Published by Institute of Physics Publishing for SISSA/ISAS

Received: *May 6, 2002*
Revised: *June 11, 2002*
Accepted: *June 12, 2002*

Matching NLO QCD computations and parton shower simulations*

Stefano Frixione†

Laboratoire d'Annecy-le-Vieux de Physique de Particules
Chemin de Bellevue, BP 110, 74941 Annecy-le-Vieux CEDEX, France
E-mail: `Stefano.Frixione@cern.ch`

Bryan R. Webber

Cavendish Laboratory, University of Cambridge
Madingley Road, Cambridge CB3 0HE, U.K.
E-mail: `webber@hep.phy.cam.ac.uk`

Abstract: We propose a method for matching the next-to-leading order (NLO) calculation of a given QCD process with a parton shower Monte Carlo (MC) simulation. The method has the following features: fully exclusive events are generated, with hadronization according to the MC model; total exclusive rates are accurate to NLO; NLO results for distributions are recovered upon expansion in $\alpha_{\rm S}$; hard emissions are treated as in NLO computations while soft/collinear emissions are handled by the MC simulation, with the same logarithmic accuracy as the MC; and matching between the hard- and soft/collinear-emission regions is smooth. A fraction of events with negative weight is generated, but unweighting remains possible with reasonable efficiency. The method is clarified using a simple toy model, and illustrated by application to the hadroproduction of $\mathrm{W}^+\mathrm{W}^-$ pairs.

Keywords: QCD, NLO Computations, Hadronic Colliders.

*Work supported in part by the UK Particle Physics and Astronomy Research Council and by the EU Fourth Framework Programme 'Training and Mobility of Researchers', Network 'Quantum Chromodynamics and the Deep Structure of Elementary Particles', contract FMRX-CT98-0194 (DG 12 - MIHT).

†On leave of absence from INFN, Sez. di Genova, Italy

Contents

---

1. Introduction

The reliable prediction of cross sections and final-state distributions for QCD processes is important not only as a test of QCD but also for the design of collider experiments and new particle searches. All systematic approaches to this problem are based on perturbation theory, usually truncated at next-to-leading order (NLO), or in a few cases next-next-to-leading order (NNLO). Such calculations yield the best available results for sufficiently inclusive observables, in kinematic regions where higher-order corrections are not enhanced. However, in many cases a more exclusive description of final states and/or a wider kinematic coverage are needed.

For the description of exclusive hadronic final states, perturbative calculations have to be combined with a model for the conversion of partonic final states into hadrons (hadronization). Existing hadronization models are in remarkably good agreement with a wide range of data, after tuning of model parameters [1, 2]. However, these models operate on partonic states with high multiplicity and low relative transverse momenta, which are obtained from a parton shower [3, 4] or dipole cascade [5] approximation to QCD dynamics and not from fixed-order calculations.

To extend the coverage of perturbative calculations into regions where higher-order contributions are enhanced, one needs to identify and resum the enhanced terms to all orders. The resummed expression can then be combined with the fixed-order result, after subtracting resummed terms of the same order to avoid double-counting. This has been done explicitly for many important observables and processes, but a general procedure has not yet been developed. In principle, the parton shower provides such a procedure, if the problem of double-counting can be overcome. Combining NLO and parton-shower computations also offers the possibility of interfacing to hadronization models, thus overcoming the main deficiencies of existing calculations.

Methods for combining fixed-order and parton-shower or dipole-cascade computations in QCD have been discussed in refs. [6]–[21]. Many of those papers concern the rather different problem of combining leading-order (tree level) matrix elements for hard multi-jet production with parton showers or dipole cascades for jet fragmentation. The only fully-implemented schemes for matching NLO computations and parton showers that have been used to produce phenomenological results [12]–[15] have been based on the so-called phase-space slicing method for regularising the infrared divergences encountered at NLO.

JHEP06(2002)029

The slicing method is simple but has the disadvantage of introducing a slicing parameter, which should in principle be sent to zero at the end of the calculation but in practice cannot be chosen too small. This can introduce uncontrolled errors and/or large negative event weights. The slicing method has therefore tended to be replaced where possible in NLO calculations by subtraction methods, in which no approximations need to be made and event weights can be controlled by improving the choice of subtraction terms.

The subtraction method has been considered in refs. [11, 18, 19, 20]. The main concern of those papers was the correct matching of NLO matrix element calculations to the scale and scheme for next-to-leading logarithmic (NLL) parton showers. Up to now, no satisfactory algorithms for NLL showering have been developed. When such algorithms are available, the methods proposed there will make it possible for event generators to generate the NLO evolution of parton distribution functions (PDFs) and hard process cross sections, without the explicit input of NLO PDFs. Existing parton shower generators, however, correspond only to LL evolution, and therefore the matching criteria of refs. [11, 18, 19, 20] operate beyond the level of precision that can be achieved at present.

In the present paper, we present a method for matching NLO matrix elements and existing parton shower generators based on the subtraction method for NLO calculations. First, in section 2, we outline the general objectives and limitations of our approach. The method is then illustrated in section 3 in the context of a simplified toy model. The case of QCD is dealt with in sects. 4 and 5. In section 4 results are presented for the process of W^+W^- production in hadronic collisions. In section 5 we discuss briefly the application of the method to other, more complex, processes. Concluding comments are given in section 6. All the technical details are collected in the appendices, where we also comment on the matching between NLO computations and parton shower approaches in the framework of the slicing method.

2. General framework

2.1 Objectives

It is undeniable that parton shower Monte Carlos (MC's) provide more general and flexible tools compared to NLO computations. By running an MC we get events, which we believe to be a faithful description of the real events detected by real experiments; this is one of the reasons why the MC's are so popular with experimentalists. On the other hand, NLO computations emphasize the role of (infrared-safe) observables; the presence of negative contributions, and the necessity of considering inclusive quantities in order to get rid of infinities, make it very problematic to even talk about events. Still, NLO computations have their virtues; they can handle hard emissions, and they can estimate total rates with better accuracy than MC's.

In this paper, we aim to construct an MC that works like a regular MC, but on top of that knows how to treat hard emissions, and can compute rates to NLO accuracy. For brevity, we call this tool an MC@NLO. Clearly, there is a lot of freedom in constructing an MC@NLO. We shall follow a minimal approach. Namely, we want to keep the original MC as unmodified as possible; all major changes have to be carried out on the NLO codes.

JHEP06(2002)029

Amongst other things, this implies that our MC@NLO will not improve the logarithmic accuracy of the original MC, in those regions of the phase space where resummation is needed. As far as the NLO calculations are concerned, we shall not perform any approximations, neither at the level of matrix elements, nor at the level of phase space; thus, we shall adopt the subtraction method in order to cancel the infrared divergences that arise in the intermediate steps of the computation.

To be more definite, we require MC@NLO to have the following characteristics:

- The output is a set of events, which are fully exclusive.
- Total rates are accurate to NLO.
- NLO results for all observables are recovered upon expansion of MC@NLO results in α_S.
- Hard emissions are treated as in NLO computations.
- Soft/collinear emissions are treated as in MC.
- The matching between hard- and soft-emission regions is smooth.
- MC hadronization models are adopted.

In section 4, we shall construct an MC@NLO fulfilling these conditions. However, we find that the technical complications of QCD tend to hide the basic ideas upon which our approach is based. Thus, we prefer to start from an oversimplified (and unrealistic) case, where the structure of the MC@NLO is apparent; this is done in section 3. The reason for doing this is that we shall see that the approach that leads to the MC@NLO in a simple case can actually be adopted in the QCD case as well, which will only be technically more involved.

2.2 Negative weights

We do not require the MC@NLO to be positive definite; events may have negative weights. However, it should be pointed out that these negative-weight events are of a completely different nature from the negative contributions that appear in NLO computations. In particular, the distributions of the positive- and negative-weight events are *separately* finite in our MC@NLO (the same is not true in an NLO computation). This implies the possibility of unweighted event generation,[1] as is customary in ordinary MC's. The situation is actually completely analogous to that occurring in polarized collisions, where physical observables are obtained as the difference of two positive-definite and finite quantities. Furthermore, we shall see that in our case the number of negative-weight events is reasonably small, and thus the number of MC@NLO events necessary to get a smooth distribution is comparable to that in ordinary MC's.

[1] In this paper, we adopt a slightly different notation from that commonly used by experimentalists. Namely, we call unweighted events any set of events whose weights are all identical, *up to a sign*. In other words, the weights of our unweighted events are of the form $\pm w$, where w is a constant.

JHEP06(2002)029

2.3 Double counting

When the MC generates events with real parton emission, it generates kinematical configurations that are also taken into account by the NLO computation: the possibility of having the same kinematical configuration from the MC *and* from the NLO may lead to double counting. We shall adopt the following quantitative definition of double counting:

An MC@NLO *is affected by double counting if its prediction for any observable, at the first order beyond the Born approximation*[2] *in the expansion in the coupling constant, is not equal to the NLO prediction.*

According to this definition, double counting may correspond to either an excess or a deficit in the prediction, at any point in phase space. This includes contributions from real emission and virtual corrections. Generally speaking, the MC fills the phase space efficiently only in the soft and collinear regions. In these regions, the leading behaviour of the kinematical distribution of one-particle emission is given by the Born cross section, times a kernel that describes the soft or collinear emission. This holds both for the MC and for the NLO. Thus, in the soft and collinear regions of the one-particle-emission phase space, the MC and NLO results coincide when only the leading terms are considered. Elsewhere, the MC is not reliable and the NLO provides the best estimate. The problem is to merge these two descriptions.

3. Toy model studies

In this section we study a toy model that allows a simple discussion of the key features of the NLO computation, of the MC approach, and of the matching between the two. We assume that a system can radiate massless particles (which we call photons), whose energy we denote by x, with $0 \leq x \leq x_s \leq 1$, x_s being the energy of the system before the radiation. After the radiation, the energy of the system is $x_s' = x_s - x$. The system can undergo several further emissions; on the other hand, one photon cannot split further.

In a perturbative computation, the Born term corresponds to no emissions. The first non-trivial order in perturbation theory gets contribution from those diagrams with one and only one emission, with either virtual or real photons. We write the corresponding contributions to the cross section as follows:

$$\left(\frac{d\sigma}{dx}\right)_{\mathrm{B}} = B\delta(x)\,, \tag{3.1}$$

$$\left(\frac{d\sigma}{dx}\right)_{\mathrm{V}} = a\left(\frac{B}{2\epsilon} + V\right)\delta(x)\,, \tag{3.2}$$

$$\left(\frac{d\sigma}{dx}\right)_{\mathrm{R}} = a\frac{R(x)}{x}\,, \tag{3.3}$$

for the Born, virtual, and real contributions respectively; a is the coupling constant, possibly times a colour factor; B and V are constant with respect to x, and

$$\lim_{x\to 0} R(x) = B\,. \tag{3.4}$$

[2] By the Born approximation we mean the lowest order at which the hard process contributes to any observable.

The constant B appears in eqs. (3.2) and (3.4) since we expect the residue of the leading singularity of the virtual and real contributions to be given by the Born term, times a suitable kernel. We take this kernel equal to 1, since this simplifies the computations, and it is not restrictive. Finally, ϵ is the parameter entering dimensional regularization in $4-2\epsilon$ dimensions.

The task of predicting an infrared-safe observable O to NLO accuracy amounts to computing the following quantity

$$\langle O \rangle = \lim_{\epsilon \to 0} \int_0^1 dx\, x^{-2\epsilon} O(x) \left[\left(\frac{d\sigma}{dx}\right)_{\mathrm{B}} + \left(\frac{d\sigma}{dx}\right)_{\mathrm{V}} + \left(\frac{d\sigma}{dx}\right)_{\mathrm{R}} \right], \tag{3.5}$$

where $O(x)$ is the observable as a function of x, possibly times a set of Θ functions defining a histogram bin. By "infrared-safe" here we simply mean that the integral in eq. (3.5) exists. The main technical problem in eq. (3.5) is due to the presence of the regularising parameter ϵ; in order to have an efficient numerical procedure, it is mandatory to extract the pole in ϵ from the real contribution, in this way cancelling analytically the pole explicitly present in the virtual contribution.

3.1 NLO: slicing versus subtraction

Within the context of this toy model, the usual methods of cancelling the pole in ϵ are as follows. In the *slicing method*, a small parameter δ is introduced into the real contribution in the following way:

$$\langle O \rangle_{\mathrm{R}} = \int_0^\delta dx\, x^{-2\epsilon} O(x) \left(\frac{d\sigma}{dx}\right)_{\mathrm{R}} + \int_\delta^1 dx\, x^{-2\epsilon} O(x) \left(\frac{d\sigma}{dx}\right)_{\mathrm{R}}. \tag{3.6}$$

In the first term on the r.h.s. of this equation we expand $O(x)$ and $R(x)$ in Taylor series around 0, and keep only the first term; the smaller δ, the better the approximation. On the other hand, the second term in eq. (3.6) does not contain any singularity, and we can just set $\epsilon = 0$. We obtain

$$\langle O \rangle_{\mathrm{R}} = aBO(0) \int_0^\delta dx\, \frac{x^{-2\epsilon}}{x} + \int_\delta^1 dx\, O(x) \left(\frac{d\sigma}{dx}\right)_{\mathrm{R}} + \mathcal{O}(\delta) \tag{3.7}$$

$$= a\left(-\frac{1}{2\epsilon} + \log\delta\right) BO(0) + a \int_\delta^1 dx\, \frac{O(x)R(x)}{x} + \mathcal{O}(\delta, \epsilon)\,. \tag{3.8}$$

Using this result in eq. (3.5), we get the NLO prediction for O as given in the slicing method:

$$\langle O \rangle_{\mathrm{slice}} = BO(0) + a\left[(B\log\delta + V)\, O(0) + \int_\delta^1 dx\, \frac{O(x)R(x)}{x} \right] + \mathcal{O}(\delta)\,. \tag{3.9}$$

The terms $\mathcal{O}(\delta)$ cannot be computed, and are neglected; in general, one must check that these neglected terms are numerically small, by plotting $\langle O \rangle_{\mathrm{slice}}$ versus δ in a suitable range in δ.

In the *subtraction method*, no approximation is performed. One writes the real contribution as follows:

$$\langle O \rangle_{\rm R} = aBO(0) \int_0^1 dx \, \frac{x^{-2\epsilon}}{x} + a \int_0^1 dx \, \frac{O(x)R(x) - BO(0)}{x^{1+2\epsilon}} \,. \tag{3.10}$$

The second term on the r.h.s. does not contain singularities, and we can set $\epsilon = 0$:

$$\langle O \rangle_{\rm R} = -a \frac{B}{2\epsilon} O(0) + a \int_0^1 dx \, \frac{O(x)R(x) - BO(0)}{x} \,. \tag{3.11}$$

Therefore, the NLO prediction as given in the subtraction method is:

$$\langle O \rangle_{\rm sub} = BO(0) + a \left[VO(0) + \int_0^1 dx \, \frac{O(x)R(x) - BO(0)}{x} \right] . \tag{3.12}$$

We rewrite this in a slightly different form, which will be more suited to discussion of the matching with the MC:

$$\langle O \rangle_{\rm sub} = \int_0^1 dx \left[O(x) \frac{aR(x)}{x} + O(0) \left(B + aV - \frac{aB}{x} \right) \right] . \tag{3.13}$$

3.2 Toy Monte Carlo

In a treatment based on Monte Carlo methods, the system can undergo an arbitrary number of emissions (branchings), with probability controlled by the *Sudakov form factor*, defined for our toy model as follows:

$$\Delta(x_1, x_2) = \exp \left[-a \int_{x_1}^{x_2} dz \frac{Q(z)}{z} \right] , \tag{3.14}$$

where $Q(z)$ is a monotonic function with the following general properties:

$$0 \le Q(z) \le 1 \,, \qquad \lim_{z \to 0} Q(z) = 1 \,, \qquad \lim_{z \to 1} Q(z) = 0 \,. \tag{3.15}$$

Note that the MC is also well defined without the last condition in eq. (3.15). We discuss later the specific functional form assumed for Q. If x_s is the energy of the system before the first branching occurs, then $\Delta(x, x_s)$ is the probability that no photon be emitted with energy z such that $x \le z \le x_s$. At fixed x, this probability tends to 1 when $a \to 0$; that is, no emission is possible in the zero-coupling limit. On the other hand, when $a \to \infty$, we have $\Delta(x, x_s) \to 0$, which means that it is impossible that the system does not emit. The toy MC code we have written implements ordered emissions in the photon energies[3] and proceeds through the following steps:

0) Define x_0 as the lower bound on the photon energy after a branching; thus x_0 plays the same role as the cutoffs used to define the Sudakov form factor in QCD MC's.

1) Define $x_{\rm M}$ to be the maximum energy available to the photon; before the first branching, we have $x_{\rm M} = x_s$, where x_s is the energy of the system.

[3]This may violate energy-momentum conservation, as is usual in a leading logarithm implementation.

2) Pick a random number $R_1 \in [0,1]$, and solve for z the equation

$$\exp\left[-a\int_z^{x_{\rm M}}\frac{dz'}{z'}\right] \equiv \left(\frac{z}{x_{\rm M}}\right)^a = R_1 \qquad \Longrightarrow \qquad z = x_{\rm M} R_1^{1/a}\,. \tag{3.16}$$

3) Pick a random number $R_2 \in [0,1]$. If $Q(z) < R_2$, let $x_{\rm M} = z$, reject the branching, and return to step 2).

4) If $z < x_0$, reject the branching and exit. If $z > x_0$, a photon with energy z has been emitted. We let $x_{\rm M} = z$, and iterate the procedure going back to 2).

This procedure ensures that the photon emissions have the distribution prescribed by the Sudakov form factor in eq. (3.14). As can be seen from that equation, the probability of having soft emission (i.e., with $z < x_0$), and thus of terminating the shower, is $\Delta(x_0, x_s)$, which tends to one when $a \to 0$ (weak coupling), and when $x_s \to x_0$ (no energy left for non-soft branching). In this toy model, the hardest emission is always the first one.

3.3 Matching NLO and MC

In order to introduce some useful notation, we start from the LO perturbative result, which does not pose any problems, since it is in fact what is implemented in ordinary MC's. At the LO, the system has not yet lost any energy to perturbative photon emission. Thus, the maximum energy available to the photon in the first branching is 1. For each event, we set $x_{\rm M} = 1$, and proceed as described in section 3.2. The total exclusive rate will have to be equal to B, according to eq. (3.1); thus, if we generate N events, each event has a weight equal to B/N.

We can formally read this procedure from eq. (3.13), by setting $a = 0$ and performing the formal substitution

$$O(x) \;\longrightarrow\; I_{\rm MC}(O, x_{\rm M}(x))\,, \tag{3.17}$$

where $I_{\rm MC}(O, x_{\rm M})$ stands for interface-to-MC, and indicates symbolically the distribution in the observable O as obtained by running the MC starting from a given $x_{\rm M}$, and giving each event the weight $1/N$; its precise definition is given in appendix B.1. For $x = 0$ we need $x_{\rm M} = 1$; we do not need to specify $x_{\rm M}(x)$ for other values of x at present. Interfacing LO with MC, we thus obtain

$$\left(\frac{d\sigma}{dO}\right)_{\rm MC@LO} = B\, I_{\rm MC}(O, 1)\,. \tag{3.18}$$

This equation reminds us of the total rate (B), and the maximum energy available for photons at the first branching ($x_{\rm M} = 1$).

3.3.1 Naive subtraction

The first attempt at extending eq. (3.18) to NLO amounts simply to substituting eq. (3.17) into eq. (3.13). We get

$$\left(\frac{d\sigma}{dO}\right)_{\rm naive} = \int_0^1 dx\left[I_{\rm MC}(O, x_{\rm M}(x))\frac{aR(x)}{x} + I_{\rm MC}(O,1)\left(B + aV - \frac{aB}{x}\right)\right]. \tag{3.19}$$

This equation suggests the following procedure:

- Pick at random $0 \leq x \leq 1$.
- Generate an MC event with $x_{\rm M}(x)$ as maximum energy available to the photon in the first branching; attach to this event the weight $w_{\rm EV} = aR(x)/x$.
- Generate another MC event (a "counter-event") with $x_{\rm M} = 1$; attach to this event the weight $w_{\rm CT} = B + aV - aB/x$.
- Repeat the first three steps N times, and normalize with $1/N$.

Unfortunately, this procedure is bound to fail, since the weights $w_{\rm EV}$ and $w_{\rm CT}$ diverge as $x \to 0$. A small cutoff ρ can then be introduced; as $x \to \rho$, we would have $w_{\rm EV} \simeq aB/\rho$ and $w_{\rm CT} \simeq -aB/\rho$. Both these weights are very large in absolute value; this means that the eventual unweighting procedure, performed by the MC, will be highly inefficient. Furthermore, as we shall see in section 3.4, eq. (3.19) has problems of double counting, defined according to the criterion in section 2.3.

3.3.2 Modified subtraction

To avoid the problems outlined above, we define the *modified subtraction method* by

$$\left(\frac{d\sigma}{dO}\right)_{\rm msub} = \int_0^1 dx \Bigg[I_{\rm MC}(O, x_{\rm M}(x)) \frac{a[R(x) - BQ(x)]}{x} + \\ + I_{\rm MC}(O, 1) \left(B + aV + \frac{aB[Q(x) - 1]}{x} \right) \Bigg] . \tag{3.20}$$

This amounts to saying that, in eq. (3.19), we subtract and add the quantity

$$I_{\rm MC}(O, x_{\rm M}) \frac{aBQ(x)}{x} , \tag{3.21}$$

using $x_{\rm M} = x_{\rm M}(x)$ in the first and $x_{\rm M} = 1$ in the second term introduced in this way. We point out that Q is the same function that appears in the definition of the Sudakov form factor, eq. (3.14). By virtue of the second condition in eq. (3.15), the weights $w_{\rm EV}$ and $w_{\rm CT}$ attached to events and counter-events are now separately convergent as $x \to 0$.

The two terms involving $Q(x)$, that we inserted in eq. (3.20), are not identical; therefore, the procedure which leads from eq. (3.19) to eq. (3.20) is not a subtraction in the usual sense of an NLO computation. Still, these two terms do not contribute to the observable O at $\mathcal{O}(a)$, because they are compensated by analogous terms due to the parton shower $BI_{\rm MC}(O, 1)$; this will be shown with explicit examples in section 3.4, and in general in appendix B.3. We point out that the function $aQ(x)/x$ that appears in eq. (3.21) does *not* depend upon the MC cutoff x_0. It follows that the dependence upon x_0 of the MC@NLO, eq. (3.20), is identical to that of the ordinary MC, being entirely due to $I_{\rm MC}$, and is therefore power-like.

It should be clear from the description of the procedure given above that the dependence upon O in eqs. (3.19) and (3.20) is purely formal, serving only to allow us to write such equations in a compact form. We stress that, in the actual implementation of *all* the

MC@NLO's described in this paper, the computation of observables proceeds exactly as in standard MC's: namely, events are generated, showered, and hadronized without reference to any specific observable.

Equation (3.20) is our master equation for the toy model. We shall show in what follows that the MC@NLO defined in this way meets the requirements given in section 2.1. In section 4, we shall use eq. (3.20) to construct a QCD MC@NLO by analogy with the toy model.

3.4 Observables

3.4.1 Exclusive observable

As an example of an "exclusive" observable in the toy model, we shall consider the quantity

$$y = \max(x_1, \ldots, x_n)\,, \tag{3.22}$$

where n denotes the number of emissions in one given event (in the case of no emission, the event is not used to fill the histograms). This observable is the energy of the hardest photon of the event. It is particularly convenient because in the MC approach it coincides with the energy of the first photon emitted in the shower; when MC@NLO is considered, y is the larger of the energy of the photon emitted perturbatively, and that of the first photon in the shower (if any). This fact allows us to obtain analytical results for MC, NLO and MC@NLO.

The NLO prediction for this observable is (here and in the following, we consider only $y > x_0$ in order to simplify the analytical results)

$$\left(\frac{d\sigma}{dy}\right)_{\rm NLO} = a\frac{R(y)}{y}\,, \tag{3.23}$$

while the MC result, normalized to the Born cross section, is

$$\left(\frac{d\sigma}{dy}\right)_{\rm MC} = aB\frac{Q(y)}{y}\Delta(y,1)\,, \tag{3.24}$$

which is just the differential cross section at $x = y$, computed using eq. (3.14), times the probability that there is no harder emission at $x > y$.

In order to study the MC@NLO results, we have to consider the quantities $I_{\rm MC}(y,1)$ and $I_{\rm MC}(y, x_{\rm M}(x))$ that will be substituted into eq. (3.19) or (3.20). In the case of $I_{\rm MC}(y,1)$, the hardest photon comes from the MC and we have

$$I_{\rm MC}(y,1) = a\frac{Q(y)}{y}\Delta(y,1)\,. \tag{3.25}$$

In the case of $I_{\rm MC}(y, x_{\rm M}(x))$, the hardest photon may come from the NLO contribution or the MC, and so we find

$$\begin{aligned} I_{\rm MC}(y, x_{\rm M}(x)) = {} & \Delta\Big(\min\{x, x_{\rm M}(x)\}, x_{\rm M}(x)\Big)\delta(y-x) + \\ & + a\frac{Q(y)}{y}\Delta(y, x_{\rm M}(x))\Theta(y-x)\Theta(x_{\rm M}(x)-y)\,. \end{aligned} \tag{3.26}$$

As a shorthand notation, in the following we shall denote the contributions to $d\sigma/dy$ from $I_{\rm MC}(y,1)$ and $I_{\rm MC}(y,x_{\rm M}(x))$ by $d\sigma_{\mathbb{S}}/dy$ and $d\sigma_{\mathbb{H}}/dy$ respectively, since they correspond to a $\mathbb{S}$tandard MC evolution (i.e., with no prior NLO real emission), and to a $\mathbb{H}$ard MC evolution (whose initial condition includes one NLO real emission). This notation will be extensively used from section 4.1 on. In the case of the naive subtraction method, eq. (3.19), these contributions are separately divergent, but their sum is not. We find

$$\left(\frac{d\sigma_{\mathbb{S}}}{dy}\right)_{\rm naive} = a\frac{Q(y)}{y}\Delta(y,1)\left[B + aV - aB\int_0^1 \frac{dx}{x}\right]. \tag{3.27}$$

The results for $d\sigma_{\mathbb{H}}/dy$ depend somewhat on the form of the function $x_{\rm M}(x)$ introduced in eq. (3.17). We define x_e to be the solution of $x = x_{\rm M}(x)$, i.e., $x_e = x_{\rm M}(x_e)$, and we assume this solution to be unique. We further assume $x_{\rm M}(x)$ to be a monotonically decreasing function; thus, $x < x_{\rm M}(x)$ for $x < x_e$. Then:

- for $y < x_e$:

$$\left(\frac{d\sigma_{\mathbb{H}}}{dy}\right)_{\rm naive} = a\frac{R(y)}{y}\Delta(y,x_{\rm M}(y)) + a^2\frac{Q(y)}{y}\int_0^y \frac{dx}{x}R(x)\Delta(y,x_{\rm M}(x))\;; \tag{3.28}$$

- for $y > x_e$:

$$\left(\frac{d\sigma_{\mathbb{H}}}{dy}\right)_{\rm naive} = a\frac{R(y)}{y} + a^2\frac{Q(y)}{y}\int_0^{x_{\rm M}^{-1}(y)} \frac{dx}{x}R(x)\Delta(y,x_{\rm M}(x))\,. \tag{3.29}$$

Expanding in a, we find

$$\left(\frac{d\sigma}{dy}\right)_{\rm naive} \equiv \left(\frac{d\sigma_{\mathbb{S}}}{dy}\right)_{\rm naive} + \left(\frac{d\sigma_{\mathbb{H}}}{dy}\right)_{\rm naive} = \frac{a}{y}[R(y) + BQ(y)] + \mathcal{O}(a^2)\,. \tag{3.30}$$

Therefore the naive subtraction method suffers from double counting, according to the definition in section 2.3, especially at small y, where $Q(y) \to 1$.

For the modified subtraction method, eq. (3.20), we find instead that both contributions are finite:

$$\left(\frac{d\sigma_{\mathbb{S}}}{dy}\right)_{\rm msub} = a\frac{Q(y)}{y}\Delta(y,1)\left[B + aV + aB\int_0^1 dx\frac{Q(x)-1}{x}\right], \tag{3.31}$$

while

- for $y < x_e$:

$$\begin{aligned}\left(\frac{d\sigma_{\mathbb{H}}}{dy}\right)_{\rm msub} &= \frac{a}{y}[R(y) - BQ(y)]\,\Delta(y,x_{\rm M}(y)) + \\ &+ a^2\frac{Q(y)}{y}\int_0^y \frac{dx}{x}[R(x) - BQ(x)]\Delta(y,x_{\rm M}(x))\,;\end{aligned} \tag{3.32}$$

- for $y > x_e$:

$$\left(\frac{d\sigma_{\mathbb{H}}}{dy}\right)_{\rm msub} = \frac{a}{y}[R(y) - BQ(y)] + a^2\frac{Q(y)}{y}\int_0^{x_{\rm M}^{-1}(y)} \frac{dx}{x}[R(x) - BQ(x)]\Delta(y,x_{\rm M}(x))\,. \tag{3.33}$$

Expanding in a, we now get

$$\left(\frac{d\sigma}{dy}\right)_{\text{msub}} = \frac{aR(y)}{y} + \mathcal{O}(a^2)\,, \tag{3.34}$$

in agreement with the NLO prediction for all values of y. Furthermore, for $y \to 0$ we have the expected resummation of leading logarithms,

$$\left(\frac{d\sigma}{dy}\right)_{\text{msub}} \longrightarrow \left(\frac{aB}{y} + \mathcal{O}(a^2)\right)\Delta(y,1) + \text{PST} \qquad \text{for } y \to 0\,, \tag{3.35}$$

where PST stands for power-suppressed terms, i.e., terms that are not singular for $y \to 0$. We stress that the behaviour of $d\sigma/dy$ in eq. (3.35) is dictated by eq. (3.31), i.e., the contribution of $\mathbb{H}$ events is power suppressed for $y \to 0$, since $(R(y) - BQ(y))/y = \mathcal{O}(1)$ in this limit. We also notice that the functional form in eq. (3.35) is the same as that in eq. (3.24). Thus, the MC@NLO resums large logarithms of y in the same way as does the MC. A general derivation of this result for a wide class of observables is given in appendix B.4.

3.4.2 Inclusive observable

As an "inclusive" observable, we shall consider the fully inclusive distribution of the photon energies; thus, each photon emitted, either at the NLO level or by the MC shower, will contribute to the physical observable, which we shall denote by z. The NLO and MC results (for $z > x_0$) are:

$$\left(\frac{d\sigma}{dz}\right)_{\text{NLO}} = a\frac{R(z)}{z}\,, \tag{3.36}$$

$$\left(\frac{d\sigma}{dz}\right)_{\text{MC}} = aB\frac{Q(z)}{z}\,. \tag{3.37}$$

Eq. (3.36) coincides with eq. (3.23), which is a trivial consequence of the fact that only one emission occurs at NLO. On the other hand, the MC result in eq. (3.37) differs from the exclusive case, eq. (3.24), since all emissions now contribute, not only the first. As a consequence, eq. (3.37) has a logarithmically-divergent integral at $z \to 0$, whereas this does not happen in eq. (3.24).

In order to obtain the predictions of the MC@NLO, we need to find the analogues of eqs. (3.25) and (3.26), in the case of the variable z. We obtain

$$I_{\text{MC}}(z,1) = a\frac{Q(z)}{z}\,, \tag{3.38}$$

$$I_{\text{MC}}(z,x_{\text{M}}(x)) = \delta(z-x) + a\frac{Q(z)}{z}\Theta(x_{\text{M}}(x) - z)\,. \tag{3.39}$$

The term with the δ-function in eq. (3.39) accounts for real emission at the NLO level, which always contributes to the inclusive spectrum when it occurs; the other term is the contribution of the MC shower.

To compute the z distribution for the naive subtraction method, we use eq. (3.38) and eq. (3.39) in eq. (3.19), to obtain

$$\left(\frac{d\sigma}{dz}\right)_{\rm naive} = \frac{a}{z}\left[R(z) + Q(z)\left(\sigma_{\rm tot} - a\int_{x_{\rm M}^{-1}(z)}^{1}\frac{dx}{x}R(x)\right)\right], \tag{3.40}$$

where $\sigma_{\rm tot}$ is the NLO total cross section

$$\sigma_{\rm tot} = B + aV + a\int_0^1 \frac{dx}{x}[R(x) - B]\,. \tag{3.41}$$

As before, we may expand in a to find

$$\left(\frac{d\sigma}{dz}\right)_{\rm naive} = \frac{a}{z}[R(z) + BQ(z)] + \mathcal{O}(a^2)\,, \tag{3.42}$$

which shows again that the naive subtraction method suffers from double counting.

For the modified subtraction method we use eq. (3.38) and eq. (3.39) in eq. (3.20) to find

$$\left(\frac{d\sigma}{dz}\right)_{\rm msub} = \frac{a}{z}\left[R(z) - BQ(z) + Q(z)\left(\sigma_{\rm tot} - a\int_{x_{\rm M}^{-1}(z)}^{1} dx\frac{R(x) - BQ(x)}{x}\right)\right]. \tag{3.43}$$

Expanding in a we now find

$$\left(\frac{d\sigma}{dz}\right)_{\rm msub} = \frac{aR(z)}{z} + \mathcal{O}(a^2)\,; \tag{3.44}$$

thus, as in the case of the observable y, the MC@NLO prediction at $\mathcal{O}(a)$ coincides with the NLO result for all z. The small-z limit of eq. (3.43) is

$$\lim_{z\to 0} z\left(\frac{d\sigma}{dz}\right)_{\rm msub} = a\sigma_{\rm tot}\,. \tag{3.45}$$

From eq. (3.37) we obtain instead

$$\lim_{z\to 0} z\left(\frac{d\sigma}{dz}\right)_{\rm MC} = aB\,. \tag{3.46}$$

Therefore, at $z \to 0$ the MC@NLO result is equal to that of the standard MC, times $\sigma_{\rm tot}/B$, which is the K-factor for the total rate.

3.5 Results

In this section we present toy model results for NLO, MC, and MC@NLO. The following parameters have been used in eqs. (3.1), (3.2) and (3.3):

$$a = 0.3\,, \qquad B = 2\,, \qquad V = 1\,. \tag{3.47}$$

For the real emission distribution in eq. (3.3) we have taken

$$R(x) = B + x\left(1 + \frac{x}{2} + 20x^2\right), \tag{3.48}$$

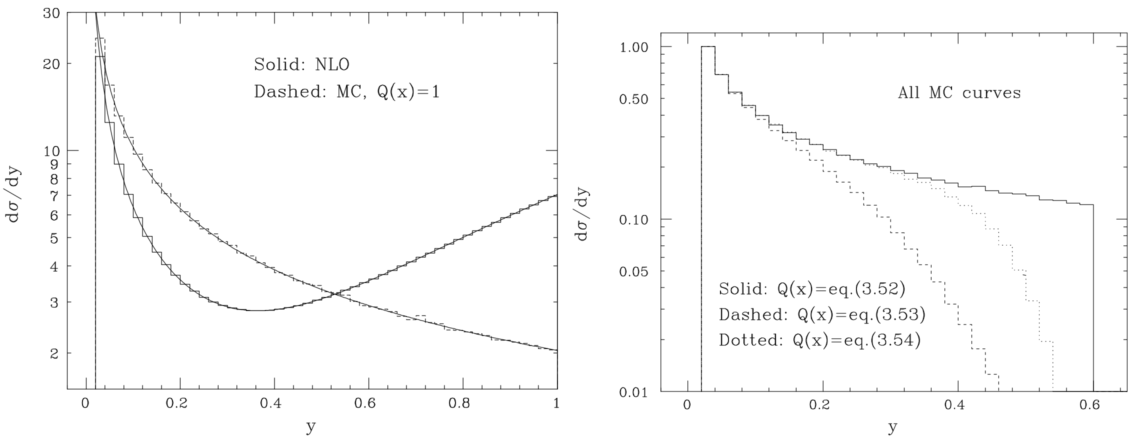

Figure 1: Left panel: Toy model NLO (solid histogram) and MC results (dashed histogram), overlaid on analytical results of eqs. (3.23) and (3.24), with $Q(x) \equiv 1$. Right panel: MC results, in the case of the dead zone with $x_{\rm dead} = 0.6$. Q is chosen as in eq. (3.52) (solid), eq. (3.53) (dashed), and eq. (3.54) (dotted); the first bin is normalized to unity.

so that the total cross section is

$$\sigma_{\rm tot} = B + a\left(V + \frac{95}{12}\right) . \tag{3.49}$$

Note that this implies a large K-factor, $\sigma_{\rm tot}/B = 2.34$. Finally, for the MC cutoff parameter we used

$$x_0 = 0.02\,, \tag{3.50}$$

and for the maximum energy function in eq. (3.17) we chose the form $x_{\rm M}(x) = 1 - x$. In the left panel of figure 1 we show the NLO result obtained with eq. (3.13) for the observable y defined in eq. (3.22) (solid histogram), compared to the result obtained by running the MC (dashed histogram), with $Q(x) \equiv 1$; the integral of the latter result has been normalized to the NLO rate (i.e., each event is given a weight equal to $\sigma_{\rm tot}/N$). As a cross check for the numerical implementation against analytical results, we also plot as solid lines the functions given in eqs. (3.23) and (3.24) (for the latter, B has been substituted with $\sigma_{\rm tot}$, in order to be consistent with the normalization of the dashed histogram), which in fact appear to be in perfect agreement with the corresponding histograms. The shapes of the NLO and MC results are completely different. In fact, the function $R(x)$ has been chosen to introduce a clear difference between the NLO and MC results, for a better understanding of the effect of the matching between NLO and MC.

To study the matching of NLO and MC, we consider an MC with a *dead zone*. In our toy model, this corresponds to an MC that cannot emit photons with energy $x > x_{\rm dead}$ (in the numerical simulations we set $x_{\rm dead} = 0.6$). We can easily get a toy MC with a dead zone by suitably choosing the function Q in eq. (3.14). We first introduce a smoothing function of the form

$$G(x) = \frac{c^2(1-x)^{2\beta}}{x^{2\alpha} + c^2(1-x)^{2\beta}}, \tag{3.51}$$

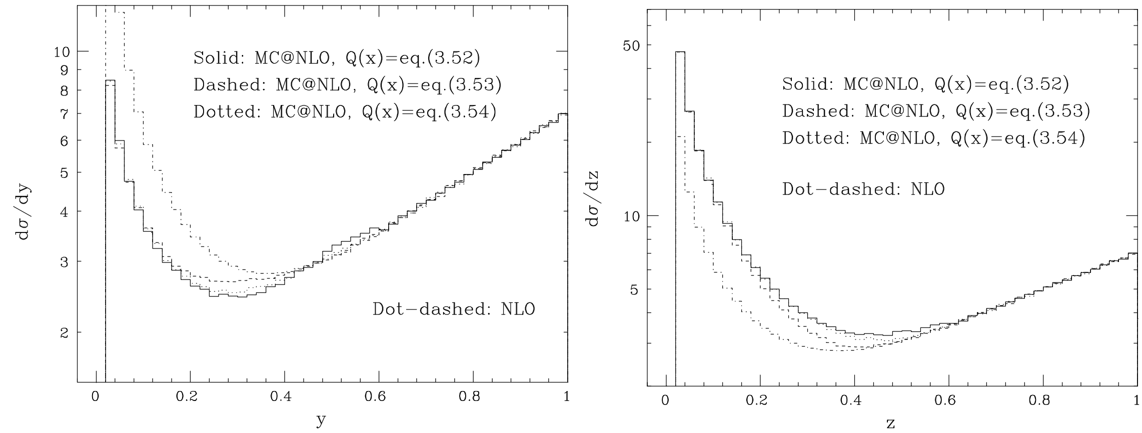

Figure 2: Toy model MC@NLO results; a dead zone is present ($x_{\rm dead} = 0.6$) and Q is as in eq. (3.52) (solid), eq. (3.53) (dashed), or eq. (3.54) (dotted). The results are for the exclusive observable y (left panel), and for the fully inclusive observable z (right panel). The normalization is absolute, and the dot-dashed histogram is the pure NLO result.

where α, β and c are free parameters, and then consider three different implementations of the dead zone:

$$1.\quad Q(x) = \Theta(x_{\rm dead} - x)\,; \tag{3.52}$$

$$2.\quad Q(x) = \Theta(x_{\rm dead} - x)G(x/x_{\rm dead}), \qquad \text{with } \alpha = 1\,,\ \beta = 1\,,\ c = 1\,; \tag{3.53}$$

$$3.\quad Q(x) = \Theta(x_{\rm dead} - x)G(x/x_{\rm dead}), \qquad \text{with } \alpha = 2\,,\ \beta = 1\,,\ c = 8\,. \tag{3.54}$$

The results for the pure MC evolution are presented in the right panel of figure 1; the dependence upon Q is actually fairly big, since we made extreme choices in eqs. (3.52)–(3.54). It is thus striking that the corresponding results obtained with the MC@NLO, shown in the left panel of figure 2, do not display big differences from each other. All three MC@NLO curves are quite close to the pure NLO result (and coincide with it for $y > x_{\rm dead}$), and differ sizably from it only in the region of small x, where the effect of multiple emission is expected to be important.[4] In this region, however, the three MC@NLO curves are close to each other, notwithstanding the differences shown in the right panel of figure 1. We are here in a region where the leading logarithmic behaviour of the MC is highly dominant, and thus the impact of Q (which is subleading in logarithmic accuracy) is moderate. We see that the curve corresponding to eq. (3.52) displays a less regular behaviour than the others at $y \simeq x_{\rm dead}$; however, the effect is very moderate, being at worst of $\mathcal{O}(a^2)$.

The results for the inclusive z distribution are presented in the right panel of figure 2. Again, we can see that the three MC@NLO curves are remarkably close to each other, in spite of the differences in the implementations of the dead zone. The ratio of the cross section in the leftmost bin of the solid histogram over the NLO result for the same quantity is 2.23. This value is fairly close to the K-factor $\sigma_{\rm tot}/B = 2.34$, that we should find according to eqs. (3.45) and (3.46); the small difference is due to fact that this bin is not quite at the limit $z \to 0$.

[4] Since y is an exclusive variable, and the total rate is that of the NLO result, the differences between the MC@NLO and NLO curves at $y > x_0$ are in fact compensated by the region $y < x_0$ in the total rate.

We can thus conclude that MC@NLO reduces the numerical dependence upon subleading logarithms, which is seen when running the MC alone; this is because those regions where the subleading terms are becoming numerically important get sizable contributions from the hard NLO emissions in the MC@NLO, which are not present in the MC, and which eventually become dominant. On the other hand, in the soft-dominated regions, the shapes of the MC@NLO curves just reproduce those of the MC curves, as we expect.

4. The QCD case: W^+W^- production

We now face the problem of constructing an MC@NLO in QCD. In order to simplify as much as possible all the technical details related to the colour structure of the hard process, we consider here the production of W^+W^- pairs in hadronic collisions. This process also has a fairly simple structure of soft and collinear singularities. However, we argue that the method we propose is actually valid for any kind of hard production process.

We shall use HERWIG [1] as our reference parton shower MC. Other choices are of course possible, and our approach is not tailored to work with HERWIG, although some of the technical details, reported in the appendices, do depend on the specific implementation of the hard process and parton showering in that program.

In order to define our MC@NLO for W pair production, we rely on the results previously presented for the toy model. First we emphasize the main points of the toy MC@NLO. Then we make a parallel between the contributions to the toy model cross section, and the contributions to the W^+W^- cross section. Using these results, we proceed to construct the QCD MC@NLO, by analogy with the toy MC@NLO.

4.1 Toy model revisited

Let us consider again our master equation for the toy model MC@NLO, eq. (3.20). The MC@NLO approach consists in running the ordinary MC, with initial conditions and weights determined by the NLO calculation. Depending upon the initial conditions, we distinguish two classes of events, which correspond to $I_{\rm MC}(O, x_{\rm M}(x))$ and $I_{\rm MC}(O, 1)$. We denote the class corresponding to $I_{\rm MC}(O, x_{\rm M}(x))$ by $\mathbb{H}$, which stands for "hard", and reminds us that one emission already occurred, due to NLO, prior to the MC run. The class corresponding to $I_{\rm MC}(O, 1)$ is denoted by $\mathbb{S}$, which stands for "standard"; this reminds us that this initial condition is the same as for the ordinary MC runs.

It is important to realize that for $\mathbb{S}$ events the finite part of the virtual contribution, aV, the real counterterm, $-aB/x$, and the term $aBQ(x)/x$, serve exclusively to fix the normalization of this class of events, relative to that of the class $\mathbb{H}$. In other words, it is irrelevant that the latter two terms explicitly depend upon x: in the MC@NLO, the spectrum of the observable O resulting from $\mathbb{S}$ events is *entirely* due to MC evolution. In contrast, the NLO hard emissions do contribute to the kinematics of $\mathbb{H}$ events, and thus help to determine O.

In order to obtain in QCD a structure analogous to that outlined above, we therefore have to collect all the terms of the NLO cross section that do *not* contribute to hard emissions, and use them to compute the normalization of class $\mathbb{S}$ events; these events result from MC evolutions whose initial conditions are identical to those of an ordinary MC. On the other hand, the kinematics of NLO hard emissions determine the initial conditions for MC evolutions for $\mathbb{H}$ events.

Finally, we have already seen in the case of the observables y and z (see also appendix B.3 for a more formal proof) that at $\mathcal{O}(a)$ the dependence upon $Q(x)$ disappears. Roughly speaking, this happens because the contribution due to the Born term B, in class $\mathbb{S}$ events, cancels at $\mathcal{O}(a)$ with the contribution of the terms $\pm aBQ(x)/x$. We observe that, in the toy model, $aBQ(x)/x$ is the $\mathcal{O}(a)$ term in the expansion of the MC result (see eqs. (3.24) and (3.37)). We therefore guess that the role of $aBQ(x)/x$ in a QCD MC@NLO is played by the $\mathcal{O}(\alpha_{\rm s})$ term in the expansion of the result of the ordinary MC, when only the shower is taken into account (i.e., the hadronization step is skipped). We shall eventually prove (see appendix B.3) that this is in fact the case.

4.2 NLO cross section

We now deal explicitly with $\mathrm{W^+W^-}$ production in hadronic collisions, and we start by writing the NLO cross section. In this paper, we use the results of ref. [22], which we rewrite in a form more suited to our present goal; it should be clear that any $\mathrm{W^+W^-}$ cross section (or at least, one based upon the subtraction method) can be cast in the same form. Owing to the factorization theorem, we have

$$d\sigma(P_1,P_2)=\sum_{ab}\int dx_1dx_2 f_a^{(H_1)}(x_1)f_b^{(H_2)}(x_2)d\hat{\sigma}_{ab}(x_1P_1,x_2P_2)\,, \tag{4.1}$$

for the collision of hadrons H_1 and H_2 with momenta P_1 and P_2 respectively. The sum runs over those parton flavours that give partonic subprocesses contributing to W pair production at NLO. We symbolically denote these processes as follows:

$$q\bar{q}\longrightarrow W^+W^-\,, \tag{4.2}$$
$$q\bar{q}\longrightarrow W^+W^-g\,, \tag{4.3}$$
$$qg\longrightarrow W^+W^-q\,, \tag{4.4}$$

where $q=u,\,d,\ldots$ and eq. (4.4) includes the corresponding antiquark processes. Equation (4.2) contributes to Born and virtual emission terms, whereas eqs. (4.3) and (4.4) contribute to real emission terms.

The information on the hard process in eq. (4.1) is contained in the subtracted partonic cross sections, $d\hat{\sigma}_{ab}$ (the hat means that counterterms are implicit in the expression). We write these cross sections as follows (eq. (3.45) of ref. [22]):

$$d\hat{\sigma}_{ab}=d\sigma_{ab}^{(b)}+d\sigma_{ab}^{(sv)}+d\hat{\sigma}_{ab}^{(c+)}+d\hat{\sigma}_{ab}^{(c-)}+d\hat{\sigma}_{ab}^{(f)}\,, \tag{4.5}$$

where:

a) $d\sigma_{ab}^{(b)}$ is the Born contribution. It has $2\to 2$ kinematics, and corresponds to eq. (3.1) of the toy model.

b) $d\sigma_{ab}^{(sv)}$ is the contribution of the virtual diagrams, and of all those terms, resulting from various subtractions, that have $2 \to 2$ kinematics and are of $\mathcal{O}(\alpha_s)$. This term corresponds to the finite part of eq. (3.2) of the toy model.

c) $d\hat{\sigma}_{ab}^{(c\pm)}$ are the contributions of the finite remainders resulting from the subtraction of the initial-state collinear divergences ($+$ and $-$ indicate radiation from partons a and b respectively). These terms do not have an analogue in the toy model, since in the toy model there are no parton densities. They have $2 \to 3$ kinematics, with the light outgoing parton always collinear to the beam-line; we refer to these configurations as quasi-$2 \to 2$ kinematics. We stress that $d\hat{\sigma}_{ab}^{(c\pm)}$ are subtracted quantities, and that the counterterms have $2 \to 2$ kinematics.

d) $d\hat{\sigma}_{ab}^{(f)}$ is the contribution of the real emission diagrams, with its divergences suitably subtracted; it has $2 \to 3$ kinematics (the counterterms have $2 \to 2$ or quasi-$2 \to 2$ kinematics). The unsubtracted part of this term corresponds to eq. (3.3) of the toy model.

We point out that all the terms on the r.h.s. of eq. (4.5) are separately finite. Also, they do not contain any large logarithms, and so no cancellation of large numbers occurs among them. This property is a feature of the subtraction method, and does not hold in the slicing method.

Items **a)**–**d)** above achieve the identification between terms contributing to the QCD cross section and to the toy model cross section. We now go into more detail, by writing the quantities appearing in eq. (4.5) more explicitly .

We denote by $d\phi_2(s)$ and $d\phi_3(s)$ the phase spaces of the W^+W^- pair and of the W^+W^- pair plus a light parton, respectively; they depend upon the kinematical variables of these particles, and upon $s = (p_1 + p_2)^2$, the c.m. energy squared of the partons that initiate the hard scattering, whose momenta we denote by $p_1 = x_1 P_1$ and $p_2 = x_2 P_2$. The Born term is

$$d\sigma_{ab}^{(b)}(p_1, p_2) = \mathcal{M}_{ab}^{(b)}(p_1, p_2) d\phi_2(s) \,, \tag{4.6}$$

where $\mathcal{M}_{ab}^{(b)}$ is the Born matrix element squared, summed over spin and colour degrees of freedom (the average factor for initial-state ones is understood), times the flux factor $1/(2s)$. The soft-virtual contribution is

$$d\sigma_{ab}^{(sv)}(p_1, p_2) = \left[\mathcal{Q}_{ab}\, \mathcal{M}_{ab}^{(b)}(p_1, p_2) + \mathcal{M}_{ab}^{(v,\mathrm{reg})}(p_1, p_2) \right] d\phi_2(s) \,, \tag{4.7}$$

where $\mathcal{Q}_{ab}$ is a constant with respect to the kinematics of the partonic process, and $\mathcal{M}_{ab}^{(v,\mathrm{reg})}$ is the finite part of the virtual contribution. As with any other finite part resulting from the cancellation of singularities, its definition is to a certain extent arbitrary; however, any redefinition is compensated by suitable changes in the other terms contributing to the cross section. The initial-state collinear terms are:

$$d\hat{\sigma}_{ab}^{(c\pm)} = d\sigma_{ab}^{(c\pm)}\Big|_{\mathrm{ev}} - d\sigma_{ab}^{(c\pm)}\Big|_{\mathrm{ct}} \,, \tag{4.8}$$

where

$$d\sigma_{ab}^{(c+)}(p_1,p_2)\Big|_{\rm ev} = \frac{\alpha_{\rm S}}{2\pi}\left[\log\frac{s(1-x)^2}{\mu^2}P_{ca}^{(0)}(x) - P_{ca}^{(1)}(x)\right]\mathcal{M}_{cb}^{(b)}(xp_1,p_2)d\phi_2(xs)dx\,, \tag{4.9}$$

$$d\sigma_{ab}^{(c-)}(p_1,p_2)\Big|_{\rm ev} = \frac{\alpha_{\rm S}}{2\pi}\left[\log\frac{s(1-x)^2}{\mu^2}P_{cb}^{(0)}(x) - P_{cb}^{(1)}(x)\right]\mathcal{M}_{ac}^{(b)}(p_1,xp_2)d\phi_2(xs)dx\,. \tag{4.10}$$

In eq. (4.8), the term $d\sigma_{ab}^{(c\pm)}|_{\rm ct}$ (the "ct" stands for counterterm) is constructed in order to cancel the divergences present in the collinear "event" contribution $d\sigma_{ab}^{(c\pm)}|_{\rm ev}$; these divergences are soft ($x \to 1$), as can be seen explicitly in eqs. (4.9) and (4.10). The counterterm has a form that is fully specified in a given subtraction scheme (see appendix A). In eqs. (4.9) and (4.10), μ is the factorization mass scale (which also enters in the parton densities), and $P_{ab}^{(i)}$ is the $i^{\rm th}$ term in the ϵ-expansion of the Altarelli-Parisi kernel in $4-2\epsilon$ dimensions, taken for $x < 1$ (i.e., no $\delta(1-x)$ terms are included). Finally, we write the real emission term as

$$d\hat\sigma_{ab}^{(f)} = d\sigma_{ab}^{(f)}\Big|_{\rm ev} - d\sigma_{ab}^{(f)}\Big|_{\rm ct}\,, \tag{4.11}$$

$$d\sigma_{ab}^{(f)}(p_1,p_2)\Big|_{\rm ev} = \mathcal{M}_{ab}^{(r)}(p_1,p_2)d\phi_3(s)\,, \tag{4.12}$$

where $\mathcal{M}_{ab}^{(r)}$ is the matrix element squared resulting from the real emission diagrams (summed over all spins and colours, and averaged over those relevant to incoming partons), times the flux factor. This quantity has soft and initial-state collinear singularities, and thus $d\sigma_{ab}^{(f)}|_{\rm ev}$ has divergences, which are cancelled by the counterterms $d\sigma_{ab}^{(f)}|_{\rm ct}$ in eq. (4.11). As in the case of $d\sigma_{ab}^{(c\pm)}|_{\rm ct}$, these counterterms are fully specified only in a given subtraction scheme. However, one should keep in mind that the physical cross section is always strictly independent of the choice of the subtraction scheme.

4.3 Observables at NLO

As in the case of the toy model, a preliminary step in the construction of the MC@NLO is the prediction of an observable at the NLO. We start by defining an unintegrated, fully-differential cross section

$$d\hat\Sigma_{ab}(x_1,x_2) = f_a^{(H_1)}(x_1)f_b^{(H_2)}(x_2)d\hat\sigma_{ab}(x_1P_1,x_2P_2)\,, \tag{4.13}$$

and analogous quantities, for all the contributions to the partonic cross section that appear on the r.h.s. of eq. (4.5). Any observable O can depend, at the NLO, upon the kinematical variables of the W^+, of the W^-, and of the light outgoing parton. We denote by $O(\mathbf{3})$ the observable computed in non-soft and non-collinear kinematics; this corresponds to $O(x)$ in the toy model. We also denote by $O(\mathbf{2})$ and $O(\tilde{\mathbf{2}})$ the observable computed in $2 \to 2$ (i.e., soft) and in quasi-$2 \to 2$ (i.e., collinear non-soft) kinematics respectively; these correspond to $O(0)$ in the toy model. In all cases, we understand the variables of the outgoing particles to be given in the laboratory frame. We then have

$$
\begin{aligned}
\langle O\rangle = \sum_{ab}\int dx_1 dx_2 d\phi_3 \Bigg[& O(\mathbf{3})\frac{d\Sigma_{ab}^{(f)}}{d\phi_3}\bigg|_{\rm ev} + O(\mathbf{2})\frac{1}{\mathcal{I}_2}\left(\frac{d\Sigma_{ab}^{(b)}}{d\phi_2}+\frac{d\Sigma_{ab}^{(sv)}}{d\phi_2}\right) + \\
& + O(\tilde{\mathbf{2}})\frac{1}{\mathcal{I}_{\tilde{2}}}\left(\frac{d\Sigma_{ab}^{(c+)}}{d\phi_2 dx}\bigg|_{\rm ev}+\frac{d\Sigma_{ab}^{(c-)}}{d\phi_2 dx}\bigg|_{\rm ev}\right) - \\
& - O(\mathbf{2})\frac{1}{\mathcal{I}_{\tilde{2}}}\left(\frac{d\Sigma_{ab}^{(c+)}}{d\phi_2 dx}\bigg|_{\rm ct}+\frac{d\Sigma_{ab}^{(c-)}}{d\phi_2 dx}\bigg|_{\rm ct}\right) - \\
& - \{O(\mathbf{2}),O(\tilde{\mathbf{2}})\}\frac{d\Sigma_{ab}^{(f)}}{d\phi_3}\bigg|_{\rm ct}\Bigg], \qquad (4.14)
\end{aligned}
$$

where in the last term $\{O(\mathbf{2}),O(\tilde{\mathbf{2}})\}$ indicates that the counterterms of the real contribution can have both $2\to 2$ and quasi-$2\to 2$ kinematics. The factors $\mathcal{I}_2$ and $\mathcal{I}_{\tilde{2}}$ guarantee that the $2\to 2$ and quasi-$2\to 2$ terms are correctly normalized upon integration over the full three-body phase space; the following equations must therefore hold

$$
\int d\phi_3(s) = \mathcal{I}_2 \int d\phi_2(s)\,, \qquad (4.15)
$$

$$
\int d\phi_3(s) = \mathcal{I}_{\tilde{2}} \int d\phi_2(xs)dx\,, \qquad (4.16)
$$

the integration being performed over the whole phase space. Eq. (4.14) should be compared with eq. (3.13). The factors $1/\mathcal{I}_2$ and $1/\mathcal{I}_{\tilde{2}}$ are not present in the toy model, simply because the integral over the phase space dx in that case is equal to 1; if this were not the case, a factor analogous to $1/\mathcal{I}_2$ would have to be inserted in eq. (3.13). The term $d\Sigma_{ab}^{(f)}/d\phi_3|_{\rm ev} = f_a f_b \mathcal{M}_{ab}^{(r)}$ corresponds to $aR(x)/x$ in eq. (3.13). The term $d\Sigma_{ab}^{(b)}/d\phi_2 = f_a f_b \mathcal{M}_{ab}^{(b)}$ corresponds to B, the Born contribution, while the term $d\Sigma_{ab}^{(sv)}/d\phi_2 = f_a f_b(\mathcal{Q}_{ab}\mathcal{M}_{ab}^{(b)} + \mathcal{M}_{ab}^{(v,{\rm reg})})$ corresponds to V, the finite part of the virtual contribution. As we said already, the contributions $d\Sigma_{ab}^{(c\pm)}$ have no analogues in the toy model, simply because there are no parton densities there. Finally, $d\Sigma_{ab}^{(f)}/d\phi_3|_{\rm ct}$ represents the counterterms necessary to cancel the divergences of the real matrix element; it is thus equivalent to the last term in eq. (3.13), aB/x.

4.4 Matching NLO and MC: the QCD MC@NLO

Eq. (4.14) is very similar to eq. (3.13). We could try to follow the same steps that brought us from eq. (3.13) to eq. (3.20): replace the observables in eq. (4.14) with suitable interface-to-MC's and then modify the equation with a quantity analogous to that introduced in eq. (3.21). There is, however, a technical problem, related to the initial conditions of class $\mathbb{S}$ events. In the toy model, these initial conditions are unambiguously determined by letting $x\to 0$ in the corresponding system-plus-one-photon configurations. This procedure cannot be automatically generalized to QCD, where matrix elements have soft *and collinear* singularities (corresponding to $\mathbf{2}$ and $\tilde{\mathbf{2}}$ configurations respectively). This may suggest the necessity of considering two $\mathbb{S}$ classes, defined as the sets of those events resulting from

MC evolutions that start from $\mathbf{2}$ and $\tilde{\mathbf{2}}$ initial conditions respectively. However, the MC cannot handle $\tilde{\mathbf{2}}$ initial conditions, since final-state light partons are strictly collinear to the beam-line in these cases.

The solution can be found by observing that, at least as far as the variables of the W^+ and W^- are concerned, $\mathbf{2}$ and $\tilde{\mathbf{2}}$ only differ by a longitudinal boost. The variables of the outgoing light parton are of no concern here, since they cannot contribute to the kinematics of any observables in the configurations $\mathbf{2}$ and $\tilde{\mathbf{2}}$, the parton being either soft of collinear to the beam line. The integration over the momentum fractions x_1 and x_2 in eq. (4.14) implicitly contains such a boost. Roughly speaking, for fixed x_1 and x_2 we have $\mathbf{2} \neq \tilde{\mathbf{2}}$, but given x_1 and x_2, there exist $\tilde{x}_1$ and $\tilde{x}_2$ such that $\mathbf{2}(x_i) = \tilde{\mathbf{2}}(\tilde{x}_i)$. In practice, we adopt the following procedure.[5] We treat each term under the integral sign in eq. (4.14) independently (we assume some regularization procedure has been adopted in the intermediate steps of the computation), performing a change of integration variables that is in general different for each term:

$$(x_1, x_2) \longrightarrow (z_1, z_2), \qquad x_1 = x_1^{(\alpha)}(z_1, z_2, \phi_3), \qquad x_2 = x_2^{(\alpha)}(z_1, z_2, \phi_3), \tag{4.17}$$

where ϕ_3 denotes here the set of independent three-body phase-space variables. The functional forms of these changes of variables will be given in appendix A.4. Here we are only interested in the fact that eq. (4.17) must be such that $\mathbf{2}(x_i) = \tilde{\mathbf{2}}(\tilde{x}_i)$. We define, for each of the functions Σ_{ab} that contribute to eq. (4.14),

$$d\overline{\Sigma}_{ab}^{(\alpha)}(z_1, z_2, \phi_3) = \frac{\partial(x_1^{(\alpha)}, x_2^{(\alpha)})}{\partial(z_1, z_2)} d\Sigma_{ab}^{(\alpha)}\left(x_1^{(\alpha)}(z_1, z_2, \phi_3), x_2^{(\alpha)}(z_1, z_2, \phi_3), \phi_3\right), \tag{4.18}$$

where $\partial(x_1^{(\alpha)}, x_2^{(\alpha)})/\partial(z_1, z_2)$ is the jacobian corresponding to eq. (4.17). Then, as we shall show in appendix A.4, eq. (4.14) can be recast in the following form:

$$\begin{aligned} \langle O \rangle = \sum_{ab} \int dz_1 dz_2 d\phi_3 \Bigg\{ & O(\mathbf{3}) \left. \frac{d\overline{\Sigma}_{ab}^{(f)}}{d\phi_3} \right|_{\mathrm{ev}} + \\ & + O(\mathbf{2}) \left[\frac{1}{\mathcal{I}_2} \left(\frac{d\overline{\Sigma}_{ab}^{(b)}}{d\phi_2} + \frac{d\overline{\Sigma}_{ab}^{(sv)}}{d\phi_2} \right) + \frac{1}{\mathcal{I}_{\tilde{2}}} \left(\left. \frac{d\overline{\Sigma}_{ab}^{(c+)}}{d\phi_2 dx} \right|_{\mathrm{ev}} + \left. \frac{d\overline{\Sigma}_{ab}^{(c-)}}{d\phi_2 dx} \right|_{\mathrm{ev}} \right) - \right. \\ & \left. - \frac{1}{\mathcal{I}_{\tilde{2}}} \left(\left. \frac{d\overline{\Sigma}_{ab}^{(c+)}}{d\phi_2 dx} \right|_{\mathrm{ct}} + \left. \frac{d\overline{\Sigma}_{ab}^{(c-)}}{d\phi_2 dx} \right|_{\mathrm{ct}} \right) - \left. \frac{d\overline{\Sigma}_{ab}^{(f)}}{d\phi_3} \right|_{\mathrm{ct}} \right] \Bigg\} . \end{aligned} \tag{4.19}$$

We stress that the only difference between eq. (4.14) and eq. (4.19) is that the latter does not contain $O(\tilde{\mathbf{2}})$ any longer, thanks to the definition of the change of variables, which ensures that $\mathbf{2}(x_i) = \tilde{\mathbf{2}}(\tilde{x}_i)$. We also stress that no approximation has been made in eq. (4.19); it is just a different way of writing the NLO cross section.

It should be clear that the necessity for the change of variables in eq. (4.17) is due to the simultaneous presence in eq. (4.14) of $O(\mathbf{2})$ and $O(\tilde{\mathbf{2}})$, which is in turn due to the form

[5] To the best of our knowledge, this method was invented in order to improve the numerical stability of the results of ref. [23].

JHEP06(2002)029

of the hadronic cross section in eq. (4.1), where the luminosity is factorized. One could directly start with an NLO cross section that does not contain $O(\tilde{\mathbf{2}})$; clearly, our formulae hold in this case as well, since one just needs to set $(z_1, z_2) \equiv (x_1, x_2)$.

It should also be clear that our goal is to obtain an expression for the NLO cross section that only contains **3** and **2** configurations. Although we find that we can easily achieve this through eq. (4.17), we stress that other methods could be devised, and that our final formula for the MC@NLO will hold, regardless of the procedure adopted to eliminate $\tilde{\mathbf{2}}$ configurations.

The quantities that contribute to $O(\mathbf{3})$ and $O(\mathbf{2})$ in eq. (4.19) correspond to those that contribute to $O(x)$ and $O(0)$ in the toy model, respectively, according to the analogies discussed in section 4.2. The only exceptions are the terms $d\hat{\sigma}_{ab}^{(c\pm)}$, which do not have analogues in the toy model. However, as shown in appendix A.4, these terms must be treated in the same way as the real matrix element collinear counterterms. This is in fact to be expected: $d\hat{\sigma}_{ab}^{(c\pm)}$ are strictly related to the collinear counterterms, being finite remainders resulting from the subtraction of the collinear singularities of the real matrix elements.

Equation (4.19) is therefore strictly analogous to eq. (3.13). This allows us to extend to the QCD case the concept of the event classes $\mathbb{H}$ and $\mathbb{S}$. In turn, this implies that we need two interface-to-MC's, as in the toy model, which we denote by $I_{\rm MC}(O, \mathbf{3})$ and $I_{\rm MC}(O, \mathbf{2})$. These are the analogues of $I_{\rm MC}(O, x_{\rm M}(x))$ and $I_{\rm MC}(O, 1)$ in the toy model. We can thus turn eq. (4.19) into a prescription for a MC@NLO, by replacing $O(\mathbf{3})$ by $I_{\rm MC}(O, \mathbf{3})$, and $O(\mathbf{2})$ by $I_{\rm MC}(O, \mathbf{2})$. In this way, we get the analogue of eq. (3.19), that is, a “naive” MC@NLO that suffers from double counting and difficulties in the generation of unweighted events. However, precisely as in the case of the toy model, we can modify the naive prescription in order to get a QCD MC@NLO that implements a modified subtraction, and fulfills the requirements given in section 2.1.

To accomplish this task, we need to find the analogue of the quantity $aBQ(x)/x$ of the toy model. As discussed in section 4.1, what we are looking for is the $\mathcal{O}(\alpha_{\rm S})$ term in the expansion of the result obtained by running the ordinary MC; we denote this quantity by

$$\left.\frac{d\overline{\Sigma}_{ab}}{d\phi_3}\right|_{\rm MC}. \tag{4.20}$$

The specific form of the quantity in eq. (4.20) is dependent upon the implementation; different MC's will induce different forms of $d\overline{\Sigma}_{ab}|_{\rm MC}$. On the other hand, $d\overline{\Sigma}_{ab}|_{\rm MC}$ is independent of the procedure through which we eliminated $\tilde{\mathbf{2}}$ configurations. The form we use in the present paper, and all the technical details concerning its determination, are reported in appendix A.5. Here, we merely need to note that $d\overline{\Sigma}_{ab}|_{\rm MC}$ must have the same leading singular behaviour as the matrix element contribution $d\overline{\Sigma}_{ab}^{(f)}$ when approaching soft or collinear configurations, in order for the former quantity to be used as a local counterterm to the latter. As we shall discuss in appendix A.5, this fact is not trivial, and we shall need to modify the $\mathcal{O}(\alpha_{\rm S})$ result obtained from the MC. For the time being, we treat $d\overline{\Sigma}_{ab}|_{\rm MC}$ in the same way as $aBQ(x)/x$ is treated in the toy model (see eq. (3.21)):

we subtract and add the quantity

$$I_{\rm MC}(O,\mathbf{n})\,\frac{d\overline{\Sigma}_{ab}}{d\phi_3}\bigg|_{\rm MC} \tag{4.21}$$

under the integral sign, setting $\mathbf{n}=\mathbf{3}$ in the negative term and $\mathbf{n}=\mathbf{2}$ in the positive one. This gives a modified subtraction formula analogous to that in eq. (3.20):

$$\begin{aligned}\frac{d\sigma}{dO} = \sum_{ab}\int dz_1dz_2d\phi_3\Bigg\{ & I_{\rm MC}(O,\mathbf{3})\left(\frac{d\overline{\Sigma}_{ab}^{(f)}}{d\phi_3}\bigg|_{\rm ev}-\frac{d\overline{\Sigma}_{ab}}{d\phi_3}\bigg|_{\rm MC}\right)+\\ & +I_{\rm MC}(O,\mathbf{2})\left[-\frac{d\overline{\Sigma}_{ab}^{(f)}}{d\phi_3}\bigg|_{\rm ct}+\frac{d\overline{\Sigma}_{ab}}{d\phi_3}\bigg|_{\rm MC}+\frac{1}{\mathcal{I}_2}\left(\frac{d\overline{\Sigma}_{ab}^{(b)}}{d\phi_2}+\frac{d\overline{\Sigma}_{ab}^{(sv)}}{d\phi_2}\right)+\right.\\ & +\frac{1}{\mathcal{I}_{\tilde{2}}}\left(\frac{d\overline{\Sigma}_{ab}^{(c+)}}{d\phi_2dx}\bigg|_{\rm ev}+\frac{d\overline{\Sigma}_{ab}^{(c-)}}{d\phi_2dx}\bigg|_{\rm ev}\right)-\\ & \left.-\frac{1}{\mathcal{I}_{\tilde{2}}}\left(\frac{d\overline{\Sigma}_{ab}^{(c+)}}{d\phi_2dx}\bigg|_{\rm ct}+\frac{d\overline{\Sigma}_{ab}^{(c-)}}{d\phi_2dx}\bigg|_{\rm ct}\right)\right]\Bigg\}.\end{aligned} \tag{4.22}$$

Eq. (4.22) is our proposal for a QCD MC@NLO. In order to use this prescription, we need to write it in a form more suited to implementation in a computer code, which will be done in the next section. We stress that, as in the case of the toy model, the dependence upon O in eq. (4.22) is only formal: the MC@NLO generates events without reference to any observable.

4.5 QCD MC@NLO: practical implementation

The straightforward implementation of eq. (4.22) results in a computer code which is not optimally efficient, in particular in the case of unweighted event generation. We therefore implement our MC@NLO through the following procedure.

We start by computing the following integrals:

$$I_{\mathbb{H}} = \sum_{ab}\int dz_1dz_2d\phi_3\left(\frac{d\overline{\Sigma}_{ab}^{(f)}}{d\phi_3}\bigg|_{\rm ev}-\frac{d\overline{\Sigma}_{ab}}{d\phi_3}\bigg|_{\rm MC}\right) \tag{4.23}$$

$$\begin{aligned}I_{\mathbb{S}} = \sum_{ab}\int dz_1dz_2d\phi_3&\left[-\frac{d\overline{\Sigma}_{ab}^{(f)}}{d\phi_3}\bigg|_{\rm ct}+\frac{d\overline{\Sigma}_{ab}}{d\phi_3}\bigg|_{\rm MC}+\frac{1}{\mathcal{I}_2}\left(\frac{d\overline{\Sigma}_{ab}^{(b)}}{d\phi_2}+\frac{d\overline{\Sigma}_{ab}^{(sv)}}{d\phi_2}\right)+\right.\\ &\left.+\frac{1}{\mathcal{I}_{\tilde{2}}}\left(\frac{d\overline{\Sigma}_{ab}^{(c+)}}{d\phi_2dx}\bigg|_{\rm ev}+\frac{d\overline{\Sigma}_{ab}^{(c-)}}{d\phi_2dx}\bigg|_{\rm ev}\right)-\frac{1}{\mathcal{I}_{\tilde{2}}}\left(\frac{d\overline{\Sigma}_{ab}^{(c+)}}{d\phi_2dx}\bigg|_{\rm ct}+\frac{d\overline{\Sigma}_{ab}^{(c-)}}{d\phi_2dx}\bigg|_{\rm ct}\right)\right].\end{aligned} \tag{4.24}$$

Note that these two integrals are finite, owing to the properties of $d\overline{\Sigma}_{ab}|_{\rm MC}$. Clearly,

$$\sigma_{\rm tot} = I_{\mathbb{S}} + I_{\mathbb{H}}\,, \tag{4.25}$$

where $\sigma_{\rm tot}$ is the total rate at the NLO level. We also compute the integrals $J_{\mathbb{H}}$ and $J_{\mathbb{S}}$, defined as in eqs. (4.23) and (4.24) respectively, except that the integrands are taken in absolute value (so that, if the integrands were positive-definite, we would have $I_i = J_i$). We use a modified version of the adaptive integration routine BASES [24], which allows one to compute at the same time the integral of a given function, and the integral of the absolute value of that function. BASES is an evolution of the original VEGAS package [25].

We then make use of the routine SPRING, which is part of the BASES package. Roughly speaking, for any positive-definite function $f(\vec{x})$, SPRING returns a set of values $\{\vec{x}_i\}_{i=1}^n$ distributed according to the function f; the number n must be given as an input. Thus, SPRING effectively achieves the generation of n unweighted events (with weights equal to 1). In our case, the absolute values of the integrands in eqs. (4.23) and (4.24) are given as input functions to SPRING. The numbers of events generated in the two cases are (possibly with a roundoff to the next integer)

$$N_{\mathbb{H}} = N_{\rm tot}\frac{J_{\mathbb{H}}}{J_{\mathbb{S}}+J_{\mathbb{H}}}\,,\qquad N_{\mathbb{S}} = N_{\rm tot}\frac{J_{\mathbb{S}}}{J_{\mathbb{S}}+J_{\mathbb{H}}}\,, \tag{4.26}$$

where $N_{\rm tot}$ is the total number of MC@NLO events that we eventually want to generate. It is easy, within SPRING, to keep track of the sign of the integrand function, before its absolute value is taken. Thus, SPRING achieves unweighted event generation, with the weights equal to $+1$ or -1.

We symbolically denote the results of the SPRING runs as follows:

$$I_{\mathbb{H}} \quad \stackrel{\rm SPRING}{\longrightarrow} \quad \{\mathbf{3}_i, \mathbf{f}_i, w_i^{(\mathbb{H})}\}_{i=1}^{N_{\mathbb{H}}}\,, \tag{4.27}$$

$$I_{\mathbb{S}} \quad \stackrel{\rm SPRING}{\longrightarrow} \quad \{\mathbf{2}_i, \mathbf{f}_i, w_i^{(\mathbb{S})}\}_{i=1}^{N_{\mathbb{S}}}\,; \tag{4.28}$$

thus, eqs. (4.27) and (4.28) correspond to events belonging to classes $\mathbb{H}$ and $\mathbb{S}$ respectively. For a given i, $\{\mathbf{3}_i, \mathbf{f}_i, w_i^{(\mathbb{H})}\}$ gives the complete description of a three-body event; $\mathbf{3}_i$ represents the kinematics (in the lab frame), $\mathbf{f}_i = \{a_i, b_i, c_i\}$ are the flavours of the initial- (a_i, b_i) and final-state (c_i) partons, and $w_i^{(\mathbb{H})} = \pm 1$ is the weight of the event returned by SPRING. The event $\{\mathbf{2}_i, \mathbf{f}_i, w_i^{(\mathbb{S})}\}$ is analogous, except that it has $2 \to 2$ kinematics (and thus $\mathbf{f}_i = \{a_i, b_i\}$).

The events in eqs. (4.27) and (4.28) are the initial conditions for MC evolution. In the case of $\mathbb{H}$ events, eq. (4.27), the kinematics $\mathbf{3}_i$ is determined by (z_1, z_2, ϕ_3). The choice of the flavours $\mathbf{f}_i$ is more involved. One possibility is to fix the pair of flavours (a, b) that appears in the sum in eq. (4.23) (and thus to fix $\mathbf{f}$, since c in this case is unambiguously determined by a and b), to compute the integral, and to repeat the computation for all possible flavour pairs. This procedure is however inconvenient from the point of view of computing time. A faster approach is simply to compute three integrals of the type in eq. (4.23), for the cases $ab = q\bar{q}$, qg and $\bar{q}g$ summed over quark flavours. For each of these integrals, the quark flavours are then determined by statistical methods; this can give a wrong assignment of flavours on event-by-event basis, but guarantees that the "distribution" in the flavours is returned correctly. The reason for not computing a single integral is that subprocesses with $q\bar{q}$, qg and $\bar{q}g$ initial states often give contributions of different signs to the cross section, which is a source of errors when statistical methods are used.

In the case of $\mathbb{S}$ events, eq. (4.28), the kinematic configuration $\mathbf{2}_i$ is unambiguously defined as the soft limit of the corresponding $2 \to 3$ kinematics, determined by (z_1, z_2, ϕ_3).

JHEP06(2002)029

For the flavour assignment, we proceed similarly to what was done in the case of $\mathbb{H}$ events. However, here the relevant flavours are those of the partons entering the $2 \to 2$ hard process which factorizes in the case of soft or collinear emissions. Thus, we use the Born contribution $d\overline{\Sigma}_{ab}^{(b)}$ to assign the flavour. Notice that this can result in errors of $\mathcal{O}(\alpha_{\rm S})$ in flavour assignment. However, the effects induced in this way are $\mathcal{O}(\alpha_{\rm S}^2)$, and thus beyond control. On the other hand, with this prescription one achieves at $\mathcal{O}(\alpha_{\rm S})$ the cancellation of the MC terms $d\overline{\Sigma}_{ab}|_{\rm MC}$ introduced in the modified subtraction, i.e., one avoids double counting. More details on this point are given in appendix B.3. It should be noticed that, since the Born contribution is always positive, there is no need to compute the integral in eq. (4.24) separately for $q\bar{q}$, qg and $\bar{q}g$ initial states.

The above statistical method has the virtue of being easy to generalize to other processes. We point out that eqs. (4.25) and (4.26), together with those in the remainder of this section, are trivially extended to the case in which more than one integral is computed for $\mathbb{H}$ and/or $\mathbb{S}$ events.

The weights in eqs. (4.27) and (4.28) have the following properties:

$$\frac{J_{\mathbb{H}}}{N_{\mathbb{H}}}\sum_{i=1}^{N_{\mathbb{H}}} w_i^{(\mathbb{H})} = I_{\mathbb{H}}\,, \qquad \frac{J_{\mathbb{S}}}{N_{\mathbb{S}}}\sum_{i=1}^{N_{\mathbb{S}}} w_i^{(\mathbb{S})} = I_{\mathbb{S}}\,, \tag{4.29}$$

and thus, using eq. (4.26),

$$\sigma_{\rm tot} = \frac{J_{\mathbb{H}} + J_{\mathbb{S}}}{N_{\rm tot}}\left(\sum_{i=1}^{N_{\mathbb{H}}} w_i^{(\mathbb{H})} + \sum_{i=1}^{N_{\mathbb{S}}} w_i^{(\mathbb{S})}\right). \tag{4.30}$$

We then define the set of events

$$\{\mathbf{k}_i, \mathbf{f}_i, w_i\}_{i=1}^{N_{\rm tot}} = \left\{\mathbf{3}_i, \mathbf{f}_i, w_i^{(\mathbb{H})}\right\}_{i=1}^{N_{\mathbb{H}}} \bigcup \left\{\mathbf{2}_i, \mathbf{f}_i, w_i^{(\mathbb{S})}\right\}_{i=1}^{N_{\mathbb{S}}}\,, \tag{4.31}$$

where $\mathbf{k}_i$ denotes either $\mathbf{3}_i$ or $\mathbf{2}_i$, and

$$w_i = w_i^{(\mathbb{S},\mathbb{H})}\frac{J_{\mathbb{H}} + J_{\mathbb{S}}}{N_{\rm tot}} \qquad \Rightarrow \qquad \sum_{i=1}^{N_{\rm tot}} w_i = \sigma_{\rm tot}\,. \tag{4.32}$$

We constructed a Fortran code whose output is the set of events $\{\mathbf{k}_i, \mathbf{f}_i, w_i\}_{i=1}^{N_{\rm tot}}$. We point out that the merging of the two event sets on the r.h.s. of eq. (4.31) is done randomly; therefore, the probability of finding a $\mathbb{S}$ ($\mathbb{H}$) event at an arbitrary place in the event list is a constant, equal to $N_{\mathbb{S}}/N_{\rm tot}$ ($N_{\mathbb{H}}/N_{\rm tot}$). This set of events is passed to HERWIG; there, it replaces the routine that generates the hard process. We stress that HERWIG has not been modified; we just added a routine that reads the parton configurations coming from the NLO code, and translates them in the standard HERWIG form. Each configuration generates a shower, and all the other operations requested by the user. Notice that the unweighting procedure is entirely dealt with by the NLO code; HERWIG reads the weights provided by the NLO code, and does not discard any events (apart from the few killed in the ordinary shower/hadronization process). According to eq. (4.32), in order to get the total cross section right, exactly $N_{\rm tot}$ events must be processed by HERWIG . However, should only $N_{ev} < N_{\rm tot}$ events be processed, the correct rate will be recovered by multiplying all the physical results by a factor of $N_{\rm tot}/N_{ev}$, since all weights have the same absolute value. In other words, any subset of the event set $\{\mathbf{k}_i, \mathbf{f}_i, w_i\}_{i=1}^{N_{\rm tot}}$ can act as an event set itself, provided that the weights are multiplied by a common factor, in order for eq. (4.32) to hold.

The presence of negative-weight events is a distinctive feature of MC@NLO's. Negative weights imply larger statistical errors than in the case of ordinary MC's, when the same total number of events is generated. In the $\mathbb{H}$ and $\mathbb{S}$ samples of eqs. (4.27) and (4.28), the fraction of negative-weight events is

$$f_i = \frac{1}{2}\left(1 - \frac{I_i}{J_i}\right), \tag{4.33}$$

for $i = \mathbb{S}, \mathbb{H}$; clearly, the closer J_i is to I_i, the smaller is f_i. We can in fact reduce the number of negative-weight events in our MC@NLO by varying $J_{\mathbb{S}}$. This quantity depends upon the form of the counterevents, or, in a given subtraction scheme, upon the free parameters that define the subtraction.[6] Notice that this is not in contradiction with the properties of the subtraction method, since $J_{\mathbb{S}}$ is not a physical quantity. Thus, we can adjust these free parameters in order to minimize the difference $J_{\mathbb{S}} - I_{\mathbb{S}}$; the solution can only be found through numerical methods, because of the complexity of the integrand appearing in $J_{\mathbb{S}}$. Further details on this point are given in appendix A.3.

We point out that the integral $J_{\mathbb{H}}$ does not depend upon any free parameters; thus, the number of negative-weight $\mathbb{H}$ events is a constant for fixed input (physical) parameters in eq. (4.22). We remind the reader that these events will mainly populate the hard emission region.

We remark finally that, owing to the procedure described in section 4.4, the $2 \to 2$ kinematic configurations resulting from the integral in eq. (4.24) can all be obtained as the soft limit of the $2 \to 3$ kinematic configurations identified by (z_1, z_2, ϕ_3). This suggests writing the integration measure in eq. (4.24) with the two-body phase space explicitly factorized, $d\phi_3 = d\phi_2 d\tilde{\phi}_3$, and integrating over $d\tilde{\phi}_3$ before passing the integrand function to SPRING. It should be clear that this intermediate integration step does not affect our capability of obtaining fully-exclusive final states; however, it does have an impact on the number of negative-weight events produced by the MC@NLO. In this paper, we do not exploit this possibility, leaving the option open for future developments.

4.6 Generalities on matched computations

We now consider the case of an observable O whose kinematically-allowed range is $O_s < O < O_h$. We assume that the region $O \simeq O_h$ receives only contributions due to hard emissions, and that large logarithms are present in any fixed-order cross section for $O \simeq O_s$, where a resummation is therefore necessary. In accordance with eq. (4.22), we

[6] These are $\tilde{\rho}$ and ω in the standard subtraction, ζ in the ζ-subtraction: see appendix A.2 and appendix A.3.

write a differential cross section obtained from the MC@NLO as follows:

$$\frac{d\sigma}{dO} = \left(\frac{d\sigma}{dO}\right)_{\rm H} - \left(\frac{d\sigma}{dO}\right)_{\rm RS0} + \left(\frac{d\sigma}{dO}\right)_{\rm RS} . \tag{4.34}$$

The quantity $(d\sigma/dO)_{\rm RS}$ is the contribution of all $\mathbb{S}$ events, i.e. of the MC evolutions $I_{\rm MC}(O, \mathbf{2})$. The quantities $(d\sigma/dO)_{\rm H}$ and $(d\sigma/dO)_{\rm RS0}$ are the contributions of the first and second terms under the integral sign in eq. (4.22) respectively, and thus both are associated with $\mathbb{H}$ events. We expect them to be divergent for $O \to O_s$, only their difference being finite. However, for $O \to O_h$, we can assume that $(d\sigma/dO)_{\rm H}$ and $(d\sigma/dO)_{\rm RS0}$ are separately finite. This is because this region is by hypothesis dominated by hard emission, which can only occur through real-emission diagrams; this prevents the emitted parton over whose variables we integrate in eq. (4.22) from being soft or collinear. The quantities introduced in eq. (4.34) have the following properties:

$$\left|\left(\frac{d\sigma}{dO}\right)_{\rm H} - \left(\frac{d\sigma}{dO}\right)_{\rm RS0}\right| \ll \left(\frac{d\sigma}{dO}\right)_{\rm RS} \qquad \text{for } O \to O_s\,; \tag{4.35}$$

$$\left(\frac{d\sigma}{dO}\right)_{\rm H} = \left(\frac{d\sigma}{dO}\right)_{\rm NLO} + \mathcal{O}(\alpha_{\rm s}^2) \qquad \text{for } O \to O_h\,; \tag{4.36}$$

$$\left(\frac{d\sigma}{dO}\right)_{\rm RS} = \left(\frac{d\sigma}{dO}\right)_{\rm RS0} + \mathcal{O}(\alpha_{\rm s}^2) \qquad \text{for } O \to O_h\,. \tag{4.37}$$

Eqs. (4.36) and (4.37) imply that, for $O \to O_h$, we recover the expected result $d\sigma/dO = (d\sigma/dO)_{\rm NLO}$, which holds up to $\mathcal{O}(\alpha_{\rm s}^2)$ (NNLO) terms. Strictly speaking, eq. (4.37) is not necessary for this to happen; $(d\sigma/dO)_{\rm RS}$ and $(d\sigma/dO)_{\rm RS0}$ include the leading (and possibly next-to-leading) tower of logarithms $\log O/O_h$, or more precisely $\log[(O - O_s)/(O_h - O_s)]$, which become small for $O \to O_h$. However, eq. (4.37) implies that $d\sigma/dO \simeq (d\sigma/dO)_{\rm NLO}$ also for those values of O which are not too close to O_h, but still outside the region where resummation is necessary. The point here is the size of the unknown $\mathcal{O}(\alpha_{\rm s}^2)$ coefficients. There is in fact no guarantee that these coefficients are small; if they are big, the MC@NLO result can differ substantially from the NLO one, although the accuracy of the two computations is the formally of the same order in perturbation theory. This problem is always encountered when a fixed-order computation is matched with a resummed computation. A possible solution (see for example refs. [26, 27]) would be to rewrite eq. (4.34) as follows

$$\frac{d\sigma}{dO} = \left(\frac{d\sigma}{dO}\right)_{\rm H} + \mathcal{D}(O)\left[\left(\frac{d\sigma}{dO}\right)_{\rm RS} - \left(\frac{d\sigma}{dO}\right)_{\rm RS0}\right], \tag{4.38}$$

where $\mathcal{D}(O)$ is a regular function, such that $\mathcal{D}(O_h) \to 0$ for $O \to O_h$, and $\mathcal{D}(O) \to 1$ for $O \to O_s$, up to terms suppressed by powers of $(O - O_s)/(O_h - O_s)$. The perturbative and logarithmic accuracies of eq. (4.38) are the same as those of eq. (4.34); however, the function $\mathcal{D}$ can be suitably chosen in order to suppress the "MC" contribution in the region in which the NLO computation can be trusted. Notice that the presence of $\mathcal{D}(O)$ implies that the total rate predicted by eq. (4.38) differs from that predicted at the NLO level, even when no cuts are applied.

The problem with eq. (4.38) is that the function $\mathcal{D}$ is specific to the observable considered; thus, eq. (4.38) cannot be implemented directly in an MC@NLO. However, one could try to devise a solution, similar to that of eq. (4.38), that works in the context of MC@NLO's. Clearly, in the hard-emission region the function $\mathcal{D}$ suppresses the contribution of $\mathbb{S}$ events, and that part of $\mathbb{H}$ events that corresponds to the $\mathcal{O}(\alpha_{\rm s})$ expansion of the MC result. In order to achieve this result without any reference to a specific observable, a modification of the showering mechanism in the MC is necessary. In particular, one could multiply the Altarelli-Parisi kernel appearing in the Sudakov by a function $Q(\xi, z)$ of the showering variables ξ and z, that tends to 1 in the soft and/or collinear limits, and to zero in the hard emission region. It should be noted that this function is completely analogous to the function $Q(x)$ that we used in the toy MC, eq. (3.14). Exactly as in the case of the toy model, in order to satisfy the requirement that the $\mathcal{O}(\alpha_{\rm s})$ expansion of the MC@NLO result be identical to the NLO result, we would need to perform the formal substitution

$$d\overline{\Sigma}_{ab}\Big|_{\rm MC} \longrightarrow Q(\xi, z) d\overline{\Sigma}_{ab}\Big|_{\rm MC}, \tag{4.39}$$

in eq. (4.22). Since this procedure would modify the integrand in eqs. (4.23) and (4.24), it would also have an impact on the number of negative-weight events generated. In this paper, we shall not pursue this issue further, since we prefer to keep the parton shower MC unmodified.

4.7 Results: $\mathbf{W^+W^-}$ and jet observables

In this section we present results obtained with our $\rm W^+W^-$ MC@NLO, implemented as described in section 4.5. We do not seek to give a phenomenological description that corresponds to a specific experimental configuration; rather, we wish to show the differences between MC@NLO, standard HERWIG MC, and NLO results. For this reason, we also present distributions obtained by integrating over the whole phase space, i.e., in most cases we do not apply any acceptance cuts. In the case in which cuts are applied, we only consider the $\rm W^+$ and $\rm W^-$ variables, and not the variables of their decay products. We also consider jet distributions, the jets being reconstructed by means of the $k_{\rm T}$-clustering algorithm [28]. We show results for pp collisions at $\sqrt{S} = 14$ TeV. We adopt the low-$\alpha_{\rm s}$ set of MRST99 parton densities [29], since this set has a $\Lambda_{\rm QCD}$ value which is rather close to that used as HERWIG default; this value of $\Lambda_{\rm QCD}$ ($\Lambda_5^{\overline{\rm MS}} = 164$ MeV) is used for all our MC@NLO, MC, and NLO runs. In the case of standard HERWIG MC runs, we give each event (we generate unweighted events) a weight equal to $\sigma_{\rm tot}/N_{\rm tot}$, with $N_{\rm tot}$ the total number of events generated. For fully exclusive distributions with no cuts applied, this is equivalent to normalizing HERWIG results to the total NLO rate $\sigma_{\rm tot}$. All the MC@NLO and MC results (but not, of course, the NLO ones) include the hadronization of the partons in the final state.

We present in figure 3 the rapidity of the $\rm W^+$, with (right panel) or without (left panel) a cut on the transverse momentum $p_{\rm T}^{\rm W^+}$. The solid, dashed and dotted histograms show the MC@NLO, MC and NLO results, respectively. This distribution is fairly inclusive, and we expect it to be predicted well by NLO QCD in a wide range. From the figure, we see that the NLO and MC@NLO results are extremely close to each other in the whole range considered, whether or not there is a $p_{\rm T}^{\rm W^+}$ cut. The MC result is also similar, but clearly broader than the other two.

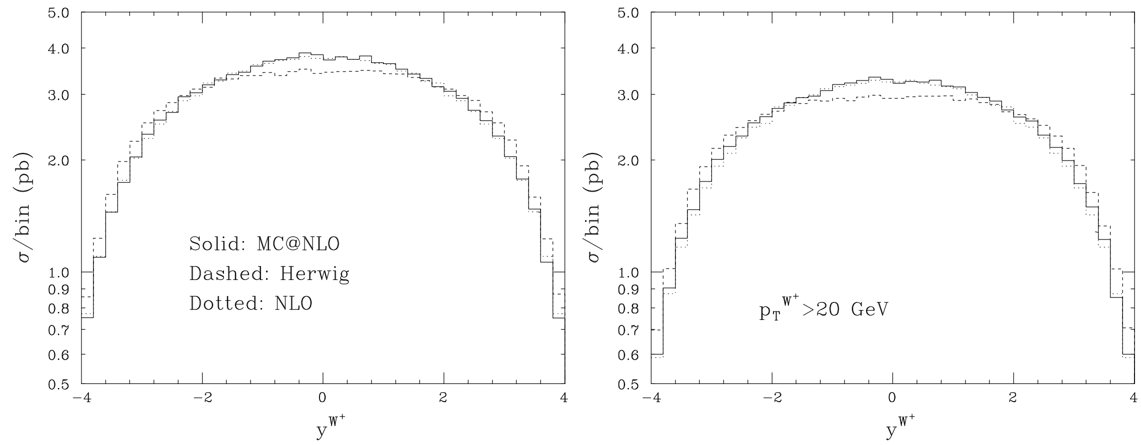

Figure 3: MC@NLO (solid), HERWIG (dashed) and NLO (dotted) results for the rapidity of the W^+. HERWIG results have been normalized as explained in the text.

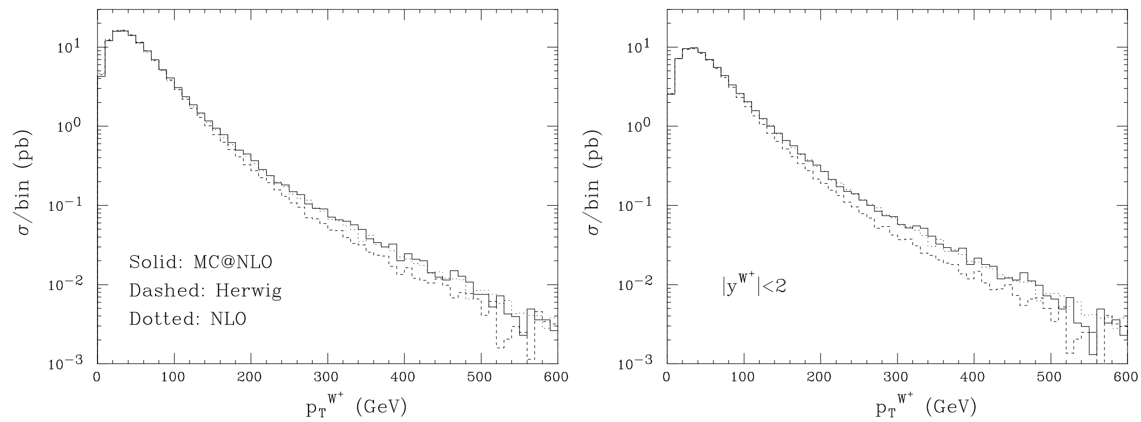

Figure 4: As in figure 3, for the transverse momentum of the W^+.

A similarly inclusive variable is presented in figure 4. The transverse momentum of the W^+ is shown with (right panel) or without (left panel) a cut on the rapidity y^{W^+}. The previous comment applies in this case as well: NLO and MC@NLO basically coincide, whereas the MC result is softer than the other two. The same pattern can for example be found for the invariant mass of the W pair. We conclude that, for this kind of observable, the $\mathcal{O}(\alpha_s^2)$ effects are very small, and NLO and MC@NLO are almost equivalent. This also implies that any possible reshuffling of the momenta, due to the hadronization phase in MC@NLO, has negligible impact on the colourless W's, as we expect.

We now turn to the case of more exclusive quantities, such as correlations between W^+ and W^- variables. In figure 5 we present $p_T^{(WW)}$, the modulus of the vector sum of the transverse momenta of the two W's. NLO computations cannot predict this observable in the region $p_T^{(WW)} \simeq 0$, because of a logarithmic divergence for $p_T^{(WW)} \to 0$; on the other hand, NLO is expected to give reliable predictions at large $p_T^{(WW)}$. The MC behaves in the opposite way; thanks to the cascade emission of soft and collinear partons, it can effectively resum the distribution around $p_T^{(WW)} = 0$; however, its results are not reliable in the large-

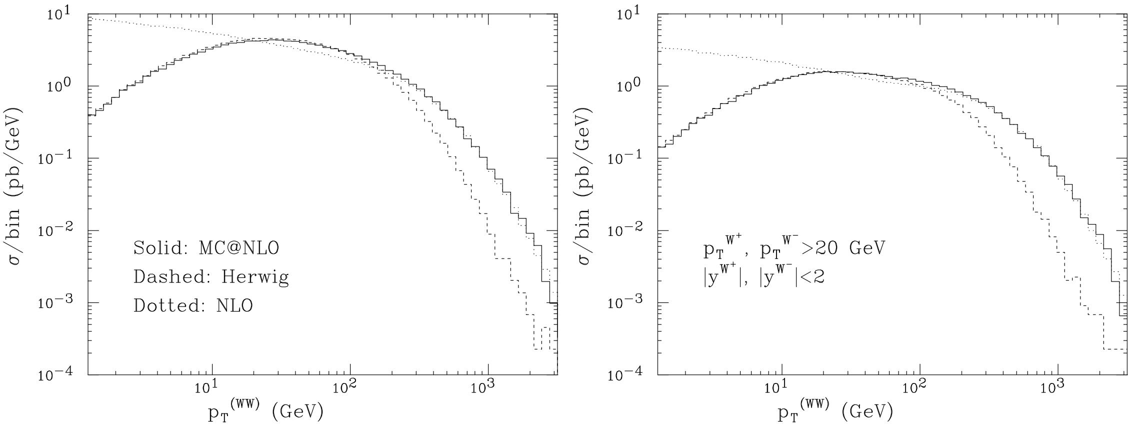

Figure 5: As in figure 3, for the transverse momentum of the W^+W^- pair.

$p_T^{(WW)}$ region, which is mainly populated by events in which a very hard parton recoils against the W pair. Apart from the unreliability of the MC shower approximation to the matrix element in the large-$p_T^{(WW)}$ region, this region cannot really be filled efficiently by the shower; this is easily understood by observing that, at fixed partonic c.m. energy, the large-$p_T^{(WW)}$ region coincides with the dead zone (the border of the dead zone in terms of $p_T^{(WW)}$ can be worked out by using the results of appendix A.5).

This complementary behaviour of NLO and MC approaches can be seen clearly in figure 5, independent of the cuts on the rapidities and transverse momenta of the bosons. In the tail, the NLO cross section is much larger than the MC, simply because hard emissions are correctly treated only in the former. For $p_T^{(WW)} \to 0$, the difference between the two histograms shows the effect of all-order resummation; clearly, no meaningful comparison between NLO and data can be attempted in this region. It is therefore reassuring that the MC@NLO result interpolates the MC and NLO results smoothly. In the small-$p_T^{(WW)}$ region, the shape of the MC@NLO curve is identical to that of the MC result. This is evidence of the fact that MC and MC@NLO resum large logarithms at the same level of accuracy, as argued in appendix B.4. When $p_T^{(WW)}$ grows large, the MC@NLO tends to the NLO result, as expected. Again, the hadronization has no significant impact on the W^+W^- system.

In figure 6 we present the distribution of the difference between the azimuthal scattering angles (i.e., those in the plane transverse to the beam direction) of the W^+ and W^-. This distribution cannot be reliably predicted by fixed-order QCD computations in the region $\Delta\phi^{(WW)} \simeq \pi$; in fact, the NLO prediction diverges logarithmically for $\Delta\phi^{(WW)} \to \pi$. This can be seen in the insets of the plots, where the variable $\Delta\phi^{(WW)}$ has been plotted versus $(\pi - \Delta\phi^{(WW)})/\pi$ on a logarithmic scale, in order to visually enhance the region $\Delta\phi^{(WW)} \simeq \pi$. We can see that in this region the MC@NLO and MC results have identical shapes, as in the case of the observable $p_T^{(WW)}$ near zero discussed above. The other end of the spectrum, i.e. the tail $\Delta\phi^{(WW)} \simeq 0$, might be considered as analogous to the region of large $p_T^{(WW)}$, in that this region is dominated, in an NLO computation, by hard single-parton emission. However, there are important differences from the point of view of MC treatment. As discussed earlier, the region of high $p_T^{(WW)}$ is not filled efficiently by the MC shower. On

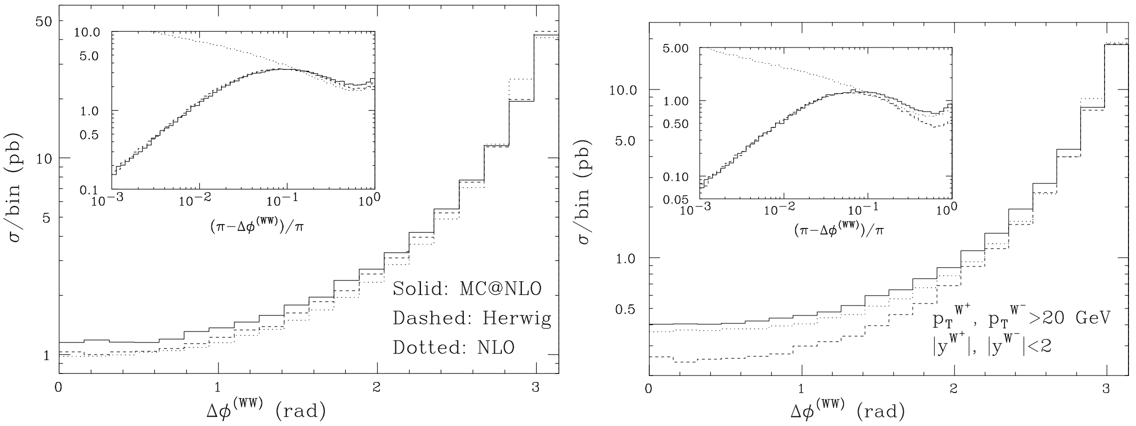

Figure 6: As in figure 3, for the difference in the azimuthal angles of the W^+ and W^-.

the other hand, the $\Delta\phi^{(\mathrm{WW})} \simeq 0$ tail for low $p_{\mathrm{T}}^{(\mathrm{WW})}$ is more easily populated by the MC. This is confirmed in the left-hand panel of figure 6, where we can see that the NLO and MC results are quite close to each other at $\Delta\phi^{(\mathrm{WW})} \simeq 0$ when no cuts are applied. Thus, the region $\Delta\phi^{(\mathrm{WW})} \simeq 0$ receives contributions both from hard parton emissions (which dominate the configurations in which the W's have large transverse momenta), and from multiple soft or collinear parton emissions (when the W's have small transverse momenta). NLO can treat the former but not the latter, and MC can treat the latter but not the former. MC@NLO, on the other hand, is expected to handle both emissions correctly, while still avoiding any double counting. When cuts on the transverse momenta of the W's are applied, the contribution from multiple soft or collinear emissions becomes less important, as may be seen in the right-hand panel of figure 6. Overall, the MC@NLO results are close to the NLO predictions in the $\Delta\phi^{(\mathrm{WW})} \simeq 0$ tail; as we shall show later, the differences are compatible with $\mathcal{O}(\alpha_{\mathrm{S}}^2)$ effects.

We now turn to the discussion of jet observables. In figure 7 we present the transverse energy distribution of the hardest jet of each event (left), and of the inclusive jets (right). In this case, we only compare MC@NLO and MC results; NLO jet results are trivial for this process, since there is only one "jet", which coincides with a parton. In the case of the hardest jet, the same discussion as in the case of $p_{\mathrm{T}}^{(\mathrm{WW})}$ applies; the MC@NLO resums large logarithms at small E_{T} in the same way as the MC does, and can also treat the large-E_{T} region, where the MC fails. In the case of inclusive jets, the spectrum diverges for $E_{\mathrm{T}} \to 0$, because of the increasing number of jets with smaller and smaller E_{T}. In this region, the MC@NLO and MC results (rescaled by the K-factor) coincide. We stress that this is actually the same pattern that we have already discussed in the toy model: see eqs. (3.45) and (3.46).

The NLO results presented so far have been obtained by setting the renormalization (μ_{R}) and factorization (μ_{F}) scales, that we collectively refer to as NLO scales (μ_{NLO}), equal to a reference scale μ_0, defined as follows

$$\mu_0^2 = \frac{1}{2}\left[\left(M_{\mathrm{T}}^{\mathrm{W}^+}\right)^2 + \left(M_{\mathrm{T}}^{\mathrm{W}^-}\right)^2\right], \tag{4.40}$$

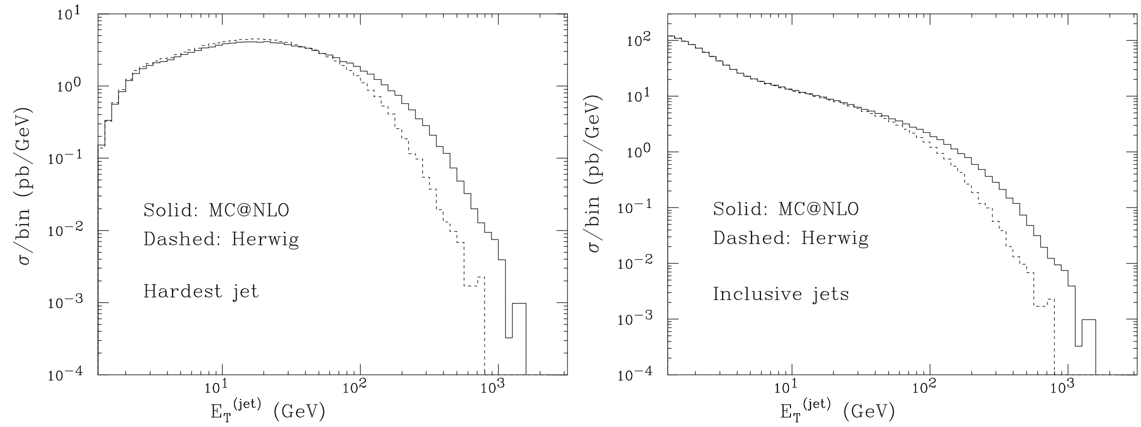

Figure 7: MC@NLO (solid) and HERWIG (dashed) results for the transverse energy of the hardest jet in each event (left panel), and of fully inclusive jets (right panel). HERWIG results have been normalized as explained in the text.

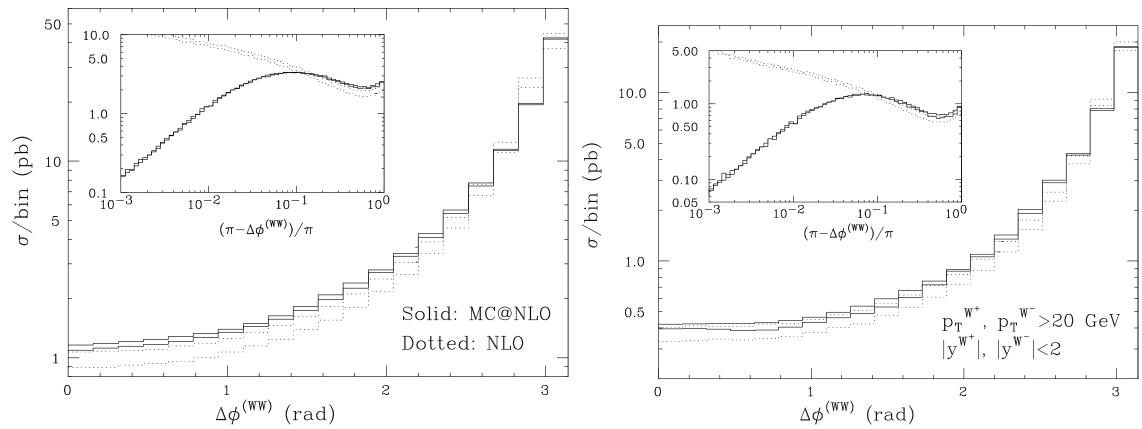

Figure 8: Scale dependence of MC@NLO (solid band) and NLO (dotted band) results, obtained by varying the renormalization and factorization scales in the range $\mu_0/2 < \mu_R, \mu_F < 2\mu_0$. See the text for details.

where $M_T^{W^+}$ and $M_T^{W^-}$ are the transverse masses of the W^+ and W^- respectively. The same scale choice has been adopted to compute the integrals given in eqs. (4.23) and (4.24), that is, to obtain the sets of $\mathbb{H}$ and $\mathbb{S}$ events eventually given to the MC as initial conditions.

In general, we refer to the scales used in the computations of eqs. (4.23) and (4.24) as MC@NLO scales ($\mu_{\text{MC@NLO}}$). We point out that the NLO and MC@NLO scales need *not* coincide. It actually seems more natural to choose different forms for them. The MC@NLO scale should in some way match the scale used in the parton shower; although any mismatch between the two is formally of higher order than we are considering, a choice of the MC@NLO scale motivated by MC considerations should result in smaller coefficients for the $\mathcal{O}(\alpha_S^2)$ terms in the regions dominated by hard emissions. Notice that, when the NLO and MC@NLO scales do not coincide, the total rates predicted by NLO and MC@NLO computations are different.

We shall not pursue the issue of an appropriate MC@NLO scale choice any further in the present paper. We shall limit ourselves to considering the variation of NLO and MC@NLO predictions obtained by setting $\mu_R = \mu_F = \mu_0/2$ and $\mu_R = \mu_F = 2\mu_0$. We stress that, while this procedure is a commonly accepted way to obtain a rough estimate of the unknown higher-order terms in an NLO computation, it cannot be used to guess the MC@NLO scale that best matches the parton shower scale; for this purpose, functional forms other than that shown in eq. (4.40) would need to be considered. Still, the procedure can give an idea of the size of the neglected higher-order terms in the MC@NLO formalism. The results are presented in figure 8 for the azimuthal correlation $\Delta\phi^{(WW)}$; we have chosen this observable since it is the one in which the scale effects are largest. The solid and dotted bands show the variation in the MC@NLO and NLO predictions respectively. We observe a sizable reduction of the scale dependence when going from NLO to MC@NLO results. Also, in the $\Delta\phi^{(WW)} \simeq 0$ tail, the two bands are rather close to each other, and overlap in the case that cuts are applied. This is an indication of the fact that NLO and MC@NLO results indeed differ by $\mathcal{O}(\alpha_S^2)$ terms, and that the difference is less significant when hard cuts are applied, as we should expect.

An issue related to the choice of scale for the parton showers is the setting of the HERWIG parameter `EMSCA`, which enforces an upper limit on the transverse momenta of partons emitted in the showers. The coherence of gluon emission is ensured in HERWIG by using an angular variable to perform the shower evolution. In order to avoid double-counting of hard processes, it is then necessary to veto emissions with transverse momenta greater than the scale of the primary hard process, which is specified by `EMSCA`. If such emissions were allowed, they would constitute the true primary hard process. When the veto is applied, the harder emission does not occur but the evolution variable is reset to the value selected for the vetoed emission. In this way the constraint that the true hard scale is `EMSCA` is correctly taken into account together with QCD coherence in the shower.

In our case, when running the standard HERWIG MC for W^+W^- production, we set `EMSCA`$=\sqrt{s}$. In the case of the MC@NLO, there is no reason to have the same `EMSCA` for events belonging to the classes $\mathbb{H}$ and $\mathbb{S}$. We have already pointed out that the class $\mathbb{S}$ must behave in the same way as the ordinary MC, in order to reproduce (shape-wise) its results in those regions where a resummation is needed. Thus, for $\mathbb{S}$ events, we set should also set `EMSCA`$=\sqrt{s}$. As far as $\mathbb{H}$ events are concerned, we notice that the initial condition for the MC in the toy model can also be interpreted as a veto scale; in the case of toy $\mathbb{H}$ events, the initial condition is $x_M(x)$, which is imposed to be a monotonic function, with $x_M(0) = 1$ (see section 3.3 and section 3.4). A possible choice that reproduces this condition in the real case is `EMSCA`$=\sqrt{s} - 2p_T$, where p_T is the transverse momentum of the final-state light parton, which is always present in $\mathbb{H}$ events. Then, in the soft or collinear limits, this scale tends to that adopted in the case of $\mathbb{S}$ events. However, such a choice implies that if a very hard emission (i.e., $p_T \simeq \sqrt{s}/2$) occurs at the NLO, only a very limited possibility is left for the MC to emit further, which can lead to problems in the hadronization process. Other choices are of course possible. In the toy model, it can be proven that different x_M's only affect logarithms beyond the accuracy of the MC. In the real case, we have verified that other reasonable choices of `EMSCA` do not induce visible differences in any of the observables studied.

JHEP06(2002)029

Our MC@NLO results are obtained without soft matrix element corrections [6], i.e. we set the HERWIG parameter `SOFTME=.FALSE.`. By doing so, we obtain the $\mathcal{O}(\alpha_s)$ HERWIG result reported in appendix A.5, which is crucial in our implementation of MC@NLO to avoid double counting. Furthermore, the purpose of soft matrix element corrections is to ensure smooth matching on the boundary between the regions dominated by parton showers and by hard emissions, whereas in our formalism there is no such hard boundary. In order to have meaningful comparisons with standard MC results, we set `SOFTME=.FALSE.` in that case as well. We verified that the choice of `SOFTME` has in general only a small impact on the MC results for the observables we studied; the only visible effects are in the shapes of distributions in those regions in which a resummation is needed: $p_T^{(WW)} \simeq 0$, $\Delta\phi^{(WW)} \simeq \pi$, and $p_T^{(jet)} \simeq 0$ for the hardest jet of the event. A non-negligible effect ($\simeq 10\%$) is also seen at $\Delta\phi^{(WW)} \simeq 0$, which is partly related to the behaviour at $\Delta\phi^{(WW)} \simeq \pi$, since the total rate is independent of the value of `SOFTME`.

We should mention finally that we performed many studies of the stability of the MC@NLO results. We studied the dependences of the physical results upon the free parameters α and β introduced in eqs. (A.86) and (A.87), and found them completely negligible. This is a rather direct confirmation of the claim in appendix A.5, that the angular distribution of soft partons is irrelevant in the determination of physical observables. We performed test MC@NLO runs with increased values of the showering cutoffs `VQCUT` and `VGCUT`, and found no significant variations with respect to the results presented here; this supports the claim that MC and MC@NLO have the same dependences upon MC showering cutoffs. Also, we find that the parton densities used for MC backward evolution do not strongly influence the MC@NLO results for this production process. The results obtained using low-α_s MRST99 in the MC evolution are statistically compatible with those obtained using MRST-LO (low-α_s MRST99 being used in both cases for the NLO part).

5. Other processes

The extension of MC@NLO to other processes, at least those with unambiguous colour flow, seems straightforward in principle. The key equation, (4.22), is formally valid for any kind of singularities. For final-state collinear singularities, the choice of the Bjorken x's and the flavour assignment are trivial: one just has to give the MC the lower-order process that preceded the collinear splitting.

In processes where there are several possible colour flows, the procedure adopted in the HERWIG MC is first to select the kinematics according to the full differential cross section, including interference between colour flows, and then to assign a colour flow according to the relative contributions to the cross section neglecting interference (i.e. in the large-N_c limit) [30]. This procedure can easily be extended to the MC@NLO case. We point out that the colour flows can be determined independently for $\mathbb{S}$ and $\mathbb{H}$ events. The correct NLO results will still be recovered upon expansion in α_s, since the cancellation of the MC terms (4.21) inserted in the modified subtraction formula, eq. (4.22), occurs on statistical basis, and not event by event. In order to construct the MC-like cross sections (4.20), the technique presented in this paper can be adopted: namely, the pure MC result is smoothly matched onto that predicted by perturbative QCD in the soft limit. In fact, independently of the choice of the showering variables, the former cross section is in general locally different from the latter, since it is only in the former that the radiation function is integrated over an azimuthal angle, in order to implement colour coherence in a MC-friendly manner. For processes involving emission from gluons, there are also azimuthal correlations in the collinear limit, which integrate to zero and are therefore not included in MC's but cannot be neglected locally.

For some processes, such as QCD jet production, the total cross section $\sigma_{\rm tot}$ is not defined without a cut; for example one may require the total amount of transverse energy to be bigger than a given quantity. The formulation of the MC@NLO described in this paper should remain valid in the presence of such a cut.

In processes involving incoming photons, pointlike and hadronic components needs to be defined in a given subtraction scheme. Within the NLO approximation, a change of scheme gives rise to terms that are of higher order. However, for consistency the same scheme should be adopted for the generation of the NLO cross section and the backward evolution of the parton shower in the MC. We also point out that, in order for the procedure described in this paper to be applicable, the photons should have a continuous energy spectrum, such as that from an incoming electron beam, in which the photon energies are distributed according to the Weizsäcker-Williams function.

6. Conclusions and outlook

We have presented a practical proposal for constructing an event generator which, we believe, displays the main advantages of the NLO and MC approaches. For brevity we call this MC@NLO. A key ingredient is the use of a modified subtraction method for dealing with infrared and collinear singularities. The method has been illustrated using a toy model and implemented in full for the process of hadroproduction of W boson pairs. The results there look encouraging (although of course comparisons with data are for the future): all the observables investigated show the expected behaviour in regions where the NLO or MC results should be most reliable, with a smooth transition between hard and soft/collinear regions.

The method generates a fraction of events with negative weights, but in the cases we have studied the fraction is small and unweighted (signed) event generation remains rather efficient. One can reduce the number of negative weights for a given process by tuning the parameters specific to that process (these parameters have nothing to do with the global MC parameters, which are tuned once and for all by comparison to data). We shall consider in future upgrades of the implementation of our method the strategies outlined in section 4.5 and in section 4.6, which should reduce the number of $\mathbb{S}$ and $\mathbb{H}$ negative-weight events respectively, using an intermediate integration step, and a modification of the showering algorithm. One could also consider more exotic possibilities, such as defining "superevents" by combining configurations that are close together in three-body phase space.

JHEP06(2002)029

The implementation presented here is based on the HERWIG MC generator, but there appears no reason why other parton shower generators such as PYTHIA should not be used. The main points to be reconsidered when changing the MC generator will be the dead zones for hard emission, the matching of local counterterms in the soft region, and the relation between the 2- and 3-body configurations connected by the backward evolution in the showering initiated by initial-state partons (corresponding to the event projection that we used in this paper).

Acknowledgments

We wish to thank CERN TH division for hospitality during the preparation of this work. Many informative discussions with S. Catani, G. Corcella, M. Dobbs, J.-Ph. Guillet, Z. Kunszt, M. Mangano, P. Nason, R. Pittau, G. Ridolfi, and M. Seymour are gratefully acknowledged. We also thank S. Catani, J. Collins, and M. Dobbs for comments on the manuscript.

A. W^+W^- cross sections

In this appendix, we give all the details relevant to cross sections at $\mathcal{O}(\alpha_s^0)$ and $\mathcal{O}(\alpha_s)$. Although we study explicitly the case of W^+W^- production, some of the results are based on fairly general ideas, which can be applied to any other hard production process.

A.1 Kinematics

We start by defining some kinematic variables for W^+W^- production. The momenta in the $2 \to 3$ partonic processes are assigned as follows:

$$a(p_1) + b(p_2) \longrightarrow c(k) + W^+W^-(k_{WW}) . \tag{A.1}$$

Following ref. [22], we introduce the quantities

$$x = M_{WW}^2/s , \qquad y = \cos\theta , \tag{A.2}$$

where $s = (p_1 + p_2)^2$, M_{WW} is the invariant mass of the W pair, and θ is the scattering angle of the outgoing light parton in the c.m. frame of the colliding partons. In this frame, $1-x$ is the energy fraction of the outgoing light parton (in units of $\frac{1}{2}\sqrt{s}$). These quantities unambiguously identify the soft and initial-state collinear regions of the three-body phase space. In particular

$$x = 1 \qquad \Rightarrow \qquad \text{soft } (k^0 = 0) , \tag{A.3}$$

$$y = 1 \qquad \Rightarrow \qquad \text{collinear } (\vec{k} \parallel \vec{p}_1) , \tag{A.4}$$

$$y = -1 \qquad \Rightarrow \qquad \text{collinear } (\vec{k} \parallel \vec{p}_2) . \tag{A.5}$$

The ranges allowed for x and y are

$$\rho \le x \le 1 , \qquad -1 \le y \le 1 , \tag{A.6}$$

where $\rho = 4m_{\mathrm{W}}^2/s$. Defining

$$\beta = \sqrt{1-\rho}\,, \qquad \beta_x = \sqrt{1-\rho/x}\,, \tag{A.7}$$

the two-body and three-body phase spaces are

$$\begin{aligned} d\phi_2(s) &= \frac{\beta}{16\pi} d\cos\theta_{\mathrm{W}}\,, & \text{(A.8)}\\ d\phi_2(xs) &= \frac{\beta_x}{16\pi} d\cos\theta_{\mathrm{W}}\,, & \text{(A.9)}\\ d\phi_3(s) &= \frac{s\beta_x}{512\pi^4}(1-x)dx\,dy\,d\cos\theta_{\mathrm{W}}\,d\varphi\,, & \text{(A.10)} \end{aligned}$$

where θ_{W} and φ are the polar and azimuthal angles of the W^+ in the $\mathrm{W}^+\mathrm{W}^-$ c.m. frame. The normalization in eq. (A.10) is such that both θ_{W} and φ range between 0 and π.

A.2 Standard subtraction

Here we give the subtraction terms using the definition of ref. [22]. The quantity $\mathcal{Q}_{ab}$ appearing in eq. (4.7) reads

$$\mathcal{Q}_{q\bar{q}} = \frac{C_{\mathrm{F}}\alpha_{\mathrm{S}}}{4\pi}\left(\log\frac{s}{\mu^2}(6+16\log\tilde{\beta}) + 32\log^2\tilde{\beta} - \frac{4}{3}\pi^2\right), \tag{A.11}$$

and $\mathcal{Q}_{ab} = 0$ for $\{a,b\} \neq \{q,\bar{q}\}$. We have introduced a free parameter $\tilde{\rho}$ such that

$$\rho \le \tilde{\rho} < 1\,, \qquad \tilde{\beta} = \sqrt{1-\tilde{\rho}}\,. \tag{A.12}$$

The collinear terms of eq. (4.8) read

$$\begin{aligned} d\hat{\sigma}_{ab}^{(c+)}(p_1,p_2) = \frac{\alpha_{\mathrm{S}}}{2\pi}\Bigg\{&\left[\log\frac{\omega s}{2\mu^2}\left(\frac{1}{1-x}\right)_{\tilde{\rho}} + 2\left(\frac{\log(1-x)}{1-x}\right)_{\tilde{\rho}}\right](1-x)P_{ca}^{(0)}(x) - \\ &- P_{ca}^{(1)}(x)\Bigg\}\mathcal{M}_{cb}^{(b)}(xp_1,p_2)d\phi_2(xs)dx\,, \end{aligned} \tag{A.13}$$

and analogously for $d\hat{\sigma}_{ab}^{(c-)}$. The real emission contribution of eq. (4.11) reads

$$\begin{aligned} d\hat{\sigma}_{ab}^{(f)}(p_1,p_2) = \frac{1}{2}\left(\frac{1}{1-x}\right)_{\tilde{\rho}}&\left[\left(\frac{1}{1-y}\right)_{\omega} + \left(\frac{1}{1+y}\right)_{\omega}\right]\times \\ &\times(1-x)(1-y^2)\mathcal{M}^{(r)}(p_1,p_2)d\phi_3(s)\,. \end{aligned} \tag{A.14}$$

The distributions introduced in eqs. (A.13) and (A.14) define the subtraction terms as follows:

$$\begin{aligned} \int_\rho^1 dx\,h(x)\left(\frac{1}{1-x}\right)_{\tilde{\rho}} &= \int_\rho^1 dx\,\frac{h(x)-\Theta(x-\tilde{\rho})h(1)}{1-x}\,, & \text{(A.15)}\\ \int_\rho^1 dx\,h(x)\left(\frac{\log(1-x)}{1-x}\right)_{\tilde{\rho}} &= \int_\rho^1 dx\,\frac{h(x)-\Theta(x-\tilde{\rho})h(1)}{1-x}\log(1-x)\,, & \text{(A.16)}\\ \int_{-1}^1 dy\,g(y)\left(\frac{1}{1\pm y}\right)_{\omega} &= \int_{-1}^1 dy\,\frac{g(y)-\Theta(\mp y-1+\omega)g(\mp 1)}{1\pm y}\,, & \text{(A.17)} \end{aligned}$$

where $0 < \omega \leq 2$, and h and g are generic regular functions. It is clear that a variation of the (arbitrary) parameters $\tilde{\rho}$ and ω amounts to a finite redefinition of the soft and collinear counterterms respectively. The cross section is strictly independent of the values of $\tilde{\rho}$ and ω in their whole range (i.e., these parameters do not need to be small, as opposed to the slicing method).

A.3 ζ-subtraction

We give here an alternative form of the subtraction terms, which has been used in our implementation of the MC@NLO. Although the results of the MC@NLO do not depend on the particular form adopted for the subtraction terms, we find that the form introduced in this section gives a superior numerical stability, in a sense that, in the context of our MC@NLO, allows to reduce the number of negative-weight events. See section 4.5 for more details. We define

$$\mathcal{P}(x,y) = (1-x)^2\,(1-y^2)\,, \tag{A.18}$$

where x and y are defined in eq. (A.2). The function $\mathcal{P}$ has a simple physical meaning; in fact

$$p_{\rm T}^{\rm (WW)} = \frac{\sqrt{s}}{2}\sqrt{\mathcal{P}(x,y)}\,, \tag{A.19}$$

where $p_{\rm T}^{\rm (WW)}$ is the transverse momentum of the W pair, which coincides with that of the outgoing light parton at NLO. The ζ-subtraction basically amounts to subtracting the counterterms when the condition

$$\mathcal{P}(x,y) < \zeta \tag{A.20}$$

is met, ζ being a numerical parameter, $0 < \zeta \leq 1$. More precisely, we define

$$\int_\rho^1 dx\, h(x)\left(\frac{1}{1-x}\right)_{\mathcal{P}} = \int_\rho^1 dx\,\frac{h(x)-\Theta(\zeta-\mathcal{P}(x,y))h(1)}{1-x}\,, \tag{A.21}$$

$$\int_\rho^1 dx\, h(x)\left(\frac{F(x)}{1-x}\right)_{\mathcal{P}} = \int_\rho^1 dx\,\frac{h(x)-\Theta(\zeta-\mathcal{P}(x,y))h(1)}{1-x}F(x)\,, \tag{A.22}$$

$$\int_{-1}^1 dy\, g(y)\left(\frac{1}{1\pm y}\right)_{\mathcal{P}} = \int_{-1}^1 dy\,\frac{g(y)-\Theta(\zeta-\mathcal{P}(x,y))\Theta(\mp y)g(\mp 1)}{1\pm y}\,, \tag{A.23}$$

where y in eqs. (A.21), (A.22), and x in eq. (A.23), are held fixed. We have

$$\mathcal{Q}_{q\bar{q}} = \frac{C_{\rm F}\alpha_{\rm S}}{4\pi}\left(\log\frac{s}{\mu^2}(6+16\log\beta)+32\log^2\beta-\frac{4}{3}\pi^2+\mathcal{F}_s\Theta(\bar{y})\right), \tag{A.24}$$

$$\begin{aligned} d\hat{\sigma}_{ab}^{(c+)}(p_1,p_2) = \frac{\alpha_{\rm S}}{2\pi}\Bigg\{&\left[\log\frac{s}{2\mu^2}\left(\frac{1}{1-x}\right)_{\mathcal{P}}+\left(\frac{\log(1-\mathcal{F}_c(x))}{1-x}\right)_{\mathcal{P}}+2\left(\frac{\log(1-x)}{1-x}\right)_{\mathcal{P}}\right]\times \\ &\times(1-x)P_{ca}^{(0)}(x)-P_{ca}^{(1)}(x)\Bigg\}\mathcal{M}_{cb}^{(b)}(xp_1,p_2)d\phi_2(xs)dx\,, \end{aligned} \tag{A.25}$$

$$\begin{aligned} d\hat{\sigma}_{ab}^{(f)}(p_1,p_2) = \frac{1}{2}&\left(\frac{1}{1-x}\right)_{\mathcal{P}}\left[\left(\frac{1}{1-y}\right)_{\mathcal{P}}+\left(\frac{1}{1+y}\right)_{\mathcal{P}}\right]\times \\ &\times(1-x)(1-y^2)\mathcal{M}^{(r)}(p_1,p_2)d\phi_3(s)\,, \end{aligned} \tag{A.26}$$

where

$$\bar{y} = \begin{cases} \sqrt{1-\zeta/\beta^4} & \sqrt{\zeta} < 1-\rho\,, \\ 0 & \sqrt{\zeta} \geq 1-\rho\,, \end{cases} \tag{A.27}$$

and

$$\begin{aligned} \mathcal{F}_s = &\ 4\log\frac{\zeta}{\beta^4}\log\frac{1+\bar{y}}{1-\bar{y}} - 2\log\frac{1+\bar{y}}{1-\bar{y}}\log(4(1-\bar{y}^2)) + \\ &+4\mathrm{Li}_2\left(\frac{1+\bar{y}}{2}\right) - 4\mathrm{Li}_2\left(\frac{1-\bar{y}}{2}\right), \end{aligned} \tag{A.28}$$

$$\mathcal{F}_c(x) = \begin{cases} \sqrt{1-\zeta/(1-x)^2} & x < 1-\sqrt{\zeta}\,, \\ 0 & x \geq 1-\sqrt{\zeta}\,. \end{cases} \tag{A.29}$$

We point out that the subtraction scheme defined in this appendix is rather general. In fact, the condition that determines the subtraction, eq. (A.20), applies to the variables of the parton that recoils against the system, regardless of the nature of the system itself.

We can now go back to the point raised in section 4.5, concerning the number of negative-weight events in our MC@NLO. As mention there, this number is a (very complicated) function of the free parameters used to define the subtractions terms: $\tilde{\rho}$ and ω for standard subtraction, ζ for ζ–subtraction. However, for a crude estimate of their best values, we can assume that the partonic cross sections do not depend upon them, and thus that the parameter dependence of $J_{\mathbb{S}}$ is only due to the size of the phase-space regions in which the subtraction of the counterterms is performed (this amounts to saying that the only parameter dependence taken into account is that due to the Θ functions in eqs. (A.15)–(A.17) for the standard subtraction, and in eqs. (A.21)–(A.23) for the ζ–subtraction). In general, there is a cancellation between the first two terms in eq. (4.24); to a certain extent, the term $-d\overline{\Sigma}_{ab}^{(f)}/d\phi_3|_{\mathrm{ct}}$ can be tuned in order to cancel with $d\overline{\Sigma}_{ab}/d\phi_3|_{\mathrm{MC}}$. This explains why ζ-subtraction is a better choice than standard subtraction in the present case: the region in which the counterterms are subtracted, eq. (A.20), is a better match to the dead zone, eq. (A.69), compared with the corresponding region for the standard subtraction.

In this paper we do not try to find an optimal value for the parameter ζ. The crude estimate mentioned above gives $\zeta = 0.275$; the fraction of negative-weight events that we get by using this value is 12%. This fraction appears to be remarkably stable against variations of ζ in the region $0.05 \lesssim \zeta \lesssim 0.5$.

A.4 Choice of Bjorken x's

Here we give the form of the change of variables in eq. (4.17). We stress that this is a convenient procedure for getting rid of the $\tilde{\mathbf{2}}$ configurations that appear in eq. (4.14), which also has the virtue of being useful for defining the MC-like terms of eq. (4.21), as we shall show in appendix A.5. These two problems could in principle be solved independently, and by different means, as it may be necessary in other MC implementations; however, our master formula, eq. (4.22), would still hold. We denote by

$$\mathbf{3}(x_1, x_2, \phi_3) \tag{A.30}$$

the three-body configuration that is obtained by using the three-body phase-space variables ϕ_3, and the momentum fractions x_1 and x_2; the momenta in eq. (A.30) are understood in the hadronic c.m. frame. Furthermore, we denote by

$$\mathbf{2}_s(x_1, x_2, \phi_3^{(s)}) \equiv \mathbf{3}(x_1, x_2, \phi_3^{(s)}) \tag{A.31}$$

the two-body configuration that is obtained by setting to zero the energy of the final-state light parton; $\phi_3^{(s)}$ means "soft limit" of ϕ_3. In the parametrization given in eq. (A.10), this limit can be obtained by letting $x \to 1$ (see eq. (A.3)). It has to be pointed out that any $\mathbf{2}$ configuration appearing in eq. (4.14) can be viewed as the soft limit of a $\mathbf{3}$ configuration; thus, eq. (A.31) is fully general. Analogously, we denote by

$$\tilde{\mathbf{2}}_{c+}(x_1, x_2, \phi_3^{(c+)}) \equiv \mathbf{3}(x_1, x_2, \phi_3^{(c+)}) \tag{A.32}$$

$$\tilde{\mathbf{2}}_{c-}(x_1, x_2, \phi_3^{(c-)}) \equiv \mathbf{3}(x_1, x_2, \phi_3^{(c-)}) \tag{A.33}$$

the three-body configurations that are obtained when the final-state light parton becomes collinear to the incoming parton coming from the left ($p_1 \parallel k$) and from the right ($p_2 \parallel k$) respectively; these limits correspond to $y \to 1$ and $y \to -1$ respectively (see eqs. (A.4) and (A.5)).

We now look for changes of variables of the type given in eq. (4.17). There, the possibility exists to have a different form for each of the terms contributing to eq. (4.14). However, since we are looking for a solution to the equation $\mathbf{2}(x_i) = \tilde{\mathbf{2}}(\tilde{x}_i)$, we shall actually have to find a change of variables (4.17) for each kinematical configuration $\mathbf{2}$ and $\tilde{\mathbf{2}}$. Since only kinematics matters here, it follows that the collinear counterterms to $d\Sigma_{ab}^{(f)}|_{\rm ev}$, and the terms $d\Sigma_{ab}^{(c\pm)}|_{\rm ev}$, must be treated exactly in the same way in eq. (4.14), since they all have $\tilde{\mathbf{2}}$ kinematics. Similarly, the Born term $d\overline{\Sigma}_{ab}^{(b)}$, the soft-virtual term $d\overline{\Sigma}_{ab}^{(sv)}$, the soft counterterm to $d\Sigma_{ab}^{(f)}|_{\rm ev}$, and the counterterms $d\Sigma_{ab}^{(c\pm)}|_{\rm ct}$, are treated in the same way, since they have $\mathbf{2}$ kinematics. In summary, our task amounts to finding the three changes of variables

$$x_1 = x_1^{(\alpha)}(z_1, z_2, \phi_3)\,, \qquad x_2 = x_2^{(\alpha)}(z_1, z_2, \phi_3)\,, \qquad \alpha = s, c+, c-\,, \tag{A.34}$$

such that, at *fixed* z_1, z_2 and ϕ_3,

$$\begin{aligned}
&\mathbf{2}_s(x_1^{(s)}(z_1, z_2, \phi_3), x_2^{(s)}(z_1, z_2, \phi_3), \phi_3^{(s)}) = \\
&\qquad = \tilde{\mathbf{2}}_{c+}(x_1^{(c+)}(z_1, z_2, \phi_3), x_2^{(c+)}(z_1, z_2, \phi_3)\,, \phi_3^{(c+)}) \\
&\qquad = \tilde{\mathbf{2}}_{c-}(x_1^{(c-)}(z_1, z_2, \phi_3), x_2^{(c-)}(z_1, z_2, \phi_3)\,, \phi_3^{(c-)})\,,
\end{aligned} \tag{A.35}$$

where the equality holds for the W^+ and W^- variables only. It is then apparent that eqs. (4.14) and (4.19) are identical; in fact, if $\mathbf{2} = \tilde{\mathbf{2}}$ for the W^+ and W^- variables, then $O(\mathbf{2}) = O(\tilde{\mathbf{2}})$, since the outgoing light parton cannot contribute to the kinematics of any infrared-safe observables, being either soft or collinear to the beam line.

The requirement $\mathbf{2} = \tilde{\mathbf{2}}$ amounts to two conditions, since in both configurations the transverse momentum of the W pair is equal to zero, and only two components of the four-momentum of the pair are left to be constrained. Notice that it is sufficient to constrain the pair momentum, since the internal variables of the pair are held fixed in eq. (A.35). Instead of imposing eq. (A.35) directly, it turns out to be convenient to follow the strategy developed in ref. [23] (where the procedure described here is called "event projection"), where $\mathbf{2}$ and $\tilde{\mathbf{2}}$ are related to $\mathbf{3}$. This is done by imposing

$$O_1(\mathbf{3}(z_1, z_2, \phi_3)) = O_1(\mathbf{2}_s(x_1^{(s)}, x_2^{(s)}, \phi_3^{(s)})) \,, \tag{A.36}$$

$$O_2(\mathbf{3}(z_1, z_2, \phi_3)) = O_2(\mathbf{2}_s(x_1^{(s)}, x_2^{(s)}, \phi_3^{(s)})) \,, \tag{A.37}$$

and analogous equations for $c+$ and $c-$; O_1 and O_2 are two arbitrary observables, defined with the variables of the $\mathrm{W^+W^-}$ pair; therefore, the conditions in eq. (A.35) are effectively implemented by relating $\mathbf{2}$ and $\tilde{\mathbf{2}}$ to $\mathbf{3}$. It must be stressed that eqs. (A.36) and (A.37) are imposed in the hadronic c.m. frame for those observables that are not invariant under longitudinal boosts.

Notice that this procedure can be applied to any production process; in the general case, the observables O_1 and O_2 are defined with the variables of the *system* that recoils against the hard parton emitted by real diagrams at the NLO level. In the present case, one obvious choice is

$$O_1 = M^2_{\mathrm{WW}} \,, \qquad O_2 = y_{\mathrm{WW}} \,, \tag{A.38}$$

where M^2_{WW} and y_{WW} are the invariant mass squared and the rapidity of the pair respectively. These two observables are easily expressed in terms of the momentum fractions $x_{1,2}$, and of the phase space variables in the parametrization given in eq. (A.10):

$$M^2_{\mathrm{WW}}(\mathbf{3}(z_1, z_2, \phi_3)) \equiv M^2_{\mathrm{WW}}(z_1, z_2, x) = x\, z_1\, z_2\, S \,, \tag{A.39}$$

$$y_{\mathrm{WW}}(\mathbf{3}(z_1, z_2, \phi_3)) \equiv y_{\mathrm{WW}}(z_1, z_2, x, y) = \log \xi + \frac{1}{2} \log \frac{z_1}{z_2} \,, \tag{A.40}$$

where S is the hadronic c.m. energy squared, and

$$\xi = \sqrt{\frac{2 - (1-x)(1+y)}{2 - (1-x)(1-y)}} \,. \tag{A.41}$$

By imposing the conditions in eqs. (A.36) and (A.37), and using the explicit forms for M^2_{WW} and y_{WW} given in eqs. (A.39) and (A.40), we get the following:

$$x_1^{(s)} = z_1\, \xi \sqrt{x} \,, \qquad x_2^{(s)} = z_2 \frac{\sqrt{x}}{\xi} \,, \tag{A.42}$$

$$x_1^{(c+)} = \frac{x_1^{(s)}}{x} \,, \qquad x_2^{(c+)} = x_2^{(s)} \,, \tag{A.43}$$

$$x_1^{(c-)} = x_1^{(s)} \,, \qquad x_2^{(c-)} = \frac{x_2^{(s)}}{x} \,. \tag{A.44}$$

It should be stressed that $x_1^{(\alpha)}$ and $x_2^{(\alpha)}$ as given in eqs. (A.42), (A.43), and (A.44) do not need to be less than 1 any longer; thus, the parton densities contained in $\overline{\Sigma}_{ab}$ have to be defined as equal to zero if their arguments are larger than 1.

Although the choice made in eq. (A.38) is arbitrary, we shall show later that, in the context of our MC@NLO, it is actually dictated by the procedure used by the MC to generate the Bjorken x's. In this procedure, the MC typically constrains either the invariant mass and the rapidity of the pair, or the invariant mass and the longitudinal momentum of the pair. In the current version of HERWIG the latter choice is adopted. We therefore also consider

$$O_1 = M^2_{\rm WW}\,, \qquad O_2 = k_{\rm WW}\,, \tag{A.45}$$

where $k_{\rm WW}$ is the longitudinal momentum of the pair. With the parametrization of eq. (A.10), we get

$$k_{\rm WW}(\mathbf{3}(z_1,z_2,\phi_3)) \equiv k_{\rm WW}(z_1,z_2,x,y) = -\frac{\sqrt{S}}{4}\left[(1-x)y(z_1+z_2)+(1+x)(z_2-z_1)\right]. \tag{A.46}$$

By imposing again eqs. (A.36) and (A.37), we now obtain

$$x_1^{(s)} = \frac{1}{2}\left(\sqrt{B^2+4A}-B\right), \qquad x_2^{(s)} = \frac{1}{2}\left(\sqrt{B^2+4A}+B\right), \tag{A.47}$$

$$x_1^{(c+)} = \frac{x_1^{(s)}}{x}, \qquad x_2^{(c+)} = x_2^{(s)}, \tag{A.48}$$

$$x_1^{(c-)} = x_1^{(s)}, \qquad x_2^{(c-)} = \frac{x_2^{(s)}}{x}, \tag{A.49}$$

where

$$A = x\,z_1\,z_2\,, \tag{A.50}$$

$$B = \frac{1}{2}\left[(1-x)y(z_1+z_2)+(1+x)(z_2-z_1)\right]. \tag{A.51}$$

Note that eq. (A.48) is identical to eq. (A.43), and eq. (A.49) is identical to eq. (A.44). This gives a consistent picture when the results given above are used in the MC@NLO, eq. (4.22). In fact, event projection has been introduced in the context of an NLO computation, and the constraints have been imposed on final-state quantities only. However, in an MC@NLO the initial-state configuration of the hard process is relevant as well, since it defines the initial conditions for the MC evolutions implicit in $I_{\rm MC}(O,\mathbf{3})$ and $I_{\rm MC}(O,\mathbf{2})$. Eqs. (A.42), (A.43), and (A.44) (and eqs. (A.47), (A.48), and (A.49)) all enter the definition of $\mathbb{S}$ events, and since in general $x_i^{(s)} \neq x_i^{(c+)} \neq x_i^{(c-)}$, there is an ambiguity for the prescription of the momenta of the initial state partons that initiate the backward evolution. However, as explained in section 4.5, the initial conditions for class $\mathbb{S}$ events are the $2\to 2$ hard processes; the initial states of these processes coincide, by definition, with those of the $2\to 2$ contributions to $\mathbb{S}$ events, but do *not* coincide in general with those of the quasi-$2\to 2$ contributions. If we denote the momentum fractions of the initial-state partons entering the $2\to 2$ processes as x_1' and x_2', owing to the fact that $1-x$ is the energy fraction of the outgoing light parton (see appendix A.1), we have

$$x_1^{(s)\prime} = x_1^{(s)}, \qquad x_2^{(s)\prime} = x_2^{(s)}, \tag{A.52}$$

$$x_1^{(c+)\prime} \equiv x\,x_1^{(c+)} = x_1^{(s)}, \qquad x_2^{(c+)\prime} \equiv x_2^{(c+)} = x_2^{(s)}, \tag{A.53}$$

$$x_1^{(c-)\prime} \equiv x_1^{(c-)} = x_1^{(s)}, \qquad x_2^{(c-)\prime} \equiv x\,x_2^{(c-)} = x_2^{(s)}. \tag{A.54}$$

Thus, $x_i^{(s)'} = x_i^{(c+)'} = x_i^{(c-)'}$, and there is no ambiguity in the momenta of the initial state partons. It should be noted that an ambiguity here would in any case affect our results at $\mathcal{O}(\alpha_s^2)$, since it would only be due to quasi-2 $\to$ 2 contributions; however, the absence of any ambiguity fits nicely in a picture where MC and NLO give different descriptions of the same physics.

In appendix A.5, we shall need to consider the soft and collinear limits of $x_i^{(s)}$ and $x_i^{(c\pm)}$. By direct computation, we obtain the following

$$x \to 1 \qquad \Longrightarrow \qquad \begin{cases} x_1^{(s)} \to z_1\,, & x_2^{(s)} \to z_2\,, \\ x_1^{(c+)} \to z_1\,, & x_2^{(c+)} \to z_2\,, \\ x_1^{(c-)} \to z_1\,, & x_2^{(c-)} \to z_2\,, \end{cases} \tag{A.55}$$

$$y \to 1 \qquad \Longrightarrow \qquad \begin{cases} x_1^{(s)} \to x z_1\,, & x_2^{(s)} \to z_2\,, \\ x_1^{(c+)} \to z_1\,, & x_2^{(c+)} \to z_2\,, \end{cases} \tag{A.56}$$

$$y \to -1 \qquad \Longrightarrow \qquad \begin{cases} x_1^{(s)} \to z_1\,, & x_2^{(s)} \to x z_2\,, \\ x_1^{(c-)} \to z_1\,, & x_2^{(c-)} \to z_2\,, \end{cases} \tag{A.57}$$

where we did not consider the $y \to 1$ limit of $x_i^{(c-)}$ and the $y \to -1$ limit of $x_i^{(c+)}$, since these limits are not relevant in the computation of singular cross sections. We stress that eqs. (A.55)–(A.57) hold regardless of whether eq. (A.38) or eq. (A.45) has been used. This is a manifestation of the factorization properties of QCD, as will become clear in appendix A.5.

A.5 Dead zone and subtraction term

Here we define the quantity $d\overline{\Sigma}_{ab}|_{\rm MC}$, introduced in eq. (4.20). We make use of some of the results of ref. [8]; although ref. [8] concerns only single vector boson production, the kinematics carry over to the present case if we simply replace the momentum of the vector boson by that of the pair. According to the discussion of section 4.1, we have to consider the cross section that we would obtain by keeping the $\mathcal{O}(\alpha_s)$ term in the expansion of the result of the MC, which we denote by $d\sigma|_{\rm MC}$. Roughly speaking, $d\sigma|_{\rm MC}$ is equal to the Born cross section, times the $\mathcal{O}(\alpha_s)$ term in the expansion of the Sudakov form factor. More precisely

$$\begin{aligned} d\sigma\Big|_{\rm MC} = \sum_{abc} & dx_1^{({\rm MC})} dx_2^{({\rm MC})} \frac{\alpha_s}{2\pi} d\sigma_{ab}^{(b)} \left(x_1^{({\rm MC})} P_1, x_2^{({\rm MC})} P_2\right) \times \\ & \times \left(\frac{d\xi_+}{\xi_+} \frac{dz_+}{z_+} \Theta(z_+^2 - \xi_+) P_{ac}^{(0)}(z_+) f_c^{(H_1)}\left(\frac{x_1^{({\rm MC})}}{z_+}\right) f_b^{(H_2)}(x_2^{({\rm MC})}) + \right. \\ & \left. + \frac{d\xi_-}{\xi_-} \frac{dz_-}{z_-} \Theta(z_-^2 - \xi_-) P_{bc}^{(0)}(z_-) f_a^{(H_1)}(x_1^{({\rm MC})}) f_c^{(H_2)}\left(\frac{x_2^{({\rm MC})}}{z_-}\right) \right). \end{aligned} \tag{A.58}$$

Equation (A.58) has been obtained directly from the QCD Sudakov form factor; in doing so, we have eliminated any cutoffs, precisely as in the case of the toy model. This implies that eq. (A.58) has soft and collinear divergences, which is what we want, since from it we derive $d\overline{\Sigma}_{ab}|_{\rm MC}$, namely a local counterterm to the real emission matrix element. The two terms on the r.h.s. account for the emissions from the incoming partons. If strongly-interacting partons were present in the final state at the Born level, corresponding terms should be added. The variables $\xi_\pm$ and $z_\pm$ are those used by the HERWIG MC [4] in the parton showering process. The Θ functions define the region where an emission can occur during the shower; therefore, from eq. (A.58) we can easily read the form of the dead zone, which is by definition the region where *no emission* can occur. We have

$$\text{Dead zone:} \qquad \xi_+ > z_+^2\,, \quad \xi_- > z_-^2\,, \tag{A.59}$$

where both conditions must be fulfilled. Finally, the $1/z_\pm$ factor that appears in the arguments of the parton densities in eq. (A.58) is due to the backward evolution in the shower from initial-state partons. In the case of emission from final-state partons, the arguments of the parton densities would simply read $x_1^{(\rm MC)}$ and $x_2^{(\rm MC)}$.

The values of $x_1^{(\rm MC)}$ and $x_2^{(\rm MC)}$ are determined by the MC (through energy-momentum conservation) at the Born level, after the generation of the **2** configuration which is used as initial condition for the shower. At the end of the shower, the momenta of the final-state particles are boosted, the boost being determined by imposing a relation between the variables of the W pair (in general, of the system produced at the Born level) before and after the shower. It follows that, after the shower, $x_1^{(\rm MC)}$ and $x_2^{(\rm MC)}$ have still the values determined by the **2** configuration; however, after the shower $x_1^{(\rm MC)}$ and $x_2^{(\rm MC)}$ cannot be interpreted any longer as Bjorken x's associated with the incoming partons.

This procedure is identical (when we consider the $\mathcal{O}(\alpha_{\rm s})$ result, that is, the emission of a single parton) to the one we defined for event projection in appendix A.4. In other words, the **3** configuration implicit in the l.h.s. of eq. (A.58) is related to the **2** configuration that initiates the shower by means of constraints such as those in eqs. (A.36) and (A.37).[7] It follows that, if we label the **3** configurations according to the conventions of this paper, $\mathbf{3} \equiv \mathbf{3}(z_1, z_2, \phi_3)$, we have

$$x_1^{(\rm MC)} = x_1^{(s)}\,, \qquad x_2^{(\rm MC)} = x_2^{(s)}\,. \tag{A.60}$$

The explicit form of $x_i^{(s)}$ depends on the observables O_i used by the MC to implement eqs. (A.36) and (A.37). A priori, these need *not* coincide with those given in eq. (A.38) or eq. (A.45). In practice, HERWIG uses the variables in eq. (A.45), and therefore the results in eq. (A.47) give the explicit form of $x_1^{(\rm MC)}$ and $x_2^{(\rm MC)}$. The relations between $x_1^{(\rm MC)}$, $x_2^{(\rm MC)}$ and z_1, z_2 allow us to recast eq. (A.58) in the following form

$$d\sigma\Big|_{\rm MC} = \sum_{ab} dz_1 dz_2 d\overline{\Sigma}_{ab}\Big|_{\rm MC}\,, \tag{A.61}$$

[7] This is the reason for introducing event projection in appendix A.4, instead of just imposing eq. (A.35). In a completely general case, one should map the three-body configurations obtained from the MC onto the **3** configurations of the NLO cross section, thus defining a $d\overline{\Sigma}_{ab}|_{\rm MC}$ that acts as a local counterterm in eq. (4.22).

where

$$d\overline{\Sigma}_{ab}\Big|_{\rm MC} = \frac{1}{z_+}\frac{\partial(x_1^{(s)},x_2^{(s)})}{\partial(z_1,z_2)} f_a^{(H_1)}\left(\frac{x_1^{(s)}}{z_+}\right) f_b^{(H_2)}(x_2^{(s)}) d\sigma_{ab}^+\Big|_{\rm MC} + \\ + \frac{1}{z_-}\frac{\partial(x_1^{(s)},x_2^{(s)})}{\partial(z_1,z_2)} f_a^{(H_1)}(x_1^{(s)}) f_b^{(H_2)}\left(\frac{x_2^{(s)}}{z_-}\right) d\sigma_{ab}^-\Big|_{\rm MC} \tag{A.62}$$

is the quantity that the appears in eq. (4.22), and

$$d\sigma_{ab}^+\Big|_{\rm MC} = \frac{\alpha_{\rm S}}{2\pi}\frac{d\xi_+}{\xi_+} dz_+ \Theta(z_+^2-\xi_+) P_{ca}^{(0)}(z_+) d\sigma_{cb}^{(b)}(x_1^{(s)}P_1, x_2^{(s)}P_2)\,, \tag{A.63}$$

$$d\sigma_{ab}^-\Big|_{\rm MC} = \frac{\alpha_{\rm S}}{2\pi}\frac{d\xi_-}{\xi_-} dz_- \Theta(z_-^2-\xi_-) P_{cb}^{(0)}(z_-) d\sigma_{ac}^{(b)}(x_1^{(s)}P_1, x_2^{(s)}P_2)\,, \tag{A.64}$$

where the sum over the parton label c is now understood.

In order to prove that $d\overline{\Sigma}_{ab}|_{\rm MC}$ in eq. (A.62) can act as a local counterterm in eq. (4.22), we have to express it in terms of the variables we use to describe the NLO cross section. This can be easily done as explained in ref. [8]: we relate the variables $\xi_\pm$ and $z_\pm$ used in the HERWIG parton showering to Lorentz invariant quantities, which we then express in terms of the phase-space variables of eq. (A.10). Following ref. [22], we define

$$t_k = (p_1-k)^2 = -\frac{s}{2}(1-x)(1-y)\,, \tag{A.65}$$

$$u_k = (p_2-k)^2 = -\frac{s}{2}(1-x)(1+y)\,, \tag{A.66}$$

where the momenta are assigned as in eq. (A.1), and x and y have been defined in eq. (A.2). We have [8]

$$t_k = -\frac{1-z_+}{z_+^2}\xi_+ M_{\rm WW}^2\,, \tag{A.67}$$

$$\frac{t_k u_k}{s} = \frac{(1-z_+)^2}{2z_+^2}\xi_+(2-\xi_+) M_{\rm WW}^2\,, \tag{A.68}$$

for a shower initiated by parton a. A shower initiated by the other incoming parton (b) is described in terms of ξ_- and z_-, which are in turn related to t_k and u_k by the same functional forms given in eqs. (A.67) and (A.68), with the formal exchange $t_k \leftrightarrow u_k$. We can now substitute eqs. (A.2), (A.65), and (A.66) into eqs. (A.67) and (A.68), to find $\xi_\pm(x,y)$ and $z_\pm(x,y)$, whose explicit form we do not report here, since it is not needed in the following.

We can now find the dead zone boundary in the $\langle x, y\rangle$ plane; from eq. (A.59) we obtain

$$\text{Dead zone:} \qquad \rho \le x \le x_{\rm DZ}\,, \quad |y| < Y_{\rm DZ}(x)\,, \tag{A.69}$$

where both conditions are met simultaneously, and

$$x_{\rm DZ} = \frac{7+\sqrt{17}}{16}\,, \tag{A.70}$$

$$Y_{\rm DZ}(x) = 1 - \frac{x}{1-x}\left(3-\sqrt{1+8x}\right). \tag{A.71}$$

We now consider eqs. (A.63) and (A.64). Using eq. (A.10), we obtain

$$\left.\frac{d\sigma^{+}_{ab}(p_1,p_2)}{d\phi_3}\right|_{\rm MC} = \frac{16\pi\alpha_{\rm S}}{s}\frac{P^{(0)}_{ca}(z_+)}{(1-x)\xi_+}\frac{\partial(z_+,\xi_+)}{\partial(x,y)}\mathcal{M}^{(b)}_{cb}(x_1^{(s)}P_1,x_2^{(s)}P_2)\Theta(y-Y_{\rm DZ}(x)), \tag{A.72}$$

$$\left.\frac{d\sigma^{-}_{ab}(p_1,p_2)}{d\phi_3}\right|_{\rm MC} = \frac{16\pi\alpha_{\rm S}}{s}\frac{P^{(0)}_{cb}(z_-)}{(1-x)\xi_-}\frac{\partial(z_-,\xi_-)}{\partial(x,y)}\mathcal{M}^{(b)}_{ac}(x_1^{(s)}P_1,x_2^{(s)}P_2)\Theta(-y-Y_{\rm DZ}(x)), \tag{A.73}$$

where $p_i = z_i P_i$ (according to the definition given in eq. (4.13)); therefore, $s = z_1 z_2 S$.

We can now consider the soft and collinear limits of $d\overline{\Sigma}_{ab}|_{\rm MC}$ given in eq. (A.62). We observe that, owing to the procedure adopted by the MC to determine $x_1^{(\rm MC)}$ and $x_2^{(\rm MC)}$, the luminosities that appear in eq. (A.62) have non-trivial soft and collinear limits. We make use of the fact, obtained by direct computation, that

$$\lim_{x\to 1} z_\pm = 1\,, \qquad \lim_{y\to 1} z_+ = x\,, \qquad \lim_{y\to -1} z_- = x\,. \tag{A.74}$$

Therefore, using eqs. (A.55)–(A.57), we obtain

$$\begin{aligned}
\lim_{x\to 1} f_a^{(H_1)}\left(\frac{x_1^{(s)}}{z_+}\right) f_b^{(H_2)}(x_2^{(s)}) &= \lim_{x\to 1} f_a^{(H_1)}(x_1^{(s)}) f_b^{(H_2)}\left(\frac{x_2^{(s)}}{z_-}\right)\\
&= \lim_{y\to 1} f_a^{(H_1)}\left(\frac{x_1^{(s)}}{z_+}\right) f_b^{(H_2)}(x_2^{(s)})\\
&= \lim_{y\to -1} f_a^{(H_1)}(x_1^{(s)}) f_b^{(H_2)}\left(\frac{x_2^{(s)}}{z_-}\right)\\
&= f_a^{(H_1)}(z_1) f_b^{(H_2)}(z_2)\,,
\end{aligned} \tag{A.75}$$

where the collinear limits have been considered only for the luminosities attached to short-distance cross sections which are singular in those limits. In the soft and collinear limits, $d\overline{\Sigma}_{ab}|_{\rm MC}$ will have to cancel the singularities of $d\overline{\Sigma}^{(f)}_{ab}$ and $d\overline{\Sigma}^{(c\pm)}_{ab}$. Owing to the event projection procedure described in appendix A.4, the luminosities attached to the latter two terms behave as follows in the soft and collinear limits:

$$\begin{aligned}
\lim_{x\to 1} f_a^{(H_1)}(x_1^{(s)}) f_b^{(H_2)}(x_2^{(s)}) &= \lim_{y\to 1} f_a^{(H_1)}(x_1^{(c+)}) f_b^{(H_2)}(x_2^{(c+)})\\
&= \lim_{y\to -1} f_a^{(H_1)}(x_1^{(c-)}) f_b^{(H_2)}(x_2^{(c-)})\\
&= f_a^{(H_1)}(z_1) f_b^{(H_2)}(z_2)\,,
\end{aligned} \tag{A.76}$$

where we used eqs. (A.55)–(A.57). Eqs. (A.75) exactly match eqs. (A.76). It is also straightforward to prove that

$$\frac{1}{z_\pm}\frac{\partial(x_1^{(s)},x_2^{(s)})}{\partial(z_1,z_2)} \longrightarrow 1 \tag{A.77}$$

in the soft and collinear limits, and that the corresponding jacobians in the NLO cross section have the same behaviour. This implies that the properties of $d\overline{\Sigma}_{ab}|_{\rm MC}$ as a local counterterm are determined by the behaviour of its short-distance parts, given in eqs. (A.72) and (A.73).

When considering eqs. (A.72) and (A.73), we notice that these formulae are valid for any parton flavours a and b. However, in our case the soft and collinear limits are all non-trivial only in the case $a = q$, $b = \bar{q}$, with $c = g$; we shall therefore explicitly discuss only this case. We expect the forms of the leading soft and collinear singularities of the cross sections given in eqs. (A.72) and (A.73) to be dictated by QCD factorization theorem. In other words, we *expect* the following equations to hold

$$\begin{aligned}
\left.\frac{d\sigma^{+}_{q\bar{q}}(p_1,p_2)}{d\phi_3}\right|_{\rm MC} + \left.\frac{d\sigma^{-}_{q\bar{q}}(p_1,p_2)}{d\phi_3}\right|_{\rm MC} &\xrightarrow{x\to 1} 8\pi\alpha_{\rm S} C_{\rm F} \frac{p_1\cdot p_2}{p_1\cdot k\; p_2\cdot k} \mathcal{M}^{(b)}_{q\bar{q}}(p_1,p_2) \\
&= \frac{64\pi\alpha_{\rm S}}{s} \frac{C_{\rm F}}{(1-x)^2(1-y^2)} \mathcal{M}^{(b)}_{q\bar{q}}(p_1,p_2)\,, \qquad \text{(A.78)} \\
\left.\frac{d\sigma^{+}_{q\bar{q}}(p_1,p_2)}{d\phi_3}\right|_{\rm MC} &\xrightarrow{y\to 1} \frac{4\pi\alpha_{\rm S}}{p_1\cdot k} P^{(0)}_{qq}(x) \mathcal{M}^{(b)}_{q\bar{q}}(xp_1,p_2) \\
&= \frac{16\pi\alpha_{\rm S}}{s} \frac{P^{(0)}_{qq}(x)}{(1-x)\,(1-y)} \mathcal{M}^{(b)}_{q\bar{q}}(xp_1,p_2)\,, \qquad \text{(A.79)} \\
\left.\frac{d\sigma^{-}_{q\bar{q}}(p_1,p_2)}{d\phi_3}\right|_{\rm MC} &\xrightarrow{y\to -1} \frac{4\pi\alpha_{\rm S}}{p_2\cdot k} P^{(0)}_{qq}(x) \mathcal{M}^{(b)}_{q\bar{q}}(p_1,xp_2) \\
&= \frac{16\pi\alpha_{\rm S}}{s} \frac{P^{(0)}_{qq}(x)}{(1-x)\,(1+y)} \mathcal{M}^{(b)}_{q\bar{q}}(p_1,xp_2)\,. \qquad \text{(A.80)}
\end{aligned}$$

Note that in the case of the soft limit we sum the short distance cross sections, thanks to eq. (A.75). Explicit computations of the collinear limits $y \to \pm 1$ in eqs. (A.72) and (A.73) give eqs. (A.79) and (A.80). On the other hand, in the soft limit $x \to 1$ we find the following:

$$\begin{aligned}
\left.\frac{d\sigma^{+}_{q\bar{q}}(p_1,p_2)}{d\phi_3}\right|_{\rm MC} + \left.\frac{d\sigma^{-}_{q\bar{q}}(p_1,p_2)}{d\phi_3}\right|_{\rm MC} &\xrightarrow{x\to 1} \frac{64\pi\alpha_{\rm S}}{s} C_{\rm F} \mathcal{M}^{(b)}_{q\bar{q}}(p_1,p_2) \times \\
\times \left(\frac{2\Theta(y+1/3)}{(1-x)^2(1-y)(3+y)} + \frac{2\Theta(-y+1/3)}{(1-x)^2(1+y)(3-y)} \right)&. \qquad \text{(A.81)}
\end{aligned}$$

It appears therefore that the choice of the showering variables in HERWIG violates the factorization properties of QCD cross sections. However, the situation is not as bad as it may seem. In fact, the kernel in eq. (A.81) has the following property:

$$\lim_{\varepsilon\to 0} \int_{-1+\varepsilon}^{1-\varepsilon} dy \left(\frac{2\Theta(y+1/3)}{(1-y)(3+y)} + \frac{2\Theta(-y+1/3)}{(1+y)(3-y)} - \frac{1}{1-y^2} \right) = 0\,. \qquad \text{(A.82)}$$

Eq. (A.82) implies that, although (in the absence of soft matrix element corrections [8]) HERWIG generates an incorrect angular distribution of soft gluons, the total amount of radiated soft energy is identical to that computed in perturbative QCD. This is reassuring, and explains why HERWIG correctly describes physical observables, eq. (A.81) notwithstanding: any IR safe observable is not sensitive to the details of the angular distribution of soft gluons.

Nevertheless, eq. (A.82) prevents us from using $d\overline{\Sigma}|_{\rm MC}$ in the construction of the MC@NLO. In fact, $d\overline{\Sigma}|_{\rm MC}$ acts as a *local* counterterm in eq. (4.22), and its soft behaviour (eq. (A.81)) prevents the necessary cancellation.

This problem can be solved in two different ways. The most straightforward procedure would be to define different showering variables, that would give short-distance cross sections with the correct soft limit. However, this would imply a modification of the existing MC code in one of its crucial parts, the showering algorithm. As already stressed, in this paper we prefer not to modify any substantial part of HERWIG.

The second solution relies on the observation that the angular distribution of soft gluons is irrelevant to MC@NLO physical results. We need a local counterterm to cancel the divergences of the real matrix element, but in the soft emission region the MC@NLO cross section is dominated by the pure MC contribution. On the other hand, for hard MC emission, we must have full control of the gluon emitted by the MC, since only in this way will the MC@NLO result match the NLO result. It follows that we can choose a form for $d\Sigma|_{\rm MC}$ identical to that we get by using eqs. (A.72) and (A.73) for non-soft emission, and a form that can act as a local counterterm for real matrix elements in the soft region. One possibility is therefore the following

$$
\begin{aligned}
\left.\frac{d\sigma^{+}_{q\bar{q}}(p_1,p_2)}{d\phi_3}\right|_{\rm MC} &= \frac{16\pi\alpha_{\rm S}}{s}\left(\frac{P^{(0)}_{qq}(z_+)}{(1-x)\xi_+}\,\frac{\partial(z_+,\xi_+)}{\partial(x,y)}\,\Theta(y-Y_{\rm DZ}(x))\mathcal{G}(x)\,+\right.\\
&\left.\qquad +\,\frac{P^{(0)}_{qq}(x)}{(1-x)\,(1-y)}(1-\mathcal{G}(x))\right)\mathcal{M}^{(b)}_{cb}(x^{(s)}_1P_1,x^{(s)}_2P_2)\,, \qquad \text{(A.83)}\\
\left.\frac{d\sigma^{-}_{q\bar{q}}(p_1,p_2)}{d\phi_3}\right|_{\rm MC} &= \frac{16\pi\alpha_{\rm S}}{s}\left(\frac{P^{(0)}_{qq}(z_-)}{(1-x)\xi_-}\,\frac{\partial(z_-,\xi_-)}{\partial(x,y)}\,\Theta(-y-Y_{\rm DZ}(x))\mathcal{G}(x)\,+\right.\\
&\left.\qquad +\,\frac{P^{(0)}_{qq}(x)}{(1-x)\,(1+y)}(1-\mathcal{G}(x))\right)\mathcal{M}^{(b)}_{ac}(x^{(s)}_1P_1,x^{(s)}_2P_2)\,, \qquad \text{(A.84)}
\end{aligned}
$$

where $\mathcal{G}$ is a largely arbitrary function, such that

$$
\lim_{x\to 1}\mathcal{G}(x)=0\,,\qquad \lim_{x\to x_{\rm DZ}}\mathcal{G}(x)=1\,. \qquad \text{(A.85)}
$$

A simple choice is

$$
\mathcal{G}(x)=\frac{(1-T(x))^{2\alpha}}{T(x)^{2\alpha}+(1-T(x))^{2\alpha}}\,,\qquad T(x)=\begin{cases}\dfrac{(x-\tilde{x}_{\rm DZ})}{(1-\tilde{x}_{\rm DZ})} & \tilde{x}_{\rm DZ}<x<1\,,\\ 0 & x<\tilde{x}_{\rm DZ}\,,\end{cases} \qquad \text{(A.86)}
$$

where

$$
\tilde{x}_{\rm DZ}=1-(1-x_{\rm DZ})\beta\,. \qquad \text{(A.87)}
$$

The parameters in eqs. (A.86) and (A.87) must be chosen in the following ranges: $\alpha\geq 1$, $0<\beta\leq 1$. The dependence of MC@NLO results upon these parameters will show the sensitivity of physical observables to the details of soft gluon emission.

JHEP06(2002)029

We point out that the procedure outlined above is not necessary when the matrix element is not singular in the soft limit. This happens for the partonic subprocesses $qg \to W^+W^-q$. In such a case, we simply set $\mathcal{G} \equiv 1$ in eqs. (A.83) and (A.84).

Finally, it is worth noting that the interesting features of eqs. (A.81) and (A.82) are of a general nature, since the hard production process is factorized into $\mathcal{M}^{(b)}_{q\bar{q}}$ in eq. (A.81). Thus, the problem raised here is not specific to $\mathrm{W^+W^-}$ production. We expect that the angular distributions of soft gluons in HERWIG do not agree with those prescribed by the factorization theorems of QCD. Therefore, the local counterterm $d\overline{\Sigma}|_{\rm MC}$ that we need in order to construct the MC@NLO will not coincide in general with the $\mathcal{O}(\alpha_{\rm s})$ HERWIG result. As in the present case, we expect that physical observables will not display any dependence upon the form chosen for $d\overline{\Sigma}|_{\rm MC}$ in the soft limit. In all cases, we must verify that the total amount of soft gluon energy given by the MC agrees with the QCD prediction, i.e., that the analogue of eq. (A.82) holds.

B. Technicalities of MC@NLO

In this appendix, we present formal proofs of some of the features of MC@NLO. We do so in the case of the toy model. However, we believe that the corresponding properties hold in the case of QCD as well, as we shall point out explicitly; the toy model just allows us to simplify the notation.

B.1 Construction of $I_{\rm MC}$

Our first aim is to give a precise definition of $I_{\rm MC}(O, x_{\rm M}(x))$. We identify an MC event with the set of energies of the n photons emitted in the shower, $\{x_k\}_{k=1}^n$; the dependence upon $x_{\rm M}(x)$ is understood. We define an MC@NLO event as the corresponding MC event plus the energy of the photon emitted at the NLO level: $\{x, x_k\}_{k=1}^n$. In the case of soft emission, $x = 0$, the MC@NLO event simply coincides with the MC event. Any observable O will be in general a multi-valued function of MC@NLO events, $O(\{x, x_k\}_{k=1}^n) = \{O_1, \ldots, O_p\}$, with $1 \le p \le n+1$, the O_j's being values in the physical range of O. We then define

$$\mathcal{N}(\{x, x_k\}_{k=1}^n; \overline{O}, \delta O) = \sum_{j=1}^{p} \Theta\left(\overline{O} + \frac{\delta O}{2} - O_j\right) \Theta\left(O_j - \overline{O} + \frac{\delta O}{2}\right), \tag{B.1}$$

where $\overline{O}$ is any given value that the observable O can assume in its physical range. $\mathcal{N}$ is then just the number of entries in the bin $(\overline{O} - \delta O/2, \overline{O} + \delta O/2)$ resulting from the MC@NLO event $\{x, x_k\}_{k=1}^n$. Therefore $0 \le \mathcal{N} \le p \le n+1$. Starting from a given $x_{\rm M}(x)$, we generate N MC events, and we define

$$I_{\rm MC}(\overline{O}, x_{\rm M}(x)) = \lim_{N\to\infty} \lim_{\delta O \to 0} \frac{1}{N} \frac{1}{\delta O} \sum_{i=1}^{N} \mathcal{N}\left(\{x, x_k^{(i)}\}_{k=1}^{n_i}; \overline{O}, \delta O\right), \tag{B.2}$$

n_i being the number of photons, with energies $x_k^{(i)}$, emitted in event number i. Thus in general $I_{\rm MC}(O, x_{\rm M}(x))$ is the differential cross section that one obtains by running the MC starting at $x_{\rm M}(x)$, defining O through all photons emitted in the shower, *plus* that emitted at the NLO level, normalized by the total number of MC events generated.

If we assume that the MC is switched off, then from eq. (B.2) we obtain

$$I_{\rm MC}(O, x_{\rm M}(x)) = \delta(O - O(x))\,, \tag{B.3}$$

$O(x)$ being the single-valued function used at the NLO level to describe the observable O. Eq. (B.3) can be proven by using eq. (B.2) in the computation of $\int dO I_{\rm MC}(O) f(O)$, for any arbitrary function $f(O)$. With the same technique, we can also find a closed form for $I_{\rm MC}(O, x_{\rm M}(x))$ in the general case. We denote by $P_n(x_1, \ldots, x_n)$ the probability that the toy MC emits n photons with energies $x_1 > x_2 > \cdots > x_n$. This probability can be obtained from eq. (3.14), and reads as follows

$$P_0 \,=\, \Delta(x_0, x_{\rm M}(x))\,, \tag{B.4}$$

$$P_n(x_1, \ldots, x_n) \,=\, a^n \Delta(x_0, x_{\rm M}(x)) \Theta(x_{\rm M}(x) - x_1) \prod_{i=1}^{n} \frac{Q(x_i)}{x_i} \Theta(x_i - x_{i+1})\,, \tag{B.5}$$

where in the case of n emissions we define $x_{n+1} = x_0$. We also denote by $\mathbb{P}_n$ the probability of having n emissions, regardless of the photon energies. Clearly, $\mathbb{P}_0 \equiv P_0$. For $n > 0$, we can compute this quantity by integrating eq. (B.5), to obtain

$$\mathbb{P}_n \equiv \int_0^1 \left(\prod_{i=1}^{n} dx_i \right) P_n(x_1, \ldots, x_n) = \frac{a^n}{n!} \left[\int_{x_0}^{x_{\rm M}(x)} dz \frac{Q(z)}{z} \right]^n \Delta(x_0, x_{\rm M}(x))\,. \tag{B.6}$$

With eqs. (B.4) and (B.6) we can check directly that the toy MC conserves probability:

$$\mathbb{P}_0 + \sum_{n=1}^{\infty} \mathbb{P}_n = 1\,. \tag{B.7}$$

We now introduce the functions $O_{n;j}(x; x_1, \ldots, x_n)$ that return the multi-valued observable O (for single-value observables, we simply have $j \equiv 1$). Observing that the sum over all events in eq. (B.2) can be split into sums over events with a fixed number of emissions, we easily get the following:

$$\begin{aligned} I_{\rm MC}(O, x_{\rm M}(x)) \,=\, & \mathbb{P}_0 \delta(O - O(x)) \,+ \\ & + \sum_{n=1}^{\infty} \int_0^1 \left(\prod_{i=1}^{n} dx_i \right) P_n(x_1, \ldots, x_n) \sum_{j=1}^{p(x;x_1,\ldots,x_n)} \delta(O - O_{n;j}(x; x_1, \ldots, x_n))\,, \end{aligned} \tag{B.8}$$

where the integer $p(x; x_1, \ldots, x_n)$ in general depends on the kinematics (this happens with jet observables, for example). Eq. (B.8) can be used, together with the explicit forms of the emission probabilities P_n, to recover the results for $I_{\rm MC}(y)$ and $I_{\rm MC}(z)$ given in eqs. (3.25), (3.26), (3.38), and (3.39); these equations have been obtained by direct computation, which constitutes a cross check of eq. (B.8).

It should be noted that eq. (B.8) is obtained from eq. (B.2) without using the explicit form of the emission probabilities P_n. Thus, both equations hold in the QCD case as well, with obvious formal substitutions (the MC@NLO event $\{x, x_k\}_{k=1}^n$ gets replaced by an $\mathbf{n+1}$ kinematical configuration, and the information on parton flavours must be inserted).

JHEP06(2002)029

B.2 Total rate for exclusive observables

We define an observable to be exclusive if its values can be obtained through a single-valued function, regardless of the number of the photons emitted in the shower. The hardest jet energy y, considered in section 3.4, is an observable of this kind, whereas the fully inclusive photon energy z is not. It is clear that in the exclusive case the function $\mathcal{N}$ defined in appendix B.1 has the following properties:

$$\mathcal{N}\left(\{x, x_k\}_{k=1}^{n}; \overline{O}, \delta O\right) = 0, 1 \qquad \forall \overline{O}\,; \tag{B.9}$$

$$\exists! \overline{O} \quad \text{for } \delta O \to 0 \quad \text{such that:} \quad \mathcal{N}(\{x, x_k\}_{k=1}^{n}; \overline{O}, \delta O) = 1\,. \tag{B.10}$$

Eq. (B.10) holds for any given MC@NLO event. Then

$$\int dO I_{\rm MC}(O, x_{\rm M}(x)) = 1\,. \tag{B.11}$$

We can also get this results directly from eq. (B.8), using eq. (B.7), and the fact that the observable O is a single-valued function (and thus $p(x; x_1, \ldots, x_n) = 1$). It is then apparent that eq. (B.11) also holds in QCD. We can use eq. (B.11) in eq. (3.20); we get

$$\begin{aligned} \int dO \left(\frac{d\sigma}{dO}\right)_{\rm msub} &= \int_0^1 dx \left[\left(\int dO I_{\rm MC}(O, x_{\rm M}(x))\right) \frac{a[R(x) - BQ(x)]}{x} + \right. \\ &\quad \left. + \left(\int dO I_{\rm MC}(O, 1)\right) \left(B + aV + \frac{aB[Q(x) - 1]}{x}\right)\right] \\ &= \int_0^1 dx \left[B + aV + a\frac{R(x) - B}{x}\right] \equiv \sigma_{\rm tot}\,. \end{aligned} \tag{B.12}$$

In the particular case of $O \equiv y$, this result could be obtained from the results for $d\sigma/dy$ given in section 3.4, by including the full x_0 dependence that has been neglected there.

Since eq. (B.11) holds in QCD, we conclude that eq. (B.12) also holds in QCD. This is in fact confirmed numerically by the results for the exclusive observables (such as $p_{\rm T}^{\rm W^+}$, $p_{\rm T}^{\rm (WW)}$, $\Delta\phi^{\rm (WW)}$) presented in section 4.7 (in the case that no cuts are applied).

B.3 Expansion to $\mathcal{O}(\alpha_{\rm s})$

We now consider the expansion to order a and $\alpha_{\rm s}$ of our formulae for the toy model and QCD MC@NLO, eqs. (3.20) and (4.22) respectively. We start with the toy model. Since we are not interested in $\mathcal{O}(a^2)$ terms, only zero- and one-photon emissions from the MC are relevant. We have

$$I_{\rm MC}(O, x_{\rm M}(x)) = \delta(O - O(x)) + \mathcal{O}(a)\,. \tag{B.13}$$

In eq. (B.13) we do not need to know the $\mathcal{O}(a)$ term, since $I_{\rm MC}(O, x_{\rm M}(x))$ is already multiplied by a in eq. (3.20). On the other hand

$$I_{\rm MC}(O, 1) = \Delta(x_0, 1)\delta(O - O(0)) + a\int_{x_0}^1 dt \frac{Q(t)}{t}\Delta(x_0, 1)\delta(O - O(t)) + \mathcal{O}(a^2) \tag{B.14}$$

$$= \left(1 - a\int_{x_0}^1 dt\frac{Q(t)}{t}\right)\delta(O - O(0)) + a\int_{x_0}^1 dt\frac{Q(t)}{t}\delta(O - O(t)) + \mathcal{O}(a^2)\,. \tag{B.15}$$

The first term on the r.h.s. of eq. (B.14) is the contribution to the observable in the case that the MC does not emit any photons. The corresponding observable must be computed in the soft limit, with a coefficient which is the probability of no MC emission, that is, the Sudakov form factor for the range $(x_0, 1)$. The second term corresponds to the one-emission contribution. We stress that eqs. (B.13) and (B.14) could have been obtained directly from eq. (B.8). We can now substitute eqs. (B.13) and (B.15) into eq. (3.20). We get:

$$\begin{aligned}\left(\frac{d\sigma}{dO}\right)_{\rm msub} = \int_0^1 dx \Bigg[& \delta(O-O(x))\frac{a[R(x)-BQ(x)]}{x} + \\ & + \delta(O-O(0))\left(B + aV - \frac{aB}{x}\right) + \\ & + aB\delta(O-O(0))\left(\frac{Q(x)}{x} - \int_{x_0}^1 dt\frac{Q(t)}{t}\right) + \\ & + aB\int_{x_0}^1 dt\delta(O-O(t))\frac{Q(t)}{t}\Bigg] + \mathcal{O}(a^2)\,. \end{aligned} \tag{B.16}$$

We thus have

$$\begin{aligned}\left(\frac{d\sigma}{dO}\right)_{\rm msub} = & \int_0^1 dx \left[\delta(O-O(x))\frac{aR(x)}{x} + \delta(O-O(0))\left(B + aV - \frac{aB}{x}\right)\right] + \\ & + aB\int_0^{x_0} dx\frac{Q(x)}{x}\Big[\delta(O-O(0)) - \delta(O-O(x))\Big] + \mathcal{O}(a^2)\,. \end{aligned} \tag{B.17}$$

If we use eq. (B.17) to compute $\langle O\rangle = \int dOO(d\sigma/dO)$, we get the NLO result of eq. (3.13), up to uncontrolled $\mathcal{O}(a^2)$ terms, and up to terms suppressed by powers of x_0, which are numerically very small, and can be eliminated by letting $x_0 \to 0$. In the case of QCD, the role of x_0 is played by some soft and collinear cutoffs, which cannot be set to zero. Thus, in the real case we have power-suppressed corrections to the $\mathcal{O}(\alpha_{\rm s})$ result. Notice that MC's do not in any case control terms of this kind, and the dependence upon the cutoff is in general power-like. Since, as discussed in section 3.3.2, the MC@NLO and the MC behave the same with respect to the cutoff, the result in eq. (B.17) implies that the dependence upon the cutoff of the MC@NLO result beyond $\mathcal{O}(\alpha_{\rm s})$ is also power-like.

Had we applied the same procedure using naive subtraction, eq. (3.19), we would have obtained the following (after letting $x_0 \to 0$):

$$\begin{aligned}\left(\frac{d\sigma}{dO}\right)_{\rm naive} = & \int_0^1 dx \left[\delta(O-O(x))\frac{aR(x)}{x} + \delta(O-O(0))\left(B + aV - \frac{aB}{x}\right)\right] + \\ & + aB\int_0^1 dx\frac{Q(x)}{x}\Big[\delta(O-O(x)) - \delta(O-O(0))\Big] + \mathcal{O}(a^2). \end{aligned} \tag{B.18}$$

The second integral on the r.h.s. of this equation vanishes upon integration over the whole range of O, and the total rate (if meaningful) is then identical to that obtained from eq. (B.17). On the other hand, if we compute $\langle O\rangle$ the coefficient of $O(x)$ reads

$a(R(x)+BQ(x))/x$, instead of the $aR(x)/x$ that we get from eq. (B.17). This proves in general that the naive subtraction method suffers from double counting, as already shown in eq. (3.30) and (3.42).

We now turn to the case of the QCD MC@NLO. The situation is completely analogous to that of the toy model. The NLO result can be recovered only if the $\mathcal{O}(\alpha_{\rm S})$ terms of the expansion of $I_{\rm MC}(O,\mathbf{2})$, multiplied by the Born term, cancel the contributions of the terms $d\overline{\Sigma}|_{\rm MC}$ in eq. (4.22). The QCD case is however more involved, owing to the fact that, in the case of emission from initial-state partons, the Sudakov factor has a non-trivial dependence upon the parton densities. The analogue of eq. (B.15) in the QCD case reads as follows

$$\begin{aligned} I_{\rm MC}(O,\mathbf{2}) = & \Bigg(1-\frac{\alpha_{\rm S}(\mu_{\rm MC})}{2\pi}\int_{q_c}\frac{d\xi_+}{\xi_+}\frac{dz_+}{z_+}\Theta(z_+^2-\xi_+)P_{ac}^{(0)}(z_+)\frac{f_c^{(H_1)}(x_1^{(\rm MC)}/z_+,\mu_{\rm MC})}{f_a^{(H_1)}(x_1^{(\rm MC)},\mu_{\rm MC})}- \\ & -\frac{\alpha_{\rm S}(\mu_{\rm MC})}{2\pi}\int_{q_c}\frac{d\xi_-}{\xi_-}\frac{dz_-}{z_-}\Theta(z_-^2-\xi_-)P_{bc}^{(0)}(z_-)\frac{f_c^{(H_2)}(x_2^{(\rm MC)}/z_-,\mu_{\rm MC})}{f_b^{(H_2)}(x_2^{(\rm MC)},\mu_{\rm MC})}\Bigg)\times \\ & \times\delta(O-O(\mathbf{2}))+\cdots \end{aligned} \tag{B.19}$$

where the dots indicate two more terms of $\mathcal{O}(\alpha_{\rm S})$, identical to those appearing in eq. (B.19) except for the formal substitution $\delta(O-O(\mathbf{2}))\to\delta(O-O(\mathbf{3}))$. In eq. (B.19) we understand that the initial-state partons that initiate the shower have flavours a and b. The lower bound on the integration range, q_c, collectively denotes the cutoffs that prevent MC emissions from being soft and/or collinear, and plays the same role as x_0 in eq. (B.15). The arguments of the parton densities have already been discussed in appendix A.5. We also indicate explicitly the dependence upon the hard scale $\mu_{\rm MC}$, used by the MC in the shower; in principle, this scale depends on the kinematics of the emission. However, this is not important in what follows, and the notation we use simplifies the discussion. The Born term that gets multiplied by eq. (B.19) reads as follows:

$$\frac{d\overline{\Sigma}_{ab}^{(b)}}{d\phi_2}=\frac{\partial(x_1^{(s)},x_2^{(s)})}{\partial(z_1,z_2)}f_a^{(H_1)}\left(x_1^{(s)},\mu_{\rm MC@NLO}\right)f_b^{(H_2)}\left(x_2^{(s)},\mu_{\rm MC@NLO}\right)\frac{d\sigma_{ab}^{(b)}}{d\phi_2}\left(x_1^{(s)}P_1,x_2^{(s)}P_2\right). \tag{B.20}$$

The $\mathcal{O}(\alpha_{\rm S})$ terms in the product of eqs. (B.19) and (B.20) must cancel the contributions of $d\overline{\Sigma}_{ab}|_{\rm MC}$ in eq. (4.22). It is apparent that a necessary condition for this to happen is $x_i^{(\rm MC)}\equiv x_i^{(s)}$; this is precisely eq. (A.60). However, the present proof adds one piece of information which could not be inferred from eq. (A.60): in fact, that equation merely states the fact that the functional form of $x_i^{(\rm MC)}$ can be obtained through the procedure of event projection; however, it does not constrain the $x_i^{(\rm MC)}$ variables to be equal to the $x_i^{(s)}$ variables used in the event projection *at the NLO level*. This logical step is made only here. Once the condition on the Bjorken x's is fulfilled, we achieve the desired cancellation only if the short-distance cross section resulting from eqs. (B.19) and (B.20) match those in $d\overline{\Sigma}_{ab}|_{\rm MC}$. The former have the form reported in eqs. (A.72) and (A.73), whereas we have chosen the latter as in eqs. (A.83) and (A.84); therefore, we seemingly have uncancelled $\mathcal{O}(\alpha_{\rm S})$ terms. However, as argued in appendix A.5, the difference between the two forms cannot induce any effects on infrared-safe observables. This is in fact what we observe in practice (see section 4.7). In other words, we do not have event-by-event cancellation, but we do have cancellation on a statistical basis. We remind the reader that the uncancelled $\mathcal{O}(\alpha_{\rm S})$ terms would disappear event by event if the MC had the correct angular distribution for soft gluon emission. We point out that, owing to the presence of the cutoff q_c in eq. (B.19), we still have a small uncancelled $\mathcal{O}(\alpha_{\rm S})$ contribution, which is the analogue of the last term on the r.h.s. of eq. (B.17). As in the case of the toy model, this contribution is suppressed by powers of the cutoff. Notice that, because of the simultaneous presence of soft and collinear singularities in QCD, these powers can in general be multiplied by powers of the logarithm of the cutoff, but they remain negligible.

We also mention the fact that the flavour structure of eq. (B.19) is non-trivial. For the cancellation at $\mathcal{O}(\alpha_{\rm S})$ to occur, we need to ensure that the flavour assignment in the initial conditions for $\mathbb{S}$ events matches the flavour structure of the Born term, eq. (B.20). This is in fact guaranteed by the prescription described in section 4.5.

We remark finally that the mismatch between $\mu_{\rm MC}$ and $\mu_{\rm MC@NLO}$ is formally of higher order. The two scales have fairly different physical meaning, and one cannot in general assume that $\mu_{\rm MC} = \mu_{\rm MC@NLO}$. However, since the evolution of the parton densities in the MC is only leading-logarithmic, in the present paper we have made that assumption. A limited study of scale dependence has been presented in section 4.7.

B.4 Resummation of large logarithms

Here we give arguments to indicate that the MC@NLO resums large logarithms in the same way as the ordinary MC. As previously, we start with the toy model, by reconsidering the results of section 3.4 for the variable z. We assume that eq. (3.37) is replaced by the more general form

$$\left(\frac{d\sigma}{dz}\right)_{\rm MC} = BS(z,a)\,. \tag{B.21}$$

The function S is largely arbitrary, but it should at least have the following properties:

- S has a power expansion in a, starting with an $\mathcal{O}(a)$ term;
- the leading behaviour at small z is a/z.

The function in eq. (3.37) has these properties. Equation (B.21) is, however, more general; we can for example assume that S correctly takes into account next-to-leading logarithmic terms. Eqs. (3.38) and (3.39) now become

$$I_{\rm MC}(z,1) = S(z,a)\,, \tag{B.22}$$

$$I_{\rm MC}(z,x_{\rm M}(x)) = \delta(z-x) + S(z,a)\Theta(x_{\rm M}(x)-z)\,. \tag{B.23}$$

We thus get

$$\left(\frac{d\sigma}{dz}\right)_{\rm msub} = S(z,a)\left[\sigma_{\rm tot} + a\frac{R(z)-BQ(z)}{zS(z,a)} - a\int_{x_{\rm M}^{-1}(z)}^{1} dx\frac{R(x)-BQ(x)}{x}\right]. \tag{B.24}$$

Thus, since $S(z, a)$ diverges as $1/z$ for $z \to 0$, we have

$$\left(\frac{d\sigma}{dz}\right)_{\text{msub}} \approx \sigma_{\text{tot}} S(z, a) + \cdots \tag{B.25}$$

where the neglected terms are either constant for $z \to 0$, or vanish in this limit. Eq. (B.25) generalizes eq. (3.45). Formally, the $\mathcal{O}(a)$ terms in σ_{tot} produce a new tower of NLL terms when multiplied by the LL tower in $S(z, a)$. But since σ_{tot} is factorized, this new tower of logarithms does not produce any change in the z spectrum of the emitted photons compared to the MC. This is in fact evident if we normalize the z distribution to the total exclusive rate, which is σ_{tot} in the case of MC@NLO, and B in the case of MC: in both cases, we get $S(z, a)$. Furthermore, the terms neglected in eq. (3.45) do not lead to any logarithmic-enhanced behaviour. Thus, the leading and whatever subleading logarithmic behaviour in z of soft photon emission is the same in MC and MC@NLO evolution.

Notice that eq. (B.25) implies that the ratio of the number of soft photons to the number of showers is the same as in the ordinary MC. This fact, and the fact that the energy spectrum of the emitted photons is also the same, suggests that, for an arbitrary observable O that gets resummed through ordinary MC methods, the resummation performed by the MC@NLO has the same logarithmic accuracy as the MC.

However, the preceding arguments are fairly closely linked to the details of the toy model, which makes it difficult to extend them to the QCD case. We therefore now give another argument, which can easily be extended to QCD; this argument works only for a given class of observables, which however includes most of the common measurable quantities. Namely, we consider the case of an exclusive observable O such that, when $O \simeq O_s$, any fixed-order cross section contains large logarithms $\log(O - O_s)$. Furthermore, we assume that, regardless of the number of photons emitted, at NLO and/or by the MC, when $O \simeq O_s$ the energies of *all* photons emitted are constrained to be small; in other words, any hard photon emission moves O away from O_s. This requirement is not very restrictive; still, there can be observables which do not meet it, and have large logarithms in the cross sections, but these logarithms will be multiplied by at least an extra factor of a relative to the case we treat.

Our assumption implies that, when we consider $I_{\text{MC}}(O, x_{\text{M}}(x))$ for $O \to O_s$, the energy x of the photon emitted at NLO is effectively a function of O:

$$x \equiv x(O) = \kappa(O - O_s) + \mathcal{O}\left((O - O_s)^2\right), \tag{B.26}$$

where we used the fact that $x(O_s) = 0$. We can use eq. (B.26) to compute the NLO integrated cross section. We have

$$\begin{aligned}\overline{\sigma}_{\text{NLO}}(O) &\equiv \int_{O_s}^{O} d\overline{O} \left(\frac{d\sigma}{d\overline{O}}\right)_{\text{NLO}} & \text{(B.27)}\\ &= \int_0^1 dx \left[\int_{O_s}^{O} d\overline{O}\delta(\overline{O} - O(x))\frac{aR(x)}{x} + \int_{O_s}^{O} d\overline{O}\delta(\overline{O} - O(0))\left(B + aV - \frac{aB}{x}\right)\right],\end{aligned}$$

where we have used eq. (3.13). With eq. (B.26) we obtain

$$\overline{\sigma}_{\text{NLO}}(O) = B + aV + \int_0^1 dx \left[\frac{aR(x)}{x}\Theta(\kappa(O - O_s) - x) - \frac{aB}{x}\right]. \tag{B.28}$$

Consistently with eq. (3.4), we assume that

$$R(x) = B + \sum_{n=1}^{\infty} R_n x^n \,, \tag{B.29}$$

and thus eq. (B.28) gives

$$\overline{\sigma}_{\rm NLO}(O) = B + aB \log\left(\kappa(O - O_s)\right) + aV + a \sum_{n=1}^{\infty} \frac{R_n}{n} \left(\kappa(O - O_s)\right)^n . \tag{B.30}$$

The first two terms on the right-hand side are the first two terms in the LL tower that is resummed by MC means. The third and fourth terms are NLL, the latter being also power suppressed. We point out that since the only large logarithm in the cross section has to be $\log(O - O_s)$, the coefficient κ must be of order unity.

We now consider the case of the MC@NLO. We denote the normalized pure MC result as follows:

$$\overline{\sigma}_{\rm MC}(O) \equiv \int_{O_s}^{O} d\overline{O} I_{\rm MC}(\overline{O}, 1) \,. \tag{B.31}$$

We also notice that, owing to the fact that $x_{\rm M}(0) = 1$, and assuming that $x_{\rm M}(x)$ has a power expansion in x around $x = 0$, we have (for $O \simeq O_s$)

$$\int_{O_s}^{O} d\overline{O} I_{\rm MC}(\overline{O}, x_{\rm M}(x)) = \int_{O_s}^{O} d\overline{O} \left(I_{\rm MC}(\overline{O}, 1) + {\rm NLL}\right) \Theta(\kappa(O - O_s) - x) \,. \tag{B.32}$$

We can now repeat the same computation that led us to eq. (B.30), for the MC@NLO, eq. (3.20). In the case of $\mathbb{S}$ events, analogously to what happens for the contribution of the 0-body events in eq. (B.27), the restricted range in O does not constrain the integration range in x. Thus

$$\begin{aligned} \overline{\sigma}_{\mathbb{S}}(O) &\equiv \int_{O_s}^{O} d\overline{O} \frac{d\sigma_{\mathbb{S}}}{d\overline{O}} \\ &= \overline{\sigma}_{\rm MC}(O) \left[B + aV + aB \int_0^1 dx \frac{Q(x) - 1}{x} \right] ; \end{aligned} \tag{B.33}$$

we therefore recover the pure MC result, up to $\mathcal{O}(a)$ terms in the normalization. On the other hand, the restricted range in O *does* constrain the integration range in x in the case of $\mathbb{H}$ events, as shown in eq. (B.32). We obtain

$$\begin{aligned} \overline{\sigma}_{\mathbb{H}}(O) &\equiv \int_{O_s}^{O} d\overline{O} \frac{d\sigma_{\mathbb{H}}}{d\overline{O}} \\ &= a\left(\overline{\sigma}_{\rm MC}(O) + {\rm NLL}\right) \sum_{n=1}^{\infty} \frac{R_n - BQ_n}{n} \left(\kappa(O - O_s)\right)^n , \end{aligned} \tag{B.34}$$

where Q_n are the coefficients of the expansion of $Q(x)$ around $x = 0$. Therefore, eq. (B.34) does not contribute to the large-logarithm resummation for $O \to O_s$. Taking into account eq. (B.33), this allows us to conclude that the MC@NLO resums large logarithms in the same way as the ordinary MC does.

It is clear that the argument given above works in the same way in the QCD case. One has to pay attention to the fact that, in QCD, soft and collinear singularities exist. In general, the restricted range in O constrains the emissions to be either soft *or* collinear, not necessarily soft *and* collinear. This implies that in QCD each term of the power series that appears in eq. (B.34) can be multiplied by a power of $\log(O - O_s)$. However, this does not change our conclusion, that $\mathbb{H}$ events do not contribute large logarithms to the physical cross section.

C. MC@NLO based upon slicing method

In this appendix, we comment on the approach of refs. [14, 15], where a formalism is presented to construct an MC@NLO using the slicing method at the NLO level. This formalism is called Φ-space Veto in ref. [15]. It has the main objective and virtue of ensuring that only positive event weights are generated. However, we argue that this is at the expense of being able to recover the NLO results upon expansion in $\alpha_{\rm s}$ (this is because the resulting MC@NLO does *not* have an expansion in $\alpha_{\rm s}$), and a smooth matching between different kinematic regions.

In the course of the discussion, we use the toy model, which is sufficient to highlight the main features of the method. We also propose a formula for an MC@NLO based upon the slicing method, which does not have any of the problems that affect the Φ-space Veto method, and fulfills the definition of MC@NLO given in section 2.1 (but as a consequence is not guaranteed to generate positive event weights).

The Φ-space Veto relies on the result for the NLO cross section obtained with the slicing method, eq. (3.9). The first step consists in fixing the parameter δ, by imposing the condition that eq. (3.9) does not contain any contribution proportional to $O(0)$. In other words, one seeks δ_0 such that[8]

$$B + a\,(B \log \delta_0 + V) = 0\,. \tag{C.1}$$

Thus

$$\log \delta_0 = -\frac{B + aV}{aB} \qquad \Longleftrightarrow \qquad \delta_0 = \exp\left(-\frac{B + aV}{aB}\right). \tag{C.2}$$

With such a choice, eq. (3.9) becomes

$$\langle O \rangle_{\rm slice} = a \int_{\delta_0}^{1} dx\, \frac{O(x)R(x)}{x} + \mathcal{O}(\delta_0)\,. \tag{C.3}$$

Given the fact that the parameter δ_0 is not free any longer, being related to the Born and virtual contributions as shown in eq. (C.2), the uncontrolled $\mathcal{O}(\delta_0)$ terms in eq. (C.3) may be non-negligible. An improvement of eq. (C.3) is proposed in ref. [14], which uses a subtraction method in a restricted part of the phase space (this procedure is called "hybrid subtraction"): however, this improvement is of no importance in what follows, and will be neglected here (also, it is not used in current implementations of refs. [14, 15]). As

[8]This idea was first introduced, in a slightly different form, in ref. [12].

anticipated, the reason for fixing the slicing parameter δ can be seen in eq. (C.3); this equation does *not* contain any negative contributions, and this fact is used in refs. [14, 15] to define an MC@NLO with no negative-weight events. However, the resulting δ_0 is not analytic around $a = 0$, which seems to us to cause problems, as will be discussed below.

The next step in the Φ-space Veto method is the introduction of a free parameter $\delta_{\rm PS}$, which fulfills the condition $\delta_0 < \delta_{\rm PS} < 1$. In ref. [15], it is suggested that one set $\delta_{\rm PS}$ equal to the boundary of the dead zone, in the case that HERWIG is used to construct the MC@NLO. In general, it seems difficult to prove that this condition is consistent with the requirement $\delta_0 < \delta_{\rm PS}$, especially since the boundary of the dead zone has a different (process dependent) shape from the region removed by slicing, and may depend on different variables.

In the language of the present paper, the introduction of $\delta_{\rm PS}$ implies the definition of a modified slicing formula:

$$\langle O\rangle_{\rm mslice} = a\int_0^1 dx\left[O(x)\frac{R(x)}{x}\Theta(x-\delta_{\rm PS}) + O(0)\frac{R(x)}{x}\Theta(x-\delta_0)\Theta(\delta_{\rm PS}-x)\right]. \tag{C.4}$$

From this equation, we can construct a toy MC@NLO as explained in section 3.3, paying attention to the fact that, in the Φ-space Veto method, the maximum energy available to the photon in the first branching is required to be $\delta_{\rm PS}$ (instead of 1) for class $\mathbb{S}$ events. Therefore

$$\left(\frac{d\sigma}{dO}\right)_{\rm mslice} = a\int_0^1 dx\Bigg[I_{\rm MC}(O,x_{\rm M}(x))\frac{R(x)}{x}\Theta(x-\delta_{\rm PS}) + + I_{\rm MC}(O,\delta_{\rm PS})\frac{R(x)}{x}\Theta(x-\delta_0)\Theta(\delta_{\rm PS}-x)\Bigg]. \tag{C.5}$$

Eq. (C.5) can now be used to compute the predictions of the Φ-space Veto method, in the same way that eq. (3.20) was used for our MC@NLO. We consider here the exclusive observable y, introduced in section 3.4. We can substitute eq. (3.26) directly into eq. (C.5); on the other hand, we need to use

$$I_{\rm MC}(y,\delta_{\rm PS}) = a\frac{Q(y)}{y}\Delta(y,\delta_{\rm PS})\Theta(\delta_{\rm PS}-y) \tag{C.6}$$

instead of eq. (3.25). We find the following:

$$\left(\frac{d\sigma_{\mathbb{S}}}{dy}\right)_{\rm mslice} = a^2\frac{Q(y)}{y}\Delta(y,\delta_{\rm PS})\Theta(\delta_{\rm PS}-y)\int_{\delta_0}^{\delta_{\rm PS}} dx\,\frac{R(x)}{x}. \tag{C.7}$$

We also find:

- for $y < x_e$:

$$\left(\frac{d\sigma_{\mathbb{H}}}{dy}\right)_{\rm mslice} = \left[a\frac{R(y)}{y}\Delta(y,x_{\rm M}(y)) + a^2\frac{Q(y)}{y}\int_{\delta_{\rm PS}}^{y}\frac{dx}{x}R(x)\Delta(y,x_{\rm M}(x))\right]\Theta(y-\delta_{\rm PS})\,; \tag{C.8}$$

- for $y > x_e$:

$$\left(\frac{d\sigma_{\mathbb{H}}}{dy}\right)_{\text{mslice}} = a\frac{R(y)}{y}\Theta(y-\delta_{\text{PS}})+a^2\frac{Q(y)}{y}\Theta(x_{\text{M}}^{-1}(y)-\delta_{\text{PS}})\int_{\delta_{\text{PS}}}^{x_{\text{M}}^{-1}(y)}\frac{dx}{x}R(x)\Delta(y,x_{\text{M}}(x)) \,. \tag{C.9}$$

At $y \to 0$, we expect the result of the ordinary MC, eq. (3.24), to be reproduced, up to terms of higher logarithmic order, and up to multiplicative factors of higher order in a, i.e., we expect an equation similar to (3.35) to hold. Eq. (C.7) appears to fail to achieve this result, since it starts at $\mathcal{O}(a^2)$. However, δ_0 also depends upon a; thus, in order to reach firm conclusions, the integral in eq. (C.7) needs to be evaluated explicitly. To do that, we assume again that $R(x)$ has a power series expansion in x, eq. (B.29). Eq. (C.7) then becomes

$$\left(\frac{d\sigma_{\mathbb{S}}}{dy}\right)_{\text{mslice}} = a^2\frac{Q(y)}{y}\Delta(y,\delta_{\text{PS}})\Theta(\delta_{\text{PS}}-y)\left[B\log\frac{\delta_{\text{PS}}}{\delta_0}+\sum_{n=1}^{\infty}\frac{R_n}{n}\left(\delta_{\text{PS}}^n-\delta_0^n\right)\right]. \tag{C.10}$$

Using eq. (C.2) we now find

$$\begin{aligned}\left(\frac{d\sigma_{\mathbb{S}}}{dy}\right)_{\text{mslice}} = a\frac{Q(y)}{y}\Delta(y,\delta_{\text{PS}})\Theta(\delta_{\text{PS}}-y)\Bigg[&B+aV+aB\log\delta_{\text{PS}}+ \\ &+a\sum_{n=1}^{\infty}\frac{R_n}{n}\left(\delta_{\text{PS}}^n-\exp\left(-n\frac{B+aV}{aB}\right)\right)\Bigg].\end{aligned} \tag{C.11}$$

This equation has the same functional form in y as eq. (3.24), up to uncontrolled NLL terms (due to the fact that the second argument of the Sudakov form factor is δ_{PS} instead of 1), and *formally* starts at $\mathcal{O}(a)$. However, it has the rather unpleasant feature that it cannot be expanded as a power series in a, since it is not analytic at $a = 0$. We therefore have the paradoxical situation that a quantity which is based on perturbation theory cannot be expanded in powers of the coupling constant. This is in fact due to the form of δ_0, eq. (C.2). Still, if we assume that $a^2\delta_0^n$ is of $\mathcal{O}(a^2)$, then eq. (C.11) has the correct Sudakov behaviour for $y \to 0$ (this is also formally consistent with the necessity of neglecting powers of δ_0 in the slicing procedure; however, we believe that this is not justified here, since δ_0 is not a free parameter any longer). However, with the same assumption we conclude that the Φ-space Veto method has double counting problems, since for $y < \delta_{\text{PS}}$ the result of eq. (C.11) does not coincide with the pure NLO result. On the other hand, for $y > \delta_{\text{PS}}$ the MC@NLO and NLO results agree up to $\mathcal{O}(a^2)$ terms (see eqs. (C.8) and (C.9)). Therefore, the MC@NLO prediction has a discontinuity at $y = \delta_{\text{PS}}$.

We now define an MC@NLO based upon the slicing method that does not have the problems outlined above. We can easily accomplish this task by noting that the slicing formulae can be obtained, at the NLO level, by making suitable approximations in the subtraction formulae. We start from eq. (3.9), which we re-write as follows

$$\langle O\rangle_{\text{slice}} = \int_{\delta}^{1}dx\left[O(x)\frac{aR(x)}{x}+O(0)\left(B+aV+aB\log\delta\right)\right]. \tag{C.12}$$

From now on, we do not write explicitly the terms of order $\mathcal{O}(\delta)$ that are always neglected in slicing formulae. Notice that eqs. (3.9) and (C.12) in fact coincide up to $\mathcal{O}(\delta)$ terms; the two equations would be exactly identical had we multiplied $O(0)$ by $1/(1-\delta)$. Using eq. (C.12) we now define an MC@NLO, following the procedure described in section 3.3. We obtain

$$\left(\frac{d\sigma}{dO}\right)_{\rm mslice} = \int_\delta^1 dx \left[I_{\rm MC}(O, x_{\rm M}(x)) \frac{a[R(x) - BQ(x)]}{x} + \right. \\ \left. + I_{\rm MC}(O,1) \left(B + aV + aB \log\delta + aB \frac{Q(x)}{x} \right) \right]. \tag{C.13}$$

We stress that this formula has the same features as a "standard" slicing formula, such as eq. (C.12): the parameter δ is free (and small), the integration over x does *not* extend down to zero, and the counterterm of the subtraction method is replaced by a term proportional to $\log\delta$. There is, however, one important difference. In eq. (C.12), the weights attached to $O(x)$ and $O(0)$ are divergent when $\delta \to 0$ (and can now be negative). This does not happen in eq. (C.13) for the weights attached to $I_{\rm MC}(O, x_{\rm M}(x))$ and $I_{\rm MC}(O,1)$. In the latter term, however, this property holds only if the integration over the phase space is performed before the limit $\delta \to 0$ is carried out. This is inconvenient, because the unweighting procedure is performed in practice *before* the integration. In the spirit of the slicing method, we should integrate the term $aBQ(x)/x$ in the weight attached to $I_{\rm MC}(O,1)$ analytically, which may or may not be possible, depending upon the form of $Q(x)$. In order to simplify the discussion, we assume

$$Q(x) = \Theta(x_{\rm dead} - x). \tag{C.14}$$

Using this form in eq. (C.13) we obtain

$$\left(\frac{d\sigma}{dO}\right)_{\rm mslice} = \int_\delta^1 dx\, I_{\rm MC}(O, x_{\rm M}(x)) \frac{a[R(x) - B\Theta(x_{\rm dead} - x)]}{x} + \\ + I_{\rm MC}(O,1) \left(B + aV + aB \log x_{\rm dead} \right). \tag{C.15}$$

In general, the weights attached to $I_{\rm MC}(O, x_{\rm M}(x))$ and $I_{\rm MC}(O,1)$ in this equation are not positive definite. However, it is clear that unweighting (in the sense of generating events with weights of equal *magnitude*) can easily be achieved, since all the cancellations between large numbers occur under the integral sign. In this sense, the MC@NLO based upon the slicing method is similar to the MC@NLO based upon the subtraction method, eq. (3.20). However, while in eq. (3.20) the cancellation is achieved by numerical methods, the generalization of eq. (C.15) can be obtained from eq. (C.13) only through an analytical integration. This is similar to what happens at the NLO level; in the subtraction method, the residue of the soft/collinear singularity is integrated numerically, whereas in the slicing method an analytical integration is performed.

In the general case in which $Q(x)$ does not give an easy analytical integral, we can use a mixture of analytical and numerical methods. Namely, we use the identity

$$Q(x) = \Theta(x_{\rm th} - x) + Q(x) - \Theta(x_{\rm th} - x) \tag{C.16}$$

in the last term of eq. (C.13), where $x_{\rm th}$ is an arbitrary constant. We readily obtain

$$\left(\frac{d\sigma}{dO}\right)_{\rm mslice} = \int_\delta^1 dx \Bigg[I_{\rm MC}(O, x_{\rm M}(x)) \frac{a[R(x) - BQ(x)]}{x} + \\ + I_{\rm MC}(O,1) \left(B + aV + aB \log x_{\rm th} + aB \frac{Q(x) - \Theta(x_{\rm th} - x)}{x} \right) \Bigg]. \tag{C.17}$$

We can now prove that eq. (C.13) has no double counting, and resums soft logarithms in the same way as standard MC does. We follow the procedures introduced in appendix B.3 and B.4. Using eqs. (B.13) and (B.15), we get

$$\left(\frac{d\sigma}{dO}\right)_{\rm mslice} = \int_\delta^1 dx \left[\delta(O - O(x)) \frac{aR(x)}{x} + \delta(O - O(0)) \left(B + aV + aB \log \delta \right) \right] + \\ + aB \int_0^1 dx \frac{Q(x)}{x} \Big[\delta(O - O(0)) - \delta(O - O(x)) \Big] \Big[\Theta(x - \delta) - \Theta(x - x_0) \Big] + \\ + \mathcal{O}(a^2), \tag{C.18}$$

which has to be compared to eq. (B.17). As in that case, we recover the NLO result by considering $\int dO\, O(d\sigma/dO)$, up to $\mathcal{O}(a^2)$ and power-suppressed terms. Notice that we also have terms of $\mathcal{O}(\delta)$ in eq. (C.18), consistently with the approximation upon which the slicing method is based. The distribution of the photon energies can be obtained as explained in appendix B.4. We get

$$\left(\frac{d\sigma}{dz}\right)_{\rm mslice} = S(z,a) \left[\sigma_{\rm tot} + a \frac{R(z) - BQ(z)}{zS(z,a)} + aB \int_{\max(\delta, x_{\rm M}^{-1}(z))}^1 dx \frac{Q(x)}{x} \right], \tag{C.19}$$

which is the analogue of eq. (B.24). It is apparent that the same conclusions as in appendix B.4 apply here for $z \to 0$.

Finally, notice that, by using eq. (B.11), we obtain, in the case of an exclusive variable

$$\int dO \left(\frac{d\sigma}{dO}\right)_{\rm mslice} = \int_\delta^1 dx \left[B + aV + aB \log \delta + \frac{aR(x)}{x} \right] \equiv \sigma_{\rm tot}, \tag{C.20}$$

which is the result that we also find in appendix B.2. We conclude that eq. (C.13) defines an MC@NLO that has the same formal properties as that defined in eq. (3.20).

References

[1] G. Marchesini, B.R. Webber, G. Abbiendi, I.G. Knowles, M.H. Seymour and L. Stanco, *HERWIG: a Monte Carlo event generator for simulating hadron emission reactions with interfering gluons. Version 5.1 - april 1991*, *Comput. Phys. Commun.* **67** (1992) 465; G. Corcella, I.G. Knowles, G. Marchesini, S. Moretti, K. Odagiri, P. Richardson, M.H. Seymour and B.R. Webber, *HERWIG 6: an event generator for hadron emission reactions with interfering gluons (including supersymmetric processes)*, *J. High Energy Phys.* **01** (2001) 010 [hep-ph/0011363]; G. Corcella et al., *HERWIG 6.3 release note*, hep-ph/0107071.

[2] T. Sjöstrand, P. Eden, C. Friberg, L. Lonnblad, G. Miu, S. Mrenna and E. Norrbin, *High-energy-physics event generation with PYTHIA 6.1*, *Comput. Phys. Commun.* **135** (2001) 238 [hep-ph/0010017].

[3] T. Sjöstrand, *The Lund Monte Carlo for jet fragmentation and e^+e^- physics: Jetset version 6.2*, *Comput. Phys. Commun.* **39** (1986) 347.

[4] G. Marchesini and B.R. Webber, *Monte Carlo simulation of general hard processes with coherent QCD radiation*, *Nucl. Phys.* **B 310** (1988) 461.

[5] L. Lonnblad, *Ariadne version 4: a program for simulation of QCD cascades implementing the color dipole model*, *Comput. Phys. Commun.* **71** (1992) 15.

[6] M.H. Seymour, *Matrix element corrections to parton shower algorithms*, *Comput. Phys. Commun.* **90** (1995) 95 [hep-ph/9410414]; *A simple prescription for first order corrections to quark scattering and annihilation processes*, *Nucl. Phys.* **B 436** (1995) 443 [hep-ph/9410244].

[7] J. Andre and T. Sjöstrand, *A matching of matrix elements and parton showers*, *Phys. Rev.* **D 57** (1998) 5767 [hep-ph/9708390]; *A matching of matrix elements and parton showers*, *Phys. Rev.* **D 57** (1998) 5767 [hep-ph/9708390]; G. Miu and T. Sjöstrand, *W production in an improved parton shower approach*, *Phys. Lett.* **B 449** (1999) 313 [hep-ph/9812455].

[8] G. Corcella and M.H. Seymour, *Matrix element corrections to parton shower simulations of heavy quark decay*, *Phys. Lett.* **B 442** (1998) 417 [hep-ph/9809451]; *Initial state radiation in simulations of vector boson production at hadron colliders*, *Nucl. Phys.* **B 565** (2000) 227 [hep-ph/9908388].

[9] S. Mrenna, *Higher order corrections to parton showering from resummation calculations*, hep-ph/9902471.

[10] C. Friberg and T. Sjöstrand, *Some thoughts on how to match leading log parton showers with nlo matrix elements*, in proceedings of Workshop on *Monte Carlo generators for HERA physics*, T.A. Doyle, G. Grindhammer, G. Ingelman and H. Jung eds., DESY, Hamburg, 1999, p. 181. [hep-ph/9906316].

[11] J.C. Collins, *Subtraction method for nlo corrections in Monte Carlo event generators for leptoproduction*, *J. High Energy Phys.* **05** (2000) 004 [hep-ph/0001040].

[12] H. Baer and M.H. Reno, *W and Z production at $p\bar{p}$ colliders: parton showers merged with $O(\alpha_s)$ Monte Carlo approach*, *Phys. Rev.* **D 45** (1992) 1503; *A complete $O(\alpha_s)$ event generator for $p\bar{p} \to W^+x \to e^+$ neutrino X with parton showering*, *Phys. Rev.* **D 44** (1991) 3375.

[13] H. Baer and M.H. Reno, *Multiple parton emission effects in next-to-leading order direct photon production*, *Phys. Rev.* **D 54** (1996) 2017 [hep-ph/9603209].

[14] B. Potter, *Combining QCD matrix elements at next-to-leading order with parton showers*, *Phys. Rev.* **D 63** (2001) 114017 [hep-ph/0007172];
B. Potter and T. Schorner, *Combining parton showers with next-to-leading order QCD matrix elements in deep-inelastic ep scattering*, *Phys. Lett.* **B 517** (2001) 86 [hep-ph/0104261].

[15] M. Dobbs, *Incorporating next-to-leading order matrix elements for hadronic diboson production in showering event generators*, *Phys. Rev.* **D 64** (2001) 034016 [hep-ph/0103174];

JHEP06(2002)029

M. Dobbs, *Phase space veto method for next-to-leading order event generators in hadronic collisions*, *Phys. Rev.* **D 65** (2002) 094011 [hep-ph/0111234].

[16] M.L. Mangano, M. Moretti and R. Pittau, *Multijet matrix elements and shower evolution in hadronic collisions: $WB\bar{B}+N$ jets as a case study*, *Nucl. Phys.* **B 632** (2002) 343 [hep-ph/0108069].

[17] S. Catani, F. Krauss, R. Kuhn and B.R. Webber, *QCD matrix elements + parton showers*, *J. High Energy Phys.* **11** (2001) 063 [hep-ph/0109231].

[18] Y. Chen, J.C. Collins and N. Tkachuk, *Subtraction method for NLO corrections in Monte-Carlo event generators for Z boson production*, *J. High Energy Phys.* **06** (2001) 015 [hep-ph/0105291].

[19] J.C. Collins, *Monte-Carlo event generators at NLO*, *Phys. Rev.* **D 65** (2002) 094016 [hep-ph/0110113].

[20] Y.-J. Chen, J.C. Collins and X.-M. Zu, *NLO corrections in MC event generator for angular distribution of Drell-Yan lepton pair production*, *J. High Energy Phys.* **04** (2002) 041 [hep-ph/0110257];
J.C. Collins and X.-M. Zu, *Parton distribution functions suitable for Monte-Carlo event generators*, hep-ph/0204127.

[21] L. Lönnblad, *Correcting the colour-dipole cascade model with fixed order matrix elements*, *J. High Energy Phys.* **05** (2002) 046 [hep-ph/0112284].

[22] S. Frixione, *A next-to-leading order calculation of the cross-section for the production of W^+W^- pairs in hadronic collisions*, *Nucl. Phys.* **B 410** (1993) 280.

[23] M.L. Mangano, P. Nason and G. Ridolfi, *Heavy quark correlations in hadron collisions at next-to-leading order*, *Nucl. Phys.* **B 373** (1992) 295.

[24] S. Kawabata, *A new version of the multidimensional integration and event generation package bases/spring*, *Comput. Phys. Commun.* **88** (1995) 309.

[25] G.P. Lepage, *A new algorithm for adaptive multidimensional integration*, *J. Comput. Phys.* **27** (1978) 192.

[26] M. Cacciari, M. Greco and P. Nason, *The p_T spectrum in heavy-flavour hadroproduction*, *J. High Energy Phys.* **05** (1998) 007 [hep-ph/9803400].

[27] M. Cacciari, S. Frixione and P. Nason, *The p_T spectrum in heavy-flavor photoproduction*, *J. High Energy Phys.* **03** (2001) 006 [hep-ph/0102134].

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

[29] A.D. Martin, R.G. Roberts, W.J. Stirling and R.S. Thorne, *Parton distributions and the LHC: W and Z production*, *Eur. Phys. J.* **C 14** (2000) 133 [hep-ph/9907231].

[30] K. Odagiri, *Color connection structure of (supersymmetric) QCD ($2 \to 2$) processes*, *J. High Energy Phys.* **10** (1998) 006 [hep-ph/9806531].

JHEP06(2002)029