Resummation prediction on the jet mass spectrum in one-jet inclusive production at the LHC

Ze Long Liu,[a] **Chong Sheng Li,**[a,b] **Jian Wang**[c] **and Yan Wang**[a]

[a]*School of Physics and State Key Laboratory of Nuclear Physics and Technology,*
Peking University,
Beijing 100871, China
[b]*Center for High Energy Physics, Peking University,*
Beijing 100871, China
[c]*PRISMA Cluster of Excellence & Mainz Institute for Theoretical Physics,*
Johannes Gutenberg University,
D-55099 Mainz, Germany

E-mail: liuzelong@pku.edu.cn, csli@pku.edu.cn, jian.wang@uni-mainz.de,
wangyanwww@pku.edu.cn

ABSTRACT: We study the factorization and resummation prediction on the jet mass spectrum in one-jet inclusive production at the LHC based on soft-collinear effective theory. The soft function with anti-k_T algorithm is calculated at next-to-leading order and its validity is demonstrated by checking the agreement between the expanded leading singular terms with the exact fixed-order result. The large logarithms $\ln^n \left(m_J^2/p_T^2 \right)$ and the global logarithms $\ln^n \left(s_4/p_T^2 \right)$ in the process are resummed to all order at next-to-leading logarithmic and next-to-next-to-leading logarithmic level, respectively. The cross section is enhanced by about 23% from the next-to-leading logarithmic level to next-to-next-to-leading logarithmic level. Comparing our resummation predictions with those from Monte Carlo tool PYTHIA and ATLAS data at the 7 TeV LHC, we find that the peak positions of the jet mass spectra agree with those from PYTHIA at parton level, and the predictions of the jet mass spectra with non-perturbative effects are in coincidence with the ATLAS data. We also show the predictions at the future 13 TeV LHC.

Contents

1 Introduction

The substructure of jets produced at the Large Hadron Collider (LHC) has become one of the hot topics for both theorists and experimentalists. The particles such as massive electroweak bosons, top quark and other possible new resonances produced with transverse momenta much greater than their masses, i.e., $p_T \gg m$, can decay to hadronic products, which are almost collinear and may be recombined into a single jet by jet algorithms. Therefore it is necessary to find a way to distinguish the interesting signal jets from the purely QCD backgrounds.

During the past few years, many studies on jet substructures have been performed [1–13], in which new techniques and observables have been designed to analyze the events. The event generators such as SHERPA [14, 15], PYTHIA [16, 17] and HERWIG++ [18, 19], can provide fully differential events, by which any observable can be predicted and compared with data. However, the various event generators employ different models for parton shower and non-perturbative effects, such as the hadronization and multiparton interactions. As a consequence, they might provide very different predictions. For instance, the jet mass spectra from the PYTHIA and the HERWIG++ do not agree with each other, as shown in ref. [20]. Moreover, there is a type of color correlation between the initial and final colored particles that is not taken into account in these event generators.

In order to obtain more precise predictions and test the validity of the Monte Carlo tools, it is important to develop a theoretical framework to study the jet substructure. Recently, various jet substructure observables have been investigated analytically based on soft-collinear effective theory (SCET) [21–29] and the traditional perturbative QCD (pQCD) resummation formalism [30–34]. For example, the factorization and resummation prediction of the jet angularity in the multijet production at e^+e^- colliders have been studied in refs. [23, 24], and the invariant mass and energy profile of jets at hadron colliders have been explored in refs. [31, 32].

The theoretical developments of prediction on jet mass spectrum at hadron colliders can be found in [28, 32, 33, 35]. In ref. [32], the jet mass was investigated with the pQCD resummation formalism by focusing on the processes independent jet function, where it was found that the nonperturbative effects are important at small jet mass. The author of ref. [33] studied the distributions of m_J/p_T^J in $pp \to$ dijet and $Z + 1$ jet processes at NLL, using the formula in refs. [36, 37], and including resummation effects of non-global logrithms (NGLs) in large-N_c approximation. The jet mass spectrum with the Higgs + 1 jet process was discussed [35] in the N-jettiness global event shape [38]. The factorization formula and resummation prediction of the jet mass spectrum for direct photon production in the framework of SCET was provided in ref. [28], where the soft function was factorized into two pieces with different scales. Thought the non-global logarithms were not resummed there, their contribution were estimated and it was found that the NGLs only affect the jet mass spectrum in the peak regions significantly.

Studies of the jet mass can not only help us understand QCD, but also be useful to search for new physics, especially in the complex QCD environment of the LHC. In particular, if we want to identify the mass peak of a highly boosted particle, the jet mass
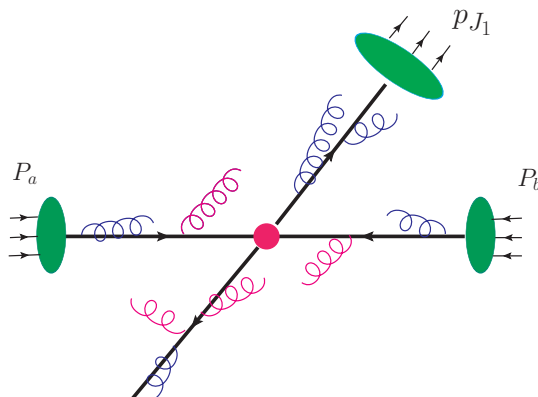
Figure 1. The illustrative picture for dijet production at the LHC. The blue and red arcs denote the collinear and soft gluons, respectively.

spectrum of QCD background must be calculated precisely. Actually, the jet invariant mass were explored in both ATALS and CMS collaborations at the 7 TeV LHC [20, 39]. From these results, we can see that the jet mass m_J peaks at about 50 GeV, which can be much smaller than the transverse momenta of jet p_T. Therefore there exist large logarithmic terms $\frac{\alpha_s^n}{m_J^2} \ln^m\left(m_J^2/p_T^2\right)$ with $m \leq 2n-1$ in the perturbative calculations near the peak region, which need to be resummed to all order in order to give reliable predictions.

In this paper, we study one of the simplest jet substructures, i.e. the invariant mass of a jet, and investigate the factorization and resummation prediction on the jet mass spectrum in SCET for one-jet inclusive production at the LHC. Compared with direct photon process [28], the factorization formula for dijet process is more complicated due to the nontrivial color structure and associating soft radiation. The illustrative picture of this process is shown in figure 1. Since the soft radiation can either be inside or outside the cone of the measured jet, there are two kinematic variables which can lead to large logarithms at threshold limit: one is the invariant mass m_J of the measured jet, and another is the invariant mass $\sqrt{s_4}$ of the partonic system that recoils against the observed jet. In the threshold region $m_J^2 \to 0$ and $s_4 \to 0$, both of the large logarithms $\ln^n\left(m_J^2/p_T^2\right)$ and $\ln^n\left(s_4/p_T^2\right)$ need to be resummed to all order. In the threshold limits, the cross section can be factorized as

$$\sigma = f_{P_a} \otimes f_{P_b} \otimes \boldsymbol{H} \otimes \boldsymbol{S} \otimes J^{\text{obs.}} \otimes J^{\text{rec.}}, \tag{1.1}$$

where $\boldsymbol{H}$, $\boldsymbol{S}$, J, f_P are the hard function, soft function, jet function and parton distribution function (PDF), respectively. Both of the hard and soft function are matrices in color space. The hard function includes the short distance contributions arising from virtual corrections. The jet function presents the collinear radiation in the jet. The indices "obs." and "rec." denote the observed jet and the recoiled one, respectively. The effects from soft gluon emission are incorporated in the soft function and its phase space is constrained by the jet algorithms. It is noteworthy that the large angle soft gluon arising from the initial state radiation (ISR) and recoiled final state radiation are taken into account in this formalism. In contrast to the cone algorithm adopted in ref. [28], we choose anti-k_T algorithm [40] to

calculate the jet and soft functions, which is boost-invariant and stable against the change of jet boundary [41]. Thus, our prediction can be valid for the jet with both small and large rapidity, and is more useful for phenomenological purposes.

This paper is organized as follows. In section 2, we analyze the kinematics of the one-jet inclusive production at hadron colliders and give the definition of the threshold region. In section 3, we derive the factorization formula. In section 4 and section 5, we show the results of hard function and jet function at NLO, respectively. We calculate the soft function at NLO and present its refactorization in section 6. In section 7, we give the final renormalization group (RG) improved cross section analytically. In section 8, we discuss the numerical results of the jet mass distribution for one-jet inclusive production at the LHC, including the leading singular distribution at threshold limit, scale uncertainties, R dependence, distinction between quark jets and gluon jets, and comparison between the RG improved predictions and ATLAS data. We conclude in section 9.

2 Analysis of kinematics and factorization

In this section, we introduce the relevant kinematical variables and the factorization formula needed in our analysis. We consider the process

$$N_1(P_a) + N_2(P_a) \rightarrow J(p_{J_1}) + X \,, \tag{2.1}$$

where J denotes the leading final jet, and m_J is its invariant mass. The partonic channels include $qq \rightarrow qq$, $gg \rightarrow qq$, $gg \rightarrow gg$ and their various crossing ones. The Feynman diagrams at leading order (LO) are shown in appendix A.

It is convenient to introduce two lightlike vectors $n_a^\mu = (1, 0, 0, 1)$ and $n_b^\mu = (1, 0, 0, -1)$ along the beam directions, and another lightlike vector $n_J = (1, \hat{n}_J)$ along the measured jet direction. In the center-of-mass (CM) frame of the initial partons, for the one-jet inclusive production, the momentum of recoiling parton to the observed jet is along the direction $\bar{n}_J = (1, -\hat{n}_J)$. In the CM frame of the hadronic collision, the momenta of the incoming hadrons are given by

$$P_a^\mu = E_{\text{CM}} \frac{n_a^\mu}{2}, \qquad P_b^\mu = E_{\text{CM}} \frac{n_b^\mu}{2}. \tag{2.2}$$

Here E_{CM} is the CM energy of the collider and we have neglected the mass of the hadrons. The momenta of the incoming partons, with a light-cone momentum fraction of the hadronic momenta, are

$$p_a = x_a E_{\text{CM}} \frac{n_a^\mu}{2}, \qquad p_b = x_b E_{\text{CM}} \frac{n_b^\mu}{2}. \tag{2.3}$$

The hadronic kinematic invariants are defined as

$$s = (P_a + P_b)^2 \,, \quad t_1 = (P_a - p_{J_1})^2 - m_{J_1}^2 \,, \quad u_1 = (P_b - p_{J_1})^2 - m_{J_1}^2 \,,$$
$$M_X^2 \equiv P_X^2 = (P_a + P_b - p_{J_1})^2 = s + t_1 + u_1 + m_{J_1}^2 \,, \tag{2.4}$$

and the partonic ones are defined as

$$\hat{s} = (p_a + p_b)^2 = x_a x_b s \,, \quad \hat{t}_1 = (p_a - p_{J_1})^2 - m_{J_1}^2 = x_a t_1 \,, \quad \hat{u}_1 = (p_b - p_{J_1})^2 - m_{J_1}^2 = x_b u_1 \,,$$
$$s_4 \equiv m_X^2 = (p_a + p_b - p_{J_1})^2 = \hat{s} + \hat{t}_1 + \hat{u}_1 + m_{J_1}^2 \,, \tag{2.5}$$

where $p_{J_1}^2 = m_{J_1}^2$. In the threshold limits, we have $p_{J_1}^2 \to 0$ and $s_4 \to 0$. The kinematic region we are interested in is

$$\hat{s}\,,\hat{t}_1\,,\hat{u}_1 \gg m_J^2\,, s_4 \gg \Lambda_{\text{QCD}}^2\,. \tag{2.6}$$

Any four vector can be decomposed along the light-like reference vector n_i

$$p^\mu = (n_i \cdot p)\frac{\bar{n}_i^\mu}{2} + (\bar{n}_i \cdot p)\frac{n_i^\mu}{2} + p_\perp^\mu = p^+ \frac{\bar{n}_i^\mu}{2} + p^- \frac{n_i^\mu}{2} + p_\perp^\mu\,. \tag{2.7}$$

Hence the momentum p^μ can be denoted by $p^\mu = (p^+, p^-, p_\perp)$. The momentum modes relevant to our discussions are the collinear mode $p_{n_J}^\mu \sim \sqrt{\hat{s}}\,(\lambda^2, 1, \lambda)$, anti-collinear mode $p_{\bar{n}_J}^\mu \sim \sqrt{\hat{s}}\,(1, \lambda^2, \lambda)$ and soft mode $p_s^\mu \sim \sqrt{\hat{s}}\,(\lambda^2, \lambda^2, \lambda^2)$, where $\lambda = m_J/\sqrt{\hat{s}}$ is treated as a small expansion parameter. In the partonic threshold limits $m_J \to 0$ and $s_4 \to 0$, the radiation is constrained to be either soft or collinear with the final-state partons.

In order to identify energetic cluster of radiation, the sequential recombination jet algorithms are used. The longitudinal boost invariant distance measures d_{ij} and d_{iB} are defined by

$$d_{ij} = \min\left(p_{T,i}^\alpha, p_{T,j}^\alpha\right) \Delta R_{ij}/R\,, \qquad \Delta R_{ij} = \sqrt{(y_i - y_j)^2 + (\phi_i - \phi_j)^2}\,,$$
$$d_{iB} = p_{T,i}^\alpha\,,$$

where R is the jet radius parameter, y_i and ϕ_i are rapidity and azimuthal angle of the jet i, respectively. $\alpha = -1$, 0 and 1 represent the inclusive anti-k_T [40], Cambridge-Aachen [42, 43] and k_T [44, 45] jet algorithms, respectively. The effects of jet algorithms on the resummation have been studied in refs. [41, 46–49], among which ref. [41] has shown that jet boundary can be changed significantly by boundary clustering for Cambridge-Aachen and k_T algorithms, while the change of the phase space is power suppressed for anti-k_T algorithm. In this paper, the anti-k_T algorithm is adopted, and the jet boundary is just a circle of radius R in $\phi - y$ plane around the jet direction.

After clustering jets, the jet invariant mass m_J receives contribution from the radiation inside the jet, whether from collinear and soft gluons. Thus we split the soft radiation k^μ to two parts, denoted by $k^\mu = k_{\text{in}}^\mu + k_{\text{out}}^\mu$. Then, the partonic threshold variables take the form

$$m_J^2 = (p_{J_1} + k_{\text{in}})^2 = m_{J_1}^2 + 2k_{\text{in}} \cdot p_{J_1}\,,$$
$$s_4 = (p_{J_2} + k_{\text{out}})^2 = m_{J_2}^2 + 2k_{\text{out}} \cdot p_{J_2}\,. \tag{2.8}$$

In the kinematic region $m_J^2, s_4 \ll \hat{s}$, the momenta of the two jets can be written as $p_{J_1}^\mu = E_{J_1} n_J^\mu$ and $p_{J_2}^\mu = E_{J_2} \bar{n}_J^\mu$ in the partonic CM frame, where $E_{J_1} = E_{J_2} = \sqrt{\hat{s}}/2$ in the threshold limit. And m_J and s_4 can be rewritten as

$$m_J^2 = m_{J_1}^2 + 2E_J(n_J \cdot k_{\text{in}})\,,$$
$$s_4 = m_{J_2}^2 + 2E_J(\bar{n}_J \cdot k_{\text{out}})\,. \tag{2.9}$$

For later convenience, we write $k_{\text{in}} \equiv n_J \cdot k_{\text{in}}$ and $k_{\text{out}} \equiv \bar{n}_J \cdot k_{\text{out}}$.

The hadronic threshold is defined as $M_X^2 \to 0$. In this limit, the final state radiations and beam remnants are highly suppressed, which leads to final states consisting of two narrow jets, as well as the remaining soft radiations. For convenience, we introduce the dimensionless variables

$$v = 1 + \frac{\hat{t}_1}{\hat{s}}, \quad w = -\frac{\hat{u}_1}{\hat{s} + \hat{t}_1}, \quad \overline{v} = 1 - v. \quad (2.10)$$

In terms of m_X, x_1, x_2 and v,

$$M_X^2 = \frac{m_X^2}{x_2} + E_{\rm CM}^2 \left[(1 - x_1)v + (1 - x_2)\overline{v} \right] + m_J^2. \quad (2.11)$$

In the limit $x_1 \to 1$, $x_2 \to 1$, $m_J^2 \to 0$ and $m_X^2 \to 0$, we have

$$M_X^2 = m_X^2 + m_J^2 + \frac{p_T^2}{v\overline{v}} \left[(1 - x_1)v + (1 - x_2)\overline{v} \right] + \dots. \quad (2.12)$$

This expression is helpful when we derive the RG equation of the soft function by using the RG invariance in section 6.

3 Factorization in SCET

To derive a factorization formula for dijet process in SCET, we first have to match the full QCD onto the effective theory [50, 51]. To illustrate the factorization in detail, we consider the process $qq' \to qq'$. The initial partons are labeled by 1 and 2 and the final partons are labeled by 3 and 4, and the relevant operator in QCD is given by [52]

$$\mathcal{O}_{I\Gamma}^{\rm QCD} = \left(\bar{\psi}_4^{a_4} \gamma_\mu \Gamma \psi_2^{a_2} \right) \left(\bar{\psi}_3^{a_3} \gamma^\mu \Gamma' \psi_1^{a_1} \right) (c_I)_{\{a\}}, \quad (3.1)$$

where c_I denotes a 4 order color tensor with color indices a_i, and Γ (Γ') denote the chirality (P_L or P_R). In SCET, the n-collinear quark field ψ_n can be written as

$$\chi_n(x) = W_n^\dagger(x)\xi_n(x), \quad \xi_n(x) = \frac{\not{n}\not{\bar{n}}}{4}\psi_n(x), \quad (3.2)$$

where $W_n^\dagger$ is the Wilson line, and χ_n is the gauge invariant combination of $W_n^\dagger$ and collinear quark field ξ_n in SCET. At the leading power in λ, only the $n \cdot A_s$ component of soft gluons can interact with the n-collinear field $\chi_n(x)$, which can be decoupled by a field redefinition [53]:

$$\chi_n(x) \to Y_n(x)\chi_n(x), \quad (3.3)$$

with

$$Y_n(x) = \mathbf{P} \exp \left(ig_s \int_{-\infty}^0 ds\, n \cdot A_s^a(x + sn)t^a \right). \quad (3.4)$$

Then the effective Lagrangian can be expressed as

$$\mathcal{L}_{\rm eff} = \sum_{I,\Gamma} \mathcal{C}_I^\Gamma \, \mathcal{O}_{I\Gamma}^{\rm SCET}, \quad (3.5)$$

with

$$\mathcal{O}_{I\Gamma}^{\mathrm{SCET}} = \sum_{\{a\}} (c_I)_{\{a\}} [O^c(x)]_{\Gamma}^{b_1 b_2 b_3 b_4} [O^s(x)]^{\{a\},\{b\}} , \tag{3.6}$$

$$[O^c(x)]_{\Gamma}^{b_1 b_2 b_3 b_4} = \bar{\chi}_{\bar{n}_J}^{b_4}(x) \gamma_\mu \Gamma \chi_{\bar{n}}^{b_2}(x) \bar{\chi}_{n_J}^{b_3}(x) \gamma^\mu \Gamma' \chi_n^{b_1}(x) , \tag{3.7}$$

$$[O^s(x)]^{\{a\},\{b\}} = \left[Y_{\bar{n}_J}^\dagger(x) \right]^{b_4 a_4} [Y_{\bar{n}}(x)]^{a_2 b_2} \left[Y_{n_J}^\dagger(x) \right]^{b_3 a_3} [Y_n(x)]^{a_1 b_1} . \tag{3.8}$$

Here $\mathcal{C}_I^\Gamma$ is the hard matching coefficient. The scattering amplitude for the $qq' \to qq'$ can be written as

$$|\mathcal{M}^\Gamma(x)\rangle = \langle X|O_\Gamma^c(x)\boldsymbol{O}^s(x)|N_1 N_2\rangle |C^\Gamma\rangle , \tag{3.9}$$

where $|C^\Gamma\rangle$ is the vector of Wilson coefficient combination in color basis $|c_I\rangle$, as following

$$|C^\Gamma\rangle = \sum_I \mathcal{C}_I^\Gamma |c_I\rangle . \tag{3.10}$$

For $qq' \to qq'$, the color basis is chosen as

$$|c_1\rangle = t_{i_3,i_1}^c t_{i_4,i_2}^c , \quad |c_2\rangle = \delta_{i_3,i_1} \delta_{i_4,i_2} . \tag{3.11}$$

The differential cross section can be written as

$$\frac{d\sigma}{dp_T dy dm_J^2} = \frac{1}{2s} \sum_X \sum_\Gamma \int d^4x \langle \mathcal{M}^\Gamma(x)|\widehat{\mathcal{M}}\left(m_J^2, p_T, y, R\right)|\mathcal{M}^\Gamma(0)\rangle , \tag{3.12}$$

where the operator $\widehat{\mathcal{M}}(p_T, y, R)$ denotes the measurement in the final state, including the jet algorithm. It acts on the final-state collinear and soft particles with momenta $\{p_c\}, \{k_s\}$ as follows

$$\widehat{\mathcal{M}}\left(m_J^2, p_T, y, R\right)|X_{c+s}\rangle = \mathcal{M}\left(m_J^2, p_T, y, R, p_c, k_s\right)|X_{c+s}\rangle , \tag{3.13}$$

where

$$\mathcal{M}\left(m_J^2, p_T, y, R, \{p_c\}, \{k_s\}\right) = \delta\left((p_c + k_s)^2 - m_J^2\right) \delta\left(|\vec{p}_{cT}| - p_T\right) \delta\left(y - \frac{1}{2}\ln\frac{p_c^+}{p_c^-}\right) \tag{3.14}$$

$$\times \Theta\left(R^2 - (y_s - y_c)^2 - (\phi_s - \phi_c)^2\right) .$$

Since the soft and collinear sectors are decoupled due to field redefinition, the matrix element in eq. (3.12) can be factorized into a product of several matrices,

$$\sum_X \langle \mathcal{M}(x)|\widehat{\mathcal{M}}\left(m_J^2, p_T, y, R\right)|\mathcal{M}(0)\rangle = \frac{1}{N_{\mathrm{init}}} \sum_\Gamma \left(\Gamma \gamma_\nu\right)_{\alpha 1 \gamma 1} \left(\gamma_\mu \Gamma\right)_{\beta 1 \sigma 1} \left(\Gamma' \gamma_\nu\right)_{\alpha 2 \gamma 2} \left(\gamma_\mu \Gamma'\right)_{\beta 2 \sigma 2}$$

$$\times \langle N_1(P_1)|\bar{\chi}_n^{\alpha 1}(x) \chi_n^{\beta 1}(0)|N_1(P_1)\rangle$$

$$\times \langle N_2(P_2)|\bar{\chi}_{\bar{n}}^{\alpha 2}(x) \chi_{\bar{n}}^{\beta 2}(0)|N_2(P_2)\rangle$$

$$\times \sum_{X_{c1}} \langle 0|\chi_{n_J}^{\gamma 1}(x)|X_{c1}\rangle \langle X_{c1}|\bar{\chi}_{n_J}^{\sigma 1}(0)|0\rangle$$

$$\times \sum_{X_{c2}} \langle 0|\chi_{\bar{n}_J}^{\gamma 2}(x)|X_{c2}\rangle \langle X_{c2}|\bar{\chi}_{\bar{n}_J}^{\sigma 2}(0)|0\rangle$$

$$\times \sum_{X_s} \langle C^\Gamma|\langle 0|\boldsymbol{O}^{s\dagger}(x)|X_s\rangle \langle X_s|\boldsymbol{O}^s(0)|0\rangle |C^\Gamma\rangle$$

$$\times \mathcal{M}\left(m_J^2, p_T, y, R, \{p_c\}, \{k_s\}\right) , \tag{3.15}$$

where $N_{\text{init}} = 1/(4N^2)$ denotes the average over the colors and spin of the initial-state partons, and $\alpha 1, \beta 1$, etc, are Dirac indices. The initial state collinear sectors match to the conventional PDFs:

$$\langle N_i(P_i)| \bar{\chi}_i^{\alpha 1}\left(n_i \cdot x \frac{\bar{n}_i^\mu}{2}\right) \chi_i^{\beta 1}(0) |N_i(P_i)\rangle = \frac{1}{2}\bar{n}_i \cdot P_i \left(\frac{\slashed{n}_i}{2}\right)^{\beta 1 \alpha 1} \int_{-1}^{1} d\xi \, f_{q/N_i}(\xi) \, e^{i\,\xi\,(n_i \cdot x)(\bar{n}_i \cdot P_i)/2}\,, \tag{3.16}$$

and the matrix elements of the collinear fields in the final state match to the quark jet function:

$$\sum_{X_{c1}} \langle 0| \chi_{n_i}^{\gamma 1}(x) |X_{c1}\rangle \langle X_{c1}| \bar{\chi}_{n_i}^{\sigma 1}(0) |0\rangle = \left(\frac{\slashed{n}_i}{2}\right)^{\gamma 1 \sigma 1} \int \frac{d^4p}{(2\pi)^3} \theta(p^0) \, (\bar{n}_J \cdot p) \, J_q(p^2) \, e^{-i\,x\,p}\,. \tag{3.17}$$

The soft function can be defined as the matrix element associated with the soft Wilson line

$$\boldsymbol{S}(x,\mu) = \langle 0| \boldsymbol{O}^{s\dagger}(x)|X_s\rangle \langle X_s|\boldsymbol{O}^s(0)|0\rangle\,, \tag{3.18}$$

which can be decomposed in the color basis

$$S_{IJ} \equiv \langle c_I|\boldsymbol{S}|c_J\rangle\,. \tag{3.19}$$

Now the matrix element appearing in eq. (3.15) can be simplified as

$$\langle C^\Gamma| \langle 0| \boldsymbol{O}^{s\dagger}(x)|X_s\rangle \langle X_s|\boldsymbol{O}^s(0)|0\rangle |C^\Gamma\rangle = \sum_{IJ} \mathcal{C}_I^{\Gamma*} \, S_{IJ} \, \mathcal{C}_J^\Gamma\,. \tag{3.20}$$

All the above components in the factorization form in eq. (3.15) satisfy certain RG equations, which we will discuss in the following sections. Combining the different parts together, we get the factorized differential cross section in the threshold limits

$$\frac{d\sigma}{dp_T dy dm_J^2} = \frac{p_T}{8\pi s} \sum_{\substack{i,j \\ \text{channels}}} \int_{x_a^{\min}}^{1} \frac{dx_a}{x_a} \int_{x_b^{\min}}^{1} \frac{dx_b}{x_b} f_{i/N_1}(x_a,\mu_f) \, f_{j/N_2}(x_b,\mu_f) \, C_{ij}(\hat{s},\hat{t}_1,\hat{u}_1,m_J^2,R,\mu_f)\,, \tag{3.21}$$

where C_{ij} is the hard-scattering kernel

$$C_{ij}(\hat{s},\hat{t}_1,\hat{u}_1,m_J^2,R) = \sum_{I,J} \int dm_{J_1}^2 \, dm_{J_2}^2 \, dk_{\text{in}} \, dk_{\text{out}} \, H_{IJ}(\hat{s},\hat{t}_1,\hat{u}_1) \, S_{JI}(k_{\text{in}},k_{\text{out}}) \tag{3.22}$$

$$\times \, J_1\left(m_{J_1}^2\right) J_2\left(m_{J_2}^2\right) \delta\left(m_J^2 - m_{J_1}^2 - 2E_J k_{\text{in}}\right) \delta\left(s_4 - m_{J_2}^2 - 2E_J k_{\text{out}}\right)\,,$$

with

$$H_{IJ} = \sum_\Gamma \mathcal{C}_I^\Gamma \mathcal{C}_J^{\Gamma*}\,. \tag{3.23}$$

And H_{IJ} is the hard function, the details of which are shown in section 4.

For other channels, such as $gg \to qq'$ or $gg \to gg$, the formula of factorization is similar to the process $qq' \to qq'$, except for the different jet functions and PDFs. The definitions of gluon PDF and jet function are given by

$$\langle N_i(P_i)|(-g_{\mu\nu}) \mathcal{A}_{i\perp}^\mu\left(n_i \cdot x\frac{\bar{n}_i^\mu}{2}\right) \mathcal{A}_{i\perp}^\nu(0) |N_i(P_i)\rangle = \int_{-1}^{1} \frac{d\xi}{\xi} f_{g/N_i}(\xi) \, e^{i\,\xi(n_i \cdot x)\,(\bar{n}_i \cdot P_i)/2}\,, \tag{3.24}$$

and

$$\langle 0| \mathcal{A}_{J\perp}^{a\,\mu}(x)\mathcal{A}_{J\perp}^{b\,\nu}(0) |0\rangle = \delta^{ab} \left(-g_\perp^{\mu\nu}\right) g_s^2 \int \frac{d^4p}{(2\pi)^3} \theta(p^0) \, J_g(p^2) \, e^{-i\,x\,p}\,. \tag{3.25}$$

$12 \to 34$	crossing	$12 \to 34$	crossing	$12 \to 34$	crossing	$12 \to 34$	crossing
$qq' \to qq'$	stu	$qq' \to q'q$	sut	$gg \to q\bar{q}$	stu	$gg \to \bar{q}q$	sut
$q\bar{q}' \to q\bar{q}'$	uts	$q\bar{q}' \to \bar{q}'q$	tus	$\bar{q}g \to g\bar{q}$	uts	$\bar{q}g \to \bar{q}g$	tus
$q\bar{q} \to \bar{q}'q'$	tsu	$q\bar{q} \to q'\bar{q}'$	ust	$qg \to qg$	tsu	$qg \to gq$	ust

Table 1. Crossing relations for the 4-quark and $gg \to q\bar{q}$ channels.

4 Hard function

The coefficient $\mathcal{C}_I^\Gamma$ can be obtained by matching the full theory onto SCET. The one loop results for all partonic $2 \to 2$ process in QCD have been available in ref. [52], which are derived in dimensional regularization and the $\overline{\text{MS}}$ renormalization scheme. In this section, we show the crossing relations for different channels and the RG evolution briefly. The explicit expressions of hard matching coefficients are shown in appendix B.

4.1 Wilsons coefficient at NLO

First, for the 4-quark processes, there are six channels if two different flavor quarks are involved (e.g. $ud \to ud$)

$$ qq' \to qq' \,, \quad q\bar{q}' \to q\bar{q}' \,, \quad q\bar{q} \to q'\bar{q}' \,, \quad qq' \to q'q \,, \quad q\bar{q}' \to \bar{q}'q \,, \quad q\bar{q} \to q'\bar{q}' \,. \tag{4.1}$$

The Wilson coefficients for the channel $qq' \to qq'$ are denoted by $\mathcal{C}_I^\Gamma(s, t, u)$ and the others can be obtained by crossing symmetries, as shown in table 1. For example, the Wilson coefficients for the channel $q\bar{q} \to q'\bar{q}'$ are $\mathcal{C}_I^\Gamma(u, s, t)$. Γ in Wilson coefficients denotes the chirality of the incoming and outgoing partons. In general, there are 16 possible chirality amplitudes. Actually, for the channel $qq' \to qq'$, only 4 chirality amplitudes are non-zero. This is because that chiralities of massless particles 1 and 3 (2 and 4) must be the same. We rewrite the Wilson coefficients as $\mathcal{C}_I^{\lambda_1, \lambda_2} \equiv \mathcal{C}_I^{\lambda_1, \lambda_2, \lambda_3, \lambda_4}$ with $\lambda_{1,2} = L$ or R. In addition, since the chirality can be changed by charge conjugation, the only two independent chirality amplitudes for $qq' \to qq'$ are $\mathcal{C}_I^{LL}$ and $\mathcal{C}_I^{LR}$.

If the 4 quarks are identical, there are two additional non-vanishing chirality amplitudes $\mathcal{C}_I^{LRRL}$ and $\mathcal{C}_I^{RLLR}$ because of the contribution of u-channel. The interference between t-channel and u-channel also makes the results different from $qq' \to qq'$ case. The results for $qq \to qq$ can be expressed as

$$\begin{aligned} \mathcal{C}_I^{LLLL} &= \mathcal{C}_I^{RRRR} = \mathcal{C}_I^{LL}(s, t, u) + B_{IJ} \mathcal{C}_J^{LL}(s, u, t) \,, \\ \mathcal{C}_I^{LRLR} &= \mathcal{C}_I^{RLRL} = \mathcal{C}_I^{LR}(s, t, u) \,, \\ \mathcal{C}_I^{LRRL} &= \mathcal{C}_I^{RLLR} = B_{IJ} \mathcal{C}_J^{LR}(s, u, t) \,, \end{aligned} \tag{4.2}$$

where

$$ B_{IJ} = \begin{pmatrix} -\frac{1}{C_A} & 2 \\ -\frac{C_F}{C_A} & \frac{1}{C_A} \end{pmatrix} \,. \tag{4.3}$$

The results of other channel associated with $qq \to qq$ can be obtained by crossing symmetry as shown in table 1.

Next, we consider the Wilson coefficients for $gg \to q\bar{q}$ channel and its crossing. There are six relevant channels

$$gg \to q\bar{q}, \quad qg \to qg, \quad \bar{q}g \to \bar{q}g, \quad gg \to \bar{q}q, \quad qg \to gq, \quad \bar{q}g \to g\bar{q} . \tag{4.4}$$

The Wilson coefficients for the channel $gg \to q\bar{q}$ are denoted by $\mathcal{C}_I^{\lambda_1,\lambda_2,\lambda_3,\lambda_4}(s,t,u)$ and the others can be obtained by crossing symmetries as shown in table 1. According to parity invariance, we have

$$\mathcal{C}_I^{\lambda_1,\lambda_2,\lambda_3,\lambda_4} = \mathcal{C}_I^{-\lambda_1,-\lambda_2,-\lambda_3,-\lambda_4} . \tag{4.5}$$

In addition, $\mathcal{C}_I^{\lambda_1,\lambda_2,\lambda_3,\lambda_4} = 0$ when $\lambda_3 = \lambda_4$. Thus, the Wilson coefficients for $gg \to q\bar{q}$ can be rewritten as $\mathcal{C}_I^{\lambda_1,\lambda_2} \equiv \mathcal{C}_I^{\lambda_1,\lambda_2;+-}$, and there are only 4 independent chirality amplitudes for each color structure, the explicit expressions of which are shown in appendix B.

Finally, we consider the process $gg \to gg$. In ref. [52], the Wilson coefficients are obtained by matching to an overcomplete basis of 9 color structures, though there are only 8 independent color structures. Then, 16 possible helicity amplitudes for each color structures give 144 matching coefficients. Basing on the symmetry, the Wilson coefficients can be expressed concisely as follows

$$\begin{aligned}
\mathcal{C}_I^\Gamma &= 4g_s^2 \mathcal{M}_I^\Gamma \left(1 + \frac{\alpha_s}{4\pi} \mathcal{Q}_I^\Gamma\right) & I &= 1 \cdots 6 & \Gamma &= 1 \cdots 6 \,, \\
\mathcal{C}_I^\Gamma &= 4g_s^2 \frac{\alpha_s}{4\pi} \mathcal{Q}_I^\Gamma & I &= 7,8,9 & \Gamma &= 1 \cdots 6 \,, \\
\mathcal{C}_I^\Gamma &= 4g_s^2 \frac{\alpha_s}{4\pi} \mathcal{Q}_I^\Gamma \,. & I &= 1 \cdots 9 & \Gamma &= 7 \cdots 16 \,.
\end{aligned} \tag{4.6}$$

The explicit expressions of $\mathcal{M}_I^\Gamma$ and $\mathcal{Q}_I^\Gamma$ are listed in appendix B for the convenience of the reader.

4.2 RG evolution of the hard function

The Wilson coefficients $\mathcal{C}_I^\Gamma$ satisfy the RG equation [54–58]

$$\frac{d}{d\ln\mu} \mathcal{C}_I^\Gamma(\mu) = \Gamma_{IJ}^H \mathcal{C}_J^\Gamma(\mu) , \tag{4.7}$$

where Γ_{IJ}^H can be expressed as

$$\Gamma_{IJ}^H(s,t,u,\mu) = \left(\gamma_{\text{cusp}} \frac{c_H}{2} \ln \frac{-t}{\mu^2} + \gamma_H - \frac{\beta(\alpha_s)}{\alpha_s}\right) \delta_{IJ} + \gamma_{\text{cusp}} M_{IJ}(s,t,u) , \tag{4.8}$$

with

$$c_H = n_q C_F + n_g C_A , \tag{4.9}$$

and

$$\gamma_H = n_q \gamma_q + n_g \gamma_g , \tag{4.10}$$

– 10 –

where $\beta(\alpha_s)$ is the QCD beta function, $\gamma_{\rm cusp}$ is the cusp anomalous dimension, and n_q and n_g is the number of external quarks and gluons involved in the process, respectively. $M_{IJ}(s,t,u)$ denotes the color mixing terms, and can be written as

$$\boldsymbol{M} = -\sum_{i \neq j} \frac{\boldsymbol{T}_i \cdot \boldsymbol{T}_j}{2} \left[L(s_{ij}) - L(t) \right] , \tag{4.11}$$

where $s_{12} = s_{34} = \hat{s}$, $s_{13} = s_{24} = \hat{t}$, $s_{14} = s_{23} = \hat{u}$, and $L(x)$ is defined as

$$L(x) = \ln \frac{|x|}{\mu^2} - i\pi\theta(x) . \tag{4.12}$$

The explicit expressions of M_{IJ} for each channel can be found in appendix B. M_{IJ} can be diagonalized with eigenvalues λ_K. For example, for $qq' \to qq'$ channel, we have

$$\left(F \cdot M \cdot F^{-1} \right)_{KK'} = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix} , \tag{4.13}$$

where $F(s,t,u)$ denotes the transform matrix, which can be calculated numerically. The Wilson coefficients in the diagonal basis are denoted by $\hat{\mathcal{C}}_K^\Gamma \equiv F_{KI}\mathcal{C}_I^\Gamma$, which satisfy the RG equation

$$\frac{d}{d\ln\mu}\hat{\mathcal{C}}_I^\Gamma(\mu) = \left[\gamma_{\rm cusp} \frac{c_H}{2} \ln \frac{-t}{\mu^2} + \gamma_H + \gamma_{\rm cusp}\lambda_K - \frac{\beta(\alpha_s)}{\alpha_s} \right] \hat{\mathcal{C}}_I^\Gamma(\mu) . \tag{4.14}$$

The hard function in the diagonal basis is denoted by $\hat{H}_{KK'} \equiv (F \cdot H \cdot F^\dagger)_{KK'}$. With eq. (3.23), the RG equation of the hard function can be obtained,

$$\frac{d}{d\ln\mu}\hat{H}_{KK'}(\mu) = \left[\gamma_{\rm cusp} \left(c_H \ln \left| \frac{\hat{t}_1}{\mu^2} \right| + \lambda_K + \lambda_{K'}^* \right) + 2\gamma_H - \frac{2\beta(\alpha_s)}{\alpha_s} \right] \hat{H}_{KK'}(\mu) . \tag{4.15}$$

Solving the RG equation, we can get the resummed hard function

$$\hat{H}_{KK'}\left(\hat{s},\hat{t},\hat{u},\mu\right) = \frac{\alpha_s(\mu_h)^2}{\alpha_s(\mu)^2} \exp\left[2c_H S(\mu_h,\mu) - 2A_H(\mu_h,\mu) \right]$$
$$\times \exp\left[-A_\Gamma(\mu_h,\mu)\left(\lambda_K\left(\hat{s},\hat{t},\hat{u}\right) + \lambda_{K'}^*\left(\hat{s},\hat{t},\hat{u}\right) + c_H \ln\left| \frac{t}{\mu_h^2} \right| \right) \right] \hat{H}_{KK'}\left(\hat{s},\hat{t},\hat{u},\mu_h\right) , \tag{4.16}$$

where $S(\nu,\mu)$ and $A_\Gamma(\nu,\mu)$ are defined as

$$S(\nu,\mu) = -\int_{\alpha_s(\nu)}^{\alpha_s(\mu)} d\alpha \frac{\gamma_{\rm cusp}(\alpha)}{\beta(\alpha)} \int_{\alpha_s(\nu)}^{\alpha} \frac{d\alpha'}{\beta(\alpha')} ,$$
$$A_\gamma(\nu,\mu) = -\int_{\alpha_s(\nu)}^{\alpha_s(\mu)} d\alpha \frac{\gamma(\alpha)}{\beta(\alpha)} . \tag{4.17}$$

Up to NNLL level, we need three loop $\gamma_{\rm cusp}$ and β function and two loop γ_H, and their explicit expressions are collected in the appendix A of ref. [59].

5 Jet function

The jet functions $J(p^2, \mu)$, defined in eqs. (3.17) and (3.25), describes a collinear quark or gluon with the invariant mass p^2. It is process independent and has been calculated at NLO in ref. [60] and NNLO in refs. [61, 62], respectively. The nonvanishing diagrams contributing to the NLO jet function in Feynman gauge are given in figure 2. The relevant diagrams corresponding to quark jet function are shown in the top row, and the ones corresponding to gluon are shown by the others. The RG evolution of the quark jet function is given by

$$\frac{dJ_q(p^2, \mu)}{d\ln\mu} = \left(-2C_F\gamma_{\text{cusp}}\ln\frac{p^2}{\mu^2} - 2\gamma^{J_q}\right)J_q(p^2, \mu) + 2C_F\gamma_{\text{cusp}}\int_0^{p^2} dq^2\,\frac{J_q(p^2, \mu) - J_q(q^2, \mu)}{p^2 - q^2}.$$

(5.1)

The gluon jet function is the same with C_F replaced by C_A and γ^{J_q} replaced by γ^{J_g}, respectively. This evolution equation can be solved by the Laplace transformation [63, 64]:

$$\widetilde{j}\left(\ln\frac{Q^2}{\mu^2}, \mu\right) = \int_0^\infty dp^2\,\exp\left(-\frac{p^2}{Q^2 e^{\gamma_E}}\right)J(p^2, \mu),$$

(5.2)

which satisfies the the RG equation

$$\frac{d}{d\ln\mu}\widetilde{j}_q\left(\ln\frac{Q^2}{\mu^2}, \mu\right) = \left(-2C_F\gamma_{\text{cusp}}\ln\frac{Q^2}{\mu^2} - 2\gamma^{J_q}\right)\widetilde{j}_q\left(\ln\frac{Q^2}{\mu^2}, \mu\right),$$
$$\frac{d}{d\ln\mu}\widetilde{j}_g\left(\ln\frac{Q^2}{\mu^2}, \mu\right) = \left(-2C_A\gamma_{\text{cusp}}\ln\frac{Q^2}{\mu^2} - 2\gamma^{J_g}\right)\widetilde{j}_g\left(\ln\frac{Q^2}{\mu^2}, \mu\right).$$

(5.3)

Thus the jet function at an arbitrary scale μ is given by

$$J_q(p^2, \mu) = \exp\left[-4C_F S(\mu_j, \mu) + 2A_{J_q}(\mu_j, \mu)\right]\widetilde{j}_q(\partial_{\eta_j}, \mu_j)\frac{1}{p^2}\left(\frac{p^2}{\mu_j^2}\right)^{\eta_{j_q}}\frac{e^{-\gamma_E\eta_j}}{\Gamma(\eta_{j_q})},$$
$$J_g(p^2, \mu) = \exp\left[-4C_A S(\mu_j, \mu) + 2A_{J_g}(\mu_j, \mu)\right]\widetilde{j}_g(\partial_{\eta_j}, \mu_j)\frac{1}{p^2}\left(\frac{p^2}{\mu_j^2}\right)^{\eta_{j_g}}\frac{e^{-\gamma_E\eta_{j_g}}}{\Gamma(\eta_{j_g})},$$

(5.4)

where $\eta_{j_q} = 2C_F A_\Gamma(\mu_j, \mu)$ and $\eta_{j_q} = 2C_A A_\Gamma(\mu_j, \mu)$. The μ-dependent part of the Laplace transformed jet function $\widetilde{j}(L, \mu)$ is determined by the anomalous dimensions of the jet function as in eq. (5.3), while the μ-independent part can be obtained by a fixed-order calculation. At NLO, it is

$$\widetilde{j}(L, \mu) = 1 + \frac{\alpha_s}{4\pi}\left(\frac{\Gamma_0^J}{2}L^2 + \gamma_0^J L + c_1^J\right).$$

(5.5)

When jet algorithm is applied, the phase space for the collinear radiation need to be constrained. For anti-k_T algorithm, the restriction of phase space is given by

$$\Theta_{\text{anti}-k_T} = \Theta\left(\tan^2\frac{R}{2} > \frac{k^+(p^-)^2}{k^-(p^- - k^-)^2}\right).$$

(5.6)

p and k denote the incoming and loop momenta, respectively. In the threshold limit $p^2 \to 0$, the integral region of momentum k can be expressed as

$$\frac{p^2}{p^-\tan^2\frac{R}{2}} < k^- < p^- - \frac{p^2}{p^-\tan^2\frac{R}{2}}.$$
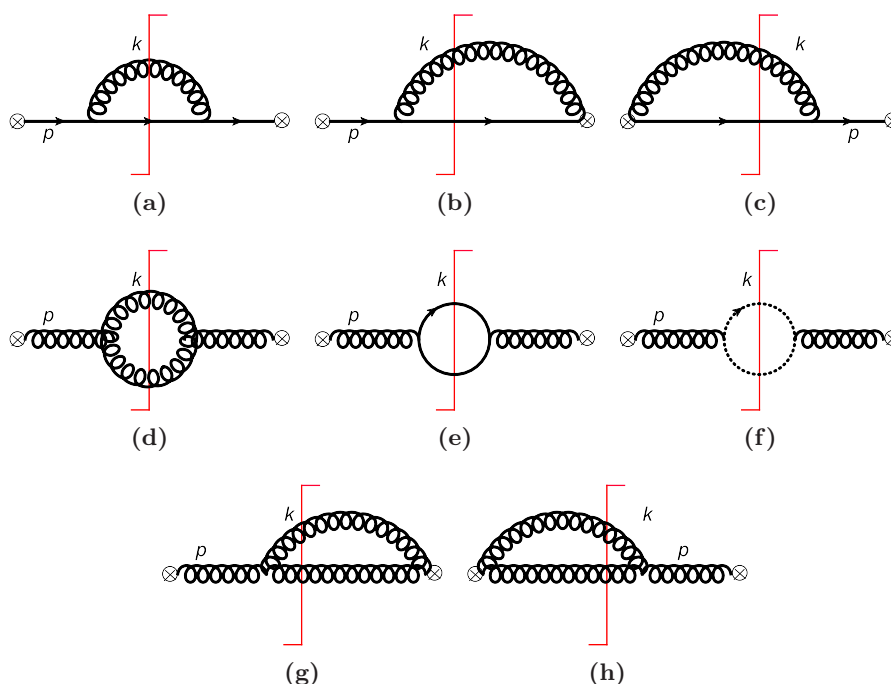
(5.7)

Figure 2. Nonvanishing diagrams contributing to the quark and gluon jet function at NLO.

To avoid the double counting of the soft region covered by soft function, the zero-bin subtraction [65] should be considered. Taking the soft limit of the restriction in eq. (5.6), the phase space of zero-bin region is given by [22, 24]

$$\Theta_{\text{anti}-\text{k}_\text{T}}^{(0)} = \Theta\left(\tan^2\frac{R}{2} > \frac{k^+}{k^-}\right). \tag{5.8}$$

After integrating the phase space and taking zero-bin substraction, the jet functions with anti-k_T algorithm are given by

$$
\begin{aligned}
J_q^{\text{anti}-\text{k}_\text{T}}\left(p^2, p^-, R, \mu\right) &= J_q\left(p^2, \mu\right) + \frac{C_F\alpha_s}{4\pi}\left(\frac{1}{p^2}\right)_\star \frac{2p^2}{(p^-)^2\tan^2\frac{R}{2}}, \\
J_g^{\text{anti}-\text{k}_\text{T}}\left(p^2, p^-, R, \mu\right) &= J_g\left(p^2, \mu\right) + \frac{\alpha_s}{4\pi}\left(\frac{1}{p^2}\right)_\star \frac{p^2}{(p^-)^2\tan^2\frac{R}{2}}(4C_A - 2n_f),
\end{aligned}
\tag{5.9}
$$

where $J_q\left(p^2, \mu\right)$ and $J_g\left(p^2, \mu\right)$ are traditional jet functions. $J_q^{\text{anti}-\text{k}_\text{T}}$ and $J_g^{\text{anti}-\text{k}_\text{T}}$ approach to the traditional ones when jet radius $R \to \infty$, which means that there is no restriction to collinear radiation. In addition, it can be seen that the R-dependent terms in eq. (5.9) are suppressed by m_J^2/p_T^2 at threshold limit, because $p^2 \sim m_J^2$, $p^- \sim 2E_J$ and $m_J \ll E_J$. We have checked numerically that the corrections from R-dependent terms to jet mass spectra are below 1% for $m_J < 100\,\text{GeV}$, so we will use the traditional jet functions in the following numerical calculation.

6 Soft function

The soft function defined in eq. (3.18) describes soft interactions between all colored particles. When calculating the soft function, we need to consider jet algorithm, which imposes a restriction on the phase space of the soft radiation. In ref. [28], the jet was defined as all the radiation in a cone of half-angle R around the jet direction, which is different from the one defined by the standard sequential recombination jet algorithms at hadron colliders. In this work, we choose anti-k_T algorithm to calculate the boost-invariant soft function. In addition, as discussed in ref. [28], there are multiple soft scales in soft function, which need to be refactorized. In this section, we first discuss the calculation of the NLO soft function with jet algorithm for all channels, and then show its refactorization. The details of the calculations can be found in appendix C.

6.1 NLO calculation

As shown in eq. (3.19), the soft function $\mathbf{S}(k_{\text{in}}, k_{\text{out}}, \mu)$ can be decomposed in color space and calculated in the eikonal approximation. Eq. (3.11) has shown the color structures for 4-quark channels. For $gg \to q\bar{q}$ and 4-gluon channels, the color structures are given by

$$|c_1\rangle = (t^{a_1} t^{a_2})_{i_3,i_4}\,, \quad |c_2\rangle = (t^{a_2} t^{a_1})_{i_3,i_4}\,, \quad |c_3\rangle = \delta^{a_1,a_2} \delta_{i_3,i_4}\,, \tag{6.1}$$

and

$$\begin{aligned}
|c_1\rangle &= \text{Tr}\,(t^{a_1} t^{a_2} t^{a_3} t^{a_4})\,, & |c_2\rangle &= \text{Tr}\,(t^{a_1} t^{a_2} t^{a_4} t^{a_3})\,, & |c_3\rangle &= \text{Tr}\,(t^{a_1} t^{a_4} t^{a_3} t^{a_2})\,, & (6.2)\\
|c_4\rangle &= \text{Tr}\,(t^{a_1} t^{a_4} t^{a_2} t^{a_3})\,, & |c_5\rangle &= \text{Tr}\,(t^{a_1} t^{a_3} t^{a_4} t^{a_2})\,, & |c_6\rangle &= \text{Tr}\,(t^{a_1} t^{a_3} t^{a_2} t^{a_4})\,, \\
|c_7\rangle &= \text{Tr}\,(t^{a_1} t^{a_4})\,\text{Tr}\,(t^{a_2} t^{a_3})\,, & |c_8\rangle &= \text{Tr}\,(t^{a_1} t^{a_2})\,\text{Tr}\,(t^{a_3} t^{a_4})\,, & |c_9\rangle &= \text{Tr}\,(t^{a_1} t^{a_3})\,\text{Tr}\,(t^{a_2} t^{a_4})\,,
\end{aligned}$$

respectively. At LO, the soft functions is given by

$$S_{IJ}^{(0)} = \tilde{s}_{IJ}^{(0)} \delta(k_{\text{in}}) \delta(k_{\text{out}})\,. \tag{6.3}$$

At NLO, the soft functions can be expressed as follows

$$S_{IJ}^{(1)}(k_{\text{in}}, k_{\text{out}}, y_J, R, \mu) = \sum_{i,j}^{i \neq j} (w_{ij})_{IJ}\, \mathcal{I}_{ij}(k_{\text{in}}, k_{\text{out}}, y_J, R, \mu)\,, \tag{6.4}$$

where i and j index the massless external partons, while I and J index the color structures. According to eq. (3.19), the color matrix $(w_{ij})_{IJ}$ can be written as

$$(w_{ij})_{IJ} = \langle c_I | \boldsymbol{T}_i \cdot \boldsymbol{T}_j | c_J \rangle\,. \tag{6.5}$$

For $\mathcal{I}_{ij}(k_{\text{in}}, k_{\text{out}}, p_T, y, R, \mu)$, we need to calculate the non-vanishing real emission diagrams in dimension regularization, as shown in figure 3, which is given by

$$\mathcal{I}_{ij}(k_{\text{in}}, k_{\text{out}}, y_J, R, \mu) = -\frac{4\pi\alpha_s}{(2\pi)^{d-1}} \left(\frac{\mu^2 e^{\gamma_E}}{4\pi}\right)^\epsilon \int d^d q\, \delta(q^2)\, \theta(q_0) \mathcal{M}_R(k_{\text{in}}, k_{\text{out}}, R, q) \frac{n_i \cdot n_j}{(n_i \cdot q)(n_j \cdot q)}\,. \tag{6.6}$$
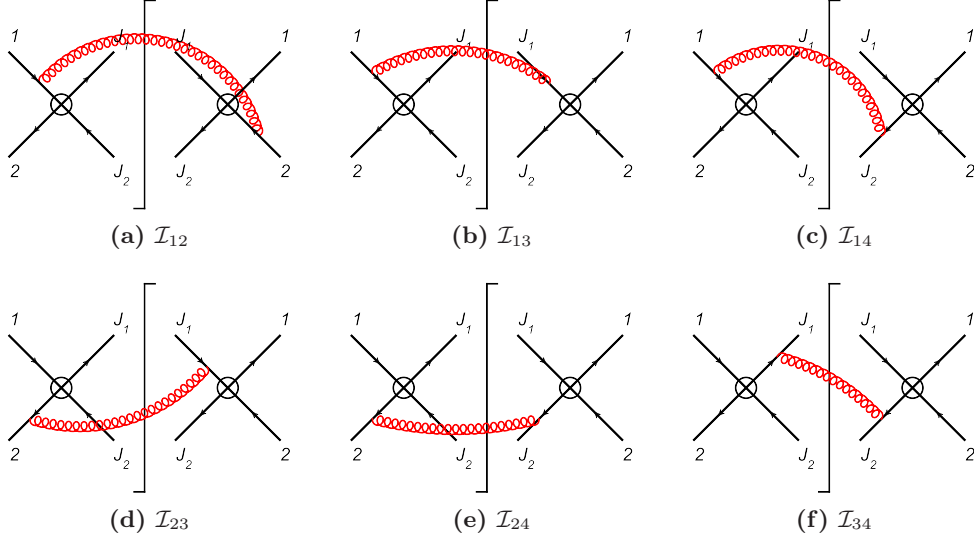
Figure 3. Non-vanishing diagrams contributing to the soft function at NLO.

In the CM of partons, the measurement function $\mathcal{M}_R(k_{\text{in}}, k_{\text{out}}, R, q)$ for anti-k_T algorithm is

$$\mathcal{M}_R(k_{\text{in}}, k_{\text{out}}, R, q) = \Theta\Big(R^2 - (y - y_J)^2 - (\phi - \phi_J)^2\Big)\delta(k_{\text{in}} - n_J \cdot q)$$
$$+ \Theta\Big((y - y_J)^2 + (\phi - \phi_J)^2 - R^2\Big)\delta(k_{\text{out}} - \bar{n}_J \cdot q)\,, \tag{6.7}$$

where y_J and ϕ_J presents the rapidity and azimuthal angle of the measured jet. And y and ϕ are the rapidity and azimuthal angle of the soft gluon with momentum q^μ, respectively. For convenience, we calculate the soft function in the CM frame of initial partons and take ϕ_J to be zero. The results of function $\mathcal{I}_{ij}$ are

$$\mathcal{I}_{12}(k_{\text{in}}, k_{\text{out}}, y_J, R, \mu) = \left(\frac{\alpha_s}{4\pi}\right)\Bigg\{\delta(k_{\text{in}})\delta(k_{\text{out}})\bigg[-4R^2\ln(2\cosh y_J) - 2R^2 + 4R^2\ln R$$
$$- 4\log^2(2\cosh y_J) + \frac{\pi^2}{6}\bigg] + \delta(k_{\text{out}})\left[-2R^2\left(\frac{1}{k_{\text{in}}}\right)_\star\right]$$
$$+ \delta(k_{\text{in}})\left[-8\left[\frac{1}{k_{\text{out}}}\ln\left(\frac{k_{\text{out}}}{\mu}2\cosh y_J\right)\right]_\star + 2R^2\left(\frac{1}{k_{\text{out}}}\right)_\star\right]\Bigg\}$$

$$\mathcal{I}_{13}(k_{\text{in}}, k_{\text{out}}, y_J, R, \mu) = \left(\frac{\alpha_s}{4\pi}\right)\Bigg\{\delta(k_{\text{in}})\delta(k_{\text{out}})\bigg[-R^2\ln(2\cosh y_J) - \frac{R^2}{2} + R^2\ln R$$
$$- 8\ln R\ln(2\cosh y_J) + 4\ln^2 R - 4y_J^2 + 4\ln(2\cosh y_J)\bigg]$$
$$+ \delta(k_{\text{out}})\left[4\left[\frac{1}{k_{\text{in}}}\ln\left(\frac{k_{\text{in}}}{\mu}2\cosh y_J\right)\right]_\star - \left(\frac{R^2}{2} + 4\ln R\right)\left(\frac{1}{k_{\text{in}}}\right)_\star\right]$$
$$+ \delta(k_{\text{in}})\left[-4\left[\frac{1}{k_{\text{out}}}\ln\left(\frac{k_{\text{out}}}{\mu}2\cosh y_J\right)\right]_\star\right.$$
$$+ \left.\left(\frac{R^2}{2} + 4\ln R + 8y_J\right)\left(\frac{1}{k_{\text{out}}}\right)_\star\right]\Bigg\}$$

– 15 –

$$
\begin{aligned}
\mathcal{I}_{14}(k_{\text{in}}, k_{\text{out}}, y_J, R, \mu) = \left(\frac{\alpha_s}{4\pi}\right) &\left\{ \delta(k_{\text{in}})\delta(k_{\text{out}}) \left[\frac{1}{32} R^2 e^{2y_J} \text{sech}^2 y_J \right. \right. \\
&\times \left[-4(R^2+8)\ln(2\cosh y_J) - R^2 + 4(R^2+8)\ln R - 16 \right] \Bigg] \\
&+ \delta(k_{\text{out}}) \left[-\frac{1}{16} e^{2y_J} \text{sech}^2 y_J R^2 (R^2+8) \left(\frac{1}{k_{\text{in}}}\right)_\star \right] \\
&\left. + \delta(k_{\text{in}}) \left[\frac{1}{16} e^{2y_J} \text{sech}^2 y_J R^2 (R^2+8) \left(\frac{1}{k_{\text{out}}}\right)_\star \right] \right\}
\end{aligned}
$$

$$
\begin{aligned}
\mathcal{I}_{34}(k_{\text{in}}, k_{\text{out}}, y_J, R, \mu) = \left(\frac{\alpha_s}{4\pi}\right) &\left\{ \delta(k_{\text{in}})\delta(k_{\text{out}}) \left[-8\ln R \ln(2\cosh y_J) + 4\ln^2 R \right. \right. \\
&+ 4\ln^2(2\cosh y_J) - \frac{\pi^2}{6} \Bigg] \\
&+ \delta(k_{\text{out}}) \left[4\left[\frac{1}{k_{\text{in}}} \ln\left(\frac{k_{\text{in}}}{\mu} 2\cosh y_J\right) \right]_\star - 4\ln R \left(\frac{1}{k_{\text{in}}}\right)_\star \right] \\
&+ \delta(k_{\text{in}}) \left(4\left[\frac{1}{k_{\text{out}}} \ln\left(\frac{k_{\text{out}}}{\mu} 2\cosh y_J\right) \right]_\star \right. \\
&\left. \left. + 2\left[-2\ln R + 4\ln(2\cosh y_J) \right] \left(\frac{1}{k_{\text{out}}}\right)_\star \right) \right\}.
\end{aligned} \tag{6.8}
$$

$\mathcal{I}_{23}$ and $\mathcal{I}_{24}$ can be obtained by the relations

$$
\begin{aligned}
\mathcal{I}_{23}(k_{\text{in}}, k_{\text{out}}, y_J, R, \mu) &= \mathcal{I}_{13}(k_{\text{in}}, k_{\text{out}}, -y_J, R, \mu)\,, \\
\mathcal{I}_{24}(k_{\text{in}}, k_{\text{out}}, y_J, R, \mu) &= \mathcal{I}_{14}(k_{\text{in}}, k_{\text{out}}, -y_J, R, \mu)\,.
\end{aligned} \tag{6.9}
$$

In the calculation, we take the limit that $R \to 0$, but we have kept all the terms up to $O(R^2)$ so that our result can can be used for a wider range of R.

6.2 RG equation of the soft function

Now, we discuss the evolution of the soft function. Its RG equation can be derived by using the RG invariance of the cross section. In the hadronic threshold limit, we have [59]

$$
\begin{aligned}
\frac{d^2\sigma_{q\bar{q}}}{dM_X^2 dy} \propto \int dx_1 \int dx_2 \int dm_{J_1}^2 \int dm_{J_2}^2 \int dk_{\text{in}} \int dk_{\text{out}} \\
\times H_{IJ}(\hat{s}, \hat{t}, \hat{u}, \mu)\, S_{JI}(k_{\text{in}}, k_{\text{out}}, \mu)\, f_{q/N_1}(x_1, \mu) f_{\bar{q}/N_2}(x_2, \mu) J_1(m_{J_1}^2, \mu)\, J_2(m_{J_2}^2, \mu) \\
\times \delta\left[M_X^2 - \left(m_{J_1}^2 + m_{J_2}^2 + 2E_J(k_{\text{in}} + k_{\text{out}}) + \frac{p_T^2}{\bar{v}}(1-x_1) + \frac{p_T^2}{v}(1-x_2) \right) \right].
\end{aligned} \tag{6.10}
$$

To transform the convolution form to a product form, using the Laplace transformation

$$
\frac{d^2\tilde{\sigma}}{dQ^2 dy} = \int_0^\infty dM_X^2 \exp\left(-\frac{M_X^2}{Q^2 e^{\gamma_E}} \right) \frac{d^2\sigma}{dM_X^2 dy}\,, \tag{6.11}
$$

we can obtain

$$
\frac{d^2\tilde{\sigma}}{dQ^2 dy} = H_{IJ}(\hat{s}, \hat{t}, \hat{u}, \mu)\, \tilde{s}_{JI}(\kappa_{\text{in}}, \kappa_{\text{out}}, \mu)\, \tilde{f}_{i_1/N_1}(\tau_1, \mu)\, \tilde{f}_{i_2/N_2}(\tau_2, \mu)\, \tilde{j}_1(Q^2, \mu)\, \tilde{j}_2(Q^2, \mu)\,, \tag{6.12}
$$

– 16 –

where $\widetilde{s}$ is the Laplace transformed soft function

$$\tilde{s}_{IJ}(\kappa_{\text{in}}, \kappa_{\text{out}}, \mu) = \int_0^\infty dk_{\text{in}} \int_0^\infty dk_{\text{out}} \exp\left(-\frac{k_{\text{in}}}{\kappa_{\text{in}}e^{\gamma_E}}\right) \exp\left(-\frac{k_{\text{out}}}{\kappa_{\text{out}}e^{\gamma_E}}\right) S_{IJ}(k_{\text{in}}, k_{\text{out}}, \mu).$$
(6.13)

The RG invariance requires

$$\frac{d}{d\ln\mu}\left[H_{IJ}(\hat{s}, \hat{t}, \hat{u}, \mu)\,\tilde{s}_{JI}(\kappa_{\text{in}}, \kappa_{\text{out}}, \mu)\,\widetilde{f}_{i_1/N_1}(\tau_1, \mu)\,\widetilde{f}_{i_2/N_2}(\tau_2, \mu)\,\widetilde{j}_1(Q^2, \mu)\,\widetilde{j}_2(Q^2, \mu)\right] = 0.$$
(6.14)

And the RG equation of PDF is

$$\frac{d\widetilde{f}_{q/N}(\tau, \mu)}{d\ln\mu} = \left[2C_F\gamma_{\text{cusp}}\ln(\tau) + 2\gamma^{f_q}\right]\widetilde{f}_{q/N}(\tau, \mu),$$
(6.15)

with

$$\tau_1 = Q^2/(-\hat{u}_1), \quad \tau_2 = Q^2/(-\hat{t}_1),$$
(6.16)

for beam N_1 and N_2, respectively. Here

$$\hat{t}_1 = -\frac{p_T^2}{v}, \quad \hat{u}_1 = -\frac{p_T^2}{\bar{v}},$$
(6.17)

in threshold limit $m_{J_1,J_2}^2 \to 0$ and $w \to 1$. The gluon PDF equation is the same with $C_F \to C_A$ and $\gamma^{f_q} \to \gamma^{f_g}$. With eqs. (4.15), (5.3), (6.14) and (6.15), the RG equation of the soft function $\hat{\tilde{s}}_{K'K}$ in the diagonal basis is given by

$$\frac{d}{d\ln\mu}\hat{\tilde{s}}_{K'K}(\kappa_{\text{in}}, \kappa_{\text{out}}, \mu) = \Big\{\gamma_{\text{cusp}}\left[2C_{i_1}L(\hat{u}_1) + (2C_{i_2} - c_H)L(\hat{t}_1) - \lambda_K - \lambda_{K'}^*\right]$$
$$- 2\gamma_{\text{cusp}}(C_{i_1} + C_{i_2} - C_{j_1} - C_{j_2})\ln\frac{Q^2}{\mu^2} - 2\gamma^S\Big\}\hat{\tilde{s}}_{K'K},$$
(6.18)

where $\gamma^S = \gamma_H + \gamma^{f_{i_1}} + \gamma^{f_{i_2}} - \gamma^{J_1} - \gamma^{J_2}$, $C_{i,j} = C_F$ and C_A for quark and gluon, respectively. The relation between the soft functions $\hat{\tilde{s}}$ and $\widetilde{s}$ is

$$\hat{\tilde{s}}_{K'K} = \left[(F^{-1})^\dagger \cdot \widetilde{s} \cdot F^{-1}\right]_{K'K}.$$
(6.19)

We have checked that the NLO soft function in eq. (6.4) satisfies the RG equation (6.18), which means our factorization is reasonable.

6.3 Refactorization of the soft function

As shown in eq. (6.4), the soft function depends on two variables k_{in} and k_{out}, which are $k_{\text{in}} \sim m_J^2/p_T$ and $k_{\text{out}} \sim s_4/p_T$, in principle. It means that we should treat the two scales separately to control the convergence of perturbative expansion. However, at two-loop level, a complicated dependence on $k_{\text{in}}/k_{\text{out}}$ will emerge [27, 66], which represents the nonglobal structure of the soft radiation. Although we could not ideally factorize the soft function into separate two pieces which depend only on k_{in}/μ and k_{out}/μ, respectively, we

can at least extract part of the soft function which depend only on a single scale [28]. We define an auxiliary soft function $\boldsymbol{S}^{\text{in}}$ which only depends on k_{in}

$$\boldsymbol{S}^{\text{in}}(k_{\text{in}}) = \langle 0|\boldsymbol{O}^{s\dagger}|X_s^{\text{in}}\rangle\langle X_s^{\text{in}}|\boldsymbol{O}^s(0)|0\rangle\delta\left(k_{\text{in}} - n_J \cdot P_{Xs}^{\text{in}}\right). \tag{6.20}$$

In color basis, the NLO $\boldsymbol{S}^{\text{in}}(k_{\text{in}})$ can be expressed as

$$S_{IJ}^{\text{in}}(k_{\text{in}}) = \sum_{i,j}^{i\neq j} (w_{ij})_{IJ}\, \mathcal{I}_{ij}^{\text{in}}(k_{\text{in}}, y_J, R, \mu)\,, \tag{6.21}$$

where $\mathcal{I}_{ij}^{\text{in}}$ can be computed by the similar integration in eq. (6.6), except for the measurement function replaced by

$$\mathcal{M}_{\text{in}}(k_{\text{in}}, R, q) = \Theta\Big(R^2 - (y - y_J)^2 - (\phi - \phi_J)^2\Big)\delta(k_{\text{in}} - n_J \cdot q)\,. \tag{6.22}$$

Besides, it is necessary to introduce the residual soft function to describe the soft radiation excluded by $\boldsymbol{S}^{\text{in}}(k_{\text{in}})$

$$\boldsymbol{S}^{\text{res}}(k_{\text{in}}, k_{\text{out}}, \mu) = \frac{\boldsymbol{S}(k_{\text{in}}, k_{\text{out}}, \mu)}{\boldsymbol{S}^{\text{in}}(k_{\text{in}}, \mu)}\,. \tag{6.23}$$

At one-loop level, we consider only one soft gluon emission, which is either inside or outside the jet. It means that $\boldsymbol{S}^{\text{res}}$ describe the soft radiation outside the jet, which only depend on k_{out} at $\mathcal{O}(\alpha_s)$, and we rewrite it by notation $\boldsymbol{S}^{\text{out}}(k_{\text{out}})$

$$S_{IJ}^{\text{out}}(k_{\text{out}}) = \sum_{i,j}^{i\neq j} (w_{ij})_{IJ}\, \mathcal{I}_{ij}^{\text{out}}(k_{\text{out}}, y_J, R, \mu)\,, \tag{6.24}$$

where $\mathcal{I}_{ij}^{\text{out}}(k_{\text{out}}, y_J, R, \mu)$ can be calculated according to eq. (6.6), through replacing the measurement function by $\mathcal{M}_{\text{out}}(k_{\text{out}}, R, q)$

$$\mathcal{M}_{\text{out}}(k_{\text{out}}, R, q) = \Theta\Big((y - y_J)^2 + (\phi - \phi_J)^2 - R^2\Big)\delta(k_{\text{out}} - \bar{n}_J \cdot q)\,. \tag{6.25}$$

Now, the soft function at $\mathcal{O}(\alpha_s)$ in diagonal basis reads

$$\hat{S}_{K'K}(k_{\text{in}}, k_{\text{out}}, \mu_{s_{\text{in}}}, \mu_{s_{\text{out}}}, \mu) = \hat{S}_{K'L}^{\text{in}}(k_{\text{in}}, \mu_{s_{\text{in}}}, \mu)\left(\hat{S}^{(0)}\right)_{LM}^{-1}\hat{S}_{MK}^{\text{out}}(k_{\text{out}}, \mu_{s_{\text{out}}}, \mu)\,. \tag{6.26}$$

$\hat{\tilde{s}}^{\text{in}}(L_{\text{in}}, \mu)$ and $\hat{\tilde{s}}^{\text{out}}(L_{\text{out}}, \mu)$ is the Laplace transformation of $\hat{S}^{\text{in}}$ and $\hat{S}^{\text{out}}$, respectively, and their RG equations are

$$\begin{aligned} \frac{d}{d\ln\mu}\hat{\tilde{s}}_{K'L}^{\text{in}}(L_{\text{in}}, \mu) &= \left[-2\widetilde{B}_{K'L}^{\text{in}}\gamma_{\text{cusp}} L_{\text{in}} - \widetilde{C}_{K'L}^{\text{in}}\gamma_{\text{cusp}} - \widetilde{\gamma}_{K'L}^{\text{in}}\right]\hat{\tilde{s}}_{K'L}^{\text{in}}\,, \\ \frac{d}{d\ln\mu}\hat{\tilde{s}}_{MK}^{\text{out}}(L_{\text{out}}, \mu) &= \left[-2\widetilde{B}_{MK}^{\text{out}}\gamma_{\text{cusp}} L_{\text{out}} - \widetilde{C}_{MK}^{\text{out}}\gamma_{\text{cusp}} - \widetilde{\gamma}_{MK}^{\text{out}}\right]\hat{\tilde{s}}_{MK}^{\text{out}}\,, \end{aligned} \tag{6.27}$$

where $\widetilde{\gamma}^{\text{in,out}}$ are anomalous dimensions depending on the jet radius R, which are given at one-loop level in appendix C. Solving the RG equation, we get the resummed soft functions

$\boldsymbol{S}^{\text{in}}$ and $\boldsymbol{S}^{\text{out}}$

$$\hat{S}^{\text{in}}_{K'L}(k_{\text{in}}, \mu_{s_{\text{in}}}, \mu) = \exp\left[-2\widetilde{B}^{\text{in}}_{K'L} S(\mu_{s_{\text{in}}}, \mu) + \widetilde{C}^{\text{in}}_{K'L} A_\Gamma(\mu_{s_{\text{in}}}, \mu) + A_{\widetilde{\gamma}^{\text{in}}_{K'L}}(\mu_{s_{\text{in}}}, \mu)\right]$$
$$\times \hat{\tilde{s}}^{\text{in}}_{K'L}(\partial_{\eta_{\text{in}}}, \mu_{s_{\text{in}}}) \frac{1}{k_{\text{in}}} \left(\frac{k_{\text{in}}}{\mu_{s_{\text{in}}}} \sqrt{\frac{2n_{12}}{n_{1J} n_{2J}}}\right)^{\eta_{\text{in}}} \frac{e^{-\gamma_E \eta_{\text{in}}}}{\Gamma(\eta_{\text{in}})}, \qquad (6.28)$$

$$\hat{S}^{\text{out}}_{MK}(k_{\text{out}}, \mu_{s_{\text{out}}}, \mu) = \exp\left[-2\widetilde{B}^{\text{out}}_{MK} S(\mu_{s_{\text{out}}}, \mu) + \widetilde{C}^{\text{out}}_{MK} A_\Gamma(\mu_{s_{\text{out}}}, \mu) + A_{\widetilde{\gamma}^{\text{out}}_{MK}}(\mu_{s_{\text{out}}}, \mu)\right]$$
$$\times \hat{\tilde{s}}^{\text{out}}_{MK}(\partial_{\eta_{\text{out}}}, \mu_{s_{\text{out}}}) \frac{1}{k_{\text{out}}} \left(\frac{k_{\text{out}}}{\mu_{s_{\text{out}}}} \sqrt{\frac{2n_{12}}{n_{1J} n_{2J}}}\right)^{\eta_{\text{out}}} \frac{e^{-\gamma_E \eta_{\text{out}}}}{\Gamma(\eta_{\text{out}})}, \qquad (6.29)$$

with

$$\eta_{\text{in}} = 2\widetilde{B}_{K'L} A_\Gamma(\mu_{s_{\text{in}}}, \mu),$$
$$\eta_{\text{out}} = 2\widetilde{B}_{MK} A_\Gamma(\mu_{s_{\text{out}}}, \mu).$$

As shown in ref. [28], the above procedure, so-called refactorization, is an approximate factorization, because the residual soft function would depend on both k_{in} and k_{out} beyond $\mathcal{O}(\alpha_s)$. At two-loop level, $\ln^n(k_{\text{in}}/k_{\text{out}})$ would emerge due to the non-global structure, which has been widely studied at the $e^+ e^-$ colliders [27, 30, 66–71], but rarely investigated at hadron colliders with a sequential recombination jet algorithm [33]. A systematical discussion of them requires the complete two-loop results of the soft function with jet algorithms, and is beyond the scope of this paper.

7 RG improved cross section

From eq. (3.22), using eqs. (4.16), (5.4) and (6.28), we can obtain

$$C_{ij}(\hat{s}, \hat{t}, \hat{u}, m_{J_1}^2, \mu) = \sum_{K,K',L,M} \frac{\alpha_s(\mu_h)^2}{\alpha_s(\mu)^2} \exp\left[2c_H S(\mu_h, \mu) - 2A_H(\mu_h, \mu)\right]$$
$$\times \exp\left[-A_\Gamma(\mu_h, \mu)\left(\lambda_K + \lambda_{K'}^* + c_H \ln\left|\frac{\hat{t}}{\mu_h^2}\right|\right)\right] H_{KK'}(\hat{s}, \hat{t}, \hat{u}, \mu_h)$$
$$\times \exp\left[-2\widetilde{B}^{\text{in}}_{K'L} S(\mu_{s_{\text{in}}}, \mu) + \widetilde{C}^{\text{in}}_{K'L} A_\Gamma(\mu_{s_{\text{in}}}, \mu) + A_{\widetilde{\gamma}^{\text{in}}_{K'L}}(\mu_{s_{\text{in}}}, \mu)\right]$$
$$\times \exp\left[-4C_1 S(\mu_{j_1}, \mu) + 2A_{J_1}(\mu_{j_1}, \mu)\right]\left(\frac{\mu_{j_1}^2}{\mu_{s_{\text{in}}} p_T}\right)^{\eta_{\text{in}}}$$
$$\times \widetilde{j}_1(\partial_{\eta_1}, \mu_{j_1}) \hat{\tilde{s}}^{\text{in}}_{K'L}\left(\ln\frac{\mu_{j_1}^2}{\mu_{s_{\text{in}}} p_T} + \partial_{\eta_1}, \mu_{s_{\text{in}}}\right) \frac{1}{m_{J_1}^2}\left(\frac{m_{J_1}^2}{\mu_{j_1}^2}\right)^{\eta_1} \frac{e^{-\gamma_E \eta_1}}{\Gamma(\eta_1)}$$
$$\times \left(\hat{S}^{(0)}\right)^{-1}_{LM} \exp\left[-4C_2 S(\mu_{j_2}, \mu) + 2A_{J_2}(\mu_{j_2}, \mu)\right]\left(\frac{\mu_{j_2}^2}{\mu_{s_{\text{out}}} p_T}\right)^{\eta_{\text{out}}}$$
$$\times \exp\left[-2\widetilde{B}^{\text{out}}_{MK} S(\mu_{s_{\text{out}}}, \mu) + \widetilde{C}^{\text{out}}_{MK} A_\Gamma(\mu_{s_{\text{out}}}, \mu) + A_{\widetilde{\gamma}^{\text{out}}_{MK}}(\mu_{s_{\text{out}}}, \mu)\right]$$
$$\times \widetilde{j}_2(\partial_{\eta_2}, \mu_{j_2}) \hat{\tilde{s}}^{\text{out}}_{MK}\left(\ln\frac{\mu_{j_2}^2}{\mu_{s_{\text{out}}} p_T} + \partial_{\eta_2}, \mu_{s_{\text{out}}}\right) \frac{1}{s_4}\left(\frac{s_4}{\mu_{j_2}^2}\right)^{\eta_2} \frac{e^{-\gamma_E \eta_2}}{\Gamma(\eta_2)}, \quad (7.1)$$

with

$$\eta_1 = \eta_{\text{in}} + \eta_{j_1}, \quad \eta_2 = \eta_{\text{out}} + \eta_{j_2}. \qquad (7.2)$$

And the resummed cross section (3.21) can be written as

$$\frac{d\sigma^{\mathrm{NNLL}_p}}{dp_T dy dm_{J_1}^2} = \frac{p_T}{8\pi s} \sum_{i,j} \int_{\frac{-u_1 - m_{J_1}^2}{s+t_1}}^{1} \frac{dx_1}{x_1} \int_0^{x_1 s + x_1 t_1 + u_1 + m_{J_1}^2} \frac{ds_4}{s_4 - x_1 t_1 - m_{J_1}^2}$$

$$\times f_{i/N_1}(x_1, \mu_f) \, f_{j/N_2}(x_2, \mu_f) \, C_{ij}\big(\hat{s}, p_T, y, m_{J_1}^2, \mu_f\big), \qquad (7.3)$$

where NNLL_p denotes the approximate NNLL resummation, which means that the NGLs are ignored in this paper. Here, we have changed the integration variables from x_2 to s_4, which have relation

$$x_2(s_4) = \frac{s_4 - x_1 t_1 - m_{J_1}^2}{x_1 s + u_1} \,. \qquad (7.4)$$

To give precise predictions, we resum the singular terms $\ln^n\big(m_J^2/p_T^2\big)$ and $\ln^n(s_4/p_T^2)$ in threshold limits to all orders and include the nonsingular terms up to NLO. And the RG improved differential cross section is given by

$$\frac{d\sigma^{\mathrm{NNLL}_p+\mathrm{NLO}}}{dp_T dy dm_{J_1}^2} = \frac{d\sigma^{\mathrm{NLO}}}{dp_T dy dm_{J_1}^2} + f(m_J) \left(\frac{d\sigma^{\mathrm{NNLL}_p}}{dp_T dy dm_{J_1}^2} - \frac{d\sigma^{\mathrm{NNLL}_p}}{dp_T dy dm_{J_1}^2}\bigg|_{\text{expanded to NLO}} \right), \quad (7.5)$$

where

$$f(m_J) = \frac{1}{1 + \big(m_J/m_J^{\mathrm{match}}\big)^i} \qquad (7.6)$$

is the weight function, as defined in refs. [72, 73]. m_J^{match} denotes the scale above which the fixed order calculation is reliable. For small m_J, $f(m_J)$ approximates to one, and σ^{NLO} and $\sigma^{\mathrm{NNLL}_p}|_{\text{expanded}}$ will cancel each other, and the resummation result dominates the cross section. With increasing m_J above m_J^{match}, $f(m_J)$ goes to zero quickly, and the main contributions are from the fixed-order results. When the power index i becomes larger, the translation from the resummation results to the fixed-order ones is faster. In this work, m_J^{match} is chosen at $100\,\mathrm{GeV}$ and i is taken as 4. But the numerical results are not sensitive to the choices of these parameters.

8 Numerical results

In this section, we discuss the numerical results for the jet mass distribution in dijet process at the LHC. Throughout the numerical calculations, we use the MSTW2008 PDF sets [74] and associated strong coupling α_s. In order to compare with Monte Carlo tools, we use PYTHIA8 [17] with its default "Tune 4C" input. FASTJET [75] is used to perform jet clustering, and the anti-k_T algorithm is chosen unless specified otherwise.

8.1 Leading singular spectrum of jet mass

To verify the correctness of the factorizatoin formula, we expand the eq. (3.22) to the leading singular terms (blue dashed line), and compare with the exact NLO results (red solid line), which are obtained from ref. [76]. From figure 4, we can see that the leading singular terms of the jet mass distribution can reproduce the exact NLO jet mass spectrum in small m_J region. As m_J increases, the difference between the leading singular terms and the exact NLO results increase. We find that in both cases of $R = 0.4$ and 1, the expanded results agree with the fixed-order ones. This means that our soft function is applicable for not only small R.
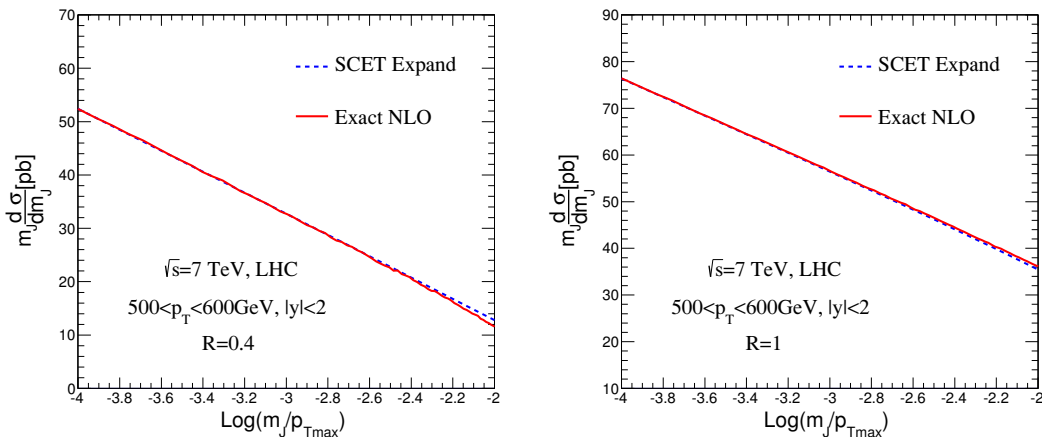
Figure 4. The jet mass distributions from the exact NLO calculation and the resummed results expanded to leading order (SCET Expand). Here, $p_{T\mathrm{max}} = 600\,\mathrm{GeV}$.
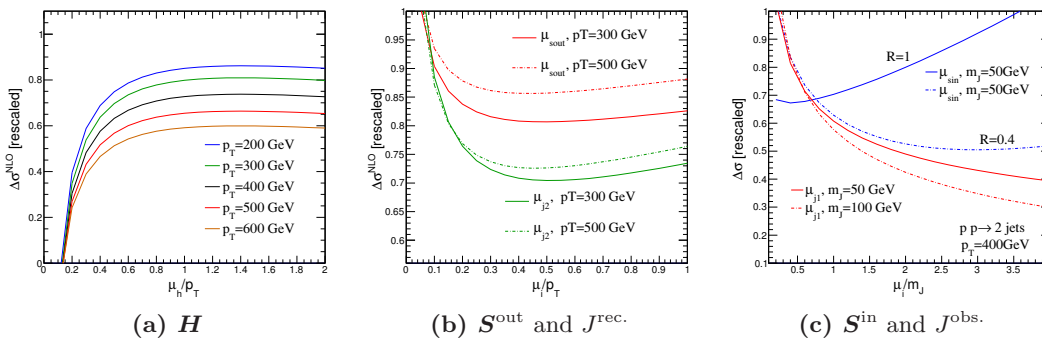

(a) $\boldsymbol{H}$ **(b)** $\boldsymbol{S}^{\mathrm{out}}$ and $J^{\mathrm{rec.}}$ **(c)** $\boldsymbol{S}^{\mathrm{in}}$ and $J^{\mathrm{obs.}}$

Figure 5. The rescaled contribution from hard, soft and jet functions, as a function of their corresponding scales.

8.2 Scale choices and uncertainties

The factorization scales are set at p_T unless specified otherwise. Besides, there are five other matching scales, μ_h, $\mu_{s_{\mathrm{in}}}$, μ_{j_1}, $\mu_{s_{\mathrm{out}}}$ and μ_{j_2}, which need to be chosen properly so that the corresponding hard, soft and jet functions have stable perturbative expansions. The matching scales can be determined by examining the contribution of the NLO matching coefficients as a function of their corresponding scales [59, 77–79]. As shown in figures 5a and 5b, the values of the scale μ_h, $\mu_{s_{\mathrm{out}}}$ and μ_{j_2} are chosen as

$$\mu_h = 1.4\,p_T\,, \quad \mu_{s_{\mathrm{out}}} = 0.2\,p_T + 80\,\mathrm{GeV}\,, \quad \mu_{j_2} = 0.5\,p_T\,, \tag{8.1}$$

where the relevant one-loop contributions get the extreme values.

However, the extreme points of the one-loop contributions of the observed jet function $J_1(\mu_{j_1})$ and soft function $\boldsymbol{S}_{\mathrm{in}}(\mu_{s_{\mathrm{in}}})$ do not exist. It can be seen from their NLO corrections

$$\Delta\sigma^{\mathrm{NLO}} \sim \frac{1}{m_J^2}\left(A\ln\frac{m_J^2}{\mu_{J_1}^2} + B\ln\frac{m_J^2}{p_T\mu_{s_{\mathrm{in}}}} + C\right)\,, \tag{8.2}$$

where A, B, C are scale independent coefficients. If we measure the jet mass m_J, it should not be integrated so that there is no quadratic logarithm term of $\mu_{s_{\mathrm{in}}}$ and μ_{j_1} in the one-loop corrections, which is different from the cases of $\boldsymbol{S}_{\mathrm{out}}(\mu_{s_{\mathrm{out}}})$ and $J_2(\mu_{j_2})$. As shown from the red line in figure 5c, the NLO corrections to J_1 always decrease with increasing μ_{j_1}. For $\mu_{j_1} = 3m_J$, we can see that the corrections decrease slowly. The blue lines in figure 5c show the variations of resummed results as a function of $\mu_{s_{\mathrm{in}}}$ for different jet radius R. The extreme points emerge because higher order contribution of $\boldsymbol{S}^{\mathrm{in}}$ are included. The variations of $\mu_{s_{\mathrm{in}}}$ are minimized at about $20\,\mathrm{GeV}$ and $120\,\mathrm{GeV}$ for $R = 1$ and 0.4, respectively. Using the method in ref. [28], we can determine $\mu_{s_{\mathrm{in}}}$ numerically by the power function of m_J

$$\mu_{s_{\mathrm{in}}} = \frac{\mu_*^2}{c_R} \frac{p_T^*}{p_T}\,, \tag{8.3}$$

where c_R is an R-dependent parameter, $p_T^* = 400\,\mathrm{GeV}$ and $\mu_* = 1.67m_J^{1.47}$ (m_J in GeV) [28]. According to the extreme points of the variations of $\boldsymbol{S}^{\mathrm{in}}$, c_R is numerically determined as 14000 and 2400 for $R = 1$ and $R = 0.4$, respectively.

After all of the natural scales involved in this process have been chosen, we discuss the scale dependence of the resummation results of jet mass spectra in figure 6. At NNLL_p level, three loop cusp anomalous dimension and two loop normal anomalous dimension are used. For the R-dependent pieces, the one-loop soft anomalous dimensions are used. At NLL level, two loop cusp anomalous dimension and one loop normal anomalous dimension are used. Figure 6 shows the scale uncertainties for variation of each scales by a factor of 2 about its default value. It can be seen that the scale uncertainties for μ_h, μ_{j_1}, μ_{j_2} and $\mu_{s_{\mathrm{out}}}$ reduce significantly from NLL to NNLL_p. But for scale $\mu_{s_{\mathrm{in}}}$, the NNLL_p bands are broader than the NLL ones at large m_J region. The reason may be that in large m_J region non-singular terms become important and the resummation results are unreliable. In addition, we can also see that the distribution is enhanced by about 23% from NLL to NNLL_p at the peak region. We confirm numerically that this enhancement mainly comes from the one-loop corrections of the hard function, which are included at NNLL_p order, but not at NLL. This means that if we want to obtain accurate theoretical predictions, the high order corrections of the hard function must be included.

8.3 R dependence

In figure 7a, the blue and red solid lines show the results of NNLL_p resummation for $R = 1$ and 0.4, respectively. We can see that the jet mass spectra shift to right with increasing R, and peak at about $20\,\mathrm{GeV}$ for $R = 0.4$ and $40\,\mathrm{GeV}$ for $R = 1$, respectively. This is due to the fact that when R increases, more large angle soft radiation can be combined into the jet, so that the invariant mass of jet $m_J = \sqrt{(p_c + k_s)^2}$ become larger. The results from PYTHIA are shown as dashed histograms. Figure 7a shows that the peak positions and shapes of our resummation results agree with the ones of PYTHIA at parton level.

8.4 The difference of jet mass spectra between quark and gluon

In order to study the difference between quark jet and gluon jet, we show the jet mass distributions for processes with quark and gluon final state separately. In figure 7b, the blue and red solid lines correspond to $q\bar{q} \rightarrow q\bar{q}$ and $q\bar{q} \rightarrow gg$, respectively. The jet mass
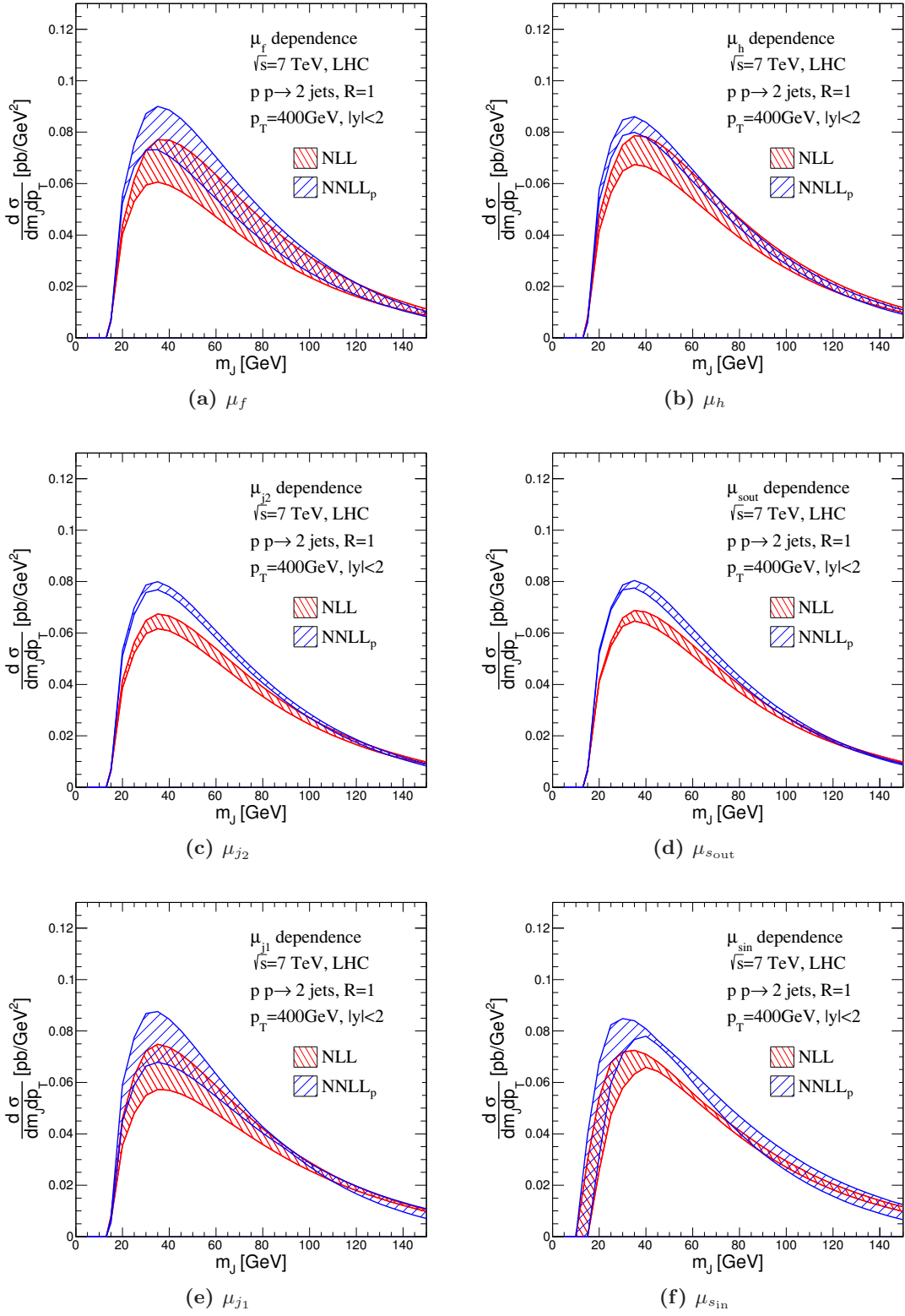
Figure 6. The scale uncertainties of the resummation results for $p_T = 400\text{GeV}$ and $R = 1$.

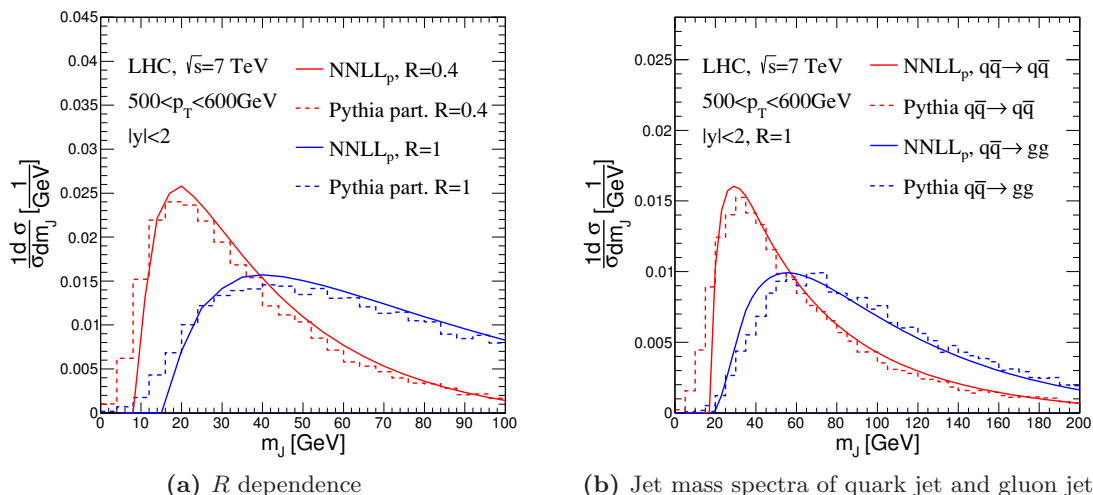**(a)** R dependence   **(b)** Jet mass spectra of quark jet and gluon jet

Figure 7. (a) R dependence of the jet mass distribution. (b) Comparison of jet mass spectrum between quark jet and gluon jet.

spectra for quark and gluon jet peak at about $30\,\text{GeV}$ and $55\,\text{GeV}$, respectively, which is helpful to distinguish between the quark and gluon jet. The peak positions and shapes of our resummation results agree with the ones of PYTHIA.

8.5 Phenomenological studies of jet mass spectrum at the LHC

In this section, we give the RG improved predictions of jet mass spectra at the LHC, and compare them with the results of PYTHIA and the ATLAS data [20]. Figure 8 shows the normalized jet mass distributions with $R = 1$ in four different p_T bins. At $\text{NNLL}_p + \text{NLO}$ level, the jet mass spectra peak around 25–40 GeV, and shift to right with increasing jet p_T. The peak positions agree with the ones of PYTHIA at parton level. In addition, we can see from the results of PYTHIA that the additional hadronization and multiparton interaction shift the spectra to right by about $10\,\text{GeV}$ and $20\,\text{GeV}$, respectively. This means that if we want to obtain predictions which are comparable to data,[1] the non-perturbative effects must be considered. Ref. [80] has computed the non-perturbative corrections to jet mass and their results have been used for $Z+1$ jet process in ref. [33], where a shift $m_J^2 \to m_J^2 + 2\Omega R\, p_T$ for jet mass has been used to account for the non-perturbative effects. However, as discussed in ref. [33], this shift in small jet mass region is not meaningful, so we truncate the spectrum in the left side of the peak. Figure 8 also shows that the $\text{NNLL}_p + \text{NLO}$ results with a shift of $\Omega = 3.0\,\text{GeV}$ (the black solid lines) are consistent with the ATLAS data [20] in all of four p_T bins. Here the shift accounts for the total effects of hadronization and multiparton interaction, so the value of Ω in this work is larger than the one in ref. [33], where only hadronization is concerned. Notice that our treatment of the non-perturbative interaction effects here is just an approximation. A precise estimate of these effects require some modification of the resummation scheme and global fitting with certain precise data. The further discussion of the non-perturbative effects is beyond the scope of this work, and left in future study.

---

[1] Here we have included the multiparton interaction to the non-perturbative effects for simplicity thought it is not necessarily true.
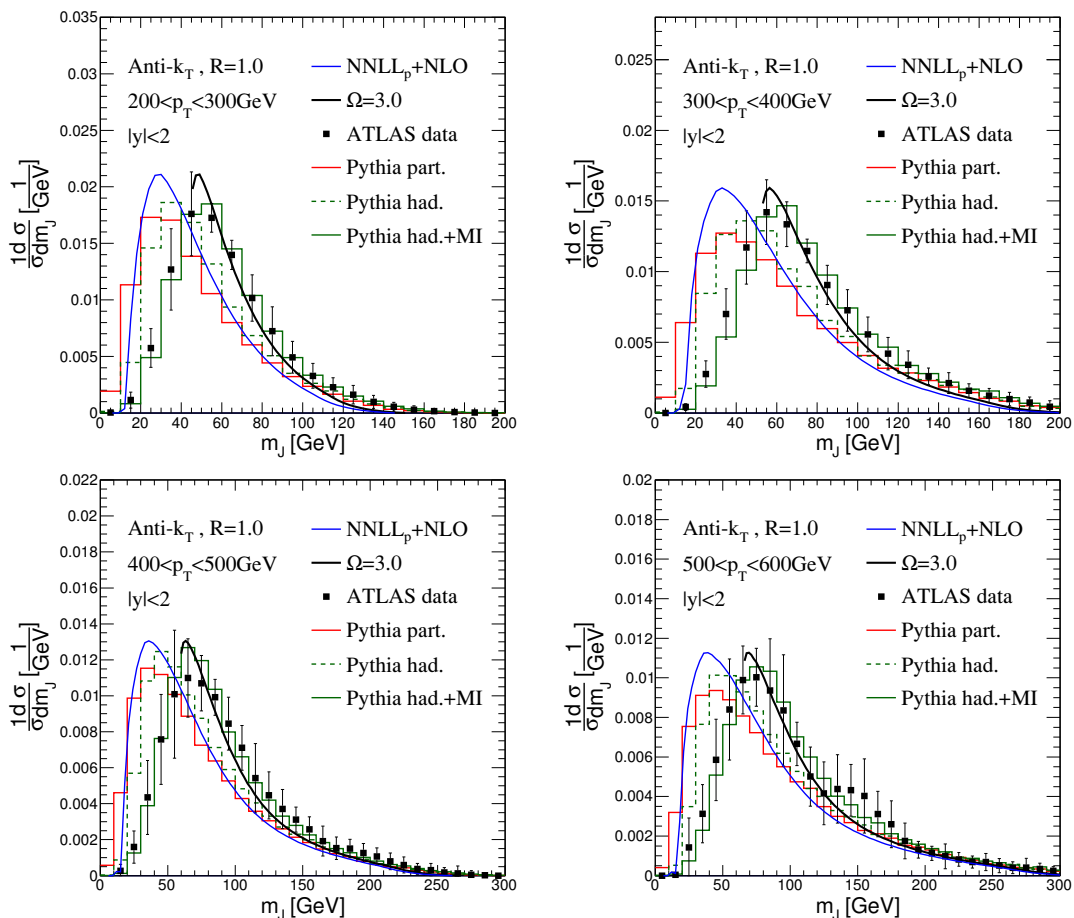
Figure 8. Comparison between theoretical predictions and ATLAS data at the 7 TeV LHC. The label MI in the plots denotes the multiparton interactions. The blue lines represent our NNLL_p + NLO predictions, and the black solid lines represent the results with non-pertubative effects. The red solid, green dashed and green solid histograms represent the results from PYTHIA.

In figure 9, we give our RG improved predictions at the 13 TeV LHC. Comparing with the results at the 7 TeV LHC, the jet mass spectra at parton level in the same kinematic region shift to right by about 5 GeV. The reason is that the dominated contributions is from $qg \to qg$ and $gg \to gg$ channel for 7 TeV and 13 TeV LHC, respectively, and the latter channel gives more gluon final states, the average jet mass of which is larger than the one of quark final states. After including the non-perturbative effects (hadronization and multiparton interaction), the PYTHIA results are closer to the SCET predictions with $\Omega = 4.5\,\text{GeV}$ than $\Omega = 3\,\text{GeV}$, which implies that the non-perturbative effects become more significant at hadron colliders with higher CM energy.

Moreover, we can see this more clearly with the mean values of the jet mass squared, defined as

$$\langle M^2 \rangle \equiv \int m_J^2 \frac{1}{\sigma} \frac{d\sigma}{dm_J} dm_J \,, \tag{8.4}$$

which can be changed by non-perturbative effects in collisions. In table 2, we list the mean jet mass squared at parton level, including hadronization, and both hadronization and
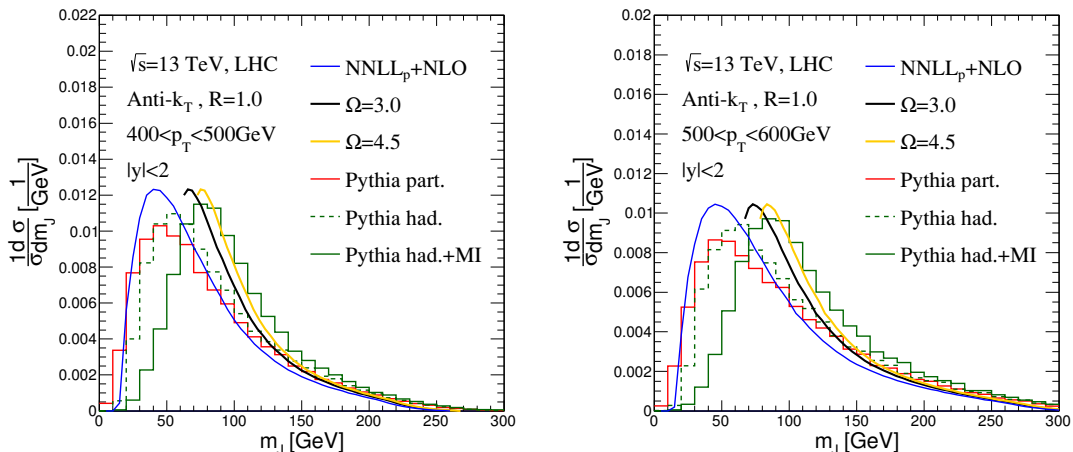
Figure 9. Predictions of jet mass spectrum at 13 TeV run of the LHC.

	$\sqrt{s}$	$\langle M^2_{\text{part.}} \rangle$	$\langle M^2_{\text{had.}} \rangle$	$\langle M^2_{\text{had.+MI}} \rangle$	$\delta M^2_{\text{had.}}$	δM^2_{UE}
$pp \rightarrow$ dijet	7 TeV	7893	8689	10460	796	1771
	13 TeV	9295	10190	12420	895	2230
$q\bar{q} \rightarrow q'\bar{q}'$	7 TeV	4777	5295	6989	518	1694
	13 TeV	5183	5731	8101	548	2370
$gg \rightarrow gg$	7 TeV	11370	12490	14060	1120	1570
	13 TeV	12020	13120	15430	1100	2310

Table 2. The comparison of mean values of jet mass squared from `PYTHIA` at the 7 TeV and 13 TeV LHC. $\delta M^2_{\text{had.}} = \langle M^2_{\text{had.}} \rangle - \langle M^2_{\text{part.}} \rangle$ and $\delta M^2_{\text{UE}} = \langle M^2_{\text{had.+MI}} \rangle - \langle M^2_{\text{had.}} \rangle$. The observed jets are selected with $400 < p_T < 500\,\text{GeV}$ and $|y| < 2$. Unit is GeV^2.

underlying event (described by multiparton interactions in `PYTHIA`), which are denoted by $\langle M^2_{\text{part.}} \rangle$, $\langle M^2_{\text{had.}} \rangle$ and $\langle M^2_{\text{had.+MI}} \rangle$, respectively. We can see that $\delta M^2_{\text{had.}}$ and δM^2_{UE} increase by about 12% and 26%, respectively, with CM energy from 7 TeV to 13 TeV. Besides, $\delta M^2_{\text{had.}}$ in gluon final states is much larger than in quark final states (because of the color factor difference between quark final state and gluon final state [80]), and nearly insensitive to the CM energy. Because the $gg \rightarrow gg$ channel is more dominant at higher CM energy collision, $\delta M^2_{\text{had.}}$ increases apparently in the $pp \rightarrow$ dijet production. In contrast, δM^2_{UE} is almost the same for the quark and gluon final states and sensitive to the CM energy, which increases by about 40%–50% from 7 TeV to 13 TeV with `PYTHIA`. This is just the improvement from $\Omega = 3\,\text{GeV}$ to $\Omega = 4.5\,\text{GeV}$ in our resummation predictions, as shown in figure 9.

9 Conclusion

We have studied the factorization and resummation of jet mass for the one-jet inclusive production at the LHC with SCET. The factorization formula is derived systematically. The NLO soft function with anti-k_T algorithm is calculated and its validity is demonstrated

by checking the agreement between the expanded leading singular terms with the fixed order results. The soft function is refactorized into two pieces corresponding two different scales. The RG invariance of the cross section is checked at NLO for all channels, which demonstrates the correctness of the factorization. By ignoring the NGLs, we first carry out the resummation at approximate NNLL level. From the numerical results, we find that the jet mass spectrum is enhanced by about 23% from NLL to NNLL_p at the peak region. The enhancement mainly comes from one-loop correction of the hard function. The jet mass spectra shift to right with increasing jet radius R and transverse momentum p_T. In addition, we show that there is a significant difference in jet mass spectra between quark and gluon jets. Finally, the normalized jet mass distributions with $R = 1$ are given in four different transverse momentum regions. We show that the $\mathrm{NNLL}_p + \mathrm{NLO}$ spectra peak at 25-40 GeV and shift to right with jet p_T increasing. The peak positions agree with the ones of PYTHIA at parton level. Including the non-perturbative effects, our results are consistent with the ATLAS data. We also give the RG improved predictions at the 13 TeV LHC and find that the peak shift to right by about 5 GeV comparing with the results at the 7 TeV LHC. Our results are helpful to precisely study jet mass spectrum at hadron colliders and test the validity of the Monte Carlo tools.

Acknowledgments

A LO Feynman diagrams

The Feynman diagrams for dijet process at LO are shown in figure 10.

B Explicit expressions of hard Wilson coefficients

All the expression of Wilson coefficients can be found in ref. [52]. We list them below the convenience of the reader.

For $qq' \to qq'$ channels, the expressions of the Wilson coefficients in eq. (4.2) are given by

$$
\mathcal{C}_1^{LL}(s,t,u) = 2g_s^2 \frac{s}{t} \left\{ 1 + \frac{\alpha_s}{4\pi} \left[-2C_F L(t)^2 + X_1(s,t,u)L(t) + Y + \left(\frac{1}{2}C_A - 2C_F \right) Z(s,t,u) \right] \right\},
$$

$$
\mathcal{C}_1^{LR}(s,t,u) = 2g_s^2 \frac{u}{t} \left\{ 1 + \frac{\alpha_s}{4\pi} \left[-2C_F L(t)^2 + X_1(s,t,u)L(t) + Y + (2C_F - C_A) Z(u,t,s) \right] \right\},
$$

$$
\mathcal{C}_2^{LL}(s,t,u) = 2g_s^2 \frac{s}{t} \left\{ \frac{\alpha_s}{4\pi} \left[X_2(s,t,u)L(t) - \frac{C_F}{2C_A} Z(s,t,u) \right] \right\},
$$

$$
\mathcal{C}_2^{LR}(s,t,u) = 2g_s^2 \frac{u}{t} \left\{ \frac{\alpha_s}{4\pi} \left[X_2(s,t,u)L(t) + \frac{C_F}{2C_A} Z(u,t,s) \right] \right\}, \tag{B.1}
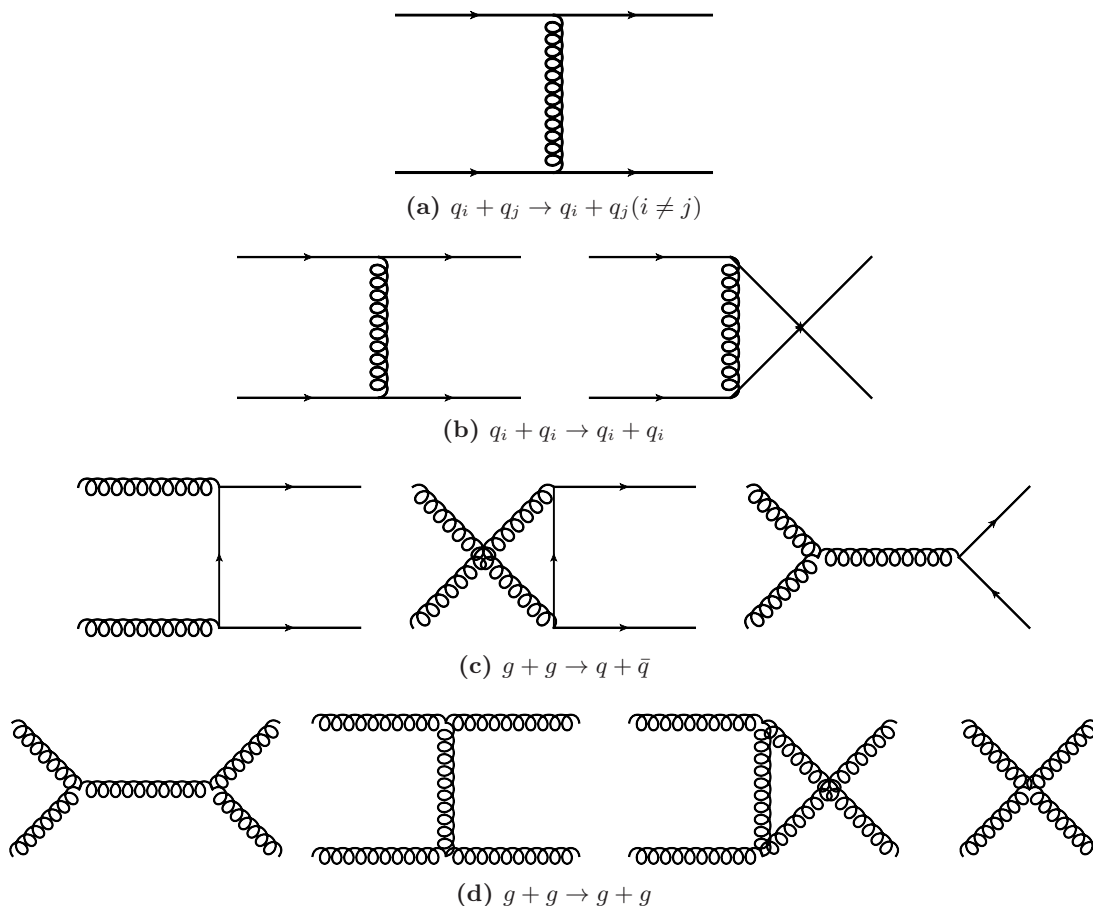$$

(a) $q_i + q_j \to q_i + q_j (i \neq j)$

(b) $q_i + q_i \to q_i + q_i$

(c) $g + g \to q + \bar{q}$

(d) $g + g \to g + g$

Figure 10. Feynman diagrams contributing to the $2 \to 2$ process at leading order.

with

$$X_1(s,t,u) = 6C_F - \beta_0 + 8C_F[L(s) - L(u)] - 2C_A[2L(s) - L(t) - L(u)]$$

$$X_2(s,t,u) = \frac{2C_F}{C_A}[L(s) - L(u)]$$

$$Y = C_A\left(\frac{10}{3} + \pi^2\right) + C_F\left(\frac{\pi^2}{3} - 16\right) + \frac{5}{3}\beta_0 \qquad \text{(B.2)}$$

$$Z(s,t,u) = \frac{t}{s}\left(\frac{t+2u}{s}[L(u) - L(t)]^2 + 2[L(u) - L(t)] + \pi^2\frac{t+2u}{s}\right) .$$

If the 4 quarks are identical, the corresponding Wilson coefficients can be obtained by using eq. (4.2). The other crossed channels, the Wilson coefficients can be obtained by using crossing relations shown in table 1.

For $gg \to q\bar{q}$ channel, the Wilson coefficients are given by

$$\mathcal{C}_1^{-+}(s,t,u) = 2g_s^2 \frac{\sqrt{tu}}{s} \left\{ 1 + \frac{\alpha_s}{4\pi} \left[-(C_A + C_F)L(s)^2 + V_1(s,t,u)L(s) + W_1(s,t,u) \right] \right\}$$

$$\mathcal{C}_1^{+-}(s,t,u) = 2g_s^2 \frac{u}{s}\sqrt{\frac{u}{t}} \left\{ 1 + \frac{\alpha_s}{4\pi} \left[-(C_A + C_F)L(s)^2 + V_1(s,t,u)L(s) + W_2(s,t,u) \right] \right\}$$

$$\mathcal{C}_1^{++}(s,t,u) = \mathcal{C}_1^{--}(s,t,u) = 2g_s^2 \sqrt{\frac{u}{t}} \frac{\alpha_s}{4\pi} W_3(s,t,u)$$

$$\mathcal{C}_2^{-+}(s,t,u) = \mathcal{C}_1^{+-}(s,u,t)$$

$$\mathcal{C}_2^{+-}(s,t,u) = \mathcal{C}_1^{-+}(s,u,t) \tag{B.3}$$

$$\mathcal{C}_2^{++}(s,t,u) = \mathcal{C}_2^{--}(s,t,u) = \mathcal{C}_1^{++}(s,u,t)$$

$$\mathcal{C}_3^{-+}(s,t,u) = 2g_s^2 \sqrt{\frac{t}{u}} \frac{\alpha_s}{4\pi} \left\{ V_2(s,t,u)L(s) + W_4(s,t,u) \right\}$$

$$\mathcal{C}_3^{+-}(s,t,u) = \mathcal{C}_3^{-+}(s,u,t)$$

$$\mathcal{C}_3^{++}(s,t,u) = \mathcal{C}_3^{--}(s,t,u) = 0\,,$$

where

$$W_1(s,t,u) = (C_A - C_F)\frac{s}{u}\left([L(s) - L(t)]^2 + \pi^2 \right) + C_A - 8C_F + \left(\frac{7C_A + C_F}{6} \right)\pi^2$$

$$W_2(s,t,u) = \left(-C_F \frac{s^3}{u^3} - C_A \frac{t^3 + u^3 - s^3}{2u^3} \right)\left([L(s) - L(t)]^2 + \pi^2 \right)$$

$$+ \left(2C_A \frac{ts}{u^2} + C_F \frac{s(2s-u)}{u^2} \right)[L(t) - L(s)] + C_F \frac{t-7u}{u} - C_A \frac{t}{u} + \left(\frac{7C_A + C_F}{6} \right)\pi^2$$

$$W_3(s,t,u) = 2C_F - 2C_A - \frac{2t}{3s}(C_A - n_f)$$

$$W_4(s,t,u) = -\frac{3u}{4t}[L(s) - L(u)]^2 - [L(s) - L(t)][L(s) - L(u)] + \frac{3\pi^2}{2}\frac{u^2}{ts}$$

$$V_1(s,t,u) = 3C_F - 2C_A[L(t) - L(s)]$$

$$V_2(s,t,u) = [L(s) - L(u)] + \frac{t}{s}[L(t) - L(u)]\,. \tag{B.4}$$

For the other crossed channels, the Wilson coefficients can be obtained by using crossing relations shown in table 1.

For 4-gluon channel, the Wilson coefficients can be obtained by eq. (4.6). The LO matching coefficients $\mathcal{M}_I^\Gamma$ can be obtained in table 3. At NLO, we also need $\mathcal{Q}$. They can be expressed in terms of $\mathcal{A}$, $\mathcal{B}$ and $\mathcal{F}$, the expressions of which are

$$\mathcal{A}(s,t,u) = -2C_A L(u)^2 + \left(-2C_A[L(s) - L(u)] + \beta_0 \right)L(u) + \left(\frac{4\pi^2}{3} - \frac{67}{9} \right)C_A + \frac{10}{9}n_f$$

$$\mathcal{B}(s,t,u) = \mathcal{A}(s,t,u) + \beta_0 \frac{u}{t}[L(u) - L(s)] - \frac{3n_f}{2}\frac{su}{t^2}\left([L(u) - L(s)]^2 + \pi^2 \right)$$

$$+ (C_A - n_f)\frac{su}{t^2}\left[\frac{s-u}{t}[L(u) - L(s)] + \left(\frac{su}{t^2} - 2 \right)\left([L(u) - L(s)]^2 + \pi^2 \right) - 1 \right]$$

$$\mathcal{F}(s,t,u) = \frac{1}{C_A}\left(\frac{s^2}{tu}\mathcal{B}(t,s,u) + \frac{s^2}{tu}\mathcal{B}(u,s,t) + \frac{2s}{u}\mathcal{A}(s,t,u) + \frac{2s}{t}\mathcal{A}(s,u,t) \right)\,. \tag{B.5}$$

$\mathcal{M}_I^\Gamma$	$\Gamma = 1,2$	$3,4$	$5,6$		$1,2$	$3,4$	$5,6$
$I = 1$	$\dfrac{s}{u}$	$\dfrac{u}{s}$	$\dfrac{t^2}{su}$	4	$\dfrac{s^2}{tu}$	$\dfrac{u}{t}$	$\dfrac{t}{u}$
2	$\dfrac{s}{t}$	$\dfrac{u^2}{st}$	$\dfrac{t}{s}$	5	$\dfrac{s}{t}$	$\dfrac{u^2}{st}$	$\dfrac{t}{s}$
3	$\dfrac{s}{u}$	$\dfrac{u}{s}$	$\dfrac{t^2}{su}$	6	$\dfrac{s^2}{tu}$	$\dfrac{u}{t}$	$\dfrac{t}{u}$

Table 3. LO matching coefficients $\mathcal{M}_1^\Gamma$ for the 4-gluon channel.

C Calculation of the soft functions

C.1 Color matrix

The color matrix of NLO soft function has been defined in eq. (6.5). At tree level, the color matrix is

$$\tilde{\boldsymbol{s}}^{(0)}_{qq'\to qq'} = \begin{pmatrix} \frac{1}{2}C_A C_F & 0 \\ 0 & C_A^2 \end{pmatrix}. \tag{C.1}$$

The NLO color matrix is

$$\boldsymbol{w}_{12} = \begin{pmatrix} -\frac{C_F}{2} & \frac{C_A C_F}{2} \\ \frac{C_A C_F}{2} & 0 \end{pmatrix}, \qquad \boldsymbol{w}_{13} = \begin{pmatrix} \frac{C_F}{4} & 0 \\ 0 & -C_A^2 C_F \end{pmatrix},$$

$$\boldsymbol{w}_{14} = \begin{pmatrix} \frac{C_F}{2} - \frac{1}{4}C_A^2 C_F & -\frac{1}{2}C_A C_F \\ -\frac{1}{2}C_A C_F & 0 \end{pmatrix},$$

$$\boldsymbol{w}_{23} = \boldsymbol{w}_{14}, \qquad \boldsymbol{w}_{24} = \boldsymbol{w}_{13}, \qquad \boldsymbol{w}_{34} = \boldsymbol{w}_{12}. \tag{C.2}$$

For $gg \to q\bar{q}$ channel, the color matrix at tree level is

$$\tilde{\boldsymbol{s}}^{(0)}_{gg\to q\bar{q}} = \begin{pmatrix} C_A C_F^2 & -\frac{C_F}{2} & C_A C_F \\ -\frac{C_F}{2} & C_A C_F^2 & C_A C_F \\ C_A C_F & C_A C_F & 2C_A^2 C_F \end{pmatrix}. \tag{C.3}$$

The NLO color matrix is

$$\boldsymbol{w}_{12} = \begin{pmatrix} -\frac{1}{4}C_A^3 C_F & 0 & -C_A^2 C_F \\ 0 & -\frac{1}{4}C_A^3 C_F & -C_A^2 C_F \\ -C_A^2 C_F & -C_A^2 C_F & -2C_A^2 C_F C_A \end{pmatrix},$$

$$\boldsymbol{w}_{13} = \begin{pmatrix} \frac{1}{12}C_A^2 C_F(1-3C_A) & \frac{1}{12}C_A^2 C_F & -\frac{1}{2}C_A^2 C_F \\ \frac{1}{12}C_A^2 C_F & \frac{1}{12}C_A^2 C_F & \frac{1}{2}C_A^2 C_F \\ -\frac{1}{2}C_A^2 C_F & \frac{1}{2}C_A^2 C_F & 0 \end{pmatrix},$$

$$\boldsymbol{w}_{14} = \begin{pmatrix} \frac{1}{12}C_A^2 C_F & \frac{1}{12}C_A^2 C_F & \frac{1}{2}C_A^2 C_F \\ \frac{1}{12}C_A^2 C_F & \frac{1}{12}C_A^2 C_F(1-3C_A) & -\frac{1}{2}C_A^2 C_F \\ \frac{1}{2}C_A^2 C_F & -\frac{1}{2}C_A^2 C_F & 0 \end{pmatrix},$$

$$\boldsymbol{w}_{34} = \begin{pmatrix} -\frac{C_F}{12} & -\frac{1}{12}\left(3C_A + 1\right)C_F & -C_A C_F^2 \\ -\frac{1}{12}\left(3C_A + 1\right)C_F & -\frac{C_F}{12} & -C_A C_F^2 \\ -C_A C_F^2 & -C_A C_F^2 & -2C_A^2 C_F^2 \end{pmatrix},$$

$$\boldsymbol{w}_{23} = \boldsymbol{w}_{14}, \qquad \boldsymbol{w}_{24} = \boldsymbol{w}_{14}. \tag{C.4}$$

For $gg \to gg$ channel the color matrix at tree level is

$$\tilde{\boldsymbol{s}}^{(0)}_{gg\to gg} = \frac{C_F}{8C_A}
\begin{pmatrix}
a_0 & b_0 & c_0 & b_0 & b_0 & b_0 & d_0 & d_0 & -e_0 \\
b_0 & a_0 & b_0 & b_0 & c_0 & b_0 & -e_0 & d_0 & d_0 \\
c_0 & b_0 & a_0 & b_0 & b_0 & b_0 & d_0 & d_0 & -e_0 \\
b_0 & b_0 & b_0 & a_0 & b_0 & c_0 & d_0 & -e_0 & d_0 \\
b_0 & c_0 & b_0 & b_0 & a_0 & b_0 & -e_0 & d_0 & d_0 \\
b_0 & b_0 & b_0 & c_0 & b_0 & a_0 & d_0 & -e_0 & d_0 \\
d_0 & -e_0 & d_0 & d_0 & -e_0 & d_0 & d_0 e_0 & e_0^2 & e_0^2 \\
d_0 & d_0 & d_0 & -e_0 & d_0 & -e_0 & e_0^2 & d_0 e_0 & e_0^2 \\
-e_0 & d_0 & -e_0 & d_0 & d_0 & d_0 & e_0^2 & e_0^2 & d_0 e_0
\end{pmatrix} ,
\tag{C.5}$$

with

$$a_0 = C_A^4 - 3C_A^2 + 3 , \quad b_0 = 3 - C_A^2 , \quad c_0 = 3 + C_A^2 , \quad d_0 = 2C_A^2 C_F , \quad e_0 = C_A . \tag{C.6}$$

The NLO color matrix is

$$\boldsymbol{w}_{12} =
\begin{pmatrix}
a & h & c & b & h & b & -f & d & 0 \\
h & a & h & b & c & b & 0 & d & -f \\
c & h & a & b & h & b & -f & d & 0 \\
b & b & b & g & b & g & f & k & f \\
h & c & h & b & a & b & 0 & d & -f \\
b & b & b & g & b & g & f & k & f \\
-f & 0 & -f & f & 0 & f & 0 & -e & e \\
d & d & d & k & d & k & -e & m & -e \\
0 & -f & 0 & f & -f & f & e & -e & 0
\end{pmatrix} , \quad
\boldsymbol{w}_{13} =
\begin{pmatrix}
g & b & g & b & b & b & f & f & k \\
b & a & b & h & c & h & 0 & -f & d \\
g & b & g & b & b & b & f & f & k \\
b & h & b & a & h & c & -f & 0 & d \\
b & c & b & h & a & h & 0 & -f & d \\
b & h & b & c & h & a & -f & 0 & d \\
f & 0 & f & -f & 0 & -f & 0 & e & -e \\
f & -f & f & 0 & -f & 0 & e & 0 & -e \\
k & d & k & d & d & d & -e & -e & m
\end{pmatrix} ,$$

$$\boldsymbol{w}_{14} =
\begin{pmatrix}
a & b & c & h & b & h & d & -f & 0 \\
b & g & b & b & g & b & k & f & f \\
c & b & a & h & b & h & d & -f & 0 \\
h & b & h & a & b & c & d & 0 & -f \\
b & g & b & b & g & b & k & f & f \\
h & b & h & c & b & a & d & 0 & -f \\
d & k & d & d & k & d & m & -e & -e \\
-f & f & -f & 0 & f & 0 & -e & 0 & e \\
0 & f & 0 & -f & f & -f & -e & e & 0
\end{pmatrix} ,$$

$$\boldsymbol{w}_{23} = \boldsymbol{w}_{14} , \quad \boldsymbol{w}_{24} = \boldsymbol{w}_{13} , \quad \boldsymbol{w}_{34} = \boldsymbol{w}_{12} , \tag{C.7}$$

with

$$a = -\frac{1}{16}\left(C_A^4 - 2C_A^2 + 2\right) C_F , \quad b = -\frac{1}{16}\left(2 - C_A^2\right) C_F , \qquad c = -\frac{1}{8}\left(C_A^2 + 1\right) C_F ,$$

$$d = -\frac{1}{4}C_A^2 C_F^2 , \qquad\qquad e = \frac{1}{8}C_A^2 C_F , \quad f = \frac{1}{16}C_A^3 C_F , \quad g = \frac{1}{4}C_A C_F^2 , \tag{C.8}$$

$$h = -\frac{C_F}{8} , \qquad\qquad\qquad k = \frac{C_A C_F}{8} , \qquad\qquad\qquad m = -\frac{1}{4}C_A^3 C_F^2 .$$

C.2 Calculation of $\mathcal{I}_{ij}$

Here, we show the detail of the calculation of the $\mathcal{I}_{ij}$ function. First, in order to compute $\mathcal{I}_{ij}^{\text{out}}$ conveniently, we define an auxiliary function $\mathcal{I}_{ij}^{\text{aux}}(k_{\text{out}})$ with the measurement function $\mathcal{M}_{\text{aux}}(k_{\text{out}}, R, q)$,

$$\mathcal{M}_{\text{aux}}(k_{\text{out}}, R, q) = \Theta\Big(R^2 - (y - y_J)^2 - (\phi - \phi_J)^2\Big)\delta(k_{\text{out}} - \bar{n}_J \cdot q)\,, \tag{C.9}$$

which is the same as $\mathcal{M}_{\text{in}}$ in eq. (6.22) except for the delta function. Then $\mathcal{I}_{ij}^{\text{out}}$ can be obtained by

$$\mathcal{I}_{ij}^{\text{out}}(k_{\text{out}}) = \mathcal{I}_{ij}^{\text{full}}(k_{\text{out}}) - \mathcal{I}_{ij}^{\text{aux}}(k_{\text{out}})\,, \tag{C.10}$$

where $\mathcal{I}_{ij}^{\text{full}}(k_{\text{out}})$ denote the soft radiation without constraints from jet algorithm, the results of which are

$$\mathcal{I}_{12}^{\text{full}}(k_{\text{out}}, \mu) = -\left(\frac{\alpha_s}{4\pi}\right)\left\{\delta(k)\left[\ln^2\frac{2n_{12}}{n_{14}\,n_{24}} - \frac{\pi^2}{6}\right] + 8\left[\frac{1}{k}\ln\left(\frac{k}{\mu}\sqrt{\frac{2n_{12}}{n_{14}\,n_{24}}}\right)\right]_\star\right\}\,,$$

$$\mathcal{I}_{13}^{\text{full}}(k_{\text{out}}, \mu) = -\left(\frac{\alpha_s}{4\pi}\right)\left\{\delta(k)\left[\ln^2\frac{2n_{13}}{n_{14}\,n_{34}} - \frac{\pi^2}{6}\right] + 8\left[\frac{1}{k}\ln\left(\frac{k}{\mu}\sqrt{\frac{2n_{13}}{n_{14}\,n_{34}}}\right)\right]_\star\right\}\,, \tag{C.11}$$

$$\mathcal{I}_{23}^{\text{full}}(k_{\text{out}}, \mu) = -\left(\frac{\alpha_s}{4\pi}\right)\left\{\delta(k)\left[\ln^2\frac{2n_{23}}{n_{24}\,n_{34}} - \frac{\pi^2}{6}\right] + 8\left[\frac{1}{k}\ln\left(\frac{k}{\mu}\sqrt{\frac{2n_{23}}{n_{24}\,n_{34}}}\right)\right]_\star\right\}\,,$$

$$\mathcal{I}_{14}^{\text{full}}(k_{\text{out}}, \mu) = \mathcal{I}_{24}^{\text{full}}(k_{\text{out}}, \mu) = \mathcal{I}_{34}^{\text{full}}(k_{\text{out}}, \mu) = 0\,,$$

where $n_{ij} = n_i \cdot n_j$.

In partonic CM frame, the four vectors of initial and final partons can be written as

$$\begin{aligned}
n_1^\mu &= (1, 0, 0, 1)\,, \\
n_2^\mu &= (1, 0, 0, -1)\,, \\
p_{J_1}^\mu &= p_T(\cosh y_J, 0, 1, \sinh y_J)\,, \\
p_{J_2}^\mu &= p_T(\cosh y_J, 0, -1, -\sinh y_J)\,, \\
q^\mu &= q_T(\cosh y, \sin\phi, \cos\phi, \sinh y)\,.
\end{aligned} \tag{C.12}$$

This choice of frame makes the measurement functions simple but leaves the complexity in delta function. The phase space integration can be written as

$$\int d^d q\, \delta(q^2)\, \Theta(q^0) = \frac{\pi^{\frac{1}{2}-\epsilon}}{\Gamma\left(\frac{1}{2} - \epsilon\right)} \int_0^\pi d\phi\, \sin^{-2\epsilon}\phi \int dy \int dq_T\, q_T^{1-2\epsilon}\,. \tag{C.13}$$

Integrating over the delta function, we can get

$$q_T = \frac{k_{\text{in}} \cosh y_J}{\cosh(y - y_J) - \cos\phi}\,, \tag{C.14}$$

for soft emission inside jet, and

$$q_T = \frac{k_{\text{out}} \cosh y_J}{\cosh(y + y_J) + \cos\phi}\,, \tag{C.15}$$

for the one outside the jet. For $\mathcal{I}_{ij}^{\text{in}}$ and $\mathcal{I}_{ij}^{\text{aux}}$, the integral region of rapidity y and azimuthal angle ϕ are constrained by measurement function is a circle with radius R. We redefine the integration variables

$$y = y' + y_J, \quad y' = r \cos \varphi, \quad \phi = r \sin \varphi \tag{C.16}$$

and then

$$\int dy \int_0^\pi d\phi \, \Theta \left(R^2 - (y - y_J)^2 - \phi^2 \right) = \int_0^R dr \, r \int_0^\pi d\varphi . \tag{C.17}$$

For $\mathcal{I}_{12}^{\text{in}}$, we can get

$$
\begin{aligned}
\mathcal{I}_{12}^{\text{in}}(k_{\text{in}}, y_J, R, \mu) = &-\frac{4\pi\alpha_s}{(2\pi)^{3-2\epsilon}} \left(\frac{e^{\gamma_E}}{4\pi} \right)^\epsilon \frac{2\pi^{\frac{1}{2}-\epsilon}}{\Gamma\left(\frac{1}{2}-\epsilon\right)} \frac{1}{k_{\text{in}}} \left(\frac{k_{\text{in}}}{\mu} \right)^{-2\epsilon} \\
&\times \int_0^R dr \, r \int_0^\pi d\phi \, \sin^{-2\epsilon} \phi \, (\cosh y' - \cos \phi)^{2\epsilon} \cosh^{-2\epsilon} y_J .
\end{aligned}
\tag{C.18}
$$

This integration can be computed analytically by approximation at small R

$$
\begin{aligned}
\sin \phi &\approx \phi = r \sin \varphi , \\
\cosh y' - \cos \phi &\approx \frac{1}{2} y'^2 + \frac{1}{2} \phi^2 = \frac{1}{2} r^2 .
\end{aligned}
\tag{C.19}
$$

From figure 4, we can see that the approximation is validity at even larger R, i.e. $R = 1$. The other I_{ij}^{in} and I_{ij}^{aux} functions can be calculated by similar method.

The results of the refactorized soft function in Laplace space are

$$
\begin{aligned}
\tilde{I}_{12}^{\text{in}}(\kappa_{\text{in}}, y_J, R, \mu) &= \left(\frac{\alpha_s}{4\pi} \right) \left(A_{I_{12}}^{\text{in}} + \gamma_{I_{12}}^{(0)} L_{\text{in}} \right) , \\
\tilde{I}_{12}^{\text{out}}(\kappa_{\text{out}}, y_J, R, \mu) &= \left(\frac{\alpha_s}{4\pi} \right) \left(A_{I_{12}}^{\text{out}} - \gamma_{\text{cusp}}^{(0)} L_{\text{out}}^2 - \gamma_{I_{12}}^{(0)} L_{\text{out}} \right) , \\
\tilde{I}_{13}^{\text{in}}(\kappa_{\text{in}}, y_J, R, \mu) &= \left(\frac{\alpha_s}{4\pi} \right) \left(A_{I_{13}}^{\text{in}} + \frac{1}{2} \gamma_{\text{cusp}}^{(0)} L_{\text{in}}^2 + \gamma_{I_{13}}^{(0)} L_{\text{in}} \right) , \\
\tilde{I}_{13}^{\text{out}}(\kappa_{\text{out}}, y_J, R, \mu) &= \left(\frac{\alpha_s}{4\pi} \right) \left(A_{I_{13}}^{\text{out}} - \frac{1}{2} \gamma_{\text{cusp}}^{(0)} L_{\text{out}}^2 + 2\gamma_{\text{cusp}}^{(0)} y_J L_{\text{out}} - \gamma_{I_{13}}^{(0)} L_{\text{out}} \right) , \\
\tilde{I}_{14}^{\text{in}}(\kappa_{\text{in}}, y_J, R, \mu) &= \left(\frac{\alpha_s}{4\pi} \right) \left(A_{I_{14}}^{\text{in}} + \gamma_{I_{14}}^{(0)} L_{\text{in}} \right) , \\
\tilde{I}_{14}^{\text{out}}(\kappa_{\text{out}}, y_J, R, \mu) &= \left(\frac{\alpha_s}{4\pi} \right) \left(A_{I_{14}}^{\text{out}} - \gamma_{I_{14}}^{(0)} L_{\text{out}} \right) , \\
\tilde{I}_{34}^{\text{in}}(\kappa_{\text{in}}, y_J, R, \mu) &= \left(\frac{\alpha_s}{4\pi} \right) \left(A_{I_{34}}^{\text{in}} + \frac{1}{2} \gamma_{\text{cusp}}^{(0)} L_{\text{in}}^2 + \gamma_{I_{34}}^{(0)} L_{\text{in}} \right) , \\
\tilde{I}_{34}^{\text{out}}(\kappa_{\text{out}}, y_J, R, \mu) &= \left(\frac{\alpha_s}{4\pi} \right) \left[A_{I_{34}}^{\text{out}} + \frac{1}{2} \gamma_{\text{cusp}}^{(0)} L_{\text{out}}^2 - 2\gamma_{\text{cusp}}^{(0)} \ln \left(2 \cosh y_J \right) L_{\text{out}} - \gamma_{I_{34}}^{(0)} L_{\text{out}} \right] , \\
\tilde{I}_{23}^{\text{in}}(\kappa_{\text{in}}, y_J, R, \mu) &= \tilde{I}_{13}^{\text{in}}(\kappa_{\text{in}}, -y_J, R, \mu) , \qquad \tilde{I}_{23}^{\text{out}}(\kappa_{\text{out}}, y_J, R, \mu) = \tilde{I}_{13}^{\text{out}}(\kappa_{\text{out}}, -y_J, R, \mu) , \\
\tilde{I}_{24}^{\text{in}}(\kappa_{\text{in}}, y_J, R, \mu) &= \tilde{I}_{14}^{\text{in}}(\kappa_{\text{in}}, -y_J, R, \mu) , \qquad \tilde{I}_{24}^{\text{out}}(\kappa_{\text{out}}, y_J, R, \mu) = \tilde{I}_{14}^{\text{out}}(\kappa_{\text{out}}, -y_J, R, \mu) ,
\end{aligned}
$$

with $L_{\rm in} = \ln\left(2\kappa_{\rm in}\cosh y_J/\mu\right)$ and $L_{\rm out} = \ln\left(2\kappa_{\rm out}\cosh y_J/\mu\right)$. The one-loop R-dependent anomalous dimensions $\gamma_{I_{ij}}$ are

$$
\gamma_{I_{12}}^{(0)}(y_J, R) = -2\,R^2\,, \qquad\qquad\qquad \gamma_{I_{13}}^{(0)}(y_J, R) = -\frac{R^2}{2} - 4\ln R\,,
$$
$$
\gamma_{I_{14}}^{(0)}(y_J, R) = -\frac{1}{16}R^2\left(R^2 + 8\right)e^{2y_J}\,{\rm sech}^2 y_J\,, \quad \gamma_{I_{34}}^{(0)}(y_J, R) = -4\ln R\,,
$$
$$
\gamma_{I_{23}}^{(0)}(y_J, R) = \gamma_{I_{13}}^{(0)}(-y_J, R)\,, \qquad\qquad \gamma_{I_{24}}^{(0)}(y_J, R) = \gamma_{I_{14}}^{(0)}(-y_J, R)\,, \qquad \text{(C.20)}
$$

and the constant terms $A_{I_{ij}}^{\rm in,out}$ are

$$
A_{I_{12}}^{\rm in}(y_J, R) = R^2\left(-1 + 2\ln R\right)\,,
$$
$$
A_{I_{12}}^{\rm out}(y_J, R) = -4R^2\ln\left(2\cosh y_J\right) + R^2(-1 + 2\ln R) - \frac{\pi^2}{2}\,,
$$
$$
A_{I_{13}}^{\rm in}(y_J, R) = \frac{1}{2}R^2\ln R + 2\ln^2 R + \frac{\pi^2}{4}\,,
$$
$$
A_{I_{13}}^{\rm out}(y_J, R) = \frac{1}{2}\left[R^2 - 16\ln(2\cosh y_J)\right]\ln R - \left(R^2 + 8y_J\right)\ln(2\cosh y_J)
$$
$$
\qquad\qquad\quad - \frac{R^2}{2} + 2\ln^2 R - 4y_J^2 + 4\ln^2(2\cosh y_J) - \frac{\pi^2}{4}\,,
$$
$$
A_{I_{14}}^{\rm in}(y_J, R) = \frac{1}{64}R^2\left[R^2 + 4\left(R^2 + 8\right)\ln R - 16\right]e^{2y_J}\,{\rm sech}^2 y_J\,,
$$
$$
A_{I_{14}}^{\rm out}(y_J, R) = \frac{1}{64}R^2\left[-8\left(R^2 + 8\right)\ln(2\cosh y_J) - 3R^2 + 4\left(R^2 + 8\right)\ln R - 16\right]e^{2y_J}\,{\rm sech}^2 y_J\,,
$$
$$
A_{I_{34}}^{\rm in}(y_J, R) = 2\ln^2 R + \frac{\pi^2}{4}\,,
$$
$$
A_{I_{34}}^{\rm out}(y_J, R) = -8\ln(2\cosh y_J)\ln R + 2\ln^2 R + 8\ln^2(2\cosh y_J) + \frac{\pi^2}{4}\,,
$$
$$
A_{I_{23}}^{\rm in}(y_J, R) = A_{I_{13}}^{\rm in}(-y_J, R)\,, \qquad A_{I_{23}}^{\rm out}(y_J, R) = A_{I_{13}}^{\rm out}(-y_J, R)\,,
$$
$$
A_{I_{24}}^{\rm in}(y_J, R) = A_{I_{14}}^{\rm in}(-y_J, R)\,, \qquad A_{I_{24}}^{\rm out}(y_J, R) = A_{I_{14}}^{\rm out}(-y_J, R)\,.
$$

References

[1] J.M. Butterworth, B.E. Cox and J.R. Forshaw, *WW scattering at the CERN LHC*, *Phys. Rev.* **D 65** (2002) 096014 [`hep-ph/0201098`] [INSPIRE].

[2] J.M. Butterworth, A.R. Davison, M. Rubin and G.P. Salam, *Jet substructure as a new Higgs search channel at the LHC*, *Phys. Rev. Lett.* **100** (2008) 242001 [`arXiv:0802.2470`] [INSPIRE].

[3] D.E. Kaplan, K. Rehermann, M.D. Schwartz and B. Tweedie, *Top Tagging: A Method For Identifying Boosted Hadronically Decaying Top Quarks*, *Phys. Rev. Lett.* **101** (2008) 142001 [`arXiv:0806.0848`] [INSPIRE].

[4] S.D. Ellis, C.K. Vermilion and J.R. Walsh, *Recombination Algorithms and Jet Substructure: Pruning as a Tool for Heavy Particle Searches*, *Phys. Rev.* **D 81** (2010) 094023 [arXiv:0912.0033] [INSPIRE].

[5] J. Thaler and L.-T. Wang, *Strategies to Identify Boosted Tops*, *JHEP* **07** (2008) 092 [arXiv:0806.0023] [INSPIRE].

[6] D. Krohn, J. Thaler and L.-T. Wang, *Jet Trimming*, *JHEP* **02** (2010) 084 [arXiv:0912.1342] [INSPIRE].

[7] J. Gallicchio et al., *Multivariate discrimination and the Higgs + W/Z search*, *JHEP* **04** (2011) 069 [arXiv:1010.3698] [INSPIRE].

[8] J. Thaler and K. Van Tilburg, *Identifying Boosted Objects with N-subjettiness*, *JHEP* **03** (2011) 015 [arXiv:1011.2268] [INSPIRE].

[9] J. Gallicchio and M.D. Schwartz, *Seeing in Color: Jet Superstructure*, *Phys. Rev. Lett.* **105** (2010) 022001 [arXiv:1001.5027] [INSPIRE].

[10] Y. Cui, Z. Han and M.D. Schwartz, *W-jet Tagging: Optimizing the Identification of Boosted Hadronically-Decaying W Bosons*, *Phys. Rev.* **D 83** (2011) 074023 [arXiv:1012.2077] [INSPIRE].

[11] J. Gallicchio and M.D. Schwartz, *Quark and Gluon Tagging at the LHC*, *Phys. Rev. Lett.* **107** (2011) 172001 [arXiv:1106.3076] [INSPIRE].

[12] A. Altheimer et al., *Jet Substructure at the Tevatron and LHC: New results, new tools, new benchmarks*, *J. Phys.* **G 39** (2012) 063001 [arXiv:1201.0008] [INSPIRE].

[13] S.D. Ellis, A. Hornig, T.S. Roy, D. Krohn and M.D. Schwartz, *Qjets: A Non-Deterministic Approach to Tree-Based Jet Substructure*, *Phys. Rev. Lett.* **108** (2012) 182003 [arXiv:1201.1914] [INSPIRE].

[14] T. Gleisberg et al., *SHERPA 1. α: A proof of concept version*, *JHEP* **02** (2004) 056 [hep-ph/0311263] [INSPIRE].

[15] T. Gleisberg et al., *Event generation with SHERPA 1.1*, *JHEP* **02** (2009) 007 [arXiv:0811.4622] [INSPIRE].

[16] T. Sjöstrand, S. Mrenna and P.Z. Skands, *PYTHIA 6.4 Physics and Manual*, *JHEP* **05** (2006) 026 [hep-ph/0603175] [INSPIRE].

[17] T. Sjöstrand, S. Mrenna and P.Z. Skands, *A Brief Introduction to PYTHIA 8.1*, *Comput. Phys. Commun.* **178** (2008) 852 [arXiv:0710.3820] [INSPIRE].

[18] M. Bahr et al., *HERWIG++ Physics and Manual*, *Eur. Phys. J.* **C 58** (2008) 639 [arXiv:0803.0883] [INSPIRE].

[19] S. Gieseke et al., *Herwig++ 2.5 Release Note*, arXiv:1102.1672 [INSPIRE].

[20] ATLAS collaboration, *Jet mass and substructure of inclusive jets in $\sqrt{s} = 7$ TeV pp collisions with the ATLAS experiment*, *JHEP* **05** (2012) 128 [arXiv:1203.4606] [INSPIRE].

[21] T. Becher and M.D. Schwartz, *A precise determination of α_s from LEP thrust data using effective field theory*, *JHEP* **07** (2008) 034 [arXiv:0803.0342] [INSPIRE].

[22] W. M.-Y. Cheung, M. Luke and S. Zuberi, *Phase Space and Jet Definitions in SCET*, *Phys. Rev.* **D 80** (2009) 114021 [arXiv:0910.2479] [INSPIRE].

[23] S.D. Ellis, A. Hornig, C. Lee, C.K. Vermilion and J.R. Walsh, *Consistent Factorization of Jet Observables in Exclusive Multijet Cross-Sections*, *Phys. Lett.* **B 689** (2010) 82 [arXiv:0912.0262] [INSPIRE].

[24] S.D. Ellis, C.K. Vermilion, J.R. Walsh, A. Hornig and C. Lee, *Jet Shapes and Jet Algorithms in SCET*, *JHEP* **11** (2010) 101 [arXiv:1001.0014] [INSPIRE].

[25] T.T. Jouttenus, *Jet Function with a Jet Algorithm in SCET*, *Phys. Rev.* **D 81** (2010) 094017 [arXiv:0912.5509] [INSPIRE].

[26] R. Kelley, M.D. Schwartz and H.X. Zhu, *Resummation of jet mass with and without a jet veto*, arXiv:1102.0561 [INSPIRE].

[27] R. Kelley, M.D. Schwartz, R.M. Schabinger and H.X. Zhu, *Jet Mass with a Jet Veto at Two Loops and the Universality of Non-Global Structure*, *Phys. Rev.* **D 86** (2012) 054017 [arXiv:1112.3343] [INSPIRE].

[28] Y.-T. Chien, R. Kelley, M.D. Schwartz and H.X. Zhu, *Resummation of Jet Mass at Hadron Colliders*, *Phys. Rev.* **D 87** (2013) 014010 [arXiv:1208.0010] [INSPIRE].

[29] Y.-T. Chien and I. Vitev, *Jet Shape Resummation Using Soft-Collinear Effective Theory*, *JHEP* **12** (2014) 061 [arXiv:1405.4293] [INSPIRE].

[30] A. Banfi, M. Dasgupta, K. Khelifa-Kerfa and S. Marzani, *Non-global logarithms and jet algorithms in high-pT jet shapes*, *JHEP* **08** (2010) 064 [arXiv:1004.3483] [INSPIRE].

[31] H.-n. Li, Z. Li and C.-P. Yuan, *QCD resummation for jet substructures*, *Phys. Rev. Lett.* **107** (2011) 152001 [arXiv:1107.4535] [INSPIRE].

[32] H.-n. Li, Z. Li and C.-P. Yuan, *QCD resummation for light-particle jets*, *Phys. Rev.* **D 87** (2013) 074025 [arXiv:1206.1344] [INSPIRE].

[33] M. Dasgupta, K. Khelifa-Kerfa, S. Marzani and M. Spannowsky, *On jet mass distributions in Z+jet and dijet processes at the LHC*, *JHEP* **10** (2012) 126 [arXiv:1207.1640] [INSPIRE].

[34] M. Dasgupta, A. Fregoso, S. Marzani and G.P. Salam, *Towards an understanding of jet substructure*, *JHEP* **09** (2013) 029 [arXiv:1307.0007] [INSPIRE].

[35] T.T. Jouttenus, I.W. Stewart, F.J. Tackmann and W.J. Waalewijn, *Jet mass spectra in Higgs boson plus one jet at next-to-next-to-leading logarithmic order*, *Phys. Rev.* **D 88** (2013) 054031 [arXiv:1302.0846] [INSPIRE].

[36] A. Banfi, G.P. Salam and G. Zanderighi, *Phenomenology of event shapes at hadron colliders*, *JHEP* **06** (2010) 038 [arXiv:1001.4082] [INSPIRE].

[37] A. Banfi, G.P. Salam and G. Zanderighi, *Principles of general final-state resummation and automated implementation*, *JHEP* **03** (2005) 073 [hep-ph/0407286] [INSPIRE].

[38] I.W. Stewart, F.J. Tackmann and W.J. Waalewijn, *N-Jettiness: An Inclusive Event Shape to Veto Jets*, *Phys. Rev. Lett.* **105** (2010) 092002 [arXiv:1004.2489] [INSPIRE].

[39] CMS collaboration, *Studies of jet mass in dijet and W/Z + jet events*, *JHEP* **05** (2013) 090 [arXiv:1303.4811] [INSPIRE].

[40] M. Cacciari, G.P. Salam and G. Soyez, *The anti-k(t) jet clustering algorithm*, *JHEP* **04** (2008) 063 [arXiv:0802.1189] [INSPIRE].

[41] R. Kelley, J.R. Walsh and S. Zuberi, *Abelian Non-Global Logarithms from Soft Gluon Clustering*, *JHEP* **09** (2012) 117 [arXiv:1202.2361] [INSPIRE].

[42] Y.L. Dokshitzer, G.D. Leder, S. Moretti and B.R. Webber, *Better jet clustering algorithms*, *JHEP* **08** (1997) 001 [hep-ph/9707323] [INSPIRE].

[43] M. Wobisch and T. Wengler, *Hadronization corrections to jet cross-sections in deep inelastic scattering*, `hep-ph/9907280` [ɪɴSPIRE].

[44] S. Catani, Y.L. Dokshitzer, M.H. Seymour and B.R. Webber, *Longitudinally invariant K_t clustering algorithms for hadron hadron collisions*, *Nucl. Phys.* **B 406** (1993) 187 [ɪɴSPIRE].

[45] S.D. Ellis and D.E. Soper, *Successive combination jet algorithm for hadron collisions*, *Phys. Rev.* **D 48** (1993) 3160 [`hep-ph/9305266`] [ɪɴSPIRE].

[46] A. Banfi and M. Dasgupta, *Problems in resumming interjet energy flows with k_t clustering*, *Phys. Lett.* **B 628** (2005) 49 [`hep-ph/0508159`] [ɪɴSPIRE].

[47] Y. Delenda, R. Appleby, M. Dasgupta and A. Banfi, *On QCD resummation with k(t) clustering*, *JHEP* **12** (2006) 044 [`hep-ph/0610242`] [ɪɴSPIRE].

[48] K. Khelifa-Kerfa, *Non-global logs and clustering impact on jet mass with a jet veto distribution*, *JHEP* **02** (2012) 072 [`arXiv:1111.2016`] [ɪɴSPIRE].

[49] R. Kelley, J.R. Walsh and S. Zuberi, *Disentangling Clustering Effects in Jet Algorithms*, `arXiv:1203.2923` [ɪɴSPIRE].

[50] C.W. Bauer, A. Hornig and F.J. Tackmann, *Factorization for generic jet production*, *Phys. Rev.* **D 79** (2009) 114013 [`arXiv:0808.2191`] [ɪɴSPIRE].

[51] C.W. Bauer, N.D. Dunn and A. Hornig, *Factorization of Boosted Multijet Processes for Threshold Resummation*, *Phys. Rev.* **D 82** (2010) 054012 [`arXiv:1002.1307`] [ɪɴSPIRE].

[52] R. Kelley and M.D. Schwartz, *1-loop matching and NNLL resummation for all partonic 2 to 2 processes in QCD*, *Phys. Rev.* **D 83** (2011) 045022 [`arXiv:1008.2759`] [ɪɴSPIRE].

[53] C.W. Bauer, D. Pirjol and I.W. Stewart, *Soft collinear factorization in effective field theory*, *Phys. Rev.* **D 65** (2002) 054022 [`hep-ph/0109045`] [ɪɴSPIRE].

[54] T. Becher and M. Neubert, *On the Structure of Infrared Singularities of Gauge-Theory Amplitudes*, *JHEP* **06** (2009) 081 [*Erratum ibid.* **1311** (2013) 024] [`arXiv:0903.1126`] [ɪɴSPIRE].

[55] E. Gardi and L. Magnea, *Factorization constraints for soft anomalous dimensions in QCD scattering amplitudes*, *JHEP* **03** (2009) 079 [`arXiv:0901.1091`] [ɪɴSPIRE].

[56] L.J. Dixon, E. Gardi and L. Magnea, *On soft singularities at three loops and beyond*, *JHEP* **02** (2010) 081 [`arXiv:0910.3653`] [ɪɴSPIRE].

[57] S. Catani, *The Singular behavior of QCD amplitudes at two loop order*, *Phys. Lett.* **B 427** (1998) 161 [`hep-ph/9802439`] [ɪɴSPIRE].

[58] G.F. Sterman and M.E. Tejeda-Yeomans, *Multiloop amplitudes and resummation*, *Phys. Lett.* **B 552** (2003) 48 [`hep-ph/0210130`] [ɪɴSPIRE].

[59] T. Becher and M.D. Schwartz, *Direct photon production with effective field theory*, *JHEP* **02** (2010) 040 [`arXiv:0911.0681`] [ɪɴSPIRE].

[60] A.V. Manohar, *Deep inelastic scattering as $x \to 1$ using soft collinear effective theory*, *Phys. Rev.* **D 68** (2003) 114019 [`hep-ph/0309176`] [ɪɴSPIRE].

[61] T. Becher and M. Neubert, *Toward a NNLO calculation of the $\bar{B} \to X_s\gamma$ decay rate with a cut on photon energy. II. Two-loop result for the jet function*, *Phys. Lett.* **B 637** (2006) 251 [`hep-ph/0603140`] [ɪɴSPIRE].

[62] T. Becher and G. Bell, *The gluon jet function at two-loop order*,
Phys. Lett. **B 695** (2011) 252 [`arXiv:1008.1936`] [ɪɴSPIRE].

[63] T. Becher and M. Neubert, *Threshold resummation in momentum space from effective field theory*, *Phys. Rev. Lett.* **97** (2006) 082001 [`hep-ph/0605050`] [ɪɴSPIRE].

[64] T. Becher, M. Neubert and B.D. Pecjak, *Factorization and Momentum-Space Resummation in Deep-Inelastic Scattering*, *JHEP* **01** (2007) 076 [`hep-ph/0607228`] [ɪɴSPIRE].

[65] A.V. Manohar and I.W. Stewart, *The Zero-Bin and Mode Factorization in Quantum Field Theory*, *Phys. Rev.* **D 76** (2007) 074002 [`hep-ph/0605001`] [ɪɴSPIRE].

[66] R. Kelley, M.D. Schwartz, R.M. Schabinger and H.X. Zhu, *The two-loop hemisphere soft function*, *Phys. Rev.* **D 84** (2011) 045022 [`arXiv:1105.3676`] [ɪɴSPIRE].

[67] K. Khelifa-Kerfa, *Non-global logs and clustering impact on jet mass with a jet veto distribution*, *JHEP* **02** (2012) 072 [`arXiv:1111.2016`] [ɪɴSPIRE].

[68] M. Dasgupta and G.P. Salam, *Resummation of nonglobal QCD observables*,
Phys. Lett. **B 512** (2001) 323 [`hep-ph/0104277`] [ɪɴSPIRE].

[69] M. Dasgupta and G.P. Salam, *Accounting for coherence in interjet E_t flow: A case study*, *JHEP* **03** (2002) 017 [`hep-ph/0203009`] [ɪɴSPIRE].

[70] A. Banfi, G. Marchesini and G. Smye, *Away from jet energy flow*, *JHEP* **08** (2002) 006 [`hep-ph/0206076`] [ɪɴSPIRE].

[71] A. Hornig, C. Lee, I.W. Stewart, J.R. Walsh and S. Zuberi, *Non-global Structure of the $O(\alpha_s^2)$ Dijet Soft Function*, *JHEP* **08** (2011) 054 [`arXiv:1105.4628`] [ɪɴSPIRE].

[72] T. Plehn, *Single stop production at hadron colliders*, *Phys. Lett.* **B 488** (2000) 359 [`hep-ph/0006182`] [ɪɴSPIRE].

[73] T. Han, I. Lewis and T. McElmurry, *QCD Corrections to Scalar Diquark Production at Hadron Colliders*, *JHEP* **01** (2010) 123 [`arXiv:0909.2666`] [ɪɴSPIRE].

[74] A.D. Martin, W.J. Stirling, R.S. Thorne and G. Watt, *Uncertainties on α_s in global PDF analyses and implications for predicted hadronic cross sections*,
Eur. Phys. J. **C 64** (2009) 653 [`arXiv:0905.3531`] [ɪɴSPIRE].

[75] M. Cacciari, G.P. Salam and G. Soyez, *FastJet User Manual*,
Eur. Phys. J. **C 72** (2012) 1896 [`arXiv:1111.6097`] [ɪɴSPIRE].

[76] R.K. Ellis and J.C. Sexton, *QCD Radiative Corrections to Parton Parton Scattering*,
Nucl. Phys. **B 269** (1986) 445 [ɪɴSPIRE].

[77] T. Becher, M. Neubert and G. Xu, *Dynamical Threshold Enhancement and Resummation in Drell-Yan Production*, *JHEP* **07** (2008) 030 [`arXiv:0710.0680`] [ɪɴSPIRE].

[78] T. Becher, C. Lorentzen and M.D. Schwartz, *Resummation for W and Z production at large pT*, *Phys. Rev. Lett.* **108** (2012) 012001 [`arXiv:1106.4310`] [ɪɴSPIRE].

[79] T. Becher, C. Lorentzen and M.D. Schwartz, *Precision Direct Photon and W-Boson Spectra at High p_T and Comparison to LHC Data*, *Phys. Rev.* **D 86** (2012) 054026 [`arXiv:1206.6115`] [ɪɴSPIRE].

[80] M. Dasgupta, L. Magnea and G.P. Salam, *Non-perturbative QCD effects in jets at hadron colliders*, *JHEP* **02** (2008) 055 [`arXiv:0712.3014`] [ɪɴSPIRE].