Stefano Frixione

MC@NLO: a tutorial

Physics at TeV Colliders, Les Houches, 6/5/2005

SF & Bryan Webber, JHEP 0206(2002)029 [hep-ph/0204244]
SF, Paolo Nason & Bryan Webber, JHEP 0308(2003)007 [hep-ph/0305252]

What is it?

- MC@NLO is a Parton Shower Monte Carlo which works just like any other PSMC: it outputs events

- The *defining feature*: partonic hard subprocesses are computed by including the full NLO QCD corrections

Why NLO corrections?

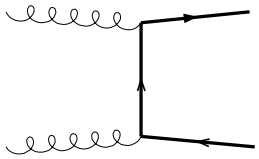
Among the many good reasons, let me mention those that are likely to have the largest impact on phenomenology

- Provides the *only* way to sensibly compute the K factors event by event, and thus to use this information in detector simulation – this is impossible with NLO parton-level codes
- The hardest p_T emission is computed *exactly*, and is in agreement with the NLO matrix element result – the correct NLO normalization is obtained upon integration over the visible spectrum
- The scale dependence of physical observables can be computed – this procedure is either meaningless or very difficult to perform with standard Monte Carlos

MC@NLO includes dynamic features that cannot be present in standard MC's – heavy flavour physics is a major example

Charm and bottom with standard MC's

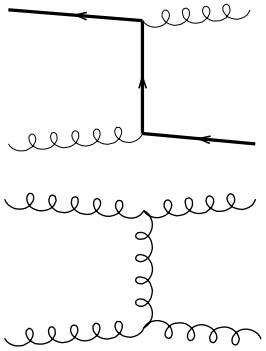
MC rule: if we aim to study any physical system, we start by producing it in the hard process $\Longrightarrow$



Flavour CReation

This is going to underestimate the rate by a factor of 4 (which is not so important), and to miss key kinematic features (which is crucial – see R. Field)

So break the rule and add other hard processes

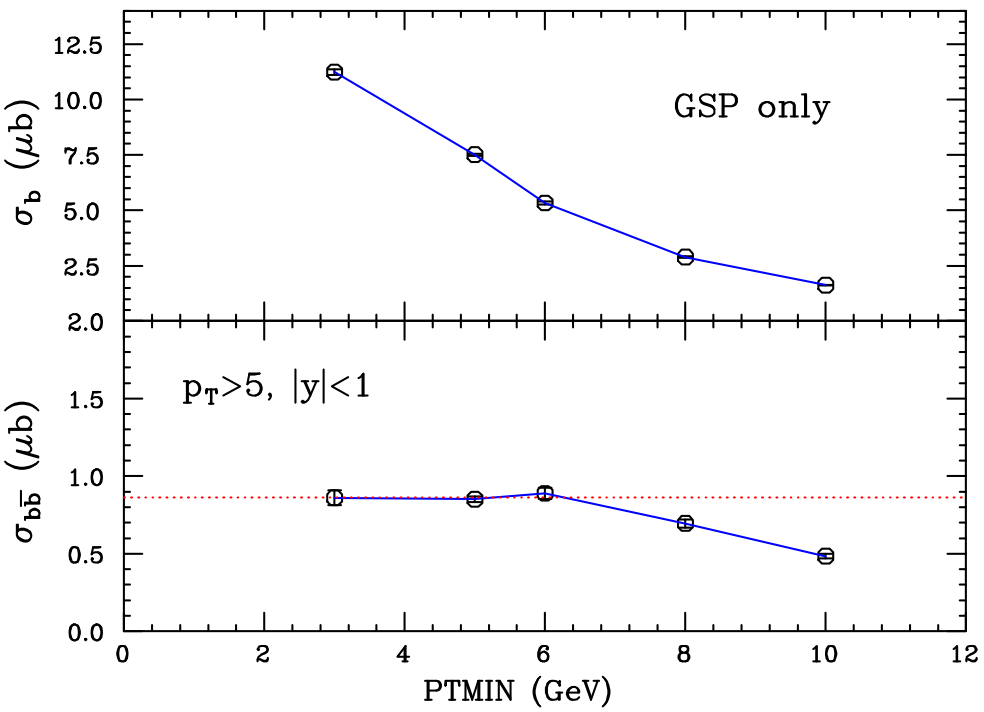


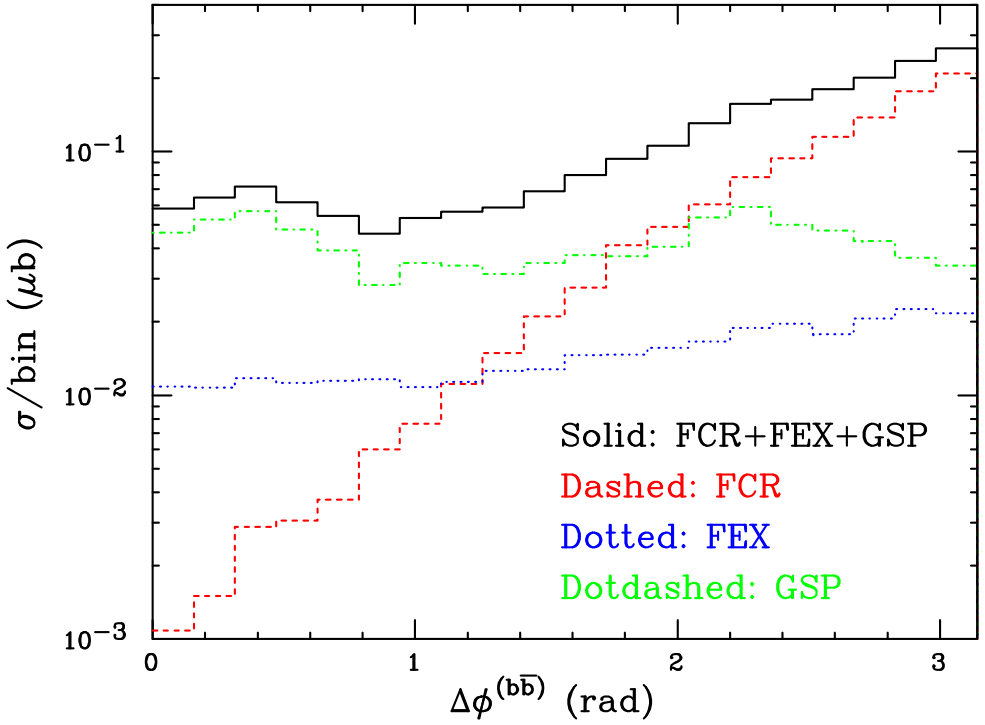
Flavour EXcitation

Gluon SPlitting

- In FEX, the missing Q or $\overline{Q}$ results from initial-state radiation. A cutoff PTMIN avoids divergences in the matrix element
- In GSP, the Q and $\overline{Q}$ result from final-state gluon splitting. PTMIN is again necessary to obtain finite results

b production with HERWIG



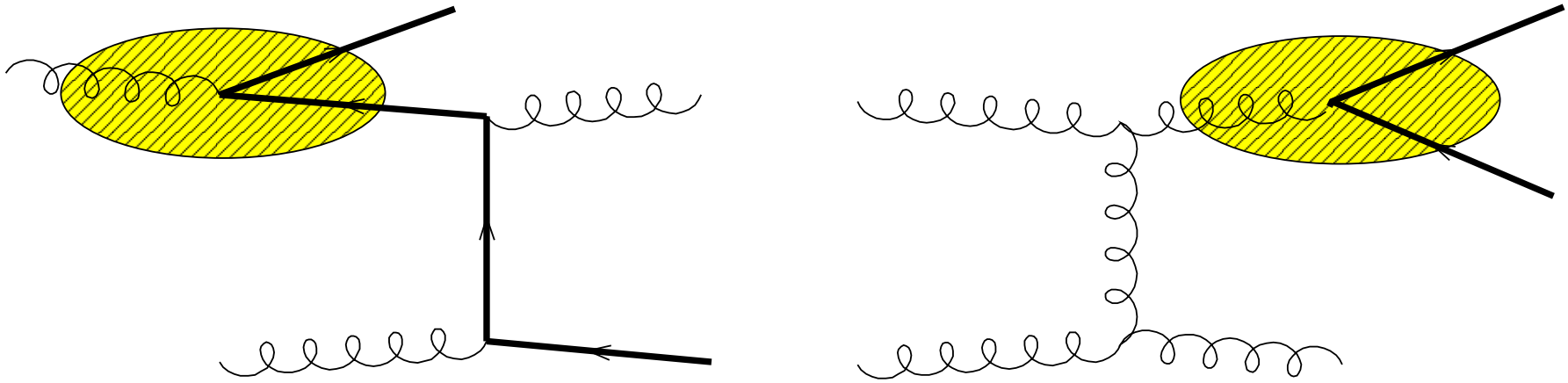


- The PTMIN dependence is worrisome in the case of single-inclusive observables
- FCR, FEX and GSP are complementary, and all must be generated
- GSP efficiency is extremely poor: 10^{-4} within cuts for correlations

Reliability and efficiency rapidly degrade for smaller p_T cuts. In FEX, the dependence on bottom PDF is problematic. No standard MC can work for $p_T \simeq 0$

In MC@NLO...

FEX and GSP are "subdiagrams" of real-emission diagrams



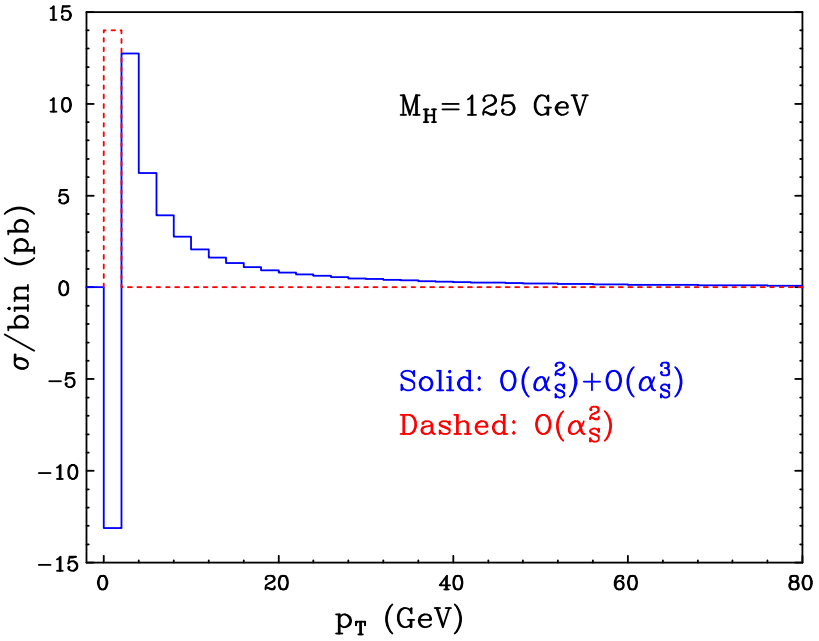
which has non-trivial consequences

- There are no divergences (since they are cancelled by the virtual contribution)
- It is QCD that dictates the right mixing of FCR, FEX, and GSP
- No need to throw events away as in GSP
- There is no dependence on the largely unknown b density

Last but not least, you get a much better prediction for rates

What does NLO mean?

Consider Higgs production:



$$\frac{d\sigma}{dp_T} = \left(A\alpha_S^2 + B\alpha_S^3\right)\delta(p_T) + C(p_T)\alpha_S^3$$

$$\int_{p_T^{min}}^{\infty} dp_T \frac{d\sigma}{dp_T} = \mathcal{C}_3\alpha_S^3, \qquad p_T^{min} > 0$$

$$= \mathcal{D}_2\alpha_S^2 + \mathcal{D}_3\alpha_S^3, \quad p_T^{min} = 0$$

$p_T^{min} > 0 \Rightarrow$ LO, $p_T^{min} = 0 \Rightarrow$ NLO

The answer depends on the observable, and even on the kinematic range considered.

So this definition cannot be adopted in the context of event generators

- N^{k}LO accuracy in event generators is defined by the number k of extra gluons (either virtual or real) wrt the LO contribution (hopefully we all agree on LO definition)

Installing MC@NLO I

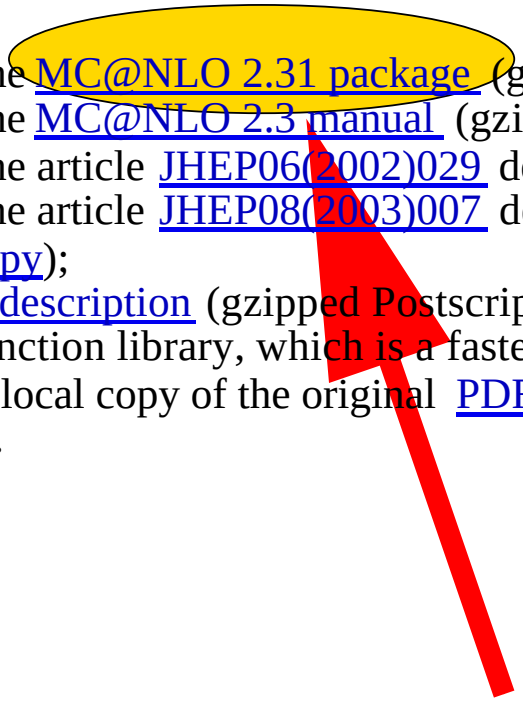
Open the web page

`http://www.hep.phy.cam.ac.uk/theory/webber/MCatNLO`

and look for the following lines

Here you can find

- The MC@NLO 2.31 package (gzipped tar file);
- The MC@NLO 2.3 manual (gzipped Postscript);
- The article JHEP06(2002)029 describing the MC@NLO formalism (local copy);
- The article JHEP08(2003)007 describing the application to heavy quark pair production (local copy);
- A description (gzipped Postscript) and the grid files (gzipped tar file) of our parton density function library, which is a faster version of PFDLIB with many newer PDFs.
- A local copy of the original PDFLIB manual (gzipped Postscript), which you may need to refer to.



Copy the tarball which you find following the link "MC@NLO v.nn package" into one of your local directories **(source directory)**. The tarball contains the source files of the package (`f77` and `c`)

Installing MC@NLO II

We have a self-contained PDF library. Some PDFs require external grid files, that you can download as a tarball

Here you can find

- The MC@NLO 2.31 package (gzipped tar file);
- The MC@NLO 2.3 manual (gzipped Postscript);
- The article JHEP06(2002)029 describing the MC@NLO formalism (local copy);
- The article JHEP08(2003)007 describing the application to heavy quark pair production (local copy);
- A description (gzipped Postscript) and the grid files (gzipped tar file) of our parton density function library, which is a faster version of PFDLIB with many newer PDFs.
- A local copy of the original PDFLIB manual (gzipped Postscript), which you may need to refer to.

Copy the tarball into one of your local directories **(data directory)**

You may also use PDFLIB or LHAPDF (the program will be slower)

Prior to running MC@NLO

You need to have

`bash` and `gmake` (`make` for Mac)

installed on your system (the package has been run so far on `Digital Unix`, `Linux`, `Sun Unix`, and `OSX v10` or higher). Mac appear to make life a bit harder

Then edit the files

`mcatnlo_hwan`*xxx*`.f` `mcatnlo_hwdriver.f` `mcatnlo_hwlhin.f`

and replace there

```
INCLUDE 'HERWIGXX.INC'
```

with

```
INCLUDE 'HERWIGYY.INC'
```

where `HERWIGYY.INC` is the include file relevant to the version of `HERWIG` you want to use (larger than 6.500). You must have copied `HERWIGYY.INC` into the source directory

Running MC@NLO

Edit the file

`MCatNLO.inputs`

and write there the inputs relevant to your run. For example

`ECM=14000`

tells the code you run at the LHC energy. You must always set

`HWPATH` = *address of the directory where* `HERWIG` *is*

`HERWIGVER` = *the version of* `HERWIG` *you use*

`PDFPATH` = *address of the data directory*

If you don't do this, the program will likely crash

Finally, the last line of MCatNLO.inputs should be

`runMCatNLO`

Execute MCatNLO.inputs from the `bash` shell. The scripts will compile and run the code

Running MC@NLO: results

The general scheme of MC@NLO is as follows

- NLO code: integrates and unweights the matrix elements
- Event file: a list of hard events, i.e. the kinematics configurations emerging from hard subprocesses (typically, $2 \to 2$ and $2 \to 3$)
- MC code: Herwig, which reads the hard events and showers them

When the script command runMCatNLO is executed, a subdirectory of the source directory **(running directory)** is created, whose name is

`Alpha` or `Linux` or `Sun` or `Darwin`

depending on the operating system you are running with

The NLO and MC codes are compiled and executed sequentially. The non-physical event file, and the (user defined) results of the analysis will be stored in the running directory

The NLO in MC@NLO

The NLO code prepares the $2 \to n$ and $2 \to (n+1)$ kinematic configurations that serve as initial conditions for the shower. This is done in a two-step procedure

- The (subtracted) NLO matrix elements are integrated with BASES (this is an advanced version of Vegas by GRACE). In the process, BASES finds the spikes of the cross section, and stores them in data files

- Using the information in the data files, SPRING (a version modified by us) samples efficiently the phase space, and produces the hard kinematic configurations with weights ± 1

In the processes implemented so far, the unweighting efficiency is large: 20–50%

A couple of things to remember

- The MC@NLO formalism is correct regardless of the parton shower code used. The implementation of the current version, however, is specifically designed to work with HERWIG. This information is hardwired in the code.
  If you use PYTHIA, you get wrong results

- The parton configurations corresponding to the hard subprocesses, stored in the event file, are non physical. They cannot be used in the same way as their analogues in the standard MC's codes

Mass scales

Scale uncertainties in an NLO computation are obtained by varying the renormalization (μ_R) and factorization (μ_F) scales

There's no theorem that tells you the range in which the scales can be varied. Typically, one chooses a default scale (μ_0) (a measurable quantity, whose value may change event-by-event), and then sets

$$1/2\,\mu_0 \leq \mu_R \leq 2\mu_0\,, \qquad 1/2\,\mu_0 \leq \mu_F \leq 2\mu_0$$

In MC@NLO, the default scale is fixed and is not accessible to the user (it is sort of optimized to reduce the impact of unknown higher orders). μ_R and μ_R are assigned via the script variables (in MCatNLO.inputs)

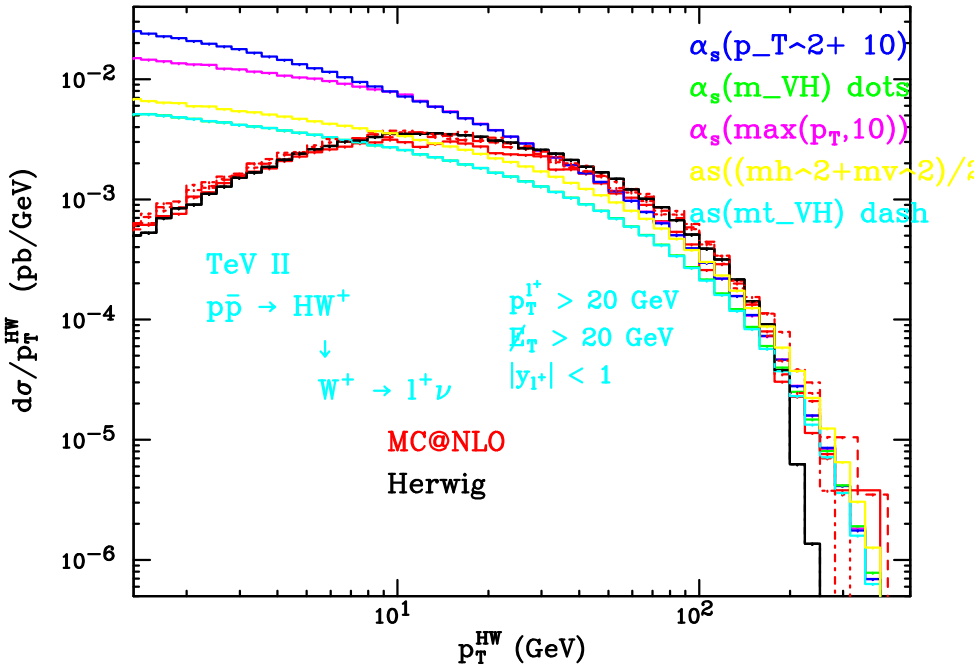
`FREN` = a number not too far from 1

`FFACT` = a number not too far from 1

which means

$$\mu_R = \texttt{FREN}\;\mu_0\,, \qquad \mu_F = \texttt{FFACT}\;\mu_0$$

Scale dependence in WH production



Plot: C. Oleari (preliminary)

- The difference between MC@NLO and NLO is not small for moderate $p_T(WH)$
- This effect has been seen elsewhere in matched computations (also with standard analytic techniques)
- May be an artifact of the scale chosen?

- For this specific observable, $m_T(WH)$ is not an ideal choice at NLO
- When many different choices of scale are explored, huge variations in NLO, no changes in MC@NLO $\Longrightarrow$ A spectacular proof of the benefits of matched computations

MC@NLO vs HERWIG: analysis

If you can run one, you can run the other. The analisys routines (HWANAL) are unchanged (except perhaps for a few particle codes that are treated in a special way in HERWIG – this mainly concerns vector bosons)

- Unweighted event generation achieved (weights: ± 1)
- Weighted event generation possible (currently not implemented)
- MC@NLO shape identical to HERWIG shape in soft/collinear regions
- MC@NLO/NLO=1 in hard regions
- There are negative-weight events

Negative weights don't mean negative cross sections. They arise from a different mechanism wrt those at the NLO, and their number is fairly limited

Negative weights

- Why are they around?

  Exact quantum mechanics computations feature interference phenomena, whose contributions don't have a definite sign. The presence of contributions of negative sign to the cross sections prevents us from having *only* $+1$ weights

- What's the difference wrt NLO?

  At the NLO, the negative-only weight distribution is divergent, while it is finite in MC@NLO. Unweighted event generation can only be achieved in MC@NLO

- Can I throw them away in MC@NLO?

  No, you can't: they are necessary in order to obtain the exact NLO results for total rates, and for differential distributions where relevant

- How do I have to use them?

  Just add -1 to *(i.e. subtract* $+1$ *from)* the histograms of physical observables. For geometric properties, treat them as you treat the positive weights

The only implication of negative weights is that you have to run a bit longer to obtain the same nominal accuracy – and in b physics you actually have to run *less*

MC@NLO 2.31 [hep-ph/0402116]

IPROC	Process
–1350–IL	$H_1H_2 \to (Z/\gamma^* \to) l_{\rm IL}\bar{l}_{\rm IL} + X$
–1360–IL	$H_1H_2 \to (Z \to) l_{\rm IL}\bar{l}_{\rm IL} + X$
–1370–IL	$H_1H_2 \to (\gamma^* \to) l_{\rm IL}\bar{l}_{\rm IL} + X$
–1460–IL	$H_1H_2 \to (W^+ \to) l^+_{\rm IL}\nu_{\rm IL} + X$
–1470–IL	$H_1H_2 \to (W^- \to) l^-_{\rm IL}\bar{\nu}_{\rm IL} + X$
–1396	$H_1H_2 \to \gamma^*(\to \sum_i f_i\bar{f}_i) + X$
–1397	$H_1H_2 \to Z^0 + X$
–1497	$H_1H_2 \to W^+ + X$
–1498	$H_1H_2 \to W^- + X$
–1600–ID	$H_1H_2 \to H^0 + X$
–1705	$H_1H_2 \to b\bar{b} + X$
–1706	$H_1H_2 \to t\bar{t} + X$
–2850	$H_1H_2 \to W^+W^- + X$
–2860	$H_1H_2 \to Z^0Z^0 + X$
–2870	$H_1H_2 \to W^+Z^0 + X$
–2880	$H_1H_2 \to W^-Z^0 + X$

- Works identically to HERWIG: the very same analysis routines can be used
- Reads shower initial conditions from an event file (as in ME corrections)
- Exploits Les Houches accord for process information and common blocks
- Features a self contained library of PDFs with old and new sets alike
- LHAPDF will also be implemented

What's going on

No major theoretical work: the MC@NLO formalism is as defined in the original paper (no need to change it – the implementation of final-state collinear singularities poses no problems, as sometimes incorrectly claimed)

We figured out a few tricks with impact on efficiency

- Alternative way of implementing spin correlations
- Cuts at the level of hard matrix elements (improved efficiency)

We made progress with the implementation of processes

- WH and ZH with full spin correlations (with C. Oleari and V. del Duca)
- Spin correlations added to W^+W^- production
- Single top at advanced stage (with E. Laenen and P. Motylinski)
- Format of hard event files will be different from v3.1 (should be irrelevant to the user, since these files are non-physical)

Adding new processes

In standard MC's, the implementation of a new process requires to:

- Compute, or get from someone else, the matrix elements
- Code them in your MC, in the appropriate format. Alternatively, code them in a standalone package, and write a routine which reads the results into your MC
- Figure out the colour and mother-daughter connections for the partons entering the hard subprocess

The very same things have to be done in MC@NLO. So the only difference between standard MC's and MC@NLO is:

- NLO matrix elements are more complicated than LO ones

This is why in general we first implement a process by neglecting the spin correlations of the decay products: spin correlations are equivalent to adding more legs to the final state, which takes time to implement

What's going on in Les Houches

We are trying to start new projects

- Jets, dijets (S. Ellis, B. Kilgore, E. Laenen, P. Motylinski)
- VBF (C. Oleari, V. del Duca)
- Anomalous TGC (A. Oh)
- CKKW to NLO? (?)

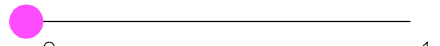
Extensive running of the code is foreseen

TECHNICAL ASPECTS

A simple way to understand MC@NLO

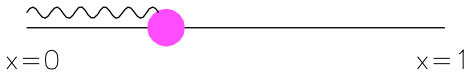
A system S moves along a line between 0 and 1. It can radiate "photons", whose energy we denote with x. S can undergo several further emissions; on the other hand, one photon cannot branch. Internal degrees of freedom of S are understood

$$\left(\frac{d\sigma}{dx}\right)_B = B\delta(x) \qquad \longleftrightarrow$$



$$\left(\frac{d\sigma}{dx}\right)_V = \alpha_S\left(\frac{B}{2\epsilon}+V\right)\delta(x) \qquad \longleftrightarrow$$

$$\left(\frac{d\sigma}{dx}\right)_R = \alpha_S\frac{R(x)}{x} \qquad \longleftrightarrow$$



where $\lim_{x\to 0} R(x) = B$ as in QCD. An NLO prediction:

$$\frac{d\sigma}{dO} = \lim_{\epsilon\to 0}\int_0^1 dx x^{-2\epsilon}\delta(O-O(S,x))\left[\left(\frac{d\sigma}{dx}\right)_B + \left(\frac{d\sigma}{dx}\right)_V + \left(\frac{d\sigma}{dx}\right)_R\right]$$

with $\lim_{x\to 0} O(S,x) = O(S,0)$ (infrared safeness). Note the kinematics:

B&V $\Longrightarrow O(S,0)$, R $\Longrightarrow O(S,x)$

The computation of the NLO cross section I

■ *SLICING*

$$\left(\frac{d\sigma}{dO}\right)_{NLOslice} = \int_{\delta}^{1} dx \Bigg\{ \delta(O - O(S,x)) \frac{\alpha_S R(x)}{x} + \delta(O - O(S,0)) \Big[B + \alpha_S \left(B \log \delta + V \right) \Big] \Bigg\}$$

■ *SUBTRACTION*

$$\left(\frac{d\sigma}{dO}\right)_{NLOsubt} = \int_{0}^{1} dx \Bigg\{ \delta(O - O(S,x)) \frac{\alpha_S R(x)}{x} + \delta(O - O(S,0)) \left(B + \alpha_S V - \frac{\alpha_S B}{x} \right) \Bigg\}$$

B&V $\Longrightarrow O(S,0)$, R $\Longrightarrow O(S,x)$

The computation of the NLO cross section II

$$\left(\frac{d\sigma}{dO}\right)_{NLOsubt} = \int_0^1 dx \Bigg\{ \delta(O - O(S,x)) \frac{\alpha_S R(x)}{x} + \delta(O - O(S,0)) \left(B + \alpha_S V - \frac{\alpha_S B}{x} \right) \Bigg\}$$

Upon integration in x, the bin of $O(S,x)$ gets a weight

$$w_{\mathbb{H}}(x) = \frac{\alpha_S R(x)}{x}$$

and the bin of $O(S,0)$ gets a weight

$$w_{\mathbb{S}}(x) = B + \alpha_S V - \frac{\alpha_S B}{x}$$

The divergence of $w_{\mathbb{H}}(x)$ and $w_{\mathbb{S}}(x)$ for $x \to 0$ is the reason for:

1) numerical instabilities
2) the impossibility of getting unweighted events in NLO computations

The toy MC

The system can undergo an arbitrary number of emissions, with probability controlled by the Sudakov form factor

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

i.e., the probability that no photon be emitted with energy $x_1 < x < x_2$. The function $Q(z)$ parametrizes beyond-LL effects, with

$$0 \leq Q(z) \leq 1, \quad \lim_{z \to 0} Q(z) = 1$$

The Born cross section

$$\left(\frac{d\sigma}{dx}\right)_B = B\delta(x)$$

gives the overall normalization (B) and initial condition ($(S, 0)$) for the shower. Apart from the trivial normalization, this can be formally embedded in the generating functional (i.e., the history of all possible showers)

$$\mathcal{F}_{\text{MC}}(S, 0)$$

NLO $\oplus$ MC $\longrightarrow$ MC@NLO?

Naive first try: use the NLO kinematic configurations as initial conditions for showers, rather than for filling the histograms

- $\delta(O - O(S,0)) \longrightarrow$ start the MC with 0 emissions: $\mathcal{F}_{\text{MC}}(S,0)$
- $\delta(O - O(S,x)) \longrightarrow$ start the MC with 1 emission at x: $\mathcal{F}_{\text{MC}}(S,x)$

$$\mathcal{F}_{\text{naive}} = \int_0^1 dx \left[\mathcal{F}_{\text{MC}}(S,x) \frac{\alpha_S R(x)}{x} + \mathcal{F}_{\text{MC}}(S,0) \left(B + \alpha_S V - \frac{\alpha_S B}{x} \right) \right]$$

It doesn't work:

- Cancellations between (S,x) and $(S,0)$ contributions occur after the shower: hopeless from the practical point of view; and, unweighting is still impossible
- $(d\sigma/dO)_{naive} - (d\sigma/dO)_{NLO} = \mathcal{O}(\alpha_S)$. In words: double counting

The problem is a fundamental one: KLN cancellation is achieved in standard MC's through unitarity, and embedded in Sudakovs. This is no longer possible: IR singularities do appear in hard ME's

MC@NLO: modified subtraction I

Get rid of the MC $\mathcal{O}(\alpha_S)$ contributions by an extra subtraction of $\mathcal{O}(\alpha_S)$

$$\mathcal{F}_{\text{MC@NLO}} = \int_0^1 dx \left[\mathcal{F}_{\text{MC}}(S,x) \frac{\alpha_S [R(x) - BQ(x)]}{x} + \mathcal{F}_{\text{MC}}(S,0) \left(B + \alpha_S V + \frac{\alpha_S B [Q(x) - 1]}{x} \right) \right]$$

where the two (one for branching, one for no-branching probability) new terms are sensibly chosen:

$$\left(\frac{d\sigma}{dx} \right)_{MC} = \alpha_S B \frac{Q(x)}{x} + \mathcal{O}(\alpha_S^2)$$

$Q(x)$ is MC-dependent (i.e., Pythia's and Herwig's differ), but $Q(x) \to 1$ for $x \to 0$ always holds

By explicit computation, $(d\sigma/dO)_{MC@NLO} - (d\sigma/dO)_{NLO} = \mathcal{O}(\alpha_S^2)$, and therefore there is no double counting

Furthermore $\longrightarrow$

MC@NLO: modified subtraction II

Let's look at the weights of $\mathcal{F}_{\mathsf{MC}}(S,x)$ and $\mathcal{F}_{\mathsf{MC}}(S,0)$

$$w_{\mathbb{H}}(x) = \frac{\alpha_S[R(x) - BQ(x)]}{x}$$

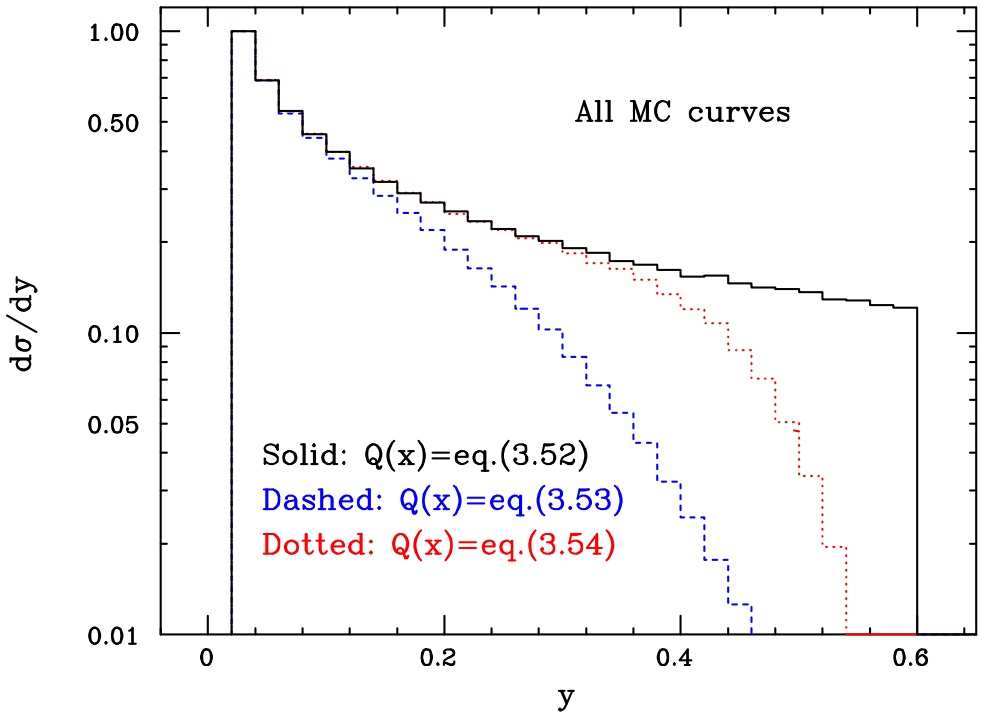
$$w_{\mathbb{S}}(x) = B + \alpha_S V + \frac{\alpha_S B[Q(x) - 1]}{x}$$

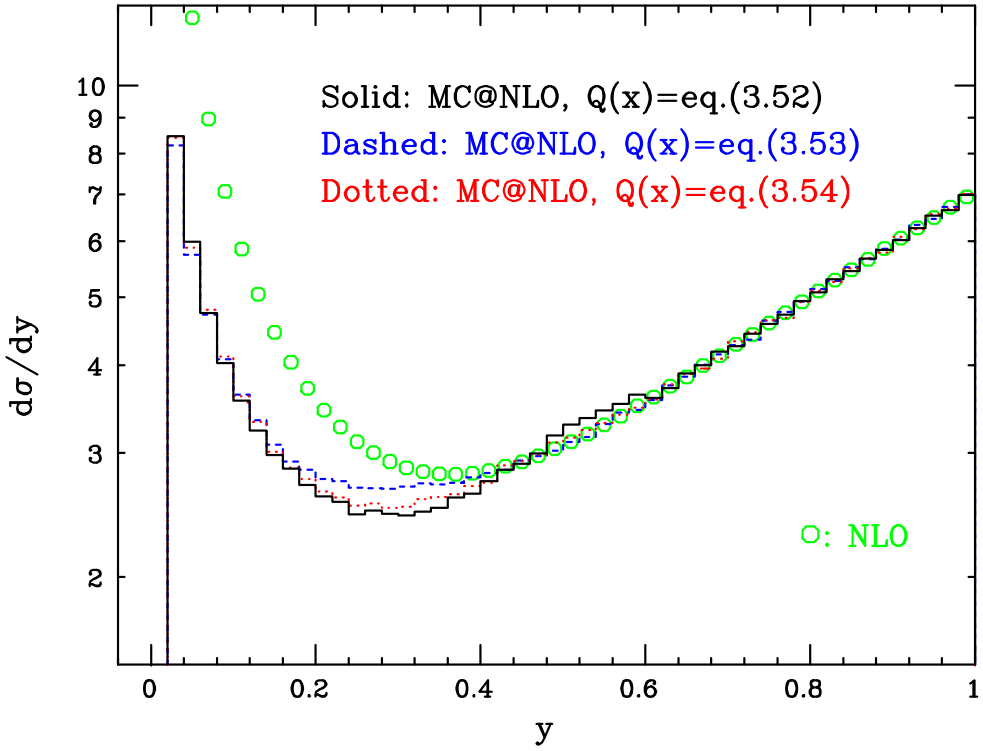
They don't diverge any longer for $x \to 0$

The MC provides local, observable-independent, counterterms $\Longrightarrow$ greater numerical stability, unweighting possible

MC@NLO can thus be minimally seen as a way to stabilize NLO computations, through the construction of a simplified MC whose only aim is to furnish the local counterterms. In this sense, the generalization to NNLO should not be too difficult

Toy model: results





1. $Q(x) = \Theta(x_{dead} - x)$;
2. $Q(x) = \Theta(x_{dead} - x) G(x/x_{dead})$, with $\alpha = 1, \beta = 1, c = 1$;
3. $Q(x) = \Theta(x_{dead} - x) G(x/x_{dead})$, with $\alpha = 2, \beta = 1, c = 8$.

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

- Very smooth transition across the dead zone border (good control beyond NLO)

Dead zone

HERWIG cannot emit in the whole region kinematically accessible for each branching. Worried?

- I think this is honest, and a blessing when matching it with NLO matrix elements

MC's are based on a collinear approximation: the harder the emissions, the larger the errors made. To forbid emissions in the hard region means to limit the size of the errors

The absence of MC emissions in the hard region implies a smaller number of negative-weight events in MC@NLO

MC@NLO: the QCD case

Strategy: follow what learned with the toy model

Toy model

$$\mathcal{F}_{\text{MC@NLO}} = \int_0^1 dx \left[\mathcal{F}_{\text{MC}}(S,x) \frac{\alpha_S [R(x) - BQ(x)]}{x} + \mathcal{F}_{\text{MC}}(S,0) \left(B + \alpha_S V + \frac{\alpha_S B[Q(x)-1]}{x} \right) \right]$$

QCD

$$\mathcal{F}_{\text{MC@NLO}} = \sum_{ab} \int dx_1\, dx_2\, d\phi_3\, f_a(x_1) f_b(x_2) \left[\mathcal{F}_{\text{MC}}^{(2\to 3)} \left(\mathcal{M}_{ab}^{(r)}(x_1,x_2,\phi_3) - \mathcal{M}_{ab}^{(\text{MC})}(x_1,x_2,\phi_3) \right) + \mathcal{F}_{\text{MC}}^{(2\to 2)} \left(\mathcal{M}_{ab}^{(b,v,c)}(x_1,x_2,\phi_2) - \mathcal{M}_{ab}^{(c.t.)}(x_1,x_2,\phi_3) + \mathcal{M}_{ab}^{(\text{MC})}(x_1,x_2,\phi_3) \right) \right]$$

Difficulties

As far as the modified subtraction is concerned, QCD is not that different from the toy model. There are however at least a couple of highly non-trivial issues

- QCD has soft *and* collinear singularities. In the case of initial state emissions, the hard $2 \to 2$ processes that factorize have *different kinematics* in the soft and collinear limits. But there is only one

$$\mathcal{F}_{\text{MC}}^{(2\to 2)}$$

  functional generator, therefore the hard configuration *must be unique*

- The computation of the MC counterterms

$$\mathcal{M}_{ab}^{(\text{MC})}(x_1, x_2, \phi_3)$$

  requires a deep knowledge of MC implementation details. The *shower variables* have to be expressed in terms of the *phase-space variables* ϕ_3 used in the NLO computation

Initial-state emissions: the problem

Take the generic $2 \to (n+1)$ configuration

$$\mathcal{M}\left\{a(p_1) + b(p_2) \longrightarrow c(k) + S_n(k1)\right\}$$

Now consider the various singular limits

- $\vec{k} \parallel \vec{p}_1 \quad \Longrightarrow \quad P_{a\to ca'} \otimes \mathcal{M}\left\{a'((1-z)p_1) + b(p_2) \longrightarrow S_n(k1)\right\}$
- $\vec{k} \parallel \vec{p}_2 \quad \Longrightarrow \quad P_{b\to cb'} \otimes \mathcal{M}\left\{a(p_1) + b'((1-z)p_2) \longrightarrow S_n(k1)\right\}$
- $k^0 = 0 \quad \Longrightarrow \quad \mathrm{Eik}_{ij}(k) \otimes \mathcal{M}_{ij}\left\{a(p_1) + b(p_2) \longrightarrow S_n(k1)\right\}$

But

$$p_1 = z_1 P_1\,, \qquad\qquad p_2 = z_2 P_2$$

and z_i are integration variables, with no proper physical meaning

Initial-state emissions: the solution

- Event projection: adopt a different form for the integration variables z_i in the collinear and soft counterterms, in such a way that

$$p_1^{(soft)} = (1-z)p_1^{(cp+)} = p_2^{(cp-)}$$
$$p_2^{(soft)} = p_2^{(cp+)} = (1-z)p_2^{(cp-)}$$

There are infinitely many ways to accomplish this

The MC has initial-state collinear branchings too, and to perform those it maps $2 \to n$ to $2 \to (n+1)$ configurations. We choose the inverse of this map to change the integration variables z_i

It follows that the implementation of MC@NLO must be based on a subtraction formalism that is able to treat separately the soft and the various collinear emissions, since this is what the Monte Carlo does. The FKS method (Frixione, Kunszt, Signer) has this property

FKS also allows to freely redefine the counterterms, which is useful in order to reduce the number of negative weights

MC counterterms

The MC counterterms are what one gets from the MC by stopping the shower at $\mathcal{O}(a)$

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

This has to be integrated over the same phase space as the NLO

- The shower variables $\xi_\pm$, $z_\pm$ are expressed in terms of invariants
- The MC Bjorken x's $x_1^{(MC)}$, $x_2^{(MC)}$ are identical to the Bjorken x's relevant to the soft countervent (this happens thanks to the definition of event projection)

These procedures *do not depend* on the process; there is some dependence on the multiplicity of the final state (i.e. n in $2 \to n$) through the shower scales

Pitfalls

It is crucial that $\mathcal{M}_{ab}^{(\mathrm{MC})}$ have the same local behaviour of the real emission and of the corresponding counterterms

- The MC counterterms $\mathcal{M}_{ab}^{(\mathrm{MC})}$ are MC-specific; in other words, HERWIG's and PYTHIA's differ. However, they are observable- and process-independent

- The straightforward construction of $\mathcal{M}_{ab}^{(\mathrm{MC})}$ leads to

$$\lim_{E\to 0}\left[\mathcal{M}_{ab}^{(r)}(\phi_3)-\mathcal{M}_{ab}^{(\mathrm{MC})}(\phi_3)\right]\neq 0 \qquad \text{for some } \phi_3$$

  This is due to the peculiar treatment of soft emissions in MC's

  - Smoothly match $\mathcal{M}_{ab}^{(\mathrm{MC})}$ with real ME's, which is equivalent to power-suppressed effects for suitable choices of the matching procedures (current strategy)
  - Change shower variables, improving the soft emission (doable)

What's the problem with the soft limit?

From perturbative computations, we expect the following formula to hold

$$d\sigma_{2\to 3} \overset{E\to 0}{\longrightarrow} \frac{\alpha_S}{E^2}\frac{1}{1-\cos^2\theta}d\sigma_{2\to 2}$$

Using the MC (HERWIG) showering variables, we find instead

$$d\sigma_{2\to 3} \overset{E\to 0}{\longrightarrow} \frac{\alpha_S}{E^2}\left[\frac{2\Theta(\cos\theta > -1/3)}{(1-\cos\theta)(3+\cos\theta)} + \frac{2\Theta(\cos\theta < 1/3)}{(1+\cos\theta)(3-\cos\theta)}\right] d\sigma_{2\to 2}$$

MC's are not designed to produce fixed-order results. Colour-coherence is implemented *after* azimuthal integration. Furthermore, there are regions not covered by shower emissions. But:

$$\int_{-1+\varepsilon}^{1-\varepsilon} d\cos\theta \left[\frac{2\Theta(\cos\theta > -1/3)}{(1-\cos\theta)(3+\cos\theta)} + \frac{2\Theta(\cos\theta < 1/3)}{(1+\cos\theta)(3-\cos\theta)} - \frac{1}{1-\cos^2\theta}\right] \overset{\varepsilon\to 0}{\longrightarrow} 0$$

i.e., the total amount of "soft" energy given by the MC is in agreement with pQCD. Physical observables must be independent of the angular distributions of soft gluons (beware of non-global logs)

MC@NLO in a nutshell

1. Choose your favourite MC (HERWIG, PYTHIA), and compute analytically the "NLO cross section", i.e., the first emission. This is an observable-independent, process-independent procedure, which is done once and for all

2. Implement the NLO matrix elements of your favourite process according to the universal, observable-independent, subtraction-based formalism of SF, Kunszt, Signer for cancelling IR divergences.

   This is the only non-trivial step necessary in order to add new processes

3. Add and subtract the MC counterterms, computed in step 1, to what computed in step 2. The resulting expression allows one to generate the hard kinematic configurations, which are eventually fed into the MC showers as initial conditions

The MC counterterms have been computed by choosing HERWIG to perform the showers. If you use PYTHIA in the showering phase, you get wrong results

Spin correlations

First compute the amplitude for the process

$$a + b \longrightarrow (P \longrightarrow) d_1 + \cdots + d_n + X \qquad \text{Full ME}$$

Then that for

$$a + b \longrightarrow P + X \qquad \text{Undecayed ME}$$

Finally, go to the rest frame of P, and perform the decay

$$P \longrightarrow d_1 + \cdots + d_n \qquad \text{Decay}$$

If the two computations give different predictions for any observable associated with any of the decay products d_i, then we have spin correlations. In general, this occurs when P has non zero spin

When one or more non-zero spin particles decay, we must therefore

- Use the full ME's
- Alternatively, compute the undecayed ME $\otimes$ decay chain for fixed polarizations of P

Spin correlations in MC@NLO

The computation of undecayed ME's for fixed polarizations is quite awkward. When two or more particles decay, a tensorial structure emerges

$\Longrightarrow$ Use full ME's. It's just another production process, which we know how to deal with

A couple of things to keep in mind

- ME must be integrated and unweighted
- The integration time increases and the unweighting efficiency decreases by increasing the number of final-state particles

One more things to keep in mind

- A young theorist will *never* get a job for doing this, in spite of (or perhaps because of) the many thanks he/she will receive from experimenters

And as far as I'm concerned

- I plead guilty: there are actually more exciting things to do...

The current situation

In spite of the previous complaints, all of the processes with spin correlations implemented so far in MC@NLO follow the "Full ME" strategy

- Single-V production ($V = W, Z, \gamma, Z/\gamma$)
- VH production ($V = W, Z$)

Remind that

- There are no spin correlations in Higgs production
  $H^0 \to W(\to l\nu)W(\to l\nu)$ is treated correctly!!

So the spin correlations left to be implemented are for

- $t\bar{t}$, V_1V_2 production

Final states are very complicated here, and it's unlikely we'd be able to achieve the usual unweighting efficiency ($\sim 30 - 50\%$) by implementing the "Full ME" strategy

- This is a good motivation to try and find an alternative to the "Full ME" strategy

Hit-and-miss

Whatever the behaviours of the decay products, the momenta of the decaying particles will not change

$\Longrightarrow$ The full ME's must be bounded from above by the undecayed ME's, times a suitable constant. Find this bound and do hit-and-miss

Advantages

- Only the undecayed ME's will be integrated: no further loss of time
- Unweighting is a two-step procedure: first get the P's momenta, then the d's momenta with hit-and-miss. Decay ME's have no spikes, and thus the hit-and-miss only marginally degrades efficiency

So far, we only studied the decays of vector bosons (i.e. not of top)

$$\frac{d\sigma_{l_1\bar{l}_1\ldots l_n\bar{l}_n}}{d\Phi_{2n+k}} \leq \left(\prod_{i=1}^{n} \frac{2\,F_{V_i}^2\,(V_{V_i l_i} + A_{V_i l_i})^2}{\Gamma_i^2}\right) \frac{d\sigma_{V_1\ldots V_n}}{d\Phi_{n+k}}$$

$$-iF_V\gamma^\mu\,(V_{Vl} - A_{Vl}\gamma_5) \quad \longleftarrow \quad V l\bar{l} \text{ vertex}$$

This bound saturates!

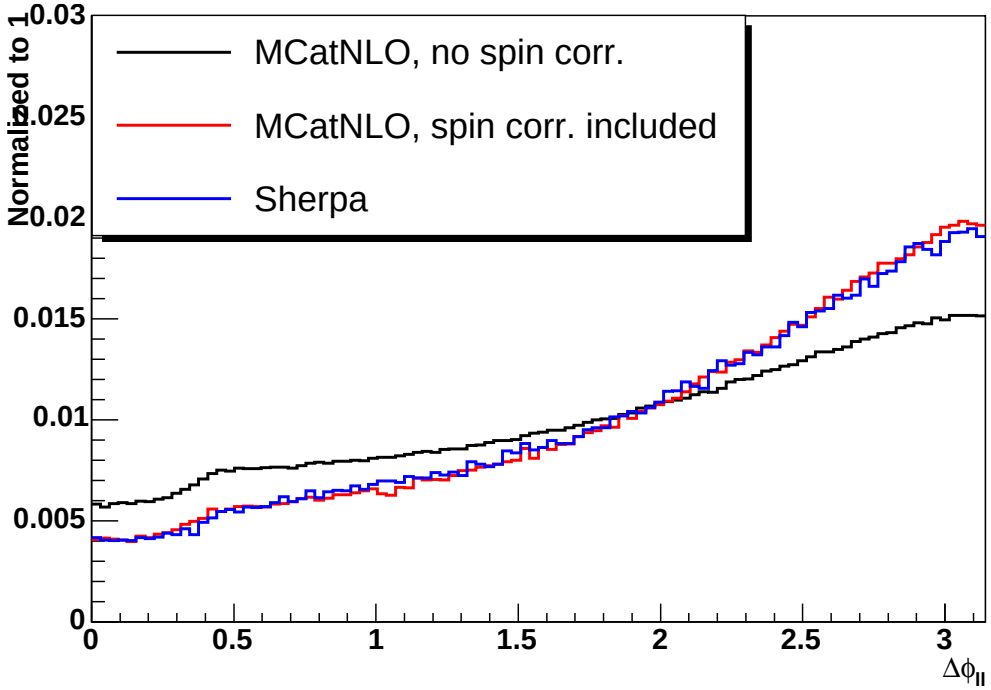
Implementation

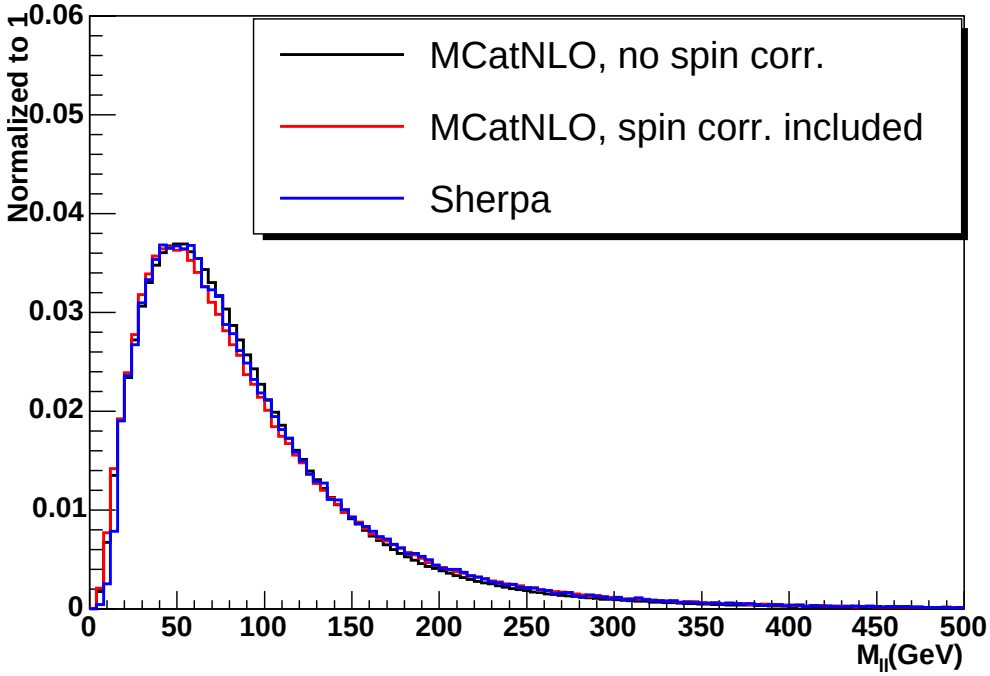
The previous bound applies only to positive-definite quantities, which is not the case for NLO computations. It also applies to those spin-correlation effects that factorize the (fully decayed) Born

The bottom line: spin correlations can't be implemented to full NLO accuracy in MC@NLO *using hit-and-miss*. Non-factorizable effects are however expected to be small

- Regardless of the size of non-factorizable effects, MC@NLO with hit-and-miss is better than standard MC's for spin correlations
- Off-shell effects can also be taken into account (we still have only doubly-resonant diagrams)
- Implemented for W^+W^- production, and tested against MCFM: no difference seen
- The time spent in hit-and-miss unweighting is negligible wrt primary unweighting

Results for W^+W^-





Plots: B. Quayle (preliminary)

- Virtual effects appear to be unimportant (apart from normalization)
- The effect of spin correlations is strictly dependent on the observable
- W^+W^- already used by ATLAS and CMS, official release with v3.1 (next month?)

Thanks to Bill Quayle and Volker Drollinger for testing a preliminary version

Efficiency in Monte Carlo simulations

Suppose one is interested in jets with $p_T^{(jet)} > 1$ TeV at the LHC

- Straightforward solution: run jet production, and event by event reconstruct the jets and impose the $p_T^{(jet)} > 1$ TeV cut

The computer will spend most of its time doing nothing, since only about 1 event in 10^5 will pass the cut. There's nothing wrong, it is just terribly inefficient

- A better solution: run jet production by requiring $p_T > p_T^{(min)}$ at the level of primary partons (hard cut), and still impose $p_T^{(jet)} > 1$ TeV for each event

Clearly, this is not an exact solution, *which does not exist* owing to the complexity of the final states produced by MC's. Thus:

The parameter $p_T^{(min)}$ must be chosen as large as possible to maximize the efficiency, and yet avoiding any bias on the physics observables

- The problem in MC@NLO: the hard events have two different kinematics

Hard cuts in MC@NLO

MC@NLO without hard cuts

$$\mathcal{F}_{\text{MC@NLO}} = \sum_{ab} \int d\phi\, f_a \otimes f_b \otimes \Big[\mathcal{F}_{\text{MC}}^{(2\to 3)} \left(\mathcal{M}_{ab}^{(r)} - \mathcal{M}_{ab}^{(\text{MC})} \right) + \mathcal{F}_{\text{MC}}^{(2\to 2)} \left(\mathcal{M}_{ab}^{(b,v,c)} - \mathcal{M}_{ab}^{(c.t.)} + \mathcal{M}_{ab}^{(\text{MC})} \right) \Big]$$

MC@NLO with hard cuts

$$\mathcal{F}_{\text{MC@NLO}} = \sum_{ab} \int d\phi\, f_a \otimes f_b \otimes \Big[\mathcal{F}_{\text{MC}}^{(2\to 3)} \left(\Theta(2\to 3) \mathcal{M}_{ab}^{(r)} - \Theta(2\to 2) \mathcal{M}_{ab}^{(\text{MC})} \right) + \mathcal{F}_{\text{MC}}^{(2\to 2)} \Theta(2\to 2) \left(\mathcal{M}_{ab}^{(b,v,c)} - \mathcal{M}_{ab}^{(c.t.)} + \mathcal{M}_{ab}^{(\text{MC})} \right) \Big]$$

- Local cancellation of singularities is preserved
- All the necessary formulae have been worked out analytically
- First implementation in $b\bar{b}$ production, but unlikely in v3.1