UCLouvain

Institut de recherche en mathématique et physique Centre de Cosmologie, Physique des Particules et Phénoménologie

Reweighting of MC sample

Olivier Mattelaer

Motivation

Motivation

Re-Weighting

- Reuse the sample (Only one Full Sim)
- Change the weight of the events

$$
W_{new} = \frac{|M_{new}|^2}{|M_{old}|^2} * W_{old}
$$
 1405.0301
1404.7129

Re-Weighting usage

- scale and pdf uncertainties (available both for LO and NLO computation)
- re-introduce top mass effect for Higgs processes
- EFT scan
- Many other application

Examples EFT

Re-Weighting Limitation

Re-Weighting Limitation

- statistical uncertainty are enhanced by the reweighting
- better to have wgt<1and small variance

Re-Weighting Limitation

- statistical uncertainty are enhanced by the reweighting
- better to have wgt<1and small variance

• You need to have the same phase-space (more exactly a subset)

Caution

LHE Additional information

Helicity

Leading color information

Intermediate particle

Caution

LHE Additional information

Helicity *• <u>Provided information is provided in the LHEF* α *in the LHEF* α *in the LHEF* α *in the acronymism of* α *in the acronymism of \alpha* $$

- Partial helicity distribution are not correct with the full re-weighting *if not done before* event file, the reweighting method will not have the e rartial nelicity distribution are not corre
	- Solution $W_{new} =$ $|M_{new}^h|$ 2 $|M_{orig}^h|$ $\frac{1}{2}W_{orig}$
- This is done by default !

given helicity *h*. Indeed in that case we can write the cross-section as

N

i=1

the presence of the second presence of the second probability do not change the total cross-section $\mathcal{L}_\mathcal{S}$

<u>r information</u> origPⁱ Leading color information

h, this probability is formally defined by *Pⁱ*

<i>p <i>h <i>h h <i>n <i><i>n <i>n <i>n <i>n <i>n <i>n *****<i>n <i>n <i>n <i>n <i>n <i>n* *****<i>n <i>n <i>n <i>n <i>n <i>n <i>n <i>n <i>n* *****<i>n <i>n <i>* Intermediate particle

^h˜ *[|]M^h*˜

orig|

^h˜ *[|]Mh*˜

*orig|*² . Note that

<u>P</u>

Caution

LHE Additional information

Helicity *• <u>Provided information is provided in the LHEF* α *in the LHEF* α *in the LHEF* α *in the acronymism of* α *in the acronymism of \alpha* $$

- Partial helicity distribution are not correct with the full re-weighting *if not done before* event file, the reweighting method will not have the e rartial nelicity distribution are not corre
	- Solution $W_{new} =$ $|M_{new}^h|$ 2 $|M_{orig}^h|$ $\frac{1}{2}W_{orig}$
- This is done by default !

<u>r information</u> origPⁱ Leading color information

di LC
Val *N Wⁱ [|]M^h orig|* 2 *^h*˜ *[|]M^h*˜ *r i l d i d d i l d i d i d i d i d i* *d i d i d i d i d i* *d i d i d i d i d i d i d i d i d i d i d i d i d i* • modify the parton-shower so not suitable.

<u>P</u>

given helicity *h*. Indeed in that case we can write the cross-section as

N

i=1

<i>p <i>h <i>h h <i>n <i><i>n <i>n <i>n <i>n <i>n <i>n *****<i>n <i>n <i>n <i>n <i>n <i>n* *****<i>n <i>n <i>n <i>n <i>n <i>n <i>n <i>n <i>n* *****<i>n <i>n <i>* Intermediate particle

 \backslash • modify the parton-shower *[|]M^h orig|* 2 P *^h*˜ *[|]Mh*˜ *orig|*² . Note that • modify the parton-shower so not suitable.

the presence of the second presence of the second probability do not change the total cross-section $\mathcal{L}_\mathcal{S}$

orig|

GPU

- Reweighting are massively repetitive task
- Ideal for GPU ➡Zenny Wettersten
- We are in need of real experimental test case
	- ➡ Number of event in sample
	- ➡ Number of benchmark

NLO Re-Weighting .u ne- $\bigcap_{i=1}^n A_i$ introduced in the context of the evaluation systematics uncertainties: *d*↵ = *f*1(*x*1*, µ^F*)*f*2(*x*2*, µ^F*) h *^W*↵ ⁰ ⁺ *^W*↵ *^F* log (*µ^F /Q*) ² + $2a$ introduced in the context of the context of the context of the evaluation of the evaluation of the evaluation of the evaluation \mathbf{r} systematics uncertainties: *d*↵ = *f*1(*x*1*, µ^F*)*f*2(*x*2*, µ^F*) h *^W*↵ ⁰ ⁺ *^W*↵ *^F* log (*µ^F /Q*)

In order to have an accurate NLO re-weighting method,

d↵

^R log (*µR/Q*)

2 i

, (9)

d↵

NLO method

–*R*– and in the finite piece of the virtual –*V*–). We spendencies in the vanous introduced in the context of the evaluation of the sing vn caal, i $intriv$ and the various color-flows (a case allows (a case allows (a case allows \mathcal{L} not handled at LO accuracy \vert More generally, the possible drawbacks and limitations on the statistical precision of the method are the ous sub-section). *Q* is the Ellis-Sexton scale and *d*↵ is ependencies The expression of the *^W*↵ ⁰ , *^W*↵ elements (born, virtual, real) of computations of the virtual by using an approximate • tracks the dependencies in the various matrix-
 tudes times a fitted parameter . It performs a separate phase-space integration to get the di↵erence between the virtual and its approximation (full descripous sub-section). *Q* is the Ellis-Sexton scale and *d*↵ is the phase-space measure. \mathbf{r} and are not repeated here. All those repeated here. All those repeated here. All those repeated here. All those repeated here.

₩

$$
d\sigma^{\alpha} = f_1(x_1, \mu_F) f_2(x_2, \mu_F) \left[\mathcal{W}_0^{\alpha} + \mathcal{W}_F^{\alpha} \log (\mu_F / Q)^2 + \mathcal{W}_R^{\alpha} \log (\mu_R / Q)^2 \right] d\chi^{\alpha},
$$

$$
\mathcal{W}_{\beta}^{\alpha} = \mathcal{B} * \mathcal{C}_{\beta, B}^{\alpha} + \mathcal{B}_{CC} * \mathcal{C}_{\beta, B_{CC}}^{\alpha}
$$

^R log (*µR/Q*)

^W↵

the following three terms:⁶

,B ⌘ *^B* ⇤ *^C*↵

^W↵

$$
\ +\ \ \mathcal{V}*\mathcal{C}^{\alpha}_{\beta,V}+\mathcal{R}*\mathcal{C}^{\alpha}_{\beta,R}
$$

ach part according to the a aterm part about anny to the a ent the corresponding expressions assessment of the corresponding expressions assessment of the corresponding of the corresponding expressions assessment of the corresponding expressions as:55 μ tion of the method is presented in Section 2.4.3 of [6]). socialcu | (*^B* ⁺ *B*) + ^Z ach nart acco *<u>D</u> C ,B* ⁺ *^BCC* ⇤ *^C*↵ *,BCC ,* (11) *^W*↵ *,V* ⌘ *^V* ⇤ *^C*↵ \mathcal{V} *new* inciated loop we give the proposed in \mathbb{R} will highly enhance the \mathbb{R} contribution of the second integral since each term of •re-weight each part according to the associated matrix-element or the matrix-element of the matrix-element. From this expression we define the matrixart according tance sampling [27]), reducing the amount of time used ed in the loop. weighting proposed in Eq. 14 will highly enhance the contribution of the second integral since each term of

$$
\mathcal{W}_{\beta,B}^{\alpha,new}=\frac{\mathcal{B}^{new}}{\mathcal{B}^{old}}*\mathcal{W}_{\beta,B}^{\alpha,old},\\\mathcal{W}_{\beta,V}^{\alpha,new}=\frac{\mathcal{V}^{new}}{\mathcal{V}^{old}}*\mathcal{W}_{\beta,V}^{\alpha,old},\\\mathcal{W}_{\beta,R}^{\alpha,new}=\frac{\mathcal{R}^{new}}{\mathcal{R}^{old}}*\mathcal{W}_{\beta,R}^{\alpha,old}.
$$

rations is assumed here and in the rest of the paper.

matrix-element is evaluated to \mathcal{C} . The matrix-element is evaluated is evaluated in evaluated in \mathcal{C} is not unique: an implicit sum over such kinematical configu*,BC* and *^W*↵

parts: *^W*↵

the Born which can lead to a breaking of the NLO

same as for the LO case. However, for NLO calculations

same as for the LO case. However, for NLO calculations

 \blacksquare

virtual contribution. This method reduces the number

of computations of the virtual by using an approximate

of the virtual contribution based on the Born ampli-

tudes times a fitted parameter . It performs a sepa-

rate phase-space integration to get the di↵erence be-

tween the virtual and its approximation (full descrip-

tion of the method is presented in Section 2.4.3 of [6]).

If it exists a value of such that *B* ⇡ *V*, the second

integral is approximately zero and does not need to be

probed as often as the first integral (thanks to impor-

the integral will be re-weighted by a di↵erent factor,

more advanced re-weighting technique. We split the

contribution proportional to the Born (*W*↵

 $T = \frac{1}{\sqrt{2}}$

having a direct impact on the statistical uncertainty.

uncertainty due to the method used to the method used to integrate the method used of the method used of

 \mathcal{L}

accuracy of the method. However such an approxima-

tion does not consist in an additional limitation of the

method since the re-weighting factors shown that \mathcal{N}

if the two theories present a di↵erence in the relative

same as for the LO case. However, for NLO calculations

tion of the method is presented in Section 2.4.3 of [6]).

Schematically it can be written as:

in MG5 aMC we face one additional source of statistical

Schematically it can be written as:

² +

, (9)

uncertainty due to the method used to integrate the

virtual contribution. This method reduces the number

(*^B* ⁺ *B*) + ^Z

(*^B* ⁺ *^V*) = ^Z

of computations of the virtual by using an approximate

of the virtual contribution based on the Born ampli-

If it exists a value of such that *B* ⇡ *V*, the second

integral is approximately zero and does not need to be

probed as often as the first integral (thanks to impor-

tance sampling [27]), reducing the amount of time used

the integral will be re-weighted by a di↵erent factor,

more advanced re-weighting technique. We split the

tional to the Born, relation to the one of the one of the countert-term of the countert-term of the countert-

,BC and *^W*↵

To reduce this e↵ect, we propose to use a slightly

having a direct impact on the statistical uncertainty.

tudes times a fitted parameter . It performs a sepa-

rate phase-space integration to get the di↵erence be-

tween the virtual and its approximation (full descrip-

If it exists a value of such that *B* ⇡ *V*, the second

integral is approximately zero and does not need to be

probed as often as the first integral (thanks to impor-

tance sampling [27]), reducing the amount of time used

contribution proportional to the Born (*W*↵

in the evaluation of the loop-diagrams. However the re-

,BB. *^W*↵

weighting proposed in Eq. 14 will highly enhance the Eq. 14 will highly enhance the Eq. 14 will highly enhance the Eq. 14 will have the Eq

NLO ISSUE

- MadGraph use some phase-space trick to avoid to compute the loop as much as possible (and replace it by the born)
	- ➡ Need smarter/complex re-weighting

NLO Re-Weighting \blacksquare troduced in \mathbf{S} is the HD and HD

Conclusion

- Re-using previous generation/computation is always a smart move.
- This methods is fully exact but not bullet proof
	- Need to check overlap of phase-space/helicity
	- not (really) suitable for mass scanning
	- helicity (need to be careful)
	- leading color information
	- intermediate particle
- NLO is ready
	- Same limitation
	- Issue with numerical precision

Why it works lies in the fact that the original weights depend linearly in the matrix-element by definition of the weight. We now introduce some notation and show how reweighting modifies the cross-section. Notice that what follows is independent follows in the context of the c the cross-section is estimated as an average over many simulated events: *Wⁱ*

 \sim

Wⁱ

Why it works lies in the fact that the original weights depend linearly in the matrix-element by definition of the weight. We now introduce some notation and show how reweighting modifies the cross-section. Notice that what follows is independent follows in the context of the c the cross-section is estimated as an average over many simulated events: *Wⁱ*

 \sim

Wⁱ

unweighting case is to reduce the number of \sim *i*=1 *Wⁱ orig,* (8) \mathcal{L} *N Wⁱ orig*

helicity case *h*, this probability is formally defined by *Pⁱ* ² *Worig,* (15)

$$
\sigma_{orig} = \sum_{i=1}^{N} W_{orig}^{i} P_{h,orig}^{i},
$$

$$
= \sum_{i=1}^{N} W_{orig}^{i} \frac{|M_{orig}^{h}|^2}{\sum_{\tilde{h}} |M_{orig}^{\tilde{h}}|^2},
$$

|M^h

orig|

$$
\begin{aligned}\n\sigma_{orig} &= \sum_{i=1}^{N} W_{orig}^{i} P_{h,orig}^{i}, \\
&= \sum_{i=1}^{N} W_{orig}^{i} \frac{|M_{orig}^{h}|^2}{\sum_{\bar{h}} |M_{orig}^{\bar{h}}|^2}, \\
&= \sum_{i=1}^{N} W_{new}^{i} \frac{|M_{new}^{h}|^2}{\sum_{\bar{h}} |M_{new}^{\bar{h}}|^2}, \\
&= \sum_{i=1}^{N} W_{orig}^{i} \frac{\sum_{\bar{h}} |M_{new}^{\bar{h}}|^2}{\sum_{h'} |M_{orig}^{\bar{h'}}|^2} \frac{|M_{new}^{\bar{h}}|^2}{\sum_{\bar{h}} |M_{new}^{\bar{h}}|^2}, \\
&= \sum_{i=1}^{N} W_{orig}^{i} \frac{1}{\sum_{h'} |M_{orig}^{\bar{h'}}|^2} \frac{|M_{new}^{\bar{h}}|^2}{\sum_{h'} |M_{orig}^{\bar{h'}}|^2}, \\
&= \sum_{i=1}^{N} W_{orig}^{i} \frac{|M_{orig}^{\bar{h}}|^2}{\sum_{h'} |M_{orig}^{\bar{h}}|^2} \frac{|M_{new}^{\bar{h}}|^2}{|M_{orig}^{\bar{h}}|^2}, \\
&= \sum_{i=1}^{N} W_{orig}^{i} P_{h,orig}^{i} \frac{|M_{new}^{\bar{h}}|^2}{|M_{orig}^{\bar{h}}|^2}.\n\end{aligned}
$$

h,orig =

P

^h˜ *[|]Mh*˜

*orig|*² . Note that

N

[|]M^h

orig|

2

[|]M^h

2