Designing and recasting LHC analyses with MadAnalysis 5

E. Conte[(a)]
eric.conte@uha.fr

B. Fuks[(b,c)]
fuks@cern.ch

T. Schmitt[(b)]
thibaut.schmitt@iphc.cnrs.fr

(a) GRPHE, Université de Haute-Alsace, IUT Colmar, 34 rue du Grillenbreit BP 50568, 68008 Colmar Cedex, France
(b) IPHC, Université de Strasbourg/CNRS-IN2P3, 23 Rue du Loess, F-67037 Strasbourg, France
(c) CERN, PH-TH, CH-1211 Geneva 23, Switzerland

FINANCÉ PAR ANR

What is MadAnalysis 5?

MadAnalysis 5 is a public program that allows high-energy physicists (theorists and experimentalists) to efficiently design and recast LHC analyses. Phenomenologists can, in this way, investigate their favorite models, and determine whether the LHC is sensitive to a given signature by either conceiving a novel analysis or by recasting existing CMS or ATLAS studies.

Natively associated with MadGraph 5, MadAnalysis 5 is now able to read the output of any Monte Carlo generator (at leading order or next-to-leading order QCD accuracy). Analyzing event samples consists in histogramming distributions of observables, applying some selection cuts and building a cut-flow chart. All results are summarized into a HTML, PS or PDF report.

According to the wishes of the user, MadAnalysis 5 can process the events before analyzing them. In particular, a jet-clustering algorithm or a detector simulation can be used effortlessly. Besides processed events can be saved in output files.

Main concepts

A unique framework

The MadAnalysis 5 package **[1]** allows one to design/recast in a same way phenomenological investigations at any step of the generation (parton, hadron and reconstructed object level), for any file format (STDHEP, HEPMC, LHE, LHCO, ROOT).

User-friendly

The user designs her/his analysis by interacting with a Python console. Settings and analyses can be written with the help of a metalanguage designed to be intuitive. Tab completion and in-line help facilitate the life of the user.

Efficient

The Python console exports the analysis encoded using the MadAnalysis 5 metalanguage to a dedicated C++ program readily to be compiled and executed.

Flexible

MadAnalysis 5 is shipped with a series of common built-in observables that includes sophisticated variables such as α_T or M_{T2}. Operations between four-momenta are also available. For more complicated selections, the user can directly write the analysis in C++ (the so-called ***expert mode***) and design her/his own observables.

Multi-interface

MadAnalysis 5 is interfaced to several packages: Gzip, Root, FastJet, FastJet-Contrib, Delphes. Installation of these packages can be done easily from the Python console. It is also distributed with Delphes-MA5tune, a modified version of Delphes.

Program summary

Current release: MadAnalysis 5 v1.1.12
Platforms: Unix, Linux, Mac OS X
Programming language: Python, C++
Requirements: Gcc, Python, Make, Root
Software License: GNU General Public License
Official web-site: https://launchpad.net/madanalysis5
Tutorials: https://madanalysis.irmp.ucl.ac.be/wiki/tutorials
Physics Analysis Database:
https://madanalysis.irmp.ucl.ac.be/wiki/PhysicsAnalysisDatabase

References

[1] E. Conte, B. Fuks, G. Serret, CPC 184 (2013) 222
[2] E. Conte, B. Dumont *et al*, *submitted to EPJC,* arXiv:1405.3982
[3] M. Cacciari, G.P. Salam, G. Soyez, EPJC 72 (2012) 1896
[4] J. De Favereau *et al*, JHEP 02 (2014) 057
[5] B. Dumont, B. Fuks *et al*, *submitted to EPJC,* arXiv:1407.3278
[6] A. Buckley, J. Butterworth *et al*, CPC 184 (2013) 2803
[7] A. Alloul, M. Frank, B. Fuks *et al*, JHEP 1310 (2013) 033

Acknowledgments

We are grateful to B. Dumont, S. Kraml and C. Wymant for helpful discussions. Many thanks to A. Alloul, J. Andrea, L. Basso, S. Bein, M. Blanke, K. de Causmaecker, E. Chabert, G. Chalons, C. Collard, I. Galon, S. Kulkarni, K. Mawatari, L. Mitzka and D. Sengupta for their help in testing and debugging the code.
This work has been partially supported by the French ANR project 12-JS05-002-01 BATS@LHC and the Theory LHC France initiative of the CNRS/IN2P3.

Analysis design based on the MadAnalysis 5 metalanguage

User side: the metalanguage

```
import DrellYan*.lhe as dy
import ttbar_semilep*.lhe as tt
import ttbar_dilep*.lhe as tt
```

optional

```
set detector.fastsim.package = fastjet
set detector.fastsim.algo = kt
set detector.fastim.bjet.efficiency = 0.5
```

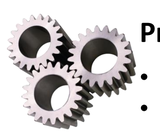
```
plot MET
plot M(mu+ mu-)
plot PT(j[1]) # the hardest jet

select 70 < M(mu+ mu-) < 110
define mu = mu+ mu-
select (mu) PT > 25
reject N(mu)< 2
```

optional

```
set main.outputfile = output.lhco
```

Behind the scene

Reading signal and background event files
Supporting any event file formats
(STDHEP, HEPMC, LHE, LHCO, ROOT)



Processing data
- Detector simulation
- Predefined analyses

Analyzing data
- Filling histograms
- Selection cuts

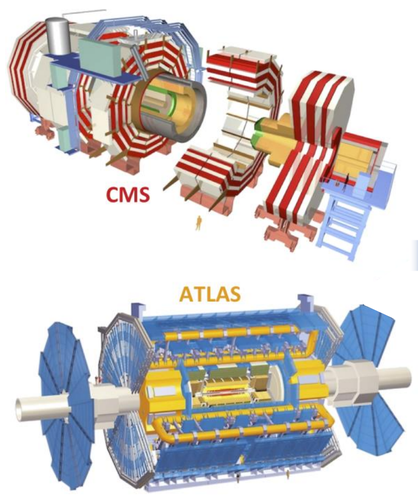
Storing events into output files
LHE, LHCO, ROOT file formats

Report Generator
three possible "graphical drivers": Root, Matplotlib or Gnuplot

available in the next release

Analysis report HTML or PDF
- histograms
- selection efficiencies
- figures of merit

Three available packages devoted to detector fast-simulation



Tunable simulation based on FastJet

Any of the jet-clustering algorithms included in the FastJet package **[3]** can be applied to the content of an hadron-level event. Options allow to switch on one or several detector effects applied on physics objects: efficiency, resolution, mis-identification.

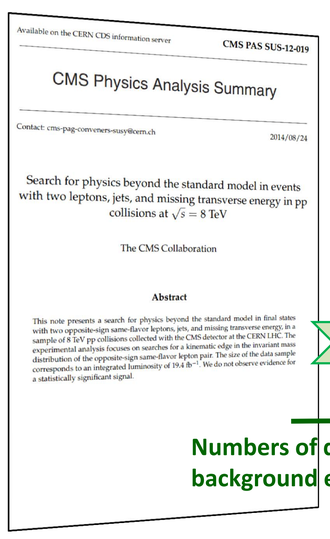
Delphes 3

MadAnalysis 5 is interfaced to Delphes 3 **[4]**. This public package offers the most realistic description of the ATLAS and the CMS detectors, including pile-up effects.

Delphes-MA5tune

This package is a modified release of Delphes 3, providing new observables and optimizing the output file. The main improvement deals with lepton (tracker/calorimeter/particle flow) isolation which is tuned and can be accounted for at the analysis level.

Recasting an existing ATLAS or CMS analysis with the MadAnalysis 5 *expert mode* [2]


Available on the CERN CDS information server — CMS PAS SUS-12-019

CMS Physics Analysis Summary

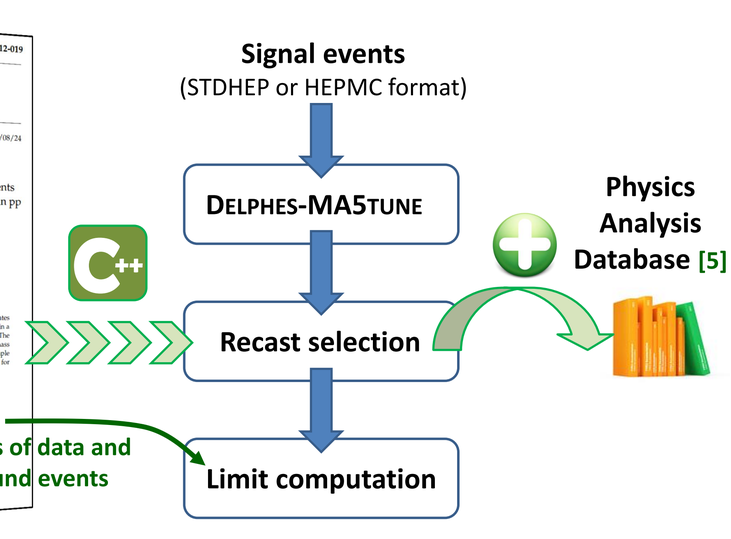
Contact: cms-pag-conveners-susy@cern.ch — 2014/08/24

Search for physics beyond the standard model in events with two leptons, jets, and missing transverse energy in pp collisions at $\sqrt{s} = 8$ TeV

The CMS Collaboration

Abstract

This note presents a search for physics beyond the standard model in final states with two opposite-sign same-flavor leptons, jets, and missing transverse energy, in a sample of 8 TeV pp collisions collected with the CMS detector at the CERN LHC. The experimental analysis focuses on searches for a kinematic edge in the invariant mass distribution of the opposite-sign same-flavor lepton pair. The size of the data sample corresponds to an integrated luminosity of 19.4 fb^{-1}. We do not observe evidence for a statistically significant signal.



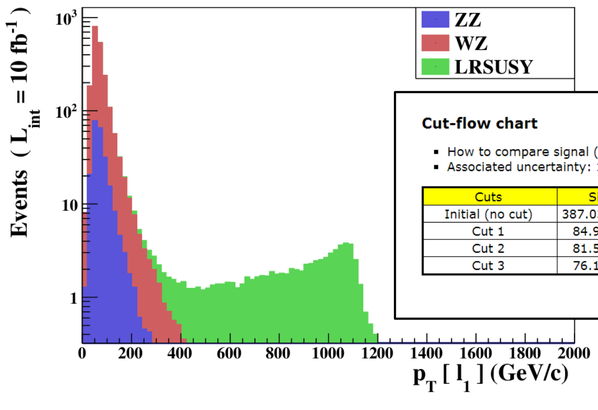
The flexibility of the *expert mode* (= writing the selection in C++) allows one to take into account any cut-based ATLAS or CMS analysis. Developer-friendly, it provides a large collection of high-energy-physics-oriented functions, services allowing one, for instance, to produce cut-flow charts and/or histograms. Analyses with multiple sub-analysis (or signal/control regions) can also be addressed.

Features:
- Complementary to the Rivet approach **[6]**.
- Efforts to improve the content and the realism of detector simulation packages.
- Helping the experimentalists to interpret their results by highlighting relevant theoretical models.
- Identification of possible topologies not scrutinized by the LHC experiments.

Examples of results

Investigating a LR-SUSY multileptonic signature for the 2015 LHC data [7]

Signal = W'production @ a Left-Right Supersymmetric (LR-SUSY) model
Background = Standard Model production of dibosons



Selection:
1. more than 3 charged leptons,
2. highest lepton P_T > 200 GeV,
3. $\not{E}_T$ > 50 GeV

Cut-flow chart

- How to compare signal (S) and background (B): S/B.
- Associated uncertainty: 1./(B**2)*sqrt(B**2*ES**2+S**2*EB**2).

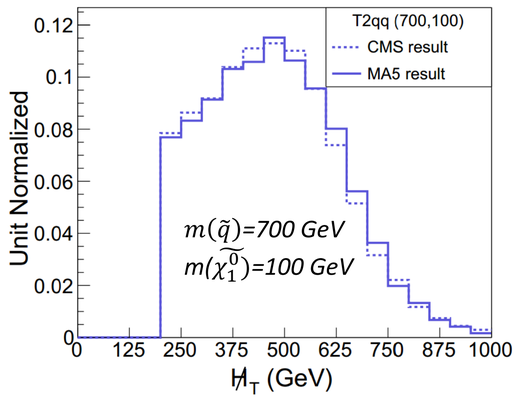
Cuts	Signal (S)	Background (B)	S vs B
Initial (no cut)	387.033 +/- 0.584	2029.72 +/- 1.44	0.190683 +/- 0.000318
Cut 1	84.95 +/- 8.14	2029.72 +/- 1.44	0.04185 +/- 0.00401
Cut 2	81.56 +/- 8.02	26.09 +/- 5.08	3.126 +/- 0.681
Cut 3	76.19 +/- 7.82	19.45 +/- 4.39	3.92 +/- 0.97

Signal and Background comparison

Extracted from a MadAnalysis 5 report

Recasting of a SUSY multijet analysis [5]

Comparison between the CMS (CMS-SUS-13-012 analysis) and MadAnalysis 5 results for the $\not{H}_T$ distribution (vector sum of the jet transverse momenta) for the "T2qq" simplified model.



16th International workshop on Advanced Computing and Analysis Techniques in physics research (ACAT), 1-5 September 2014, Prague (Czech Republic)