MG5aMC
Software and Operation

Olivier Mattelaer
IPPP/Durham

Disclaimer

Current situation on MPI

- MadGraph do not provide any support for MPI
- No work in that direction so far

Personally

- I’m interested to learn how to efficiently use this protocol
- I’m basically here to learn and see how to design a good strategy

Plan of my talk

plan

- What is MadGraph5_aMC@NLO?
- new feature to improve generation
- Phase-space integration methods
- Cluster support
- Idea for MPI

What is MG5aMC?

- This is a matrix element provider
  - tree-level and one loop
- This is a Framework of tools
  - cross-section and event generation
    - LO
    - NLO(MC@NLO)
    - Loop-induced
  - Utility tools
    - MadSpin
    - Re-weighting
  - Plugin output mode

What is MG5aMC?

- MG5aMC is a meta-code
  - ➡ the various code are independent
    - ❑ various structure/integration method
  - ➡ MG5aMC provides some common API
    - ❑ LHEF parser
    - ❑ in particular the handling of jobs on cluster

- MG5aMC is an App-Store
  - ➡ without the control/coherence of Apple

New features

- 3 new features
  - ➡ systematics
  - ➡ BSM re-weighting
  - ➡ biased generation
- released in beta version 2.5.0

Th. Uncertainty

$$\sum_{a,b} \int dx_1 dx_2 d\Phi_{\mathrm{FS}}\, f_a(x_1,\mu_F) f_b(x_2,\mu_F)\, \hat{\sigma}_{ab\to X}(\hat{s},\mu_F,\mu_R)$$

Phase-space integral — Parton density functions — Parton-level cross section

- The scale are arbitrary
- varying scale/PDF gives an **estimations** of the error

- Direct computation for different pdf/scale choice is very slow
- method of choice:
  - store the required information event by event
  - recompute the cross-section from the generated events
- We replace **SysCalc** at LO

Th. Uncertainty

- generate p p > t t~
- output
- launch

```
INFO: # events generated with PDF: NNPDF23_lo_as_0130_qed (247000)
INFO: #Will Compute 144 weights per event.
INFO: #***********************************************************************
#
# original cross-section: 505.430386046
#     scale variation: +30.2% -28.6%
#     central scheme variation: +3.36e-09% -25.9%
# PDF variation: +3.77% -3.77%
#
# dynamical scheme # 1 : 479.713 +29.9% -21.6% # \sum ET
# dynamical scheme # 2 : 395.728 +27.7% -20.5% # \sum\sqrt{m^2+pt^2}
# dynamical scheme # 3 : 505.43 +30.2% -21.7% # 0.5 \sum\sqrt{m^2+pt^2}
# dynamical scheme # 4 : 374.366 +27.1% -20.1% # \sqrt{\hat s}
#***********************************************************************
```

Th uncertainty @NLO

- generate p p > t t~ [QCD]
- output
- launch
  - set store_rwgt_info T

- systematics run_01 (**OFFLINE**)

```
INFO: # events generated with PDF: NNPDF23_nlo_as_0119_qed (244800)
INFO: #Will Compute 144 weights per event.
INFO: #*********************************************************************
#
# original cross-section: 704.418156719
#     scale variation: +9.75% -10.7%
#     central scheme variation: + 0% -28.2%
# PDF variation: +1.55% -1.55%
#
# dynamical scheme # 1 : 637.396 +7.98% -9.89% # \sum ET
# dynamical scheme # 2 : 628.485 +11.9% -11.7% # \sum\sqrt{m^2+pt^2}
# dynamical scheme # 3 : 703.534 +9.67% -10.7% # 0.5 \sum\sqrt{m^2+pt^2}
# dynamical scheme # 4 : 505.65 +4.63% -7.76% # \sqrt{\hat s}
#*********************************************************************
```

BSM Re-Weighting

Re-weighting are everywhere

- scale and pdf uncertainties (available both for LO and NLO computation)
- loop induced processes
- matching/merging
- experimental re-weighting

BSM Re-weighting

- Change the events weights of a LHEF for various BSM theories.
- Re-use the same parton shower and detector simulation

LO

Re-Weighting

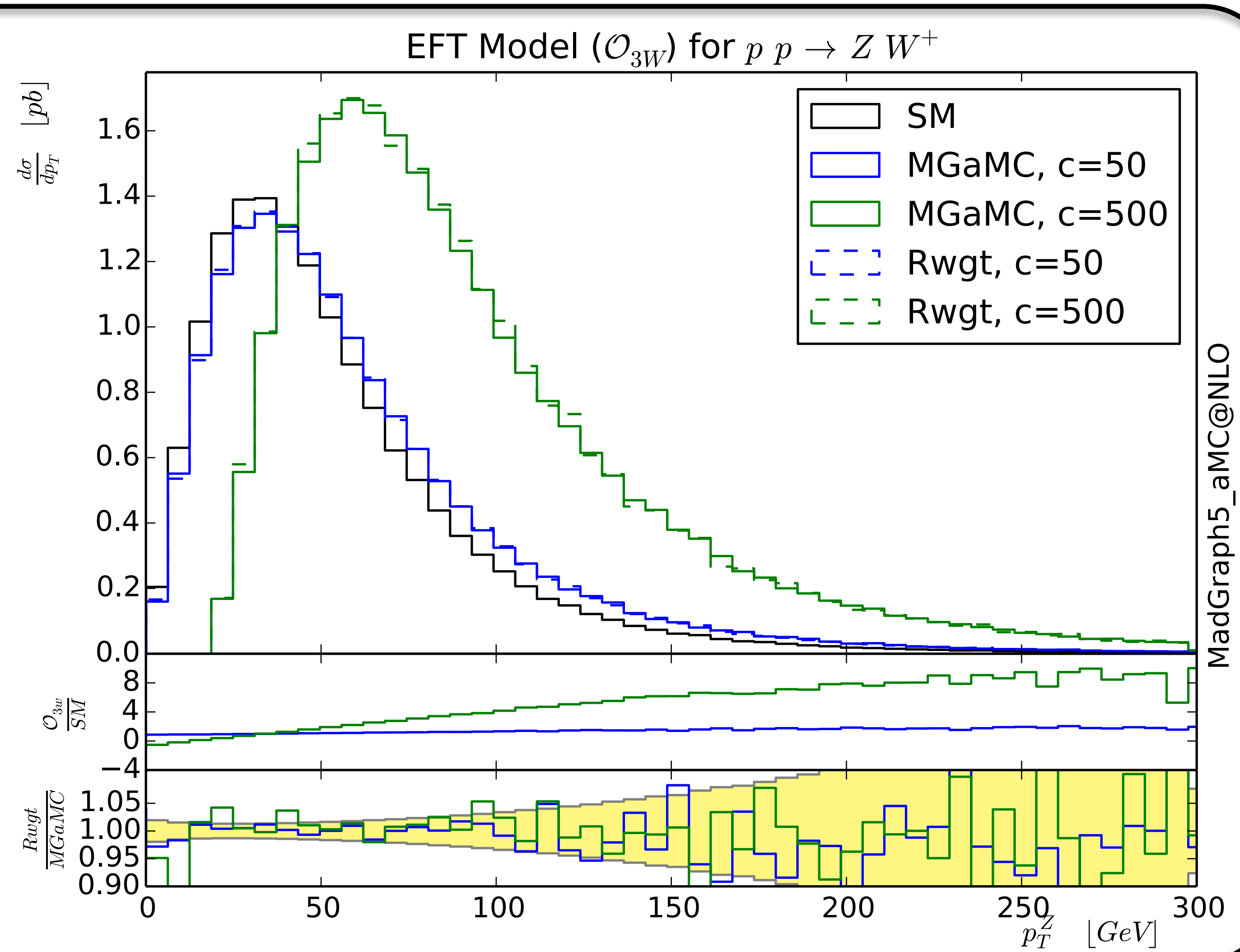
1404.7129
1607.00763

- Change the weight of the events

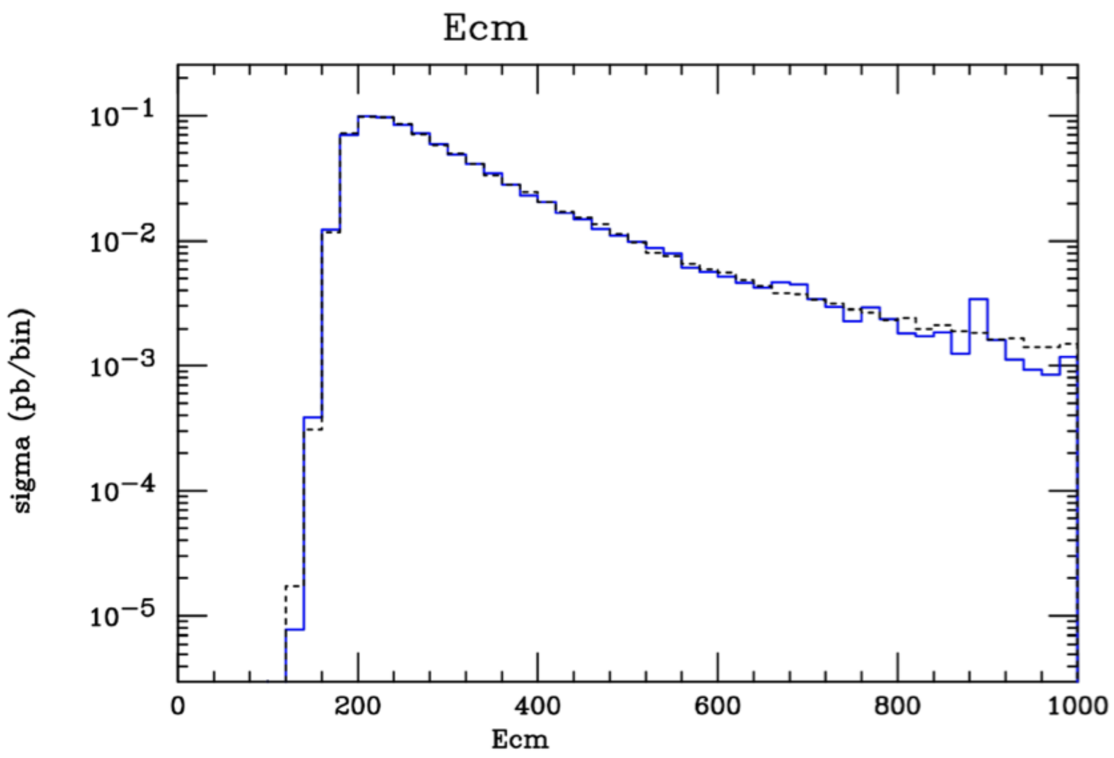
$$W_{new} = \frac{|M_{new}|^2}{|M_{old}|^2} * W_{old}$$

EFT Case

$$\mathcal{O}_{3W} = Tr\left[W_{\mu\nu}W^{\nu\rho}W_{\rho}{}^{\mu}\right]$$



Re-Weighting Limitation



$$\Delta\sigma_{new} = \frac{\sigma_{new}}{\sigma_{old}} \Delta\sigma_{new} + \mathrm{Var}_{wgt}\, \sigma_{old}$$

- statistical uncertainty can be enhanced by the re-weighting
- better to have wgt<1

- You need to have the same phase-space (more exactly a subset)
- Mass scan are possible only in special case
  - only for internal propagator
  - for small mass variation (order of the width)

NLO Re-Weighting

NLO method

- tracks the dependencies in the various matrix-elements (born, virtual, real)

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

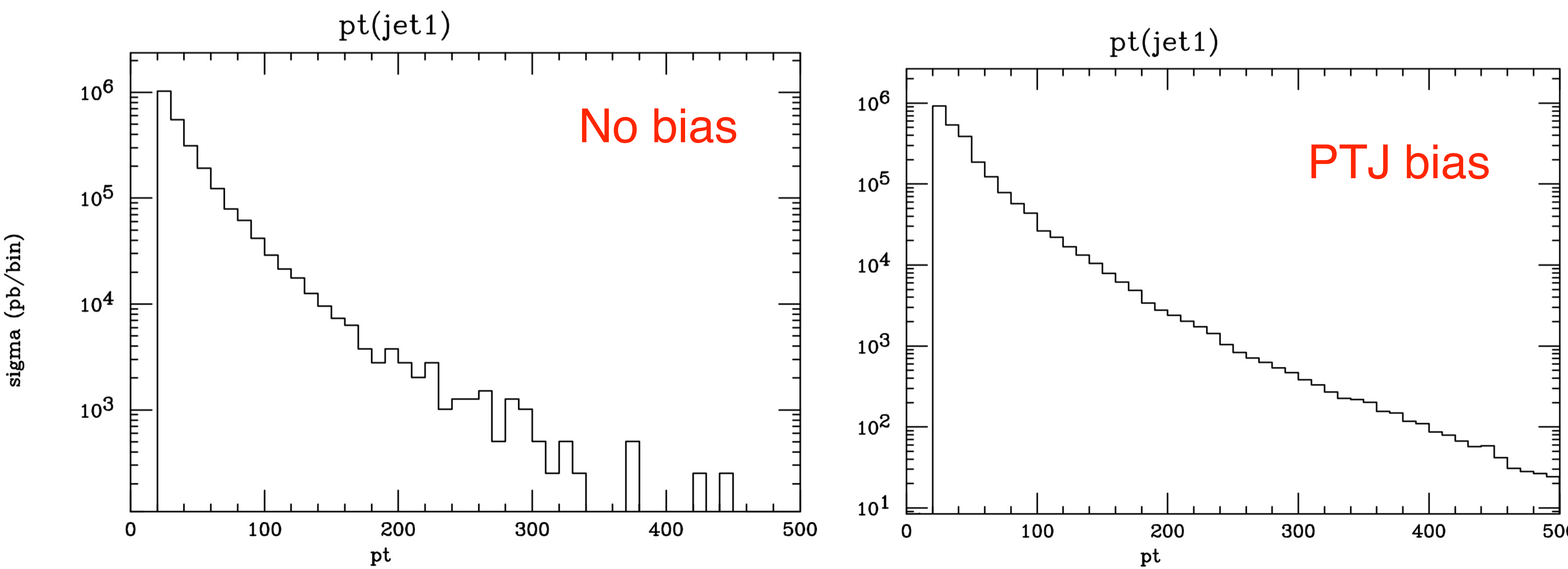
$$\mathcal{W}_\beta^\alpha = \quad \mathcal{B} * \mathcal{C}_{\beta,B}^\alpha + \mathcal{B}_{CC} * \mathcal{C}_{\beta,B_{CC}}^\alpha + \mathcal{V} * \mathcal{C}_{\beta,V}^\alpha + \mathcal{R} * \mathcal{C}_{\beta,R}^\alpha$$

- re-weight each part according to the associated matrix-element
- compute the weight

LO Bias Module

Idea

- modify the function to integrate
  - ➡ change your cross-section
  - ➡ possibility to preserve the cross-section
    - weighted events (usefull for tail



Matrix-Element

Calculate a given process (e.g. gluino pair)

- Determine the production mechanism

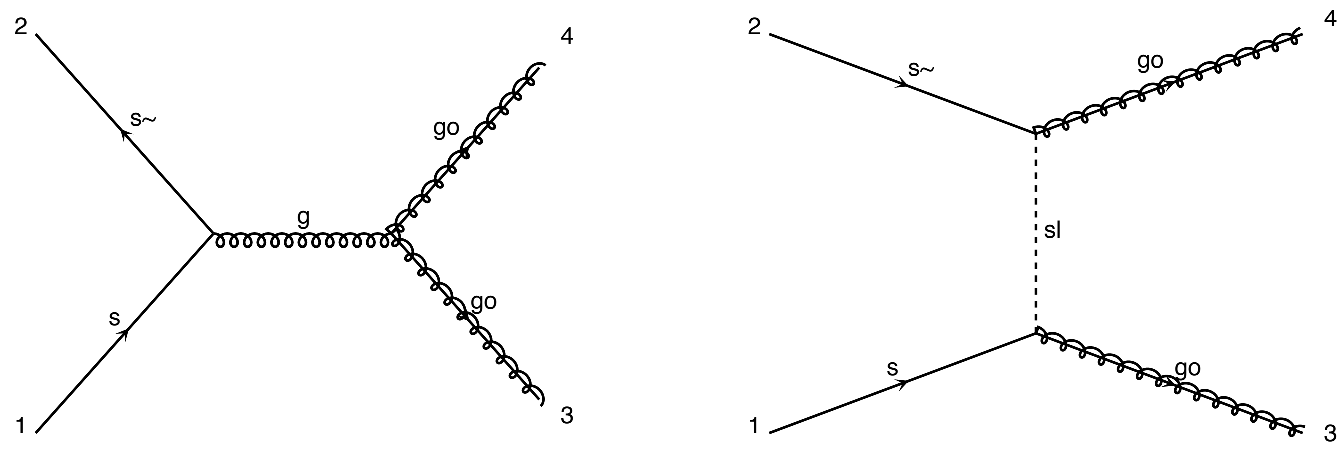


diagram 1 QCD=2, QED=0   diagram 2 QCD=2, QED=0

- Evaluate the matrix-element

$|\mathcal{M}|^2$ ➡Numerical method (faster than analytical computation)

- Phase-Space Integration

$$\sum_{a,b} \int dx_1 dx_2 d\Phi_{\mathrm{FS}}\, f_a(x_1,\mu_F) f_b(x_2,\mu_F)\, \hat{\sigma}_{ab\to X}(\hat{s},\mu_F,\mu_R)$$

Diagram enhancement

$$\int |M|^2 = \int \frac{|M_1|^2}{|M_1|^2 + |M_2|^2} |M|^2 + \int \frac{|M_2|^2}{|M_1|^2 + |M_2|^2} |M|^2$$

≈ 1 ≈ 1

Easy to integrate

LO

$$\int |M|^2 = \int \frac{|M_1|^2}{|M_1|^2 + |M_2|^2} |M|^2 + \int \frac{|M_2|^2}{|M_1|^2 + |M_2|^2} |M|^2$$

Node

scheduler

- Survey
  - probing
  - grid refinement

- Survey
  - probing
  - grid refinement

- Determine precision (number of event) per channel

- Refine
  - probing
  - grid refinement

- Refine
  - probing
  - grid refinement

NLO

- Additional splitting based on the FKS phase-space splitting (at most one soft/collinear contribution by channel)
- Use mint procedure for the phase-space integration
- splitting on node basically the same as LO

virtual optimisation

$$\int (\mathcal{B}+\mathcal{V}) = \int (\mathcal{B}+\kappa\mathcal{B}) \; + \; \int (\mathcal{V}-\kappa\mathcal{B}).$$

LO Gridpack

$$\int |M|^2 = \int \frac{|M_1|^2}{|M_1|^2 + |M_2|^2}|M|^2 + \int \frac{|M_2|^2}{|M_1|^2 + |M_2|^2}|M|^2$$

High accuracy

- Survey
  - probing
  - grid refinement

- Survey
  - probing
  - grid refinement

End gridpack

Running it

- channel run sequentially based on the survey

Each of those run provides physical distributions

Loop-Induced

$$\int |M|^2 = \int \frac{|M_1|^2}{|M_1|^2+|M_2|^2}|M|^2 + \int \frac{|M_2|^2}{|M_1|^2+|M_2|^2}|M|^2$$

Survey

- Iteration 1
- Grid Refinement
- Iteration 2
- Grid Refinement

Survey

- Iteration 1
- Grid Refinement
- Iteration 2
- Grid Refinement

Loop-Induced

$$\int |M|^2 = \int \frac{|M_1|^2}{|M_1|^2 + |M_2|^2}|M|^2 + \int \frac{|M_2|^2}{|M_1|^2 + |M_2|^2}|M|^2$$

Survey

Survey

number of node equals to cluster size

- Iteration 1
- Grid Refinement

number of node equals to cluster size

- Iteration 2
- Grid Refinement

- Iteration 1
- Grid Refinement
- Iteration 2
- Grid Refinement

Job scheduler

- We natively support the most common job scheduler
  - ➡ pbs
  - ➡ sge
  - ➡ condor
  - ➡ lsf
  - ➡ slurm
- They all share the same methods (common with the multi-core mode)
- i.e. this is fully factorise in our code

Job scheduler

- Possibility to customize/modify a job scheduler via the plugin module of MG5aMC
- All job scheduler (but condor) assumes the presence of a shared filesystem
- Condor is transferring the input first on the node (with the associate protocol) and can therefore run without shared filesystem
  - So the capability is available for all cluster
- LHAPDF data can be accessible via cvmfs

Idea for MPI

Plan of implementation

- First implement it at LO / Loop-induced
  - ➡ perfect playground to learn the system
  - ➡ partial work available due to the loop-induced improvement
- then NLO and/or MadSpin

Idea for MPI

Idea 1

- keep the loop-induced approach
  - ➡ random generation for N phase-space in a node
  - ➡ combine results before launching next iteration
- improve initialization phase
- change method of communication
- add intermediate events unweighting

Idea for MPI

Idea 2

- The node receive one kinematical configuration and ONLY compute the matrix elements (+primary unweighting)
    - ➡ close to standalone on the node
    - ➡ all random number generated by the scheduler

Conclusion

plan

- MadGraph5_aMC@NLO is a framework of tools
  - multiple code -> multiple method
- Multiple re-weighting method to improve speed
- Describe the job splitting used in the code
- MPI idea