On the reduction of negative weights in MC@NLO-type matching procedures

R. Frederix,[a] **S. Frixione,**[b] **S. Prestel,**[a] **P. Torrielli**[c]

[a] *Theoretical Particle Physics, Department of Astronomy and Theoretical Physics, Lund University, Sölvegatan 14 A, SE-223 62 Lund, Sweden*

[b] *INFN, Sezione di Genova, Via Dodecaneso 33, I-16146, Genoa, Italy*

[c] *Dipartimento di Fisica and Arnold-Regge Center, Università di Torino and INFN, Sezione di Torino, Via P. Giuria 1, I-10125, Turin, Italy*

E-mail: rikkert.frederix@thep.lu.se, Stefano.Frixione@cern.ch, stefan.prestel@thep.lu.se, torriell@to.infn.it

ABSTRACT: We show how a careful analysis of the behaviour of a parton shower Monte Carlo in the vicinity of the soft and collinear regions allows one to formulate a modified MC@NLO-matching prescription that reduces the number of negative-weight events with respect to that stemming from the standard MC@NLO procedure. As a first practical application of such a prescription, that we dub MC@NLO-Δ, we have implemented it in the MADGRAPH5_AMC@NLO framework, by employing the PYTHIA8 Monte Carlo. We present selected MC@NLO-Δ results at the 13 TeV LHC, and compare them with MC@NLO ones. We find that the former predictions are consistent with the latter ones within the typical matching systematics, and that the reduction of negative-weight events is significant.

Contents

1 Introduction

It is inconceivable that modern experimental high-energy particle physics be done without the massive use of event generators, which are also enjoying an increasing number of applications in theoretical phenomenology. This has stimulated a vigorous research activity in the past two decades, the upshot of which is that event generators, while maintaining their traditional flexibility, have significantly increased their predictive power (and its most important spinoff, the reduction of systematics), through matching and merging with perturbative matrix-element computations. Apart from a few selected cases, matching and merging are either carried out at the tree level or at the next-to-leading order (NLO). In this work, we shall consider only the latter simulations, that are more accurate than the former, but require more involved computations (which nowadays can fortunately be automated). On top of that, one must bear in mind that NLO cross sections are not positive-definite locally in the phase space. This implies that some of the hard events that will eventually be showered have negative weights. It is convenient to introduce an efficiency associated with the fraction of negative weights; denoting the latter by f, such an efficiency is defined as follows:

$$\varepsilon(f) = 1 - 2f, \qquad 0 \le f < 0.5 \qquad \Longrightarrow \qquad 0 < \varepsilon(f) \le 1. \qquad (1.1)$$

This efficiency stems from matrix-element computations (see e.g. eq. (4.33) of ref. [1]), but it affects the overall performance of the latter only insofar that it is correlated with other

efficiencies relevant to them; for example, it is usually the case that phase-space integration converges slightly faster in the case of a positive-definite integrand than in the case of an integrand whose sign can change; similar considerations are valid e.g. for the unweighting efficiency. However, for all practical purposes such correlations can be neglected: by far, the most significant impact of $\varepsilon(f)$ being smaller than one is that it implies that, in order to obtain the same statistical accuracy at the level of physical observables as the one of a positive-definite simulation, a number of events must be generated which is larger than in the latter case. In order to quantify this statement, let us denote by:

$$N\,, \qquad N_+ = (1-f)N\,, \qquad N_- = fN \qquad \Longrightarrow \qquad N = N_+ + N_-\,, \qquad (1.2)$$

the total number of events, the number of events with positive weights, and the number of events with negative weights, respectively. Furthermore, we assume such events to be unweighted[1]: in other words, their weights are equal to $\pm\omega$, with $\omega > 0$ a constant. The resulting cross section and its associated error will thus be:

$$\begin{aligned} \sigma &= \omega\left(N_+ - N_- \pm \sqrt{N_+ + N_- + 2C_\pm\sqrt{N_+}\sqrt{N_-}}\right) \\ &= \omega\left(\varepsilon(f)N \pm \sqrt{1 + C_\pm\sqrt{1-\varepsilon(f)^2}}\sqrt{N}\right), \end{aligned} \qquad (1.3)$$

where by $C_\pm$ we have denoted the non-negative correlation between positive- and negative-weight events (this number is small, and is typically neglected). In the context of a positive-definite simulation, where M unweighted events are generated with weights all equal to $\omega' > 0$, the analogue of eq. (1.3) reads:

$$\sigma = \omega'\left(M \pm \sqrt{M}\right). \qquad (1.4)$$

By imposing the cross sections in eqs. (1.3) and (1.4) to have the same relative error we obtain:

$$N = c(f)\,M\,, \qquad (1.5)$$

where

$$c(f) = \frac{1 + C_\pm\sqrt{1-\varepsilon(f)^2}}{\varepsilon(f)^2}\,. \qquad (1.6)$$

The fact that $N \geq M$ (with $N = M$ if and only if $f = 0$) formalises what was stated before eq. (1.2): a simulation that features events of either sign can attain the same statistical accuracy as one that is positive definite only by generating a number of events which is larger by a factor $c(f)$ w.r.t. that of the latter. For this reason, we call $c(f)$ the *relative cost* of the simulation; we plot this quantity as a function of f in fig. 1, for $C_\pm = 0, 0.5, 1$.

The discussion thus far has been rather schematic. For example, we have implicitly assumed f to be independent of the kinematics, which is never the case. More correctly,

---

[1]Discussing efficiencies becomes significantly more complicated in the case of weighted events. Of course, this is not the (main) reason why unweighted events must be preferred to weighted events whenever possible: rather, they constitute a much more realistic representation of actual physical events, and their samples are much smaller, for any given accuracy target, than those relevant to weighted events.
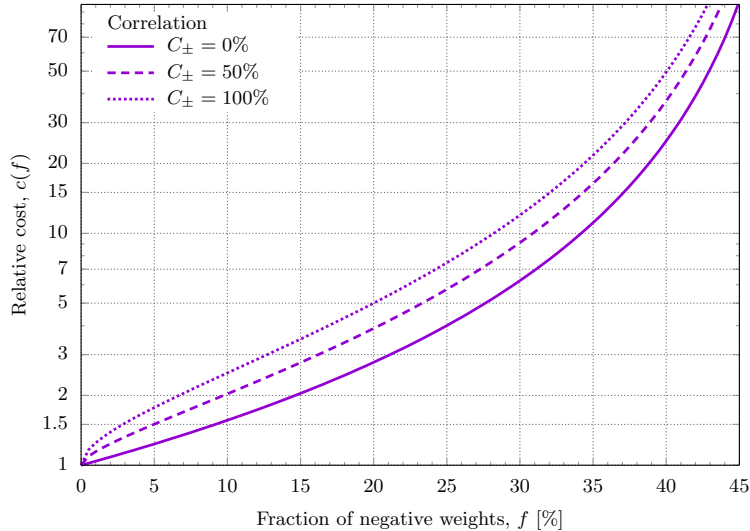
Figure 1: Relative cost as a function of the fraction of negative weights, eq. (1.6), for three different values of the correlation parameter $C_\pm$.

one would need to use eqs. (1.3) and (1.4) locally in the phase space, thus defining a local relative cost, and subsequently construct the global relative cost as the weighted (by number of events) average of the local ones. In practice, eq. (1.6), with f the overall fraction of negative-weight events, does characterise well enough the behaviour of simulations with events of either sign, and we shall often use it in the following.

The problem with $c(f) > 1$ for any $f > 0$ is not statistics *per se*, but the fact that it generally implies additional *financial* costs: longer running times, hence larger power consumption (events with negative weights contribute to climate change!), and bigger storage space, to name just the most important ones. Denoting by p (p') the overall price tag for the generation, full simulation, analysis, and storage of an individual event resulting from a positive-definite (non-positive-definite) simulation, the additional costs alluded to before are:

$$Np' - Mp = M\left[c(f)p' - p\right]. \tag{1.7}$$

With all other things being equal (and chiefly among them, the control of the theoretical systematics), it is therefore advantageous to make f as small as possible, so as to minimise the additional costs[2] of eq. (1.7). This is the goal of the present work, in the context of the MC@NLO matching formalism [1].

Before proceeding, we remind the reader that currently the vast majority of theoretical studies, and essentially all of the NLO+parton shower simulations performed by experimental collaborations, are based on either the MC@NLO or the POWHEG [2] methods,

---

[2]Note that $p' - p$ can have either sign, although when NLO and LO calculations are taken as examples of non-positive- and positive-definite simulations, respectively, most likely $p' > p$. In any case, in the context of a complete experimental analysis the contribution to the cost due to the generation phase alone is minor, and thus $p' \simeq p$.

and on their closely-related variants [3, 4]; alternative approaches (see e.g. refs. [5–10]) are much less well-developed, and rarely used in practice. At variance with MC@NLO, the POWHEG formalism is characterised by a very small fraction of negative weights[3]. Such an appealing feature relies, among other things, on the exponentiation of real-emission matrix elements. This induces a few characteristic features, in particular because neither these matrix elements nor their associated phase spaces exponentiate away from the soft and collinear regions. If we denote by α_S^b the Born-level perturbative order relevant to a given computation, at $\mathcal{O}(\alpha_S^{b+2})$ (and beyond) the POWHEG cross section will contain potentially large terms whose origin is due to matrix elements of $\mathcal{O}(\alpha_S^{b+1})$, in addition to those naturally resulting from Monte-Carlo (MC henceforth) parton showers. This is avoided by construction in the MC@NLO matching, where all terms of $\mathcal{O}(\alpha_S^{b+2})$ and beyond are solely of MC origin[4].

The central idea of the present work is that of showing that, by a suitable modification of the $\mathcal{O}(\alpha_S^k)$ terms ($k \geq 2$) in an MC@NLO cross section by means of contributions of sole MC origin, one arrives at a matching prescription, that we call MC@NLO-Δ, which preserves the good features of MC@NLO while significantly reducing the fraction of negative weights. In particular, the formal $\mathcal{O}(\alpha_S^{b+1})$ expansion of the MC@NLO-Δ cross section coincides with that of MC@NLO and hence, by construction, with the NLO result. In the infrared regions, the logarithmic behaviour of MC@NLO-Δ is the same as that of the MC one matches to, as is the case for MC@NLO. Furthermore, the MC@NLO-Δ formulation leads in a natural manner to a richer structure of the hard events to be showered by the MC. Namely, a (generally different) shower scale is associated with each oriented colour line, and is passed to the MC; this is in contrast to what is done currently, where a single scale characterises the production process, and is regarded as a global property of the event. Finally, we note that in NLO-matched MCs the fraction of negative-weight events with Born kinematics can be reduced by means of a procedure that has been called *folding* in ref. [12], and implemented in practice in POWHEG in ref. [13][5]. Here we shall show that, on top of the intrinsic reduction of the number of negative-weight events in the MC@NLO-Δ matching, the folding technique is rather effective there, whereby such a reduction is more pronounced than in the case of the standard MC@NLO matching.

This paper is organised as follows. In sect. 2 we review the basic properties of the MC@NLO matching that lead to negative-weight events, and classify the latter. In sect. 3 we introduce the new matching prescription, MC@NLO-Δ, that constitutes the core of this work. This is based on a scalar quantity, Δ, whose construction we describe in detail in sect. 3.1; the hard-event multi-scale structure it induces is introduced in sect. 3.2. Some

---

[3]This fraction is not necessarily equal to zero, and depending on the running conditions is not necessarily small – see e.g. ref. [11].

[4]A more precise version of these statements is the following: if no showers are performed, terms of order higher than α_S^{b+1} are (are not) equal to zero in an MC@NLO (a POWHEG) cross section. Note, also, that we are not implying that higher-order terms of MC origin are necessarily "small" or "correct" – however, the behaviour of the MC is a given for matching techniques, and addressing it is beyond the scope of the latter procedures.

[5]In MC@NLO, this possibility had been envisaged in the original paper [1] (see the bottom of sect. 4.5 there), but not implemented in computer codes so far.

peculiar features of the practical implementation of MC@NLO-Δ are reported in sect. 4. In sect. 5 we present sample hadroproduction MC@NLO-Δ results, which we systematically compare with their MC@NLO counterparts. We draw our conclusions in sect. 6. Finally, some technical information on the PYTHIA8 Sudakov form factors are collected in appendix A.

2 Anatomy of negative weights in MC@NLO

We start by pointing out that in MC@NLO the exact amount of negative-weight events and their distribution in the phase space depend on several elements. Among these, the most important are the following: the parton shower MC one matches to, and in particular its shower variables; the technique, which is typically a subtraction procedure, used to compute the underlying NLO cross section; and the choice of the phase-space parametrisation employed in the latter computation. Therefore, in a bottom-up approach to the reduction of negative weights, one would *construct* the shower and the short-distance computations with the specific goal of minimising f. This is a potentially very interesting strategy, which however appears to be quite complicated; we shall not pursue it here. Rather, we shall follow a top-down approach, where both the shower and the NLO calculations are considered as given, and it is the matching between them which is responsible for the minimisation of f. This can be done thanks to the fact that, in spite of their specific features mentioned above, negative weights possess universal characteristics which one can exploit to reduce their number. In order to discuss such universal characteristics, we now sketch out the basic MC@NLO formulae, simplifying them as much as possible, lest the details obscure the basic ideas. If the reader wants to be definite, explicit expressions based on FKS subtraction [14, 15] can be found e.g. in refs. [16–18] for the MADGRAPH5_AMC@NLO implementation (MG5_AMC henceforth).

The key simplification from a notational viewpoint stems from one of the basic features of the FKS subtraction and of the MC@NLO implementations based on it. Namely, for any given real-emission process the phase space is partitioned in an effective manner by means of the $\mathcal{S}$ functions, so that one ultimately deals with a linear combination of short-distance cross sections which have, at most, one soft and one collinear singularity. Such a partition singles out two partons, called the FKS parton and its sister, with which the soft and collinear singularities are associated. We shall thus work by using the rule:

R.1: The following formulae assume that the real-emission process, the FKS parton (labelled by i), and the sister of the latter (labelled by j) are given and fixed. In order to obtain the physical cross sections, one must sum over these quantities.

Bearing the above condition in mind, the MC@NLO generating functional is written as follows[6]:

$$\mathcal{F}_{\text{MC@NLO}} = \mathcal{F}_{\text{MC}}\Big(\mathcal{K}^{(\mathbb{H})}\Big)\, d\sigma^{(\mathbb{H})} + \mathcal{F}_{\text{MC}}\Big(\mathcal{K}^{(\mathbb{S})}\Big)\, d\sigma^{(\mathbb{S})}\,, \tag{2.1}$$

---

[6]As an example of the simplifications induced by rule *R.1*, the reader is encouraged to compare eq. (2.1) with eq. (2.121) of ref. [17].

where $\mathcal{F}_{\mathrm{MC}}$ is the generating functional of the MC one matches to. By $\mathcal{K}^{(\mathbb{H})}$ and $\mathcal{K}^{(\mathbb{S})}$ we have denoted $\mathbb{H}$- and $\mathbb{S}$-event kinematic configurations, respectively. For example, if Born-level processes for the cross section of interest feature n final-state particles, $\mathcal{K}^{(\mathbb{H})}$ and $\mathcal{K}^{(\mathbb{S})}$ correspond to $2 \to n+1$ and $2 \to n$ configurations, respectively[7]. The short-distance cross sections on the r.h.s. of eq. (2.1) are:

$$d\sigma^{(\mathbb{H})} = d\sigma^{(\mathrm{NLO},E)} - d\sigma^{(\mathrm{MC})}, \qquad (2.2)$$

$$d\sigma^{(\mathbb{S})} = d\sigma^{(\mathrm{MC})} + \sum_{\alpha=S,C,SC} d\sigma^{(\mathrm{NLO},\alpha)}. \qquad (2.3)$$

Here, we have denoted by $d\sigma^{(\mathrm{MC})}$ the MC counterterms; the other contributions are identical to those that enter an NLO *fixed-order* cross section:

$$\frac{d\sigma^{(\mathrm{NLO})}}{d\mathcal{K}} = \delta\left(\mathcal{K} - \mathcal{K}^{(\mathbb{H})}\right) d\sigma^{(\mathrm{NLO},E)} + \delta\left(\mathcal{K} - \mathcal{K}^{(\mathbb{S})}\right) \sum_{\alpha=S,C,SC} d\sigma^{(\mathrm{NLO},\alpha)}. \qquad (2.4)$$

Thus, $d\sigma^{(\mathrm{NLO},E)}$ is the real-emission contribution, while $d\sigma^{(\mathrm{NLO},\alpha)}$, $\alpha = S, C, SC$ collect all of the other terms (the Born, and contributions of virtual, soft, collinear, and soft-collinear origin; in a non-FKS language, the latter are therefore the integrated and unintegrated fixed-order counterterms). We point out that the cross sections on the r.h.s. of eqs. (2.2) and (2.3) have support in an $(n+1)$-body phase space. We write the latter as follows[8]:

$$d\phi_{n+1} = \Phi^{(n+1)}\left(\chi_{n+1}\right) d\chi_{n+1}, \qquad (2.5)$$

where χ_{n+1} denotes the set of the chosen $3n-1$ integration variables, whose nature need not be specified here, except for the fact that its has the following general form:

$$\chi_{n+1} = \chi_n \bigcup \chi_r, \qquad (2.6)$$

$$\chi_r = \left\{\xi, y, \varphi\right\}. \qquad (2.7)$$

By χ_n we have denoted $3n-4$ integration variables that define n-body (i.e. Born-level) configurations, and by χ_r the variables that parametrise the extra radiation that occurs at the real-emission level. In an FKS framework (where one works in the c.m. frame of the incoming partons), ξ is the rescaled FKS-parton energy, and y the cosine of the angle between the FKS parton and its sister; φ is an azimuthal angle. Thus, $\xi \to 0$ and $y \to 1$ correspond to the soft and collinear limits, respectively. One can always construct the phase spaces so that (see e.g. ref. [16]):

$$\mathcal{K}^{(\mathbb{S})} = \mathcal{K}^{(\mathbb{S})}(\chi_n), \qquad \mathcal{K}^{(\mathbb{H})} = \mathcal{K}^{(\mathbb{H})}(\chi_{n+1}) \equiv \mathcal{K}^{(\mathbb{H})}(\chi_n, \xi, y, \varphi), \qquad (2.8)$$

and

$$\mathcal{K}^{(\mathbb{S})}(\chi_n) = \mathcal{K}^{(\mathbb{H})}(\chi_n, 0, y, \varphi) = \mathcal{K}^{(\mathbb{H})}(\chi_n, \xi, 1, \varphi). \qquad (2.9)$$

---

[7]In order to simplify the notation, we assume here that all of the particles relevant to our processes are strongly interacting. It is easy to include *a posteriori* extra particles which are not strongly interacting.

[8]Throughout this paper, we understand the integration over Bjorken x's and the definition of the integration variables associated with them.

Equation (2.9) gives an unambiguous meaning to the connection between a real-emission configuration and its underlying Born-level configuration. Furthermore:

$$d\phi_{n+1} = J(\chi_n, \chi_r)\, \Phi^{(n)}(\chi_n)\, d\chi_n d\chi_r\,. \tag{2.10}$$

Finally, we define the pull:

$$P(\mathcal{K}^{(\mathbb{H})}) \equiv P(\chi_n, \xi, y, \varphi)\,, \tag{2.11}$$

as a variable that measures the distance (in phase space) between a real-emission configuration and its underlying Born-level configuration. Therefore, the pull must be such that:

$$\lim_{\xi \to 0} P(\mathcal{K}^{(\mathbb{H})}) = \lim_{y \to 1} P(\mathcal{K}^{(\mathbb{H})}) = 0\,. \tag{2.12}$$

For example, in Drell-Yan production P can be identified with the transverse momentum of the lepton pair. We note that, for any given process, there is ample freedom to define the pull. However, for the sake of the present discussion its precise definition is irrelevant; what matters is that, by assuming that P has canonical dimensions equal to one (which is not restrictive), and by denoting by M_H the typical hard scale of the process, owing to eq. (2.12) the regions:

$$P(\mathcal{K}^{(\mathbb{H})}) \ll M_H\,, \tag{2.13}$$

$$P(\mathcal{K}^{(\mathbb{H})}) \sim M_H \quad \bigcup \quad P(\mathcal{K}^{(\mathbb{H})}) > M_H\,, \tag{2.14}$$

correspond to $\mathcal{K}^{(\mathbb{H})}$ being a soft- and/or collinear-emission configuration, and an intermediate- or hard-emission configuration, respectively.

We classify negative-weight events in MC@NLO as follows:

N.1 $\mathbb{H}$ events with $P(\mathcal{K}^{(\mathbb{H})}) \ll M_H$;

N.2 $\mathbb{H}$ events with $P(\mathcal{K}^{(\mathbb{H})}) \sim M_H$;

N.3 $\mathbb{S}$ events.

Events of both classes **N.1** and **N.2** are due to the fact that the MC counterterms might overestimate the real-emission cross section, and thus the linear combination in eq. (2.2) is negative. **N.1** events will be cancelled after showering (i.e. at the level of physical cross sections) by $\mathbb{S}$ events; being in an MC-dominated region and thanks to the fact that the number of $\mathbb{S}$ events is generally much larger than that of $\mathbb{H}$ events, such a cancellation occurs with high efficiency[9]. By far and large, this also implies that they affect very mildly the shape of kinematical distributions[10], their main impact being on the absolute normalisation (we remind the reader that the MC@NLO and fixed-order NLO total cross sections, before acceptance cuts, are identical).

A similar cancellation mechanism is at play in the case of **N.2** events. However, since the region $P(\mathcal{K}^{(\mathbb{H})}) \sim M_H$ is not dominated by MC effects, the cancellation efficiency is

---

[9]The presence of events of class **N.3** just lowers this efficiency, but does not hamper the cancellation.

[10]As for all $\mathbb{H}$ events in MC-dominated regions.

lower than that relevant to **N.1** events. However, the number of **N.2** events is smaller than that of **N.1** events: the larger the pull, the smaller the probability that the MC will be able to generate the corresponding kinematic configuration. While this is always qualitatively true, quantitatively it depends on the running conditions of the MC, and in particular on the choice of the shower starting scale. We point out that the rationale behind such a choice is different if one does not or does match the MC to matrix elements. In the former case, larger scales are preferred, in order to fill the phase space more effectively, at the cost of stretching the simplifying approximations upon which the MC is based. This is not necessary in the latter case, since hard emissions are already provided by the matrix elements; thus, when performing a matching smaller shower starting scales are a more appealing option. Note that while this conclusion is purely due to physics arguments, one of its by-products is that it helps reduce the number of events of class **N.2**.

If the existence of **N.1** and **N.2** events stems from the basic physics of the MC@NLO cross section, that of **N.3** events is predominantly due to the technical procedure that is (normally) used to generate such events. Note that $\mathbb{S}$ events have an n-body kinematics (see the rightmost term on the r.h.s. of eq. (2.1)), but their associated short distance cross section, eq. (2.3), has support in an $(n + 1)$-body phase space. This implies that, if weighted events are defined, their kinematic configurations depend solely on the variables χ_n, while their associated weights depend on both χ_n and χ_r (see eqs. (2.8)–(2.10)). In the case of unweighted-event generation, the hit-and-miss procedure is carried out in the $(n + 1)$-body phase space. Therefore, the same χ_r-independent configuration $\mathcal{K}^{(\mathbb{S})}$ may be obtained multiple times, depending on the sampling of the χ_r subspace; this is a major source of negative weights in $\mathbb{S}$ events. On the other hand, it should be clear that the $\mathbb{S}$-event generating functional can be rewritten as follows[11]:

$$\mathcal{F}_{\mathrm{MC}}\left(\mathcal{K}^{(\mathbb{S})}\right) \int_{\chi_r} d\sigma^{(\mathbb{S})}. \tag{2.15}$$

In other words, by adopting eq. (2.15) one first integrates the short-distance cross section $d\sigma^{(\mathbb{S})}$ over the χ_r subspace, and then generates (weighted or unweighted) events. This loss of locality in χ_r, which has no consequence on physics since $\mathcal{K}^{(\mathbb{S})}$ is independent of χ_r, is advantageous because it reduces the number of negative weights. In fact, the integrated short-distance cross section in eq. (2.15), at variance with its un-integrated counterpart $d\sigma^{(\mathbb{S})}$, is generally positive. This can be understood by a simple perturbative argument: linear combinations of quantities that have the same kinematic structure (which is strictly the case of the $\mathbb{S}$-event cross section upon integration over χ_r) are dominated, in a well-behaved perturbative expansion, by the lowest-order terms, in this case the Born, which is positive-definite.

In summary, the reduction of negative-weight $\mathbb{S}$ events can be achieved without changing the MC@NLO prescription, simply by means of eq. (2.15). Conversely, the case of $\mathbb{H}$ events is more involved, and requires a modification of the matching procedure, which we shall detail in the next section. In the context of this new prescription, that we call

---

[11]The general idea that underpins eq. (2.15) is that of the folding, that has been already alluded to in sect. 1; see sect. 4 for more details.

MC@NLO-Δ, the reduction of class-**N.3** events can again by achieved thanks to the analogue of eq. (2.15).

3 The MC@NLO-Δ matching prescription

Consider the following generating functional:

$$\mathcal{F}_{\text{MC@NLO-}\Delta} = \mathcal{F}_{\text{MC}}\left(\mathcal{K}^{(\mathbb{H})}\right) d\sigma^{(\Delta,\mathbb{H})} + \mathcal{F}_{\text{MC}}\left(\mathcal{K}^{(\mathbb{S})}\right) d\sigma^{(\Delta,\mathbb{S})}, \tag{3.1}$$

where:

$$d\sigma^{(\Delta,\mathbb{H})} = \left(d\sigma^{(\text{NLO},E)} - d\sigma^{(\text{MC})}\right)\Delta, \tag{3.2}$$

$$d\sigma^{(\Delta,\mathbb{S})} = d\sigma^{(\text{MC})}\Delta + \sum_{\alpha=S,C,SC} d\sigma^{(\text{NLO},\alpha)} + d\sigma^{(\text{NLO},E)}\left(1 - \Delta\right). \tag{3.3}$$

The quantity Δ is understood to have support in the $(n+1)$-body phase-space, and to obey the condition $0 \leq \Delta \leq 1$. Bearing in mind the discussion on both the characteristics of class-**N.1** events and of all $\mathbb{H}$ events in MC-dominated regions (see sect. 2), we expect a reduction of the former ones with negligible effects on the shapes of physical distributions in the latter regions if Δ is such that:

$$\Delta \longrightarrow 0 \qquad \text{soft and collinear limits.} \tag{3.4}$$

Note, in fact, that with eq. (3.4) one manifestly obtains $\mathcal{F}_{\text{MC@NLO-}\Delta} \propto \mathcal{F}_{\text{MC}}\left(\mathcal{K}^{(\mathbb{S})}\right)$ in the soft and collinear regions, analogously to what happens with the standard MC@NLO matching[12]. Conversely, we know that $\mathbb{H}$ events are the sole responsible for giving the NLO-accurate shapes and normalisations in hard-emission regions. Thus, we must have:

$$\Delta \longrightarrow 1 \qquad \text{hard regions.} \tag{3.5}$$

Equations (3.4) and (3.5) are strongly reminiscent of the behaviour of a Sudakov form factor. Indeed, we shall later give the definition of Δ in terms of a combination of MC no-emission probabilities (which are governed by Sudakovs). Before going into further details, we note that Sudakovs have a well-defined perturbative expansion, such that:

$$\Delta = 1 + \mathcal{O}(\alpha_S). \tag{3.6}$$

By using eq. (3.6), a straightforward computation then shows that:

$$d\sigma^{(\Delta,\mathbb{H})} = d\sigma^{(\mathbb{H})} + \mathcal{O}\left(\alpha_S^{b+2}\right), \tag{3.7}$$

$$d\sigma^{(\Delta,\mathbb{S})} = d\sigma^{(\mathbb{S})} + \mathcal{O}\left(\alpha_S^{b+2}\right), \tag{3.8}$$

---

[12]Interestingly, and contrary to the case of MC@NLO, this property would hold even if $d\sigma^{(\text{MC})}$ and $d\sigma^{(\text{NLO},E)}$ did *not* have the same behaviours in such regions, which is appealing in view of the patterns of soft emissions by MCs, and their implications for NLO matchings [1]. We shall not elaborate on this point any further here, and only briefly comment on it at the end of sect. 3.1.

which imply that the generating functionals of eqs. (2.1) and (3.1) have the same expression at the NLO (while in general they differ at the NNLO and beyond). Furthermore:

$$\sigma_{\text{NLO}} \equiv \int_{\chi_{n+1}} \left(d\sigma^{(\mathbb{S})} + d\sigma^{(\mathbb{H})} \right) = \int_{\chi_{n+1}} \left(d\sigma^{(\Delta,\mathbb{S})} + d\sigma^{(\Delta,\mathbb{H})} \right) , \qquad (3.9)$$

where σ_{NLO} is the total NLO cross section, prior to any acceptance cuts.

We point out that eq. (3.6), which obviously is *not* a property that is uniquely associated with a Sudakov, is a sufficient condition for eqs. (3.7)–(3.9) to be fulfilled. However, eq. (3.6) is not sufficient for $\mathcal{F}_{\text{MC@NLO}}$ and $\mathcal{F}_{\text{MC@NLO-}\Delta}$ to produce comparable physical results[13]; for this to happen other conditions, in particular that of eq. (3.5) and the requirement that $0 \le \Delta \le 1$, are needed too. Taken together, then, all of these conditions constrain the functional form of Δ to be, if not that of a Sudakov, at least Sudakov-like. Indeed, in order to sketch out the basic physics ideas that underpin the MC@NLO-Δ prescription, we need only assume that the dependence of such form of Δ upon the kinematical $(n+1)$-body degrees of freedom can be parametrised in terms of two scales[14], as follows:

$$\Delta\big(\mathcal{K}^{(\mathbb{H})}\big) = \Delta\big(t(\mathcal{K}^{(\mathbb{H})}), \mu^2(\mathcal{K}^{(\mathbb{S})})\big). \qquad (3.10)$$

Here, $\mathcal{K}^{(\mathbb{S})}$ denotes the n-body configuration underlying the given $\mathcal{K}^{(\mathbb{H})}$ (see eq. (2.9)). We call $\mu^2(\mathcal{K}^{(\mathbb{S})})$ and $t(\mathcal{K}^{(\mathbb{H})})$ the *starting* and *stopping* scales (squared), respectively. We require them to have the following properties:

$$\mu(\mathcal{K}^{(\mathbb{S})}) \sim M_H \,, \qquad (3.11)$$

$$\sqrt{t(\mathcal{K}^{(\mathbb{H})})} \ll M_H \qquad \mathcal{K}^{(\mathbb{H})} \text{ soft/collinear} \,, \qquad (3.12)$$

$$\sqrt{t(\mathcal{K}^{(\mathbb{H})})} \sim M_H \qquad \mathcal{K}^{(\mathbb{H})} \text{ intermediate/hard} \,. \qquad (3.13)$$

Then, owing to the properties of eq. (3.10), eqs. (3.11) and (3.12) imply eq. (3.4), while eqs. (3.11) and (3.13) imply eq. (3.5). Furthermore, by comparing eqs. (3.12) and (3.13) with eqs. (2.13) and (2.14) one sees that the values assumed by the stopping scale and the pull are correlated. From the classification of negative-weight events given in sect. 2, it then follows that eq. (3.1) is expected to reduce the number of both **N.1** and **N.2** events.

We can now return to a point raised in sect. 1, namely that the $\mathcal{O}(\alpha_S^{b+2})$ and beyond terms in an MC@NLO cross section are of purely MC origin, i.e. they stem from the soft and collinear regions in the multi-parton radiation phase space. This is guaranteed in MC@NLO-Δ as well if it is the MC one matches to that provides the ingredients[15] for the construction of Δ. This means the definition of both the Sudakovs and the stopping scales; as far as the starting scales are concerned, we point out that these are arbitrary to a large

---

[13]In other words, not only results associated with a formal perturbative expansion, but also those at the observable level that include shower effects.

[14]In the fully realistic case we shall soon discuss, these two scales will be replaced by two *sets* of scales.

[15]These considerations remain valid if one replaces the "MC" with suitable resummed computations. Conversely, Δ must not be defined through quantities that depend on hard emissions (such as the exact matrix elements relevant to fixed-order perturbative calculations).

extent, and under the control of the user. In the current context, we shall therefore make sure that our choices are consistent with the condition of eq. (3.11).

A key physics point in the actual definition of the Δ factor, which we shall soon give, is the relationship between eqs. (3.2) and (3.4). This tells one that the faster Δ approaches zero, the smaller the number of $\mathbb{H}$ events (and, thus, those of the **N.1** class). In other words, bearing in mind that we aim to construct Δ with MC no-emission probabilities, we shall have a stronger suppression of the short-distance contributions to $\mathbb{H}$ events in those regions where the MC is expected to have a higher probability to emit. One may suspect that such a suppression, due to the overall factor Δ in eq. (3.2), while leading to a smaller number of **N.1** events will also imply a larger number of **N.3** events, owing to the first term on the r.h.s. of eq. (3.3), which is positive, being suppressed as well. However, this is the case only very marginally (if at all), because the suppression of that first term is compensated by the presence of the last term on the r.h.s. of eq. (3.3). Such a compensation is not exact, and the sum of these two terms tends to be still smaller than the MC-counterterm contribution without the Δ factor. Crucially, however, the difference between these two expressions is much smaller than the Born contribution, which dominates the $\mathbb{S}$-event cross section. Therefore, ultimately the fractions of **N.3** events are remarkably similar in MC@NLO and MC@NLO-Δ; if there is any difference between the two, the latter tends to be slightly larger than the former. As was already anticipated, both can be reduced by means of eq. (2.15), i.e. by folding, and this explains why the folding is typically more effective in MC@NLO-Δ than in MC@NLO. Namely, MC@NLO-Δ has a much smaller fraction of negative-weight $\mathbb{H}$ events w.r.t. MC@NLO, with only a slight increase of negative-weight $\mathbb{S}$ events; while the former are unaffected by folding, the latter are (in large part) eliminated by it.

3.1 Construction of the Δ factor

We have established the general guidelines for the definition of the Δ factor. In order to proceed, we introduce the MC Sudakov, which we write as follows:

$$\Delta_a \left(q^2; \left\{Q_\beta^2, \ell_\beta\right\}_{\beta=1}^{N_\ell(a)}, \mathbb{M}\right) = \prod_{\beta=1}^{N_\ell(a)} \Delta_a^{(*)} \left(q^2; Q_\beta^2, \ell_\beta, \mathbb{M}\right). \tag{3.14}$$

Here, a denotes the identity of the particle that potentially undergoes a branching (i.e. a gluon or a quark of given flavour). The variable q^2, conventionally assumed to have canonical dimensions equal to two, is associated with the quantity in which the shower we match to is ordered (e.g. a transverse momentum squared for p_T-ordered showers, and an angle weighted with an energy squared for angular-ordered showers). In the following, we shall often refer to this quantity as the "shower variable". We take into account the fact that a gluon (a quark or antiquark) has two (one) colour partners by setting:

$$N_\ell(a) = 2 \qquad a = \text{gluon}, \tag{3.15}$$

$$N_\ell(a) = 1 \qquad a = \text{(anti)quark}. \tag{3.16}$$

These partner(s) are the end (in the case of colour) and/or the beginning (in the case of anticolour) of the colour line(s) that begin or end at particle a; in eq. (3.14), we have

denoted such lines by ℓ_β. The upper value accessible to the shower variable for emissions stemming from the ℓ_β colour connection has been denoted by Q_β^2. Finally, by $\mathbb{M}$ we have denoted the set of any extra variables upon which the definition of the Sudakov may depend – we shall give an explicit example in appendix A.

The r.h.s. of eq. (3.14) is defined as follows:

$$\Delta_a^{(*)}\left(q^2; Q^2, \ell, \mathbb{M}\right) = \exp\left[-\frac{1}{N_\ell(a)}\,\delta_a\left(q^2; Q^2, \ell, \mathbb{M}\right)\right], \tag{3.17}$$

where:

$$\delta_a\left(q^2; Q^2, \ell, \mathbb{M}\right) = \frac{\Theta\left(Q^2 - q^2\right)}{N_a}\sum_b\int_{q^2}^{Q^2}\frac{dt}{t}\frac{\alpha_S(t)}{2\pi}\int_{\epsilon(t,\mathbb{M},\ell)}^{1-\epsilon(t,\mathbb{M},\ell)}dz\,\hat{P}_{ba}(z)\,. \tag{3.18}$$

Here, $\hat{P}_{ba}(z)$ is the unsubtracted Altarelli-Parisi lowest-order kernel relevant to the branching of a in which parton b carries a fraction equal to z of the longitudinal momentum of a. The r.h.s. of eq. (3.18) is summed over b, and the values assumed by b must take into account the relationships among the Altarelli-Parisi kernels. We find it convenient to work with symmetric quantities; hence, when $a = q_f$, with q_f a(n) (anti)quark of flavour f, then $b \in \{g, q_f\}$, whereas if $a = g$, then $b \in \{g, u, \bar{u}, \ldots\}$. This implies that we need to set:

$$N_a = 2\,, \tag{3.19}$$

irrespective of the identity of particle a. In eq. (3.18) we have denoted by ϵ the quantity the controls the phase-space boundaries of the z integration; as the notation suggests, such boundaries depend on the shower variable, on the variables in $\mathbb{M}$, and on the colour line ℓ; more details specific to the case of PYTHIA8 [19] are given in appendix A. Finally we remark that, depending on the type of MC showers one considers, the Θ function on the r.h.s. of eq. (3.18) might have a more complicated functional dependence than that indicated (for example, by enforcing the constraints relevant to angular-ordered showers), and because of this it might need to be moved under the integration signs. Likewise, the argument of α_S might be assigned by means of a different prescription (see e.g. refs. [20, 21]). However, for the sake of the current argument there is no loss of generality in using eq. (3.18).

We briefly comment on the physical meaning of eq. (3.14). When a and b are gluons, that equation and eqs. (3.17)–(3.18) imply that each of the two $\Delta_a^{(*)}$ factors is responsible for an evolution of effective strength equal to $C_A/2$ – in other words, each colour line contributes "half" of the radiation pattern[16]. As a consequence of that, the evolution from $\max(Q_1^2, Q_2^2)$ to $\min(Q_1^2, Q_2^2)$ is driven by branchings of effective strength equal to $C_A/2$ (and not C_A). Conversely, when a is a quark Δ_a and $\Delta_a^{(*)}$ coincide, consistently with the fact that in this case a single colour connection is relevant.

In order to employ the Sudakovs introduced above for the definition of Δ, we must first define the sets of the starting and stopping scales (see eq. (3.10) and footnote 14).

---

[16]Note that $C_A/2$ and C_F coincide at the leading-colour level – it is the colour line, rather than the particle identity, that controls the radiation.

The choices of the former scales are motivated by the form of the MC counterterms:

$$d\sigma^{(\mathrm{MC})} = D\Big(q(\mathcal{K}^{(\mathbb{H})}), \mu_1, \mu_2\Big) \sum_c \sum_{\ell \in c} d\sigma_{c\ell}^{(\mathrm{MC})}, \qquad (3.20)$$

where by $q(\mathcal{K}^{(\mathbb{H})})$ we have denoted the square root of the shower variable relevant to the kinematic configuration $\mathcal{K}^{(\mathbb{H})}$, for the branching in which the FKS parton i and its sister j emerge from their mother[17]. The reader can find more details on eq. (3.20) in ref. [17] (see in particular eqs. (2.111)–(2.116) there). Here, it suffices to say that by $d\sigma_{c\ell}^{(\mathrm{MC})}$ we have denoted the contribution to the MC counterterms due to a given colour flow c and colour line ℓ; the former is identified with the set of its colour lines $c = \{\ell_1, \ldots \ell_m\}$. In turn, each of these lines is conveniently represented by an ordered pair of labels, with the first (second) being associated with the beginning (end) of the line, and the colour (as opposed to the anticolour) flowing from the beginning of the line to its end[18]. We point out that such colour flow and colour lines are relevant to $2 \to n$ configurations. The function D in eq. (3.20) is largely arbitrary, and is parametrised as follows:

$$D(\mu; \mu_1, \mu_2) = \begin{cases} 1 & \mu \leq \mu_1, \\ \text{monotonic} & \mu_1 < \mu \leq \mu_2, \\ 0 & \mu > \mu_2, \end{cases} \qquad (3.21)$$

with μ_α, $\alpha = 1, 2$, two parameters of the order of the hardness of the process[19]:

$$\mu_\alpha = f_\alpha R, \qquad f_\alpha = \mathcal{O}(1), \qquad R \sim M_H. \qquad (3.22)$$

The presence of the function D in eq. (3.20) implies that the shower starting scales associated with unweighted events are generated by means of the following formula:

$$\mu = D^{-1}(r; \mu_1, \mu_2), \qquad (3.23)$$

with $r \in [0, 1]$ a flat random number.

The procedure outlined above motivates the definition of the starting scales for the MC@NLO-Δ prescription. Firstly, we adopt the following rule:

R.2: Among the colour flows that contribute to eq. (3.20), we select one of them at random by using $\sum_\ell d\sigma_{c\ell}^{(\mathrm{MC})}$ as relative probabilities, and denote it by d. Henceforth, by ℓ_β ($1 \leq \beta \leq m$) we understand the colour lines that belong to such a flow.

Secondly, we generalise the reference scale R introduced in eq. (3.22) by turning it into a set of scales, as follows:

$$R_{\bar{k}\bar{p}}^2 = (-1)^{\varsigma(\bar{k}) + \varsigma(\bar{p})} \Big(\varsigma(\bar{k}) \, \bar{K}_{\bar{k}} + \varsigma(\bar{p}) \, \bar{K}_{\bar{p}}\Big)^2 \qquad \Longleftrightarrow \qquad \text{if } \exists \ell_\beta \text{ s.t. } \bar{k} \in \ell_\beta \text{ and } \bar{p} \in \ell_\beta, \qquad (3.24)$$

---

[17]In general, $q(\mathcal{K}^{(\mathbb{H})})$ might also depend on the end(s) of the colour line(s) to which the mother is connected. This dependence is trivial for all of the MCs used in MG5_AMC at present (among which is PYTHIA8 with the *global recoil* option).

[18]Note that, when working with a given process and kinematical configuration as we are doing here, one can unambiguously choose a labeling convention.

[19]For example, in the current version of MG5_AMC, by default $f_1 = 0.1$, $f_2 = 1$, and $R = H_T/2$.

where $1 \leq \bar{k}, \bar{p} \leq n$ are particle labels, and $\varsigma(\bar{k}) = 1$ if the particle with label $\bar{k}$ is in the final state, while if it is in the initial state the two choices $\varsigma(\bar{k}) = \pm 1$ are both sensible; in the simulations of this paper we shall use $\varsigma(\bar{k}) = 1$, keeping the option $\varsigma(\bar{k}) = -1$ for future studies. By convention, we denote Born-level quantities by barred symbols; thus, $\mathcal{K}^{(\mathbb{S})} = \{\bar{K}_1, \ldots \bar{K}_n\}$, where $\bar{K}_{\bar{\alpha}}$ is the momentum of the particle labelled by $\bar{\alpha}$. The condition in eq. (3.24) states that the particles with labels $\bar{k}$ and $\bar{p}$ are colour connected in the colour flow d. Since each particle has either one colour partner (if it is a quark or an antiquark) or two of them (if it is a gluon), it follows that depending on the process there are at least n and at most $2n$ non-trivial reference scales defined by eq. (3.24). We can now introduce the analogues of the μ_α parameters of eq. (3.22), namely:

$$\mu_{\alpha, \bar{k}\bar{p}} = f_\alpha \, R_{\bar{k}\bar{p}}, \qquad \alpha = 1, 2, \tag{3.25}$$

and finally the sought starting scales by analogy with eq. (3.23):

$$\mu_{\bar{k}\bar{p}} = D^{-1}(r_{\bar{k}\bar{p}}; \mu_{1,\bar{k}\bar{p}}, \mu_{2,\bar{k}\bar{p}}), \tag{3.26}$$

where $r_{\bar{k}\bar{p}}$ are flat random numbers. As far as the stopping scales are concerned, we simply define them as the square roots of the MC shower variables associated with the given kinematical configuration. More precisely, at given $\bar{k}$ and $\bar{p}$, we denote by:

$$t_{\bar{k}\bar{p}} \qquad \Longleftrightarrow \qquad \text{if } \exists \ell_\beta \text{ s.t. } \bar{k} \in \ell_\beta \text{ and } \bar{p} \in \ell_\beta \tag{3.27}$$

the shower variable relevant to the emission of the FKS parton from the branching of parton $\bar{k}$, when such a parton is colour-connected to parton $\bar{p}$. It is important to note that while for a specific $\bar{k}$ the branching parton coincides with the mother of the FKS parton and its sister, in general this is not the case: $\bar{k}$ is the mother of the FKS parton and of another parton, not necessarily equal to the sister of the FKS parton. The definition in eq. (3.27) implies that there are as many non-trivial stopping scales as there are starting scales. Note that in general:

$$t_{\bar{k}\bar{p}} \neq t_{\bar{p}\bar{k}}. \tag{3.28}$$

It is useful to have a closer look at eq. (3.28) by means of an example. Consider a Drell-Yan process, which at the Born level features a quark and an antiquark (labelled by 1 and 2, respectively) in the initial state, with a single colour line $(1, 2)$. Suppose that a gluon is emitted, and employ a p_T-ordered MC such as PYTHIA8. Thus, t_{12} is the gluon transverse momentum squared *relative* to the quark, while t_{21} is the same quantity relative to the antiquark. If for example the kinematical configuration is such that the gluon is collinear to the quark, and therefore anti-collinear to the antiquark, then $t_{12} \ll t_{21}$.

In order to arrive at the definition of no-emission probabilities which will enter the sought Δ factor, we finally need to introduce some auxiliary quantities. We denote by:

$$\mathcal{C}_\gamma(\bar{k}), \qquad 1 \leq \gamma \leq N_\ell(\mathcal{I}_{\bar{k}}), \tag{3.29}$$

the labels of the particles[20] which are colour connected with the particle labelled $\bar{k}$ and, following the conventions of ref. [17], by $\mathcal{I}_{\bar{k}}$ the identity of such a particle; $N_\ell(\mathcal{I}_{\bar{k}})$ is set

---

[20]In the case of $\bar{k}$ being a gluon, the choice of which particle is the first ($\gamma = 1$) or the second ($\gamma = 2$) colour partner is arbitrary. None of the formulae presented later depends on such a choice.

according to eqs. (3.15) and (3.16). One must bear in mind that we are working at a fixed colour flow (d) here; thus, the colour partners in eq. (3.29) are unambiguously defined. This implies that, for a given γ, there exists a single colour line to which both $\bar{k}$ and $\mathcal{C}_\gamma(\bar{k})$ belong. We denote such a line as follows:

$$\ell\big[\bar{k}, \mathcal{C}_\gamma(\bar{k})\big] \equiv \ell_\beta \,, \qquad \exists!\, \ell_\beta \ \ \text{s.t.} \ \ \ell_\beta = \big(\bar{k}, \mathcal{C}_\gamma(\bar{k})\big) \ \ \text{or} \ \ \ell_\beta = \big(\mathcal{C}_\gamma(\bar{k}), \bar{k}\big) \,. \tag{3.30}$$

If $\bar{k}$ labels a quark or an antiquark, we construct its no-emission probability as follows:

$$\Pi_{\bar{k}} = F_{\bar{k}}\Big(t_{\bar{k}\mathcal{C}_1(\bar{k})}; \mu^2_{\bar{k}\mathcal{C}_1(\bar{k})}\Big) \, \Delta_{\mathcal{I}_{\bar{k}}}\Big(t_{\bar{k}\mathcal{C}_1(\bar{k})}; \mu^2_{\bar{k}\mathcal{C}_1(\bar{k})}, \ell\big[\bar{k}, \mathcal{C}_1(\bar{k})\big], \mathbb{M}\Big) \,, \tag{3.31}$$

where:

$$F_{\bar{k}}(q^2, Q^2) = \begin{cases} f^{(\bar{k})}_{\mathcal{I}_{\bar{k}}}(x, q^2) / f^{(\bar{k})}_{\mathcal{I}_{\bar{k}}}(x, Q^2)\, \Theta(Q^2 - q^2) + \Theta(q^2 - Q^2) & 1 \le \bar{k} \le 2\,, \\ 1 & \bar{k} > 2\,. \end{cases} \tag{3.32}$$

In eq. (3.32) $f^{(\bar{k})}_{\mathcal{I}_{\bar{k}}}$ is the PDF of parton $\mathcal{I}_{\bar{k}}$ inside the hadron coming from the left ($\bar{k} = 1$) or from the right ($\bar{k} = 2$). The momentum fraction x that enters the PDF is the same as that which is used in the evaluation of the MC counterterms (see ref. [17] for more details). Note that, owing to eqs. (3.18) and (3.32), we have:

$$\text{if} \quad t_{\bar{k}\mathcal{C}_1(\bar{k})} \ge \mu^2_{\bar{k}\mathcal{C}_1(\bar{k})} \quad \implies \quad \Pi_{\bar{k}} = 1\,, \tag{3.33}$$

which formalises the notion that there is no parton-shower emission in the hard region, whose lower boundary coincides by definition with the starting scale.

The case where $\bar{k}$ labels a gluon is more complicated. In a non-matched case, the MC uses the Sudakov of eq. (3.14) (or rather, a no-emission probability that differs from the Sudakov in a way which is irrelevant in this discussion) to extract randomly a value q^2 of the shower variable. In doing so, the two colour lines bid to emit. Once q^2 has been generated, the corresponding kinematical configuration is reconstructed; however, in order to do so the MC must choose a definite colour line, since in general the kinematic configurations associated with the same q^2 that emerge from different colour lines are different. In the present case, this logic must be reversed, since we start from a definite kinematical configuration; this implies that the two colour lines might give rise to different values of the shower variable. We stress that this need not be so; but we work in this scenario in order to be as general as possible. We then define the analogue of the no-emission probability of eq. (3.31) in the case of a gluon as follows:

$$\begin{aligned} \Pi_{\bar{k}} =\ & \frac{g_1}{g_1 + g_2} F_{\bar{k}}\Big(t_{\bar{k}\mathcal{C}_1(\bar{k})}; \mu^2_{\bar{k}\mathcal{C}_1(\bar{k})}\Big) \, \Delta_{\mathcal{I}_{\bar{k}}}\Big(t_{\bar{k}\mathcal{C}_1(\bar{k})}; \big\{\mu^2_{\bar{k}\mathcal{C}_\gamma(\bar{k})}, \ell\big[\bar{k}, \mathcal{C}_\gamma(\bar{k})\big]\big\}^2_{\gamma=1}, \mathbb{M}\Big) \\ & + \frac{g_2}{g_1 + g_2} F_{\bar{k}}\Big(t_{\bar{k}\mathcal{C}_2(\bar{k})}; \mu^2_{\bar{k}\mathcal{C}_2(\bar{k})}\Big) \, \Delta_{\mathcal{I}_{\bar{k}}}\Big(t_{\bar{k}\mathcal{C}_2(\bar{k})}; \big\{\mu^2_{\bar{k}\mathcal{C}_\gamma(\bar{k})}, \ell\big[\bar{k}, \mathcal{C}_\gamma(\bar{k})\big]\big\}^2_{\gamma=1}, \mathbb{M}\Big), \end{aligned} \tag{3.34}$$

where:

$$g_\gamma = \delta_{\mathcal{I}_{\bar{k}}}\Big(t_{\bar{k}\mathcal{C}_\gamma(\bar{k})}; \mu^2_{\bar{k}\mathcal{C}_\gamma(\bar{k})}, \ell\big[\bar{k}, \mathcal{C}_\gamma(\bar{k})\big], \mathbb{M}\Big)\,, \tag{3.35}$$

and $\delta_{\mathcal{I}}$ is defined in eq. (3.18). Apart from the g_γ-dependent prefactor, each of the two terms on the r.h.s. of eq. (3.34) is the standard gluon no-emission probability. The difference between them is solely due to the choice of the stopping scale that enters the PDF factor and the Sudakov, $t_{\bar{k}\mathcal{C}_1(\bar{k})}$ versus $t_{\bar{k}\mathcal{C}_2(\bar{k})}$, and to that of the starting scale that enters the PDF, $\mu^2_{\bar{k}\mathcal{C}_1(\bar{k})}$ versus $\mu^2_{\bar{k}\mathcal{C}_2(\bar{k})}$. The g_γ-dependent prefactors have the role of combining these two no-emission probabilities by means of a weighted average, in which the weights are MC-driven, and are such that the term associated with the smallest stopping scale is multiplied by the largest factor. The analogue for eq. (3.34) of eq. (3.33) reads as follows:

$$\text{if} \quad \min\left(t_{\bar{k}\mathcal{C}_1(\bar{k})}, t_{\bar{k}\mathcal{C}_2(\bar{k})}\right) \geq \max\left(\mu^2_{\bar{k}\mathcal{C}_1(\bar{k})}, \mu^2_{\bar{k}\mathcal{C}_2(\bar{k})}\right) \quad \Longrightarrow \quad \Pi_{\bar{k}} = 1. \tag{3.36}$$

We must also mention the fact that it is possible that, for given kinematical configuration, emitter, and colour partner (or partners in the case of a gluon), the FKS parton is in an MC dead zone (i.e. a phase-space region which is not accessible by MC radiation; when $\bar{k}$ labels a gluon, this condition is fulfilled only if the FKS parton is in a dead zone according to *both* of the colour connections of $\bar{k}$). In such a case we set $\Pi_{\bar{k}} = 1$, since that kinematical configuration could not have been generated by the MC.

By using eqs. (3.31) and (3.34) we can finally define the Δ factor we need for the MC@NLO-Δ matching:

$$\Delta = \prod_{\bar{k}=1}^{n} \Pi_{\bar{k}}. \tag{3.37}$$

Although the construction of the no-emission probabilities that enter eq. (3.37) is involved, the physics content of the Δ factor is easy to understand, because it is quite literally that of eq. (3.10). In fact, in a typical situation all but one of the $\Pi_{\bar{k}}$ terms on the r.h.s. of eq. (3.37) are of order one. Furthermore, the single no-emission probability for which $\Pi_{\bar{k}} \ll 1$ is most likely that where $\bar{k}$ identifies the particle that branches into the FKS parton and its sister. This is because by working in the FKS sector defined by a given $\mathcal{S}_{ij}$ function, one suppresses all collinear configurations except that where $K_i \parallel K_j$ [14, 15]; for the former the stopping scales are "large" (i.e. as in eq. (3.13)), while for the latter they are "small" (i.e. as in eq. (3.12)). It is interesting to observe that in MC@NLO there is a non-zero probability that, within the $\mathcal{S}_{ij}$ sector, the FKS parton is closer to a particle different from its sister (we denote by p its label) than to its sister; let us call anticollinear such a configuration. This behaviour is driven by the fact that while the real-emission component of the short-distance cross sections is multiplied by $\mathcal{S}_{ij}$, the MC counterterms are not since, at variance with the matrix elements, in anticollinear configurations they are naturally suppressed. However, they are not equal to zero, which induces the small but non-zero probability, alluded to before, of generating events there. Furthermore, since these are $\mathbb{H}$ events predominantly stemming from the MC counterterms, it is likely that they will have negative weights (see eq. (2.2)). While this is the mechanism at work in the context of MC@NLO, in MC@NLO-Δ these anticollinear configurations are suppressed – by the no-emission probability $\Pi_{\bar{k}}$ with $\bar{k}$ now identifying the particle that branches into the FKS parton and the particle labelled by p. Therefore, thanks to the symmetry of eq. (3.37), the MC@NLO-Δ prescription reduces the number of events of class **N.1** in both collinear

and anticollinear configurations. This is just as well, because the separation into collinear and anticollinear regions stems from adopting the FKS subtraction for writing the short-distance cross sections, whereas the ideas that underpin MC@NLO-Δ are independent of the FKS formalism.

The Δ factor of eq. (3.37) may feature a suppression due to multiple no-emission probabilities in the case of soft configurations. We expect the strength of this suppression to depend significantly on the MC one matches to. For example, soft emissions in angular-ordered MCs have *typically* a pattern whereby the contributions due to different emitters have a smaller overlap than in the case of p_T-ordered showers; this implies that the value of Δ associated with soft configurations will be typically larger (i.e. resulting in a smaller suppression) in angular-ordered showers than in p_T-ordered ones. In any case, the suppression induced by Δ is interesting because of the technical issues related to soft events in MC@NLO (see in particular appendix A.5 of ref. [1] and eq. (2.117) of ref. [17]). By construction, these issues become much less relevant in the context of MC@NLO-Δ, to the extent that in such a formulation one might get rid of the function $\mathcal{G}$ in the definition of the MC counterterms.

We conclude this discussion with a comment on the colour structure of the MC@NLO-Δ cross section. As is the case for all quantities stemming from a random unweighting of competing contributions, the Δ factor constructed with d will also multiply short-distance terms that feature colour flows different from d. This is perfectly fine, and in fact every MC is based on several applications of a similar nature. In the specific example of Δ, however, a possible alternative procedure that would allow one to bypass the unweighting would be that of decomposing *all* matrix elements as is done for Born-level quantities in standard MCs, e.g. by applying the prescription of ref. [22]. We shall not pursue any further this option in the present paper.

3.2 Assignments of scales in hard events

The definition of the Δ factor in the MC@NLO-Δ prescription naturally suggests how to improve the structure of the hard scales which are associated with hard events. At present, in NLO simulations a single hard scale per event is employed; we propose to exploit the starting and stopping scales in order to assign either one (for quarks) or two (for gluons) hard scale(s) *per particle*. In particular, in $\mathbb{S}$ events the scales associated with the particle labelled by $\bar{k}$ are the following:

$$\bar{k} \text{ particle}, \quad \mathbb{S} \text{ events} \quad \longrightarrow \quad \left\{ \mu_{\bar{k}\mathcal{C}_\gamma(\bar{k})} \right\}_{\gamma=1}^{N_\ell(\mathcal{I}_{\bar{k}})}. \tag{3.38}$$

In other words, in the case of $\mathbb{S}$ events each particle is associated with as many hard scales as its colour connections; each of these scales is equal to the starting scale stemming from the corresponding colour line.

In the case of $\mathbb{H}$ events, the obvious analogue of eq. (3.38) is that in which one sets the hard scales equal to the stopping scales. However, in order to do so some care is required: in fact, although stopping scales are constructed in terms of $(n+1)$-body kinematical configurations, the particle indices they depend upon are those of an n-body process (see

eq. (3.27)). We start by observing that, within a given FKS sector (i.e. for a given $\mathcal{S}_{ij}$ function) the particle indices at the $(n+1)$-body and n-body levels can unambiguously be mapped onto each other: we simply use the same symbol, barred (unbarred) in the case of the n-body $((n+1)$-body) process to denote corresponding particles (for example, $\bar{k}$ and k denote the index of a particle before and after, respectively, the branching that turns an n-body configuration into an $(n+1)$-body one). The FKS parton i and its sister j constitute a special case, since they have no analogues at the n-body level; there, their mother appears instead. We denote it by $\overline{i \oplus j}$.

The colour flow of each $\mathbb{H}$ event is constructed from that of the underlying n-body configuration (d), by breaking the colour line (in the case of a quark), or one of the two colour lines (in the case of a gluon; for a $g \to q\bar{q}$ branching, the lines "split"), associated with the branching of $\overline{i \oplus j}$; the other colour lines that belong to d are left untouched (bar the obvious particle relabeling). This operation is fully deterministic, except in the case of a $g \to gg$ branching, for which two different $\mathbb{H}$ colour flows can be constructed: we choose one of them randomly (with equal probabilities).

After the definition of the $\mathbb{H}$-event colour flow, in keeping with eq. (3.38) we assign one or two hard scale(s) per particle, according to the colour connection(s) of such a particle. We do that by a suitable mapping of the n-body indices relevant to the stopping scales onto the $(n+1)$-body ones. Explicitly:

$$
\begin{aligned}
&\text{if } \exists \ell[k,p] &\implies& \quad t_{kp} = t_{\bar{k}\bar{p}} & k \neq i\,, k \neq j\,, p \neq i\,, p \neq j\,, & \quad (3.39)\\
&\text{if } \exists \ell[k,j] &\implies& \quad t_{kj} = t_{\bar{k}\overline{i\oplus j}} & k \neq i\,, k \neq j\,, & \quad (3.40)\\
&\text{if } \exists \ell[k,i] &\implies& \quad t_{ki} = t_{\bar{k}\overline{i\oplus j}} & k \neq i\,, k \neq j\,, & \quad (3.41)\\
&\text{if } \exists \ell[i,k] &\implies& \quad t_{ik} = t_{\overline{i\oplus j}\bar{p}} & \bar{p} \text{ s.t. } i \text{ colour connected to } p\,, & \quad (3.42)\\
&\text{if } \exists \ell[j,k] &\implies& \quad t_{jk} = t_{\overline{i\oplus j}\bar{p}} & \bar{p} \text{ s.t. } i \text{ colour connected to } p\,. & \quad (3.43)
\end{aligned}
$$

In eqs. (3.39)–(3.43) we have employed the same symbol $\ell[a,b]$ introduced in eq. (3.30) to denote the colour line (which, in this case, belongs to the $\mathbb{H}$-event colour flow) that connects particle a and particle b, so that t_{ab} is a meaningful quantity. Owing to the construction of the $\mathbb{H}$-event colour flow described before, it is easy to see that the pairs of indices on the r.h.s.'s of eqs. (3.39)–(3.43) are also associated with two particles which are colour connected at the n-body level; therefore, the corresponding scales belong to the set of stopping scales previously determined. We point out that the conditions in eqs. (3.40) and (3.41), and those in eqs. (3.42) and (3.43), need not be simultaneously satisfied; in particular, this is never the case if k labels a quark. When they are, the corresponding scales are assigned the same value, in turn equal to a stopping scale relevant to the mother of the FKS parton and its sister. This is in keeping with the special roles played by the latter partons, a fact that has been already exploited in the construction of the $\mathbb{H}$-event colour flow.

In summary, in $\mathbb{H}$ events we set the scales associated with the particle labelled by k as follows:

$$
k \text{ particle}\,, \quad \mathbb{H} \text{ events} \quad \longrightarrow \quad \left\{ \sqrt{t_{k\mathcal{C}_\gamma(k)}} \right\}_{\gamma=1}^{N_\ell(\mathcal{I}_k)}\,, \quad (3.44)
$$

where $\mathcal{C}_\gamma(k)$ now denotes the label of the γth colour partner of the particle with label k according to the $\mathbb{H}$ colour flow.

The scale structure defined by eqs. (3.38) and (3.44) can be easily included in Les Houches event (LHE henceforth) files [23, 24], which in turn are fully compatible with modern MCs; more details on this point are given in sect. 4. This is expected to give the MC options for showering scales which are, on an *event-by-event basis*, more sensible than those relevant to single-scale hard events (i.e. those which are presently in use). The reasons that inform this argument are general, and thus valid for both $\mathbb{S}$ and $\mathbb{H}$ events, but are easier to understand if one uses an $\mathbb{H}$ event as a practical example. This is because particles in $\mathbb{S}$ events tend to be hard and well separated from each other, while this is typically not the case for $\mathbb{H}$ events – in other words, while both $\mathbb{S}$ and $\mathbb{H}$ events constitute multi-scale environments, the hierarchy among the scales of the latter events will typically be stronger than that of the former ones. Let us then consider an $\mathbb{H}$-event configuration in which two particles have a small angular separation. By adopting an MC-driven picture, these two particles are naturally seen as emerging from the branching of a mother particle, and thus any subsequent shower emission off them should use a "small" scale (owing to the small angular separation) as a reference. Conversely, shower emissions from the other particles of the event (at least those not colour-connected with the former two) would tend to use a "large" scale (these other particles being well-separated from each other and hard) as a reference. This small-scale large-scale hierarchy may be handled in single-scale LHE files in a variety of ways (by means of an average assignment, by using a biased random assignment, and so forth), but it is unavoidable that, for some individual events and subsequent shower histories, a sub-optimal choice will be made, with a better physical modeling recovered, in principle, only with large statistics. This issue is simply not relevant if LHE files incorporate a multi-scale structure, such as the one we advocate in eqs. (3.38) and (3.44) and implement in MC@NLO-Δ.

4 Implementation

In order to achieve an MC@NLO-Δ matching, any existing MC@NLO program must be given the capability to compute the Δ factor of eq. (3.37). As is discussed in sect. 3, although perturbatively there is a significant freedom in the definition of Δ, the ideal situation is that in which it is the MC one matches to that provides the ingredients that enter Δ. From the conceptual point of view, this is quite analogous to the situation of the MC counterterms, which are constructed *analytically* through a formal expansion in α_S of the MC cross sections. However, although some of the quantities that help define Δ also enter the MC counterterms (specifically, the stopping scales and dead zones associated with $\bar{k}$ being the label of the mother of the FKS parton and its sister), the majority of them would have to be computed from scratch if one had to follow the same strategy that relies on analytical results. We believe this to be an error-prone procedure that also lacks flexibility, and we therefore pursue a different approach (which, in due time, could also be applied the MC counterterms themselves). Namely, we construct Δ *numerically*, by essentially employing the same modules as the MC does when showering.

We point out that this approach entails non-trivial modifications to the structure of the interface between MG5_AMC and the MC. In particular, in the current MC@NLO implementation of MG5_AMC (i.e. one based on the analytical forms of the MC counterterms), the phase-space integration of the short-distance cross sections and the unweighting of the events (thus, the creation of the LHE files) do not require the use of the MC. The MC is run independently of MG5_AMC and after it, using as inputs the LHE files produced by the latter program. Conversely, with the MC@NLO-Δ implementation envisaged previously, MC routines must be called for each phase-space point generated by the integration routine.

In view of this, the MG5_AMC and Pythia8 versions used to obtain the results presented here include the following features:

- The executable relevant to the phase-space integration and unweighting-event phases is constructed by linking the MG5_AMC and Pythia8 programs.

- New modules have been included in Pythia8, that call native Pythia8 modules relevant to showering, and which in turn are called by MG5_AMC for the construction of Δ in the definition of the short-distance cross sections $d\sigma^{(\Delta,\mathbb{H})}$ and $d\sigma^{(\Delta,\mathbb{S})}$. For any given phase-space point, these new modules return the stopping scales and the information on the dead zones.

- Analogous modules are constructed that return the Pythia8 Sudakovs. However, in view of the fact that the Sudakovs depend on kinematical configurations in a way that can be parametrised, such modules are called prior to the phase-space integration, by a program that defines the Sudakovs as look-up tables. It is such tables that are used by MG5_AMC during integration and unweighting, which allows one to save a significant amount of CPU time. More details on this procedure, that we dub *pre-tabulation*, are given in appendix A.

As was already said, this structure paves the way to the numerical construction of the MC counterterms themselves. We have not pursued this option in the course of the present work, where we employ the same MC counterterms as those originally computed for the MC@NLO matching.

For the MC@NLO and MC@NLO-Δ matchings, we have also implemented the $\mathbb{S}$-event generating functional according to eq. (2.15) (with $d\sigma^{(\mathbb{S})} \to d\sigma^{(\Delta,\mathbb{S})}$ there in the case of MC@NLO-Δ). This is done by expressing the integral on the r.h.s. of that equation through Riemann sums:

$$\mathcal{F}_{\mathrm{MC}}\Big(\mathcal{K}^{(\mathbb{S})}\Big) \int_{\chi_r} d\sigma^{(\mathbb{S})} \simeq \mathcal{F}_{\mathrm{MC}}\Big(\mathcal{K}^{(\mathbb{S})}\Big) \sum_{i_\xi=1}^{n_\xi} \sum_{i_y=1}^{n_y} \sum_{i_\varphi=1}^{n_\varphi} \frac{w_{i_\xi i_y i_\varphi}}{n_\xi n_y n_\varphi} d\sigma^{(\mathbb{S})} \Big(\mathcal{K}^{(\mathbb{S})}, \xi_{i_\xi}, y_{i_y}, \varphi_{i_\varphi}\Big). \quad (4.1)$$

This implies that, for each n-body configuration, $n_\xi \times n_y \times n_\varphi$ points are generated that span the χ_r subspace. The capability of choosing such points $((\xi_{i_\xi}, y_{i_y}, \varphi_{i_\varphi}))$ and the corresponding weights $(w_{i_\xi i_y i_\varphi})$ is fortunately already available in the integration routine (MINT [12]) that has been adopted in MG5_AMC. Following MINT, we call folding the

procedure on the r.h.s. of eq. (4.1). The integers n_ξ, n_y, and n_φ are called folding parameters, and are under the user's control. We shall discuss their use in sect. 5; for the time being, note that by choosing all of the folding parameters equal to one we recover the standard $\mathbb{S}$-event generating functional.

We conclude this section with a few words on the shower scales associated with hard events in the context of MC@NLO-Δ. These are defined as is described in sect. 3.2, and included in LHE files by means of the `<scales>` tag. In particular, a typical LHE will read as follows:

```
<event>
  ....
  <scales muf='1.0E+01' mur='1.0E+01' ... scalup_a_b='X' ...>
  </scales>
</event>
```

For each event, there are as many `scalup_a_b` entries as is necessary, with `a` and `b` being equal to either $\bar{k}$ and $\mathcal{C}_\gamma(\bar{k})$ for $\mathbb{S}$ events (see eq. (3.38)), or k and $\mathcal{C}_\gamma(k)$ for $\mathbb{H}$ events (see eq. (3.44))[21]. The values $\mathtt{X}$ written above represent either the starting scales $\mu_{\bar{k}\mathcal{C}_\gamma(\bar{k})}$ for $\mathbb{S}$ events or the stopping scales $\sqrt{t_{k\mathcal{C}_\gamma(k)}}$ for $\mathbb{H}$ events. As is shown in the example above, the `<scales>` tag is the last entry of a LHE. This structure is compatible with the guidelines of the Les Houches accord [23, 24], and with any recent PYTHIA8 8.2 version.

5 Results

In this section we present MC@NLO-Δ predictions for several hadroproduction processes, and compare them with their MC@NLO counterparts. All of these results have been obtained by means of MG5_AMC [17] and PYTHIA8 [19]; the implementation of the MC@NLO-Δ prescription in the former, and the corresponding utility modules in the latter, will become publicly available shortly after the release of this paper.

We have performed our runs by adopting the default MG5_AMC parameters; all results are relevant to pp collisions at $\sqrt{S} = 13$ TeV. The particle masses and widths are set as follows:

$$m_t = 173 \ \text{GeV}\,, \qquad m_H = 125 \ \text{GeV}\,, \tag{5.1}$$

$$m_W = 80.385 \ \text{GeV}\,, \qquad m_Z = 91.188 \ \text{GeV}\,, \tag{5.2}$$

$$\Gamma_W = 2.047600 \ \text{GeV}\,, \qquad \Gamma_Z = 2.441404 \ \text{GeV}\,. \tag{5.3}$$

Both the Higgs and the top-quark widths have been set equal to zero, the latter choice being allowed by the specific production processes we have considered, that do not feature internal top-quark propagators which might go on-shell. We have adopted the central NNPDF2.3 PDF set [25], that is associated with the value

$$\alpha_s(m_Z) = 0.119\,. \tag{5.4}$$

---

[21]In other words, {a,b}={I,J}, where I and J label two particles in the LHE file such that ICOLUP(K,I)=ICOLUP(L,J), with K, L $\in \{1, 2\}$.

We have also set:

$$G_F = 1.16639 \cdot 10^{-5} \text{ GeV}^{-2}, \qquad \frac{1}{\alpha} = 132.507 \,. \tag{5.5}$$

The central values of the renormalisation and factorisation scales have been taken equal to the reference scale:

$$\mu = \frac{H_T}{2} \equiv \frac{1}{2} \sum_i \sqrt{m_i^2 + p_T^2(i)} \,, \tag{5.6}$$

where the sum runs over all final-state particles. In this paper, we have not considered the theoretical systematics associated with the variations of these scales. The PYTHIA8 parameters are the default ones, with the possible exception of those settings specific to MC@NLO matching[22]; these apply to MC@NLO-Δ as well. The hadronisation, underlying events, and QED showers have been turned off; the top, Higgs, and electroweak vector bosons emerging from the hard processes have been treated as stable, and thus left undecayed.

We have considered the following processes:

$$pp \longrightarrow e^+ e^- \,, \tag{5.7}$$

$$pp \longrightarrow e^+ \nu_e \,, \tag{5.8}$$

$$pp \longrightarrow H \,, \tag{5.9}$$

$$pp \longrightarrow H b\bar{b} \,, \tag{5.10}$$

$$pp \longrightarrow W^+ j \,, \tag{5.11}$$

$$pp \longrightarrow W^+ t\bar{t} \,, \tag{5.12}$$

$$pp \longrightarrow t\bar{t} \,, \tag{5.13}$$

that constitute a sufficiently diverse set as far as underlying partonic production mechanisms, complexity of phase-space, induced parton-shower dynamics, and behaviour under matching (including the fractions of negative-weight events) are concerned. This allows one to obtain a reasonably complete comparison between MC@NLO and MC@NLO-Δ results, as well as to have a first idea of the main features of the latter matching prescription. The process in eq. (5.10) has been computed, with $m_b = 4.7$ GeV, in a four-flavour scheme; thus, there is a slight inconsistency due to the usage of the (five-flavour scheme) NNPDF2.3 PDFs, which is however irrelevant for the purpose of the present study. The results of the process in eq. (5.11) have been obtained by imposing a $p_T \geq 50$ GeV cut on the hardest jet of the event; jets are reconstructed by means of FASTJET [26], with an $R = 0.5$ anti-k_T algorithm [27]. We remind the reader that the starting scales are determined as is explained in sect. 3.1; in particular, see eq. (3.22) (for MC@NLO) and eq. (3.25) (for MC@NLO-Δ), where f_α are free parameters, whose values we are soon going to specify. In order to do that, in view of what is implemented in the MG5_AMC code it is customary to define the f_α's in a redundant way, namely:

$$f_\alpha = \kappa \hat{f}_\alpha \,. \tag{5.14}$$

---

[22]See http://amcatnlo.web.cern.ch/amcatnlo/list_detailed2.htm#showersettings.

	MC@NLO			MC@NLO-Δ		
	111	221	441	Δ-111	Δ-221	Δ-441
$pp \to e^+e^-$	6.9% (1.3)	3.5% (1.2)	3.2% (1.1)	5.7% (1.3)	2.4% (1.1)	2.0% (1.1)
$pp \to e^+\nu_e$	7.2% (1.4)	3.8% (1.2)	3.4% (1.2)	5.9% (1.3)	2.5% (1.1)	2.3% (1.1)
$pp \to H$	10.4% (1.6)	4.9% (1.2)	3.4% (1.2)	7.5% (1.4)	2.0% (1.1)	0.5% (1.0)
$pp \to H b\bar{b}$	40.3% (27)	38.4% (19)	38.0% (17)	36.6% (14)	32.6% (8.2)	31.3% (7.2)
$pp \to W^+j$	21.7% (3.1)	16.5% (2.2)	15.7% (2.1)	14.2% (2.0)	7.9% (1.4)	7.4% (1.4)
$pp \to W^+t\bar{t}$	16.2% (2.2)	15.2% (2.1)	15.1% (2.1)	13.2% (1.8)	11.9% (1.7)	11.5% (1.7)
$pp \to t\bar{t}$	23.0% (3.4)	20.2% (2.8)	19.6% (2.7)	13.6% (1.9)	9.3% (1.5)	7.7% (1.4)

Table 1: Fractions of negative-weight events, f, and the corresponding relative costs, $c(f)$ (in round brackets), for the processes in eqs. (5.7)–(5.13), computed with MC@NLO (columns 2–4) and with MC@NLO-Δ (columns 5–7), for three different choices of the folding parameters.

The default choices of these parameters for all of the processes in eqs. (5.7)–(5.13), except for that in eq. (5.10), are the following:

$$\kappa = 1\,, \qquad \hat{f}_1 = 0.1\,, \qquad \hat{f}_2 = 1\,, \tag{5.15}$$

while in the case of eq. (5.10) we set:

$$\kappa = \frac{1}{2}\,, \qquad \hat{f}_1 = 0.1\,, \qquad \hat{f}_2 = 1\,. \tag{5.16}$$

The reduced value of the κ parameter in eq. (5.16) w.r.t. that of eq. (5.15) is in keeping with the findings of ref. [28]. In the case of MC@NLO, we use (see eq. (3.22)):

$$R = \frac{H_T}{2}\,, \tag{5.17}$$

which has been the MG5_AMC default since version 2.5.3[23].

   We start by reporting, in table 1, the overall fractions f of negative-weight events, expressed in percentage, for the processes in eqs. (5.7)–(5.13). For each such fraction, we also give the value of the corresponding relative cost, defined in eq. (1.6) and computed with $C_\pm = 0$ – these are the entries in round brackets. For each of the processes that we have considered, there are six results; those in columns 2 to 4 are obtained with MC@NLO, while those in columns 5 to 7 are obtained with MC@NLO-Δ. The three results relevant to a given matching prescription (either MC@NLO or MC@NLO-Δ) differ by the choices of the folding parameters

$$n_\xi\,, \qquad n_y\,, \qquad n_\varphi\,, \tag{5.18}$$

---

[23]This renders a direct comparison of the present MC@NLO results for $H b\bar{b}$ production with those of ref. [28] impossible, in view of a different choice of the scale R made in that paper. However, the combination of eqs. (5.16) and (5.17) leads to predictions which are rather similar to those obtained with the $\alpha = 1/4$ choice of ref. [28], i.e. the default there. See that paper for more details.

introduced in eq. (4.1); the labellings of the columns correspond to the three integers in eq. (5.18), in that order. Note that $(n_\xi, n_y, n_\varphi) = (1, 1, 1)$ is equivalent to not performing any folding.

While for any given matching prescription and folding choice there are large differences among the results, presented in table 1, associated with the various processes considered, the pattern of the reduction of the fraction of negative-weight events is remarkably similar across processes. Namely, by fixing the folding parameters MC@NLO-Δ has a smaller fraction of negative weights w.r.t. MC@NLO; and by fixing the matching prescription, an increase of the folding parameters leads to a decrease of the fraction of negative weights. This implies that, as far as the relative cost is concerned, MC@NLO without folding and MC@NLO-Δ with maximal folding sit at the opposite ends of the spectrum, the former being the worst (largest $c(f)$) and the latter being the best (smallest $c(f)$). A closer inspection of such a pattern, that involves considering the fractions of negative weights separately for $\mathbb{S}$ and $\mathbb{H}$ events (not shown in the table), confirms that the mechanism that underpins the reduction of negative weights is the one sketched out just before sect. 3.1; more details on this point will be presented in sect. 5.1, using $t\bar{t}$ production as an example.

We point out that the qualification "maximal" applied above to folding obviously refers to the set of parameter choices we have employed. While this is far from being complete, we can confidently make a few observations. Firstly, we have verified that by increasing n_φ there is a negligible reduction of negative weights, while of course the running time increases; this is the reason why we have only presented $n_\varphi = 1$ results. Secondly, heuristically it appears that symmetric choices $n_\xi = n_y$ lead to smaller relative costs than asymmetric ones $n_\xi \neq n_y$. Thirdly, while we did consider $n_\xi = n_y > 4$ for the simplest processes, the amount of reduction of negative weights did not seem worth it, in view of the corresponding increase of the running time. This is likely due to the fact that, for the simplest processes, the fraction of negative-weight events is already quite small for our maximal-folding choice, so that any further reduction is difficult to achieve. Conversely, depending on the user's computing resources, in the case of more involved processes the exploration of stronger foldings than those considered here might be envisaged.

We now turn to considering differential distributions. We point out that, for each process, we have studied the behaviour of dozens of observables, with the largest numbers in the cases of the processes with richer final states. For the sake of this paper, in view of the limitations of space it entails, we have restricted ourselves to discussing the case of an observable which is very often used as a test case in the context of matching procedures, namely the transverse momentum of the set of particles that are present in the final state at the Born level (for the process in eq. (5.11) such a quantity is ill-defined; we use the W-hardest jet system instead, which is its infrared-safe analogue). This observable has several appealing features. For example, it constitutes a practical choice for the pull introduced in eq. (2.11); thus, it is very directly connected to the mechanism of the reduction of negative-weight events which is central to this work. Furthermore, by simply considering a sufficiently large domain, one can separately assess the behaviour of the matched predictions in the soft/collinear, intermediate, and hard regions. This is helpful also because the theoretical systematics due to matching (at least, in MC@NLO-type prescriptions) is
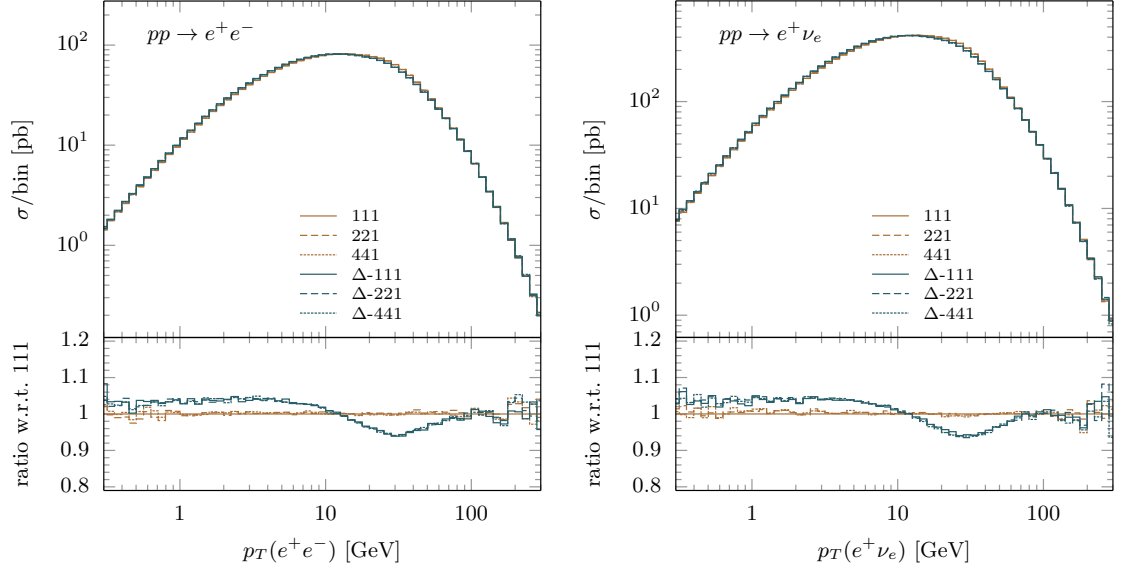
Figure 2: Transverse momentum of the e^+e^- pair (left panel) and of the $e^+\nu_e$ pair (right panel), for the process in eqs. (5.7) and (5.8), respectively. Both MC@NLO (brown histograms) and MC@NLO-Δ (blue histograms) results are shown, for three choices of folding parameters.

mostly concentrated in the intermediate region; therefore, such an observable is essentially a worst-case scenario as far as this systematics is concerned. Indeed, for the purpose of a comparison between MC@NLO and MC@NLO-Δ predictions, we have verified that the conclusions that can be drawn by studying the transverse momentum of the Born system have a rather general validity, in that by far and large they encompass the cases of the other observables that we have considered.

Our results are displayed in figs. 2–4 for the processes in eqs. (5.7)–(5.12); the case of $t\bar{t}$ production will be dealt with later (see sect. 5.1). These figures feature two panels, each relevant to a specific process, composed of a main frame and a lower inset, and all have the same layout. In the main frame there are six histograms, three of which are relevant to MC@NLO predictions (brown lines) and the other three to MC@NLO-Δ predictions (blue lines). The three results in each set differ from each other by the choice of the folding parameters: $(n_\xi, n_y, n_\varphi) = (1, 1, 1)$ (i.e. no folding; solid), $(n_\xi, n_y, n_\varphi) = (2, 2, 1)$ (dashed), and $(n_\xi, n_y, n_\varphi) = (4, 4, 1)$ (dotted). Thus, the six curves in the main frames correspond to the six results associated with each process in table 1. The lower insets of the figures show the bin-by-bin ratios of each of the histograms that appear in the main frames, over the histogram relevant to the corresponding MC@NLO result with no folding; the same plotting patterns as in the main frames are employed.

In fig. 2 we present the predictions for the Z- (left panel) and W^+-mediated (right panel) first-family lepton pair production of eqs. (5.7) and (5.8); not surprisingly, they look very similar to each other, and the only reason to show both of them is their prominence

for both SM measurements and new-physics searches at the LHC. The MC@NLO and MC@NLO-Δ results are within $\pm 5\%$ of each other in the whole range considered. They coincide in absolute value at large p_T, in keeping with one of the defining features of the MC@NLO matching, namely that in hard regions both the shape and the normalisation must coincide, up to small shower effects, with the underlying fixed-order NLO prediction. This shows that such a feature is also exhibited by MC@NLO-Δ, as is expected by construction. At small transverse momenta the MC@NLO and MC@NLO-Δ results have the same shapes, which in the case of the former matching prescription is known to coincide with that of the underlying parton-shower MC; again, this shows that this property holds for MC@NLO-Δ as well. As was anticipated, the differences between MC@NLO and MC@NLO-Δ are visible in the intermediate-p_T region, and must be attributed to matching systematics; in fact, they are of the same order as those one would obtain e.g. by varying the f_α parameters that control the starting scale(s) of the parton showers; we shall give in sect. 5.1 an explicit example on this point. We finally remark that, for both of the matching prescriptions, the differences between the results obtained with different folding parameters are statistically compatible with zero, which confirms the nature of the folding as a technical tool which has no impact on physics predictions.
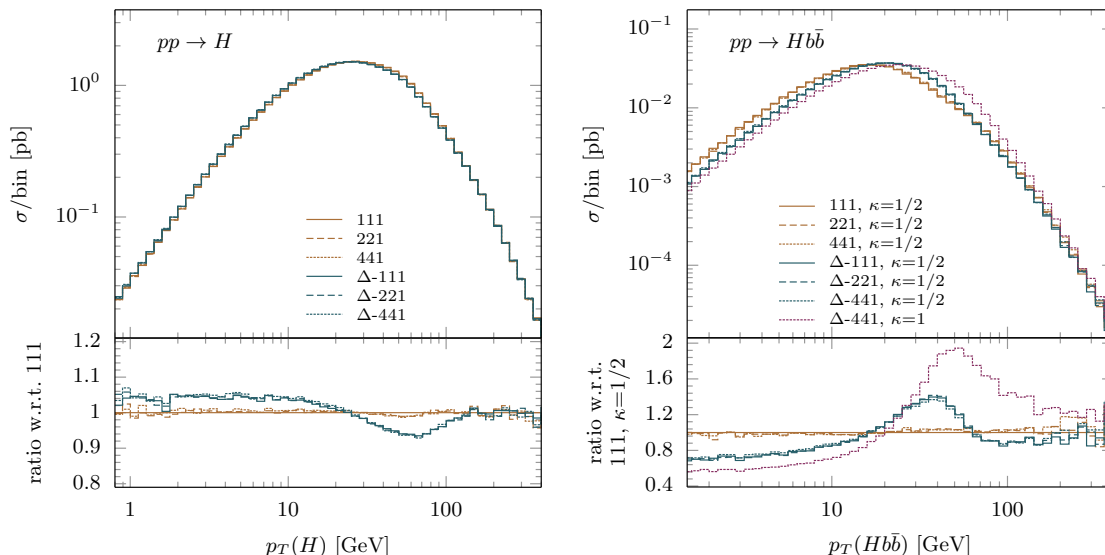

Figure 3: As in fig. 2, for the transverse momentum of the Higgs (left panel) and of the $Hb\bar{b}$ system (right panel), for the processes in eqs. (5.9) and (5.10), respectively. The latter displays an additional MC@NLO-Δ curve (red dotted); see the text for details, in particular eq. (5.14) for the definition of the parameter κ.

In fig. 3 we show the results for production processes that feature an SM Higgs – single-inclusive (eq. (5.9)) in the left panel, and in association with a $b\bar{b}$ pair (eq. (5.10)) in the right panel. As far as single-inclusive production is concerned, the same comments as for the lepton-pair processes of fig. 2 can be repeated verbatim. While this fact could have been expected, it is nevertheless an excellent test of the MC@NLO-Δ machinery. In

fact, the production of lepton pairs is dominated by $q\bar{q}$ annihilation, that of Higgs by gg fusion; as a consequence of this, the structures of the Δ factors relevant to the two cases are very different from each other.

The situation visibly changes for $H b\bar{b}$ production, which is a process notoriously affected by very large systematics (see e.g. ref. [29]). Thus, although the qualitative pattern of the MC@NLO vs MC@NLO-Δ comparison is similar to that of the other processes considered so far, in absolute value the differences are larger, up to $\pm 30\%$. We observe that the MC@NLO-Δ results are harder than the MC@NLO ones (i.e. the peak of the distribution is at larger p_T values). A by-product of this fact is that the transition between the peak and the tail regimes appears to be more regular for MC@NLO-Δ than for MC@NLO – in the latter case, the histograms are flatter in the region 20 GeV $\lesssim p_T \lesssim$ 70 GeV. We also point out that, following ref. [28], $H b\bar{b}$ production has been simulated with the parameters of eq. (5.16). Interestingly, in the case of MC@NLO, the choice of eq. (5.15) in conjunction with the reference scale of eq. (5.17) (i.e. a combination that had not been considered in ref. [28]) leads, for some observables, to a pathological behaviour – the distribution and the large fraction of negative-weight events render their cancellation a very difficult task, so that the results are always affected by dominant statistical uncertainties. It is therefore reassuring that this is *not* the case for MC@NLO-Δ: in fig. 3 we show the result for this matching prescription (and one choice of folding parameters, the others being statistically identical to it) obtained with eq. (5.15) (red dotted histogram); this result is as well behaved as those obtained with eq. (5.16). It is clear that large matching uncertainties remain, but this appears to be characteristic of this production process at this perturbative order; if anything, the present MC@NLO-Δ systematics are smaller than their MC@NLO counterparts of ref. [28] (see in particular fig. 5 there).

We finally turn to discussing W^+ production in association with a light jet (eq. (5.11), left panel of fig. 4) and with a $t\bar{t}$ pair (eq. (5.12), right panel of fig. 4). In the case of the former process, a $p_T \geq 50$ GeV cut on the $R = 0.5$, anti-k_T hardest jet is imposed. The MC@NLO and MC@NLO-Δ predictions for $W^+ j$ production are rather close to each other; differences are generally smaller than $\pm 5\%$. We note that the ratio plots are flatter than those relevant to the processes considered so far; this is in part because for the present observable the separation between the various regimes is not as clear-cut as in the other cases (owing to the jet-p_T cut, one is more inclusive here). As far as $W^+ t\bar{t}$ production is concerned, the differences between the MC@NLO and MC@NLO-Δ results are again quite small. However, at variance with what happens in the other processes, the MC@NLO-Δ large-p_T tail is harder than the MC@NLO one. In order to investigate this point further, in fig. 4 we also display the fixed-order (FO) result (red solid histogram). This shows that the matched predictions are significantly different w.r.t. the FO one for up to very large transverse momenta; in other words, the asymptotic regime (where MC@NLO-type and FO result are expected to coincide with each other) is approached in a very slow manner; essentially, one is not yet there at the rightmost end of the range shown in fig. 4. Furthermore, the extreme steepness of the distribution at large p_T implies that even small parton-shower effects might induce visible bin migrations; this is a timely reminder of the fact that, in this region, the predictions we are considering are LO-accurate in perturbation
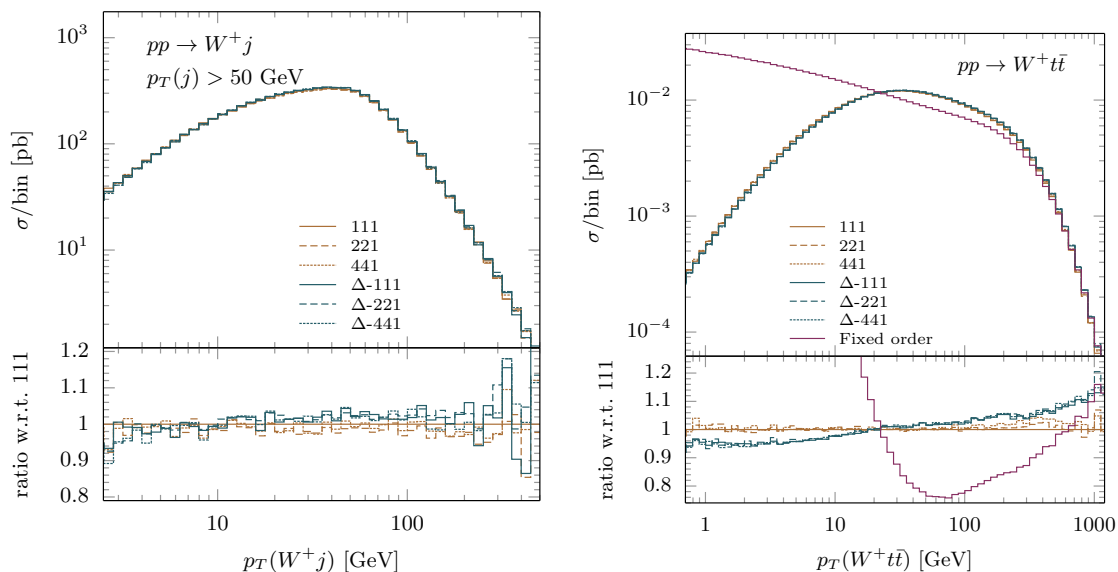
Figure 4: As in fig. 2, for the transverse momentum of the W-hardest jet system (left panel) and of the $W^+t\bar{t}$ system (right panel), for the processes in eqs. (5.11) and (5.12), respectively. In the right panel, the fixed-order result (red solid) is also shown.

theory. We have verified that, by plotting this distribution at the level of the hard events, the MC@NLO and MC@NLO-Δ results are on top of each other, and on top of the FO one; thus, the differences between the former two predictions are indeed stemming from different shower-scale assignments.

5.1 A closer look at $t\bar{t}$ production

In this section we discuss in more details the case of $t\bar{t}$ hadroproduction (eq. (5.13)); this will give us the opportunity to re-consider the arguments of sect. 2 in light of a specific example. We start by considering the analogue of the observables displayed in figs. 2–4, namely the transverse momentum of the $t\bar{t}$ pair. This is shown in fig. 5, whose layout is identical to that of the figures considered so far. The conclusions are also qualitatively fairly similar to those relevant to the other processes. We note that the differences between the MC@NLO and MC@NLO-Δ predictions are larger here in the intermediate region; we anticipate (see fig. 6) that this is not the signal of a genuine discrepancy between the two prescriptions, but it rather reflects a matching uncertainty that affects both matching techniques in a similar manner. The large-p_T region also exhibits some of the features relevant to $W^+t\bar{t}$ production that we have described in fig. 4; they are however milder here, owing to the distribution being less steep and the final-state less massive than in the previous case (thus, the asymptotic regime occurs at p_T values relatively smaller than in the case of $W^+t\bar{t}$ production).

We now study how the predictions after parton showers (e.g. those in fig. 5) are related to the hard events they stem from; the reader must bear in mind that, as always in
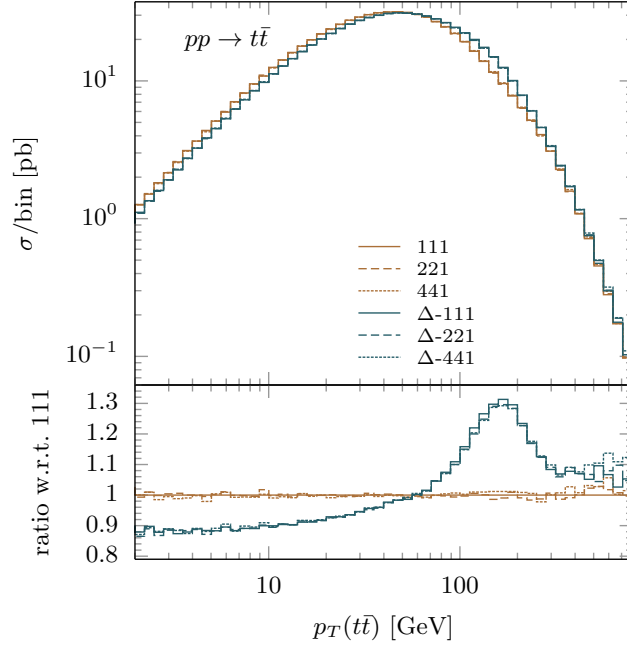
Figure 5: As in fig. 2, for the transverse momentum of the $t\bar{t}$ pair, for the process in eq. (5.13).

the context of MC@NLO-type matchings, the latter quantities are meaningful only in a technical, as opposed to a physical, sense. In order to do so, on top of considering the results obtained with the parameters in eq. (5.15), we also use what follows:

$$\kappa = 1, \qquad \hat{f}_1 = 0.1, \qquad \hat{f}_2 = 0.55, \tag{5.19}$$

$$\kappa = 1, \qquad \hat{f}_1 = 0.55, \qquad \hat{f}_2 = 1. \tag{5.20}$$

We call the settings of eqs. (5.15), (5.19), and (5.20) the *baseline*, *lower*, and *upper* choices respectively, and abbreviate these names where necessary as *b*, *l*, and *u*. We stress that this notation has no phenomenological implications. In other words, we do not claim that the baseline choice should be regarded as the default option when a comparison to data is performed; it is simply the choice of settings which is formally closest to the one usually adopted in MG5_AMC MC@NLO simulations. Whether it is the baseline, the lower, or the upper settings (or something else altogether) that are best for phenomenological MC@NLO-Δ applications is an open question, which is beyond the scope of the present paper.

In fig. 6 we consider again the transverse momentum of the $t\bar{t}$ pair. There are six panels in the figure, and all of them have the same layout. The black histograms are the MC@NLO (upper row) and MC@NLO-Δ (lower row) results; by definition, they are obtained after parton showers. The brown histograms represent results at the hard-event level, i.e. before parton showers; the dark-brown (peaking at relatively larger p_T's) and light-brown (peaking at relatively smaller p_T's) ones are associated with positive-weight
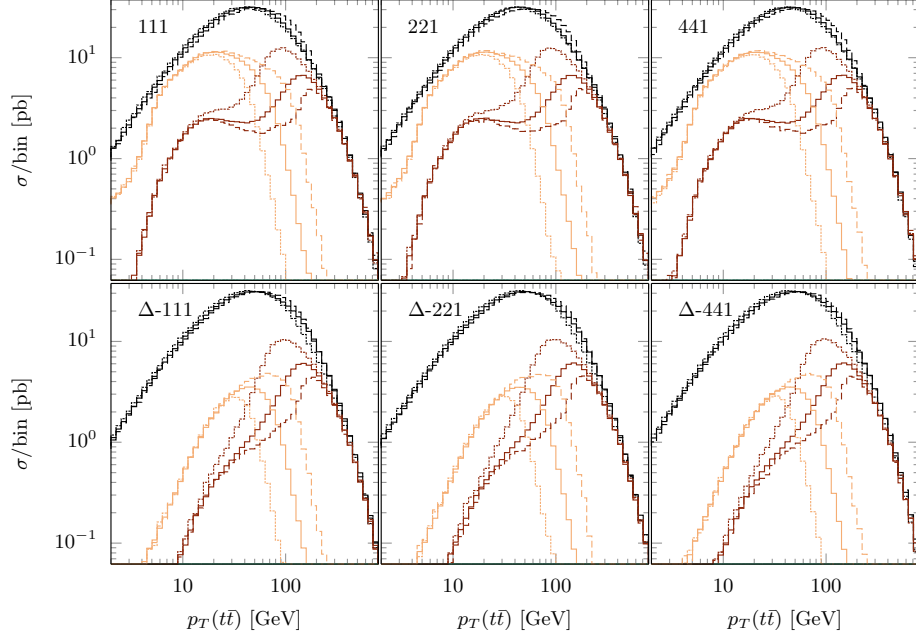
Figure 6: Transverse momentum of the $t\bar{t}$ pair, after parton showers (black histograms), and before parton showers for positive-weight (dark brown histograms) and negative-weight (light brown histograms) events; see the text for details.

and negative-weight events, respectively; in order to display them together, the latter are plotted by flipping their signs. Note that only $\mathbb{H}$ events contribute to the hard-event histograms, since for $\mathbb{S}$ events one has $p_T(t\bar{t}) = 0$, which is outside the domain of fig. 6. For each type of histogram (after parton showers and before parton showers with event weights of either sign) there are three curves; the dotted, solid, and dashed ones are obtained with the lower (eq. (5.19)), baseline (eq. (5.15)), and upper (eq. (5.20)) settings, respectively. It follows that the solid black histograms are the same predictions as those already shown in fig. 5 as blue curves. The three panels in each of the two rows of fig. 6 differ by the choice of the folding parameters.

There are several pieces of information that can be extracted from fig. 6. We start by observing that the three panels on each row are essentially indistinguishable from one another. This is what we expect, since they differ by the choice of the folding parameters; such a choice must induce differences statistically compatible with zero for physical distributions (black histograms), and must not affect $\mathbb{H}$-event distributions (brown histograms) – note that the latter statement would not be true in the case of $\mathbb{S}$ events, as we shall document below. The choice of the $\hat{f}_\alpha$ parameters gives a measure of the matching systematics after parton showers; we note that such systematics (identified as the cumulative difference among the three black histograms that appear on any given panel) are very similar, but not identical, in MC@NLO and MC@NLO-Δ; this confirms that, for precision studies, MC@NLO-Δ matching uncertainties are best assessed independently of the corresponding MC@NLO ones. As was anticipated, the MC@NLO and MC@NLO-Δ systematics emerg-

ing from fig. 6 are consistent with the differences between the MC@NLO and MC@NLO-Δ predictions of fig. 5. At variance with the physical results, those at the hard-event level (brown histograms) exhibit a very strong dependence on the $\hat{f}_\alpha$ parameters: this need not be surprising, owing to their unphysical nature. However, such a large dependence offers one the possibility of a better understanding of the interplay between positive weights, negative weights, and scale assignments.
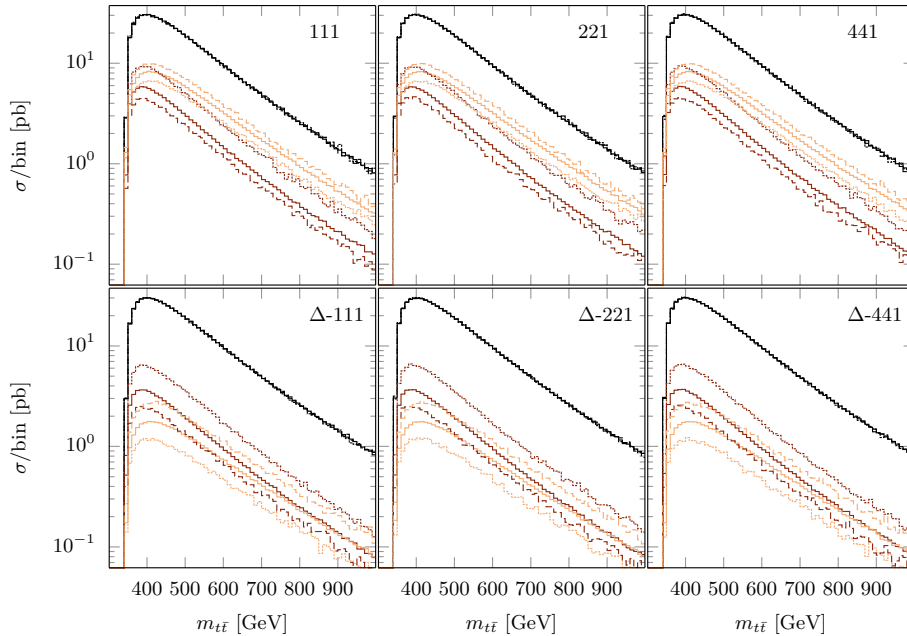

Figure 7: As in fig. 6 for the invariant mass of the $t\bar{t}$ pair. For hard events before parton shower, only $\mathbb{H}$ ones are shown.

In order to proceed, we note preliminarily that the hard-event distributions in fig. 6 are nothing but the distribution in $p_T(t\bar{t})$ as is predicted solely by the short-distance cross section $d\sigma^{(\mathbb{H})}$ (for MC@NLO, see eq. (2.1)) and $d\sigma^{(\Delta,\mathbb{H})}$ (for MC@NLO-Δ, see eq. (3.1)); one simply plots the positive and negative contributions separately. This follows from the fact that by not performing parton showers one leaves the parton-level kinematical configurations unchanged. Furthermore, we remark that the value of $p_T(t\bar{t})$ is a good estimate of the stopping scale relevant to the Sudakov form factor that dominates in the construction of Δ (i.e. the smallest one); as far as the corresponding starting scale is concerned, its value will be in the range $(\kappa\hat{f}_1\mu_t, \kappa\hat{f}_2\mu_t)$ (see eqs. (3.25) and (3.26)), with μ_t predominantly of the order of the top mass; on average, the starting scale will be roughly equal to $\kappa(\hat{f}_1 + \hat{f}_2)\mu_t/2$.

As the comparison of eq. (2.2) with eq. (3.2) shows, we have:

$$d\sigma^{(\Delta,\mathbb{H})} = d\sigma^{(\mathbb{H})}\Delta \equiv \left(d\sigma^{(\mathrm{NLO},E)} - d\sigma^{(\mathrm{MC})}\right)\Delta. \tag{5.21}$$

One of the implications of what has been said about the stopping and starting scales is that $\Delta \to 0$ for small $p_T(t\bar{t})$. Thus, the cross section in this region must be much smaller
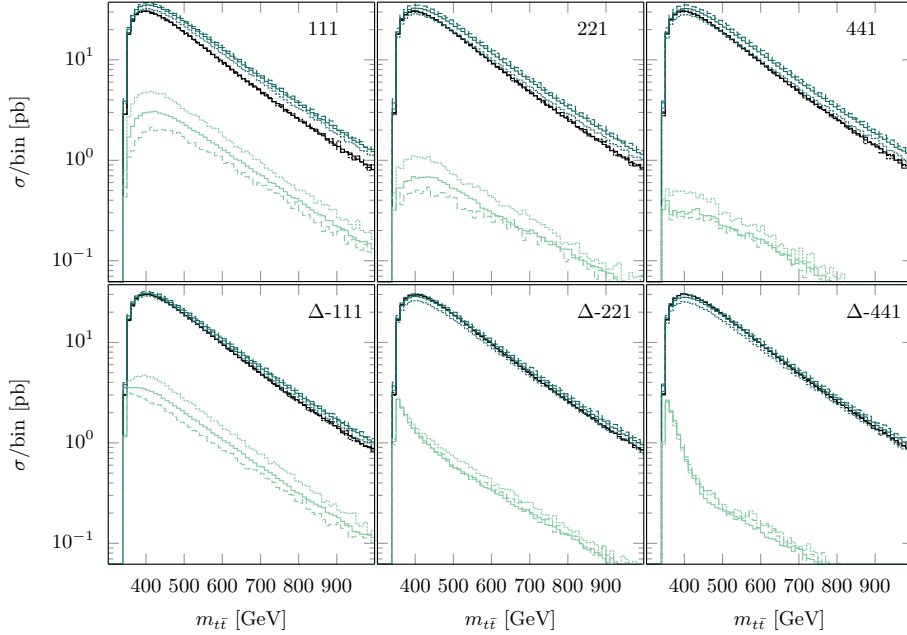
Figure 8: As in fig. 7, for $\mathbb{S}$ events rather than for $\mathbb{H}$ events.

in MC@NLO-Δ than in MC@NLO – this is what one sees by comparing the brown histograms in the lower row of fig. 6 with the corresponding upper-row ones, the differences being increasingly large as one moves towards $p_T(t\bar{t}) = 0$. This mechanism is the anticipated suppression of **N.1** events in MC@NLO-Δ w.r.t. those in MC@NLO; as is expected, positive-weight events are suppressed as well, but this does not lead to a degradation of the statistical accuracy as is explained just before sect. 3.1.

The comparisons between the dotted, solid, and dashed brown histograms relevant to the same type of events (i.e. with positive or negative weights) show that the dependence on the choice of the lower, baseline, and upper settings is negligible at small $p_T(t\bar{t})$, but becomes visible at larger transverse momenta; this is a feature present in both MC@NLO and MC@NLO-Δ predictions, and can be understood as follows. The values of the $\hat{f}_\alpha$ parameters affect the starting scales, but not the stopping scale; accordingly, there is a corresponding dependence of the Δ factor. However, such a dependence is never particularly large, being roughly equal to the ratio of the Sudakovs that dominate in the Δ factor, computed at the different starting scales (see eq. (A.2)); since the latter are typically large, these Sudakovs are then of order one, and so is their ratio. Furthermore, as the stopping scale tends towards zero, the values of Δ become quite small, so that $d\sigma^{(\Delta,\mathbb{H})}$ is vanishingly small as well; this implies that the modest dependence of Δ on $\hat{f}_\alpha$ hardly changes the cross section (in other words, a large suppression factor is as effective as the same suppression factor multiplied by a number of order one). Things do change at larger $p_T(t\bar{t})$: here, the situation is complicated by the fact that the dependence of the short-distance cross section on $\hat{f}_\alpha$ is due not only to the Δ factor, but also to the function D that appears in the MC counterterms (see eq. (3.20); note that $D = 1$ for small transverse momenta).
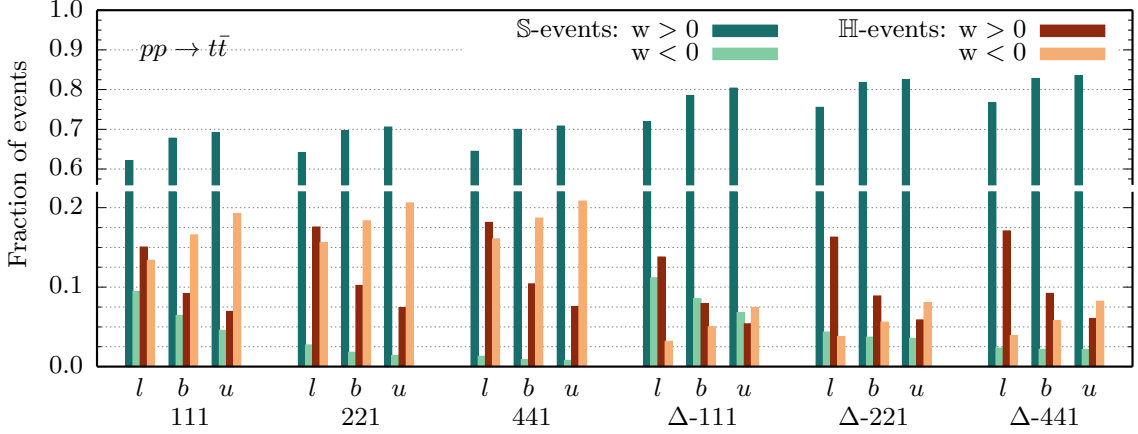
Figure 9: Fractions of positive- and negative-weight events, for both $\mathbb{S}$ and $\mathbb{H}$ events, in $t\bar{t}$ production.

For any given $p_T(t\bar{t})$, the $\hat{f}_\alpha$ dependencies of these quantities result in competing effects in the case of MC@NLO-Δ, while only the latter is present in MC@NLO (owing to the absence of Δ there): larger $\hat{f}_\alpha$ give smaller Δ (i.e. more suppression of both positive- and negative-weight events in MC@NLO-Δ) and larger D values (i.e. larger MC counterterms and therefore more negative-weight events in both MC@NLO-Δ and MC@NLO; see eq. (3.21)). Which of the two mechanisms is dominant in MC@NLO-Δ depends on the kinematical regions one is interested in. In general, given the functional form of the Sudakov form factors, and in particular the way in which they approach the value of one, as $p_T(t\bar{t})$ increases the effect of the Δ factor is subdominant w.r.t. that of the D function, *a fortiori* in the case of MC@NLO. This appears to imply that, perhaps contrary to naive expectations, smaller $\hat{f}_\alpha$ are more effective in reducing negative-weight events not only in MC@NLO, but in MC@NLO-Δ as well. However, it is important to stress that, while the mechanisms described here have a general validity, their relative balance in MC@NLO-Δ is delicate, and dependent on the process. As far as $t\bar{t}$ production is concerned, we shall later document how the conclusion above can be quantified.

In terms of the classification of negative-weight events given in sect. 2, we can recapitulate the discussion above by saying that the factor Δ primarily affects the number of **N.1** events and, to a lesser extent, that of **N.2** events. As far as the D function is concerned, it mostly affects the number of **N.2** events, and only marginally that of **N.1** events. Thus, in MC@NLO-Δ by decreasing (increasing) the $\hat{f}_\alpha$ values one eliminates less (more) **N.1** events, while simultaneously eliminating more (less) **N.2** events. In MC@NLO, only the **N.2** events are affected.

Before summarising our findings, we present plots which are the analogues of those in fig. 6, but are relevant to the invariant mass of the $t\bar{t}$ pair rather than to $p_T(t\bar{t})$. The results are shown in fig. 7 and fig. 8. The layout of fig. 7 is exactly the same as that of fig. 6, namely predictions after parton shower (black histograms) and before parton shower for $\mathbb{H}$ events (brown histograms). In fig. 8 we display the same physical results as
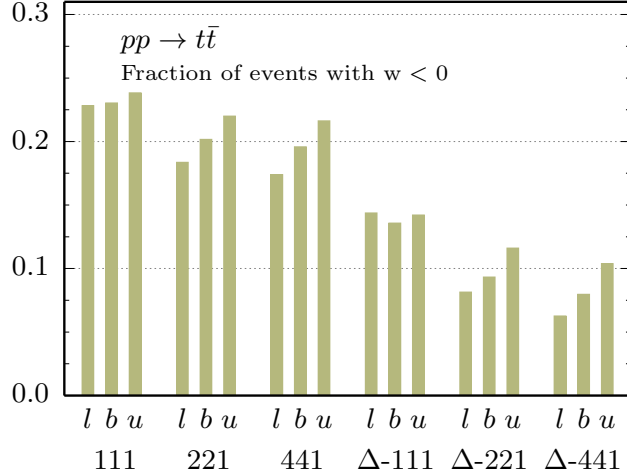
Figure 10: Fractions of negative-weight events, in $t\bar{t}$ production.

in fig. 7, but the distributions at the hard-event level are obtained with $\mathbb{S}$ events, rather than with $\mathbb{H}$ events – at variance with the case of the $t\bar{t}$ pair transverse momentum, $\mathbb{S}$-event contributions are non-trivial in the whole $m_{t\bar{t}}$ range, and are in fact the dominant ones. Positive- and negative-weight $\mathbb{S}$-event distributions are displayed as dark- (relatively larger cross sections) and light-green (relatively smaller cross sections) histograms, respectively. The interpretation of figs. 7 and 8 is not as easy as that of fig. 6, owing to the fact that the pair invariant mass is not a pull variable. Therefore, all of the effects that we have discussed for $p_T(t\bar{t})$ simultaneously contribute to the same $m_{t\bar{t}}$ value, and are impossible to disentangle, although for example the pattern associated with the choices of the $\hat{f}_\alpha$ parameters can still be seen. Furthermore, fig. 8 presents one with the direct evidence of the effects of the folding, which cannot be seen in the $\mathbb{H}$-event distributions of figs. 6 and 7, in particular for what concerns the reduction of **N.3** events.

We conclude this section by presenting a summary of the situation of the fractions of positive- and negative-weight events in $t\bar{t}$ production, fully inclusive over all kinematical variables. In other words, we simply count the number of positive and negative weights present in the various LHE files, and divide them by the total number of events; by using the notation introduced in sect. 1, these ratios are therefore equal to $1 - f$ and f, respectively, when $\mathbb{S}$ and $\mathbb{H}$ events are considered together. By keeping the $\mathbb{S}$- and $\mathbb{H}$-event contributions separate from each other, we obtain the results shown in fig. 9. In that figure, there are three groups of four bins for each of the three MC@NLO and MC@NLO-Δ scenarios (no folding, 221 folding, and 441 folding) that we have considered. Each one of such groups corresponds to a choice of the $\hat{f}_\alpha$ parameters, and is labelled accordingly: lower choice (l, left), baseline choice (b, center), and upper choice (u, right). Finally, the four bins in each group give the fractions of the various events w.r.t. the total number of events (i.e. the sum of the contents of the four bins is always equal to one); from left to right, positive-weight $\mathbb{S}$ events (dark green), negative-weight $\mathbb{S}$ events (light green), positive-weight $\mathbb{H}$ events (dark brown), and negative-weight $\mathbb{H}$ events (light brown). In summary, the information

contained in each of the bins of fig. 9 is the phase-space integrated version of that given, at the differential level, by the brown and green histograms in figs. 6–8. It should be noted that, by normalising all fractions to the total number of events, in both MC@NLO and MC@NLO-Δ matchings when folding is applied both of the fractions relevant to $\mathbb{H}$ events increase (w.r.t. the case of no folding), while either the negative- or the positive-weight fraction of $\mathbb{S}$ events decreases; in non-pathological situations, it is the former that decreases and the latter that increases. We point out that this is just a by-product of the chosen normalisation, and that the short-distance cross sections have the behaviour already described in sects. 2 and 3, namely that the $\mathbb{H}$-event cross section is unaffected by folding, while the negative contributions to the $\mathbb{S}$-event one are reduced. Furthermore, when $\mathbb{S}$ and $\mathbb{H}$ events are considered together, the total fraction of negative weights always decreases with folding (see fig. 10). Figure 9 summarises what we have discussed so far. For a given choice of the $\hat{f}_\alpha$ and folding parameters, we see a marked decrease in the fraction of negative-weight $\mathbb{H}$ events, and a slight increase in that of negative-weight $\mathbb{S}$ events, in MC@NLO-Δ w.r.t. MC@NLO. For a given choice of the $\hat{f}_\alpha$ parameters, in both MC@NLO and MC@NLO-Δ by folding one visibly reduces the fraction of negative-weight $\mathbb{S}$ events. The overall fraction (f) of negative-weight events that results from the various scenarios is presented in fig. 10, which re-iterates the anticipated fact that folding is always advantageous. In conclusion, figs. 9 and 10 confirm in more details and for the specific case of $t\bar{t}$ production the message that emerges from table 1: the reduction of the relative cost is achieved by moving from MC@NLO to MC@NLO-Δ, and by increasing the amount of folding. The relative cost also depends on the parameters that control the assignments of the starting scales; although the latter might thus be used to further reduce it, this operation must follow their tuning to data, and within the uncertainties associated with such a tuning.

6 Conclusions

In this paper we have introduced a new NLO-accurate matching prescription, that we have called MC@NLO-Δ, by starting from the MC@NLO cross section and by augmenting it with further elements of parton-shower MC origin, with the ultimate goal of reducing the number of negative-weight events that are present in MC@NLO simulations. We have achieved a practical implementation of the MC@NLO-Δ matching, by including it in the MADGRAPH5_AMC@NLO framework in conjunction with the PYTHIA8 MC. We have also taken this opportunity to implement in MADGRAPH5_AMC@NLO the folding [12] technique, which induces a (further) reduction of negative-weight events in both MC@NLO-Δ and MC@NLO. Finally, we have shown how the MC@NLO-Δ matching naturally results in the inclusion, in the hard events eventually to be processed by the MC, of as many shower scales as there are oriented colour connections; thus, there are one (in the case of quarks) or two (in the case of gluons) scales per external strongly-interacting particle. This is expected to give a better modelling w.r.t. the one which is currently in use, based on a single shower scale per event.

We have presented a systematic comparison of the MC@NLO and MC@NLO-Δ predictions for seven processes in hadronic collisions at $\sqrt{S} = 13$ TeV, which constitute a typical set of applications of NLO-matched computations to LHC physics. The conclusion emerging from such a comparison is that the combination of the MC@NLO-Δ matching and of folding leads to a significant reduction of the number of negative weights that affect MC@NLO simulations.

Although the physical results we have presented are based on the PYTHIA8 MC, as is the case for MC@NLO the MC@NLO-Δ matching has a general validity, and can be applied to any parton shower MC. However, one must bear in mind that technically the structure of the implementation in MADGRAPH5_AMC@NLO of the MC@NLO-Δ matching is more involved w.r.t. that of MC@NLO, since at variance with the latter it requires calling MC modules also during the event-generation phase. Therefore, in order to replicate this structure in the case of another MC, the latter should be flexible enough to allow for this (which should be the case for any modern C++-based MC). Alternatively, the relevant modules should be extracted from the MC and included in the MADGRAPH5_AMC@NLO (or, in general, the matrix-element provider) code. For completeness, we mention that such modules must include the capability, for any given kinematical configuration, of computing the Sudakov form factors that enter the definition of Δ, of returning the stopping shower scales, and of deciding whether the given configuration is in an MC dead zone or not.

Reducing the number of negative-weight events is always advantageous from a statistical viewpoint; an elementary quantitative analysis is given in sect. 1. However, as far as the associated financial costs are concerned, one must take into account that such a reduction is generally accompanied by an increase of the cost per event (see eq. (1.7)); importantly, this is limited to the event-generation phase, and it stems from the longer running times necessary to achieve the reduction. This is the case for both the folding (which affects both MC@NLO and MC@NLO-Δ; the increase in running time is limited from above by the product of the folding parameters, and is usually about 75% of that number), and the Δ prescription itself (in our PYTHIA8 implementation, the running time increases by a factor 2–3, with the largest increases affecting the simplest processes). Neither of these aspects should be important in the balance costs/benefits in the context of full experimental simulations, while they deserve consideration for purely-theoretical studies. Both are liable to be improved by streamlining their underlying software structures.

We conclude by observing that while MC@NLO-Δ originates from the MC@NLO prescription, it is a novel matching technique, whose implications must not be assumed to be identical to those of MC@NLO. In particular, we have shown that the physics predictions for several hadroproduction processes that originate from MC@NLO and MC@NLO-Δ are fairly similar to each other. However, exactly because they are not identical, the parameter settings that are optimal for phenomenology studies are most likely not the same in the two cases; this fact has to be taken into account when comparing MC@NLO-Δ results with experimental data. From the theoretical viewpoint, on top of the differences inherent to the matching, those stemming from the novel multi-scale structure of hard-event files will also deserve careful consideration.

Acknowledgements

A   The Sudakovs in PYTHIA8

In this appendix we report on some of the technical details relevant to the pre-tabulation of the PYTHIA8 Sudakov form factors, which we have introduced in sect. 4.

We start by observing that the Sudakovs are defined by means of eq. (3.14); in view of this, it is convenient to construct the look-up tables for $\Delta_a^{(*)}$ rather than for Δ_a, owing to the simpler functional dependence of the former w.r.t. that of the latter. The direct dependence of $\Delta_a^{(*)}$ on the colour line ℓ amounts to knowing whether the beginning and the end of ℓ are in the initial or in the final state[24]. There are therefore four possible cases, namely initial-initial, initial-final, final-initial, and final-final, that we shorten as II, IF, FI, and FF, and call *types*; the first (second) letter identifies the beginning (end) of the colour line. We also have:

$$\mathbb{M} = \{m_a, m_{dip}\}, \tag{A.1}$$

where m_a is the mass of particle a according to the internal MC settings, and m_{dip} is the generalised invariant mass of the pair colour-connected by ℓ, defined as the (square root of the) r.h.s. of eq. (3.24). In practice, the dependence of Sudakovs on m_a is essentially negligible w.r.t. the other dependencies; therefore, in our simulations we have set the value of m_a to a fixed value, equal to the PYTHIA8 default for particle a. We also exploit the fact that:

$$\Delta_a^{(*)}\left(q^2; Q^2, \ell, \mathbb{M}\right) = \frac{\Delta_a^{(*)}\left(q^2; Q_{\max}^2, \ell, \mathbb{M}\right)}{\Delta_a^{(*)}\left(Q^2; Q_{\max}^2, \ell, \mathbb{M}\right)}. \tag{A.2}$$

In other words, thanks to eq. (A.2) we can obtain the functional dependence of the Sudakov on both q^2 and Q^2 in terms of a single parameter, provided that the value of $Q_{\max}^2$ is sufficiently large. We thus arrive at what follows:

- For each Sudakov type (II, IF, FI, FF) and particle identity (u, d, ... t, g – Sudakovs for antiquarks are the same as for quarks) we set up a look-up table. There are therefore 28 tables in total.

- Each entry of a given table is the value of the corresponding $\Delta_a^{(*)}$ factor computed at pre-selected values of $\sqrt{q^2}$ and m_{dip}, called *nodal points*. In other words, a look-up

---

[24]The reader must bear in mind that a belongs to ℓ.

table is a set of numbers associated with a two-dimensional grid, defined by the nodal points in the two-dimensional space spanned by $\sqrt{q^2}$ and m_{dip}.

The nodal points for both the $\sqrt{q^2}$ and m_{dip} variables are chosen by means of the same functional form, namely:

$$\log_{10}\left(\frac{r(j)}{1\text{ GeV}}\right) = r_0\left(\left(\frac{j}{j_{\max}}\right)^b - a\right), \qquad 1 \leq j \leq j_{\max}, \tag{A.3}$$

with r_0 and a such that:

$$r(1) = r_{\min}, \qquad r(j_{\max}) = r_{\max}. \tag{A.4}$$

In eqs. (A.3) and (A.4) the parameters $j_{\max}$, b, $r_{\min}$, and $r_{\max}$ are fixed; they must be chosen so as to optimise the distributions of the nodal points. For the simulations presented in this paper we have used the following settings:

$$b = 1, \qquad r_{\min} = 1\text{ GeV}, \qquad r_{\max} = 7\text{ TeV}, \tag{A.5}$$

for both the $\sqrt{q^2}$ and m_{dip} variables, and

$$j_{\max} = 100 \text{ for } \sqrt{q^2}, \qquad j_{\max} = 50 \text{ for } m_{dip}. \tag{A.6}$$

Equation (A.3) is such that there is a higher density of nodal points towards the lower end of the range, $r_{\min}$; this takes into account the fact that Sudakovs vary faster for small values of q^2 and m_{dip}. In order to increase (decrease) such a density one must choose larger (smaller) values of b (in any case, it is best to have $b = \mathcal{O}(1)$).

We have constructed a program (`gridsudgen.f`), and included it in the MG5_AMC package, that loops over the Sudakov types and particle identities; for each of these, it calls the PYTHIA8 module (which we briefly describe below) that computes the Sudakov for all possible pairs of nodal points. These nodal values are then written in a new program (`sudakov.f`) that includes interpolating instructions as well (in turn, these exploit the `dfint` routine of the CERN libraries, that performs a double linear interpolation). The interpolation is protected in the case of inputs which are outside of the grid of nodal points. In particular: a) if m_{dip} is smaller than its minimal value or larger than its maximal value, an error is returned and the run is stopped; b) if q^2 is larger than its maximal value $Q_{\max}^2$, the Sudakov is set equal to one; c) if q^2 is smaller than its minimal value $q_{\min}^2$, the Sudakov is set either equal to zero if $q^2 < (0.5\text{ GeV})^2$ (with 0.5 GeV taken to be the typical low-energy MC cutoff), or equal to the value resulting from a linear interpolation between zero and the value assumed by the Sudakov at $q^2 = q_{\min}^2$, if $(0.5\text{ GeV})^2 \leq q^2 < q_{\min}^2$. Lest the interpolating `dfint` routine actually extrapolates, we also set (see eq. (A.4)) $q_{\min}^2 = r_{\min}^2$ and $Q_{\max}^2 = r_{\max}^2$. Finally, `sudakov.f` is automatically linked to MG5_AMC when creating the executable relevant to the phase-space integration and the unweighting of events.

The user of MG5_AMC can create his/her own Sudakov pre-tabulation by running `gridsudgen.f` prior to any proper MG5_AMC runs. For convenience, we shall include in

the package a template of `sudakov.f`, that should be appropriate for all applications to LHC physics (this relies on the numerical values on the r.h.s.'s of eqs. (A.5) and (A.6)).

We conclude this appendix by returning to the problem of the actual computation of the Sudakov values that constitute the entries of the pre-tabulated grids. Any such entry is equal to $\Delta_a^{(*)}\left(q^2; Q_{\max}^2, \ell, \mathbb{M}\right)$ (see eq. (A.2)); in turn, at given particle identity (a), Sudakov type (that fixes ℓ), and nodal points (that fix q^2 and $\mathbb{M}$), this is set equal to $\mathcal{P}$, the no-emission probability as is computed directly by PYTHIA8 with the specified conditions (note that this implies that, in the case of a gluon, only one colour line, ℓ, is considered). To this end a new PYTHIA8 module, called `pythia_get_no_emission_prob`, has been included in the code, whose return value is $\mathcal{P}$ and that serves as a wrapper to call lower-lying (and mostly pre-existing) PYTHIA8 modules.

There are two crucial aspects in the computation of $\mathcal{P}$ that must be stressed. We comment on each of them in turn.

Firstly, $\mathcal{P}$ is obtained by employing trial showering [30, 31]. In such a procedure, the parton shower, as is implemented in the event generator, is directly used to sample the no-emission probability, in the following way. Given the two values of the shower variable, $Q_{\max}^2$ and q^2, a shower is generated, at the end of which one sets:

$$\mathcal{P}_{\text{trial}} = 0 \quad \Longleftrightarrow \quad \text{an emission with } q^2 < t < Q_{\max}^2 \text{ did occur}, \tag{A.7}$$

$$\mathcal{P}_{\text{trial}} = 1 \quad \Longleftrightarrow \quad \text{an emission with } q^2 < t < Q_{\max}^2 \text{ did not occur}. \tag{A.8}$$

The procedure is repeated N times, by only changing the random numbers inherent to the showering. Note that the procedure is unitary: the number of showers that result in $\mathcal{P}_{\text{trial}} = 0$, plus the number of those leading to $\mathcal{P}_{\text{trial}} = 1$, is equal to N. One then defines:

$$\mathcal{P} = \frac{1}{N} \sum_{i=1}^{N} \mathcal{P}_{\text{trial},i}, \tag{A.9}$$

with i labelling the trial showers.

Secondly, usually the no-emission probability associated with showers is *not* equal to $\Delta_a^{(*)}$ if a is in the initial state. For this to happen, the evolution needs to be done forwards rather than backwards[25], as is typically the case in order not to lose efficiency given the constraint imposed by the value of the Bjorken x. Since such a constraint is irrelevant for the sake of pre-tabulation, a forward initial-state shower has been implemented, so that indeed $\Delta_a^{(*)} = \mathcal{P}$ regardless of whether a is in the initial or final state.

The use of forward evolution for both initial and final-state showers is convenient for pre-tabulation, but poses a technical problem. Consider the fact that any such shower is driven, among many other things, by the form of the z integral in eq. (3.18). For reasons that will soon become clear, we re-write that integral with a more generic notation for the

---

[25]The validity of some of the properties of backward evolution [32, 33] which are usually assumed to hold has very recently been called into question in ref. [34]. It is beyond the scope of this work to discuss the findings of the latter paper. We limit ourselves to observing that small differences, if any, between backward and forward initial-state evolution are perfectly acceptable in the construction of the Δ factor.

integration range. In the case of (anti)quark and gluon it reads, respectively, as follows:

$$S_q = \int_{z^-}^{z^+} dz \left[\hat{P}_{qq}(z) + \hat{P}_{gq}(z) \right], \tag{A.10}$$

$$S_g = \int_{z^-}^{z^+} dz \left[\hat{P}_{gg}(z) + \sum_f \left(\hat{P}_{q_f g}(z) + \hat{P}_{\bar{q}_f g}(z) \right) \right], \tag{A.11}$$

with $z^\pm$ defined by phase-space and momentum-conservation constraints specific to the given Sudakov type (roughly speaking, one has $z^+ = 1 - \epsilon = 1 - \sqrt{t}/Q + \mathcal{O}(t/Q^2)$). As a matter of fact, in the case of the standard backwards initial-state showers z^- is not relevant, because the lower bound of the integration range of the integrals playing the roles of S_q and S_g in backward evolution is actually equal to the Bjorken x. Thus, for initial-state showers the definition of z^- is simply not implemented in PYTHIA8. Conversely, by evolving forwards a value $z^- > 0$ is necessary, in order to prevent the integrals of eqs. (A.10) and (A.11) from diverging owing to the $z \to 0$ soft-gluon singularity of $\hat{P}_{gq}(z)$ and $\hat{P}_{gg}(z)$. There are at least three different ways to address this issue, which we now enumerate.

1. As has been advocated in eq. (3.18) and in the accompanying text, set $z^+ = 1 - \epsilon$ and $z^- = \epsilon$, by analogy with the final-state case, and because of symmetry considerations driven by the behaviour of the Altarelli-Parisi kernels under a $z \to 1 - z$ transformation. By explicit computation we obtain:

$$\frac{1}{2} \int_\epsilon^{1-\epsilon} dz \left[\hat{P}_{qq}(z) + \hat{P}_{gq}(z) \right] = -2C_F \log \epsilon - \frac{3}{2} C_F + \mathcal{O}(\epsilon), \tag{A.12}$$

$$\frac{1}{2} \int_\epsilon^{1-\epsilon} dz \left[\hat{P}_{gg}(z) + \sum_f \left(\hat{P}_{q_f g}(z) + \hat{P}_{\bar{q}_f g}(z) \right) \right] = -2C_A \log \epsilon - \beta_0 + \mathcal{O}(\epsilon), \tag{A.13}$$

where β_0 is the first coefficient of the QCD β function, and the prefactor $1/2 = 1/N_a$ takes eq. (3.19) into account.

2. Set $z^+ = 1 - \epsilon$ and $z^- = 1/2$. This is another way to exploit the $z \to 1 - z$ properties of the kernels, which allows one to map the $z = 0$ soft singularity onto the $z = 1$ one, with equivalent physical meaning. Then:

$$\int_{1/2}^{1-\epsilon} dz \left[\hat{P}_{qq}(z) + \hat{P}_{gq}(z) \right] = -2C_F \log \epsilon - \frac{3}{2} C_F + \mathcal{O}(\epsilon), \tag{A.14}$$

$$\int_{1/2}^{1-\epsilon} dz \left[\hat{P}_{gg}(z) + \sum_f \left(\hat{P}_{q_f g}(z) + \hat{P}_{\bar{q}_f g}(z) \right) \right] = -2C_A \log \epsilon - \beta_0 + \mathcal{O}(\epsilon). \tag{A.15}$$

Note that the $1/2$ prefactor that appears in eqs. (A.12) and (A.13) must not be used here, owing to the reduced integration range.

3. Set $z^+ = 1 - \epsilon$ and $z^- = 0$, and damp the $z = 0$ soft singularity by means of the formal replacements $\hat{P}_{ij}(z) \to z\hat{P}_{ij}(z)$. We obtain:

$$\int_0^{1-\epsilon} dz\, z \left[\hat{P}_{qq}(z) + \hat{P}_{gq}(z)\right] = -2C_F \log \epsilon - \frac{3}{2}C_F + \mathcal{O}(\epsilon)\,, \qquad (A.16)$$

$$\int_0^{1-\epsilon} dz\, z \left[\hat{P}_{gg}(z) + \sum_f \left(\hat{P}_{q_f g}(z) + \hat{P}_{\bar{q}_f g}(z)\right)\right] = -2C_A \log \epsilon - \beta_0 + \mathcal{O}(\epsilon)\,. \quad (A.17)$$

Note that the multiplication by z of the kernels effectively de-symmetrises the integrands, such that a symmetry factor equal to $1/2$ must not be employed.

As the r.h.s.'s of eqs. (A.12)–(A.17) show, these strategies are equivalent at the leading and subleading terms in ϵ, and differ only by terms of $\mathcal{O}(\epsilon)$. We have implemented all of them, and indeed have verified that the differences among their predictions are very small (note that ϵ is typically a small parameter). Heuristically, we have observed that strategy #3 leads to results which on average are the closest to those of the standard backwards shower, and we have therefore adopted it for our pre-tabulation.

We conclude this appendix by remarking that the use of trial showers, being binary in nature (see eqs. (A.7) and (A.8)) might lead to a relatively coarse result for $\mathcal{P}$ defined as in eq. (A.9). In order to alleviate this problem, we have employed $N = 10^4$ in our pre-tabulation. However, even that level of averaging may not be sufficiently smooth to be considered continuous for our purposes. In view of this, we also make use of the shower-enhancement mechanism described in ref. [31]. This stems from multiplying the splitting kernels by a factor C; such an artificial enhancement has to be compensated at the level of no-emission probability – eq. (A.7) becomes:

$$\mathcal{P}_{\text{trial}} = \left(1 - \frac{1}{C}\right)^n \quad \Longleftrightarrow \quad n \geq 1 \text{ emissions did occur in the } (q^2, Q_{\text{max}}^2) \text{ range}\,, \quad (A.18)$$

while eq. (A.8) is unchanged. One can then still employ eq. (A.9). The fine-graining of the enhanced trial-shower procedure is particularly important for values of q^2 that would result in a high emission rate, i.e. in the soft/collinear regions, since there the Sudakov is a fastly-varying function. To this end, we have observed that $C \geq 3$ is sufficient for the applications we have considered in this paper.

References

[1] S. Frixione and B. R. Webber, *Matching NLO QCD computations and parton shower simulations*, *JHEP* **06** (2002) 029, [hep-ph/0204244].

[2] P. Nason, *A New method for combining NLO QCD with shower Monte Carlo algorithms*, *JHEP* **0411** (2004) 040, [hep-ph/0409146].

[3] S. Hoeche, F. Krauss, M. Schonherr and F. Siegert, *A critical appraisal of NLO+PS matching methods*, *JHEP* **1209** (2012) 049, [1111.1220].

[4] S. Platzer and S. Gieseke, *Dipole Showers and Automated NLO Matching in Herwig++*, *Eur. Phys. J.* **C72** (2012) 2187, [1109.6256].

[5] Z. Nagy and D. E. Soper, *Matching parton showers to NLO computations*, *JHEP* **0510** (2005) 024, [`hep-ph/0503053`].

[6] C. W. Bauer and M. D. Schwartz, *Event Generation from Effective Field Theory*, *Phys.Rev.* **D76** (2007) 074004, [`hep-ph/0607296`].

[7] Z. Nagy and D. E. Soper, *Parton showers with quantum interference*, *JHEP* **0709** (2007) 114, [`0706.0017`].

[8] W. T. Giele, D. A. Kosower and P. Z. Skands, *A simple shower and matching algorithm*, *Phys.Rev.* **D78** (2008) 014026, [`0707.3652`].

[9] C. W. Bauer, F. J. Tackmann and J. Thaler, *GenEvA. I. A New framework for event generation*, *JHEP* **0812** (2008) 010, [`0801.4026`].

[10] S. Jadach, W. Placzek, S. Sapeta, A. Siodmok and M. Skrzypek, *Matching NLO QCD with parton shower in Monte Carlo scheme - the KrkNLO method*, *JHEP* **10** (2015) 052, [`1503.06849`].

[11] C. Oleari and L. Reina, *W +- b b̄ production in POWHEG*, *JHEP* **08** (2011) 061, [`1105.4488`].

[12] P. Nason, *MINT: A Computer program for adaptive Monte Carlo integration and generation of unweighted distributions*, `0709.2085`.

[13] S. Alioli, P. Nason, C. Oleari and E. Re, *A general framework for implementing NLO calculations in shower Monte Carlo programs: the POWHEG BOX*, *JHEP* **06** (2010) 043, [`1002.2581`].

[14] S. Frixione, Z. Kunszt and A. Signer, *Three jet cross-sections to next-to-leading order*, *Nucl.Phys.* **B467** (1996) 399–442, [`hep-ph/9512328`].

[15] S. Frixione, *A General approach to jet cross-sections in QCD*, *Nucl.Phys.* **B507** (1997) 295–314, [`hep-ph/9706545`].

[16] R. Frederix, S. Frixione, F. Maltoni and T. Stelzer, *Automation of next-to-leading order computations in QCD: The FKS subtraction*, *JHEP* **0910** (2009) 003, [`0908.4272`].

[17] J. Alwall, R. Frederix, S. Frixione, V. Hirschi, F. Maltoni, O. Mattelaer et al., *The automated computation of tree-level and next-to-leading order differential cross sections, and their matching to parton shower simulations*, *JHEP* **07** (2014) 079, [`1405.0301`].

[18] R. Frederix, S. Frixione, V. Hirschi, D. Pagani, H. S. Shao and M. Zaro, *The automation of next-to-leading order electroweak calculations*, *JHEP* **07** (2018) 185, [`1804.10017`].

[19] T. Sjostrand, S. Mrenna and P. Z. Skands, *A Brief Introduction to PYTHIA 8.1*, *Comput.Phys.Commun.* **178** (2008) 852–867, [`0710.3820`].

[20] D. Amati, A. Bassetto, M. Ciafaloni, G. Marchesini and G. Veneziano, *A Treatment of Hard Processes Sensitive to the Infrared Structure of QCD*, *Nucl. Phys.* **B173** (1980) 429–455.

[21] S. Catani, B. R. Webber and G. Marchesini, *QCD coherent branching and semiinclusive processes at large x*, *Nucl. Phys.* **B349** (1991) 635–654.

[22] K. Odagiri, *Color connection structure of supersymmetric QCD (2 → 2) processes*, *JHEP* **10** (1998) 006, [`hep-ph/9806531`].

[23] J. Alwall, A. Ballestrero, P. Bartalini, S. Belov, E. Boos et al., *A Standard format for Les Houches event files*, *Comput.Phys.Commun.* **176** (2007) 300–304, [`hep-ph/0609017`].

[24] J. Butterworth, A. Arbey, L. Basso, S. Belov, A. Bharucha et al., *The Tools and Monte Carlo working group Summary Report*, 1003.1643.

[25] R. D. Ball et al., *Parton distributions with LHC data*, *Nucl. Phys.* **B867** (2013) 244–289, [1207.1303].

[26] M. Cacciari, G. P. Salam and G. Soyez, *FastJet User Manual*, *Eur.Phys.J.* **C72** (2012) 1896, [1111.6097].

[27] M. Cacciari, G. P. Salam and G. Soyez, *The Anti-k(t) jet clustering algorithm*, *JHEP* **04** (2008) 063, [0802.1189].

[28] M. Wiesemann, R. Frederix, S. Frixione, V. Hirschi, F. Maltoni and P. Torrielli, *Higgs production in association with bottom quarks*, *JHEP* **02** (2015) 132, [1409.5301].

[29] LHC Higgs Cross Section Working Group collaboration, D. de Florian et al., *Handbook of lhc higgs cross sections: 4. deciphering the nature of the higgs sector*, 1610.07922.

[30] L. Lonnblad, *Correcting the color dipole cascade model with fixed order matrix elements*, *JHEP* **05** (2002) 046, [hep-ph/0112284].

[31] L. Lönnblad, *Fooling Around with the Sudakov Veto Algorithm*, *Eur. Phys. J.* **C73** (2013) 2350, [1211.7204].

[32] T. Sjostrand, *A Model for Initial State Parton Showers*, *Phys. Lett.* **157B** (1985) 321–325.

[33] T. D. Gottschalk, *BACKWARDS EVOLVED INITIAL STATE PARTON SHOWERS*, *Nucl. Phys.* **B277** (1986) 700–738.

[34] Z. Nagy and D. E. Soper, *Evolution of parton showers and parton distribution functions*, 2002.04125.