,3,‡]

[1]*Institute of Physics, Academia Sinica, Taipei, Taiwan 115, Republic of China,*
[2]*Department of Physics & Astronomy, Michigan State University, E. Lansing, MI 48824, USA*
[3]*Center for High Energy Physics, Peking University, Beijing 100871, China*

We provide a novel development in jet physics by predicting the energy profiles of light-quark and gluon jets in the framework of perturbative QCD. Resumming large logarithmic contributions to all orders in the coupling constant, our predictions are shown to agree well with Tevatron CDF and Large-Hadron-Collider CMS data. We also extend our resummation formalism to the invariant mass distributions of light-quark and gluon jets produced in hadron collisions. The predicted peak positions and heights in jet mass distributions are consistent with CDF data within uncertainties induced by parton distribution functions.

PACS numbers: 12.38.Cy,12.38.Qk,13.87.Ce

It has been a long-standing challenge to predict substructures (including energy profiles and masses) of light-quark and gluon jets in the perturbative QCD (pQCD) theory. During the Tevatron Run 1 era in early 90's, it was found that next-to-leading-order (NLO) QCD calculations cannot describe experimental data on jet substructures. Hence, it has been a custom for experimentalists to compare jet substructures measured at the Tevatron, either at Run 1 or Run 2, with predictions from full event generators such as PYTHIA or HERWIG. While the full event generators (usually with specific tuning) could describe data, it remains desirable to develop a theoretical framework for the study of jet substructures. For that, the soft-collinear effective theory was adopted in the literature, such as [1, 2]. In this Letter, we propose a novel approach to predicting jet substructures based on the pQCD resummation formalism[3]. We show that results of the resummation formalism for light-quark and gluon jets are well consistent with the energy profiles measured by CDF at Tevatron [4] and CMS at Large Hadron Collider (LHC) [5], and with the mass distributions measured by CDF [6].

It is known that top quark is dominantly produced at rest at the Tevatron and can be clearly identified by detecting three (or more) isolated jets from its hadronic decays. However, this strategy will not work for identifying a highly boosted top quark [7–10] at the LHC, which results in a single jet. Furthermore, it has been pointed out [11, 12] that an energetic QCD (light-quark or gluon) jet can have an invariant mass around the top quark mass to fake a top-quark jet. That is, a boosted top-quark jet is difficult to be discriminated against an ordinary QCD jet. This difficulty also appears in the identification of a highly boosted Higgs boson decaying into a bottom-quark pair [13, 14], or a highly boosted W or Z boson decaying into hadronic final states, for they can all produce a single-jet experimental signature. In order to improve jet identification at the LHC, additional information from jet energy profiles is needed, because jets initiated by different parent particles usually produce different energy profiles.

Denote the jet energy function as $J_f^E(M_J^2, P_T, \nu^2, R, r)$ for defining a light-quark ($f = q$) or gluon ($f = g$) jet with mass M_J, transverse momentum P_T, and cone size R, which describes the all-order energy distribution within a smaller cone of size $r < R$. J_f^E is constructed by inserting a sum of the step functions $\sum_i k_{iT}\Theta(r - \theta_i)$ into the usual jet definition [15], where k_{iT} and θ_i are the transverse momentum and the angle of the final-state particle i with respect to the jet axis. At leading-order (LO), it is a δ-function, i.e., $J_f^{E(0)} = P_T\delta(M_J^2)$, which is independent of r, because $\theta = 0$ at this order. The jet definition contains a Wilson line along the light cone, which collects gluons emitted from other parts of the collision process and collimated to the parent particle of the jet. To employ the resummation technique, we vary the Wilson line into an arbitrary direction n^μ with $n^2 \neq 0$ [16]. The dependence of J_f^E on n^μ and the jet momentum P_J appears through the invariants n^2, $P_J^2 = M_J^2$, and $P_J \cdot n$ (which is related to P_T). When r approaches to zero, the phase space of real radiation is strongly constrained, so the infrared enhancement in real radiation does not cancel completely with that in virtual correction. The resultant large logarithms of the ratio $(P_J \cdot n)^2/(n^2 r^2)$, which is conveniently defined as $[R^2 P_T^2/(4r^2)]\nu^2$, should be resummed to all orders in the coupling constant α_s. It is easy to see from the above ratio that the variation in n, i.e., ν^2, can turn into the variation in r. To compare with present experimental data on jet energy profile , we consider the jet energy function with the jet invariant mass being integrated out, which corresponds to taking the $N = 1$ moment in the Mellin space and is denoted as $\bar{J}_f^E(1, P_T, \nu^2, R, r)$. By definition, different choices of n^μ yield the same collinear divergences associated wth the jet. Hence, the effect of varying n, i.e., ν^2, does not involve the collinear divergences and can be factorized out of the jet energy function, leading to an evolution equation

$$\nu^2 \frac{d}{d\nu^2}\bar{J}_f^E = \left[G^{(1)} + K^{(1)}\right]\bar{J}_f^E. \quad (1)$$

The one-loop kernel $G^{(1)}$ ($K^{(1)}$) absorbs the hard (soft) dynamics of the variational effect, whose expressions are similar to those derived in [17], but with the step function in angle being inserted into the real-gluon piece of $K^{(1)}$.

The solution describing the evolution of the jet energy function from the initial value $\nu_{\rm in}^2 = C_1^2 r^2/(C_2^2 R^2)$ to the final value $\nu_{\rm fi}^2 = 1$ is written as

$$\bar{J}_f^E(1, P_T, \nu_{\rm fi}^2, R, r) = \bar{J}_f^E(1, P_T, \nu_{\rm in}^2, R, r) \times \exp\left\{-\int_{C\nu_{\rm in}^2}^{C} \frac{dy}{y}\left[\frac{1}{2}\int_{y\nu_{\rm in}^2}^{y^2} \frac{d\omega}{\omega} A(\alpha_s(\omega C_2^2 R^2 P_T^2)) - \frac{C_f}{\pi}\alpha_s\left(y^2 C_2^2 R^2 P_T^2\right)\left(\frac{1}{2} + \ln\frac{C_2}{C_1}\right)\right]\right\}, \quad (2)$$

with the cusp anomalous dimension

$$A = \frac{\alpha_s}{\pi} C_f + \frac{\alpha_s^2}{\pi^2} C_f \left[\frac{67}{12} - \frac{\pi^2}{4} - \frac{5 n_f}{18} - \frac{\beta_0}{2}\ln\frac{C_2}{C_1}\right]. \quad (3)$$

The color factor C_f is equal to $C_F(=4/3)$ and $C_A(=3)$ for the light-quark and gluon jet, respectively, β_0 is the QCD Beta function [18], and n_f is the number of active light quark flavors. The value of $\nu_{\rm in}^2$ diminishes the large logarithms in the initial condition $\bar{J}_f^E(1, P_T, \nu_{\rm in}^2, R, r)$, which is then evaluated up to NLO including non-logarithmic-r terms. The value $\nu_{\rm fi}^2 = 1$ implies the presence of the large logarithms in $\bar{J}_f^E(1, P_T, \nu_{\rm fi}^2, R, r)$, which have been summed into the Sudakov integral in Eq. (2).

We set the $\mathcal{O}(1)$ constants $C_1 = C_2 = 1$ and $C = \exp(5/2)$ ($C = \exp(17/6)$) for quark (gluon) jet in order to reproduce the large logarithms $\alpha_s \ln^2 r$ and $\alpha_s \ln r$ in the NLO calculations. The variation of these $\mathcal{O}(1)$ constants reflects theoretical uncertainty in our formalism. The value of r in lower bound is taken to be larger than 0.1, so that it is safe to evaluate the Sudakov integral perturbatively. We then derive the energy profile $\Psi(r)$ [4] as the energy fraction accumulated within the cone of size $r < R$ in terms of the solution in Eq. (2),

$$\Psi(r) = \sum_f \int \frac{dP_T}{P_T}\frac{d\sigma_f}{dP_T}\bar{J}_f^E(1, P_T, \nu_{\rm fi}^2, R, r) \times \left[\sum_f \int \frac{dP_T}{P_T}\frac{d\sigma_f}{dP_T}\bar{J}_f^E(1, P_T, \nu_{\rm fi}^2, R, R)\right]^{-1} \quad (4)$$

which respects the normalization $\Psi(r = R) = 1$. Eq. (4) contains the convolution of the LO differential cross section $d\sigma_f/dP_T$ and the quark and gluon jet energy functions. Using the CTEQ6L parton distribution functions (PDFs) [19], we compare the resummation and NLO predictions in Fig. 1, with the Tevatron CDF data [4]. The agreement between the resummation predictions and the CDF data is obvious for all P_T values. As P_T increases, the accumulation of energy inside jets becomes faster. The NLO predictions derived from $\bar{J}_f^{E(1)}(1, P_T, \nu_{\rm fi}^2, R, r)$ are also displayed, which overshoot the data. Figure 2 shows the agreement of the resummation predictions for the jet energy profiles with the LHC CMS data at 7 TeV [5], and the overshooting of the NLO predictions. The above consistency indicates that our resummation formalism has captured the dominant dynamics in a jet energy profile, and can give a direct and reliable prediction for this observable.

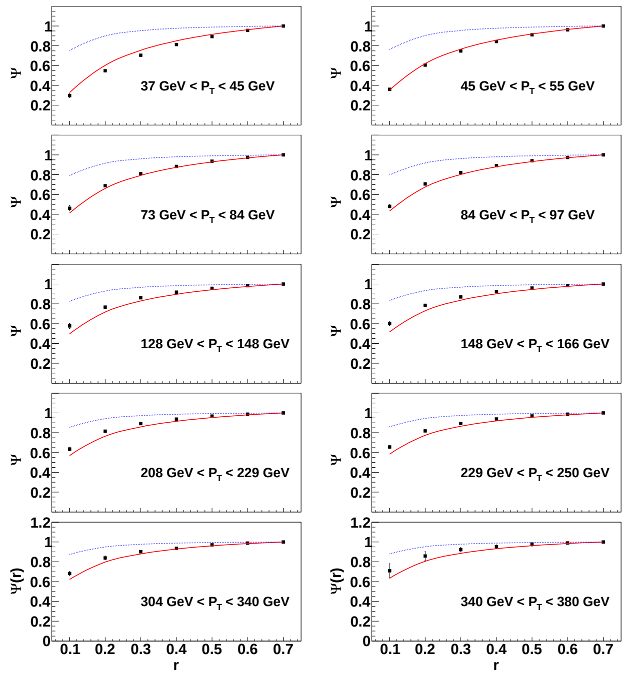

FIG. 1: Resummation (solid) and NLO (dashed) predictions for jet energy profiles compared with CDF data [4].

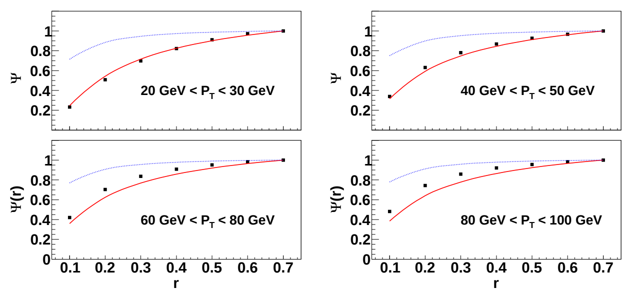

FIG. 2: Resummation (solid) and NLO (dashed) predictions for jet energy profiles compared with CMS data [5].

As stated before, the invariant mass distribution of an energetic jet can be utilized as an experimental signature of new physics. We shall demonstrate below that our formalism is also applicable to this jet substructure by predicting the mass distributions of the light-quark and gluon jets. Following the previous analysis on jet energy profiles, we vary the Wilson line into an arbitrary direction n^μ with $n^2 \neq 0$, when implementing the resummation technique [16] to derive the evolution equation for the jet function $J_f(M_J^2, P_T, \nu^2, R)$. In this case, the overlap of the collinear and soft enhancements generate the double logarithms of the ratio $(P_J \cdot n)^2/(M_J^2 n^2) \equiv (R^2 P_T^2/(4M_J^2))\nu^2$, so that the variation of n, i.e., ν^2, can turn into the variation of M_J. The dependence on the jet cone size R is introduced through the Mellin transformation with respect to the variable $x \equiv M_J^2/(R^2 P_T^2)$. The solution describing the evolution of the jet function from the initial value $\nu_{\rm in}^2 = C_1/(C_2 \bar{N})$, where $\bar{N} = N\exp(\gamma_E)$ with γ_E being the Euler's constant, to the final value $\nu_{\rm fi}^2 = 1$ is given by

$$\bar{J}_f(N, P_T, \nu_{\rm fi}^2, R) \;=\; \bar{J}_f(N, P_T, \nu_{\rm in}^2, R)\exp[S(N)], \quad (5)$$

in which the Sudakov exponent is written as

$$\begin{aligned} S(N) \;=\; &-\int_{C_1/\bar{N}}^{C_2} \frac{dy}{y}\left\{A(\alpha_s(y^2R^2P_T^2))\ln\left(\frac{C_2}{y}\right)\right. \\ &\left. -\frac{C_f}{\pi}\alpha_s\left(y^2R^2P_T^2\right)\left[\frac{1}{2}+\ln\frac{C_2}{C_1}\right]\right\}. \end{aligned} \quad (6)$$

The value $\nu_{\rm fi}^2 = 1$ implies the presence of the large logarithms (in powers of $\ln N$) in $\bar{J}_f(N, P_T, \nu_{\rm fi}^2, R)$, which have been summed into the Sudakov exponent. The initial condition $\bar{J}_f(N, P_T, \nu_{\rm in}^2, R)$ can then be evaluated perturbatively. At NLO, they have the expressions

$$\begin{aligned} \bar{J}_q^{(1)} &= \frac{1}{R^2P_T^2}\left\{1+\frac{C_F}{\pi}\alpha_s\left(C_2^2R^2P_T^2/C_1^2\right)\left[\frac{1}{2}\ln\frac{C_1}{C_2}\right.\right. \\ &\left.\left. -\frac{1}{2}\ln^2\frac{C_1}{C_2}+\frac{1}{2}\gamma_E-\frac{\pi^2}{4}-\frac{9}{8}\right]\right\}, \quad (7) \\ \bar{J}_g^{(1)} &= \frac{1}{R^2P_T^2}\left\{1+\frac{C_A}{\pi}\alpha_s\left(C_2^2R^2P_T^2/C_1^2\right)\left[\frac{1}{2}\ln\frac{C_1}{C_2}\right.\right. \\ &\left.\left. -\frac{1}{2}\ln^2\frac{C_1}{C_2}-\frac{5}{12}\gamma_E-\frac{\pi^2}{4}+\frac{1}{2}(\ln 2-3)+\frac{1}{36}\right]\right\} \quad (8) \end{aligned}$$

for the light-quark and gluon jets, respectively. The remaining N-dependent terms are suppressed by $1/N$ and their effect is small.

For a fixed P_T, the scale RP_T/N involved in Eq. (6) becomes so low at extremely large N that the perturbative analysis fails and nonperturbative contributions, arising from hadronization and underlying events, need to be included in order to predict jet mass distribution in the small M_J region. For convenience, we introduce

$$S^{NP}(N) = \frac{N^2Q_0^2}{R^2P_T^2}(C_f\alpha_0\ln N+\alpha_1)+C_f\alpha_2\frac{NQ_0}{RP_T}, \quad (9)$$

into the Sudakov exponent with Q_0 being set to 1 GeV. It consists of a logarithmic term and a Gaussian smearing term (as suggested by $S(N)$ in Eq. (6)), as well as a linear term in N (for describing a final state jet [20]). The parameters $\alpha_{0,1,2}$ can be determined by a fit to experimental data for certain jet momentum and jet cone. In this work, we perform fits to full event generators PYTHIA [21] and SpartyJet [22] for the quark (gluon) jet produced at the Tevatron Run 2 energy 1.96 TeV with $P_T = 600$ GeV and $R = 0.7$, which yields $\alpha_0 = -0.35$, $\alpha_1 = 0.50$ (-4.59), and $\alpha_2 = -1.66$. With this nonperturbative contribution, we are ready to predict jet mass distribution for arbitrary values of P_T and R using the improved resummation solution:

$$\begin{aligned} \bar{J}_f^{\rm RES}(N, P_T, \nu_{\rm fi}^2, R) \;=\; &\bar{J}_f(N, P_T, \nu_{\rm in}^2, R) \\ &\times\exp\left[S(N)+S^{NP}(N)\right]. \end{aligned} \quad (10)$$

The jet function in M_J space is derived via the inverse Mellin transformation

$$J_f^{\rm RES}(M_J^2) = \frac{1}{2\pi i}\int_{\rm C} dN(1-x)^{-N}\bar{J}_f^{\rm RES}(N), \quad (11)$$

where $x = M_J^2/(R^2P_T^2)$ and the contour C runs along negative real axis in complex-N plane and circles the origin counterclockwise with a finite radius. To avoid the Landau pole, we flatten the running behavior of α_s [18, 23] at certain scale μ_c:

$$\alpha_s(\mu) = \left\{\begin{array}{ll} \alpha_s(\mu_c\exp[{\rm iArg}(\mu)]), & |\mu| < \mu_c \\ \alpha_s(\mu), & |\mu| > \mu_c \end{array}\right. . \quad (12)$$

In the small M_J region, all moments in N are equally important, and those containing powers of $\ln N$ dominate. Therefore, our resummation formalism, in which these large logarithms are summed up, can give reliable predictions. In the large M_J region, the non-logarithmic terms are not negligible, so we have to improve the resummation formula by including the non-logarithmic terms (say, up to NLO), namely, by matching fixed-order results to the resummation formula as done for Drell-Yan processes [24].

The jet mass distribution is calculated by convoluting the LO differential cross section of inclusive (quark or gluon) jet production with the corresponding quark or gluon jet function,

$$\frac{d\sigma}{dM_J^2} \;=\; \sum_f \int dP_T\frac{d\sigma_f}{dP_T}J_f(M_J^2, P_T, \nu_{\rm fi}^2, R). \quad (13)$$

We are now ready to compare our predictions for the jet mass distributions with the Tevatron and LHC data, by choosing $C_1 = \exp(\gamma_E)$, $C_2 = \exp(-\gamma_E)$ and $\mu_c = 0.3$ GeV in the numerical analysis with the CTEQ6L PDFs [19]. Results from different choices of C_1, C_2 and μ_c reveal theoretical uncertainty, which will be investigated in a forthcoming paper. The comparison to the Tevatron data [6], with the kinematic cuts $P_T > 400$ GeV and $0.1 < |Y| < 0.7$, is presented in Fig. 3, where the label NLL/NLO denotes the prediction of NLL resummation with the NLO initial conditions given in Eqs. (7) and (8). It shows that our predictions agree well with the CDF data for $R = 0.4$ and 0.7 in the intermediate jet mass region. In the small invariant mass region, the predictions can describe the overall shape of the jet distributions within the uncertainty induced by PDFs [6], though the peak positions are slightly lower than the CDF data. However, for large M_J, e.g. $M_J > 100(200)$ GeV for $R = 0.4(0.7)$, the resummation prediction drops off quickly and deviates from data. This can be further improved by matching to exact calculations of NLO and beyond. The resummation predictions for the jet mass

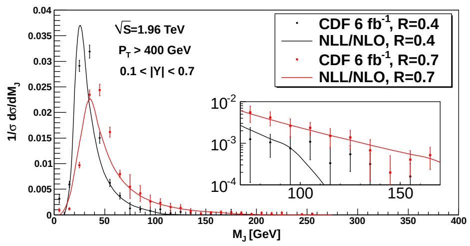

FIG. 3: Comparison between resummation predictions and Tevatron data.

distributions at Tevatron with $R = 0.3$ and at LHC with $R = 0.7$ are shown in Fig. 4, which can be tested by future measurements.

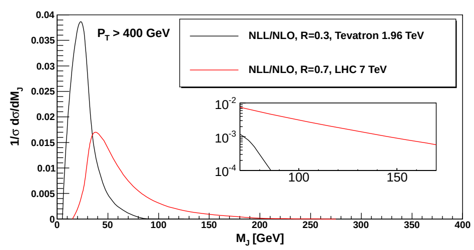

FIG. 4: Resummation predictions for jet mass distributions at Tevatron and LHC.

In conclusion, we have developed a theoretical framework based on the pQCD theory for analyzing the substructures of the light-quark and gluon jets. This is the first time in the literature that pQCD is shown to describe well the jet energy profiles and mass distributions, which are the most commonly discussed physical observables to describing the substructure of a jet signal at hadron colliders. We have demonstrated that the resummation predictions for the jet energy profiles, with the jet invariant mass being integrated out, are insensitive to nonperturbative inputs, and in excellent agreement with both the Tevatron CDF and LHC CMS data, for arbitrary values of jet momentum. In view that the jet energy profile is a useful feature for jet identification at the LHC, the energy profiles of boosted top-quark jets, and jets from boosted Higgs boson and weak gauge boson (W, Z or Z', *etc.*) hadronic decays will be studied in a forthcoming paper. We shall also calculate other jet substructures, such as the angularity [1]. These studies are crucial for LHC physics program in terms of testing the QCD theory and identifying new physics signals. We have also applied the resummation formalism to predict the light-quark and gluon jet mass distributions. To describe jet mass distributions in the low mass region, we need to introduce some nonperturbative contributions in the resummation formalism. The relevant parameters can be fixed at one jet energy scale, and then employed to make predictions for other energy scales, which are found to agree with the Tevatron data. It is expected that our formalism will apply to upcoming LHC data of jet mass distributions successfully.

This work was supported by National Science Council of R.O.C. under Grant No. NSC 98-2112-M-001-015-MY3, and by the U.S. National Science Foundation under Grand No. PHY-0855561. CPY and ZL thank the hospitality of Academia Sinica and National Center for Theoretical Sciences in Taiwan, where part of this work was done. We thank Pekka Sinervo and Raz Alon (private communication) for providing CDF jet mass distribution data.

* Electronic address: `hnli@phys.sinica.edu.tw`
† Electronic address: `zhaoli@pa.msu.edu`
‡ Electronic address: `yuan@pa.msu.edu`

[1] S. D. Ellis, C. K. Vermilion, J. R. Walsh, A. Hornig, and C. Lee, JHEP **1011**, 101 (2010).
[2] R. Kelley, M. D. Schwartz, and H. X. Zhu (2011), 1102.0561.
[3] J. C. Collins, D. E. Soper, and G. F. Sterman, Nucl.Phys. **B250**, 199 (1985).
[4] D. E. Acosta et al. (CDF), Phys. Rev. **D71**, 112002 (2005).
[5] Report CMS–PAS–QCD–10–014 (2010).
[6] Report CDF–PUB–JET–PUBLIC–10119 (2011).
[7] K. Agashe et al., Phys. Rev. **D77**, 015003 (2008).
[8] A. L. Fitzpatrick et al., JHEP **09**, 013 (2007).
[9] U. Baur and L. H. Orr, Phys. Rev. **D76**, 094012 (2007).
[10] G. H. Brooijmans et al. (2008).
[11] W. Skiba and D. Tucker-Smith, Phys. Rev. **D75**, 115010 (2007).
[12] B. Holdom, JHEP **08**, 069 (2007).
[13] J. M. Butterworth et al., Phys. Rev. Lett. **100**, 242001 (2008).
[14] E. Gabrielli et al., Nucl. Phys. **B781**, 64 (2007).
[15] L. G. Almeida et al., Phys. Rev. **D79**, 074012 (2009).
[16] H.-n. Li and H.-L. Yu, Phys. Rev. **D53**, 4970 (1996).
[17] H.-n. Li, Phys. Rev. **D55**, 105 (1997).
[18] C. Amsler et al. (Particle Data Group), Phys. Lett. **B667**, 1 (2008).
[19] J. Pumplin et al., JHEP **07**, 012 (2002).
[20] M. Dasgupta, G. E. Smye, and B. R. Webber, JHEP **04**, 017 (1998).
[21] T. Sjostrand, S. Mrenna, and P. Z. Skands, Comput. Phys. Commun. **178**, 852 (2008).
[22] S. D. Ellis et al., Prog. Part. Nucl. Phys. **60**, 484 (2008).
[23] A. Vogt, Phys. Lett. **B497**, 228 (2001).
[24] C. Balazs and C.-P. Yuan, Phys. Rev. **D56**, 5558 (1997).