MadAnalysis 5 - status and plans

Benjamin Fuks (IPHC Strasbourg / Université de Strasbourg)

In collaboration with E. Conte & G. Serret.

MadGraph spring meeting @ Pittsburgh
May 10-12, 2012

Outline

1. Introduction.
2. Overview of MadAnalysis 5.
3. Examples for normal and expert users.
4. Development plans.
5. Conclusions.

Comprehensive particle physics phenomenology.

1. Implementation of a new physics model in FeynRules.
   [Christensen, Duhr (CPC '09); Christensen, Degrande, Duhr, BenjF (in prep)]
2. Automated derivation of the UFO **files**.
   [Degrande, Duhr, BenjF, Grellscheid, Mattelaer, Reiter (CPC '12)]
3. Event generation with MadGraph **5**. [Alwall, Herquet, Mattelaer, Stelzer (JHEP '11)]
   - or with any **matrix element generator**.

   Parton-level phenomenology.
4. Parton showering and hadronization with Pythia or Herwig.
   [Sjostrand, Mrenna, Skands (JHEP '06; CPC '08); Corcella *et al.* (JHEP '01); Bahr *et al.* (EPJC '08)]
   - or with any **parton showering tool**.

   Hadron-level phenomenology.
5. Fast detector simulation with Delphes or PGS.
   [Ovyn, Rouby, Lemaitre ('09); Conway ('06)]
   - or with any **fast detector simulation algorithm**.

   Reconstructed-level phenomenology.

Need for a new framework for collider phenomenology.

- **Several levels of sophistication for the phenomenological analyses**.
  * **Parton**-level.
  * **Hadron**-level.
  * **Reconstructed**-level.
- **Analysis skeleton**.
  * **Reading** of signal and background event files.
  * **Selection cuts** on both signal and background events.
  * Creation of **histograms** and **cut-flow charts**.
  * **Extraction of information on the signal** swamped by the backgrounds.
- **Drawbacks**.
  * The procedure above is in general based on **home-made tools**.
    - ▶**Lack of traceability**.
    - ▶**Validation of the tools**?
    - ▶**Reproducibility of the results**?
  * These tools can in general only be used at a **specific sophistication level**.
    - ▶**Lack of flexibility**.
  * These tools can in general only be used with **specific event file format**.
    - ▶**Lack of flexibility**.

Introducing MadAnalysis 5.

Alleviation of these issues.

- A **new unique** framework for phenomenological analyses.
  - **Any sophistication level** (parton, hadron, reconstructed).
  - **Any event file format** (StdHep, HepMC, LHE, ...).
  - **User-friendly** $\Rightarrow$ professional analyses in a simple way.
  - **Flexible** $\Rightarrow$ no limit on the analysis complexity.
  - **Easy to maintain**.
  - **Easy to validate**.

This framework is called MadAnalysis 5.

[Conte, BenjF, Serret (June '12)]

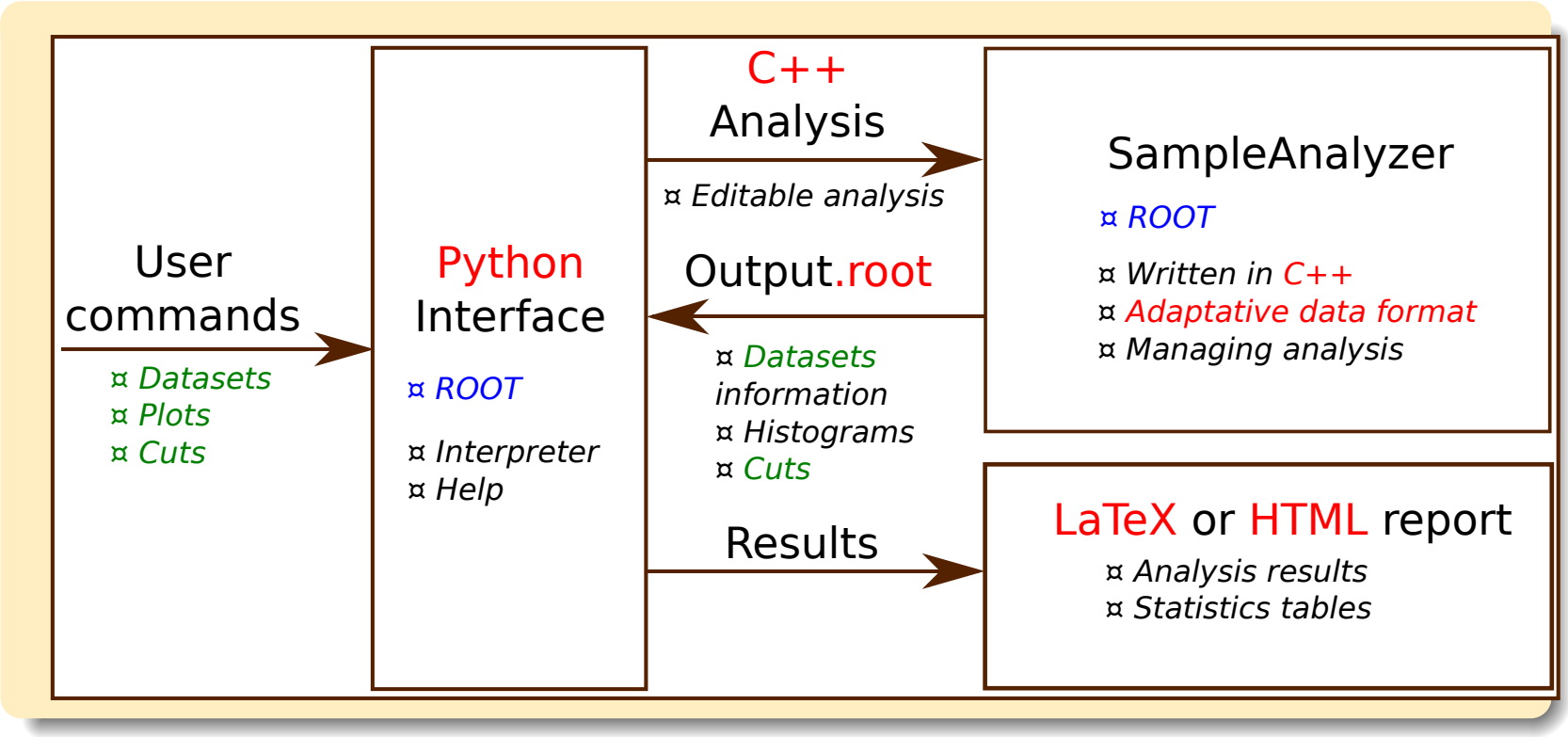
Outline.

1. Introduction.
2. Overview of MadAnalysis 5.
3. Examples for normal and expert users.
4. Development plans.
5. Conclusions.

From MadAnalysis 4 to MadAnalysis 5.

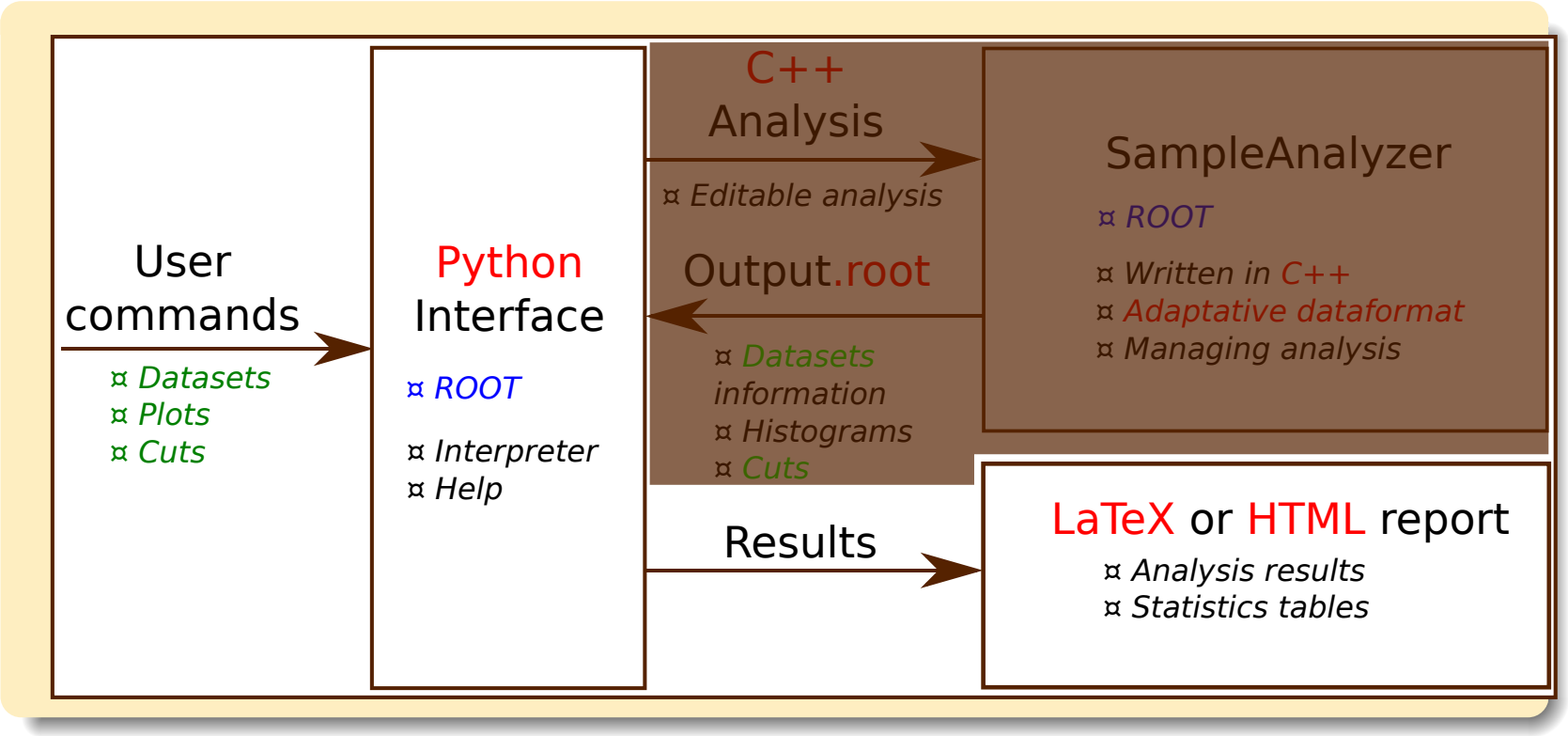
- **Object-oriented programming language**.
  * MadAnalysis 4: Fortran.
  * MadAnalysis 5: C++ core; Python interface; uses Root.
- **Flexibility**.
  * MadAnalysis 4: **No**.
  * MadAnalysis 5: **Yes**.
- **User-friendly**.
  * MadAnalysis 4: **A complicated plot card**.
  * MadAnalysis 5: **Intuitive Python commands**.
- **Limitations**.
  * MadAnalysis 4: **What is implemented**.
  * MadAnalysis 5: **The user's imagination**.
- **MadAnalysis 5 is going beyond the Mad-suite of programs.**
  * Can be used as a **standalone package**.

The MadAnalysis 5 scheme (1).



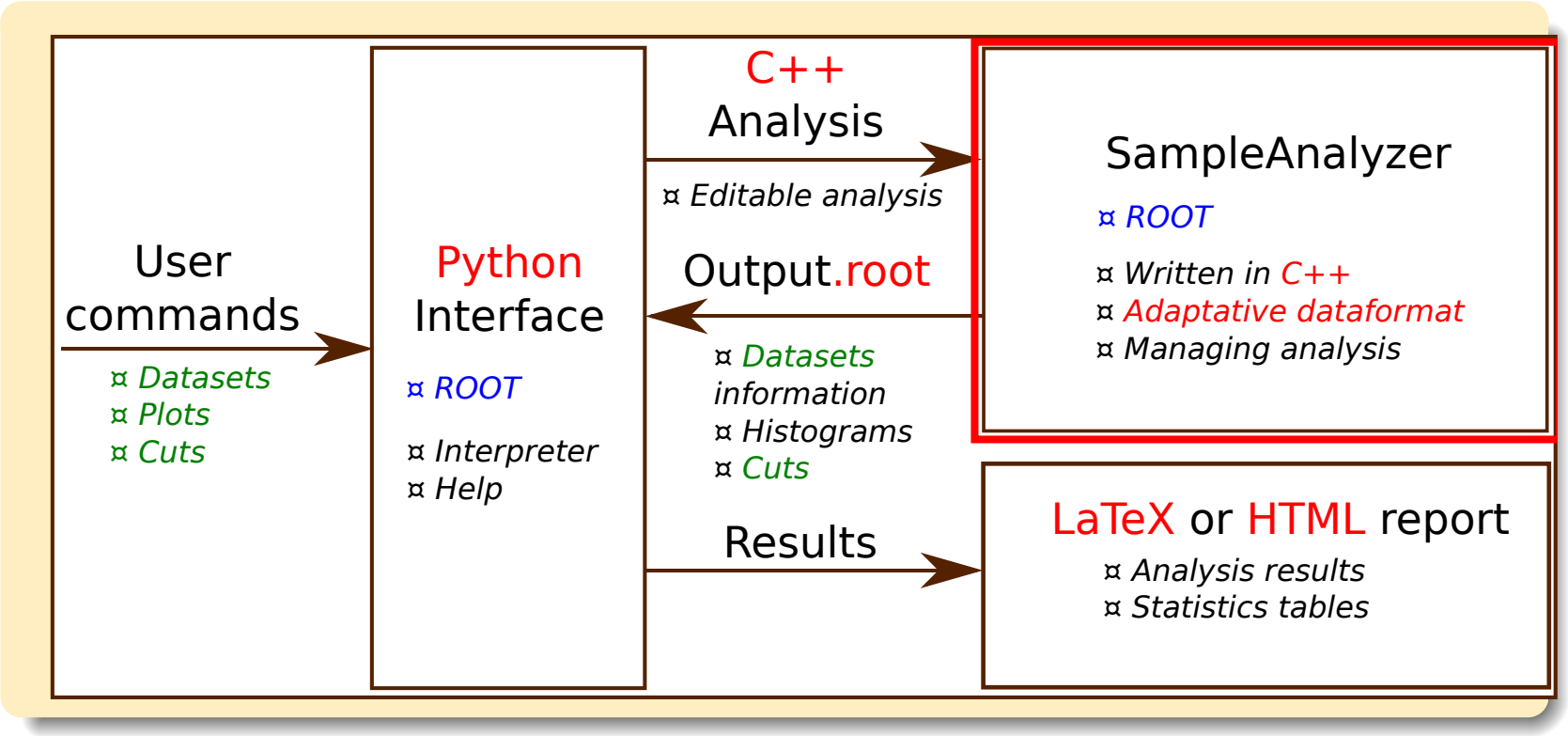
- **Main features**.
  - **Basic concepts**: datasets, plots, cuts.
  - **Python interface**: from commands to a C++/ROOT analysis.
  - **Human readable output**: HTML, LaTeX.

The MadAnalysis 5 scheme (2).



- **Normal users**.
  - The core is a **black box**.
  - The knowledge of ROOT **not necessary**.
  - Communication through PYTHON **commands**.

The MadAnalysis 5 scheme (3).



- **Expert users**.
  - Implementation of the analysis within the SampleAnalyzer **framework**.
  - C++ **and** Root **skills required** ⇒ developer-friendly.
  - The Python interface creates a **blank analysis as a starting point**.

Basic concepts.

- **Command line interface**.
  * **In-line help**.
  * **Auto-completion**.

```
ma5> help <command>
```

- **Particles and multiparticles**.
  * Particle are defined by **labels**.
  * A label points to one or several **PDG-id(s)**.
  * **MSSM + SM labels**: automatic.
  * Can be loaded from **UFO files**.
  * Labels can be **created and deleted**.
    ▶`define` and `remove`.

```
define tau = tau+ tau-
define mytau+ = -15
remove mytau+
```

- **Datasets**.
  * A dataset is a **label**.
  * **Collects** similar event samples.
  * Treated **in the same way** by MadAnalysis 5.
  * **Formats**: LHE, LHCO, StdHep, HepMC.

```
import tt1.hep as ttbar
import tt2.hep as ttbar
import Wj1.hep as Wjets
import Wj2.hep as Wjets
```

Plots and cuts.

- **The command** `plot`.
  - Creation of an **histogram**.
  - **Global observables** $\Leftrightarrow$ the entire event.
  - **Properties of the particles** in the event.
  - **Ordering** of the particles.
  - **Combining** particles
    - ►Sum and differences.
    - ►Vectorial or scalar.
  - Linear or logarithmic scales.

```
plot MET
plot N(mu)
plot PT(mu[1])
plot ETA(mu) [logY]
plot M(mu[1] mu[2])
plot dM(mu+ mu-)
```

- **Cuts**.
  - **Selecting/rejecting** events.
  - **Selecting/rejecting** particles.
    - ►not rejecting the event.
  - Still **under development**.

```
reject MHT < 50
select (mu) PT > 50
```

- **Executing the analysis**: `submit`.
- **Reports**.
  - HTML reports.
  - LaTeX reports.

```
generate_html <dir>
generate_latex <dir>
generate_pdflatex <dir>
```

Outline.

1 Introduction.

2 Overview of MadAnalysis 5.

3 Examples for normal and expert users.

4 Development plans.

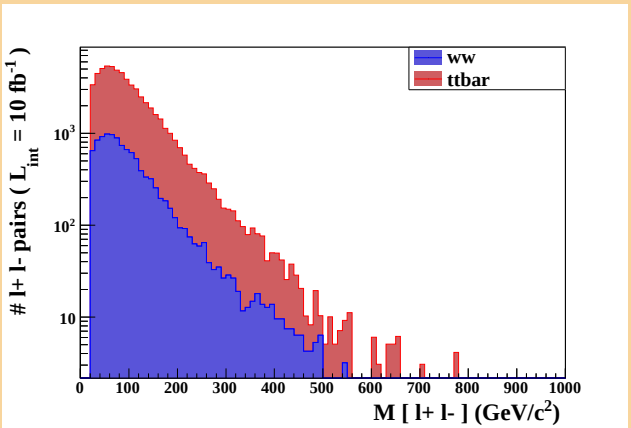
5 Conclusions.

Particle properties.

- **Kinematical distributions related to particle species**.
  * **Intuitive** commands.
  * **Available observables**:
    `BETA`, `DELTAR`, `E`, `ET`, `ETA`, `GAMMA`, `M`, `MT`, `P`, `PHI`, `PT`, `PX`, `PY`, `PZ`, `R`, `THETA`, `Y`.
  * Scalar and vectorial sums/differences are implemented.

- ▶$t\bar{t}$ (dileptonic mode).
- ▶WW (dileptonic mode).
- ▶LHC @ 8 TeV; 10 fb^{-1}.
- ▶Parton-level.
- ▶Dilepton invariant mass `M(l+ l-)`.

```
import ttbar_ll.lhe.gz as ttbar
import ww_ll.lhe.gz as ww
plot M(l+ l-) [logY]
submit tempdir
generate_latex temp_tex
```

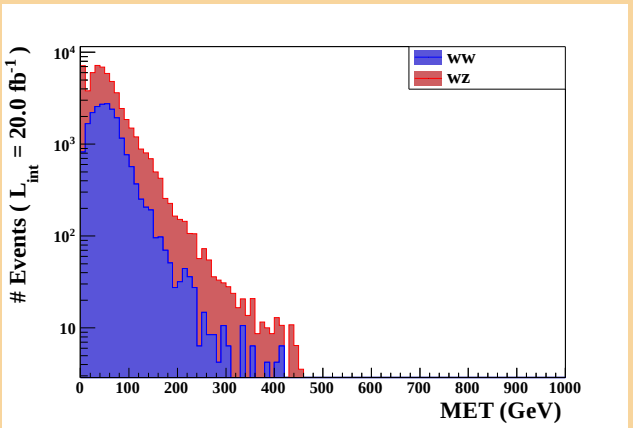


Global event observables.

- **Global event kinematical observables**.
  - * **Missing and visible energy** of the event `MET`, `ET`.
  - * **Missing and visible hadronic energy** of the event `MHT`, `HT`
  - * **Partonic center-of-mass energy** `SQRTS`.

►*WW* (dileptonic mode).
►*WZ* (decay to at least one lepton).
►LHC @ 8 TeV; 20 fb^{-1}.
►Parton-level.
►Missing energy distribution.

```
import ww_ll.lhe.gz as ww
import wz_l.lhe.gz as wz
plot MET [logY]
set main.lumi = 20
submit tempdir
generate_latex temp_tex
```

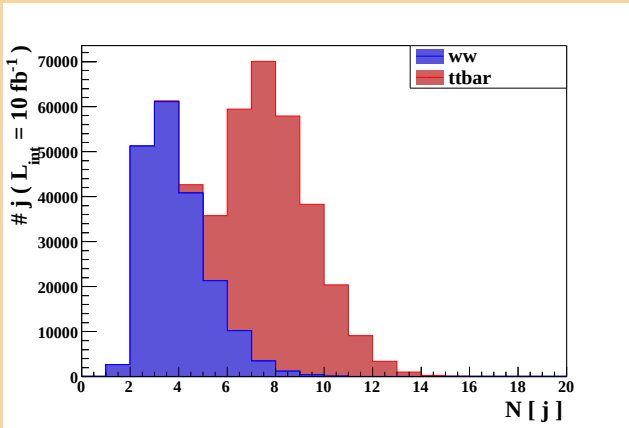


Multiplicities.

- **Particle content**.
  - * **Particle content** of the event `NPID`, `NAPID`.
  - * **Particle multiplicity** `N`

►$t\bar{t} + 0, 1, 2$ jets (hadronic mode).
►$WW + 0, 1, 2$ jets (semileptonic mode).
►LHC @ 8 TeV; 10 fb^{-1}.
►Hadron-level.
►Jet multiplicity.

```
import ttbar_hh.lhe.gz as ttbar
import ww_l.lhe.gz as ww
define j = j b b~
plot N(j)
submit tempdir
generate_latex temp_tex
```

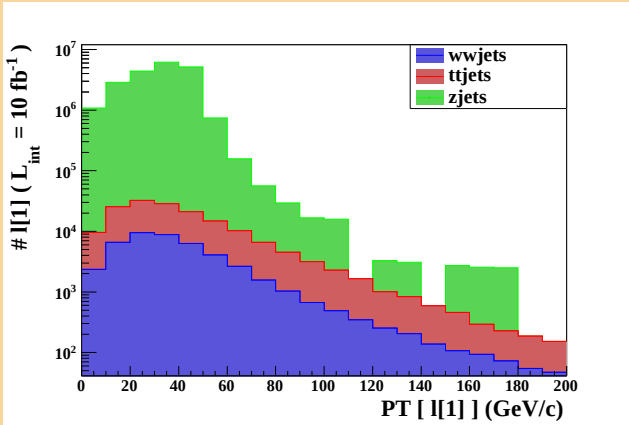


Leading lepton properties.

- **Particle ordering**.
  - Can be access with the **squared brackets** `[<i>]` .
  - Several possible **ordering variables**. `E`, `ET`, `ETA`, `P`, `PT`, `PX`, `PY`, `PZ`.

- $Z + 0, 1, 2, 3, 4$ jets (dileptonic mode).
- $WW + 0, 1, 2$ jets (dileptonic mode).
- $t\bar{t} + 0, 1, 2$ jets (dileptonic mode).
- LHC @ 8 TeV; 10 fb^{-1}.
- Hadron-level.
- **Energy ordering**.
- **Leading lepton p_T**.
- The **binning is specified**.

```
import z.lhe.gz as zjets
import ttbar.lhe.gz as ttjets
import ww.lhe.gz as wwjets
define l = l+ l-
plot PT(l[1]) 20 0 200 [logY]
set selection[1].rank = Eordering
submit tempdir
generate_latex temp_tex
```



Expert users: W-boson polarization (1).

- **Property to be investigated**.
  - **Polarization of the W issued from a top leptonic decay**.
  - Process: **$t\bar{t}$ in the semileptonic decay channel**.
  - Property investigated to an **angular distribution** $\mathrm{d}\sigma/\mathrm{d}\cos\theta^\star$.
- **The angle $\theta^\star$ is the angle between**:
  - The momentum of the W evaluated in the top rest frame.
  - The momentum of the lepton evaluated in the W rest frame.
- **Developer-friendly implementation**:
  - Only the **relevant part** of the analysis is presented here.
    - ▶ Event processing, particle identification, histogram creation $\Rightarrow$ manual.
  - Employing the built-in `ToRestFrame` and `angle` methods.
  - The momentum of the **lepton** is evaluated in the W rest frame.

```
PHYSICS->ToRestFrame(lepton,w);
```

  - The momentum of the W is evaluated in the top rest frame.

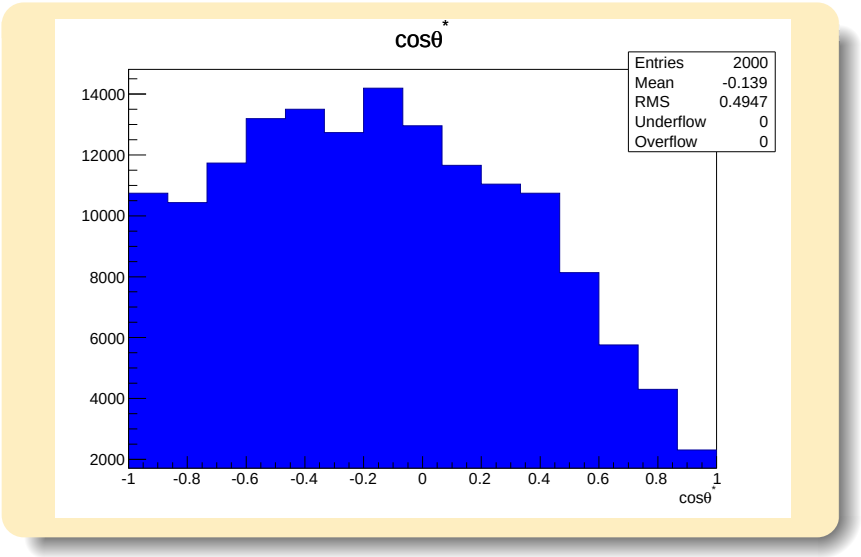
```
PHYSICS->ToRestFrame(w,top);
```

  - Filling the **histogram**:

```
histo->Fill( cos(lepton.angle(w)));
```

Expert users: W-boson polarization (2).

- **Property to be investigated**.
  - **Polarization of the W issued from a top leptonic decay**.
  - Process: **$t\bar{t}$ in the semileptonic decay channel**.
  - Property investigated to an **angular distribution** $\mathrm{d}\sigma/\mathrm{d}\cos\theta^\star$.
- **The angle $\theta^\star$ is the angle between**:
  - The momentum of the W evaluated in the top rest frame.
  - The momentum of the lepton evaluated in the W rest frame.
- **Parton-level results**.



Outline.

1. Introduction.
2. Overview of MadAnalysis 5.
3. Examples for normal and expert users.
4. Development plans.
5. Conclusions.

Development plans.

- **Three milestones**.
  * **Now**: beta version 0.6.5 available on request.
    - `http://www.madanalysis5.com`
    - `ma5team@iphc.cnrs.fr`
  * **June '12**: public release, version 5.1.0.
  * **Fall '12**: version 5.1.2.
  * **Christmas '12**: version 5.2.0 (beta).
- **On the way to MadAnalysis 5.1**
  * **Beta-testing**: bugs appeared (and will appear) $\Rightarrow$ to be fixed.
  * **Cuts**: some unsupported features remain to be implemented.
  * **Implementation of ΔR**: fixes required.
  * **Style and layout**: improvements necessary here and there.
  * **Licensing** (GPL adopted).
  * Achievement of the **manual**.

Fall '12: release of MadAnalysis 5.1.2.

- **Major improvements are scheduled**.
  * **Interface with MadGraph 5**.
    - Automatic installation.
    - Automatic plots after event generation.
  * **Matching** plots.
  * **Interface with FastJet**.
    - New hep2lhe-like package.
  * **Tutorials**.
    - FeynRules-MadGraph 5 school in Natal.
  * **Timing** service
    - optimization of the code.
  * **Exception** service.

Christmas '12: to a fast detector simulation (1).

- **Why a fast detector simulation?**
  - There are already two codes: PGS **and** DELPHES.
  - Each of those have **limitations** with respect to our needs.
    - Jet energy scale.
    - Complicated efficiency functions.
    - *etc...*
  - **Only two options**.
    1. **Hack** those codes according to our needs.
    2. **Implement** a fastsim in MADANALYSIS 5 with the required features.
- **Our choice: option #2.**
  - **No need to dig into other's code**.
  - More **flexibility, traceability**.
  - **We can use the strength of the MADANALYSIS 5 framework**.
    - A user-friendly PYTHON interface.
    - More control.
    - Easy development of new features.
  - **We can be ready for the LHC shutdown**.
    - Not guaranteed with option #1.

Christmas '12: to a fast detector simulation (2).

- **Main development plans**
  * **Detector definition** from the command line interface.
    - Existing scripts for common detectors (CMS, ATLAS):
      `import CMS`
    - Definition of new detectors (including geometry):
      `define detector ILCdet; set ILCdet.shape = cylinder; ...`
  * **Resolution effects** from the command line interface.
    `smear(mu) = 'formula'` (can depend on p_T, η, ...)
  * **Efficiencies**.
    `set efficiency(tau) = 'formula'`
  * **Object tagging (b, c, τ,...)**.
  * **Scaling**.
    `set scaling(E(j)) = 'formula'`
- **Secondary development plans**
  * Fakes.
  * Electronic noise.
  * Cosmic rays.
  * Pile up.

Outline.

1. Introduction.
2. Overview of MadAnalysis 5.
3. Examples for normal and expert users.
4. Development plans.
5. Conclusions.

Summary.

- MadAnalysis **5 is a new framework for collider phenomenology**.
  - **Unique** ⇒ partonic, hadronic or reconstructed events.
  - **User-friendly** ⇒ Python command line interface.
  - **Flexible** ⇒ a C++ kernel.
- **A special mode for expert users exists**.
  - **Developer-friendly** ⇒ C++ and Root skills required.
  - **No limitations** ⇒ *e.g.*, the W polarization.
- **Major development plans**.
  - Interface with MadGraph **5**.
  - **Matching** plots.
  - Interface with `FastJet`.
  - **Fast detector simulation**.

Ask for the beta-version.

`http://www.madanalysis5.com`

`ma5team@iphc.cnrs.fr`