Eric Conte (GRPHE / Université de Haute-Alsace)

in collaboration with B. Fuks & G. Serret

FeynRules 2012 workshop @ Mont Sainte-Odile
March 26-30, 2012

Outline

MadAnalysis overview

Scope

- Analysis of event files produced by Monte Carlo tools at parton level, hadron level or after detector simulation.
- Definition of various selection cuts on the input samples.
- Production of histograms for different distributions.
- Results of the analysis summed up by a S/B-like ratio table.

Computing details

- Interface written in Python and ROOT ; kernel in C++.
- Possible output in ROOT, HTML, LaTeX.

Website

- *https : //server06.fynu.ucl.ac.be/projects/madanalysis*
- Please send us your comments and suggestions (tickets on the wiki).

Download and installation

A tarball of the program source will be available on the website.

Two installation procedures are possible : untaring the tarball

- inside MadGraph 5 directory (MG5 settings will be used).
- in an independent directory (stand alone running).

MadAnalysis depends essentially on 3 programs which must be present on the system (libraries & headers) :

- Python version $\geq$ 2.6 but not 3.X.
- Boost C++ libraries.
- Root release $\geq$ 5.27.

If the dependencies are installed in a local directory, please set the variables $\$CPLUS\_INCLUDE\_PATH$ and $\$LD\_LIBRARY\_PATH$.

First session of MadAnalysis

```
bin/ma5
Checking ROOT libraries ...
Loading ROOT libraries ...
Checking g++ libraries ...
Checking Boost libraries ...
Checking RPC/XDR libraries ...
Checking MadAnalysis library ...
First use of MadAnalysis detected (or the library is missing)
   Creating (or overwriting) folder 'lib'...
   Copying 'SampleAnalyzer' source files...
   Creating a 'Makefile'...
   Compiling the MadAnalysis library...
   Linking the MadAnalysis library...
   Checking the MadAnalysis library presence...

************************************************************
*                                                          *
*        W E L C O M E  to  M A D A N A L Y S I S  5       *
*                                                          *
*                      ______  ______                      *
*                 /'\_/'\/\  __ \/\  ___\                  *
*                /\      \ \ \_\ \ \ \__/                  *
*                \ \ \__\ \ \  __ \ \___``\                *
*                 \ \ \_/\ \ \ \/\ \/\ \_\ \               *
*                  \ \_\\ \_\ \_\ \_\ \____/               *
*                   \/_/ \/_/\/_/\/_/\/___/                *
*                                                          *
*   MA5 release : 0.5.74                        2012/03/22 *
*                                                          *
*   The MadAnalysis Development Team - Please visit us at  *
*   https://server06.fynu.ucl.ac.be/projects/madanalysis   *
```

First session of MadAnalysis

```
************************************************************
*                                                          *
*       W E L C O M E  to  M A D A N A L Y S I S  5        *
*                                                          *
*                      ______  ______                      *
*                 /'\_/'\/\  __ \/\  ___\                  *
*                /\      \ \ \_\ \ \ \__/                  *
*                \ \ \__\ \ \  __ \ \___``\                *
*                 \ \ \_/\ \ \ \/\ \/\ \_\ \               *
*                  \ \_\\ \_\ \_\ \_\ \____/               *
*                   \/_/ \/_/\/_/\/_/\/___/                *
*                                                          *
*   MA5 release : 0.5.74                         2012/03/22  *
*                                                          *
*   The MadAnalysis Development Team - Please visit us at  *
*   https://server06.fynu.ucl.ac.be/projects/madanalysis   *
*                                                          *
*              Type 'help' for in-line help.               *
*                                                          *
************************************************************
MadGraph 5 NOT found => default particle names from the file:
  /home/econte/madanalysis/input/particles_name_default.txt
84 particles have been successfully exported.
MadGraph 5 NOT found => default multiparticle definitions from the file:
  /home/econte//madanalysis/input/multiparticles_default.txt
Creation of a multiparticle labelled by 'invisible' (related to missing energy).
Creation of a multiparticle labelled by 'hadronic' (related to jet transverse energy).
8 multiparticles have been successfully exported.
ma5>
```

MadGraph-like interface

Mimic the design concepts of the user-friendly MadGraph interface

- inline help
- autocompletion with tabulation key
- history of the commands
- possibility to launch shell commands
- several actions in only one command line (actions separated by ;)
- allowing to have comments starting with #
- coloured logger with several levels of criticality.
- alternative to interactive interface : scripts

Writing my first analysis with MadAnalysis

Step 1 : importing event files

This proceeds through the command import :

- Supported format : LHE, STDHEP, HEPMC, LHCO (home-made readers). Gzip files are also supported.
- Wildcards are allowed → several files can be imported at a time

Storage :

- Imported files are stored as datasets.
- Default set name : defaultset ; otherwise specified by the user.

```
ma5>import diboson* as diboson
ma5>import ttbar* as ttbar
```

Two types of datasets : signal and background

Aim : comparing signal and background distributions

```
ma5>set diboson.type = background
ma5>set ttbar.type = signal
```

Step 2 : defining new multiparticles

MadAnalysis reuses the concept of *multiparticle* defined by MadGraph.

Display the list of particles and multiparticles :

```
ma5>display_particles
a b b~ c c~ d d~ e+ e- g h s s~ t t~ ta+ ta- u u~
ve ve~ vm vm~ vt vt~ w+ w- z
ma5>
ma5>display_multiparticles
hadronic invisible j l+ l- p vl vl~
```

Defining your own multiparticles :

- from other (multi)particles.

```
ma5>define mu = mu+ mu-
```

- from PDG-id codes.

```
ma5>define mu = +13 -13
```

Note : 2 special labels : ***invisible*** *(missing transverse energy) and* ***hadronic*** *(hadronic activity).*

Step 3 : defining a selection

Plot :

The command plot allows to define the distributions to be investigated.

Syntax : plot <observable> nbin xmin xmax

```
ma5>plot MET 100 0 1000
```

Cut :

The command reject allows to reject events which satisfies a given condition.

Syntax : reject <condition1> [and/or <condition2> and/or <condition3> ...]

```
ma5>reject MET < 20
```

The command accept is similar but keeps events which satisfy a given condition.

```
ma5>accept MET > 20
```

List of observables implemented :

- total and missing transverse energy (TET and MET)
- total and missing transverse hadronic energy (THT and MHT)
- final particles present in the samples (NPID and NAPID)

Step 3 : defining a selection

plot, accept and reject can be applied on a (multi)particle.
The syntax is quite different :

plot <observable> (<particle>) nbins xmin xmax

reject/accept (<particle>) <condition1> [and/or ...]

Examples :

```
ma5>plot PT ( mu+ ) 100 0 100
ma5>accept ( mu+ ) PT > 20 and PT < 100
```

List of observables implemented :

- multiplicity (N),
- energy (E and ET),
- mass (M and MT),
- momentum magnitude and components (P, PT, PX, PY, PZ),
- angles (THETA, ETA and PHI),
- relativist factors (Y, BETA and GAMMA).

Step 3 : defining a selection

Moreover, analysists can find useful to plot (or apply a cut on) invariant mass of n-particles state. MadAnalysis does the job !

```
ma5>plot M ( mu- mu+ ) 100 0 150
```

In this example, all combinations μ^+ μ^- are performed and their mass fills the histogram.

To be the most generic as possible, commands plot, accept and reject can be applied to a combination of several particles. Particle momenta are summed vertorially before calculating the observable.

Note : MadAnalysis is very careful when it performs the different combinations. In particular, it avoids possible double-counting.

Step 4 : launching the analysis

The command **submit** allows to perform the selected analysis :

- reading of the Monte Carlo event files.
- updating the information associated to each dataset.
- creating a ROOT file with the analysis itself.
- syntax : submit <dirname>

Example :

```
ma5>submit toto
   Creating folder '/home/econte/myAnalysis'...
   Copying 'SampleAnalyzer' source files...
   Inserting your selection into 'SampleAnalyzer'...
   Writing the list of datasets...
   Creating a 'Makefile'...
   Compiling 'SampleAnalyzer'...
   Linking 'SampleAnalyzer'...
   Running 'SampleAnalyzer' over dataset 'defaultset'...
    ************************************************
    * SampleAnalyzer 1.5 for MadAnalysis 5 - Welcome.
    * Option choices: selecting analysis 'MadAnalysis5job'.
```

Step 5 : diplaying results

Information related to datasets

Displaying a dataset via MadAnalysis interface gives information about Monte Carlo samples.

```
ma5>display ttbar
   ****************************************
   Name of the dataset = ttbar (signal)
   User-imposed cross section = 0.0
   User-imposed weight of the set = 1.0
   Line color in histograms = auto
   Background color in histograms = auto
   List of LHE files included in this dataset:
    - /home/econte/madanalysis/ttbar_dilep.1.lhe.gz
   ****************************************
   Cross section = 3.2 pb
   Total number of events = 15059
   ****************************************
```

Generating a report

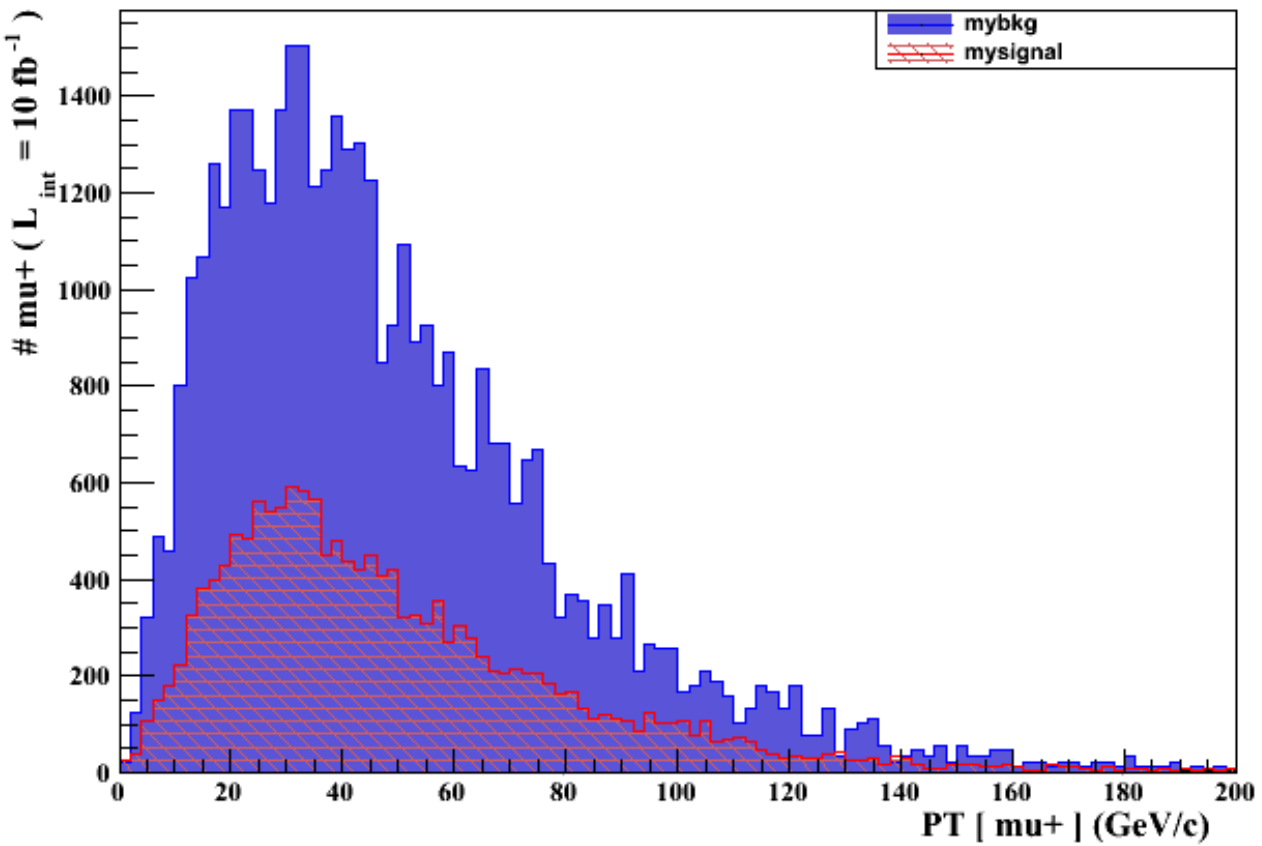
All the selection results can be gathered in a report. MadAnalysis supplies 3 commands corresponding to the format of the report :

- generate_latex : produce a report compilable by latex
- generate_pdflatex : produce a report compilable by pdflatex
- generate_html : produce a report in HTML format

Step 5 : diplaying results

Dataset	# events	Mean	RMS	% Underflow	% Overflow
mybkg	42751	48.9768	31.5	0.0	0.4688
mysignal	15939	49.4274	31.7	0.0	0.5338

Histogram number 1 - Statistics



Interactive demo 1

1. opening a MadAnalysis session
2. displaying (multi)particles and creating new ones
3. importing samples and gathering them in datasets
4. defining a selection including plots and cuts
5. launching the selection on samples
6. generating a HTML report

Advanced functionnalities

MadAnalysis 5 modes

By default, MadAnalysis is configured for analyzing MC samples generated at *parton level*. Configuring the program in *hadron level* or *reconstructed level* mode requires to open a new session with the appropriate argument.

mode	argument for launching MadAnalysis	shortcut
parton level	bin/ma5 −−partonlevel	bin/ma5 -P
hadron level	bin/ma5 −−hadronlevel	bin/ma5 -H
reconstructed level	bin/ma5 −−recolevel	bin/ma5 -R

Main consequences :

- New observables are available in the case of reconstructed object :
  - ISOL : isolated lepton ?,
  - HE_EE : hadronic energy over electomagnetic energy
  - NTRACKS : number of tracks in a jet.
- The initial list of particles loaded at the session start differs.

Import : a multi-purpose command

`import` command has been designed to extract information from external files and to fill MadAnalysis objects with them. The syntax is :

```
import [file/directory]
```

The action carried out by this command depends on the type or the content of the input file/directory :

- importing Monte Carlo samples and gathering them into datasets.
- importing particles from UFO model. Stable, electrically and colored neutral are also included into 'invisible' multiparticle (except photon).
- restoring a MadAnalysis configuration from a submitted job. Available soon.

Different kinds of particle combination

The command below creates an histogram related to the transverse momentum of the vectorial sum of the μ^- and the μ^+ momenta. This can be changed by adding a **'prefix'** to the observable.

```
plot PT(mu+ mu-) 100 0 100
```

- vPT : PT of the vectorial sum of the muon momenta
- sPT : scalar sum of the muon PT
- vdPT or dvPT or dPT : PT of the vectorial difference of the muon momenta
- sdPT or dsPT : scalar difference of the muon PT
- rPT : ratio defined by the difference between PT(mu+) and PT(mu-) over PT(mu+)

Selecting a particle according to its PT rank

Squared brackets **[ ]** allow to select a particle according to its rank in PT.

```
plot PT(mu+[1]) 100 0 100     # leading muon
plot PT(mu+[2]) 100 0 100     # next-to leading muon
```

The muons with the smallest PT can be selected by negative index.

```
plot PT(mu+[-1]) 100 0 100   # the muon with the smallest PT
plot PT(mu+[-2]) 100 0 100
```

Ordering observables can be changed :

```
set selection[1].rank = ETordering
```

Mother-to-daughter relations

Another way to select only one particle is to use the history of the particle. Two operators allows to do that : $<$ and $<<$.

- **operator** $<$
  allows to specifiy the identity of the mother

```
plot PT(mu+ < Z) 100 0 100     # positive muon coming
                               # from a Z boson
```

- **operator** $<<$
  allows to specifiy one of the descendants of the particle

```
plot PT(mu+ << t) 100 0 100     # positive muon coming
                                # from the cascade-decay
                                # of a top quark
```

Several mother-to-daughter operators could be combined. For instance,

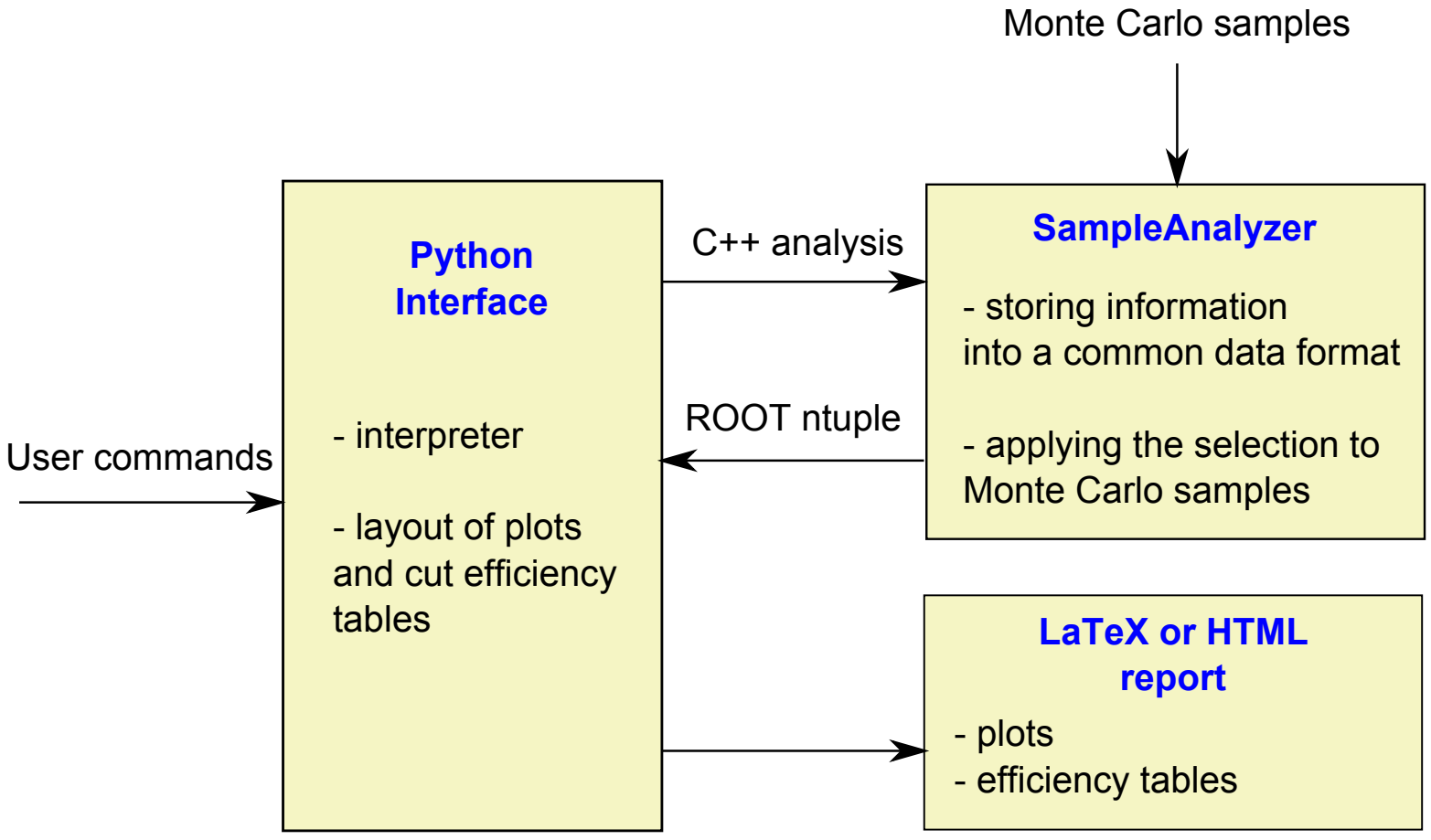
```
plot PT(mu+ << t < st1) 100 0 100
```

Warning : this option is forbidden in reconstructed level mode.

Developer-friendly mode

What is the developer-friendly mode ?

Writing the code of your analysis inside the SampleAnalyzer kernel, without using the Python interface. → **C++ skills are required !**

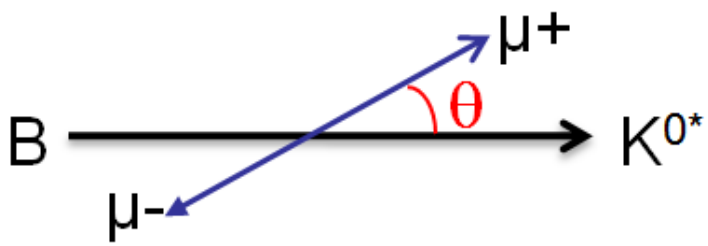


Why a developer-friendly mode ?

Several reasons could motivate the use of the expert mode :

- performing plots more sophisticated than 2D histograms.
- plotting (or cutting on) an observable which is not implemented in the program.

  example : new physics research in $B_d^0 \to K^{0*}\mu^+\mu^-$.
  asymmetry between forward and backward moving μ^+ versus the B meson direction in the $\mu^+\mu^-$ rest frame.



$$A_{FB}(s = m^2_{\mu+\mu-}) = \frac{N_F - N_B}{N_F + N_B}$$

- producing a result file in a specific format
- ...

Framework possibilities

SampleAnalyzer can be seen as a genuine framework which has been designed in order to be efficient and very simple to be used.

Developer-friendly qualities of the framework :

- compilation and linking recipe is fully automated including ROOT libraries.
- event information is stored in a common data format whatever the input sample format.
- library of physics functions is available.
- several services facilitate the task of the developer (logger, exception handle, ...).
- doxygen documentation on the web site available soon.

Interactive demo 2

1. access to Expert mode
2. setting the environment
3. editing the existing analysis
4. implementing a new analysis
5. launching and selecting an analysis

Summary

Main functionalities are implemented ... but some points must be finalized :

- Minor options such as restoring a MadAnalysis session with the command 'import'.
- Code validation : J. Andrea and master students test and use MadAnalysis for their private phenomenological investigations. A test suite will be soon available for checking the program in a exhaustive way
- A user's guide (50 pages already written)

Beta release

Be available, with the draft of the paper, for MadGraph and FeynRules collaboration : April, the 6^{th} (Easter day).

Public release

Planned for the end of April, including potential fixes resulting from the first feedback of the beta testers.

What to expect for the next public releases

Current to-do list :

- Making MadAnalysis 5 independent from libraries BOOST and XDR.
- Automating the treatment for binning of histograms.
- Determining precisely the time budget of typical jobs and optimizing the algorithms.
- Interfacing MadAnalysis to MadGraph (collaboration work with Olivier).
- Matching (collaboration work with Rick).
- Analyzing on-flight events produced by Next-To-Leading order generator.
- Implementing jet clustering algorithms (FastJet) and basic detector simulation.