Infrared sensitivity of single jet inclusive production at hadron colliders

James Currie,[a] **Aude Gehrmann-De Ridder,**[b,c] **Thomas Gehrmann,**[c] **Nigel Glover,**[a] **Alexander Huss**[d] **and João Pires**[e]

[a] *Institute for Particle Physics Phenomenology, Durham University,*
Durham, DH1 3LE U.K.

[b] *Institute for Theoretical Physics, ETH,*
Zürich, CH-8093 Switzerland

[c] *Physik-Institut, Universität Zürich,*
Winterthurerstrasse 190, Zürich, CH-8057 Switzerland

[d] *Theoretical Physics Department, CERN,*
Geneva 23, 1211 Switzerland

[e] *Centro de Física Teórica de Partículas — CFTP, Instituto Superior Técnico IST,*
Universidade de Lisboa, Av. Rovisco Pais, Lisboa, P-1049-001 Portugal

E-mail: james.currie@durham.ac.uk, gehra@phys.ethz.ch,
thomas.gehrmann@uzh.ch, e.w.n.glover@durham.ac.uk,
alexander.huss@cern.ch, joao.ramalho.pires@tecnico.ulisboa.pt

ABSTRACT: Jet production at hadron colliders is a benchmark process to probe the dynamics of the strong interaction and the structure of the colliding hadrons. One of the most basic jet production observables is the single jet inclusive cross section, which is obtained by summing all jets that are observed in an event. Our recent computation of next-to-next-to-leading order (NNLO) QCD contributions to single jet inclusive observables uncovered large corrections in certain kinematical regions, which also resulted in a sizeable ambiguity on the appropriate choice of renormalization and factorization scales. We now perform a detailed investigation of the infrared sensitivity of the different ingredients to the single jet inclusive cross section. We show that the contribution from the second jet, ordered in transverse momentum p_{T}, in the event is particularly sensitive to higher order effects due to implicit restrictions on its kinematics. By investigating the second-jet transverse momentum distribution, we identify large-scale cancellations between different kinematical event configurations, which are aggravated by certain types of scale choice. Taking perturbative convergence and stability as selection criteria enables us to single out the total partonic transverse energy $\hat{H}_{\mathrm{T}}$ and twice the individual jet transverse momentum $2\,p_{\mathrm{T}}$ (with which $\hat{H}_{\mathrm{T}}$ coincides in Born kinematics) as the most appropriate scales in the perturbative description of single jet inclusive production.

Contents

1 Introduction

At hadron colliders, the factorised form of the inclusive cross section is given by,

$$\mathrm{d}\sigma = \sum_{i,j} \int_0^1 \mathrm{d}x_1 \int_0^1 \mathrm{d}x_2 \, f_i(x_1, \mu_{\mathrm{F}}) \, f_j(x_2, \mu_{\mathrm{F}}) \, \mathrm{d}\hat{\sigma}_{ij}(\alpha_{\mathrm{s}}(\mu_{\mathrm{R}}), \mu_{\mathrm{R}}^2, \mu_{\mathrm{F}}^2), \qquad (1.1)$$

where $\mathrm{d}\hat{\sigma}_{ij}$ is the parton-level scattering cross section for parton i to scatter off parton j and the sum runs over the possible parton types i and j. The probability of finding a parton of type i in the proton carrying a momentum fraction x is described by the parton distribution function (PDF) $f_i(x)\mathrm{d}x$. By applying suitable cuts, one can study more exclusive observables such as the transverse momentum distribution or the rapidity distribution of the hard objects (jets or vector bosons, Higgs bosons or other new particles) produced in the hard scattering. In eq. (1.1), one has to fix the renormalization scale μ_{R} for the strong coupling $\alpha_{\mathrm{s}}(\mu_{\mathrm{R}})$, and the mass factorization scale μ_{F} for the parton distribution functions $f_i(x, \mu_{\mathrm{F}})$.

In this paper we study jet production at hadron colliders, in particular the single jet inclusive cross section in proton-proton collisions, $\mathrm{d}\sigma(\mathrm{p}+\mathrm{p} \rightarrow \mathrm{jet}+X)$, which is obtained by

summing over all jets in the event. The observable is inclusive over all additional radiation as no further kinematical constraints are imposed on the final state particles beyond the requirement of observing at least a single jet. In this way, the full event can contain multiple jets and all jets that lie in a given range of rapidity y and transverse momentum p_T are taken into account in determining the single jet inclusive cross section for that bin.

Large-p_T jet production at hadron colliders has been studied in particle accelerators over a period of many years by the UA1 [1] and UA2 [2] experiments at the Sp$\bar{p}$S collider ($\sqrt{s} = 546\,\text{GeV}$ and $630\,\text{GeV}$) and by the CDF [3] and D0 [4] experiments at the Tevatron ($\sqrt{s} = 1.96\,\text{TeV}$). At the Large Hadron Collider (LHC) at CERN, the ALICE, ATLAS and CMS collaborations have measured inclusive jet cross sections in proton-proton collisions at centre-of-mass energies of $\sqrt{s} = 2.76\,\text{TeV}$ [5–7], $7\,\text{TeV}$ [8, 9], $8\,\text{TeV}$ [10, 11] and $13\,\text{TeV}$ [12, 13]. These precise measurements are crucial for understanding physics at hadron colliders as jet cross sections provide valuable information about the strong coupling constant α_s, the non-perturbative structure of the proton encoded in the PDFs, and probe the shortest distance scales that are experimentally attainable. More recently, jet substructure techniques have been applied to understand the internal dynamics of QCD jets in order to identify discriminator variables which can more easily disentangle jets originating in the QCD parton scattering process from those produced by the hadronic decay of new heavy beyond-Standard-Model particles [14].

Hadron collider jet observables can be computed at a given fixed order in α_s in perturbative QCD, by retaining the corresponding terms in the series expansion in α_s for the parton-level cross sections and the PDFs, as presented in eq. (1.1). Next-to-leading-order (NLO) QCD corrections to jet production at hadron colliders were computed in [15–17] and later combined with a parton shower in [18, 19]. First-order corrections in the electroweak (EW) coupling have been derived in [20, 21], and the combination of NLO QCD and EW corrections was studied in [22]. A study of joint jet radius and threshold resummation has been presented in [23]. Progress in next-to-next-to-leading-order (NNLO) QCD calculations has been made over the past several years [24–27]. After the completion of the first calculations of the gluons-only subprocess [28, 29], the complete leading-colour and leading-N_F NNLO QCD corrections to the single jet inclusive production [30] and to di-jet production [31] were obtained recently.

The recent NNLO calculation provides new opportunities for QCD studies at hadron colliders, it enables precise theoretical predictions for jet observables to be compared with the wealth of experimental jet data which have similar precision. More formally, the knowledge of three orders in the perturbative expansion in α_s for these jet observables provides a testing ground for the impact of the higher order corrections through the notion of perturbative convergence and the reduction of theoretical scale uncertainties.

One issue which requires particular attention is the role of the renormalization and factorization scales in the theoretical predictions. At a formal level, the parameters μ_R and μ_F are introduced as auxiliary quantities which allow meaningful predictions to be calculated at each order in perturbation theory. As auxiliary quantities, an all-order prediction would be independent of these parameters. However, truncating at fixed order yields a residual dependence, formally of one order higher in the strong coupling. Varying

the numerical value of the scale (usually in an interval around a pre-defined central scale choice) is frequently used to quantify the uncertainty on the theory prediction due to the uncomputed higher orders. The huge dynamical range of jet production at the LHC and the three available perturbative orders in the theoretical prediction provide the opportunity to test thoroughly the commonly used arguments about scale dependence in perturbative calculations.

For these reasons we have recently provided jet cross section predictions for the LHC at NNLO using both the leading jet transverse momentum in an event $\mu_{\rm R} = \mu_{\rm F} = p_{\rm T,1}$ [30] or each individual jet transverse momentum $\mu_{\rm R} = \mu_{\rm F} = p_{\rm T}$ [32] as a central scale choice. We have observed an overall reduction in the scale dependence of the prediction at NNLO with respect to the NLO result with either of these scale choices. However, comparing the two predictions (which are both based on well-motivated central scale choices) against each other, we noticed a substantial difference in their quantitative behaviour [32], which can be viewed as a further uncertainty on the theory prediction. It is therefore important to arrive at a sensible central scale choice, which covers the range of jet kinematics accessible at the LHC.

In this paper, we perform a detailed study of the perturbative behaviour of the individual contributions to the single jet inclusive production cross section for a given set of sensibly chosen dynamical scales. In section 2 we present the structure and the scale dependence of the single jet inclusive cross section computed through to NNLO in QCD. We discuss possible functional forms for the scale choice in terms of kinematical variables, thereby carefully distinguishing scales which are based on individual jet kinematics (jet-based) or on full event kinematics (event-based). Particular attention is paid to the effects of the jet clustering algorithm and the jet resolution parameter on the kinematical variables used in the different scale choices.

In section 3 we subsequently perform a detailed investigation of the infrared sensitivity of the different ingredients to the single jet inclusive cross section for the jet-based scale choice $\mu_{\rm R} = \mu_{\rm F} = p_{\rm T}$ and the event-based scale choice $\mu_{\rm R} = \mu_{\rm F} = p_{\rm T,1}$. It is the aim of this section to identify the source of the different quantitative behaviour in the NNLO predictions between the $\mu_{\rm R} = \mu_{\rm F} = p_{\rm T}$ and $\mu_{\rm R} = \mu_{\rm F} = p_{\rm T,1}$ scale choices.

In section 4 we analyse the behaviour of the perturbative expansion of the single jet inclusive observable, for the different functional forms for the central scale choice established in section 2. This allows us to assess the perturbative stability and convergence properties for each scale choice up to NNLO, thereby identifying the most appropriate candidates. Subsequently we compare our predictions at NLO and NNLO to the available CMS 13 TeV jet data for the first time in section 5. Finally in section 6 we present our conclusions.

2 Renormalization and factorization scales in the single jet inclusive cross section

When calculating jet cross sections to fixed order in perturbation theory (1.1), one has to fix the renormalization scale $\mu_{\rm R}$ for the strong coupling $\alpha_{\rm s}(\mu_{\rm R})$, and the mass factorization scale $\mu_{\rm F}$ for the parton distribution functions $f_i(x, \mu_{\rm F})$.

The behaviour of the coupling constant and parton distributions under scale variations is determined by evolution equations. After fixing a central reference scale, all scale-dependent terms of hard scattering cross sections can be inferred by expanding the solutions of the evolution equations in powers of the strong coupling constant. These are collected in section 2.1 below. For processes involving massive particles (vector bosons, Higgs bosons or top quarks), the particle mass provides a natural candidate for the central reference scale. In contrast, no natural fixed scale is present in jet production processes, which involve only massless objects at parton level. Consequently, the central scale for jet production must be chosen dynamically, based on the kinematics of the final-state objects (jets or full events) under consideration. Section 2.2 discusses different prescriptions for the central scale in single jet inclusive production, based on the kinematics of each individual jet, or of the whole event. These kinematical variables depend on the jet algorithm and the jet resolution parameter. In section 2.3 we define the individual jet contributions to the single jet inclusive production cross section which are individually infrared safe only if they are inclusive in the jet rapidity. Analysing the different possible final state configurations up to NNLO, we finally discuss the impact of the jet resolution on the event properties and on the scale choices in section 2.4.

2.1 Scale dependent terms up to NNLO

2.1.1 Renormalization scale dependence

The renormalization group equation describing the running of α_{s} as a function of the renormalization scale μ_{R} reads:

$$\mu_{\mathrm{R}}^2 \frac{\mathrm{d}\alpha_{\mathrm{s}}(\mu_{\mathrm{R}})}{\mathrm{d}\mu_{\mathrm{R}}^2} = -\alpha_{\mathrm{s}}(\mu_{\mathrm{R}}) \left[\beta_0 \left(\frac{\alpha_{\mathrm{s}}(\mu_{\mathrm{R}})}{2\pi} \right) + \beta_1 \left(\frac{\alpha_{\mathrm{s}}(\mu_{\mathrm{R}})}{2\pi} \right)^2 + \beta_2 \left(\frac{\alpha_{\mathrm{s}}(\mu_{\mathrm{R}})}{2\pi} \right)^3 + \mathcal{O}(\alpha_{\mathrm{s}}^4) \right],$$

(2.1)

with the $\overline{\mathrm{MS}}$-scheme coefficients [33, 34]

$$\begin{aligned}
\beta_0 &= \frac{11 C_{\mathrm{A}} - 4 T_{\mathrm{R}} N_{\mathrm{F}}}{6}, \\
\beta_1 &= \frac{17 C_{\mathrm{A}}^2 - 10 C_{\mathrm{A}} T_{\mathrm{R}} N_{\mathrm{F}} - 6 C_{\mathrm{F}} T_{\mathrm{R}} N_{\mathrm{F}}}{6}, \\
\beta_2 &= \frac{1}{432} \big(2857 C_{\mathrm{A}}^3 + 108 C_{\mathrm{F}}^2 T_{\mathrm{R}} N_{\mathrm{F}} - 1230 C_{\mathrm{F}} C_{\mathrm{A}} T_{\mathrm{R}} N_{\mathrm{F}} - 2830 C_{\mathrm{A}}^2 T_{\mathrm{R}} N_{\mathrm{F}} \\
&\quad + 264 C_{\mathrm{F}} T_{\mathrm{R}}^2 N_{\mathrm{F}}^2 + 316 C_{\mathrm{A}} T_{\mathrm{R}}^2 N_{\mathrm{F}}^2 \big),
\end{aligned}$$

(2.2)

where $C_{\mathrm{A}} = 3$, $C_{\mathrm{F}} = 4/3$, $T_{\mathrm{R}} = 1/2$ and N_{F} is the number of light quark flavours.

Using the solution of this equation, the coupling at a fixed scale μ_{R_0} can be truncated in terms of the coupling at μ_{R} by introducing

$$L_{\mathrm{R}} = \log \left(\frac{\mu_{\mathrm{R}}^2}{\mu_{\mathrm{R}_0}^2} \right)$$

(2.3)

as

$$\alpha_{\mathrm{s}}(\mu_{\mathrm{R}_0}) = \alpha_{\mathrm{s}}(\mu_{\mathrm{R}}) \left[1 + \beta_0 L_R \frac{\alpha_{\mathrm{s}}(\mu_{\mathrm{R}})}{2\pi} + \left[\beta_0^2 L_R^2 + \beta_1 L_R \right] \left(\frac{\alpha_{\mathrm{s}}(\mu_{\mathrm{R}})}{2\pi} \right)^2 + \mathcal{O}(\alpha_{\mathrm{s}}^3) \right]. \quad (2.4)$$

The perturbative expansion of the single jet inclusive cross section starts at order α_{s}^2. In evaluating the expansion coefficients $\sigma^{(n)} = \sigma^{(n)}(\mu_{\mathrm{R}_0})$, the renormalization scale is fixed to a value μ_{R_0} (which can be dynamically evaluated event-by-event). Rescalings can then be made for a fixed ratio $\mu_{\mathrm{R}}/\mu_{\mathrm{R}_0}$ for all events; e.g. if $\mu_{\mathrm{R}_0} = p_{\mathrm{T},1}$, we can rescale to $\mu_{\mathrm{R}} = 2\, p_{\mathrm{T},1}$ or $\mu_{\mathrm{R}} = p_{\mathrm{T},1}/2$, but not to $\mu_{\mathrm{R}} = M_Z$ or $\mu_{\mathrm{R}} = H_T$).

2.1.2 Factorization scale dependence

The evolution of parton distributions associated to a variation of the factorization scale μ_{F} is determined by the Altarelli-Parisi equations [35] which read (omitting for simplicity the dependence on the Bjorken scaling variable x):

$$\mu_{\mathrm{F}}^2 \frac{\mathrm{d}}{\mathrm{d}\mu_{\mathrm{F}}^2} f_i(\mu_{\mathrm{F}}, \mu_{\mathrm{R}}) = \sum_j P_{ij}(\alpha_{\mathrm{s}}(\mu_{\mathrm{R}}), \mu_{\mathrm{F}}, \mu_{\mathrm{R}}) \otimes f_j(\mu_{\mathrm{F}}, \mu_{\mathrm{R}}). \quad (2.5)$$

The expansion to the third order of the splitting functions $P_{ij}^{(n)}$ computed for $\mu_{\mathrm{F}} = \mu_{\mathrm{R}}$ is [36, 37]:

$$P_{ij}(\alpha_{\mathrm{s}}(\mu_{\mathrm{R}}), \mu_{\mathrm{F}}, \mu_{\mathrm{R}}) = \frac{\alpha_{\mathrm{s}}(\mu_{\mathrm{R}})}{2\pi} P_{ij}^{(0)} + \left(\frac{\alpha_{\mathrm{s}}(\mu_{\mathrm{R}})}{2\pi} \right)^2 \left[P_{ij}^{(1)} + \beta_0 l P_{ij}^{(0)} \right]$$
$$+ \left(\frac{\alpha_{\mathrm{s}}(\mu_{\mathrm{R}})}{2\pi} \right)^3 \left[P_{ij}^{(2)} + \left(\beta_1 P_{ij}^{(0)} + 2\beta_0 P_{ij}^{(0)} \right) l + \beta_0^2 l^2 P_{ij}^{(0)} \right] + \mathcal{O}(\alpha_{\mathrm{s}}^4), \quad (2.6)$$

where we introduced

$$l = \log\left(\frac{\mu_{\mathrm{R}}^2}{\mu_{\mathrm{F}}^2} \right). \quad (2.7)$$

Note that (2.6) can be rewritten as

$$P_{ij}(\alpha_{\mathrm{s}}(\mu_{\mathrm{R}}), \mu_{\mathrm{F}}, \mu_{\mathrm{R}}) = \frac{\alpha_{\mathrm{s}}(\mu_{\mathrm{F}})}{2\pi} P_{ij}^{(0)} + \left(\frac{\alpha_{\mathrm{s}}(\mu_{\mathrm{F}})}{2\pi} \right)^2 P_{ij}^{(1)} + \left(\frac{\alpha_{\mathrm{s}}(\mu_{\mathrm{F}})}{2\pi} \right)^3 P_{ij}^{(2)} + \mathcal{O}(\alpha_{\mathrm{s}}^4), \quad (2.8)$$

which implies that $f_i(\mu_{\mathrm{F}}, \mu_{\mathrm{R}})$ and $f_i(\mu_{\mathrm{F}}, \mu_{\mathrm{F}})$ fulfil the same evolution equation to all perturbative orders. The finite scheme transformation between both possible choices for μ_{F}, equal or different from μ_{R}, is thus vanishing to all orders and both functions can at most vary in their non-perturbative boundary conditions. For all perturbative purposes, we thus have

$$f_i(\mu_{\mathrm{F}}, \mu_{\mathrm{R}}) = f_i(\mu_{\mathrm{F}}, \mu_{\mathrm{F}}) \equiv f_i(\mu_{\mathrm{F}}), \quad (2.9)$$

which we will normally use in what follows (except if the scale transformation of the parton distribution is not expanded in $\alpha_{\mathrm{s}}(\mu_{\mathrm{F}})$, but in $\alpha_{\mathrm{s}}(\mu_{\mathrm{R}})$).

The parton distribution at a fixed scale μ_{F_0} can be expressed in terms of parton distributions at μ_F by expanding the solution of (2.5). We distinguish the expansion in powers of $\alpha_s(\mu_R)$ and in powers of $\alpha_s(\mu_F)$ and introduce

$$L_F = \log\left(\frac{\mu_F^2}{\mu_{F_0}^2}\right).\tag{2.10}$$

The expansion in $\alpha_s(\mu_R)$ of the parton distribution at μ_{F_0} reads then:

$$
\begin{aligned}
f_i(\mu_{F_0}) = {}& f_i(\mu_F) - \frac{\alpha_s(\mu_R)}{2\pi} P_{ij}^{(0)} \otimes f_j(\mu_F) L_F \\
& - \left(\frac{\alpha_s(\mu_R)}{2\pi}\right)^2 \left[P_{ij}^{(1)} \otimes f_j(\mu_F) L_F - \frac{1}{2} P_{ij}^{(0)} \otimes P_{jk}^{(0)} \otimes f_k(\mu_F) L_F^2 \right. \\
& \left. + P_{ij}^{(0)} \otimes f_j(\mu_F) \beta_0 L_F \left(l + \frac{1}{2} L_F \right) \right] + \mathcal{O}(\alpha_s^3).
\end{aligned}
\tag{2.11}
$$

The expansion in powers of $\alpha_s(\mu_F)$ is obtained from the above by setting $\mu_R = \mu_F$ in α_s to yield

$$
\begin{aligned}
f_i(\mu_{F_0}) = {}& f_i(\mu_F) - \frac{\alpha_s(\mu_F)}{2\pi} P_{ij}^{(0)} \otimes f_j(\mu_F) L_F \\
& - \left(\frac{\alpha_s(\mu_F)}{2\pi}\right)^2 \left[P_{ij}^{(1)} \otimes f_j(\mu_F) L_F - \frac{1}{2} P_{ij}^{(0)} \otimes P_{jk}^{(0)} \otimes f_k(\mu_F) L_F^2 \right. \\
& \left. + \frac{1}{2} P_{ij}^{(0)} \otimes f_j(\mu_F) \beta_0 L_F^2 \right] + \mathcal{O}(\alpha_s^3).
\end{aligned}
\tag{2.12}
$$

In both expressions, a summation over indices appearing twice is implicit.

2.1.3 Hadron collider jet cross section

Using the results presented above, one can compute the perturbative coefficients of the hadron collider cross section with default values of $\mu_{F_0} = \mu_{R_0} = \mu_0$. The perturbative expansion to NNLO reads:

$$
\begin{aligned}
\sigma(\mu_{R_0}, \mu_{F_0}, \alpha_s(\mu_{R_0})) = {}& \left(\frac{\alpha_s(\mu_{R_0})}{2\pi}\right)^2 \hat{\sigma}_{ij}^{(0)} \otimes f_i(\mu_{F_0}) \otimes f_j(\mu_{F_0}) \\
& + \left(\frac{\alpha_s(\mu_{R_0})}{2\pi}\right)^3 \hat{\sigma}_{ij}^{(1)} \otimes f_i(\mu_{F_0}) \otimes f_j(\mu_{F_0}) \\
& + \left(\frac{\alpha_s(\mu_{R_0})}{2\pi}\right)^4 \hat{\sigma}_{ij}^{(2)} \otimes f_i(\mu_{F_0}) \otimes f_j(\mu_{F_0}) + \mathcal{O}(\alpha_s^5).
\end{aligned}
\tag{2.13}
$$

The full scale dependence of this expression, for μ_F and μ_R different from each other, can be recovered by inserting (2.4) and (2.11) into the above equation. It yields

$$
\begin{aligned}
\sigma(\mu_R, \mu_F, &\alpha_s(\mu_R), L_R, L_F) = \\
& \left(\frac{\alpha_s(\mu_R)}{2\pi}\right)^2 \hat{\sigma}_{ij}^{(0)} \otimes f_i(\mu_F) \otimes f_j(\mu_F) + \left(\frac{\alpha_s(\mu_R)}{2\pi}\right)^3 \hat{\sigma}_{ij}^{(1)} \otimes f_i(\mu_F) \otimes f_j(\mu_F)
\end{aligned}
$$

$$+L_{\mathrm{R}}\left(\frac{\alpha_{\mathrm{s}}(\mu_{\mathrm{R}})}{2\pi}\right)^3 2\beta_0\,\hat{\sigma}_{ij}^{(0)}\otimes f_i(\mu_{\mathrm{F}})\otimes f_j(\mu_{\mathrm{F}})+L_{\mathrm{F}}\left(\frac{\alpha_{\mathrm{s}}(\mu_{\mathrm{R}})}{2\pi}\right)^3$$

$$\cdot\left[-\hat{\sigma}_{ij}^{(0)}\otimes f_i(\mu_{\mathrm{F}})\otimes\left(P_{jk}^{(0)}\otimes f_k(\mu_{\mathrm{F}})\right)-\hat{\sigma}_{ij}^{(0)}\otimes\left(P_{ik}^{(0)}\otimes f_k(\mu_{\mathrm{F}})\right)\otimes f_j(\mu_{\mathrm{F}})\right]$$

$$+\left(\frac{\alpha_{\mathrm{s}}(\mu_{\mathrm{R}})}{2\pi}\right)^4\hat{\sigma}_{ij}^{(2)}\otimes f_i(\mu_{\mathrm{F}})\otimes f_j(\mu_{\mathrm{F}})$$

$$+L_{\mathrm{R}}\left(\frac{\alpha_{\mathrm{s}}(\mu_{\mathrm{R}})}{2\pi}\right)^4\left(3\,\beta_0\,\hat{\sigma}_{ij}^{(1)}+2\,\beta_1\,\hat{\sigma}_{ij}^{(0)}\right)\otimes f_i(\mu_{\mathrm{F}})\otimes f_j(\mu_{\mathrm{F}})$$

$$+L_{\mathrm{R}}^2\left(\frac{\alpha_{\mathrm{s}}(\mu_{\mathrm{R}})}{2\pi}\right)^4 3\,\beta_0^2\,\hat{\sigma}_{ij}^{(0)}\otimes f_i(\mu_{\mathrm{F}})\otimes f_j(\mu_{\mathrm{F}})$$

$$+L_{\mathrm{F}}\left(\frac{\alpha_{\mathrm{s}}(\mu_{\mathrm{R}})}{2\pi}\right)^4\left[-\hat{\sigma}_{ij}^{(1)}\otimes f_i(\mu_{\mathrm{F}})\otimes\left(P_{jk}^{(0)}\otimes f_k(\mu_{\mathrm{F}})\right)-\hat{\sigma}_{ij}^{(1)}\otimes\left(P_{ik}^{(0)}\otimes f_k(\mu_{\mathrm{F}})\right)\otimes f_j(\mu_{\mathrm{F}})\right.$$

$$\left.-\hat{\sigma}_{ij}^{(0)}\otimes f_i(\mu_{\mathrm{F}})\otimes\left(P_{jk}^{(1)}\otimes f_k(\mu_{\mathrm{F}})\right)-\hat{\sigma}_{ij}^{(0)}\otimes\left(P_{ik}^{(1)}\otimes f_k(\mu_{\mathrm{F}})\right)\otimes f_j(\mu_{\mathrm{F}})\right]$$

$$+L_{\mathrm{F}}^2\left(\frac{\alpha_{\mathrm{s}}(\mu_{\mathrm{R}})}{2\pi}\right)^4\left[\hat{\sigma}_{ij}^{(0)}\otimes\left(P_{ik}^{(0)}\otimes f_k(\mu_{\mathrm{F}})\right)\otimes\left(P_{jl}^{(0)}\otimes f_l(\mu_{\mathrm{F}})\right)\right.$$

$$+\frac{1}{2}\hat{\sigma}_{ij}^{(0)}\otimes f_i(\mu_{\mathrm{F}})\otimes\left(P_{jk}^{(0)}\otimes P_{kl}^{(0)}\otimes f_l(\mu_{\mathrm{F}})\right)$$

$$+\frac{1}{2}\hat{\sigma}_{ij}^{(0)}\otimes\left(P_{ik}^{(0)}\otimes P_{kl}^{(0)}\otimes f_l(\mu_{\mathrm{F}})\right)\otimes f_j(\mu_{\mathrm{F}})$$

$$+\frac{1}{2}\beta_0\,\hat{\sigma}_{ij}^{(0)}\otimes f_i(\mu_{\mathrm{F}})\otimes\left(P_{jk}^{(0)}\otimes f_k(\mu_{\mathrm{F}})\right)$$

$$\left.+\frac{1}{2}\beta_0\,\hat{\sigma}_{ij}^{(0)}\otimes\left(P_{ik}^{(0)}\otimes f_k(\mu_{\mathrm{F}})\right)\otimes f_j(\mu_{\mathrm{F}})\right]$$

$$+L_{\mathrm{F}}L_{\mathrm{R}}\left(\frac{\alpha_{\mathrm{s}}(\mu_{\mathrm{R}})}{2\pi}\right)^4\left[-3\,\beta_0\,\hat{\sigma}_{ij}^{(0)}\otimes f_i(\mu_{\mathrm{F}})\otimes\left(P_{jk}^{(0)}\otimes f_k(\mu_{\mathrm{F}})\right)\right.$$

$$\left.-3\,\beta_0\,\hat{\sigma}_{ij}^{(0)}\otimes\left(P_{ik}^{(0)}\otimes f_k(\mu_{\mathrm{F}})\right)\otimes f_j(\mu_{\mathrm{F}})\right]+\mathcal{O}(\alpha_{\mathrm{s}}^5)\,. \tag{2.14}$$

2.2 Scale choices

Inclusive jet observables accumulate each reconstructed jet in the event to the same kinematic distribution, resulting in multiple bookings of the event into a given histogram. The set of possible scale choices is consequently large and we shall distinguish two generic types: event-based and jet-based scales. A *jet-based scale* only uses kinematic information from the individual jet to determine the scale associated with the contribution from this jet to the cross section. In a given event, the event weight is thus evaluated at several different scales, one scale for each jet. In contrast to this, an *event-based scale* uses information from the full final state of the event to set a common scale for all binnings of the jets that are contained in this event.

In this paper we will consider the following set of functional forms for the scale choice (and multiples thereof):

the individual jet transverse momentum p_T: when this jet-based scale choice is used for the inclusive p_T distribution, the observable is directly aligned with the scale itself, making it a convenient choice for PDF fits. It mimics kinematical hierarchies in an event, where multiple jets can be reconstructed with very different p_T. However, this can lead to the scale being set to values that are not at all representative of the underlying hard scattering process.

the leading-jet transverse momentum $p_{T,1}$: this event-based scale uses the p_T of the hardest jet in the event, which is a better proxy for the scale of the hard interaction compared to the $\mu = p_T$ choice. For multi-jet events comprising many hard resolved jets, $p_{T,1}$ can still underestimate the scale of the hard interaction. Moreover, $p_{T,1}$ does not take account of scale hierarchies in an event.

the scalar sum of the transverse momenta of all reconstructed jets H_T: with this event-based scale one incorporates the kinematics of all individual jets by summing up their respective transverse momenta, $H_T = \sum_{i \in \text{jets}} p_{T,i}$. As such, it constitutes the hardest scale discussed so far and for the Born-level $2 \to 2$ process, it is related to the p_T scales as $H_T = 2\, p_T = 2\, p_{T,1}$. It however suffers from a discontinuous behaviour, when the number of reconstructed jets changes:

$$n_{\text{jets}} = 1 \Rightarrow H_T = p_{T,1}, \qquad n_{\text{jets}} = 2 \Rightarrow H_T = p_{T,1} + p_{T,2}, \qquad \ldots. \qquad (2.15)$$

For this reason, it displays a large displacement at the phase-space boundaries where $(n+1)$-jet events migrate to n-jet events. As a consequence, higher order corrections for values of p_T close to the minimum jet acceptance $p_{T,\text{min}}$ become unstable and we will no longer consider this scale in the remainder of this paper.

the scalar sum of the transverse momenta of all partons $\hat{H}_T$: the undesirable discontinuous behaviour of H_T can be alleviated if the transverse momentum sum is not based on the reconstructed jets, but instead obtained as the transverse momentum sum of all partons in the event: $\hat{H}_T = \sum_{i \in \text{partons}} p_{T,i}$. This event-based scale choice also has the advantage of being insensitive to the jet reconstruction applied in the analysis and is an infrared-safe event shape variable.

Any scale choice that is based on the kinematics of the reconstructed jets, i.e. p_T, $p_{T,1}$, and H_T from the list above, inherits a dependence on the jet cuts and the details of the clustering employed in the analysis [38]. This means that for a given partonic configuration and the same scale definition, the determined value for the scale depends on the details of the jet algorithm, the allowed rapidity range and the rapidity and p_T range probed by the experiment. In particular, the scale choice introduces an indirect dependence on the cone size R of the jet algorithm and on the jet cuts:

scale		type	$n_{\text{jets}} \to n_{\text{jets}} + 1$	analysis dependence
p_{T}	$(2\,p_{\text{T}})$	jet based	continuous	yes
$p_{\text{T},1}$	$(2\,p_{\text{T},1})$	event based	continuous	yes
$H_{\text{T}}/2\,(H_{\text{T}})$		event based	discontinuous	yes
$\hat{H}_{\text{T}}/2\,(\hat{H}_{\text{T}})$		event based	continuous	no

Table 1. Possible scale choices in inclusive jet production and their properties.

- A sensitivity of event-based scales on the jet cuts induces the unwanted property that a variation of rapidity cuts can impact the predictions in the other rapidity regions well away from the variation.

- The dependence of the scale on the jet-clustering algorithm can introduce an indirect sensitivity on the cone size R. Such an effect becomes hard to disentangle from the purely *kinematical* dependence on R, which is discussed in section 2.4 and which induces potentially enhanced corrections of the form $\log(R)$.

As an example of the above, consider the event-based scale $\mu = p_{\text{T},1}$ and a configuration in which the leading jet is relatively forward and thus does not contribute in a central rapidity slice of the single jet inclusive cross section. If the detector rapidity coverage includes the jet, the scale will be the p_{T} of this forward jet. On the other hand, if the forward jet lies outside of the detector coverage it will not be identified as the leading jet, and the event-based scale will be different. As a consequence, predictions for the jet cross section in the central region of the detector will depend on the rapidity coverage of the detector when the event-based scale $p_{\text{T},1}$ is used.

In contrast, the $\mu = p_{\text{T}}$ scale choice always uses the transverse momentum of the jet in the rapidity slice where the jet is observed and therefore its predictions are not sensitive to the jet-defining cuts. However, as will be detailed in section 2.4, the scales $\mu = p_{\text{T},1}$ and $\mu = p_{\text{T}}$ show a different sensitivity to the jet cone size.

Both of these issues are avoided for $\mu = \hat{H}_{\text{T}}$ that is defined on the basis of the parton kinematics. While not being directly accessible in the experimental measurement, $\hat{H}_{\text{T}}$ is infrared-safe and theoretically well-defined. Its use for scale settings is not problematic, since the renormalization and factorization scales are simply auxiliary quantities in the theoretical prediction.

Table 1 summarises the different scale choices together with their respective properties discussed in this section.

2.3 Individual jet contributions to inclusive jet production

To illustrate the difference between an event-based and a jet-based scale choice, we consider two of the most common scale choices in studies of jet production at hadron colliders, i.e. $\mu = p_{\text{T},1}$ and $\mu = p_{\text{T}}$. To this end, it is instructive to look at the composition of the single jet inclusive cross section in terms of contributions from individual jets in an

event (ordered in decreasing transverse momentum). With the event-based scale choice $\mu_{R_0} = \mu_{F_0} = \mu = p_{T,1}$, then through to $\mathcal{O}(\alpha_s^4)$ we have:

$$\frac{d\sigma}{dp_T}(\mu = p_{T,1}) = \frac{d\sigma}{dp_{T,1}}(\mu = p_{T,1}) + \frac{d\sigma}{dp_{T,2}}(\mu = p_{T,1}) + \frac{d\sigma}{dp_{T,3}}(\mu = p_{T,1}) + \frac{d\sigma}{dp_{T,4}}(\mu = p_{T,1}).$$

Predictions for the jet-based scale choice $\mu = p_T$ can subsequently be obtained in the following way,

$$\begin{aligned}
\frac{d\sigma}{dp_T}(\mu = p_T) &= \frac{d\sigma}{dp_{T,1}}(\mu = p_{T,1}) + \frac{d\sigma}{dp_{T,2}}(\mu = p_{T,2}) + \frac{d\sigma}{dp_{T,3}}(\mu = p_{T,3}) + \frac{d\sigma}{dp_{T,4}}(\mu = p_{T,4}) \\
&= \frac{d\sigma}{dp_T}(\mu = p_{T,1}) \\
&\quad + \frac{d\sigma}{dp_{T,2}}(\mu = p_{T,2}) - \frac{d\sigma}{dp_{T,2}}(\mu = p_{T,1}) \\
&\quad + \frac{d\sigma}{dp_{T,3}}(\mu = p_{T,3}) - \frac{d\sigma}{dp_{T,3}}(\mu = p_{T,1}) \\
&\quad + \frac{d\sigma}{dp_{T,4}}(\mu = p_{T,4}) - \frac{d\sigma}{dp_{T,4}}(\mu = p_{T,1}),
\end{aligned} \tag{2.16}$$

such that the difference between the $\mu = p_{T,1}$ and $\mu = p_T$ results can be identified in the last three lines in equation (2.16). It will therefore be important to numerically study the individual sub-leading jet contributions to the inclusive jet sample and in particular the effects that can arise from changing the scale from an event-based scale to a jet-based scale.

When decomposing the inclusive jet cross section in terms of the contributions from leading and subleading jets, the individual jet distributions are well-defined and infrared-safe only if they are inclusive in the jet rapidity (with the same global rapidity cuts applied to all jets). Since the notion of leading and sub-leading jet is not well defined at leading order ($p_{T,1} = p_{T,2}$ at LO), the rapidity assignment to the leading and subleading jet is ambiguous for leading-order kinematics. When computing higher-order corrections, the rapidity of the leading and subleading jet may thus be interchanged between event and counter-event, causing them to end up in different rapidity bins, thereby obstructing their cancellation in infrared-divergent limits. On the other hand, in the inclusive jet transverse momentum distribution (which sums over all jets in the event) IR-safety is restored in differential distributions in rapidity y, since leading and subleading jet contributions are treated equally.

2.4 Dependence on the jet resolution parameter R

In this subsection we discuss the effects stemming from the jet definition itself, in particular the jet resolution. For the sake of illustration, we represent the jets by cones of radius R in rapidity and azimuthal angle, as obtained [39] by either a cone algorithm or the commonly used anti-k_T clustering/recombination algorithm [40].

Figure 1 shows some illustrations of various jet configurations at LO and NLO where solid arrows represent partons and cones represent jets resulting from the jet algorithm. Figure 1a shows a dijet event at leading order where two back-to-back partons form two
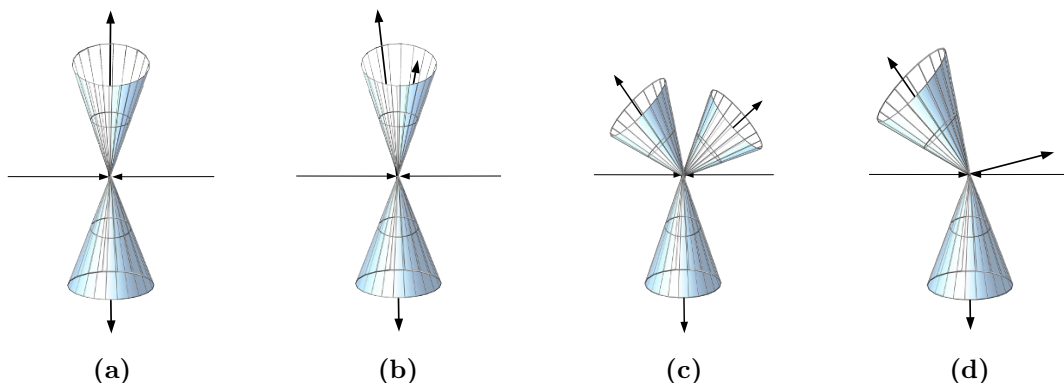
Figure 1. Illustrations of jet events at LO and NLO with arrows representing partons and cones representing jets.

jets and $p_{T,1} = p_{T,2}$. In this case there is no difference in scale choice between $p_{T,1}$ and p_T. Figure 1b shows a dijet event where three partons are clustered by the jet algorithm into two jets such that the jets are still balanced in p_T and the scale choice is identical. Figure 1c shows a trijet event where three partons are sufficiently hard and separated to form three distinct jets. In this configuration $p_{T,1} \neq p_{T,2} \neq p_{T,3}$ and so the scale choice does make a difference, although the three-jet contribution makes up only a very small fraction of the inclusive jet cross section as we will observe in section 3. Figure 1d depicts a dijet event where the third parton falls outside the jet radius and is not clustered but also is not sufficiently hard to form a jet on its own; such configurations typically lead to a small imbalance in the leading and subleading jet p_T and their description is sensitive to the scale parameterization.

At NNLO there are more configurations to consider due to the presence of four final-state partons in the double real contribution forming either two, three or four jets. Once again, many configurations do not contribute to the difference in scale parameterization. Whenever the jet algorithm clusters two, three or four partons into two jets then the jets are balanced in p_T and there is no difference between the $\mu = p_{T,1}$ and $\mu = p_T$ scale choices. The only NNLO configurations that can contribute to the difference are: three- or four-jet events (for which the cross section is very small) or two jet events where additional radiation falls outside of the jet radius, see figure 2.

As illustrated in figure 3 the choice of the R parameter in the jet algorithm can have an effect on how the partons are clustered into jets. We can take the same three-parton configuration and consider the clustering for different values of R, figure 3a, and a larger value of R, figure 3b. For the smaller R value the most subleading parton is more likely to fall outside the jet radius of the two leading jets and so generate a difference between the $\mu = p_{T,1}$ and $\mu = p_T$ scale parameterizations. Therefore, when using the $\mu = p_T$ scale choice, the value of the scale can vary with R for a fixed event. On the other hand, with the choice $\mu = p_{T,1}$, the scale for the event is R-independent at NLO, where the leading jet is not sensitive to radiation outside the cone, and becomes R-dependent only at NNLO.

This difference between the two scale choices grows significantly for small R, decreases for large R, and is moderate for the phenomenologically relevant values used at the LHC, for $R = 0.4$ (0.7) as we will observe in section 3.
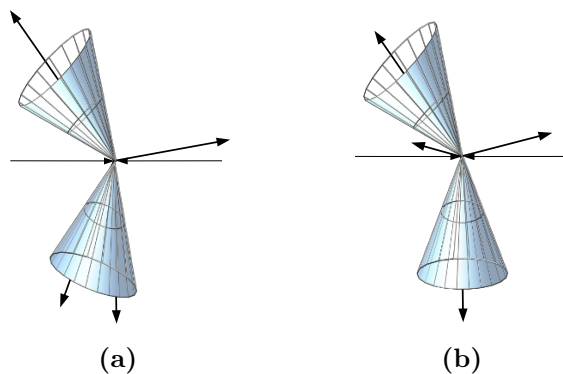
Figure 2. An illustration of 2-jet NNLO configurations that contribute to the p_T imbalance between the leading and subleading jet with arrows representing partons and cones representing jets.
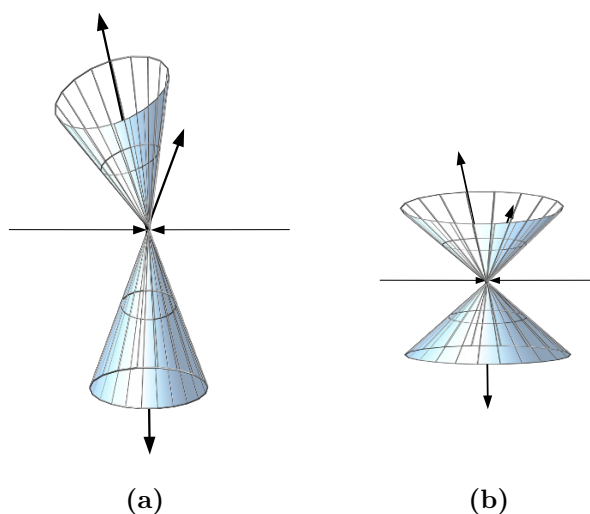

Figure 3. An illustration of clustering the same three-parton event with smaller (a) and larger (b) values of R.

3 The scale choices $\mu = p_{T,1}$ and $\mu = p_T$

As observed in ref. [32], the spread in the NNLO predictions for single jet inclusive production between using the dynamical scales $\mu = p_{T,1}$ and $\mu = p_T$ can be comparable or even larger in size than the respective uncertainties estimated through scale variations. The significant effect of this scale ambiguity on the NNLO predictions, and the lack of a theoretically well-motivated preference motivates us to revisit these results and to further study this issue.

For the leading jet in the event, the scale $\mu = p_T$ is identical to $\mu = p_{T,1}$ and its contribution is therefore insensitive to the scale choice between p_T and $p_{T,1}$. Furthermore, two-jet events where the jets are balanced in p_T cannot generate any difference as $p_T = p_{T,1} = p_{T,2}$. Away from these jet configurations, the subleading jets will have a smaller p_T than the leading jet in the event so that $p_{T,2},\ p_{T,3},\ \ldots < p_{T,1}$.

For these reasons, at LO the two scale choices generate the same prediction and similarly, for all events at higher order that have LO kinematics there is no difference between the two scale choices. In particular at high p_{T} the scale choices once again converge as is to be expected for the largely back-to-back configurations encountered at high p_{T}. Kinematical configurations where the scale choices do not coincide are events with three or more hard jets and events with hard emissions outside the jet fiducial cuts that generate an imbalance in p_{T} between the leading and subleading jets in the event. For this reason, we can expect also that for larger jet cone sizes the difference in the predictions using $\mu = p_{\mathrm{T}}$ or $\mu = p_{\mathrm{T},1}$ will be smaller, since the increased number of parton clusterings driven by a larger cone size promotes final state jets balanced in p_{T}.

It is the aim of this section to scrutinise how the contributions to the single jet inclusive transverse momentum distribution behave according to the choice of the functional form of the scale. To this end, we use the two central scale choices $\mu = p_{\mathrm{T}}$ and $\mu = p_{\mathrm{T},1}$ as representatives for jet-based and event-based scale settings. After describing our calculational set up in section 3.1, in sections 3.2 and 3.3 we study the impact on the transverse momentum distribution of the individual jet fractions at LO, NLO and NNLO level for the two central scale choices $\mu = p_{\mathrm{T}}$ and $\mu = p_{\mathrm{T},1}$ and for the two cone sizes $R = 0.7$ and $R = 0.4$. Having identified the crucial role of the second jet distribution from this analysis, in section 3.4 we focus our attention to this particular contribution and present how it behaves at a given perturbative order.

3.1 Calculational setup

In order to investigate the differences between the scale choices $\mu = p_{\mathrm{T}}$ and $\mu = p_{\mathrm{T},1}$ and their origin, we perform a numerical study for the single jet inclusive cross section at a center of mass energy $\sqrt{s} = 13\,\mathrm{TeV}$. The jets are identified using the anti-k_{T} algorithm [40] and results are presented for both $R = 0.4$ and 0.7 to further allow to inspect the dependence on the jet cone size.

Jets are accepted within the fiducial volume defined through the cuts

$$|y^j| < 4.7, \qquad p_{\mathrm{T}}^j > 114\,\mathrm{GeV}, \tag{3.1}$$

covering jet-p_{T} values up to $2\,\mathrm{TeV}$, and ordered in transverse momentum.[1] As explained in section 2.3, we can not apply a rapidity binning to leading and subleading jet distributions.

We systematically use the `PDF4LHC15_nnlo_100` PDF set [41] for the evaluation of the LO, NLO and NNLO contributions. This choice of a fixed PDF across the different perturbative orders allows us to quantify the effects of the two scale choices at the partonic cross section level rendering our conclusions independent of the PDF set used. The value for the strong coupling constant is given by $\alpha_{\mathrm{s}}(M_{\mathrm{Z}}) = 0.118$, as provided by the PDF set.

3.2 Corrections to the transverse momentum distribution

As a first step, we investigate the impact of including NLO and NNLO corrections to the single jet inclusive transverse momentum distribution. Figure 4 shows the size of the

---

[1]Unless otherwise stated, we use the p_{T} binning and rapidity bin widths used by the CMS collaboration in their $13\,\mathrm{TeV}$ jet measurement [12].
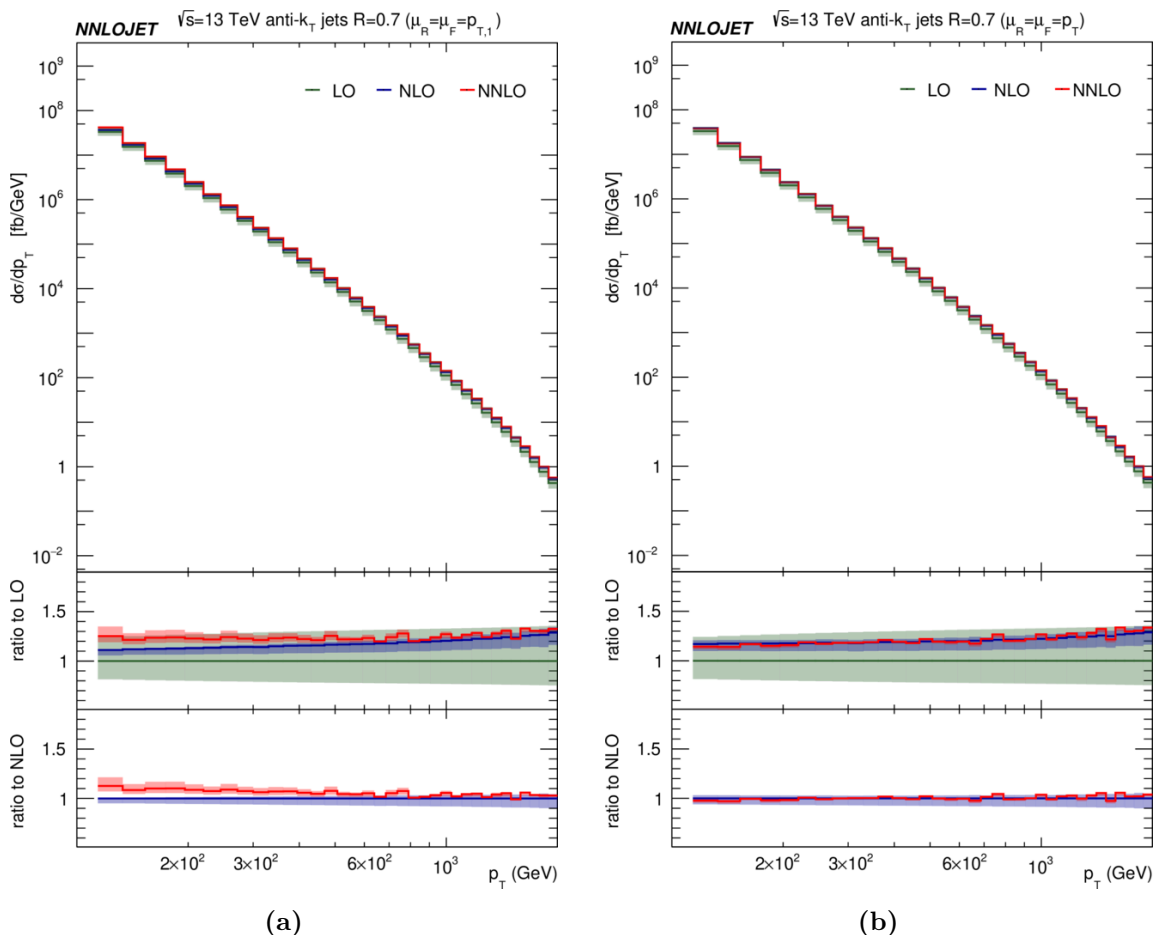
Figure 4. Perturbative corrections to the single jet inclusive distribution at 13 TeV (CMS cuts, $|y| < 4.7$, $R = 0.7$), integrated over rapidity and normalised to lower order predictions. Central scale choice: (a) $\mu = p_{T,1}$, (b) $\mu = p_T$.

higher order corrections to the single jet inclusive cross section obtained with the scale choice $\mu = p_{T,1}$ (left) and $\mu = p_T$ (right) for a fixed cone size $R = 0.7$. The top and bottom panels show respectively ratios of perturbative predictions to the LO and NLO results and the shaded bands represent the theoretical uncertainty estimated by varying the central scale choices by factors of 2 and $1/2$ and taking the envelope of the resulting cross sections. For all event-based scales in the remainder of this paper the variation includes doubling and halving the central value of the scale independently for μ_R and μ_F, with the constraint $1/2 \leq \mu_R/\mu_F \leq 2$. For jet-based scales the reevaluation of the event at several different scales is increasingly expensive to compute. For this reason we restrict the scale variation for event-based scales to 3-point symmetric μ_R, μ_F scale variations noting that the bulk of the scale dependence comes from μ_R variations and that no significant differences are observed with respect to a 7-point scale variation.

As expected, we can observe that at high p_T the NLO and NNLO effects are small and similar using either of the two central scale choices while more pronounced and different effects can be observed at low p_T. In the low p_T region we can observe larger NLO corrections with the scale $\mu = p_T$ than with the scale $\mu = p_{T,1}$, while one observes smaller

NNLO corrections with the $\mu = p_{\mathrm{T}}$ scale than with the scale $\mu = p_{\mathrm{T},1}$. As a result we see a faster convergence of the perturbative expansion when using the scale $\mu = p_{\mathrm{T}}$, where in particular the NNLO result lies inside the NLO scale uncertainty band, which itself lies inside the LO scale band. Furthermore, the scale uncertainty at NNLO displays a greater reduction for the scale choice $\mu = p_{\mathrm{T}}$.

It is instructive to compare what happens when using a smaller jet cone size. In this case, we fix the jet cone size to $R = 0.4$ and present the results in figure 5. Similarly to the $R = 0.7$ case, we observe identical higher order effects at high p_{T} between the two scale choices while more pronounced effects can be seen at low p_{T}. In this case, we observe that the NLO corrections using the central scale $\mu = p_{\mathrm{T}}$ are smaller than those corrections obtained for $R = 0.7$. The NLO scale uncertainty band is artificially small with the central scale choice sitting at the top of the band and the overlap between the NNLO result and the NLO scale band is no longer observed. Looking at the $\mu = p_{\mathrm{T},1}$ results, we observe an almost identical NLO scale band as for the results obtained with $R = 0.7$ and again non-overlapping NNLO and NLO scale bands. By comparing the NNLO/NLO K-factors for the two scale choices we observe that the NNLO cross section decreases (increases) with respect the NLO result with $\mu = p_{\mathrm{T}}$ ($\mu = p_{\mathrm{T},1}$). This is not unexpected since we have anticipated that for smaller jet cone sizes the effects of changing the scale from $\mu = p_{\mathrm{T}}$ to $\mu = p_{\mathrm{T},1}$ would be more pronounced. In particular, by comparing the NNLO/LO curves for $\mu = p_{\mathrm{T},1}$ and $\mu = p_{\mathrm{T}}$ (remembering that the LO result is identical for the two scale choices) we see that in the $R = 0.7$ case the NNLO predictions lie significantly closer to each other than for $R = 0.4$.

In order to demonstrate the last effect more clearly, figure 6 shows the ratios of the predictions at NLO (in blue) and NNLO (in red) for the two scale choices. We see that at NLO and NNLO, the impact of changing the scale from $\mu = p_{\mathrm{T},1}$ to $\mu = p_{\mathrm{T}}$ is more pronounced for the smaller jet size $R = 0.4$ (right) than it is for the larger jet size $R = 0.7$ (left). Interestingly, we also observe that for the two jet sizes, the impact of this change is bigger at NNLO than it is at NLO, which contradicts our expectation that the higher order corrections should lead to a smaller scale dependence.

It is worth noting that we do not necessarily expect that a change in the form of the central scale choice from $\mu = p_{\mathrm{T},1}$ to $\mu = p_{\mathrm{T}}$ can be captured by varying the values taken for renormalization and factorization scales in the predictions computed at a given fixed order. When the scale variation is performed, all the events are shifted simultaneously by a rescaling of the μ_{R} and μ_{F} scales. On the other hand, when we change the central scale from $\mu = p_{\mathrm{T},1}$ to $\mu = p_{\mathrm{T}}$, events with LO kinematics are unchanged while events with higher order kinematics can change significantly.

The renormalization group equations (see section 2) can be used to predict a change in the cross section due to a multiplication of the scales by a constant shift factor, but are otherwise unable to predict the behaviour of the cross section with another functional form for the central scale choice. For this reason we can expect the potentially different behaviour of the two scales used to compute IR sensitive observables (which are subject to delicate cancellations between real and virtual corrections) to be the underlying cause of the discrepancy in the results at NNLO between $\mu = p_{\mathrm{T},1}$ and $\mu = p_{\mathrm{T}}$.
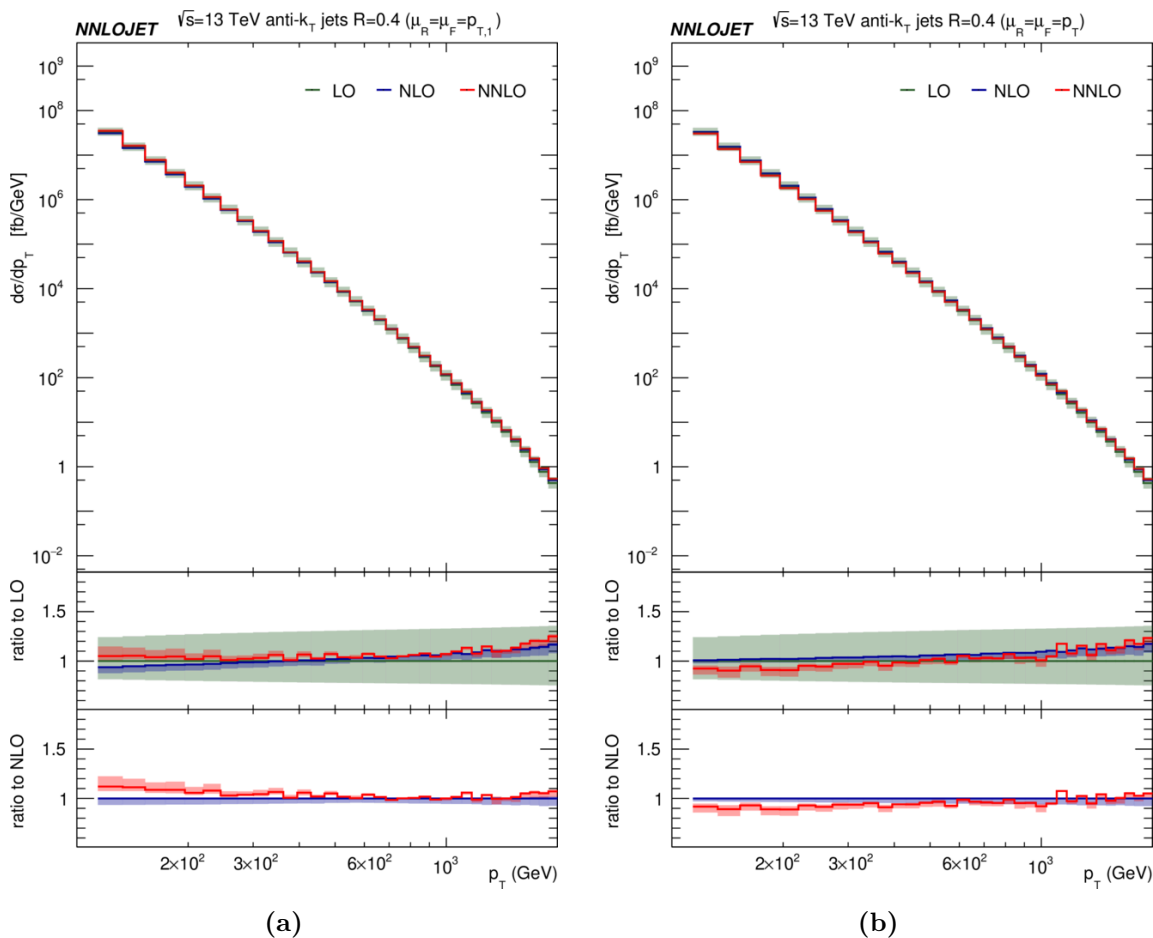
Figure 5. Perturbative corrections to the single jet inclusive distribution at $13\,\mathrm{TeV}$ (CMS cuts, $|y| < 4.7$, $R = 0.4$), integrated over rapidity and normalised to lower order predictions. Central scale choice: (a) $\mu = p_{\mathrm{T},1}$ and (b) $\mu = p_{\mathrm{T}}$.
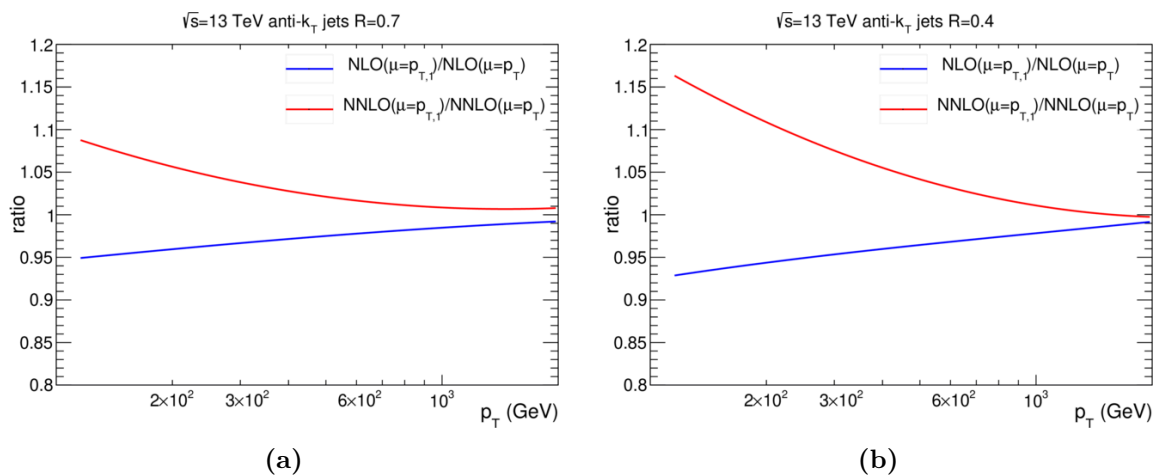

Figure 6. Ratio between predictions for the single jet inclusive distribution integrated over rapidity obtained with $\mu = p_{\mathrm{T},1}$ and $\mu = p_{\mathrm{T}}$ at NLO (blue) and NNLO (red) for different jet resolution: (a) $R = 0.7$; (b) $R = 0.4$.

3.3 Jet fractions in the single jet inclusive distribution

In order to explore this idea further, it is instructive to observe the breakdown of the single jet inclusive transverse momentum distribution into leading and subleading jet fractions, which is shown in figure 7 for the scale $\mu = p_{\mathrm{T},1}$ and jet sizes $R = 0.4$ (left) and $R = 0.7$ (right) at LO (top), NLO (middle) and NNLO (bottom). Beyond the trivial LO result, which as expected shows an equality between the first and second jet transverse momentum distributions, we observe interesting effects at higher orders. In particular, at NLO, we find that the leading jet contribution dominates the inclusive jet p_{T} spectrum for both jet sizes, while the contribution from the third jet is negligible. As expected for the larger cone size we produce more events with jets that are balanced in p_{T} and the jet fractions for the first and second jet are closer to the symmetric LO result. We can also identify a significant depletion of the second jet contribution in the NLO result for the jet cone size $R = 0.4$ at low p_{T} with the scale choice $\mu = p_{\mathrm{T},1}$. Finally, the NNLO results show a substantial increase in the second jet fraction for both jet sizes with respect to the NLO case, thereby coming closer to the LO result of similar-size first and second jet fractions.

With these results in mind, we can conclude that a small change in the second jet p_{T} distribution can have a potentially larger impact on the inclusive jet transverse momentum distribution at NNLO than at NLO, since the second jet contributes significantly more to the inclusive jet sample at NNLO than it does at NLO. It is therefore plausible that a change in scale from $\mu = p_{\mathrm{T},1}$ to $\mu = p_{\mathrm{T}}$ which affects the second jet p_{T} distribution produces a larger shift in the prediction of the inclusive jet p_{T} distribution at NNLO than it does at NLO (as shown in figure 6).

For comparison, figure 8 shows the corresponding jet fractions for the $\mu = p_{\mathrm{T}}$ scale choice and jet sizes $R = 0.4$ (left) and $R = 0.7$ (right) at LO (top), NLO (middle) and NNLO (bottom). As expected, when we compare with the results obtained with the scale $\mu = p_{\mathrm{T},1}$ we do not see significant differences in the jet fractions for the larger jet size of $R = 0.7$. On the other hand, for $R = 0.4$ we observe an increase in the second jet contribution at low p_{T} at NLO and a reduction in the same region at NNLO with respect to the results for $\mu = p_{\mathrm{T},1}$.

3.4 The second jet transverse momentum distribution

Given its potential impact on the scale uncertainty of the NNLO single jet inclusive cross section, we now focus our attention on the second jet transverse momentum distribution. Figure 9 shows the perturbative expansion of the second jet p_{T} distribution for the jet cone size of $R = 0.7$ with the scale choice $\mu = p_{\mathrm{T},1}$ (left) and $\mu = p_{\mathrm{T}}$ (right). For the two scale choices, we observe that this distribution is subject to very large perturbative corrections indicating potentially IR-sensitive effects. In particular we identify the presence of very large negative NLO corrections and large positive NNLO corrections generating an alternating series expansion with large coefficients. It is reassuring that the results at NNLO for the two scale choices are still largely identical despite this effect. We can nonetheless discern a significantly improved behaviour in the perturbative expansion when the scale
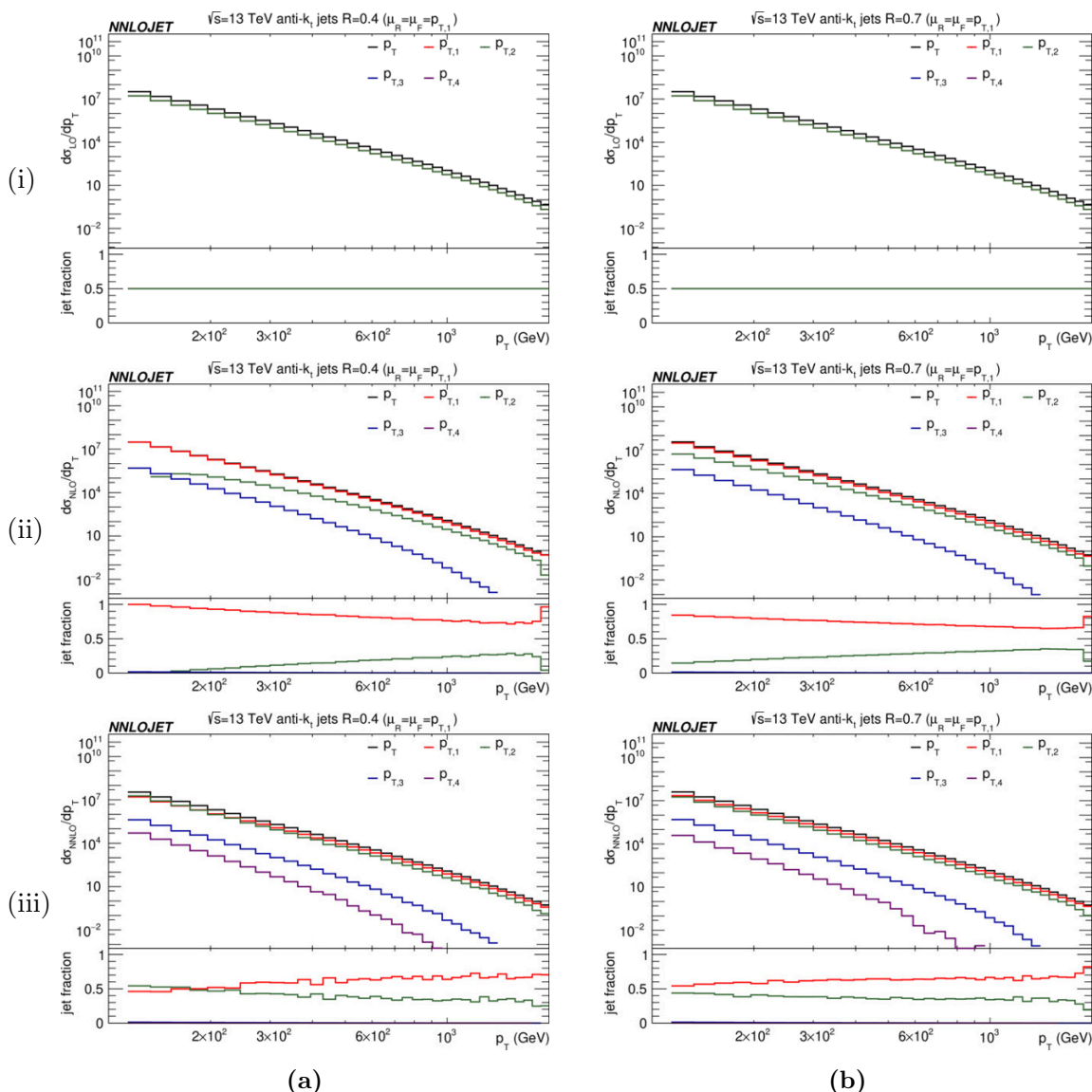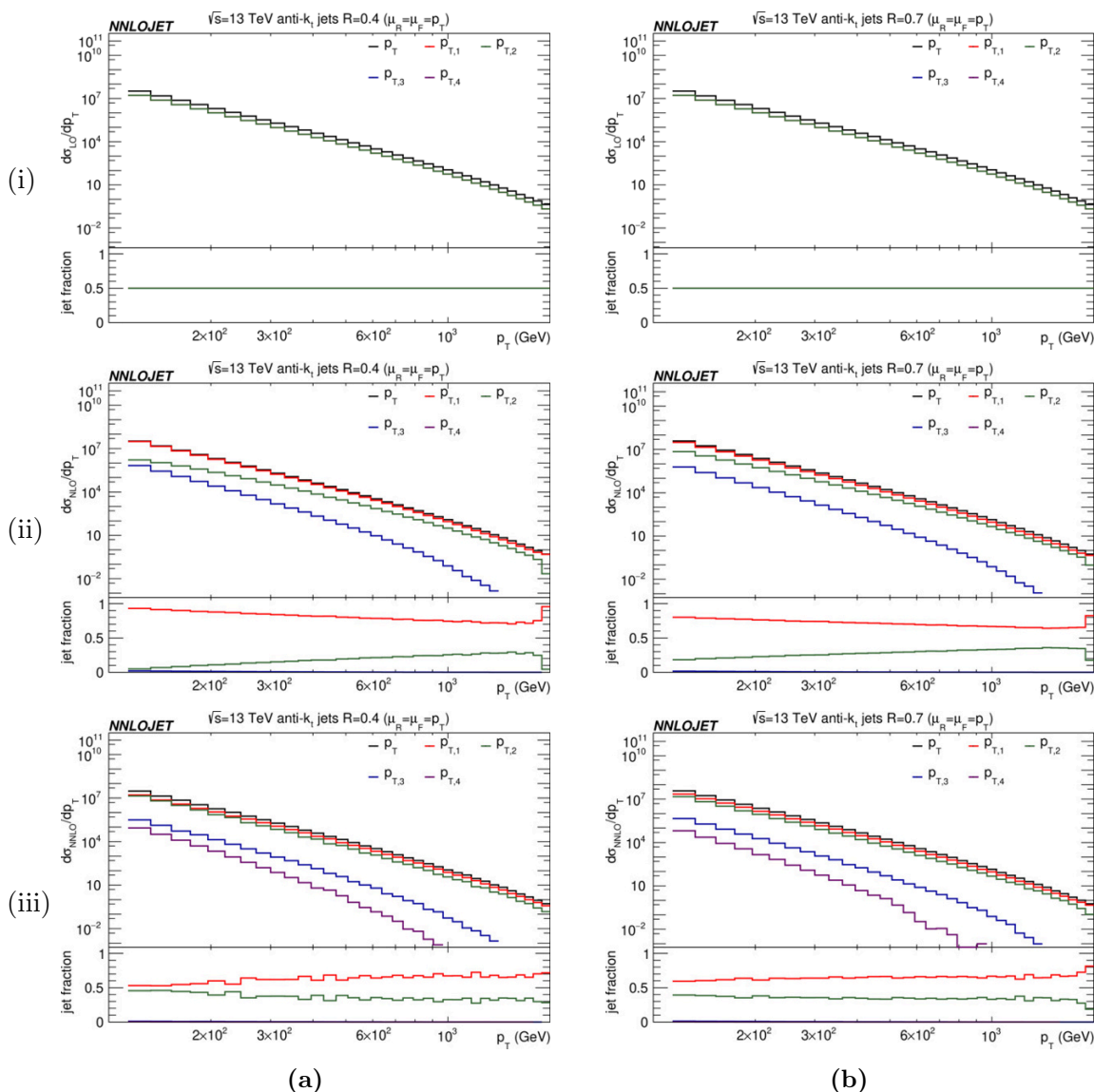
Figure 7. Breakdown of single jet inclusive cross section integrated over rapidity into contributions from first, second, third and fourth jet at (i) LO, (ii) NLO and (iii) NNLO evaluated for $\mu = p_{\mathrm{T},1}$ for jet cone sizes (a) $R = 0.4$ and (b) $R = 0.7$.

$\mu = p_{\mathrm{T}}$ is used. Both NLO and NNLO K-factors are significantly reduced and the NLO and NNLO scale uncertainty bands are also closer to each other for the $\mu = p_{\mathrm{T}}$ case.

The same behaviour can be observed for the smaller jet cone size of $R = 0.4$ in figure 10 where the sensitivity to IR effects is even more pronounced. In this case, we find a negative NLO cross section for the scale choice $\mu = p_{\mathrm{T},1}$ which is clearly exhibiting a pathological behaviour. The NNLO corrections fix this unphysical behaviour even when the scale $\mu = p_{\mathrm{T},1}$ is used, but similarly to the $R = 0.7$ case, we see a significantly better convergence of the perturbative series using $\mu = p_{\mathrm{T}}$ as the central scale choice.

Interestingly enough we observe also for both cone sizes, that in this contribution the NNLO scale band (in red) is larger than the LO scale band (in green). As explained in

Figure 8. Breakdown of single jet inclusive cross section integrated over rapidity into contributions from first, second, third and fourth jet at (i) LO, (ii) NLO and (iii) NNLO evaluated for $\mu = p_{\mathrm{T}}$ for jet cone sizes (a) $R = 0.4$ and (b) $R = 0.7$.

section 2.3, the notion of leading and subleading jet is not well defined at leading order ($p_{\mathrm{T},1} = p_{\mathrm{T},2}$ at LO) and for this reason the NLO result is the first non-trivial contribution sensitive to the difference in p_{T} between the leading and subleading jet. As such, the single jet inclusive observable is decomposed into IR-sensitive leading and subleading jet contributions and the functional form of the scale can have an impact on the final result, when the kinematics of the scale choice affects the IR cancellations between the different contributions.

In order to understand the source of the IR sensitivity in the second-jet contribution figure 11 shows the fractional contribution to the second jet p_{T} distribution in a given $p_{\mathrm{T},2}$ interval ($133\,\mathrm{GeV} < p_{\mathrm{T},2} < 153\,\mathrm{GeV}$) for particular $p_{\mathrm{T},1}$ slices plotted along the x-axis,
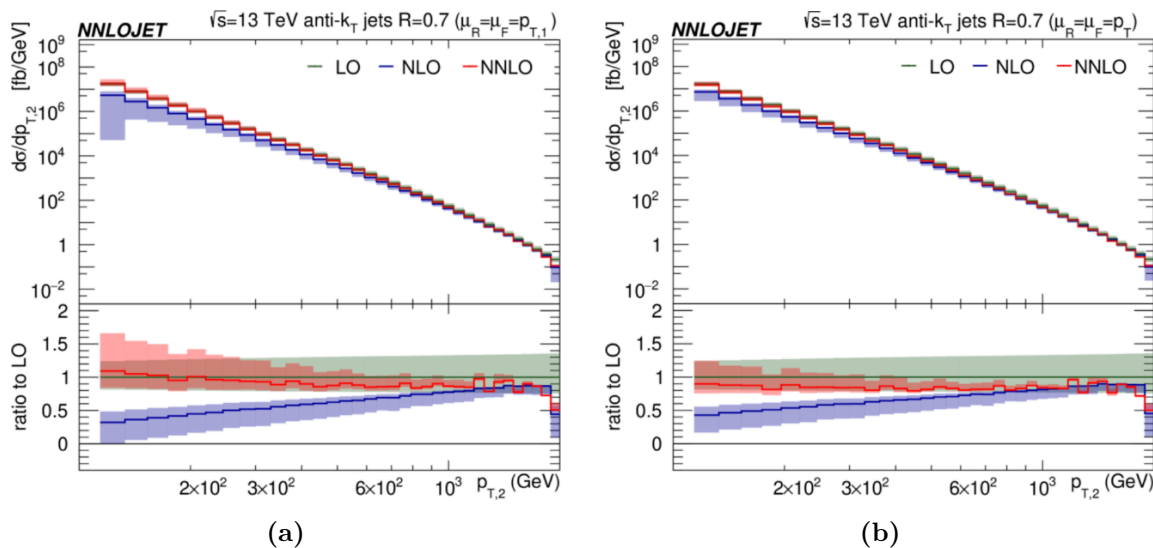
Figure 9. Perturbative corrections to the transverse momentum distribution of the second jet at 13 TeV (CMS cuts, $|y| < 4.7$, $R = 0.7$), integrated over rapidity and normalised to the LO prediction. Central scale choice: (a) $\mu = p_{T,1}$ and (b) $\mu = p_T$. Shaded bands represent the theory uncertainty due to the variation of the factorization and renormalization scales.
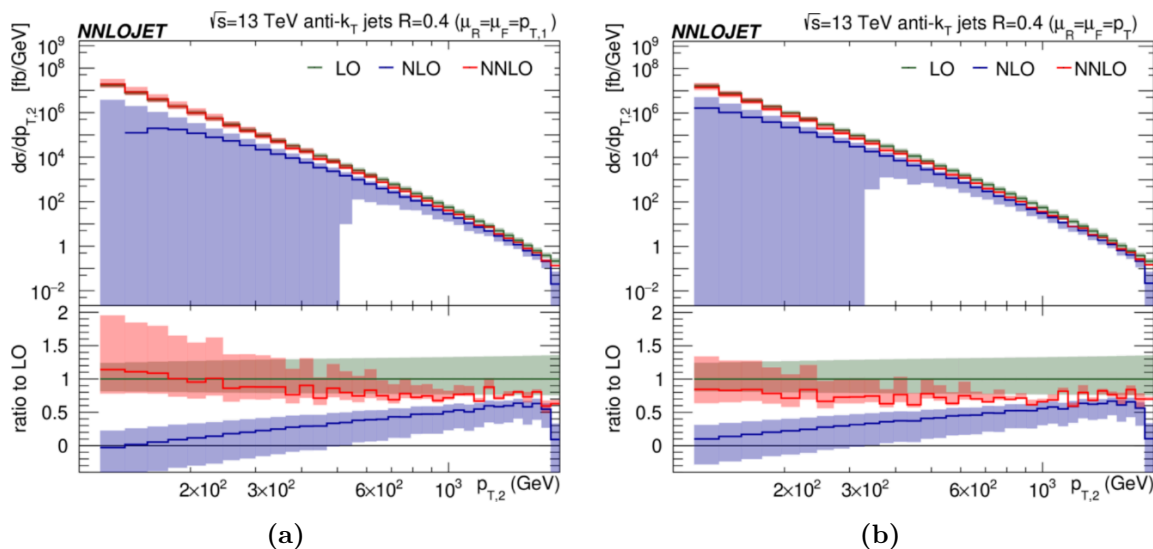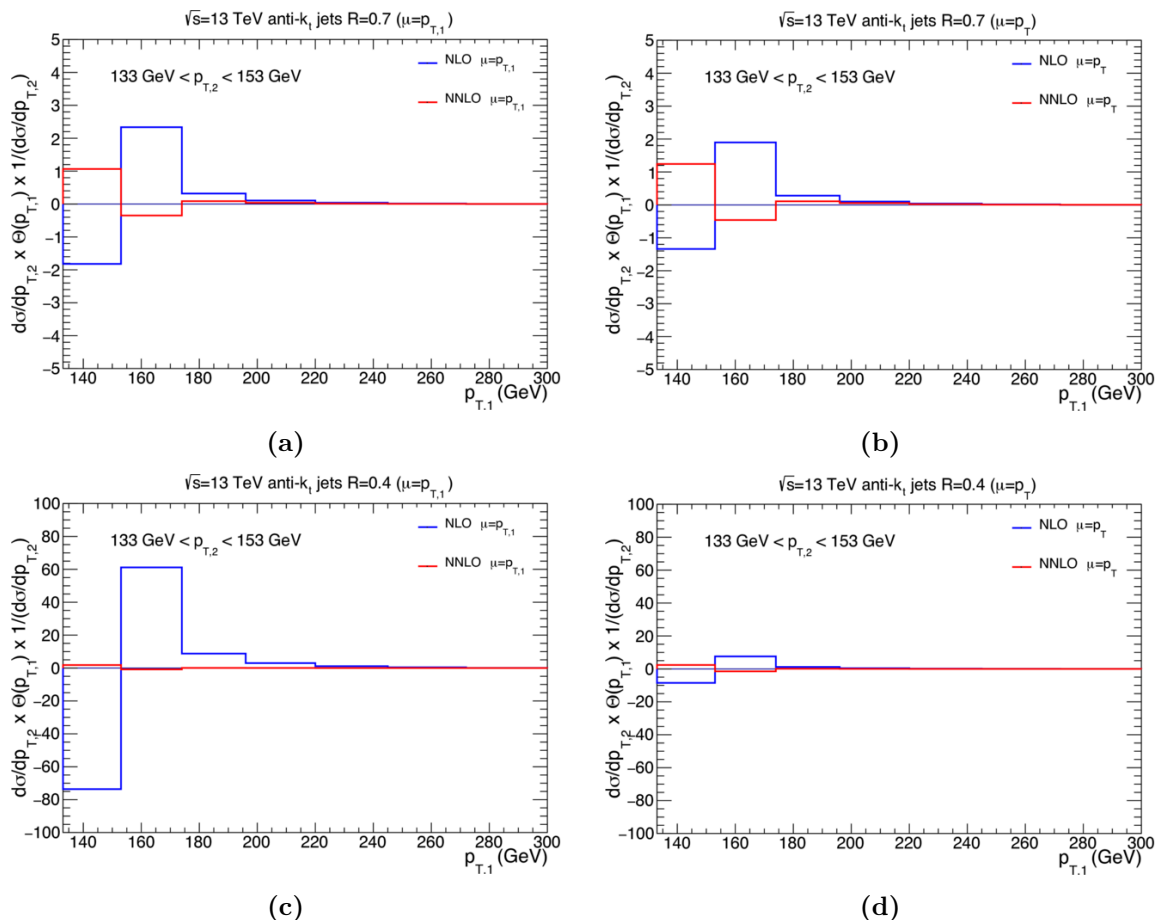

Figure 10. Perturbative corrections to the transverse momentum distribution of the second jet at 13 TeV (CMS cuts, $|y| < 4.7$, $R = 0.4$), integrated over rapidity and normalised to the LO prediction. Central scale choice: (a) $\mu = p_{T,1}$ and (b) $\mu = p_T$. Shaded bands represent the theory uncertainty due to the variation of the factorization and renormalization scales.

Figure 11. Decomposition of events contributing to a single bin in $p_{T,2}$ according to the transverse momentum of the leading jet in the event $p_{T,1}$. CMS cuts at 13 TeV with jet resolution $R = 0.7$ and scale choice (a) $\mu = p_{T,1}$; (b) $\mu = p_T$; jet resolution $R = 0.4$ and scale choice (c) $\mu = p_{T,1}$; (d) $\mu = p_T$;.

for either $\mu = p_{T,1}$ (left frames) or $\mu = p_T$ (right frames), and using $R = 0.7$ (upper frames) and $R = 0.4$ (lower frames). The bin content is constrained to sum to unity by construction. We observe that this is achieved from a large cancellation (for both scale choices) between the first bin of the distribution (where $p_{T,1} = p_{T,2}$) and the adjacent bin where ($p_{T,1} \gtrsim p_{T,2}$). In particular at NLO (in blue) the entire second bin content is filled from the NLO real emission (where $p_{T,1}$ can be larger than $p_{T,2}$ for the first time) while the virtual correction contributes to the first bin only. When comparing the behaviour of the two scale choices we note that for $\mu = p_{T,1}$ the scale is increasing along the x-axis. On the other hand for $\mu = p_T$, the scale is fixed to be equal to $p_{T,2}$ for all contributions and the cancellation between the large positive real emission and large negative virtual correction is improved (as shown by the height of the bins). This effect is even more pronounced for the $R = 0.4$ jet size as shown in the two lower frames of figure 11.

We can therefore infer that we observe an instability at higher order in the second jet p_T distribution when additional radiation is not recombined into the outgoing jet and

generates an imbalance between $p_{T,1}$ and $p_{T,2}$. In this case, relatively soft emissions do not outbalance fully with virtual corrections and large logarithms appear.[2] This effect has been observed to be particularly relevant for the smaller jet cone size distributions. After employing the $\mu = p_T$ scale choice we see an improved convergence for the second jet p_T distribution. The observed stabilisation for the jet-based scale $\mu = p_T$ as opposed to the event-based scale $\mu = p_{T,1}$ is at first sight counter-intuitive, as one should expect an event-based scale to lead to an improved infrared stability [38], since all contributions from a single parton-level event are evaluated at the same scale. The situation is somewhat different for the jet inclusive p_T distribution, since its infrared sensitivity stems only from the contribution from the second jet, which has implicit restrictions on its allowed phase space. If the second-jet cross section in a fixed kinematical bin is broken down according to the event properties that contribute to it, then the jet-based $\mu = p_T$ is a fixed scale, while the event-based $\mu = p_{T,1}$ becomes a dynamical scale.

We conclude that by employing the scale $\mu = p_T$ we improve the stability of the second jet transverse momentum distribution with respect to $\mu = p_{T,1}$ by improving the cancellation at fixed order between the real and virtual corrections. Since the leading jet $p_{T,1}$ contribution is identical with either $\mu = p_T$ and $\mu = p_{T,1}$, the single jet inclusive cross section is potentially more stable when using the jet based scale $\mu = p_T$.

4 Comparison of different scale choices

The renormalization and factorization scales are arbitrary dimensionful parameters and any scale is *a priori* an equally valid choice. Moreover, any ambiguity induced by different choices of the scales should ideally reduce as higher order terms in the perturbative expansion are included. As was shown in the previous section, however, the inclusive p_T distribution suffers from an infrared sensitivity that exhibits a strong dependence on the scale that is used and a suboptimal choice can introduce pathological behaviours in the predictions.

It is the aim of this section to go beyond the two scale choices $\mu = p_T$ and $\mu = p_{T,1}$ of the previous section and to study predictions for single jet inclusive production based on the comprehensive set of functional forms introduced in section 2.2. In particular, we will study the scale $\mu = \hat{H}_T$ and the appropriate scaling factor in front of the central scale choice. To this end, we introduce a set of criteria that define desirable properties for a suitable scale choice:

(a) *perturbative convergence*: we require that the size of the corrections reduces at each successive order in the perturbative expansion.

(b) *scale uncertainty as error estimate*: in order to have a reliable estimate of theory uncertainties due to missing higher-order corrections, we require overlapping scale-uncertainty bands between the last two orders, i.e. between the NLO and NNLO pre-

---

[2]The same instability is present in the inclusive dijet cross section. In that case, the use of asymmetric p_T cuts on the first and second jet's p_T increases the phase space available for soft emissions, suppressing the appearance of large effects from soft gluon emission.

dictions. Ideally, the central prediction with the highest accuracy should lie within the scale variation of the order that precedes it.

(c) *perturbative convergence of the individual jet spectra*: based on the observation of the previous section, where the p_T spectra of the individual jets receive large corrections with cancellations in the inclusive distribution, we further demand the convergence of the corrections to the individual $p_{\mathrm{T},1}$ and $p_{\mathrm{T},2}$ distributions.

(d) *stability of the second jet distribution*: the comparison between the scales $\mu = p_\mathrm{T}$ and $p_{\mathrm{T},1}$ has exposed the second jet distribution to be especially sensitive to the scale choice, sometimes even exhibiting unphysical behaviour where the scale variation predicts negative cross section. We therefore introduce an additional criterion based on the second jet distribution and its associated scale uncertainty and require the predictions to provide physical, positive cross sections.

In this way, a careful assessment of the behaviour of each scale can be made purely based on the behaviour of the predictions in perturbation theory, prior to any comparisons with experimental data (which are deferred to section 5).

Section 4.1 is devoted to a comparison of the different scales on the basis of the criteria defined above and identifying the choices that satisfy our requirements on transverse momentum distributions integrated over rapidity. It is the aim of this section to arrive at a sensible scale choice for single jet inclusive production. In section 4.2, we validate the optimal scale choices we made by further looking at the inclusive jet p_T distribution differentially in rapidity.

4.1 Assessment of the convergence criteria

In order to test the convergence criteria (a)–(d) defined in the introduction of this section on a more quantitative level, we define the following correction factors for the individual jet p_T spectra

$$\delta k_i^\mathrm{NLO} = \frac{\mathrm{d}\sigma^\mathrm{NLO}/\mathrm{d}p_{\mathrm{T},i}}{\mathrm{d}\sigma^\mathrm{LO}/\mathrm{d}p_\mathrm{T}} - 1\,,$$

$$\delta k_i^\mathrm{NNLO} = \frac{\mathrm{d}\sigma^\mathrm{NNLO}/\mathrm{d}p_{\mathrm{T},i}}{\mathrm{d}\sigma^\mathrm{NLO}/\mathrm{d}p_\mathrm{T}} - 1\,, \tag{4.1}$$

where $p_{\mathrm{T},i}$ denotes the i-th leading jet in the event. The K-factor for the inclusive jet distribution can be expressed in terms of the δk_i as follows,

$$K^\mathrm{NLO} = \frac{\mathrm{d}\sigma^\mathrm{NLO}/\mathrm{d}p_\mathrm{T}}{\mathrm{d}\sigma^\mathrm{LO}/\mathrm{d}p_\mathrm{T}} = 1 + \delta k_\Sigma^\mathrm{NLO}\,, \qquad \delta k_\Sigma^\mathrm{NLO} = \sum_{i\in\mathrm{jets}} \delta k_i^\mathrm{NLO}\,,$$

$$K^\mathrm{NNLO} = \frac{\mathrm{d}\sigma^\mathrm{NNLO}/\mathrm{d}p_\mathrm{T}}{\mathrm{d}\sigma^\mathrm{NLO}/\mathrm{d}p_\mathrm{T}} = 1 + \delta k_\Sigma^\mathrm{NNLO}\,, \qquad \delta k_\Sigma^\mathrm{NNLO} = \sum_{i\in\mathrm{jets}} \delta k_i^\mathrm{NNLO}\,, \tag{4.2}$$

and conditions (a), (c) are then given by

$$\begin{aligned}
&\text{(a)} \quad \left|\delta k_\Sigma^\mathrm{NNLO}\right| < \left|\delta k_\Sigma^\mathrm{NLO}\right|\,, \\
&\text{(c)} \quad \left|\delta k_i^\mathrm{NNLO}\right| < \left|\delta k_i^\mathrm{NLO}\right| \qquad \forall i\,.
\end{aligned}$$

Given that the measured single jet inclusive sample receives contributions predominantly from the two leading jets in the event, it is sufficient to test condition (c) only for $i = 1, 2$.

Figures 12 and 13 show the correction factors δk_1, δk_2, and δk_Σ for the set of scale choices of section 2.2 at NLO (solid lines) and NNLO (dashed lines) for the cone sizes $R = 0.7$ and $R = 0.4$, respectively. As anticipated, we find large cancellations between the leading (blue) and the subleading jet contributions (red) at each order in the perturbative expansion for any scale choice.

In particular, we can observe a very large negative (positive) NLO coefficient for the second (first) jet contributions in solid red (blue) respectively. This effect explains why the second jet contribution to the inclusive jet p_{T} sample at NLO is significantly reduced and the first jet fractions dominates (as shown in figures 7, 8). At the next order, the sign of the NNLO coefficient is reversed for the leading and subleading jet (dashed blue and red respectively), resulting in the leading and subleading fractions to become similar over the whole p_{T} range at NNLO.

Given that the aforementioned feature is common for all scale choices, we can now apply the criteria (a,c) to assess which scale choices show the most stable behaviour in the perturbative expansion. Criterion (c) is concerned with the *spread* of the blue and red curves, associated with δk_1 and δk_2, respectively. Going from NLO to NNLO, we require the size of the corrections to the individual jet p_{T} to become smaller and therefore that the dashed curves exhibit a smaller spread than the corresponding solid ones. We observe that for the scale choices $\mu = p_{\mathrm{T}}$, $\mu = p_{\mathrm{T},1}$, and $\mu = \hat{H}_{\mathrm{T}}/2$ this condition is not fulfilled, specially at low p_{T} and in particular for the smaller jet cone size $R = 0.4$.

The net effect on the inclusive p_{T} spectrum is given by the correction factors δk_Σ, shown as the black lines. With criterion (a), we require the K-factor at NNLO to be smaller than at NLO, i.e. the dashed black lines to be closer to zero than the solid ones. Here, we observe that the scales p_{T} and $p_{\mathrm{T},1}$ give rise to sizeable NNLO corrections that are larger in magnitude than the corresponding corrections at NLO for $R = 0.4$, while for the bigger cone size of $R = 0.7$ the same effect is observed to be true again for the $p_{\mathrm{T},1}$ scale choice. The remaining scales $2\,p_{\mathrm{T},1}$, $2\,p_{\mathrm{T}}$, $\hat{H}_{\mathrm{T}}$, and $\hat{H}_{\mathrm{T}}/2$ fulfil criterion (a), with NNLO corrections at the level of 5–10%.

In figure 14 we examine criterion (b) on the theory error estimate by plotting the predictions at a given order with their respective scale uncertainty bands normalised to the NLO prediction. Given the potentially large impact of the cone size, we present results for both $R = 0.7$ (top) and $R = 0.4$ (bottom). For both cone sizes we observe that the scale choices $p_{\mathrm{T},1}$ and $2\,p_{\mathrm{T},1}$ give rise to scale uncertainty bands at NLO and NNLO that do not overlap in the low-p_{T} region. For the scale p_{T}, the conclusion depends strongly on the cone size where we observe overlapping bands for $R = 0.7$ but not for $R = 0.4$. The remaining three scales $2\,p_{\mathrm{T}}$, $\hat{H}_{\mathrm{T}}$, and $\hat{H}_{\mathrm{T}}/2$, on the other hand, exhibit good convergence with overlapping scale uncertainty bands independently on the cone size.

Finally, we study criterion (d) by investigating the perturbative behaviour of the second jet distribution and its associated scale uncertainties. In figures 15 and 16 we show the corrections to the p_{T} distribution of the second jet for the cone sizes $R = 0.7$ and 0.4, respectively. As was already mentioned in the study of criterion (c), we clearly observe
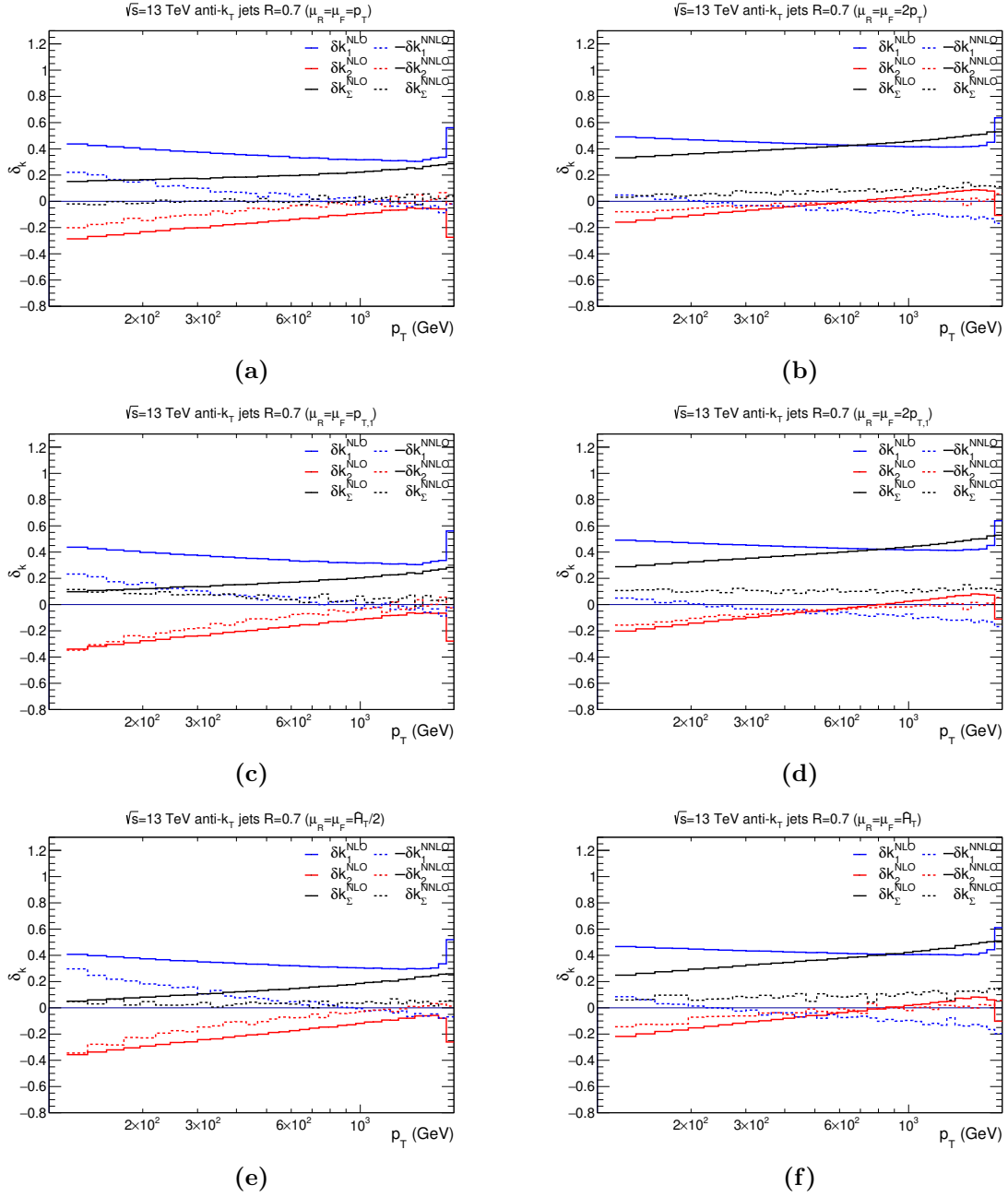
Figure 12. Differential correction factors at $\sqrt{s} = 13\,\text{TeV}$ for the leading jet ($\delta k_1^{\text{N}^k\text{LO}}$, blue), subleading jet ($\delta k_2^{\text{N}^k\text{LO}}$, red) and the inclusive jet distribution ($\delta k_\Sigma^{\text{N}^k\text{LO}}$, black) with $R = 0.7$ and integrated over rapidity for the scale choices (a) $\mu = p_\text{T}$, (b) $\mu = 2\,p_\text{T}$, (c) $\mu = p_{\text{T},1}$, (d) $\mu = 2\,p_{\text{T},1}$, (e) $\mu = \hat{H}_\text{T}/2$, (f) $\mu = \hat{H}_\text{T}$.
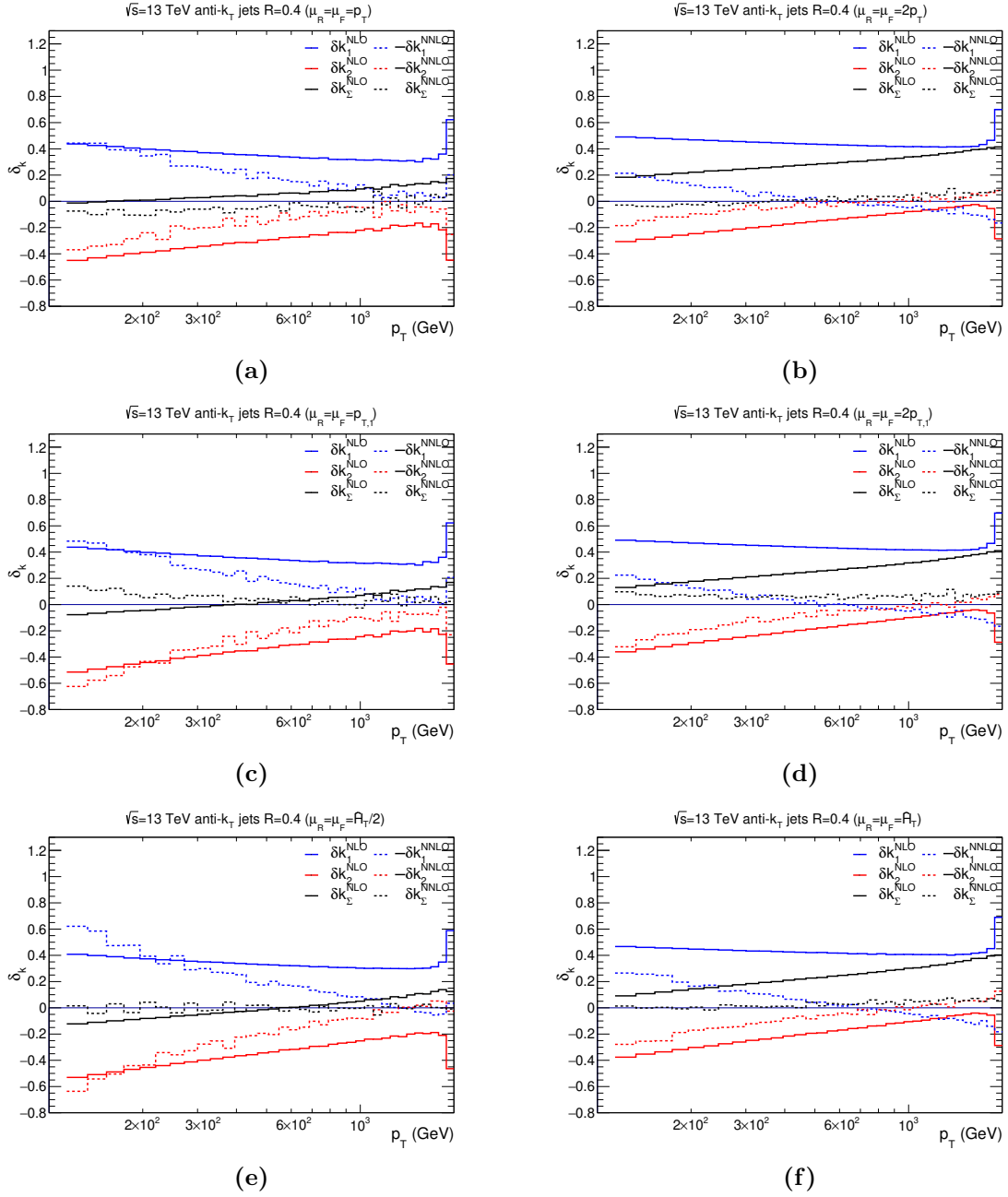
Figure 13. Differential correction factors at $\sqrt{s} = 13\,\mathrm{TeV}$ for the leading jet ($\delta k_1^{\mathrm{N}^k\mathrm{LO}}$, blue), subleading jet ($\delta k_2^{\mathrm{N}^k\mathrm{LO}}$, red) and the inclusive jet distribution ($\delta k_\Sigma^{\mathrm{N}^k\mathrm{LO}}$, black) with $R = 0.4$ and integrated over rapidity for the scale choices (a) $\mu = p_{\mathrm{T}}$, (b) $\mu = 2\,p_{\mathrm{T}}$, (c) $\mu = p_{\mathrm{T},1}$, (d) $\mu = 2\,p_{\mathrm{T},1}$, (e) $\mu = \hat{H}_{\mathrm{T}}/2$, (f) $\mu = \hat{H}_{\mathrm{T}}$.
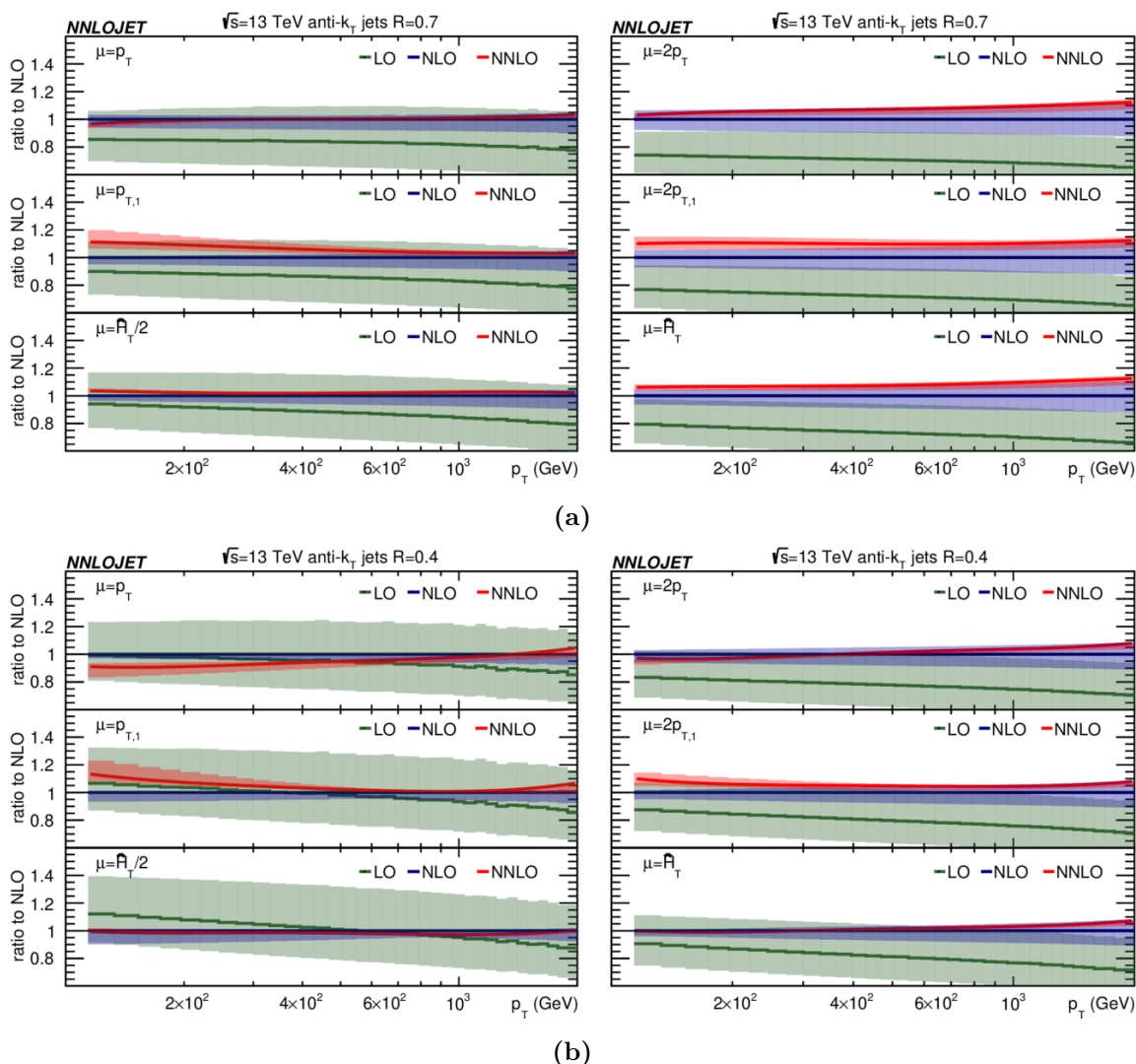
Figure 14. Inclusive jet p_T spectrum integrated over rapidity at LO (green), NLO (blue) and NNLO (red) normalised to the NLO prediction as a function of the central scale choice for (a) cone size $R = 0.7$ and (b) cone size $R = 0.4$.

an improved perturbative behaviour with smaller higher-order corrections and scale uncertainties for the three harder scale choices $2\,p_T$, $2\,p_{T,1}$, and $\hat{H}_T$ compared to their respective counterparts that are smaller by a factor of a half. For $R = 0.4$, we find that only the scale $\mu = 2\,p_T$ is able to predict positive NLO cross sections across the entire p_T range, both for the central value as well as for its variation. Although the scale choices $2\,p_{T,1}$ and $\hat{H}_T$ give rise to NLO scale uncertainties that extend to negative cross section values, this behaviour is less critical as it only occurs in the very first bin(s) below $p_T \lesssim 150\,\text{GeV}$ and the central predictions remain positive. The situation is much more severe for the remaining three scales p_T, $p_{T,1}$, and $\hat{H}_T/2$, where the NLO prediction exhibits the unphysical behaviour of negative cross sections already starting from $p_T \sim 400$–$600\,\text{GeV}$. In the case of $\mu = p_{T,1}$ and $\hat{H}_T/2$, even then central prediction turns negative in the lowest p_T

	criterion			
scale	(a)	(b)	(c)	(d)
$p_{\mathrm{T},1}$	—	—	✓	✓
$2\,p_{\mathrm{T},1}$	✓	—	✓	✓
p_{T}	—	✓	✓	✓
$2\,p_{\mathrm{T}}$	✓	✓	✓	✓
$\hat{H}_{\mathrm{T}}/2$	✓	✓	✓	—
$\hat{H}_{\mathrm{T}}$	✓	✓	✓	✓

(a) $R = 0.7$.

	criterion			
scale	(a)	(b)	(c)	(d)
$p_{\mathrm{T},1}$	—	—	—	—
$2\,p_{\mathrm{T},1}$	✓	—	✓	(✓)
p_{T}	—	—	—	—
$2\,p_{\mathrm{T}}$	✓	✓	✓	✓
$\hat{H}_{\mathrm{T}}/2$	✓	✓	—	—
$\hat{H}_{\mathrm{T}}$	✓	✓	✓	(✓)

(b) $R = 0.4$.

Table 2. Summary of scales vs. criteria for (a) $R = 0.7$ and (b) $R = 0.4$ cone sizes.

bin(s) below $\sim 150\,\mathrm{GeV}$. For the larger cone size of $R = 0.7$, on the other hand, the issue of negative cross sections at NLO is largely alleviated, where only the choice $\mu = \hat{H}_{\mathrm{T}}/2$ exhibits this unphysical behaviour.

We summarise the findings of this section in tables 2a, 2b for cone sizes of $R = 0.7$ and $R = 0.4$ respectively. By comparing the two tables we see that, as expected, the various scale choices behave in a much more similar way for the larger cone size than for $R = 0.4$. It is interesting to note that the two most commonly used scales $\mu = p_{\mathrm{T}}$ and $p_{\mathrm{T},1}$ perform by far the worst among the set of scale choices considered here. In particular, they are not able to fulfil any of the criteria for the smaller cone size of $R = 0.4$. On the other hand, the scale $\mu = 2\,p_{\mathrm{T}}$ fulfils all the requirements we identified at the beginning of this section, while the scale $\mu = \hat{H}_{\mathrm{T}}$ satisfies all of the criteria for $p_{\mathrm{T}} > 150\,\mathrm{GeV}$. We therefore identify $\mu = 2\,p_{\mathrm{T}}$ and $\mu = \hat{H}_{\mathrm{T}}$ as the two theoretically best-motivated scale choices for single jet inclusive production, noting that the former belongs to the class of jet-based scales and the latter is an event-based scale.

4.2 Results for central and forward rapidity slices

Having discussed at length the behaviour of the leading and subleading jet contributions as a function of the scale choice integrated over rapidity, for the remainder of this section, we will focus on the single jet inclusive observable for the different rapidity bin intervals used by the CMS collaboration [12].

Figure 17 shows the perturbative corrections for the single jet inclusive cross section at NLO and at NNLO for the same six scale choices discussed earlier: $\mu = p_{\mathrm{T},1}$, $\mu = p_{\mathrm{T}}$, $\mu = 2\,p_{\mathrm{T},1}$, $\mu = 2\,p_{\mathrm{T}}$, $\mu = \hat{H}_{\mathrm{T}}$, and $\mu = \hat{H}_{\mathrm{T}}/2$ for a jet cone size of $R = 0.7$ and for jets produced at (a) central rapidity ($|y| < 0.5$) and (b) forward rapidity ($2.5 < |y| < 3.0$). The shaded bands represent the scale variation around the respective central scale choice.

Focussing first on the central rapidity region shown in figure 17a, we see that the shape and size of the LO/NLO K-factor (green) for the $\mu = p_{\mathrm{T},1}$, $\mu = p_{\mathrm{T}}$ and $\mu = \hat{H}_{\mathrm{T}}/2$ scales are fairly similar. However, we observe larger NLO radiative corrections when these central choices are rescaled by a factor of 2.

Inspection of the NNLO/NLO K-factor (red) reveals that the size and shape of the NNLO corrections are generally smaller than the NLO ones, but that there is some depen-
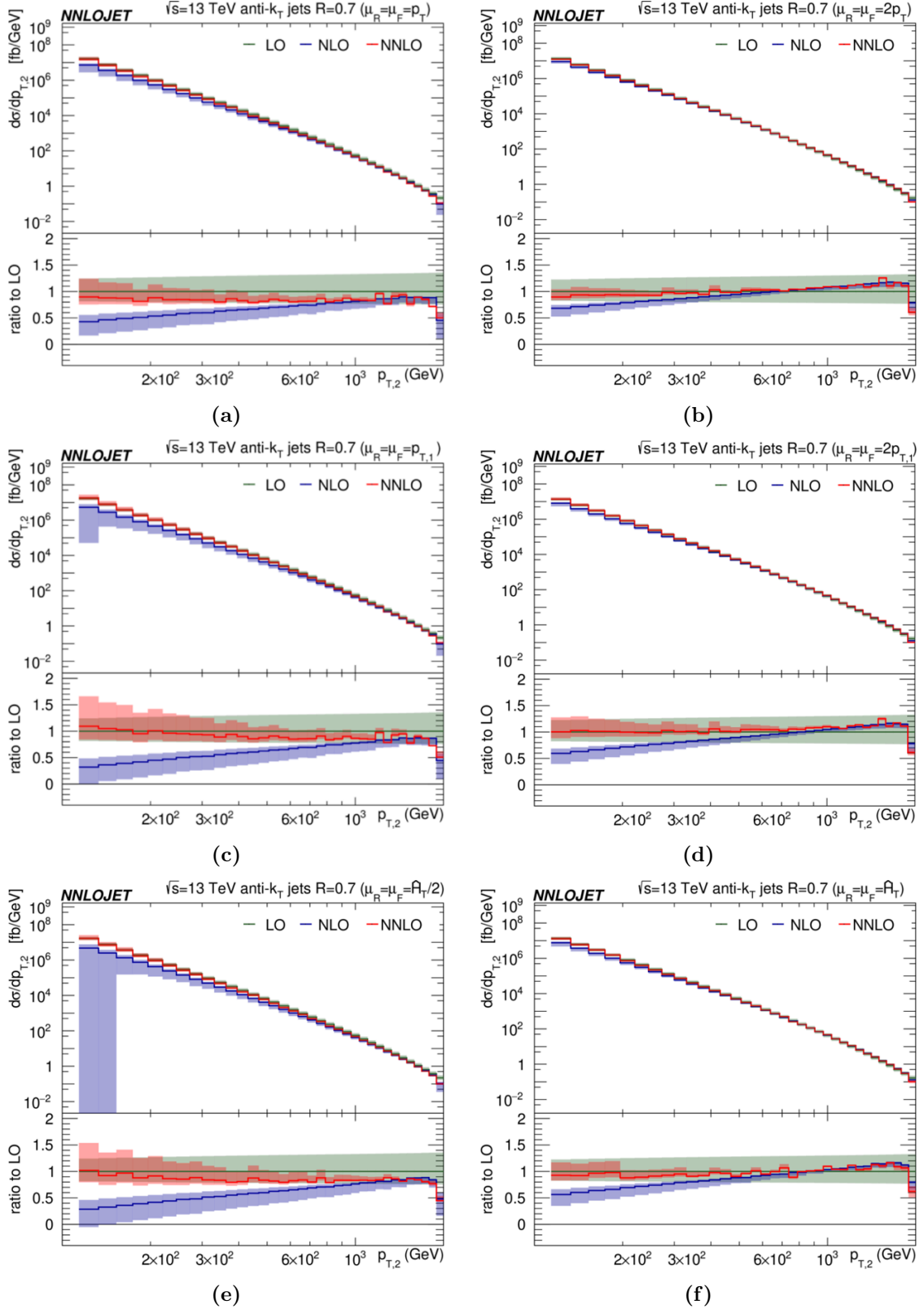
Figure 15. The transverse-momentum distribution of the subleading jet with $R = 0.7$ for (a) $\mu = p_{\mathrm{T}}$, (b) $\mu = 2\,p_{\mathrm{T}}$, (c) $\mu = p_{\mathrm{T,1}}$, (d) $\mu = 2\,p_{\mathrm{T,1}}$, (e) $\mu = \hat{H}_{\mathrm{T}}/2$, (f) $\mu = \hat{H}_{\mathrm{T}}$.
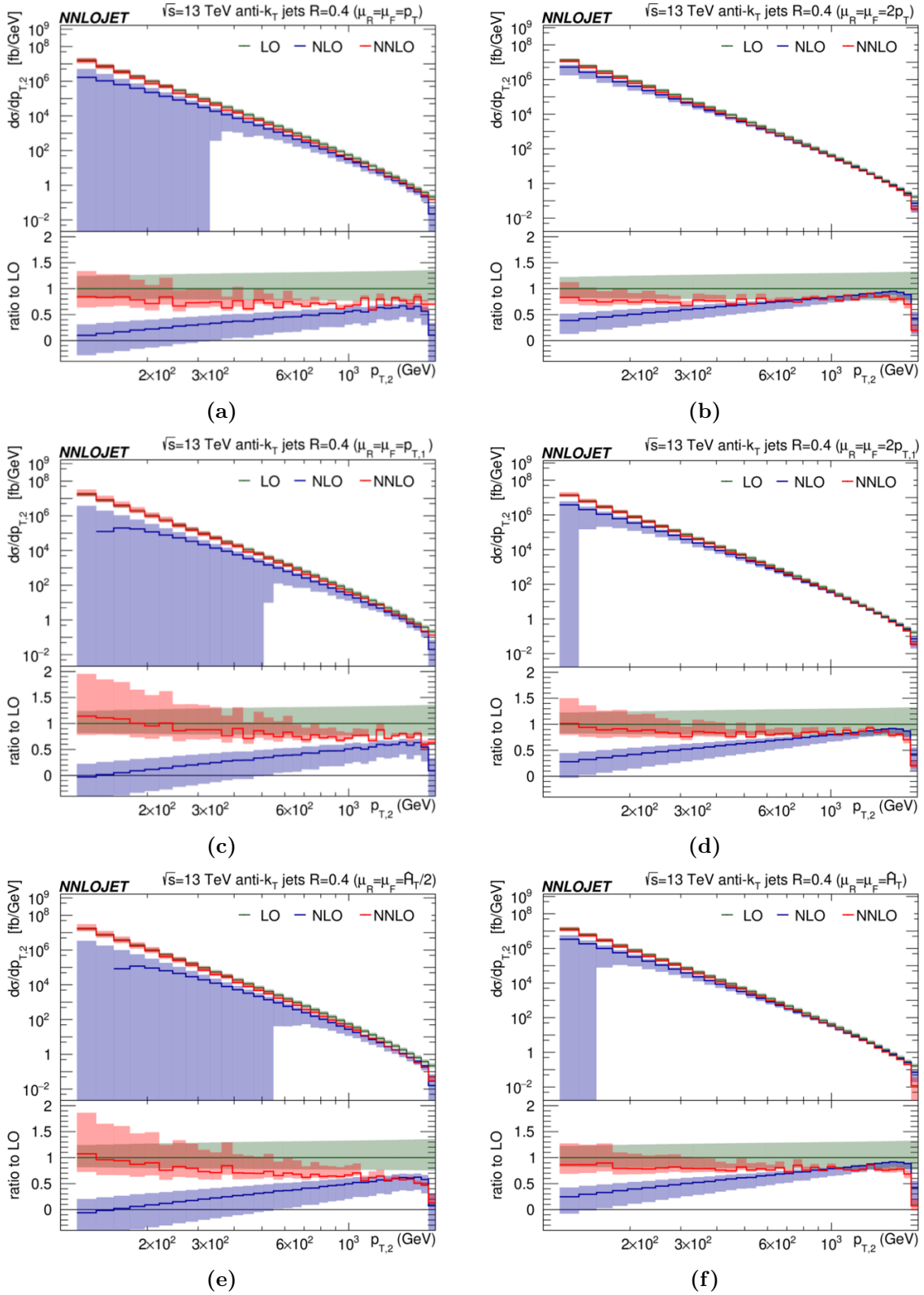
Figure 16. The transverse-momentum distribution of the subleading jet with $R = 0.4$ for (a) $\mu = p_{\mathrm{T}}$, (b) $\mu = 2\,p_{\mathrm{T}}$, (c) $\mu = p_{\mathrm{T},1}$, (d) $\mu = 2\,p_{\mathrm{T},1}$, (e) $\mu = \hat{H}_{\mathrm{T}}/2$, (f) $\mu = \hat{H}_{\mathrm{T}}$.
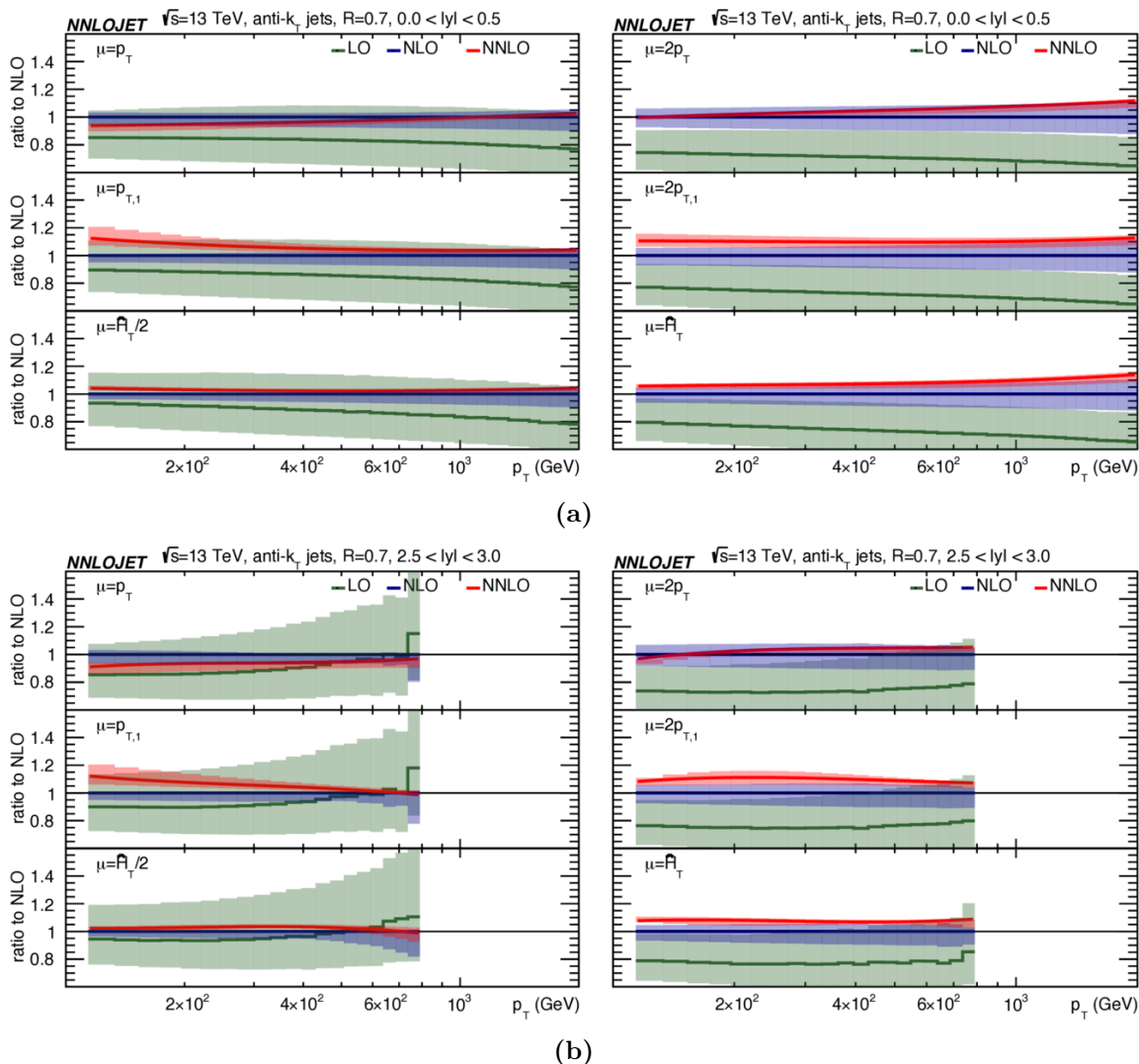
Figure 17. LO/NLO (green), NLO/NLO (blue) and NNLO/NLO (red) K-factors at $\sqrt{s} = 13\,\text{TeV}$ for (a) central rapidity, $0.0 < |y| < 0.5$, and (b) forward rapidity, $2.5 < |y| < 3.0$, as a function of the central scale choice for $R = 0.7$ and CMS cuts.

dence on the functional form of the scale choice. While the NNLO/NLO K-factor is never more than $\pm 20\%$ for any of the scale choices, the dependence on p_{T} is quite varied. For $\mu = p_{\text{T}}$ ($\mu = 2\,p_{\text{T}}$), the corrections grow from -10% (0%) at low p_{T} to a few percent (10%) at large p_{T}, while for $\mu = p_{\text{T,1}}$ ($\mu = 2\,p_{\text{T,1}}$), the corrections fall from $+15\%$ (12%) at low p_{T} to a few percent (10%) at large p_{T}. For $\mu = \hat{H}_{\text{T}}$, the corrections are always positive, growing from a few percent at low p_{T} to 12% at large p_{T}. In the case of $\mu = \hat{H}_{\text{T}}/2$, the NNLO/NLO K-factor is always small. The same qualitative behaviour can be observed in the predictions for jet production at forward rapidity ($2.5 < |y| < 3.0$), shown in figure 17b.

Because of the significantly different behaviour of the perturbative expansion for each scale choice, it is instructive to compare the respective absolute cross sections in the central rapidity region with a fixed normalisation. Figure 18a shows the NLO results for all six
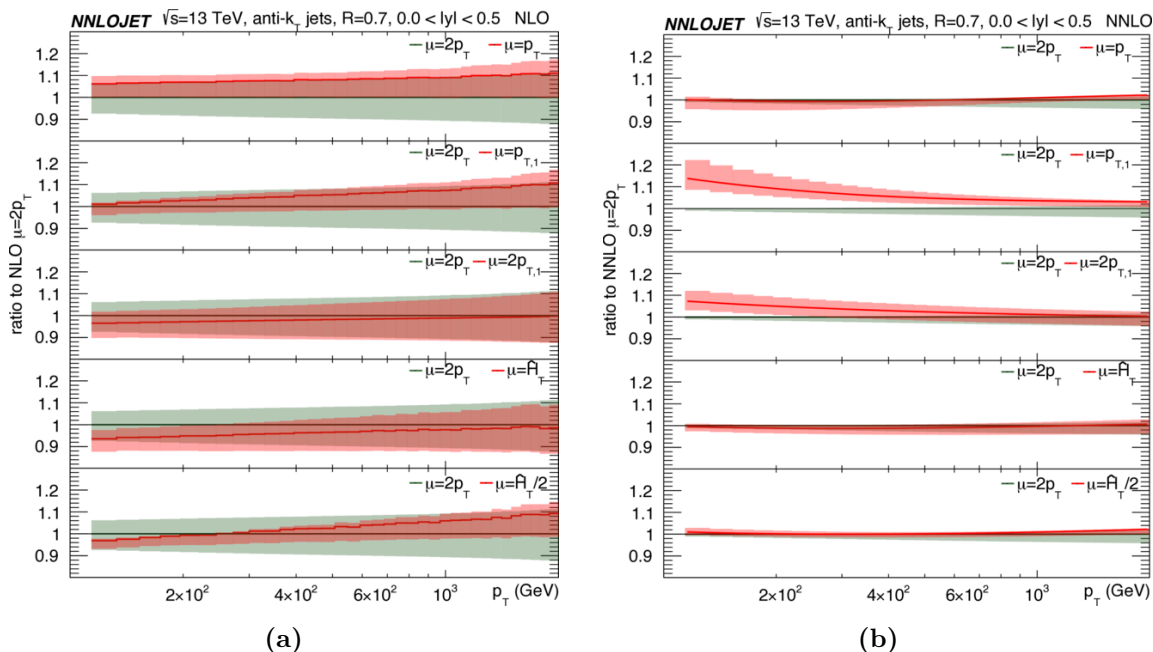
Figure 18. Ratio of 13 TeV single jet inclusive cross sections to the $\mu = 2\,p_\mathrm{T}$ scale choice at (a) NLO and (b) NNLO with $R = 0.7$ and CMS cuts.

scales normalised to one common NLO prediction, namely that for $\mu = 2\,p_\mathrm{T}$. For $R{=}0.7$ we see at NLO, that the scale uncertainty bands of the various NLO predictions (red band) are largely overlapping with the scale uncertainty for $\mu = 2\,p_\mathrm{T}$ (green band) indicating little scale choice ambiguity in the NLO predictions. In other words, the change in functional form of the scale choice is largely captured by the scale uncertainty of the NLO result.

Performing the same comparison for the different scale choices at NNLO in figure 18b and normalising to the NNLO prediction with $\mu = 2\,p_\mathrm{T}$, we observe the anticipated dramatic reduction in the scale variation with respect to NLO (as indicated by the reduction in the thickness of the red and green bands compared to figure 18a). We also conclude that the NNLO predictions are generally all in good agreement, particularly at high p_T, and independently of the scale choice. However, at low-p_T we do observe larger differences, where the scales $\mu = p_{\mathrm{T},1}$, $\mu = 2\,p_{\mathrm{T},1}$ tend to look similar and predict a larger NNLO cross section of approximately 10% with respect to the scales $\mu = p_\mathrm{T}$, $\mu = 2\,p_\mathrm{T}$, $\mu = \hat{H}_\mathrm{T}$, $\mu = \hat{H}_\mathrm{T}/2$. The size of this effect combined with a significant reduction in the scale uncertainty of the NNLO prediction introduces an ambiguity because the scale variation of the NNLO cross section no longer captures the predictions of different functional forms for the central scale choice. This has an important interplay with PDF extractions [42] using jet data and NNLO predictions, and also a significant impact when comparing the NNLO predictions with jet data.

We present the study of the perturbative corrections for the smaller jet cone size $R = 0.4$ in figure 19. Compared to the results with $R = 0.7$ (figure 17) we observe smaller NLO scale uncertainty bands and smaller NLO corrections. In particular at low-p_T, for $R = 0.4$ the central scale choices tend to sit at the upper edge of the band, and
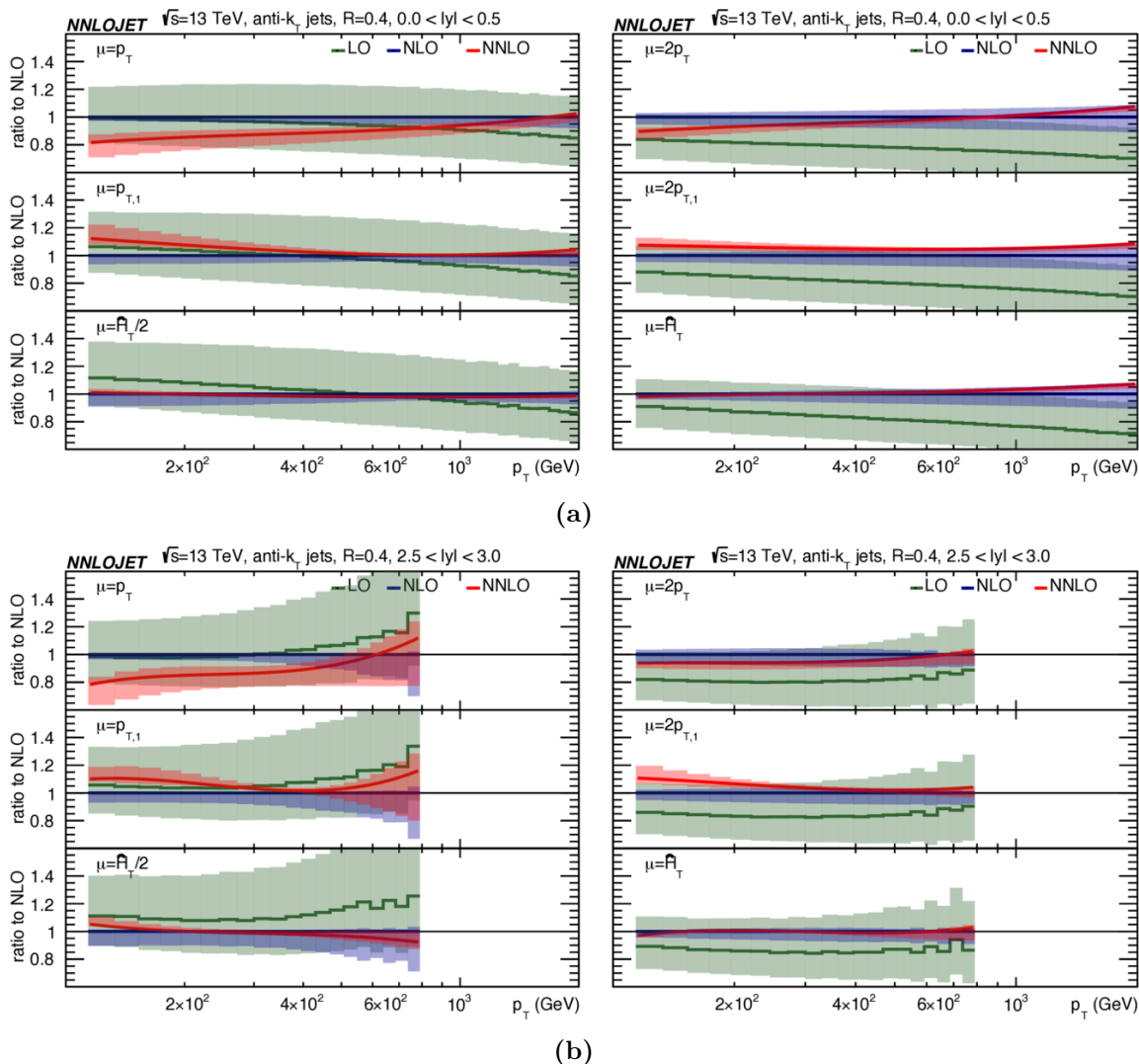
Figure 19. LO/NLO (green), NLO/NLO (blue) and NNLO/NLO (red) K-factors at $\sqrt{s} = 13\,\text{TeV}$ for (a) central rapidity, $0.0 < |y| < 0.5$, and (b) forward rapidity, $2.5 < |y| < 3.0$, as a function of the central scale choice for $R = 0.4$ and CMS cuts.

accidentally minimise the scale uncertainty. At NNLO there is a more symmetric scale variation, however the NNLO/NLO K-factors behave rather differently. The effect of the NNLO radiative corrections is positive for the scales $\mu = p_{\text{T},1}$, $\mu = 2\,p_{\text{T},1}$, negligible for the scales $\mu = \hat{H}_{\text{T}}$, $\mu = \hat{H}_{\text{T}}/2$, and negative for the scales $\mu = p_{\text{T}}$, $\mu = 2\,p_{\text{T}}$. As expected, the magnitude of the ambiguity in the scale choice for inclusive jet production is more severe for the smaller jet cone sizes.

The respective absolute cross sections at NLO and NNLO for all six scale choices compared to the prediction at the same order computed with $\mu = 2\,p_{\text{T}}$ for the jet size of $R{=}0.4$ are shown in figure 20a and 20b, respectively. As in figure 18, the red band reflects the scale uncertainty for the various scale choices, while the green band shows the scale uncertainty for $\mu = 2\,p_{\text{T}}$. As with the larger cone size, $R = 0.7$, figure 20a shows scale
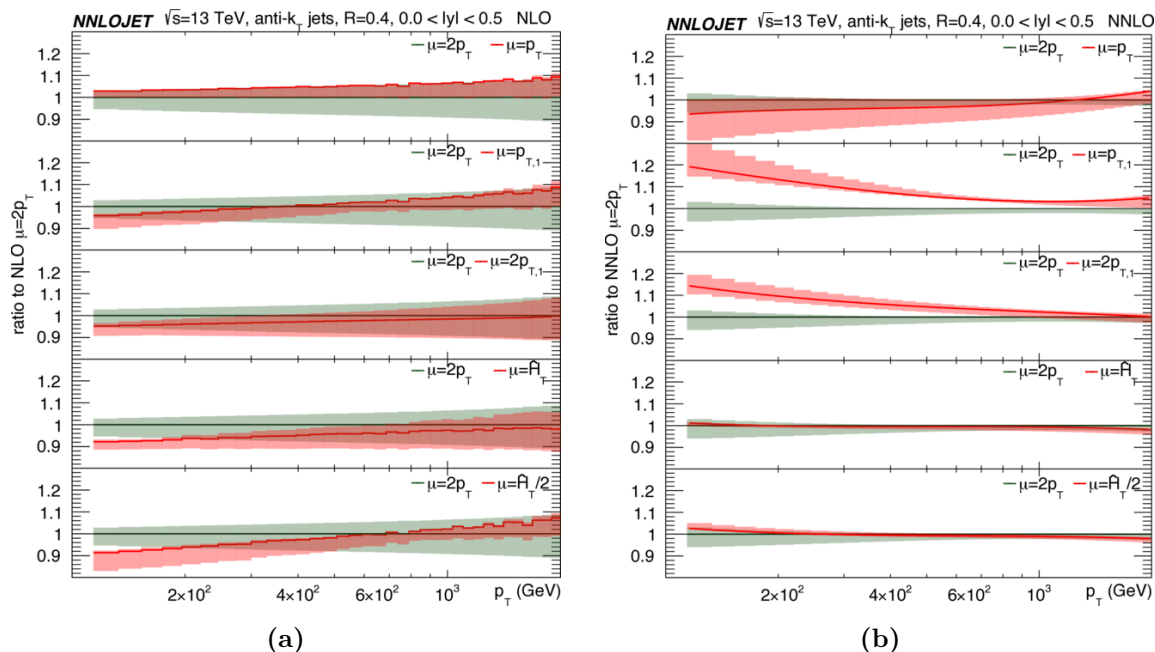
Figure 20. Ratio of 13 TeV single jet inclusive cross sections to the $\mu = 2\,p_T$ scale choice at (a) NLO and (b) NNLO with $R = 0.4$ and CMS cuts.

variations at NLO whose uncertainty largely captures the effects of changing the functional form of the scale. That is to say that the red and green bands largely overlap.

The same comparison is shown at NNLO in figure 20b where, as for $R = 0.7$, we observe a dramatic reduction in the scale variation with respect to NLO (as indicated by the reduction in the thickness of the red and green bands compared to figure 20a), except at low-p_T for the scale choices $\mu = p_T$, $\mu = p_{T,1}$ and $\mu = 2\,p_{T,1}$. For this reason, we can conclude that the NNLO predictions are generally in very good agreement at high p_T, independently of the scale choice. At low p_T we find larger differences, where in particular the scales $\mu = p_{T,1}$, $\mu = 2\,p_{T,1}$ tend to look similar and predict a larger NNLO cross section of approximately 15%–20% with respect to $\mu = p_T$, $\mu = 2\,p_T$, $\mu = \hat{H}_T$, $\mu = \hat{H}_T/2$.

The instability of the single jet inclusive cross section at low p_T has been thoroughly discussed in sections 3 and 4.1. Due to implicit restrictions on its kinematics, it was found that the contribution from the second jet in the event is particularly sensitive to higher order effects, and that the perturbative stability of the predictions can be improved to some extent by adopting sensible scale choice criteria. Moreover, the largest difference in cross section and NNLO scale uncertainty is associated to using either $\mu = p_T$, $\mu = p_{T,1}$ or $\mu = 2\,p_{T,1}$ as central scale choices. As documented in tables 2a, 2b, these scale choices introduce pathological behaviours in the perturbative expansion of the single jet inclusive observable. Since the spread in the NNLO predictions including these scale choices is larger in size than the NNLO scale variation, their inclusion (and associated pathological behaviours) is therefore overestimating the residual scale uncertainty at NNLO. It is therefore sensible to adopt well-motivated criteria for fixing the scale choice that best maximise the impact of the knowledge of the higher order QCD corrections to the observable, to the extent

that pathological behaviours are avoided. We have observed that the best perturbative stability can be obtained for $\mu = 2\,p_\text{T}$ or $\mu = \hat{H}_\text{T}$, where the perturbative convergence of the individual jet contributions is vastly improved with respect to the other functional forms of the scale choice. It is therefore not surprising that these scales tend to show smaller NNLO corrections and lead to smaller residual NNLO scale uncertainties.

In the remainder of this paper we will employ these two functional forms of the central scale choice to compare our predictions with jet data from the CMS dataset at $\sqrt{s} = 13\,\text{TeV}$ for the first time.

5 Comparison with CMS jet measurements at $\sqrt{s} = 13\,\text{TeV}$

Having discussed how the jet kinematics at the LHC differently affects each of the event-based and jet-based scale choices, in this section we present predictions for the double differential jet cross section at NLO and NNLO for the CMS measurement at $\sqrt{s} = 13\,\text{TeV}$ [12]. We use the same numerical setup as described in section 3.1 and do not include non-perturbative effects from underlying event and hadronization in our predictions. An assessment of the size of the non-perturbative contributions has been presented in [12] and we note that these can vary significantly with the jet p_T and the R cone size. In the study in [12] the non-perturbative corrections are expected to be negligible for $R = 0.4$ but can reach up to 10%–15% for $R = 0.7$ at low-p_T.

Figure 21 displays the NLO and NNLO predictions for the jet-based scale choice $\mu = 2\,p_\text{T}$, as well as for the event-based scale choice $\hat{H}_\text{T}$, compared to the CMS 13 TeV data [12] with a jet cone size of $R = 0.7$. For both scale choices we observe small positive NNLO corrections across all rapidity slices, that improve the agreement with the CMS data, as compared to the NLO prediction. In addition we identify a reduction in the scale uncertainty from NLO to NNLO across the entire p_T range.

Figure 22 shows the NLO and NNLO predictions for the smaller jet cone size of $R = 0.4$ (where non-perturbative corrections are expected to be less important than for $R = 0.7$ [12]). Similarly to the $R = 0.7$ case, we see that both scale choices provide reasonable predictions and that the agreement with data is improved at NNLO. For the $\mu = 2\,p_\text{T}$ scale choice this is achieved by having small negative NNLO corrections while for $\mu = \hat{H}_\text{T}$ the NNLO corrections are flat leading to a smaller residual scale variation at NNLO than for $\mu = 2\,p_\text{T}$.

6 Summary and conclusions

In this paper we have studied single jet inclusive production at hadron colliders and the jet transverse momentum distribution obtained by adding up the contributions from all jets that are observed in an event. Our predictions include the most up-to-date second order NNLO corrections in the perturbative expansion of the observable.

In detail we presented a breakdown of the inclusive jet-p_T sample into leading and sub-leading jet contributions and found large radiative corrections to the first and second jet contributions (that dominate the inclusive jet sample) that largely cancel each other. By
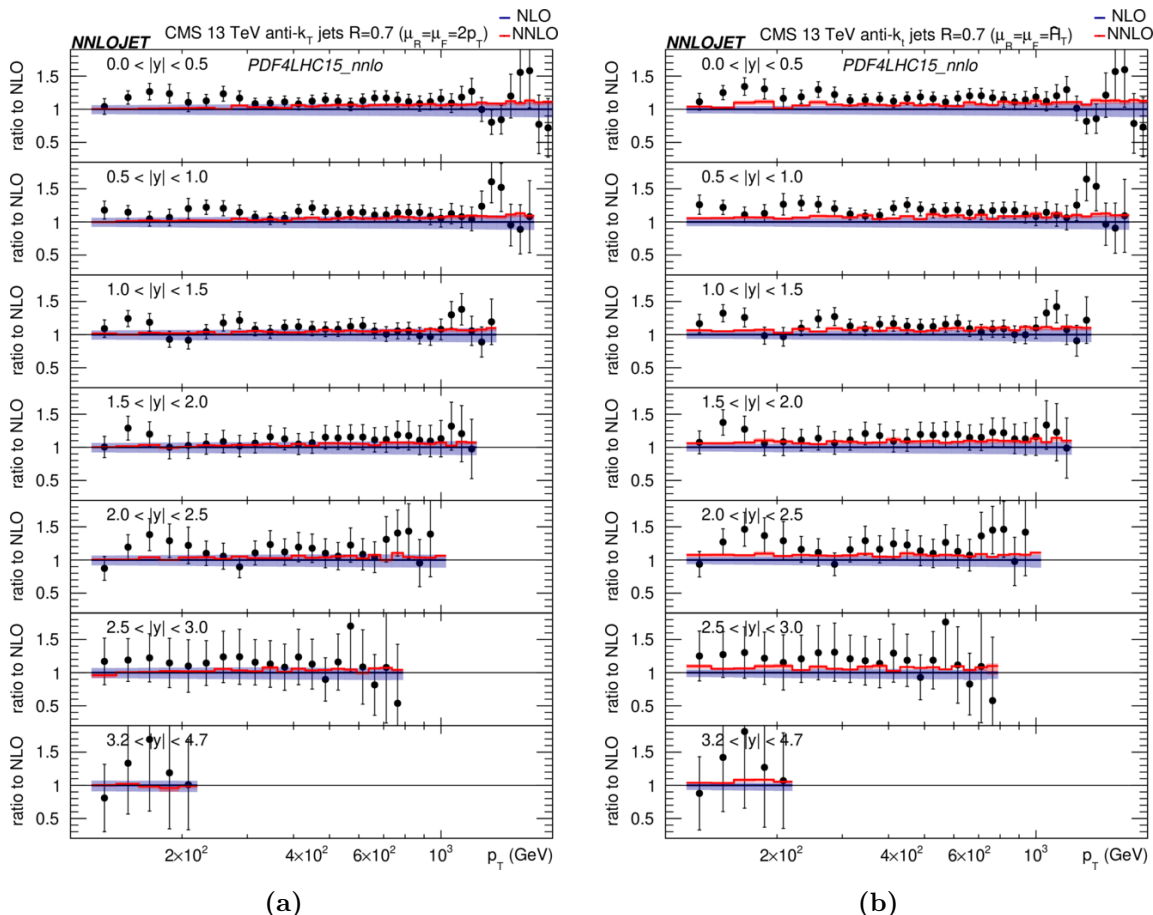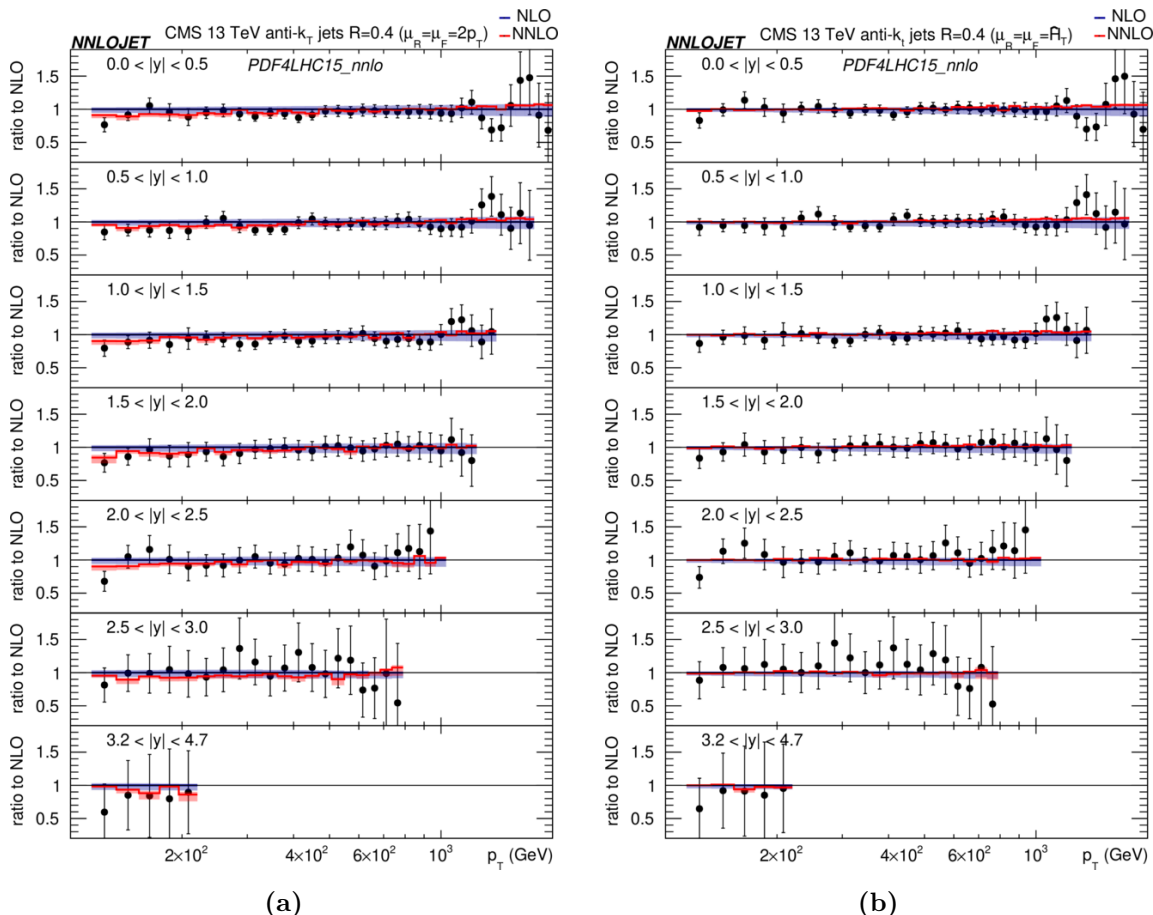
Figure 21. Double-differential single jet inclusive cross-sections measurement by CMS [12] and NNLO perturbative QCD predictions as a function of the jet p_T in slices of rapidity, for anti-k_T jets with $R = 0.7$ normalised to the NLO result for (a) $\mu = 2\,p_{\rm T}$, (b) $\mu = \hat{H}_{\rm T}$ scales. The shaded bands represent the scale uncertainty.

investigating the second-jet transverse momentum distribution we identified large cancellations between different kinematical event configurations, which are aggravated by certain types of scale choices. Since the notion of leading and subleading jet is not well defined at leading order ($p_{\rm T,1} = p_{\rm T,2}$ at LO), the single jet inclusive observable is decomposed into IR-sensitive leading and subleading jet contributions and the functional form of the scale can have an impact on the final result, when the kinematics of the scale choice affects the IR cancellations between the different contributions. We have found this effect to be worse for the smaller jet cone size $R = 0.4$ than for $R = 0.7$.

The smaller cone size increases the contribution from events where relatively soft emissions are not recombined with outgoing jets. These do not cancel fully with virtual corrections, leading to an imbalance between $p_{\rm T,1}$ and $p_{\rm T,2}$. Since the second jet contribution to the inclusive jet sample is increased at NNLO with respect to NLO we have identified this effect to be the cause in the mismatch between inclusive jet predictions at NNLO that employ $\mu = p_{\rm T}$ or $\mu = p_{\rm T,1}$ as central scale choices. By investigating the kinematical

Figure 22. Double-differential single jet inclusive cross-sections measurement by CMS [12] and NNLO perturbative QCD predictions as a function of the jet p_T in slices of rapidity, for anti-k_T jets with $R = 0.4$ normalised to the NLO result for (a) $\mu = 2\,p_T$, (b) $\mu = \hat{H}_T$ scales. The shaded bands represent the scale uncertainty.

properties of events that contribute to a fixed bin in $p_{T,2}$ (as function of $p_{T,1}$ in the event), we found that the imbalance between real and virtual emissions is much more serious for $\mu = p_{T,1}$ than for $\mu = p_T$, which can be understood from the fact that the former is changing event-by-event in this distribution, while the latter remains constant.

We have observed that the spread in the NNLO predictions that use different functional forms of the scale is larger in size than the NNLO scale variation in the low-p_T region of the transverse momentum distribution. For this reason we have introduced a sensible set of criteria that define desired properties for a suitable scale choice and that can maximise the impact of the knowledge of the higher order QCD corrections to the observable, to the extent that scale choices that introduce pathological behaviours can be identified and avoided.

We have identified $\mu = 2\,p_T$ and $\mu = \hat{H}_T$ as the two scales that fulfil all the criteria that we have defined, observing that they lead to important cancellations between the leading and subleading jet contributions which result in an improved perturbative convergence on the transverse momentum distributions, with overlapping scale uncertainty bands.

Subsequently we used these two functional forms of the central scale choice to compare our NNLO predictions with jet data from the CMS dataset at $\sqrt{s} = 13\,\text{TeV}$ [12] for the first time. We have observed that both recommended scale choices are stable and provide reasonable predictions for the two jet cone sizes employed in the measurement across the entire p_T and rapidity region where the observable is defined. In particular we find an improved agreement with data at NNLO with respect to NLO with a significant reduction in scale uncertainty by roughly more than a factor of 2 in a wide range of p_T and rapidity. We have refrained from comparing the measurement to predictions that employ scale choices that contain pathological behaviours since these scale choices are not recommended on the grounds introduced in this study.

The central scale choices $\mu = 2\,p_\text{T}$ and $\mu = \hat{H}_\text{T}$ are clearly found to be favoured in terms of stability and convergence of the predictions for single jet inclusive production. Both yield very similar predictions at NNLO. We expect that our findings will enable improved precision studies based on single jet inclusive production data, especially in using them as precision probes of the parton distributions in the proton and for a determination of QCD parameters.

Acknowledgments

References

[1] UA1 collaboration, G. Arnison et al., *Measurement of the Inclusive Jet Cross-Section at the CERN $p\bar{p}$ Collider*, *Phys. Lett.* **B 172** (1986) 461 [INSPIRE].

[2] UA2 collaboration, M. Banner et al., *Observation of Very Large Transverse Momentum Jets at the CERN $\bar{p}p$ Collider*, *Phys. Lett.* **B 118** (1982) 203 [INSPIRE].

[3] CDF collaboration, T. Aaltonen et al., *Measurement of the Inclusive Jet Cross Section at the Fermilab Tevatron $p\bar{p}$ Collider Using a Cone-Based Jet Algorithm*, *Phys. Rev.* **D 78** (2008) 052006 [*Erratum ibid.* **D 79** (2009) 119902] [arXiv:0807.2204] [INSPIRE].

[4] D0 collaboration, V.M. Abazov et al., *Measurement of the inclusive jet cross section in $p\bar{p}$ collisions at $\sqrt{s} = 1.96\,TeV$*, *Phys. Rev.* **D 85** (2012) 052006 [arXiv:1110.3771] [INSPIRE].

[5] ALICE collaboration, *Measurement of the inclusive differential jet cross section in pp collisions at $\sqrt{s} = 2.76\,TeV$*, *Phys. Lett.* **B 722** (2013) 262 [arXiv:1301.3475] [INSPIRE].

[6] ATLAS collaboration, *Measurement of the inclusive jet cross section in pp collisions at $\sqrt{s} = 2.76\,TeV$ and comparison to the inclusive jet cross section at $\sqrt{s} = 7\,TeV$ using the ATLAS detector*, *Eur. Phys. J.* **C 73** (2013) 2509 [arXiv:1304.4739] [INSPIRE].

[7] CMS collaboration, *Measurement of the inclusive jet cross section in pp collisions at $\sqrt{s} = 2.76\,TeV$*, *Eur. Phys. J.* **C 76** (2016) 265 [arXiv:1512.06212] [INSPIRE].

[8] ATLAS collaboration, *Measurement of inclusive jet and dijet cross sections in proton-proton collisions at 7 TeV centre-of-mass energy with the ATLAS detector*, *Eur. Phys. J.* **C 71** (2011) 1512 [arXiv:1009.5908] [INSPIRE].

[9] CMS collaboration, *Measurements of differential jet cross sections in proton-proton collisions at $\sqrt{s} = 7\,TeV$ with the CMS detector*, *Phys. Rev.* **D 87** (2013) 112002 [*Erratum ibid.* **D 87** (2013) 119902] [arXiv:1212.6660] [INSPIRE].

[10] ATLAS collaboration, *Measurement of the inclusive jet cross-sections in proton-proton collisions at $\sqrt{s} = 8\,TeV$ with the ATLAS detector*, *JHEP* **09** (2017) 020 [arXiv:1706.03192] [INSPIRE].

[11] CMS collaboration, *Measurement and QCD analysis of double-differential inclusive jet cross sections in pp collisions at $\sqrt{s} = 8\,TeV$ and cross section ratios to 2.76 and 7 TeV*, *JHEP* **03** (2017) 156 [arXiv:1609.05331] [INSPIRE].

[12] CMS collaboration, *Measurement of the double-differential inclusive jet cross section in proton-proton collisions at $\sqrt{s} = 13\,TeV$*, *Eur. Phys. J.* **C 76** (2016) 451 [arXiv:1605.04436] [INSPIRE].

[13] ATLAS collaboration, *Measurement of inclusive jet and dijet cross-sections in proton-proton collisions at $\sqrt{s} = 13\,TeV$ with the ATLAS detector*, *JHEP* **05** (2018) 195 [arXiv:1711.02692] [INSPIRE].

[14] A.J. Larkoski, I. Moult and B. Nachman, *Jet Substructure at the Large Hadron Collider: A Review of Recent Advances in Theory and Machine Learning*, arXiv:1709.04464 [INSPIRE].

[15] S.D. Ellis, Z. Kunszt and D.E. Soper, *Two jet production in hadron collisions at order α_s^3 in QCD*, *Phys. Rev. Lett.* **69** (1992) 1496 [INSPIRE].

[16] W.T. Giele, E.W.N. Glover and D.A. Kosower, *The Two-Jet Differential Cross Section at $\mathcal{O}(\alpha_s^3)$ in Hadron Collisions*, *Phys. Rev. Lett.* **73** (1994) 2019 [hep-ph/9403347] [INSPIRE].

[17] Z. Nagy, *Three jet cross-sections in hadron hadron collisions at next-to-leading order*, *Phys. Rev. Lett.* **88** (2002) 122003 [hep-ph/0110315] [INSPIRE].

[18] S. Alioli, K. Hamilton, P. Nason, C. Oleari and E. Re, *Jet pair production in POWHEG*, *JHEP* **04** (2011) 081 [arXiv:1012.3380] [INSPIRE].

[19] S. Höche and M. Schönherr, *Uncertainties in next-to-leading order plus parton shower matched simulations of inclusive jet and dijet production*, *Phys. Rev.* **D 86** (2012) 094042 [arXiv:1208.2815] [INSPIRE].

[20] S. Dittmaier, A. Huss and C. Speckner, *Weak radiative corrections to dijet production at hadron colliders*, *JHEP* **11** (2012) 095 [arXiv:1210.0438] [INSPIRE].

[21] J.M. Campbell, D. Wackeroth and J. Zhou, *Study of weak corrections to Drell-Yan, top-quark pair and dijet production at high energies with MCFM*, *Phys. Rev.* **D 94** (2016) 093009 [arXiv:1608.03356] [INSPIRE].

[22] R. Frederix, S. Frixione, V. Hirschi, D. Pagani, H.-S. Shao and M. Zaro, *The complete NLO corrections to dijet hadroproduction*, *JHEP* **04** (2017) 076 [arXiv:1612.06548] [INSPIRE].

[23] X. Liu, S.-O. Moch and F. Ringer, *Phenomenology of single-inclusive jet production with jet radius and threshold resummation*, *Phys. Rev.* **D 97** (2018) 056026 [arXiv:1801.07284] [INSPIRE].

[24] E.W. Nigel Glover and J. Pires, *Antenna subtraction for gluon scattering at NNLO*, *JHEP* **06** (2010) 096 [arXiv:1003.2824] [INSPIRE].

[25] A. Gehrmann-De Ridder, E.W.N. Glover and J. Pires, *Real-Virtual corrections for gluon scattering at NNLO*, *JHEP* **02** (2012) 141 [arXiv:1112.3613] [INSPIRE].

[26] A. Gehrmann-De Ridder, T. Gehrmann, E.W.N. Glover and J. Pires, *Double Virtual corrections for gluon scattering at NNLO*, *JHEP* **02** (2013) 026 [arXiv:1211.2710] [INSPIRE].

[27] J. Currie, E.W.N. Glover and S. Wells, *Infrared Structure at NNLO Using Antenna Subtraction*, *JHEP* **04** (2013) 066 [arXiv:1301.4693] [INSPIRE].

[28] A. Gehrmann-De Ridder, T. Gehrmann, E.W.N. Glover and J. Pires, *Second order QCD corrections to jet production at hadron colliders: the all-gluon contribution*, *Phys. Rev. Lett.* **110** (2013) 162003 [arXiv:1301.7310] [INSPIRE].

[29] J. Currie, A. Gehrmann-De Ridder, E.W.N. Glover and J. Pires, *NNLO QCD corrections to jet production at hadron colliders from gluon scattering*, *JHEP* **01** (2014) 110 [arXiv:1310.3993] [INSPIRE].

[30] J. Currie, E.W.N. Glover and J. Pires, *Next-to-Next-to Leading Order QCD Predictions for Single Jet Inclusive Production at the LHC*, *Phys. Rev. Lett.* **118** (2017) 072002 [arXiv:1611.01460] [INSPIRE].

[31] J. Currie, A. Gehrmann-De Ridder, T. Gehrmann, E.W.N. Glover, A. Huss and J. Pires, *Precise predictions for dijet production at the LHC*, *Phys. Rev. Lett.* **119** (2017) 152001 [arXiv:1705.10271] [INSPIRE].

[32] J. Currie, E.W.N. Glover, T. Gehrmann, A. Gehrmann-De Ridder, A. Huss and J. Pires, *Single Jet Inclusive Production for the Individual Jet p_{T} Scale Choice at the LHC*, *Acta Phys. Polon.* **B 48** (2017) 955 [arXiv:1704.00923] [INSPIRE].

[33] O.V. Tarasov, A.A. Vladimirov and A.Y. Zharkov, *The Gell-Mann-Low Function of QCD in the Three Loop Approximation*, *Phys. Lett.* **B 93** (1980) 429 [INSPIRE].

[34] S.A. Larin and J.A.M. Vermaseren, *The Three loop QCD β-function and anomalous dimensions*, *Phys. Lett.* **B 303** (1993) 334 [hep-ph/9302208] [INSPIRE].

[35] G. Altarelli and G. Parisi, *Asymptotic Freedom in Parton Language*, *Nucl. Phys.* **B 126** (1977) 298 [INSPIRE].

[36] S. Moch, J.A.M. Vermaseren and A. Vogt, *The Three loop splitting functions in QCD: The Nonsinglet case*, *Nucl. Phys.* **B 688** (2004) 101 [hep-ph/0403192] [INSPIRE].

[37] A. Vogt, S. Moch and J.A.M. Vermaseren, *The Three-loop splitting functions in QCD: The Singlet case*, *Nucl. Phys.* **B 691** (2004) 129 [hep-ph/0404111] [INSPIRE].

[38] M. Dasgupta, F.A. Dreyer, G.P. Salam and G. Soyez, *Inclusive jet spectrum for small-radius jets*, *JHEP* **06** (2016) 057 [arXiv:1602.01110] [INSPIRE].

[39] M. Cacciari, G.P. Salam and G. Soyez, *The Catchment Area of Jets*, *JHEP* **04** (2008) 005 [arXiv:0802.1188] [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] J. Butterworth et al., *PDF4LHC recommendations for LHC Run II*, *J. Phys.* **G 43** (2016) 023001 [arXiv:1510.03865] [INSPIRE].

[42] L.A. Harland-Lang, A.D. Martin and R.S. Thorne, *The Impact of LHC Jet Data on the MMHT PDF Fit at NNLO*, *Eur. Phys. J.* **C 78** (2018) 248 [arXiv:1711.05757] [INSPIRE].