Prepared for submission to JHEP Cavendish-HEP-25/06

IRC-safe jet flavour without modifying anything

Terry Generet

Cavendish Laboratory, University of Cambridge, Cambridge CB3 0HE, United Kingdom

E-mail: generet@hep.phy.cam.ac.uk

Abstract: In this work, we describe how infrared-collinear safety can be restored perturbatively for standard definitions of jets and jet flavour. We will explicitly study this approach at next-to-next-to-leading order in QCD, where we will discuss the cause of the violation of infrared safety, argue that it is merely an artefact of the calculation, and explain how a consistent treatment of flavour removes this violation. The most important feature of our approach is that it does not require any changes to the definition of the jet or its flavour, nor does it modify the definition of the cross section. Consequently, the predictions can be directly compared to measurements performed using the standard anti-k_T jet clustering algorithm, without the need for experimental collaborations to adapt their analyses to some new, infrared-collinear-safe definition of jet flavour, as would be the case for most - if not all - solutions presented in the literature thus far.

Contents

1 Introduction

It has long been known [1] that standard definitions of jet flavour break down starting at next-to-next-to-leading order (NNLO) in QCD. Usually, the jets at hadron colliders are defined using the anti-k_T [2], k_T [3, 4] or Cambridge-Aachen [5, 6] jet clustering algorithms. The flavour of the jet is then assigned according to the number of quarks of a specific flavour contained in a given jet. For ease of discussion, but without loss of generality, we will assume throughout this work that bottom is the flavour of interest.

Broadly speaking, there are two schemes: in the first scheme, a jet is considered a b-jet if it contains at least one b-quark or anti-b-quark. From now on, we shall call this scheme the 'any-flavour' scheme, as was done in ref. [7]. This definition is useful for measurements[1], since it is often difficult to experimentally distinguish jets containing two b-quarks from jets containing just one, while distinguishing either from jets containing no b-quarks at all is not. However, this definition is rarely used in higher-order calculations, because it is not infrared-collinear (IRC) safe. We say a jet definition is IRC safe if its use does not lead to the non-cancellation of divergences related to soft and/or collinear massless QCD partons. The any-flavour scheme is not IRC safe, because it does not remove all collinear divergences.

[1] Of course, experimental collaborations do not define flavour in terms of unobservable quarks. Instead, they look for secondary vertices indicative of the presence of B-hadrons or study other properties of the jet which correlate with the presence of b-quarks. However, this is essentially equivalent to the parton-level scheme described here.

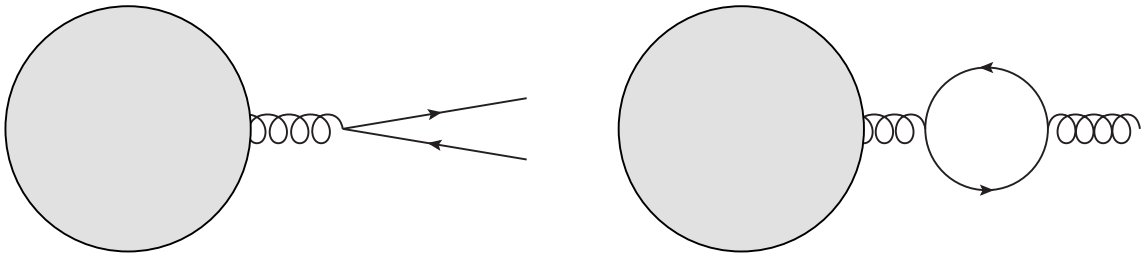

Figure 1. Left: the emission of a collinear $b\bar{b}$-pair. Right: the emission of a gluon with a b-quark loop. The grey blobs represent some arbitrary process.

Indeed, if the final state contains a $b\bar{b}$-pair and this pair becomes collinear (see left panel of fig. 1), the quarks will be clustered into the same jet, which will then be classified as flavoured, since it contains at least one flavoured quark. However, this collinear limit is singular, and thus the cross section diverges (or, if dimensional regularisation [8] with $d = 4 - 2\varepsilon$ is used, it contains uncancelled poles in ε).

The second scheme assigns flavour based on whether a jet contains an even or odd number of (anti-)b-quarks: a jet with an even number of (anti-)b-quarks is considered unflavoured, while a jet containing an odd number of (anti-)b-quarks is considered flavoured. Again following ref. [7], we will refer to this as the 'flavour modulo-2' scheme. Using this definition, the poles in ε cancel: the collinear limit mentioned before still generates a pole, but now the resulting jet is unflavoured and the pole cancels against the pole generated by the contribution shown in the right panel of fig. 1, which likewise generates an unflavoured jet. An analysis of all other singular limits reveals that this definition of jet flavour is IRC safe through next-to-leading order (NLO). For this reason, this scheme has been commonly used in NLO calculations. Since this scheme cannot easily be used for measurements – as explained above – this creates a mismatch between the measured cross sections (which use the any-flavour scheme) and the calculated ones (which use the flavour modulo-2 scheme). This mismatch is accounted for using unfolding techniques. This correction step has been a standard part of many analyses for years now and is well understood.[2]

However, starting at NNLO, neither definition of jet flavour is IRC safe when using any of the standard jet clustering algorithms. To illustrate this, we consider the anti-k_T algorithm. However, the analysis for other (IRC-unsafe) algorithms is completely analogous.[3] Consider the configuration shown in fig. 2: a final state containing two hard partons (assumed to be unflavoured) and a soft $b\bar{b}$-pair. Let us assume that the cross section of interest is a two-b-jet

[2]It is worth pointing out that, regardless of how well-practised the community is in this approach, this unfolding step is highly sensitive to the modelling of gluon splitting (see e.g. ref. [9]). Since this modelling as currently implemented in parton showers still leaves something to be desired, it would be ideal to use the any-flavour scheme for theory calculations and eliminate this step. This is a worthwhile avenue for future research, but goes beyond the scope of what this work is trying to achieve.

[3]In fact, we will assume the anti-k_T algorithm as the 'default' algorithm from now on. In practice, the solution proposed in this paper works for all commonly used algorithms.

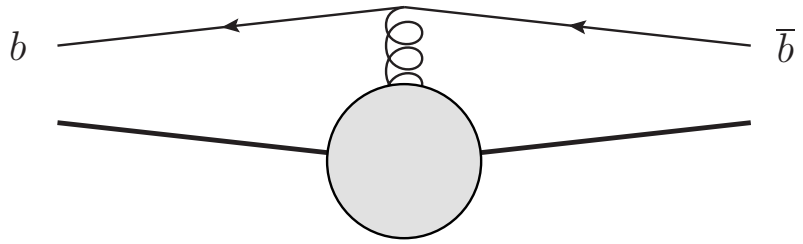

Figure 2. The problematic configuration at NNLO: the two soft b-quarks are clustered with different hard partons (thick lines) into two separate jets.

cross section and that the jets corresponding to the two hard partons satisfy any required cuts. The only way this configuration contributes to the cross section is if the two b-quarks are clustered with different hard partons into two separate jets. The angular phase space for such a configuration is of course non-zero, so the integration over the angles of the b-quarks will yield a finite (i.e. non-zero) factor. This leaves the integration over the energies of the b-quarks. However, while there is no divergence associated with either quark on its own becoming soft, if both quarks become soft simultaneously, this leads to the well-known double-soft singularity for quark-anti-quark pairs. If this divergence is dimensionally regulated, the corresponding pole will not cancel when summing over all contributions.

This issue is well-known and was first identified in ref. [1]. Since then, and especially in recent years, a number of solutions have been proposed [1, 7, 10–14].[4] These all involve modifying the jet (flavour) definition somehow: either the jet definition is modified in a way that guarantees that a soft $b\bar{b}$-pair is clustered into the same jet [1, 12], in which case the kinematics of the jet(s) will not be identical to what would have been obtained using the anti-k_T algorithm, or the definition of jet flavour is changed [7, 10, 11, 13], in which case the jet kinematics might not change, but this usually involves some extra step(s). There have even been proposals to modify the definition of the cross section itself [14]. Generally, these solutions either introduce some new parameter(s) used to define the jet (on top of the usual jet radius R) and/or an extra step in which the partons are clustered a second time using a different procedure to determine the flavour of the jets initially found using anti-k_T. Needless to say, the introduction of extra free parameters should be avoided. Adding extra steps to the jet algorithm is also not desirable, as it increases the computational complexity. Moreover, essentially by construction, the order in which partons are clustered is typically flavour-dependent and it is not clear how this should be translated to experiment. In essence, the parton-hadron correspondence valid for the anti-k_T algorithm is violated for these proposed solutions. Finally, there are many solutions, all with their own merits and drawbacks,

[4] Recent progress in fragmentation calculations [15–18] has also opened up the possibility to use identified heavy-flavour hadrons instead of flavoured jets to establish the presence of heavy flavour in processes through NNLO.

and it is unclear which should be chosen as the new standard moving forward.[5] And once this choice has been made, adapting to this new standard may involve significant effort. In the meantime, the resulting uncertainty has different groups using different algorithms, complicating comparisons of results.

In this work, we propose how this issue can be avoided entirely, allowing the community to keep using the tried and trusted algorithms. We will show that this involves only minor effort from the side of theory collaborations, while absolutely nothing changes for experimental collaborations with respect to the 'old' strategy of clustering using anti-k_T and then unfolding from the any-flavour scheme to the flavour modulo-2 scheme. In section 2, we will explain the proposed solution. Some numerical checks will be presented in section 3, followed by two phenomenological studies in section 4. Finally, we will conclude in section 5.

2 Nature of the problem and its solution

As mentioned in the introduction, the problem (at NNLO) boils down to this: there is an uncancelled singularity coming from soft $b\bar{b}$-pairs. Clearly, this singularity is regulated in reality by the non-zero b-quark mass. In fact, the only reason why b-jet (and c-jet) cross sections can be calculated perturbatively is because of the non-zero b-quark (and c-quark) mass. Strange jets, for example, will get contributions from non-perturbative $g \to s\bar{s}$ splittings, which need to be modelled empirically, using e.g. a parton shower. Thus, in both cases[6] where perturbative calculations can be used in the first place, the problem can be avoided if the corresponding quarks are simply given their physical mass. This shows that, in essence, the problem is a consequence of the choice to treat these heavy quarks as massless.

Of course, this is no news to the theory community. In fact, the $pp \to W + 2b-$jet cross section was computed in ref. [19] at NNLO using massive quarks precisely for this reason. However, there are two arguments commonly used to argue why this cannot be the standard. First of all, keeping the b-quark mass non-zero introduces another scale to the problem. It is well-known that the more scales are present in a process, the more difficult the computation of the loop amplitudes becomes. The two-loop amplitudes are typically the bottleneck for modern NNLO computations, so making them more complicated to compute would delay the entire NNLO computation.[7] Additionally, a non-zero mass can also significantly affect the numerical stability and convergence of the calculations. For these reasons, it is clear that just doing the calculation for massive quarks is not a practical solution.

The second reason is that cross sections involving massive quarks contain logarithms of the b-quark mass over the hard scale of the process, $\ln\left(Q^2/m_b^2\right)$, which can grow arbitrarily large at high energies, spoiling the perturbative convergence of the cross section. In fact,

[5] See also ref. [9] for a study comparing most of these proposals quantitatively.

[6] Or all three cases, if one counts top-jets.

[7] To get around this, a massification procedure was used in ref. [19] to obtain – up to power corrections – the two-loop amplitudes involving massive b-quarks from those with massless b-quarks.

depending on the definition of the cross section, the leading logarithms at order n can behave as poorly as $\alpha_s^n \ln^{2n}(Q^2/m_b^2)$.

Despite these concerns, our solution to the jet flavour problem is to keep the b-quark mass non-zero, at least at leading power (LP). In the following, we will argue how the resulting additional computational complexity can be minimised to be essentially negligible and how the logarithmic behaviour is actually unproblematic for standard jet definitions.

To begin, we recall that the dependence of the cross section on the b-quark mass factorises, up to power corrections. This factorisation allows one to obtain the massive result from the massless one using a few process-independent ingredients. Most importantly, this restores the logarithmic dependence of the cross section on the b-quark mass and we will be referring to these terms as 'logs', though it is understood that the massification also introduces mass-independent terms. To give a common example, there are logs coming from loops associated with the renormalisation of the strong coupling, which can be reintroduced using the α_s decoupling constant [20, 21]. There are also logs associated with quasi-collinear configurations in the initial or final state, which can be reintroduced using the corresponding matching conditions for parton distribution functions (PDFs) [22–27] or fragmentation functions (FFs) [28–32], respectively. In this way, the leading dependence on the b-quark mass can be reintroduced in a process-independent (i.e. automated) way, without the need to compute any truly massive cross sections. Importantly, these process-independent ingredients (e.g. the perturbative FFs) also contain poles in ε which exactly cancel the corresponding poles (e.g. collinear poles in the final state) in the massless cross section. In general, there is a correspondence between poles in the massless cross section and logs in the massive one. Intuitively, this correspondence arises because the process-independent ingredients only contain (aside from some arbitrary renormalisation scale μ) one scale: m_b. Therefore, their entire dependence on m_b is of the form $(\mu^2/m_b^2)^{n\varepsilon}$, which produces logs according to the same pattern as the ε-poles present in the process-independent ingredient, which in turn match the poles of the massless cross section.

Many of these logs are already routinely included and resummed in cross sections. In particular, this includes the logs coming from the α_s decoupling constant and the logs coming from the PDFs. Both are included and resummed in all standard PDF sets available. Therefore, they are already included in massless cross section calculations and do not spoil the perturbative convergence (since they have been resummed). The case of collinear final state logs (i.e. FFs) is even simpler: they can be omitted entirely.

To see this, consider all possible cases of collinear logs at NLO. The first pair of cases was already shown in fig. 1: a gluon splitting into a collinear $b\bar{b}$-pair and a gluon with a $b\bar{b}$ loop. In the first case, both quarks will be clustered into the same jet, since they are collinear. Since we are using the flavour modulo-2 scheme, they thus do not affect the flavour of any jet. In the second case, the jet flavours are also unaffected, since the particle is a gluon. Thus the logs of these two contributions cancel, provided that they are the same up to a sign. This is of course guaranteed by the Kinoshita-Lee-Nauenberg (KLN) theorem [33, 34] and can be easily checked using the explicit expressions for the one-loop FF matching conditions. The

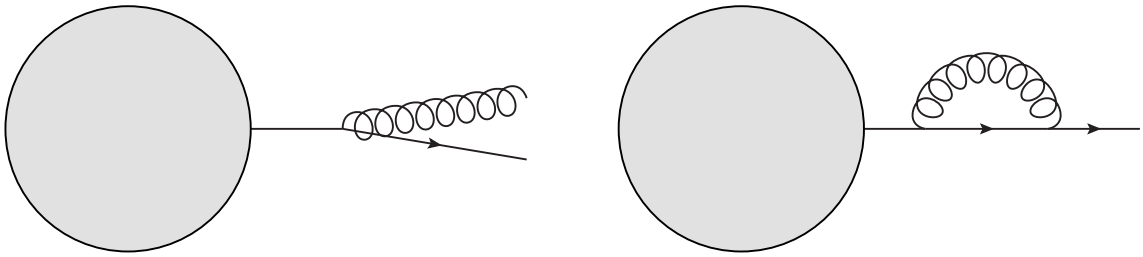

Figure 3. Left: the collinear emission of a b-quark and a gluon. Right: the emission of a b-quark with a QCD loop. The grey blobs represent some arbitrary process.

same holds for the other pair of cases, shown in fig. 3: the logs generated by the emission of a real gluon off a heavy quark exactly cancel those generated by virtual contributions (after integration over the momentum fraction). By the KLN-theorem, this holds to any order: all final-state (soft-)collinear logs cancel as long as one is insensitive to the distribution of collinear momentum among the collinear partons (i.e. one integrates over the momentum fractions) and to collinear $g \to b\bar{b}$ splittings (i.e. one uses the flavour modulo-2 scheme).[8] Thus, all collinear logs either cancel or are already resummed by default. In particular, this also includes soft-collinear logs.[9] This only leaves soft logs, which first show up at NNLO and thus the cross section only behaves as $\alpha_s^n \ln^{2n-3}(Q^2/m_b^2)$ at worst, as will be discussed further below. This formally makes them a next-to-next-to-next-to-leading logarithmic contribution, which is typically considered safe to include at fixed order, by current standards.

Since all (soft-)collinear logs have been addressed, this only leaves the soft logs. In the following, we will completely ignore the presence of the other logs, since they have already been addressed. Although this has never been discussed in the literature to the best of our knowledge, the soft logs can also be factorised in analogy to perturbative fragmentation functions. The factorisation looks as follows:

$$\begin{aligned}\mathcal{F}\big(\sigma(m_b)\big) = \mathcal{F}\big(S_{\emptyset}(m_b) \otimes \sigma(m_b = 0)\big) + \mathcal{F}\big(S_{b\bar{b}}(m_b, p_b, p_{\bar{b}}) \otimes \sigma(m_b = 0)\big) \\ + \mathcal{F}\big(S_{b\bar{b}b\bar{b}}(m_b, p_{b1}, p_{\bar{b}1}, p_{b2}, p_{\bar{b}2}) \otimes \sigma(m_b = 0)\big) + ... + \mathcal{O}(m_b/Q)\,,\end{aligned} \tag{2.1}$$

where the ellipses are an infinite sequence of analogous contributions with an increasing number of $b\bar{b}$-pairs. $\mathcal{F}$ is a measurement function, which among other things implements the jet (flavour) definition and applies a cut on the number of flavoured jets. The $\otimes$ symbol indicates two things. First of all, it indicates that the particles in the subscript of the function S are inserted into the final state observed by $\mathcal{F}$ with corresponding momenta as shown in

[8] While these logs do not cancel for the any-flavour scheme (and thus need to be included), this is not a problem in practice, since the unfolding done by experimentalists is performed using partons showers, which can treat the b-quark as massive (avoiding the need to massify anything) and resum the logs at the same time (avoiding the presence of large collinear logs).

[9] Soft-collinear logs related to the initial state (i.e. soft splittings to b-quarks at low x) are typically not resummed, but they are phenomenologically irrelevant.

the arguments of the function S. I.e.: the function $S_\emptyset$ leaves the final state untouched, while $S_{b\bar{b}}(p_b, p_{\bar{b}})$ adds a $b\bar{b}$-pair to the final state with momenta p_b and $p_{\bar{b}}$, etc. Important to note is that these momenta should be treated as infinitesimal, i.e. they do not affect the kinematics of the other particles and enter the clustering algorithm and cuts as soft partons. This is again in analogy to perturbative fragmentation, where the fragment is taken to be strictly collinear to the fragmenting parton. Second of all, the $\otimes$ symbol indicates that the S functions factorise from the cross section up to colour correlations. This is in full analogy to the way soft singularities of cross sections factorise up to colour correlations. The integration over all momenta is implicit. Since this factorisation assumes that the b-quark mass and the momenta introduced by the S functions are small compared to the hard scale of the process, the momenta of the other partons can equivalently be considered to be infinitely large. As such, the integration of the momenta introduced by the S functions should extend to infinity and there is no need to consider the effect of the recoil of the soft quarks on the other partons. The former point will be further explained when the computation of the S functions is discussed below, but is once again fully analogous to the case of perturbative fragmentation, where, despite being 'collinear', the relative transverse momenta of the collinear partons are integrated from zero to infinity when computing the perturbative fragmentation functions.

Clearly, the S function for n b-quark pairs is only non-zero starting at $\mathrm{N}^{2n}\mathrm{LO}$. At the same time, a flavour-insensitive cross section is insensitive to mass effects, so

$$\sum_{i=0}^{\infty} \int dp_{b1} \int dp_{\bar{b}1} \ldots \int dp_{bi} \int dp_{\bar{b}i} S_{\underbrace{b\bar{b}\ldots b\bar{b}}_{\text{i pairs}}} = \mathbf{1} \,, \tag{2.2}$$

where we suppressed the arguments of the S functions and $\mathbf{1}$ is the identity matrix in colour space, i.e. it acts trivially on σ. Through NNLO, this gives

$$S_\emptyset = \mathbf{1} - \int S_{b\bar{b}} dp_b dp_{\bar{b}} + \mathcal{O}(\alpha_s^3) \,. \tag{2.3}$$

At NNLO, there are three contributions introducing soft logs: one double-virtual, one real-virtual and one double-real, shown in the top-left, top-right and bottom panel of fig. 4, respectively. Here, eikonal vertices should be used to attach the gluons to the external partons i and j. The double-virtual and real-virtual contributions will modify a cross section without affecting the flavour assignments of the Born jets. These correspond to contributions to $S_\emptyset$. The double-real emission changes the flavour assignments depending on the direction (and possibly the relative magnitude) of the b-quark momenta. This is a contribution to $S_{b\bar{b}}$.

The real-virtual contribution is trivial: the b-quark bubble reduces to a multiplicative factor. The rest corresponds to the emission of a single gluon. The integral over its energy is scaleless, so the contribution vanishes.

Moving on to the double-virtual contribution, it can in principle be computed analytically using the diagram shown in the top-left panel of fig. 4.[10] However, we here refrain from

[10] Note that two distinct cases would need to be considered: one where $i \neq j$ and one where $i = j$.

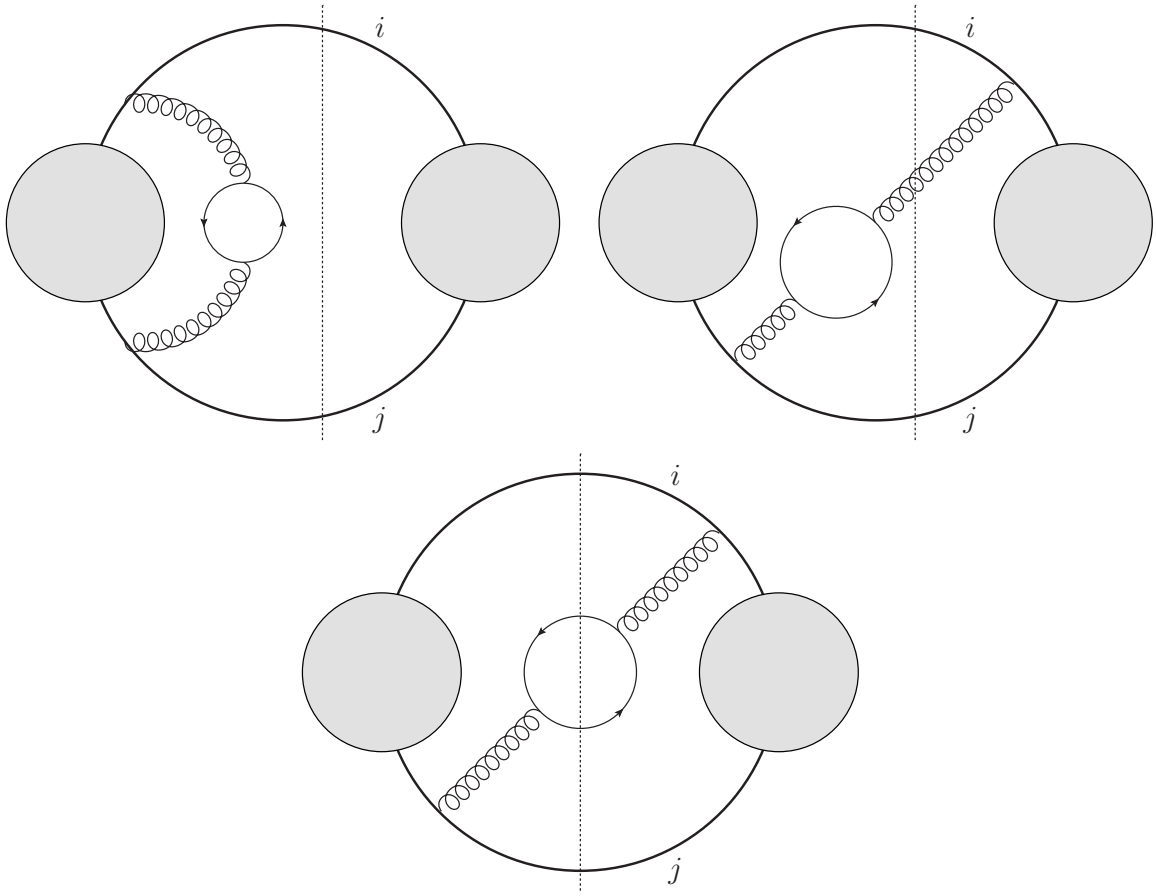

Figure 4. Top-left: a double-virtual diagram which contributes soft mass logs. The thick lines correspond to arbitrary external partons i and j, while the dashed line represents the cut. Top-right: the corresponding real-virtual diagram. Bottom: the corresponding double-real diagram. The existence of complex conjugate diagrams is understood for the virtual contributions.

performing this exercise. Instead, we use the relation given in eq. (2.3) to obtain it numerically. I.e. after deriving the NNLO expression for $S_{b\bar{b}} \otimes \sigma$, we will compute $S_{\emptyset} \otimes \sigma$ by negation and integration over the soft momenta p_b and $p_{\bar{b}}$, while leaving the final state intact (i.e. without inserting the soft quarks into the final state). In part, this is to reduce the amount of analytical integration that needs to be performed in order to implement our scheme – indeed, no analytical integration will be necessary at all – but there is also a more significant reason which will become apparent when discussing the implementation of $S_{b\bar{b}} \otimes \sigma$ further below.

Moving on to the real contribution, the squared matrix element factorises as follows:

$$|M_{b,\bar{b},a_1\ldots}|^2 \sim (4\pi\alpha_s)^2 T_F \sum_{i,j} \mathcal{I}_{ij}(p_b, p_{\bar{b}}, m_b) \langle M_{a_1,\ldots}|\mathbf{T}_i \cdot \mathbf{T}_j|M_{a_1,\ldots}\rangle \,, \tag{2.4}$$

where we used the same notation for colour correlators as in ref. [35]. This form is of course identical to the case of a massless quark pair. The double-soft function for massive quarks

$\mathcal{I}_{ij}(q_1, q_2, m)$ is given by

$$\mathcal{I}_{ij}(q_1, q_2, m) = \frac{(p_i \cdot q_1)(p_j \cdot q_2) + (p_j \cdot q_1)(p_i \cdot q_2) - (p_i \cdot p_j)[m^2 + (q_1 \cdot q_2)]}{[m^2 + (q_1 \cdot q_2)]^2 [p_i \cdot (q_1 + q_2)][p_j \cdot (q_1 + q_2)]} \,. \tag{2.5}$$

Note that for $m = 0$, the massless result is trivially recovered. $\mathcal{I}_{ij}(q_1, q_2, m)$ can be derived in full analogy to the massless version by studying the behaviour of amplitudes when both heavy-quark momenta and the quark mass uniformly approach zero, i.e. $q_1 \to \lambda q_1$, $q_2 \to \lambda q_2$ and $m \to \lambda m$ with $\lambda \to 0$.

At NNLO, the function $S_{b\bar{b}}$ is simply the operator which replaces

$$|M_{a_1 \ldots}|^2 \to (4\pi\alpha_s)^2 T_F \sum_{i,j} \mathcal{I}_{ij}(p_b, p_{\bar{b}}, m_b) \langle M_{a_1,\ldots} | \mathbf{T}_i \cdot \mathbf{T}_j | M_{a_1,\ldots} \rangle \,, \tag{2.6}$$

and adds the two b-quarks to the final state. Technically, the S functions should correspond to the difference between the massive and massless cross sections. Therefore $S_{b\bar{b}}$ should correspond to the difference between the double-soft limit of the massive and massless cross sections. In this case, the cross section is proportional to $\mathcal{I}_{ij}(p_b, p_{\bar{b}}, m_b) - \mathcal{I}_{ij}(p_b, p_{\bar{b}}, 0)$. This also highlights that the massification removes the double-soft poles from the massless cross section, since the double-soft limit is explicitly subtracted. This view should be compared to the definition of perturbative fragmentation functions employed in refs. [29, 30], which is likewise given by the difference between the massive and massless cases. At sufficiently high energies of the soft b-quarks (but by assumption still much smaller than the hard scale of the process), the difference between the massive and massless soft functions gives a vanishing contribution, so what is changed by including S truly is only the soft behaviour. At the same time, the massless contribution vanishes when integrated over the energies, since it is scaleless. It can therefore be omitted, leaving only the contribution proportional to $\mathcal{I}_{ij}(p_b, p_{\bar{b}}, m_b)$.

Keeping the massless term for now, the function $S_{b\bar{b}}$ contains a soft pole. In the case of perturbative FFs, the corresponding collinear poles could be treated analytically, since the calculation of the perturbative FFs can be performed analytically: all momenta are integrated over, up to the integration over the longitudinal momentum fraction. In the case of $S_{b\bar{b}}$ at NNLO, the momenta cannot be integrated over analytically, since their directions and relative magnitude[11] affect the result of the jet clustering. At the same time, the only integration which contains a soft singularity is the integral over the absolute magnitude of the soft momenta. One might e.g. parametrise the momenta using x_S and E_S like $|\vec{p}_b| = x_S E_S$, $|\vec{p}_{\bar{b}}| = (1 - x_S) E_S$, $0 \leq x_S \leq 1$, $E_S \geq 0$, in which case the soft divergence is purely due to the integration over E_S. Since the absolute magnitude of the soft momenta is kinematically irrelevant, this integral could in principle be performed analytically, enabling the use of dimensional regularisation to deal with the divergence. However, we instead prefer to perform the integral numerically, using subtraction techniques to deal with the singularity.

[11] While the absolute magnitude of the soft momenta is irrelevant (since they are kinematically treated as infinitesimal), it is possible to construct jet algorithms where their relative magnitude matters. However, this is not the case for any of the common algorithms.

When omitting the scaleless massless contribution, the soft singularity is traded for a UV singularity at $E_S = \infty$. Studying the integrand reveals the scaling of the cross section in this limit:

$$\begin{aligned}\frac{d^{d-1}p_b}{E_b}\frac{d^{d-1}p_{\bar{b}}}{E_{\bar{b}}}\mathcal{I}_{ij}(p_b,p_{\bar{b}},m) &= d^{d-2}\Omega_b d^{d-2}\Omega_{\bar{b}}\frac{d|\vec{p}_b||\vec{p}_b|^{d-2}}{E_b}\frac{d|\vec{p}_{\bar{b}}||\vec{p}_{\bar{b}}|^{d-2}}{E_{\bar{b}}}\mathcal{I}_{ij}(p_b,p_{\bar{b}},m)\\ &\to d^{d-2}\Omega_b d^{d-2}\Omega_{\bar{b}}\frac{dE_S}{E_S^{1+4\varepsilon}}dx_S\big(x_S(1-x_S)\big)^{1-2\varepsilon}\mathcal{I}_{ij}(x_S\hat{n}_b,(1-x_S)\hat{n}_{\bar{b}},0)\ , \end{aligned} \tag{2.7}$$

where we performed the variable transformation to x_S and E_S and introduced the versors $\hat{n}_b$ and $\hat{n}_{\bar{b}}$, defined via $p_i = p_i^0\hat{n}_i$ in the massless limit. In the UV, the soft momenta can of course be treated as massless. To perform this integral, we use a variation on the usual trick

$$\int_0^1\frac{f(x)}{x^{1+a\varepsilon}}dx = \int_0^1\frac{(f(x)-f(0))}{x^{1+a\varepsilon}}dx - \frac{1}{a\varepsilon}f(0) = \int_0^1\left[-\frac{1}{a\varepsilon}\delta(x)+\frac{1}{x_+^{1+a\varepsilon}}\right]f(x)dx\ , \tag{2.8}$$

where we introduced the usual plus-distribution. In our case, the integral extends from 0 to infinity and the subtraction needs to happen at infinity, not zero. Additionally, the high-energy approximation shown in eq. (2.7) contains both a UV and a soft divergence, and so is not suitable as a subtraction term. Instead, we can use the following formula:

$$\begin{aligned}&\int E_S^{3-4\varepsilon}\mathcal{I}_{ij}(p_b,p_{\bar{b}},m)dE_S = \\ &\qquad\int\left[E_S^{3-4\varepsilon}\mathcal{I}_{ij}(p_b,p_{\bar{b}},m) - \frac{E_S^A}{(R^B+E_S^B)^{(A+1+4\varepsilon)/B}}\mathcal{I}_{ij}(x_S\hat{n}_b,(1-x_S)\hat{n}_{\bar{b}},0)\right]dE_S\\ &\qquad + R^{-4\varepsilon}\frac{\Gamma(\frac{A+1}{B})\Gamma(\frac{4\varepsilon}{B})}{B\,\Gamma(\frac{A+1+4\varepsilon}{B})}\mathcal{I}_{ij}(x_S\hat{n}_b,(1-x_S)\hat{n}_{\bar{b}},0)\ ,\end{aligned} \tag{2.9}$$

where R, A and B are free parameters. A correct result does not depend on them, which is a valuable test of any implementation of this scheme. The first term is free of singularities and poles in ε, so it can be integrated numerically in four dimensions. The integrated subtraction term does contain an ε-pole, and so should be evaluated in d dimensions. However, the singular E_S integration has been performed analytically and is no longer an issue. The d-dimensional nature of this term means great care must be taken to correctly include all ε-dependent factors coming from the phase space measure. It also means that the integration over the angles of the soft quarks must be performed in d dimensions. This can be done by promoting one of the momenta to five dimensions and the second to six dimensions, as described in ref. [35].

This last point provides another valuable cross check of results. The definition of normal four-dimensional observables beyond four dimensions is ambiguous. A jet algorithm typically clusters particles based on some angular distance between momenta. It is ambiguous whether - and how - the five- and six-dimensional components should be included in these definitions. Such choices can certainly have an impact on the contribution coming from the massification term $S_{b\bar{b}}\otimes\sigma$. However, the exact same ambiguity shows up in the massless cross section.

There, the integrated subtraction terms for soft $b\bar{b}$-pairs likewise involve higher-dimensional momenta. Clearly, the physical cross section cannot depend on the choice of analytical continuation of four-dimensional observables, so the dependence on this choice must drop out in the combination of all terms.

There is a subtlety regarding the choice of analytical continuation of observables. It might appear as if one has complete freedom regarding how the fifth and sixth momentum components enter the observable definition, as long as the usual four-dimensional definition is recovered when the higher-dimensional components are set to zero. This is, however, not the case. Dimensional regularisation can be applied as long as the four-dimensional case is recovered in the limit $\varepsilon \to 0$ up to higher-order terms in the ε-expansion. However, the angular integrations beyond four dimensions receive, in general, non-trivial contributions which do not vanish at $\varepsilon = 0$. This can be seen in eq. (88) of ref. [35]: while the integral over the fifth component is $\mathcal{O}(\varepsilon)$ for a trivial sixth component, the integral over the sixth component is $\mathcal{O}(\varepsilon^0)$. For physical cross sections, this is unproblematic, because there can only be a non-trivial dependence on the sixth component of the second unresolved momentum if the first has a non-zero fifth component. If this is the case, then the whole contribution is $\mathcal{O}(\varepsilon)$ on account of the non-zero fifth component of the first unresolved momentum. If the fifth component of the first unresolved momentum does vanish, then the integrand does not depend on the sixth component of the second unresolved momentum, leaving just the plus distribution from the phase space measure in eq. (88) of ref. [35], which integrates to zero. Because of this, it is sufficient to use four-dimensional momenta for the integration over the real emission phase space for contributions absent any ε-poles. However, if an explicit dependence on the $(d-4)$-dimensional components of momenta is introduced that involves more than just physical quantities, i.e. scalar products of the $(d-4)$-dimensional parts of the momenta, then this can introduce an $\mathcal{O}(\varepsilon^0)$ difference w.r.t. the four-dimensional case, yielding an incorrect result. Here, we will limit ourselves to the class of analytical continuations of four-dimensional observables which depend on higher-dimensional components of momenta only through scalar products of their $(d-4)$-dimensional parts, and in such a way that the four-dimensional definition is recovered when all these scalar products vanish. Such analytical continuations are guaranteed to only induce effects of ε beyond four dimensions. Within this class, one has complete freedom regarding the choice of analytical continuation and the dependence on this choice drops out in the final result, as discussed above.

Having dealt with the soft singularity, one important point remains: the massless contribution also contains collinear singularities. When this contribution is omitted, these divergences are replaced with collinear singularities at infinity. In practice, these are collinear singularities in the (integrated) subtraction term introduced to regulate the UV divergence. However, note that these collinear configurations cannot affect the flavour of jets. Indeed, the only singularities occur when the two b-quarks become collinear or when they additionally become collinear to one of the hard partons. Since the b-quarks must be collinear, they are necessarily clustered into the same jet and thus leave its flavour unchanged. This flavour-unaffected cross section also gets a contribution from $S_\emptyset$, which can be obtained by integrating

the contribution of $S_{b\bar{b}} \otimes \sigma$ with the b-quarks removed with a minus sign. If we compute the contribution from $S_\emptyset$ numerically in this way at the same time as $S_{b\bar{b}} \otimes \sigma$, it thus acts as a local subtraction term for the collinear singularities in $S_{b\bar{b}} \otimes \sigma$.

In practice, the implementation of the soft massification terms thus looks as follows at NNLO. A Born phase space configuration and the momenta p_b and $p_{\bar{b}}$ are generated. The soft $b\bar{b}$-pair is inserted into the final state and the contributions from $S_{b\bar{b}}$ and its UV subtraction term are computed and added to the cross section. Then these same contributions are subtracted, after removing the soft quarks from the final state. The integrated subtraction term is treated analogously. This will simultaneously compute the contributions from $S_\emptyset$ and $S_{b\bar{b}}$, while regulating all singularities. Additionally, it is important that the computation of the massless cross section is kept fully differential in the momenta of the soft b-quark pair, including the five- and six-dimensional components where needed. The implementation of the sector-improved residue subtraction scheme [35–38] in the C++ library STRIPPER has been extended to support our proposed solution to the jet flavour problem. STRIPPER has the advantage of having been designed with the aim of avoiding analytical integrations. In contrast to most implementations of subtraction schemes, the integrations over soft partons' angles are thus performed numerically, keeping the framework completely differential w.r.t. such typically unobserved degrees of freedom, making the present extension and others like it essentially trivial.

While the focus here is on NNLO – which is, after all, the only case of phenomenological relevance at this time – it is apparent that our solution can easily be extended to higher orders. Since the IRC poles are theoretically cancelled locally (though the practical implementation used here is non-local), it is not affected by any of the problematic cases described in ref. [7], as long as the higher-order extension is performed consistently. It should also be evident that the additional effort needed to implement our solution at a given order will be just a fraction of the effort needed to perform calculations at that order in the first place. The implementation at NNLO requires no analytical integration whatsoever, only requiring studying one of the simplest double-unresolved limits of amplitudes and some minor modifications to the subtraction scheme. At N^{3}LO, some one-loop integrations will need to be performed analytically, but the effort required to do so is minor compared to the effort needed to construct an N^{3}LO subtraction scheme and compute all its necessary ingredients. We therefore see no well-justified barrier to using this solution at higher orders.

2.1 Logarithmic structure and perturbative convergence

As with any factorisation, the factorisation of mass logs into the S functions in principle introduces a new arbitrary renormalisation scale, we shall call it μ_S, in analogy to μ_R for the strong coupling, μ_F for PDFs and μ_{Fr} for FFs. Our discussion so far has ignored this detail – choosing effectively to set $\mu_S = \mu_R$ – since the present work only considers fixed-order cross sections, where the dependence on this scale drops out exactly as long as the ε-poles cancel.[12]

[12] Which of course happens if the implementation is correct. Thus, generalising to $\mu_S \neq \mu_R$ and checking that the μ_S-dependence drops out does not constitute an independent test of the correctness of the implementation.

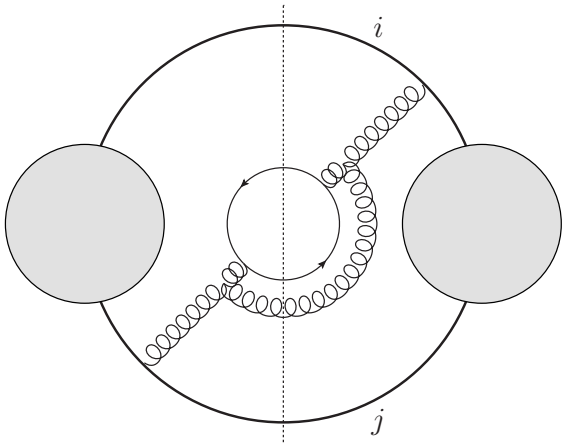

Figure 5. An example of an N^{3}LO contribution to $S_{b\bar{b}}$ which generates a non-cancelling double log.

That being said, if the soft mass logs are resummed, then the dependence on μ_S only cancels up to higher-order terms, as for all renormalisation scales.

Due to the more complicated nature of the S functions, resummation of the soft mass logs cannot be done in a (semi-)analytical way as for couplings, FFs or PDFs. Realistically, a parton-shower-like approach will need to be used. Indeed, the flavour-changing soft logs are a type of non-global logarithm [39, 40], which are notoriously tricky to resum and were first resummed in ref. [39] using a numerical Monte Carlo approach. We will not spend a lot of time musing about how the resummation for the present case might be done in this work. Instead, we will argue that resummation is not necessary, at least for the foreseeable future. To this end, we consider the structure of the soft mass logs through the first few orders in perturbation theory. Obviously, the leading-order and NLO contributions are trivial. The first non-trivial order is NNLO, where a single log appears, corresponding to the simple ε-pole caused by the double-soft limit of a massless quark-anti-quark pair. The scaling through NNLO is thus $\alpha_s^2 \ln\left(m_b^2/\mu_S^2\right)$, where μ_S should be set to the hard scale of the process. Due to the extra suppression by one power of α_s compared to e.g. collinear mass logs, we expect that the soft mass logs at NNLO will be quite small. This will indeed be confirmed numerically in the next section.

In general, the behaviour at N^{n}LO is $\alpha_s^n \ln^{2n-3}(m_b^2/\mu_S^2)$. Assuming $\alpha_s \ln^2(m_b^2/\mu_S^2)$ to be $\mathcal{O}(1)$, the soft mass logs at each consecutive order will give a contribution similar in size to the one at NNLO. While in theory this is the classic situation requiring resummation, in practice the resummation may not be relevant until a sufficiently high order is reached. This can be seen by considering the leading logarithmic contribution at a given order. An example diagram at N^{3}LO yielding a (non-cancelling) triple-logarithm is shown in fig. 5. The triple-logarithm originates from iterating soft limits: first, the $b\bar{b}$-pair is taken to be soft, and then the gluon can exhibit a soft singularity as well. Finally, the gluon can also become collinear to a hard parton, while the soft $b\bar{b}$-pair remains non-collinear. In general, the leading logarithmic

contribution at any order receives contributions from such iterated limits. At N^{n}LO, one can obtain a leading-log contribution by first emitting $n-1$ soft-collinear gluons, each emitted by the previous one and each parametrically softer than the previous, and then splitting the last one (which is not collinear) into a (non-collinear) $b\bar{b}$-pair. Each consecutive gluon introduces a factor $\alpha_s \ln^2(m_b^2/\mu_S^2)$, except for the last gluon splitting into a $b\bar{b}$-pair, which only contributes $\alpha_s^2 \ln(m_b^2/\mu_S^2)$. This gives the $\alpha_s^n \ln^{2n-3}(m_b^2/\mu_S^2)$ scaling mentioned before. This also highlights how these logs might be resummed. One defines generalisations of the S functions to remain differential in all soft partons (i.e. introduces S_g at NLO, etc.) and then iteratively adds soft radiation according to these S functions in the style of a parton shower. In practice, this (leading-logarithmic) resummation will be significantly simpler than standard parton showers, since we are only interested in iterating strict soft limits, without momentum recoil, etc. Again, we do not wish to go into further detail about this here.

The point is that the leading log at N^{n}LO will thus, schematically, always look like $S_{b\bar{b}} \times S_g^{(n-2)}$, which is resummed into $S_{b\bar{b}} \times \mathcal{O}(1)$. Thus if the contribution from the NNLO soft mass logs is small, then the entire leading-log contribution will be of a similar magnitude. Thus, if we find that the NNLO soft mass logs are significantly smaller than the other theory uncertainties, then whether or not the logs are resummed is inconsequential. The resummation will only need to be performed once the total contribution from soft mass logs up to a given order starts reaching a magnitude similar to the scale uncertainties (or other uncertainties) at that order. Considering how small the contribution is at NNLO (as we will see in section 4), resummation is definitely not necessary at that order. Additionally, once such a precision has been reached that one is sensitive to these effects, power corrections in the mass will be relevant as well, in which case the massification discussed here is irrelevant, since one has to work with massive quarks anyway. We will show this explicitly in section 4.

2.2 Renormalisation and non-perturbative flavour

As discussed before, the ε-poles produced by the S functions are analogous to the ε-poles present in perturbative PDFs and FFs. In those cases, they are typically removed by collinearly renormalising those objects. Then, collinear renormalisation counterterms are introduced in massless cross section computations to cancel the homologous divergences seen there. In effect, one is not truly renormalising in the traditional sense, but merely shifting terms around to make two separately divergent quantities separately finite. One could similarly perform a 'soft renormalisation' of the S functions and introduce by hand 'soft renormalisation counterterms' to the massless cross section. We spare ourselves this exercise here, since it is pointless for purely perturbative fixed-order quantities.

That being said, it is perfectly valid to define non-perturbative versions of the S functions, just like one defines (non-perturbative) fragmentation functions for hadronisation. This would introduce a set of non-perturbative S functions, which can be softly-renormalised to render e.g. strange-jet cross sections finite. They would be subject to renormalisation group equations, which however would have to be solved using parton-shower techniques, as mentioned above. As always, they must be fitted to data, but are universal, i.e. process-independent.

While an interesting concept, we refrain from going into more detail here, not only because a proper exploration of this topic would be beyond the scope of this introductory work, but also because, to our knowledge, there would be no phenomenological application at this time.

We will now present some tests of the implementation of our scheme in STRIPPER and consider its phenomenology.

3 Numerical checks

To test our idea and its implementation in STRIPPER, we first consider the single-inclusive b-jet cross section for the 13 TeV LHC, where partons are clustered using the anti-k_T algorithm with $R = 0.4$ and b-jets are defined using the flavour modulo-2 scheme. We will consider the p_T spectrum from 25 GeV to 1600 GeV and require $|y_{b\text{-jet}}| < 2$. Our default choice for the parameters of the UV subtraction term, cf. eq. (2.9), will be $A = 3$, $B = 2$ and $R = m_b$.[13] For the analytic continuation of four-dimensional observables to five or six dimensions, our default choice will be to define azimuthal angles in terms of p_1 and p_2 and rapidities in terms of p_0 and p_3 using the usual definitions. Explicitly:

$$y(p) = \frac{1}{2}\ln\left(\frac{p_0 + p_3}{p_0 - p_3}\right), \quad \Delta\phi(p_a, p_b) = \operatorname{acos}\left(\frac{p_{a,1}p_{b,1} + p_{a,2}p_{b,2}}{\sqrt{\left(p_{a,1}^2 + p_{a,2}^2\right)\left(p_{b,1}^2 + p_{b,2}^2\right)}}\right). \tag{3.1}$$

In practice, these are the only observables of five- and six-dimensional momenta that enter the calculation.[14]

We choose the central scale $\mu_R = \mu_F = p_T$, where p_T is the transverse momentum of the jet being binned. We use the NNLO NNPDF3.1 PDF+α_s set [41], taken from the LHAPDF interface [42], fixing $m_b = 4.92$ GeV to be consistent with that set.

We present four tests of our implementation. The first test is to make sure the S function contribution is independent of the choice of the parameters A, B and R in eq.(2.9). We can compute the four-dimensional finite part and the d-dimensional integrated subtraction term separately for two different choices of the parameters, verify that the individual contributions change and then check that their sum does not change. To perform this test, we compute the difference between our default choice and the choice $A = 0$, $B = 1$ and $R = 2m_b$. We explicitly present the result for the sum over all partonic configurations in the left panel of fig. 6. Clearly, the difference between the two parameter choices is non-zero for both the four-dimensional contribution and the integrated subtraction term, while this dependence cancels in their sum. It was checked that this also holds for each $2 \to 2$ partonic Born configuration separately. This test has thus been passed.

[13] The justification for this choice is as follows. Choosing $B = 2$ seems the most natural, while choosing $A = 3$ matches the behaviour of the S function for small three-momenta. Natural choices of R should be proportional to m_b. Which one of $R = m_b$ and $R = 2m_b$ is more natural is less obvious, but in the end these choices only affect numerical convergence and even there their impact should not be huge, provided they are chosen within reason.

[14] Note that the choice of analytic continuation of (infinitesimal) p_T has no impact on anti-k_T jets.

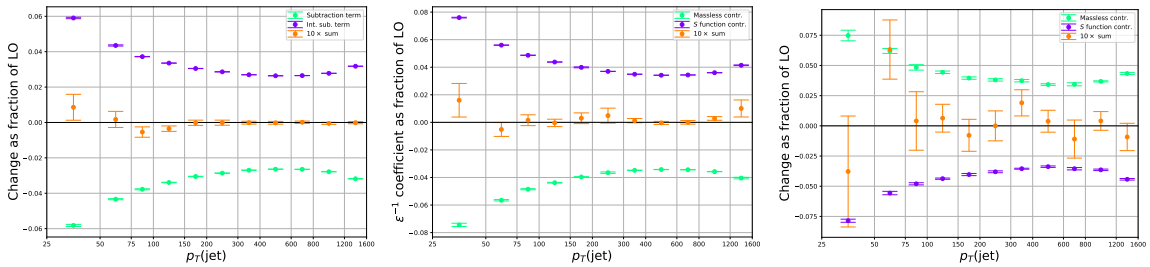

Figure 6. Left: the effect of changing the parameters A, B and R as described in the text on the subtraction term contribution (green), the integrated subtraction term contribution (purple) and their sum (orange, magnified 10×). Shown are the effects at NNLO as fractions of the LO cross section. Centre: the ε^{-1} contribution of the massless cross section (green), the S function contribution (purple) and their sum (orange, magnified 10×) as fractions of the LO cross section. Right: the effect of changing the d-dimensional definitions of y and $\Delta\phi$ as described in the text on the massless cross section (green), the S function contribution (purple) and their sum (orange, magnified 10×) as fractions of the LO cross section.

The second test is to make sure that the poles in ε cancel between the massless jet cross section and the contribution from the S functions. The central panel of fig. 6 shows this for the full NNLO result. We can clearly see that the massless cross section contains a leftover simple pole in ε - this is precisely the result of the non-IRC safety of the standard definition of jet flavour.[15] Adding the S function contribution cancels this leftover pole, rendering the whole calculation finite. Again, while only the results for the sum over all channels are shown here, it was checked that the poles also cancel for each combination of initial-state partons separately. Thus, this test has also been passed. Pole cancellation was also verified for the other processes discussed in the remainder of this work.

The third test is to verify that the dependence on how observables are defined beyond four dimensions drops out in the sum over contributions. In this case, we calculate the difference between results obtained using our default choice and those obtained using the following definitions of rapidity and $\Delta\phi$:

$$y(p) = \frac{1}{2}\ln\left(\frac{\sqrt{p^2+p_1^2+p_2^2+p_3^2}+p_3}{\sqrt{p^2+p_1^2+p_2^2+p_3^2}-p_3}\right) = \frac{1}{2}\ln\left(\frac{\sqrt{p_0^2-p_4^2-p_5^2}+p_3}{\sqrt{p_0^2-p_4^2-p_5^2}-p_3}\right), \tag{3.2}$$

$$\Delta\phi(p_a,p_b) = \operatorname{acos}\left(\frac{\sum_{i\in\{1,2,4,5\}} p_{a,i}p_{b,i}}{\sqrt{\left(\sum_{i\in\{1,2,4,5\}} p_{a,i}^2\right)\left(\sum_{i\in\{1,2,4,5\}} p_{b,i}^2\right)}}\right). \tag{3.3}$$

[15] Note that while in previous versions of Stripper, soft wide-angle flavour-changing emissions were not subtracted in a way that preserves all information on the soft partons, leading to non-integrable integrands, this updated version of Stripper fixes this limitation, fully locally subtracting the singularity and correctly adding the corresponding ε-pole fully differentially. As a result, this new version can be used to numerically test the safety of jet flavour definitions through NNLO by confirming the cancellation of ε-poles. Alternatively - and more efficiently - it can be checked that the S function contribution vanishes identically, i.e. does not contribute for any phase space point.

The result of this test is shown in the right panel of fig. 6. Again, the massless cross section clearly depends on the choice of analytic continuation of observables. The dependence of the S function contribution is equal and opposite, so the full cross section is independent of these choices. As before, only the results for the sum over all $2 \to 2$ partonic Born configurations are shown here, but it was checked that the dependence drops out also for each Born configuration separately. Therefore, this third test has also been passed.

Note that these first three tests are mathematically equivalent, since the terms being added to yield zero are always proportional to the non-cancelling ε-pole of the massless cross section. However, they each test completely independent parts of the practical implementation. In fact, combined with the final test, they cover all aspects of the implementation.

The final test is obvious: the massified cross section should match the massive one up to power corrections. In order to eliminate the effect of power corrections and maximise the size of the quasi-soft logarithms, this test should be performed for a very small value of the massive quark mass. However, this will lead to very large cancellations between different contributions to the massive calculation: the collinear (and part of the soft) logarithms cancel in the final result, but are present in the individual contributions to the cross section. This means that the individual cross section contributions need to be determined with unusually high numerical precision, i.e. a small Monte Carlo uncertainty. While not impossible, it would be unnecessarily computationally expensive to perform this test for an LHC cross section. For the sake of the environment, we will therefore instead consider the cross section for producing top-jets at a hypothetical 1 TeV electron-positron collider, setting $m_t = 1$ GeV and including only the QED contribution for simplicity. This process actually constitutes a very rare exception which cannot be handled by the original formulation of the sector-improved residue subtraction scheme. To solve this, we have extended the scheme to cover such edge cases, as described in appendix A. The small mass also makes the cross section hard to integrate using Monte Carlo techniques due to the presence of quasi-singularities. This is also solved by the strategy described in appendix A.

We cluster jets using the Durham algorithm with $y = 10^{-2}$ and study the distribution of the zenith angle of the highest-energy top-jet. The left panel of fig. 7 shows the different NNLO contributions to the leading power cross section: the purely massless contribution (blue), the α_s decoupling contribution (pink) and the S function contribution (orange). Also shown are the NNLO contributions to the fully massive cross section, distinguishing between final states with 2 (green) or 4 (purple) top quarks and plotted with an additional minus sign. If the LP contributions have all been computed correctly, then they should - up to negligible power corrections - add up to the fully massive cross section. Thus, adding all of the plotted contributions listed above should yield zero. Note that all plotted contributions have a different shape, requiring a non-trivial interplay in order for them to sum up to zero. Their sum is shown in grey, and is indeed consistent with zero.

This can be seen more clearly in the right panel of fig. 7. Shown there are the S function contribution in orange and the sum of all other contributions in blue. Clearly, both are constant w.r.t. the zenith angle in units of the LO cross section and their sum is zero. Among

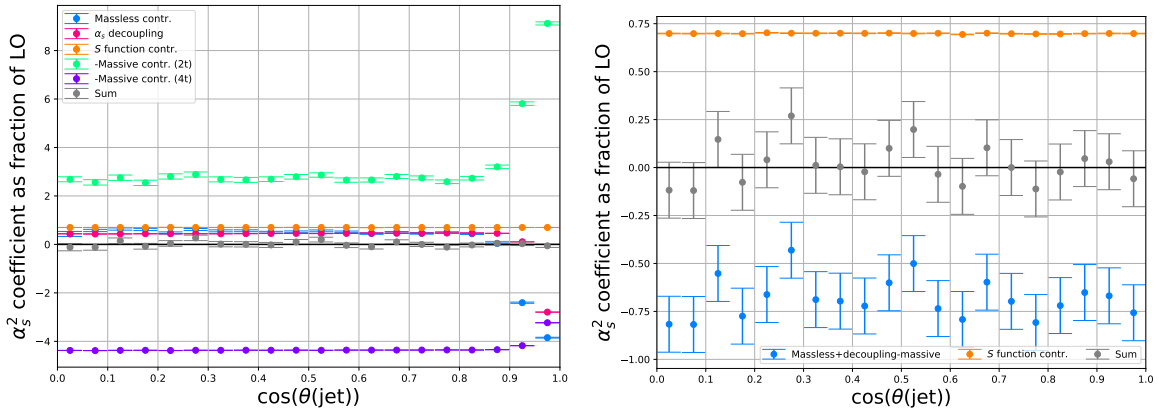

Figure 7. Left: shown are the different NNLO contributions to the LP hypothetical e^+e^- to top-jet cross section described in the text, namely the purely massless contribution (blue), the α_s decoupling contribution (pink) and the S function contribution (orange), as well as the NNLO contributions to the fully massive cross section, distinguishing between final states with 2 (green) or 4 (purple) top quarks and plotted with an additional minus sign. The sum of all of these contributions is shown in grey. Right: the S function contribution to the LP cross section (orange), the sum over all contributions shown in the left panel except for the S function contribution (blue) and the sum of both of these (grey).

those presented here, this is the most powerful test of the implementation of our idea - and indeed of the concept itself. And it, too, has been passed.

Having confirmed that everything is working as expected, we now turn to a pair of phenomenological applications.

4 Phenomenology

4.1 b-jet production in association with a Z-boson at the LHC

We will first consider the transverse momentum spectrum of the leading b-jet produced in association with a Z-boson (decaying to a pair of muons) at the 13 TeV LHC. This process was studied extensively in ref. [9] using several different IRC-safe definitions of jet flavour. To facilitate a direct comparison with the results of that study, we will use an identical setup for our calculation. This means we will apply the following cuts:[16]

- Both muons must satisfy $p_T(\mu) > 20$ GeV and $|y(\mu)| < 2.4$
- $p_T(\mu\mu) > 20$ GeV and 71 GeV $< m(\mu\mu) <$ 111 GeV
- At least one flavour-modulo-2 b-tagged $R = 0.5$ anti-k_T jet with $p_T(b\text{-jet}) > 30$ GeV, $|y(b\text{-jet})| < 2.4$ and $\Delta R(\mu, b\text{-jet}) > 0.5$ w.r.t. both muons.

[16]Note that the lower invariant mass cut is erroneously listed as 77 GeV in ref. [9] instead of 71 GeV.

Furthermore, we will use the PDF4LHC21 PDF and α_s set [43] and use the G_F scheme for the electroweak coupling, with $m_W = 80.352$ GeV, $m_Z = 91.1535$ GeV, $\Gamma_W = 2.084$ GeV, $\Gamma_Z = 2.4943$ GeV and $G_F = 1.16638 \cdot 10^{-5}$ GeV^{-2}. As before, we set $m_b = 4.92$ GeV. We will perform a 7-point scale variation around $\mu_R = \mu_F = \sqrt{m^2(\mu\mu) + p_T^2(\mu\mu)}$. The only difference between our calculation and the calculations of ref. [9] is the approach to flavour tagging.

We will particularly focus on comparing our results obtained using anti-k_T to results obtained using the CMP jet algorithm [12].[17] This algorithm is identical to anti-k_T, except it modifies the distance measure for pairs of pseudojets with non-zero net flavour numbers that are equal in magnitude but opposite in sign as follows:

$$\begin{aligned} d_{ij}^{\text{flavoured}} &= d_{ij}^{\text{anti-}k_T} \times \left[1 - \theta(1-\kappa_{ij}) \cos\left(\frac{\pi}{2}\kappa_{ij}\right)\right] , \\ \kappa_{ij} &= \frac{1}{a} \frac{p_{T,i}^2 + p_{T,j}^2}{p_{T,\text{max}}^2} \sqrt{\frac{\Omega_{ij}^2}{\Delta R_{ij}^2}} , \qquad \Omega_{ij}^2 = 2\left[\frac{1}{\omega^2}\big(\cosh(\omega \Delta y_{ij}) - 1\big) + \big(\cos(\Delta\phi_{ij}) - 1\big)\right] , \end{aligned} \tag{4.1}$$

where a is a new free parameter, $p_{T,\text{max}}$ is the highest transverse momentum among the pseudojets at a given point in the clustering procedure and $\omega > 1$ is another free parameter. We will choose $\omega = 2$ throughout this work and fix $a = 0.1$ in this subsection.

The results are shown in the left panel of fig. 8. We explicitly compare our approach (red) to the CMP (yellow) and IFN [7] (lime) algorithms, which together give a good impression of the general spread of results produced by the different algorithms tested in ref. [9]. At high transverse momentum, our approach yields results very similar to those of the other two algorithms. For lower transverse momenta, however, our approach yields a cross section that is about 10% lower than the other algorithms. This is noteworthy, since all the algorithms tested in ref. [9] yielded results that lie within just a few percent of each other at low p_T.

One might suspect that this deviation is the result of the 'large' logarithm $\ln p_T^2/m_b^2$ introduced by the S function contribution. However, if this were the case, one would expect a deviation at high transverse momentum, where this logarithm is largest, not at low transverse momentum. Indeed, the logarithm is not to blame, as can be concluded from the right panel of fig. 8, which directly compares the NNLO results obtained using CMP, using our approach and the result obtained using our approach when increasing the b-quark mass by a factor of 2.[18] First of all, the effect is very small - at most only about 2%. Since the relevant 'hard' scale in the first p_T bin (where the largest deviation relative to the other flavour definitions is observed) is $R \cdot p_T \sim 15$ GeV, the 'large' logarithm would be eliminated if $m_b \sim 15$ GeV. Since a factor of 2 already yields a mass reasonably close to that, it seems unlikely that the 'large' logarithm could give a contribution beyond just a few percent. More importantly,

[17] The original formulation of the CMP algorithm had some flaws, which were pointed out in ref. [7]. In that work, a modification of the algorithm was suggested to resolve these issues. Later, the algorithm was further modified in ref. [9]. We will employ this latest version of the CMP algorithm throughout this work.

[18] We only change the b-quark mass in the S function contribution, since this is the only source of non-resummed (and therefore potentially problematic) logarithms of the mass.

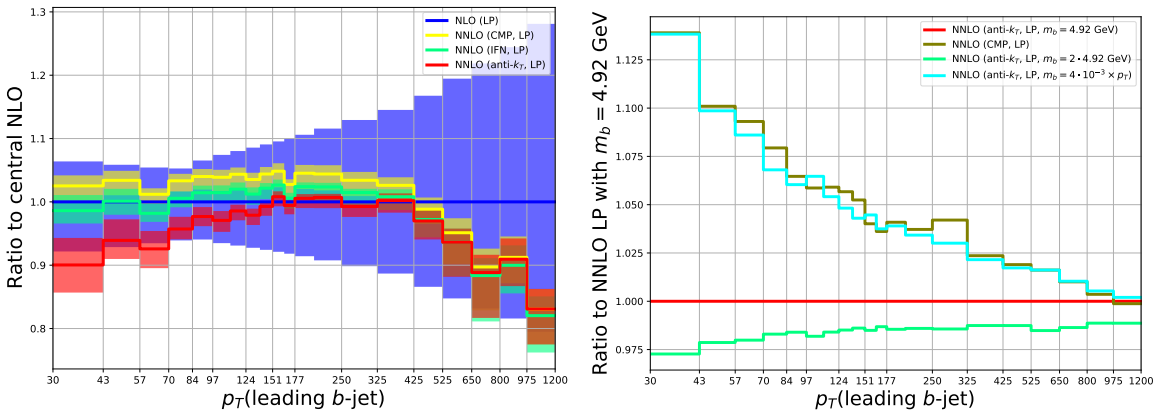

Figure 8. Left: the leading-b-jet transverse momentum spectrum for $Z + b$-jet production at NLO and NNLO, using different jet algorithms. The results obtained using CMP and IFN were taken from ref. [9]. Right: a comparison of the difference between the NNLO predictions obtained using anti-k_T and CMP at LP, and the impact of doubling the b-quark mass. Also shown is a result using anti-k_T with a small, dynamical mass.

increasing the b-quark mass - and thus reducing the 'large' logarithm - actually decreases the cross section, further increasing the gap between our approach the those tested in ref. [9].

This now suggests that one might be able to reproduce the results of the other methods by using a smaller b-quark mass, increasing the 'large' logarithm. Most of the flavoured jet algorithms in the literature effectively introduce a cutoff energy, below which clustering two flavoured jets is favoured. E.g. in the case of CMP, where clustering two flavoured jets is increasingly favoured for transverse momenta below roughly $\sqrt{a} \cdot p_{T,\text{max}}$. This puts an effective cutoff on the energy of soft flavoured quark pairs, regulating the singularity and introducing the logarithm $\ln a$. If one now imagines taking a to be incredibly small, then the cross section is obviously sensitive to unphysical regions of the phase space, where the flavoured quarks are allowed to be softer than they could physically be, due to their finite mass. In the physical cross section, which our approach reproduces up to power corrections, the finite b-quark mass sets the effective energy cutoff. Any flavoured jet algorithm should thus ensure that the effective energy cutoff it implements is never below this physical effective energy cutoff. This is especially important at low p_T, where probing unphysically small scales can easily lead to inaccurate cross sections, as has been known for e.g. inclusive heavy quark production for a long time [44].

The cyan curve in fig. 8 shows the cross section obtained when using anti-k_T with a dynamical b-quark mass of $4 \cdot 10^{-3}$ times the p_T of the b-jet. For this dynamical choice of the mass, it would appear that our approach can reproduce the CMP results, suggesting this is the 'effective' b-quark mass implemented by that algorithm in this case for $a = 0.1$. This would certainly qualify as being too low for transverse momenta below 1 TeV. However, it is worth remembering that all of these results - including the literature results we are comparing

to - are only accurate up to power corrections in the b-quark mass. It is, in principle, possible that different approaches lead to wildly different power corrections, which could also explain the observed differences.

To test this, we can compare massless results to massive ones and study the dependence of the transverse momentum spectrum on the b-quark mass. To this end we will consider a slightly simplified setup to aid with numerical convergence. This time, the Z-boson will be on-shell and we will apply the following cuts:

- $p_T(Z) > 20$ GeV and $|y(Z)| < 2.4$
- At least one b-tagged jet with $p_T(b\text{-jet}) > 30$ GeV and $|y(b\text{-jet})| < 2.4$.

We will use the $n_f = 4$ variant of PDF4LHC21 [43]. For the massless results, we will include - in addition to the S function - the PDF and α_s [45–49] threshold matching conditions through fixed-order NNLO, as previously implemented in STRIPPER in ref. [17]. This ensures that our massless results are LP accurate. Note that, since we are working with $n_f = 4$, the leading contribution to the cross section is one order higher than before: whereas before, the Born process corresponded to the final state $Z + b/\bar{b}$, this is no longer allowed. The leading contribution thus corresponds to the final state $Z + b + \bar{b}$. We will thus only need to consider NLO massive cross sections. We will label the LP cross sections in this case as 'NLO', since it corresponds to the LP approximation of the NLO massive cross section, but its singularity structure is that of an NNLO cross section, with collinear singularities already present at 'LO'. Crucially, this also means that the problematic double-soft b-quark-pair singularity is present at 'NLO'.

The results are shown in the left panel of fig. 9. Shown are NLO predictions for three values of the b-quark mass: 5.0 GeV (solid lines), 0.5 GeV (dashed lines) and 0.05 GeV (dotted lines). All predictions are shown as ratios w.r.t. the corresponding fully massive computations obtained using standard anti-k_T. For a mass of 5.0 GeV, the results obtained using CMP (yellow) differ from the LP results obtained using standard anti-k_T by around 20%, in line with the previously observed 10% deviation. The fact that we do not get the exact same deviation is not surprising, since the relative weight of different contributions for $n_f = 4$ and $n_f = 5$ can be very different. What is interesting, however, is that the 'correct' result, i.e. the fully massive result, happens to lie almost exactly in the middle between the two LP predictions. This is of course a numerical coincidence, but it highlights that neither prediction is accurate at low p_T: the presence of power corrections on the order of $5 - 10\%$ limits the accuracy of any massless computation in this regime. This means that, to some extent, it is pointless to argue over which approach to defining jet flavour is best, if different approaches typically only differ by a few percent - an effect which is negligible compared to the size of power corrections which all approaches neglect. Furthermore, this also means that if a precision greater than about 5% is desired, then this leaves us with no choice but to compute (at least part of) the cross section using massive quarks.

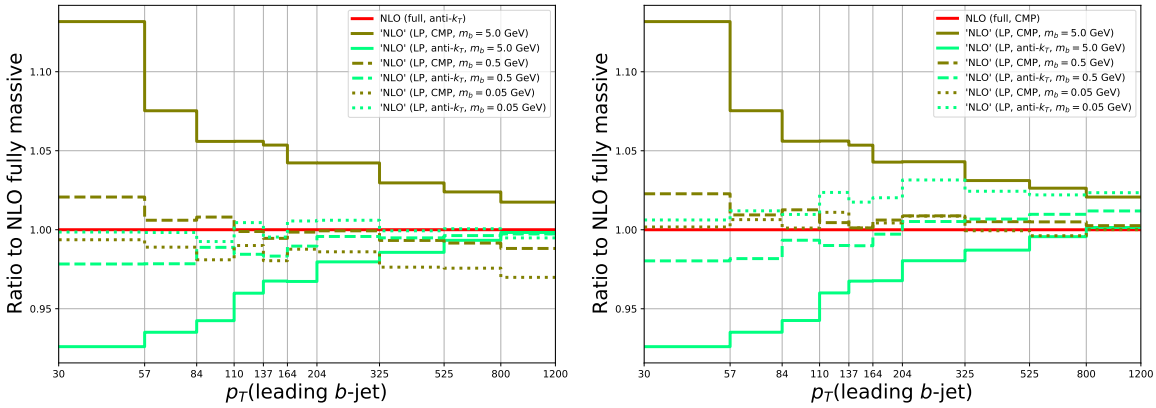

Figure 9. Left: ratios of the 'NLO' LP anti-k_T and CMP cross sections to the NLO fully massive anti-k_T cross section for $m_b = 5.0$, 0.5 and 0.05 GeV. Right: the same as on the left, but as ratios to the NLO fully massive CMP cross section.

Considering now the curves corresponding to other values of the b-quark mass shown in the left panel of fig. 9, we can see that, as the mass is reduced, the LP anti-k_T cross section approaches the fully massive one, as expected. It is interesting to note that the rate at which the two curves approach each other, however, is extremely slow. One would naively expect the power corrections to already be completely negligible for a mass of 0.5 GeV, even for the lowest transverse momenta shown. However, as it turns out, they are still clearly visible. It would appear - perhaps unsurprisingly - that this cross section suffers from linear power corrections, which at higher orders are further enhanced by logarithms of the b-quark mass. As a result, the power corrections only decrease by a factor of 4 or so as the mass is reduced from 5 GeV to 0.5 GeV.

Shifting focus to the CMP curves, we can see that it does not approach the fully massive result in the small-mass limit. Instead, each bin is shifted down at high p_T by a roughly fixed amount for each factor of ten by which the mass is reduced. This is exactly the behaviour expected from results which differ by a single logarithm of the mass. We can also see that, at low p_T, the first factor of ten has a larger effect than the second. This suggests that the power corrections are negative for CMP, while they were positive for anti-k_T. This thus already explains a large portion of the 10% difference between the two algorithms observed before.

The power corrections being negative for CMP can also be explicitly seen in the right panel of fig. 9, which shows the same curves as the left panel, but this time as ratios to the fully massive CMP cross section. As expected, the LP CMP curves properly converge to the fully massive result, while the LP anti-k_T curves do not. This shows that for both anti-k_T and CMP, the LP results are indeed LP-accurate. This constitutes a further test of the implementation, along the same lines as the lepton-collider test in the previous section, except now for an LHC process. The $m_b = 0.05$ GeV results would not have been feasible

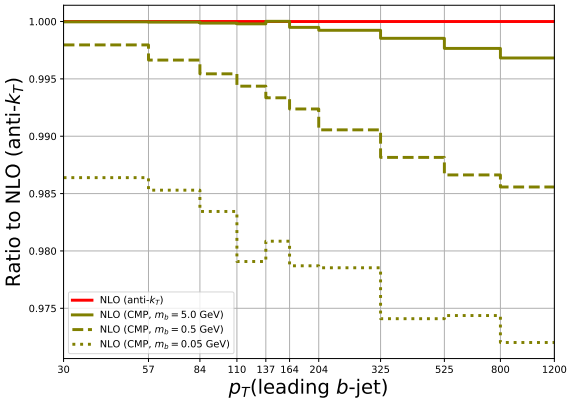

Figure 10. Comparison of the NLO fully massive cross sections obtained using anti-k_T and CMP for $m_b = 5.0$, 0.5 and 0.05 GeV.

without the use of the optimisation technique described in appendix A, which sped up the calculation by roughly two orders of magnitude.[19]

The use of anti-k_T and CMP leading to opposite-sign power corrections in the mass at least partially explains the 10% difference between the corresponding LP cross sections. However, this does not rule out the possibility that the effective energy cutoff of CMP for $a = 0.1$ is too low at low p_T. To test this conclusively, we can study the difference between CMP and anti-k_T jets for fully massive computations. If the effective energy cutoff introduced by CMP exceeds that of the mass, then switching from anti-k_T to CMP reduces the cross section. If it is lower or the same, then the cross section remains unaffected by this change. Such a study is shown in of fig. 10, which shows the ratio between fully massive results obtained using CMP and anti-k_T. For a mass of 5.0 GeV, there is no difference between the two cross sections up until around 160 GeV, at which point the CMP cross section starts to be suppressed relative to the anti-k_T one. This suggests that around 160 GeV, the effective energy cutoff on soft b-quarks introduced by CMP using $a = 0.1$ matches the physical energy cutoff. For lower transverse momenta, a massless calculation would thus be allowed to probe an unphysical region of the phase space, yielding overestimated cross sections. This statement of course depends on the process, phase space cuts, etc., but it highlights that great care must be taken when employing such algorithms, since their effective cutoff might be significantly lower than initially expected. Note that these results are consistent with some of the results of ref. [9]: e.g. fig. 16 of that work also shows that the impact of switching jet flavour definitions is much smaller for cross sections based on massive quarks than for those based on massless quarks, hinting at the same problem.

The curve for $m_b = 0.5$ GeV suggests that for that mass, the point where CMP potentially

[19]Within this optimisation approach, the NLO fully massive calculation in this particular case is treated as an 'NNLO' calculation, since its (quasi-)singularity structure is that of an NNLO cross section. This is the same reason why the 'NLO' LP cross section had to be computed using NNLO techniques, as mentioned before.

starts probing flavoured quarks with unphysical kinematics lies around 40 GeV. This is roughly in line with our earlier observation that the effective mass introduced by CMP for $a = 0.1$ is about $4 \cdot 10^{-3} \cdot 40$ GeV $= 0.16$ GeV for $p_T = 40$ GeV. Correcting this estimate for the different power corrections of CMP and anti-k_T has increased this to 0.5 GeV, which is still very low.

Finally, looking at the curve for $m_b = 0.05$ GeV, we can see that it is shifted relative to the curve for $m_b = 0.5$ GeV by a constant. In fact, at the highest transverse momenta, the relative shift is the same every time the mass is reduced by a factor of ten. This is again as expected from cross sections which differ by a single logarithm of the quark mass. Since the curve for $m_b = 5.0$ GeV smoothly transitions to unity for decreasing transverse momentum, a better estimate of where CMP with $a = 0.1$ starts probing unphysical kinematics is obtained by finding where the 5.0 GeV curve would intersect unity if it were obtained by adding the relative shift between the 0.05 GeV and 0.5 GeV curves to the 0.5 GeV one. This point would lie around roughly 600 GeV. This is again consistent with our earlier estimate of the effective mass introduced by CMP, up to a factor of two or so.

4.2 Single-inclusive b-jet production at the LHC

We will now consider the inclusive b-jet p_T spectrum at the LHC, using the same setup already discussed in the previous section in the context of numerical checks. To our knowledge, this cross section has not been considered at NNLO before. As before, missing-higher-order uncertainties are estimated by performing standard 7-point scale variation. Since we will not be considering fully massive cross sections in this subsection, it is understood that all results correspond to LP cross sections.

The left panel of fig. 11 shows the results obtained using anti-k_T through NNLO. The NLO curve is only marginally consistent with the LO result and the NNLO curve shows very poor perturbative convergence below 300 GeV. The latter is evident both from the scale uncertainties, which do not decrease after adding the NNLO corrections, and from the size of the NNLO corrections, which exceed the size of the NLO scale band. Additionally, the NNLO/NLO K-factor is significantly less flat than the NLO/LO K-factor. We confirmed that increasing the mass again decreases the NNLO cross section, worsening the perturbative convergence. The 'large' logarithm of the b-quark mass is therefore again not to blame. Instead, it is known [50–53] that single-inclusive jet cross sections exhibit poor perturbative convergence, or rather, that the scale uncertainties give an inaccurate impression of the size of missing higher orders for these processes. It would seem that our results confirm this to be the case for single-inclusive flavoured jet cross sections as well.

Interestingly, the situation looks different for the cross section obtained using the CMP algorithm. This is shown in the right panel of fig. 11. We consider cross sections for $a = 0.01$, 0.1 and 0.5. Clearly, the cross section for $a = 0.01$ converges very poorly: the NNLO scale uncertainties are much larger than at NLO and the NNLO curve lies well outside the NLO scale band. This is of course to be expected for such a small value of a, since it effectively introduces a large $\log \ln a$ with a coefficient roughly proportional to the LO cross section. Interestingly, the NNLO/NLO K-factor is very flat for all shown values of a. As the value of

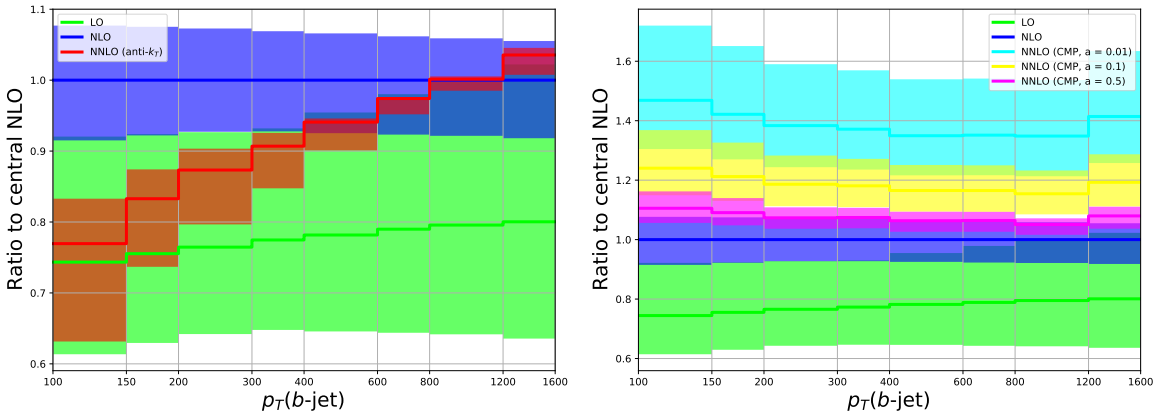

Figure 11. Left: the single-inclusive b-jet cross section using anti-k_T at LO (green), NLO (blue) and NNLO (red). Error bands indicate the scale uncertainties. Right: as on the left, but at NNLO using CMP with $a = 0.01$ (cyan), $a = 0.1$ (yellow) and $a = 0.5$ (purple).

a increases, the apparent convergence significantly improves. For the unusually large value $a = 0.5$, the series naively appears to be convergent. However, given the long history of incorrect conclusions based on apparent perturbative convergence that has plagued this class of processes, we do not wish to jump to any conclusions based on such criteria alone.

Another interesting aspect is the strong dependence of the NNLO cross section on a. In ref. [12], the effect on the $Z + b$-jet cross section of changing a from 0.1 to 0.5 was found to be just a few percent at high p_T. For single-inclusive b-jet production, the effect is roughly 15% - significantly more sizeable. This suggests this process is much more sensitive to the choice of jet flavour definition, making it a potentially interesting test bed for future comparisons of different approaches to jet flavour.

Comparing now the two algorithms, we reach similar conclusions as in the previous subsection. First, CMP tends to yield higher cross sections than anti-k_T. At low p_T, the difference is probably again mostly the result of power corrections. But at high p_T, this seems quite unlikely. Interestingly, we found that changing m_b to $\mathcal{O}(\text{few } 10^{-2}) \times p_T$ again roughly reproduces the CMP result for $a = 0.1$, despite all the differences between this process and $p\,p \to Z\,b$-jet.

Second, one needs to choose a to be unexpectedly large in order for the two algorithms to give the same result. In fact, the differences are much larger for single-inclusive b-jet production than they are for $Z + b$-jet production: even for $a = 0.5$, the two algorithms roughly agree for only the highest p_T bin. Since such a large value of a surely sufficiently shields against soft b-quark pairs with unremarkable kinematic configurations, this suggests that maybe a certain unusual limit is the cause of these differences. Comparing the two algorithms sector by sector, we find that the biggest differences by far occur in the process $gg \to ggb\bar{b}$ for triple-collinear sectors involving a gluon reference and two unresolved b-quarks. These sectors obviously contain the problematic soft-quark-pair limit, but the differences for

other sectors containing that singularity are much milder. This suggests that the unphysical kinematics probed by CMP are related to configurations where a gluon and a b-quark pair are close to collinear. Further investigating the details of this phenomenon would require a dedicated study beyond the scope of this work. Whatever the cause of the differences, their large size further supports the idea that this process is highly sensitive to the definition of jet flavour.

As a concluding remark, it would be interesting to extend the results of ref. [53] to the flavoured case to see if including small-jet-radius resummation also improves the perturbative behaviour for single-inclusive flavoured jet production. Now that it is possible to use standard anti-k_T clustering, this should be significantly more manageable.

5 Conclusion

We have introduced a new approach to jet flavour which renders standard definitions of jets and their flavour infrared-collinear safe by correctly reintroducing the flavoured quark's mass at leading power. By analysing all sources of mass dependence, we were able to reduce the required amount of theory effort to a minimum: only simple modifications to suitably-constructed subtraction schemes are needed and a single simple additional contribution has to be implemented and computed. The numerical integration of this extra contribution converges much more quickly than the rest of the calculation, meaning there is not even any extra computational cost. Meanwhile, since our approach is compatible with all commonly used jet clustering algorithms and yields - up to power corrections - the cross section one would obtain if one were to compute the cross section using the physical quark mass, there is no need for experimental collaborations to change their procedures at all. They can simply keep using anti-k_T (or any other jet algorithm of their choice) and unfold from particle-level any-flavour scheme jets to parton-level flavour modulo-2 scheme jets, as they have done for many years.

Through arguments based on the logarithmic structure of the cross section and using explicit phenomenological studies, we have demonstrated that this approach does not lead to logarithms of the heavy quark mass sufficiently large to pose a threat to perturbative convergence. Instead, we have found that power corrections in the mass easily exceed the size of such logarithms. This implies that, if one is to achieve a precision sufficiently high to notice these 'large' logarithms, then one needs to take into account the full mass dependence of the cross section, at which point there is no need to use special flavoured jet algorithms. Additionally, we have found that such flavoured jet algorithms may not necessarily prevent unphysical regions of the phase space from being probed. This issue certainly deserves a dedicated study testing multiple jet flavour definitions in this regard, which we leave for future work.

We have also extended the sector-improved residue subtraction scheme and its implementation in Stripper to be capable of handling processes without massless partons at the Born level. Not only can this extension be used for phenomenological studies of such processes in

the future, it can also be leveraged as an extremely powerful optimisation technique for cross sections containing quasi-singularities.

In the future, our approach can be used to obtain results for a wide range of jet processes at NNLO, which can be directly compared to measurements made using standard jet algorithms. We also expect our work to play an important role in ongoing discussions between theory and experiment on how to move forward regarding the jet flavour issue.

Acknowledgments

We are especially grateful to Michał Czakon for making the STRIPPER library available to us. We would also like to thank René Poncelet for a careful reading of the manuscript, for helpful discussions about literature results on jet flavour and for providing some of the numerical results of ref. [9]. T.G. has been supported by STFC consolidated HEP theory grants ST/T000694/1 and ST/X000664/1. This work was performed using the Cambridge Service for Data Driven Discovery (CSD3), part of which is operated by the University of Cambridge Research Computing on behalf of the STFC DiRAC HPC Facility (www.dirac.ac.uk). The DiRAC component of CSD3 was supported by STFC grants ST/P002307/1, ST/R002452/1 and ST/R00689X/1.

A Massive sectors

One of the numerical checks presented in section 3 involved the computation of fully differential NNLO cross sections for the process $e^+e^- \to Q\overline{Q}$, where Q is a massive quark. While such computations have been available for a while [54–56], they pose an interesting problem for the sector-improved residue subtraction scheme. Indeed, while the scheme and its implementation in STRIPPER are fully general, meaning they can be used to compute almost any cross section at NNLO, there is a very small set of processes they previously could not describe. This is because the construction of the subtraction scheme relies on defining sectors relative to massless reference partons. In very rare cases of phenomenological interest - essentially only $e^+e^- \to Q\overline{Q}$ - there are no massless partons at leading order. As a result, no sectors can be defined and the subtraction terms cannot be constructed according to the sector-improved residue subtraction scheme.

Thus, for the purpose of the present study, we extended the definition of the subtraction scheme and its implementation in STRIPPER to be able to handle Born processes without massless partons, covering this final edge case and finally rendering the subtraction scheme truly general - without exceptions. To do this, we use the following two-step procedure:

1. Pretend the massive partons are massless and construct the $n+2$ phase space accordingly, using the conventional phase space parametrisation of the subtraction scheme.

2. Massify the phase space using the RAMBO algorithm [57].

While seemingly incredibly simple, there are several technical details that require great care. This approach also has a major numerical advantage, which we will discuss later.

First, let us consider this approach at NLO. The premise is that there are only massive partons at LO, so at NLO there are only massive partons and one single gluon. Clearly, by first taking the massive partons to be massless, at least one sector can always be defined, allowing us to use the usual construction of the subtraction scheme to generate the massless phase space. The resolved $(n+1)$-particle phase space generated by the two-step procedure is trivially correct: we start from a correct and complete parametrisation of the massless phase space, so the RAMBO procedure is guaranteed to yield the correct corresponding massive one. This only leaves the (integrated) subtraction terms. Clearly, there are no (soft-)collinear singularities, so the corresponding (integrated) subtraction terms will have to be removed. There is a soft singularity, however due to the simple and fully numerical nature of the subtraction scheme - defining each subtraction term in a single strict limit and integrating over all other variables numerically - there is no need to redefine subtraction terms: they are defined point-by-point in the phase space using the particles' actual momenta in the limit, instead of relying on complicated momenta mappings. Thus, they do not care about how a phase space point was generated, only what it looks like in the relevant limit. Since the integrated subtraction terms involve no analytical integration to begin with, there is also no need to recompute any analytical integrals. The only point of care here is the strict limit: although the gluon is infinitely soft and does not contribute to observables, its momentum must also be rescaled during the RAMBO procedure, just like all the non-vanishing momenta. This affects both the phase space weight of the subtraction term (through the RAMBO reweighting factor) and the factorised soft singularity, since it depends on the regulated gluon energy (u^0/ξ in the notation of ref. [35]). Similarly, the gluon should be treated as d-dimensional for the integrated subtraction term. This only affects the phase space weight through the substitution[20]

$$\xi^{2n-3} \to \xi^{2n-3-2\varepsilon} \tag{A.1}$$

in the reweighting factor of the RAMBO procedure.

At NNLO, two things change. First, there are now up to two soft partons at a time. However, there are no extra modifications necessary beyond combining those for each individual soft parton. The second difference is the presence of a collinear singularity between the two additional massless partons. However, as before, the implementation of the corresponding subtraction terms in Stripper is done directly based on the generated momenta, not on the initial phase space parameters, so once again no further modifications are necessary. One should just take care to reweight the phase space using the RAMBO formula for $n+2$ partons, rather than treating the two collinear partons as a single parton.

This strategy is thus incredibly simple and elegant. Perhaps its greatest feature, however, is not that it finally enables NNLO computations of processes with no massless partons at

[20] Here 'ξ' is the factor by which 3-momenta are rescaled during the RAMBO massification, not the variable mentioned a few lines earlier parametrising the unresolved parton's energy within Stripper.

the Born level using the sector-improved residue subtraction scheme. Instead, its greatest strength is its application as a powerful optimisation technique for cross section calculations involving very light massive quarks. For example: for $e^+e^- \to Q\overline{Q}$, it introduces two sectors at NLO: one sector parametrising the gluon relative to the quark and one sector parametrising the gluon relative to the anti-quark. While there is no collinear singularity between the massive quarks and the gluon, there is a quasi-collinear singularity if the mass of the quark is small compared to its energy. While there is no need to define the gluon's direction relative to anything in this process - one could just sample its direction uniformly and only worry about parametrising its energy in a way that facilitates the use of a subtraction scheme - it is incredibly useful to do so. If the gluon's direction were sampled uniformly, then these quasi-collinear singularities would be located on complicated hyperplanes within the $(0,1)^m$ hypercube of integration variables. From a Monte Carlo perspective, if the mass is small enough, this would appear as very strongly peaked structures within the hypercube, which are only sampled rarely and randomly. This leads to very poor numerical convergence of the Monte Carlo integration.

The construction of the phase space within the sector-improved residue subtraction scheme directly parametrises the angle between the reference and unresolved partons in a sector. Additionally, the selector functions guarantee that a given quasi-collinear singularity is contained within a single sector. A given quasi-collinear singularity is thus located entirely on one of the faces of the hypercube of a single sector, making it trivial to importance sample the peaked behaviour of the cross section. This is especially impactful at NNLO, where the quasi-singularity structure is much more complicated, with nested quasi-collinear limits and double-quasi-soft massive quark pairs.

Using this massification procedure makes importance sampling all these quasi-singularities trivial, speeding up the Monte Carlo convergence by several orders of magnitude - without any extra effort from the user - if the mass is small enough. This is often the case for numerical checks, like those performed in this paper, where minimising the Monte Carlo errors is very important, but typically challenging. However, even for phenomenological studies of e.g. massive b- or c-quark production at high energies, this technique can easily yield significant speed-ups.

This now suggests applications of the massification procedure to processes at hadron colliders, even though this is strictly speaking not necessary, since in those cases there will always be massless partons at the Born level. Indeed, this was used in section 4 to compute NLO cross sections for $pp \to Zb\bar{b}$ with very light b-quarks. In principle, this is straightforward: one simply adds the extra sectors involving massive partons, treating these massive sectors as described above. The original sectors without any massive reference or unresolved partons can be treated as usual, without the need to massify the phase space using a two-step procedure.[21]

However, applying the procedure to hadron colliders comes with a few more minor complications. First, the centre-of-mass energy of the massless phase space point might be below

[21] Although one certainly could do so, if desired.

the threshold of the massive phase space. One might be tempted to simply sample the PDF fractions in a way that guarantees that the total centre-of-mass energy of the initial massless phase space is sufficient. However, this is not that trivial in practice. If the sector involves final-state references only, then this will work trivially. However, if one of the reference partons is in the initial state, then, within the sector-improved residue subtraction scheme, the cross section and its subtraction terms will have different centre-of-mass energies, complicating the minimum-energy requirement. While it is not terribly difficult to simply come up with alternate phase space parametrisations for these cases, we instead opted to simply sample all centre-of-mass energies and reject contributions with insufficient energy. While technically less efficient - sometimes generating phase space points that cannot contribute to the cross section - in practice this inefficiency is negligible. Indeed, the premise for using this technique is that the massive quarks are light relative to their energy. This makes the difference between the initial massless phase space point and its massified version tiny. Thus, configurations with insufficient energy will not or only rarely be sampled anyway.

The final complication caused by initial-state references is due to certain assumptions usually built into subtraction scheme codes. Typically, if a parton is taken to be collinear to the initial state, it is essentially removed from the final state - it cannot be observed anyway. However, for the purpose of the massification, it has to be kept, just like for final-state collinear singularities. Furthermore, since a gluon becoming soft or collinear to an initial-state reference usually yields the same observable phase space point within the sector-improved residue subtraction scheme, there is typically no need to recompute the observable momenta for each of those limits separately. However, when using the massification procedure, they do differ: the energy of the gluon affects the massification of the entire phase space. Thus, there is an observable difference between the configurations with a non-soft collinear gluon and a soft gluon. This is not a problem for a final-state collinear singularity, since in that case the sum of the energies of the reference and unresolved partons does not depend on the softness of the collinear gluon, and so the observable phase space is unaffected.

References

[1] A. Banfi, G. P. Salam and G. Zanderighi, Eur. Phys. J. C **47**, 113-124 (2006) doi:10.1140/epjc/s2006-02552-4 [arXiv:hep-ph/0601139 [hep-ph]].

[2] M. Cacciari, G. P. Salam and G. Soyez, JHEP **04**, 063 (2008) doi:10.1088/1126-6708/2008/04/063 [arXiv:0802.1189 [hep-ph]].

[3] S. Catani, Y. L. Dokshitzer, M. H. Seymour and B. R. Webber, Nucl. Phys. B **406**, 187-224 (1993) doi:10.1016/0550-3213(93)90166-M and references therein

[4] S. D. Ellis and D. E. Soper, Phys. Rev. D **48**, 3160-3166 (1993) doi:10.1103/PhysRevD.48.3160 [arXiv:hep-ph/9305266 [hep-ph]].

[5] Y. L. Dokshitzer, G. D. Leder, S. Moretti and B. R. Webber, JHEP **08**, 001 (1997) doi:10.1088/1126-6708/1997/08/001 [arXiv:hep-ph/9707323 [hep-ph]].

[6] M. Wobisch and T. Wengler, [arXiv:hep-ph/9907280 [hep-ph]].

[7] F. Caola, R. Grabarczyk, M. L. Hutt, G. P. Salam, L. Scyboz and J. Thaler, Phys. Rev. D **108**, no.9, 094010 (2023) doi:10.1103/PhysRevD.108.094010 [arXiv:2306.07314 [hep-ph]].

[8] G. 't Hooft and M. J. G. Veltman, Nucl. Phys. B **44**, 189-213 (1972) doi:10.1016/0550-3213(72)90279-9

[9] A. Behring, S. Caletti, F. Giuli, R. Grabarczyk, A. Hinzmann, A. Huss, J. Huston, E. D. Lesser, S. Marzani and D. Napoletano, *et al.* JHEP **09**, 149 (2025) doi:10.1007/JHEP09(2025)149 [arXiv:2506.13449 [hep-ph]].

[10] S. Caletti, A. J. Larkoski, S. Marzani and D. Reichelt, Eur. Phys. J. C **82**, no.7, 632 (2022) doi:10.1140/epjc/s10052-022-10568-7 [arXiv:2205.01109 [hep-ph]].

[11] S. Caletti, A. J. Larkoski, S. Marzani and D. Reichelt, JHEP **10**, 158 (2022) doi:10.1007/JHEP10(2022)158 [arXiv:2205.01117 [hep-ph]].

[12] M. Czakon, A. Mitov and R. Poncelet, JHEP **04**, 138 (2023) doi:10.1007/JHEP04(2023)138 [arXiv:2205.11879 [hep-ph]].

[13] R. Gauld, A. Huss and G. Stagnitto, Phys. Rev. Lett. **130**, no.16, 161901 (2023) [erratum: Phys. Rev. Lett. **132**, no.15, 159901 (2024)] doi:10.1103/PhysRevLett.130.161901 [arXiv:2208.11138 [hep-ph]].

[14] A. J. Larkoski, [arXiv:2510.06031 [hep-ph]].

[15] M. Czakon, T. Generet, A. Mitov and R. Poncelet, JHEP **10**, 216 (2021) doi:10.1007/JHEP10(2021)216 [arXiv:2102.08267 [hep-ph]].

[16] M. Czakon, T. Generet, A. Mitov and R. Poncelet, JHEP **03**, 251 (2023) doi:10.1007/JHEP03(2023)251 [arXiv:2210.06078 [hep-ph]].

[17] M. Czakon, T. Generet, A. Mitov and R. Poncelet, Phys. Rev. Lett. **135**, no.16, 16 (2025) doi:10.1103/b6pf-rj4h [arXiv:2411.09684 [hep-ph]].

[18] T. Generet, R. Poncelet and M. Muškinja, [arXiv:2510.24525 [hep-ph]].

[19] L. Buonocore, S. Devoto, S. Kallweit, J. Mazzitelli, L. Rottoli and C. Savoini, Phys. Rev. D **107**, no.7, 074032 (2023) doi:10.1103/PhysRevD.107.074032 [arXiv:2212.04954 [hep-ph]].

[20] S. Weinberg, Phys. Lett. B **91**, 51-55 (1980) doi:10.1016/0370-2693(80)90660-7

[21] B. A. Ovrut and H. J. Schnitzer, Phys. Lett. B **100**, 403-406 (1981) doi:10.1016/0370-2693(81)90146-5

[22] M. A. G. Aivazis, J. C. Collins, F. I. Olness and W. K. Tung, Phys. Rev. D **50**, 3102-3118 (1994) doi:10.1103/PhysRevD.50.3102 [arXiv:hep-ph/9312319 [hep-ph]].

[23] M. Buza, Y. Matiounine, J. Smith, R. Migneron and W. L. van Neerven, Nucl. Phys. B **472**, 611-658 (1996) doi:10.1016/0550-3213(96)00228-3 [arXiv:hep-ph/9601302 [hep-ph]].

[24] M. Buza, Y. Matiounine, J. Smith and W. L. van Neerven, Eur. Phys. J. C **1**, 301-320 (1998) doi:10.1007/BF01245820 [arXiv:hep-ph/9612398 [hep-ph]].

[25] I. Bierenbaum, J. Blumlein and S. Klein, Nucl. Phys. B **780**, 40-75 (2007) doi:10.1016/j.nuclphysb.2007.04.030 [arXiv:hep-ph/0703285 [hep-ph]].

[26] I. Bierenbaum, J. Blumlein, S. Klein and C. Schneider, Nucl. Phys. B **803**, 1-41 (2008) doi:10.1016/j.nuclphysb.2008.05.016 [arXiv:0803.0273 [hep-ph]].

[27] I. Bierenbaum, J. Blumlein and S. Klein, Phys. Lett. B **672**, 401-406 (2009) doi:10.1016/j.physletb.2009.01.057 [arXiv:0901.0669 [hep-ph]].

[28] B. Mele and P. Nason, Nucl. Phys. B **361**, 626-644 (1991) [erratum: Nucl. Phys. B **921**, 841-842 (2017)] doi:10.1016/0550-3213(91)90597-Q

[29] K. Melnikov and A. Mitov, Phys. Rev. D **70**, 034027 (2004) doi:10.1103/PhysRevD.70.034027 [arXiv:hep-ph/0404143 [hep-ph]].

[30] A. Mitov, Phys. Rev. D **71**, 054021 (2005) doi:10.1103/PhysRevD.71.054021 [arXiv:hep-ph/0410205 [hep-ph]].

[31] M. Cacciari, P. Nason and C. Oleari, JHEP **10**, 034 (2005) doi:10.1088/1126-6708/2005/10/034 [arXiv:hep-ph/0504192 [hep-ph]].

[32] M. Neubert, [arXiv:0706.2136 [hep-ph]].

[33] T. Kinoshita, J. Math. Phys. **3**, 650-677 (1962) doi:10.1063/1.1724268

[34] T. D. Lee and M. Nauenberg, Phys. Rev. **133**, B1549-B1562 (1964) doi:10.1103/PhysRev.133.B1549

[35] M. Czakon and D. Heymes, Nucl. Phys. B **890**, 152-227 (2014) doi:10.1016/j.nuclphysb.2014.11.006 [arXiv:1408.2500 [hep-ph]].

[36] M. Czakon, Phys. Lett. B **693**, 259-268 (2010) doi:10.1016/j.physletb.2010.08.036 [arXiv:1005.0274 [hep-ph]].

[37] M. Czakon, Nucl. Phys. B **849**, 250-295 (2011) doi:10.1016/j.nuclphysb.2011.03.020 [arXiv:1101.0642 [hep-ph]].

[38] M. Czakon, A. van Hameren, A. Mitov and R. Poncelet, JHEP **10**, 262 (2019) doi:10.1007/JHEP10(2019)262 [arXiv:1907.12911 [hep-ph]].

[39] M. Dasgupta and G. P. Salam, Phys. Lett. B **512**, 323-330 (2001) doi:10.1016/S0370-2693(01)00725-0 [arXiv:hep-ph/0104277 [hep-ph]].

[40] A. Banfi, G. Marchesini and G. Smye, JHEP **08**, 006 (2002) doi:10.1088/1126-6708/2002/08/006 [arXiv:hep-ph/0206076 [hep-ph]].

[41] R. D. Ball *et al.* [NNPDF], Eur. Phys. J. C **77**, no.10, 663 (2017) doi:10.1140/epjc/s10052-017-5199-5 [arXiv:1706.00428 [hep-ph]].

[42] A. Buckley, J. Ferrando, S. Lloyd, K. Nordström, B. Page, M. Rüfenacht, M. Schönherr and G. Watt, Eur. Phys. J. C **75**, 132 (2015) [arXiv:1412.7420 [hep-ph]].

[43] R. D. Ball *et al.* [PDF4LHC Working Group], J. Phys. G **49**, no.8, 080501 (2022) doi:10.1088/1361-6471/ac7216 [arXiv:2203.05506 [hep-ph]].

[44] M. Cacciari, M. Greco and P. Nason, JHEP **05**, 007 (1998) doi:10.1088/1126-6708/1998/05/007 [arXiv:hep-ph/9803400 [hep-ph]].

[45] W. Wetzel, Nucl. Phys. B **196**, 259-272 (1982) doi:10.1016/0550-3213(82)90038-4

[46] W. Bernreuther and W. Wetzel, Nucl. Phys. B **197**, 228-236 (1982) [erratum: Nucl. Phys. B **513**, 758-758 (1998)] doi:10.1016/0550-3213(82)90288-7

[47] W. Bernreuther, Annals Phys. **151**, 127 (1983) doi:10.1016/0003-4916(83)90317-2

[48] W. Bernreuther, Z. Phys. C **20**, 331-333 (1983) doi:10.1007/BF01407825

[49] S. A. Larin, T. van Ritbergen and J. A. M. Vermaseren, Nucl. Phys. B **438**, 278-306 (1995) doi:10.1016/0550-3213(94)00574-X [arXiv:hep-ph/9411260 [hep-ph]].

[50] J. Currie, E. W. N. Glover and J. Pires, Phys. Rev. Lett. **118**, no.7, 072002 (2017) doi:10.1103/PhysRevLett.118.072002 [arXiv:1611.01460 [hep-ph]].

[51] J. Currie, A. Gehrmann-De Ridder, T. Gehrmann, E. W. N. Glover, A. Huss and J. Pires, JHEP **10**, 155 (2018) doi:10.1007/JHEP10(2018)155 [arXiv:1807.03692 [hep-ph]].

[52] J. Bellm, A. Buckley, X. Chen, A. Gehrmann-De Ridder, T. Gehrmann, N. Glover, A. Huss, J. Pires, S. Höche and J. Huston, *et al.* Eur. Phys. J. C **80**, no.2, 93 (2020) doi:10.1140/epjc/s10052-019-7574-x [arXiv:1903.12563 [hep-ph]].

[53] T. Generet, K. Lee, I. Moult, R. Poncelet and X. Zhang, JHEP **08**, 015 (2025) doi:10.1007/JHEP08(2025)015 [arXiv:2503.21866 [hep-ph]].

[54] J. Gao and H. X. Zhu, Phys. Rev. D **90**, no.11, 114022 (2014) doi:10.1103/PhysRevD.90.114022 [arXiv:1408.5150 [hep-ph]].

[55] J. Gao and H. X. Zhu, Phys. Rev. Lett. **113**, no.26, 262001 (2014) doi:10.1103/PhysRevLett.113.262001 [arXiv:1410.3165 [hep-ph]].

[56] L. Chen, O. Dekkers, D. Heisler, W. Bernreuther and Z. G. Si, JHEP **12**, 098 (2016) doi:10.1007/JHEP12(2016)098 [arXiv:1610.07897 [hep-ph]].

[57] R. Kleiss, W. J. Stirling and S. D. Ellis, Comput. Phys. Commun. **40**, 359 (1986) doi:10.1016/0010-4655(86)90119-0