Designing and recasting physics analyses with MADANALYSIS 5

Application to Higgs EFT

Fuks Benjamin

CERN - IPHC - U. Strasbourg

MADANALYSIS 5: with E. Conte (Colmar), B. Dumont (Grenoble) & C. Wymant (ex-Annecy)
HEFT: with A. Alloul (Strasbourg) & V. Sanz (Sussex)

MC4BSM 2014 @ Daejeon, South Korea

May 23, 2014

Outline

Effective field theories for Higgs physics

✦ The effective field theory (EFT) approach

- ✤ All new phenomena are assumed to appear at a scale Λ
- ✤ No assumption on the form of new physics
  - ★ Addition of higher-dimensional operators
  - ★ We restrict ourselves to dimension six
- ✤ Renormalizable order by order in the scale Λ
- ✤ Not predictive at scales larger than Λ (loss of unitarity)

✦ In the context of the Higgs boson: construction of simple EFTs

- ✤ EFT are excellent approaches to characterize the properties of any new state
- ✤ There are two equivalent possibilities
- ✤ First possibility: [Artoisenet *et al.* (JHEP'13); Maltoni, Mawatari, Zaro (EPJC.14)]
  - ★ Couplings of the physical Higgs boson to the Standard Model (physical) states
  - ★ One operator associated with a single coupling (and Lorentz structure)
  - ★ No assumption on the Higgs boson spin

[see K. Mawatari's talk]

- ✤ Second possibility: [Alloul, BenjFuks, Sanz (JHEP'14; in prep)]
  - ★ Only using Standard Model gauge-eigenstates
  - ★ Several operators may be associated with a single coupling (in the mass basis)
  - ★ One operator associated with several couplings (in the mass basis)

[approach of this talk]

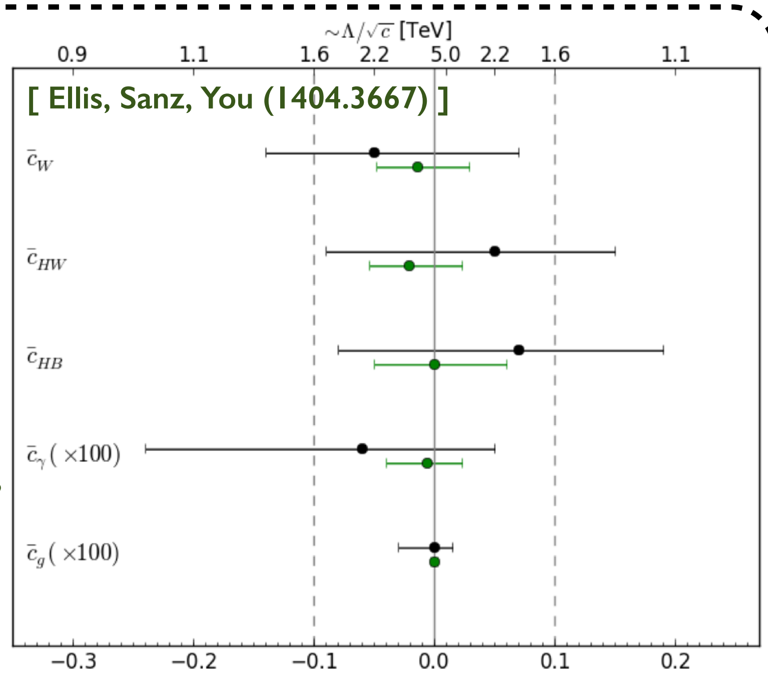
Effective operators relevant for diboson production

- In this presentation: we focus on a subset of all dimension-six operators
  - Inducing new physics effects in diboson production
  - All other operators are neglected

$$\frac{ig\,\bar{c}_W}{m_W^2}\big[\Phi^\dagger T_{2k}\overleftrightarrow{D}^\mu\Phi\big]D^\nu W^k_{\mu\nu} + \frac{ig'\,\bar{c}_B}{2m_W^2}\big[\Phi^\dagger\overleftrightarrow{D}^\mu\Phi\big]\partial^\nu B_{\mu\nu}$$
$$+\frac{2ig\,\bar{c}_{HW}}{m_W^2}\big[D^\mu\Phi^\dagger T_{2k}D^\nu\Phi\big]W^k_{\mu\nu} + \frac{ig'\,\bar{c}_{HB}}{m_W^2}\big[D^\mu\Phi^\dagger D^\nu\Phi\big]B_{\mu\nu}$$
$$+\frac{g'^2\,\bar{c}_\gamma}{m_W^2}\Phi^\dagger\Phi B_{\mu\nu}B^{\mu\nu} + \frac{g_s^2\,\bar{c}_g}{m_W^2}\Phi^\dagger\Phi G^a_{\mu\nu}G_a^{\mu\nu}$$

6 new physics parameters

- Strong constraints on those operators from Higgs data
  - Could these operators be further constrained from diboson data?
  - Total rates?
  - Differential distributions?



FeynRules [Christensen, Duhr (CPC'09)] [Alloul, Christensen, Degrande, Duhr, BenjFuks (CPC'14)]

MadGraph5_aMC@NLO [Alwall, Frederix, Frixione, Hirschi, Maltoni, Mattelaer, Shao, Stelzer, Torrielli, Zaro (1405.0301)]

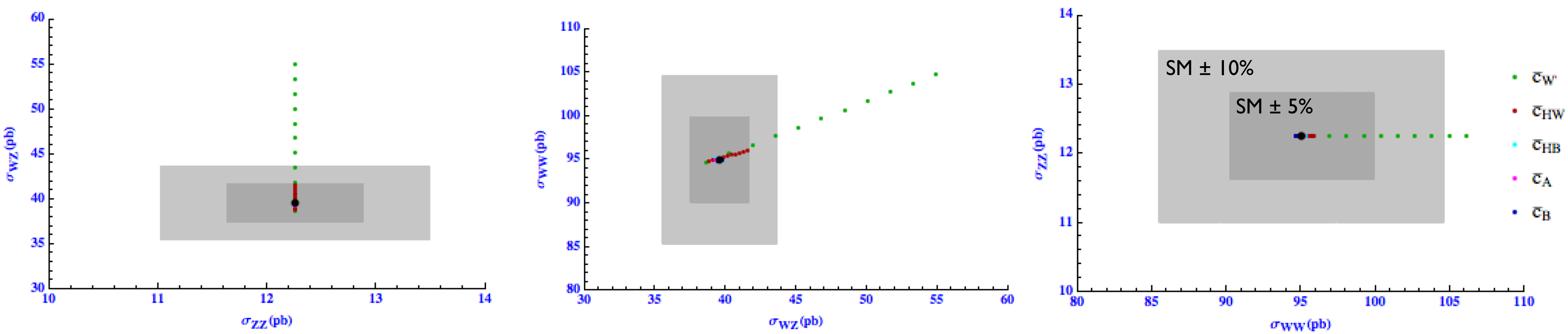
MadAnalysis 5 [Conte, BenjFuks, Seret (CPC'13)] [Conte, Dumont, BenjFuks, Wymant (1405.3982)]

Can total cross sections give hints of new physics?

✦ Study of the correlations in diboson total rates

[Alloul, BenjFuks, Sanz (in prep)]

- ✤ Induce new physics effects in diboson production
- ✤ Is an investigation of total rates enough? **NO!** (except for larger values of $\bar{c}_W$)



★ Results obtained with MadGraph 5, including next-to-leading order K-factors from aMC@NLO
★ HEFT model publicly available in FeynRules: http://feynrules.irmp.ucl.ac.be/wiki/HEL

Can differential distributions enable a more striking discrimination of operators?

→ **MadAnalysis 5**

Outline

1. Higgs effective field theories and diboson production
2. **Overview of MADANALYSIS 5 and basic concepts**
3. Analyzing dimension-six operators effects with simulated events
4. An extension of the expert mode for recasting LHC analyses
5. Summary

MADANALYSIS 5 in a nutshell

[Conte, BenjFuks, Serret (CPC '13); Conte, BenjFuks (arXiv:1309.7831); Conte, Dumont, BenjFuks, Wymant (1405.3982)]

✦ What is MADANALYSIS 5?

- ✤ A framework for phenomenological analyses
- ✤ Any level of sophistication: partonic, hadronic, detector, reconstructed
- ✤ Several input format: STDHEP, HEPMC, LHE, LHCO, ROOT (from DELPHES)
- ✤ User-friendly, flexible and fast
- ✤ Interfaces to several HEP packages to process events (fast detector simulation, jet clustering, *etc.)*

✦ Two modules

- ✤ A PYTHON command line interface (interactive; including inline help)
- ✤ A C++/ROOT core module, SAMPLEANALYZER

✦ Normal mode

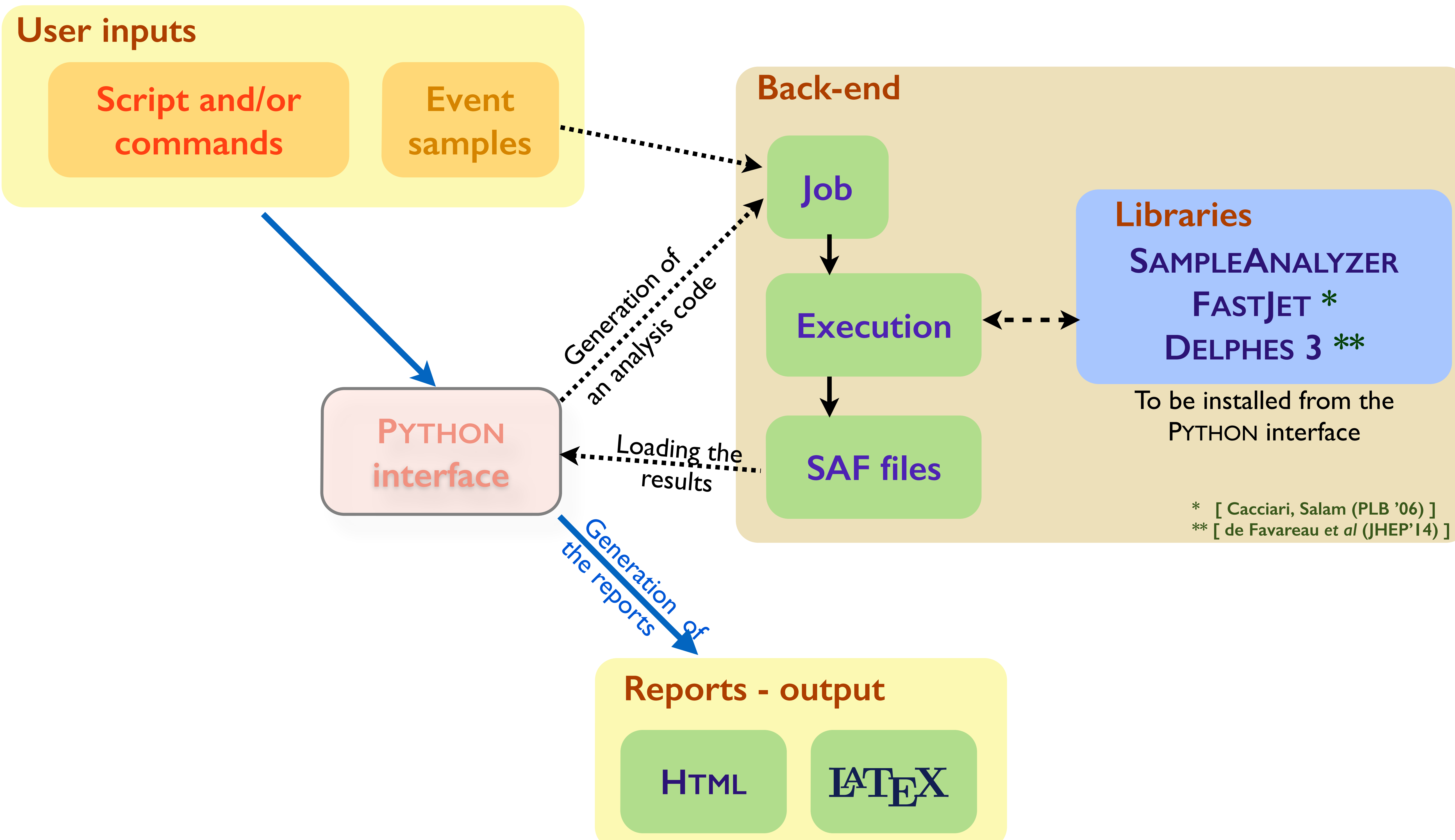
- ✤ Intuitive commands typed in the PYTHON interface
- ✤ Analysis performed behind the scenes (black box)
- ✤ Human readable output: HTML and LaTeX

✦ Expert mode: recently extended for recasting existing LHC analyses

- ✤ C++/ROOT programming within the SAMPLEANALYZER framework
- ✤ Support for multiple sub-analyses, an efficient way for handling cuts and histograms, *etc.*

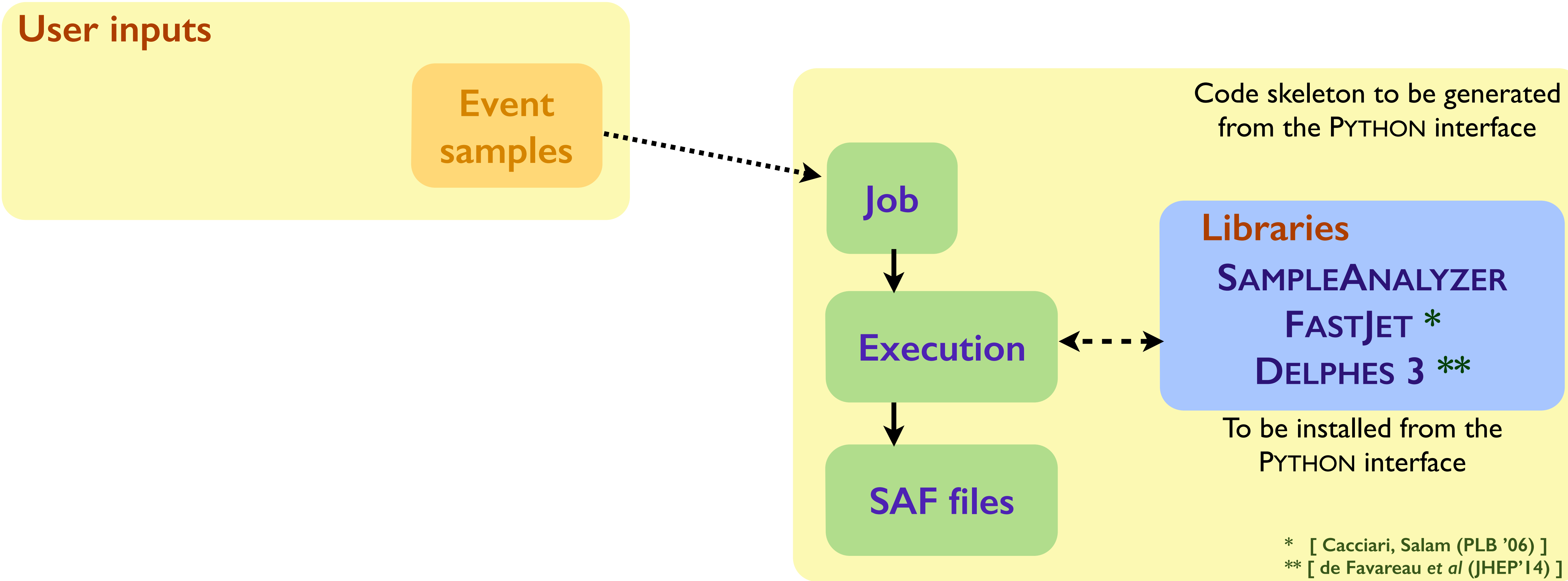
MadAnalysis 5: normal running mode

[Conte, BenjFuks, Serret (CPC '13); Conte, BenjFuks (arXiv:1309.7831); Conte, Dumont, BenjFuks, Wymant (1405.3982)]



MADANALYSIS 5: expert running mode

[Conte, BenjFuks, Serret (CPC '13); Conte, BenjFuks (arXiv:1309.7831); Conte, Dumont, BenjFuks, Wymant (1405.3982)]



Basic concepts of the normal mode (1)

Datasets

- Event file format automatically detected
- Events files associated with a label
- Supported file formats: LHE, STDHEP, HEPMC, LHCO, ROOT (from DELPHES 3)
- Several samples can be grouped (e.g., to increase statistics)

```
ma5>import ttbar* as top-antitop
   -> Storing the file 'ttbar.hep.gz' in the dataset 'top-antitop'.
   -> Storing the file 'ttbar2.hep.gz' in the dataset 'top-antitop'.
ma5>import Wjets.lhe.gz as Wboson
   -> Storing the file 'Wjets.lhe.gz' in the dataset 'Wboson'.
ma5>import VV.hep as diboson
   -> Storing the file 'VV.hep' in the dataset 'diboson'.
```

Particles and multiparticles

- Particles and multiparticles are defined via their PDG code (*labels*)
- Default: SM, MSSM, invisible, hadronic
- Can be imported from a UFO model

```
ma5>define TheMuon = 13
ma5>define TheAntiMuon = -13
ma5>define AllMuons = TheMuon TheAntiMuon
ma5>display l+
   The multiparticle 'l+' is defined by the PDG-ids -15 -13 -11.
ma5>display e+
   The particle 'e+' is defined by the PDG-id -11.
ma5>display invisible
   The multiparticle 'invisible' is defined by the PDG-ids -16 -14 -12 12 14 16 1000022 1000039.
ma5>remove TheMuon
ma5>display TheMuon
 ** ERROR: no object called 'TheMuon' found.
```

Basic concepts of the normal mode (2)

Histograms - the command *plot*

- *plot*: creation of an histogram
- Global observables
  related to the full event (MET, H_T, *etc.*)
- Properties of a particle species (p_T, E, *etc.*)
- Particle ordering can be used
- Particles can be combined
- Virtual particle properties can be studied
- Log scales can be employed
- Different ways to normalize an histogram

```
ma5>plot MET [
]            ETAordering   initialstate  logY          PTordering   PZordering
allstate     ETordering    interstate    normalize2one PXordering   stack
Eordering    finalstate    logX          Pordering     PYordering   superimpose
ma5>plot MET [ logY ]
ma5>plot N(mu)
ma5>plot PT(mu[1])
ma5>plot ETA(t) [ interstate ]
ma5>plot M(t t~)
ma5>plot dPHI(mu[1] mu[2]) [ logX logY ]
```

- Remark: the list available observables depend on the level of simulation (parton, reconstructed, *etc.*)

Selection cuts - the commands *reject/select*

- Events can be selected/rejected according to a cut condition
- Particles can be selected/rejected from an analysis according to a cut condition

```
ma5>reject MHT < 200
ma5>select N(j) > 3
ma5>reject (j) PT < 20
ma5>reject (j) DELTAR(mu) < 0.4
```

Basic concepts of the normal mode (3)

Executing the analysis - the command *submit*

- First: create a C++ code with the analysis
- Second: compile and execute the code
- Create all the histograms
- Apply all the cuts
- Generate the reports

```
ma5>submit
   Creating folder 'ANALYSIS_4'...
   Copying 'SampleAnalyzer' source files...
   Inserting your selection into 'SampleAnalyzer'...
   Writing the list of datasets...
   Writing the command line history...
   Creating Makefiles...
   Compiling 'SampleAnalyzer'...
   Linking 'SampleAnalyzer'...
   Running 'SampleAnalyzer' over dataset 'defaultset'...
    *******************************************************
    * SampleAnalyzer for MadAnalysis 5 - Welcome.
    * Initializing all components
      - version: 1.1.10.35 (2014/05/09) [ python interface version: "1.1.10.35 (2014/05/09") ]
      - general: everything is default.
      - extracting the list of event samples...
      - analyzer 'MadAnalysis5job'
```

Interfaces to other HEP tools

- MadAnalysis 5 has been interfaced to FastJet and Delphes 3 [Cacciari, Salam (PLB '06); de Favareau *et al* (JHEP'14)]
- Starts from events at the hadron level and produces LHE/LHCO files (FastJet) or Root files (Delphes)
- Delphes is modular ➢ MadAnalysis 5 includes some extra modules [Les Houches 2013 proceedings (1405.1617)]
  - ★ extra information on lepton isolation
  - ★ track information
  - ★ exported to the output file and in the analysis code
  - ★ smaller output Root files (Delphes)

```
ma5>install
delfes  delphes fastjet samples zlib
```

Basic concepts of the normal mode (4)

Jet clustering and basic detector effects

- Running of FastJet via the MadAnalysis 5 interpreter (in the *reco* mode)
- B-tagging efficiencies/mistagging rates, tau-tagging efficiencies/mistagging rates, *etc.*, can be included
- The reconstructed events can be redirected to a file
- Can also be used for checking the merging procedure

```
ma5>set main.fastsim.package =
delfes  delphes fastjet none
ma5>set main.fastsim.package = fastjet
ma5>set main.fastsim.algorithm =
antikt      cambridge   cdfjetclu   cdfmidpoint genkt       gridjet     kt          none        siscone
ma5>set main.fastsim.algorithm = antikt
ma5>set main.fastsim.
main.fastsim.algorithm          main.fastsim.bjet_id.matching_dr main.fastsim.exclusive_id       main.fastsim.radius
main.fastsim.bjet_id.efficiency  main.fastsim.bjet_id.misid_cjet  main.fastsim.package            main.fastsim.tau_id.efficiency
main.fastsim.bjet_id.exclusive   main.fastsim.bjet_id.misid_ljet  main.fastsim.ptmin              main.fastsim.tau_id.misid_ljet
ma5>set main.fastsim.bjet_id.efficiency = 0.60
ma5>set main.fastsim.bjet_id.misid_cjet = 0.10
ma5>set main.fastsim.bjet_id.misid_ljet = 0.01
ma5>set main.outputfile = blabla.lhe
```

Fast simulation of the detector with Delphes 3

- Running of Delphes via the MadAnalysis 5 interpreter (in the reco mode)
- Choice of ATLAS or CMS; pile-up can be included
- The Root output file is stored

```
ma5>set main.fastsim.package = delfes
ma5>set main.fastsim.
main.fastsim.detector main.fastsim.output   main.fastsim.package  main.fastsim.pileup
ma5>set main.fastsim.detector = cms
```

Outline

Analyzing parton-level events: basic cuts

✦ Include reasonable selections to get closer to a real detector

- ✤ No soft object:
  - ➢ remove any soft jet, lepton, *etc.*
- ✤ Objects not in the detector acceptance:
  - ➢ remove object with a too large pseudorapidity
- ✤ Not used here: however, objet overlaps could be removed (cut on the ΔR)

```
ma5>define l = l+ l-
ma5>select (l) PT>10
ma5>select (j) PT>20
ma5>select (l) -2.5 < ETA < 2.5
ma5>select (j) -2.5 < ETA < 2.5
```

New multiparticle labels (more compact commands later)

Soft objects (units: GeV)

Outside the detector

✦ Cut on the lepton multiplicity

- ✤ 2 leptons for WW
- ✤ 3 leptons for WZ
- ✤ 4 leptons for ZZ

```
ma5>select N(l)==2
ma5>select N(l)==3
ma5>select N(l)==4
```

Cut: reject N (l) != 2.0

Dataset	Events kept: K	Rejected events: R	Efficiency: K / (K + R)	Cumul. efficiency: K / Initial
cbm05	269967 +/- 636	12286653 +/- 17066	2.15e-02 +/- 4.09e-05	2.15e-02 +/- 4.09e-05
chbm05	266392 +/- 636	12275638 +/- 17506	2.12e-02 +/- 4.07e-05	2.12e-02 +/- 4.07e-05
chbp01	263008 +/- 643	12297077 +/- 18493	2.09e-02 +/- 4.04e-05	2.09e-02 +/- 4.04e-05
chwm03	270141 +/- 638	12400645 +/- 17415	2.13e-02 +/- 4.06e-05	2.13e-02 +/- 4.06e-05
cwp03	268868 +/- 634	12019437 +/- 16708	2.19e-02 +/- 4.17e-05	2.19e-02 +/- 4.17e-05

§

✦ Some observables are related to the full event (called global)

```
ma5>plot MET 50 0 1000 [logY]
```

- ✤ Missing and visible energy (*MET, TET*)
- ✤ Missing and visible hadronic energy (*MHT, THT*)
- ✤ The partonic center-of-mass energy (*SQRTS*)
- ✤ The α_T variable (*ALPHAT*): depends on the missing energy, H_T and the jet configuration
- ✤ The particle content of the event (*NPID, NAPID, N*)

✦ General setup for the histograms

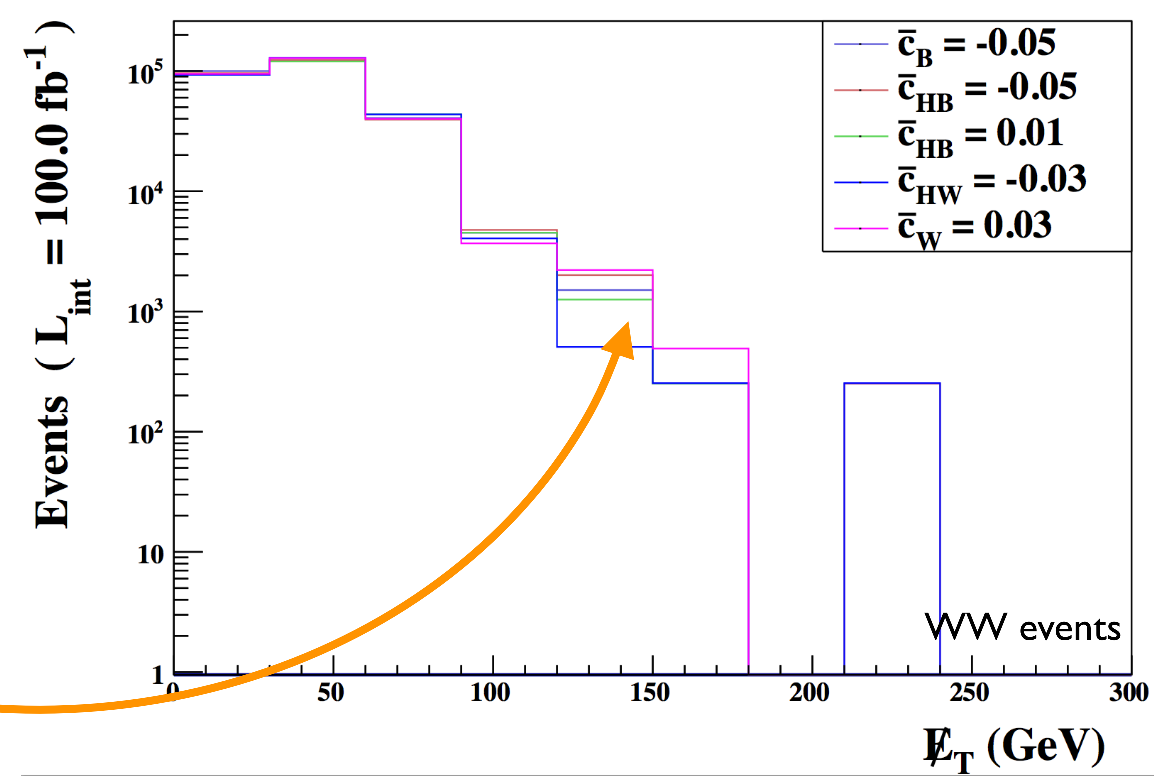
```
ma5>set main.stacking_method = superimpose
ma5>set main.lumi = 100
```

- ✤ The luminosity in fb^{-1}
- ✤ How to superimpose the curves on a single histogram (shape analysis here)

✦ Executing the analysis and browsing the results

- ✤ The command *submit*
  (the progress can be followed on the screen)
- ✤ The command *open*
  (open the HTML report containing all results)

Effects of different dimension-six operators (in the tail)
➢ Measurements could make their distinction possible

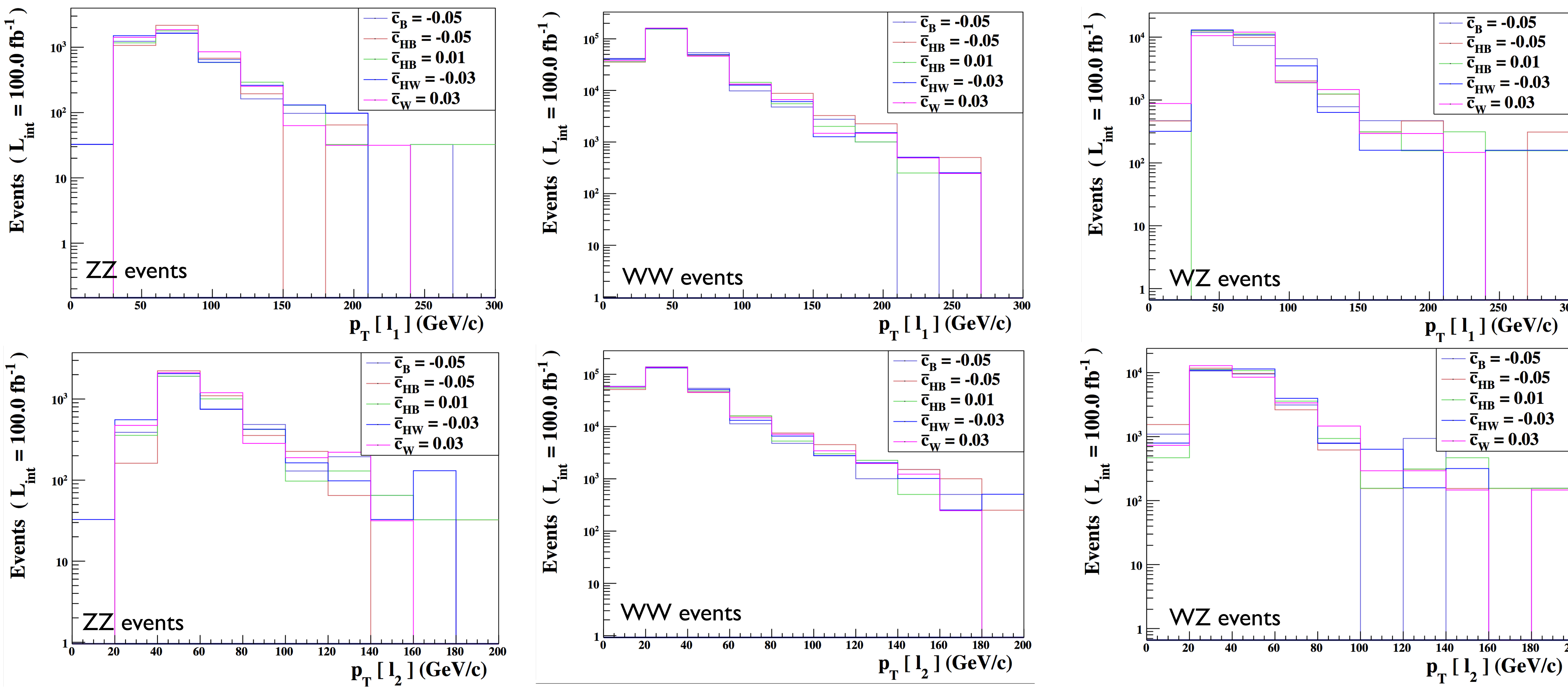


Investigating particle properties (I)

- **Many kinematical properties of a given particle can be studied**
  - *BETA, E, ET, ETA, GAMMA, M, MT, P, PHI, PT, PX, PY, PZ, R, THETA, Y*
  - Each of these functions take a single argument (a particle)
- **The particles are ordered**
  - Use squared brackets to select the right one
  - Several ordering variables are available (*PT, E, PX, etc.*)

```
ma5>plot PT(l[1]) 10 0 300 [logY]
ma5>plot PT(l[2]) 10 0 200 [logY]
```

Tails: where new physics could be hidden



Investigating particle properties (2)

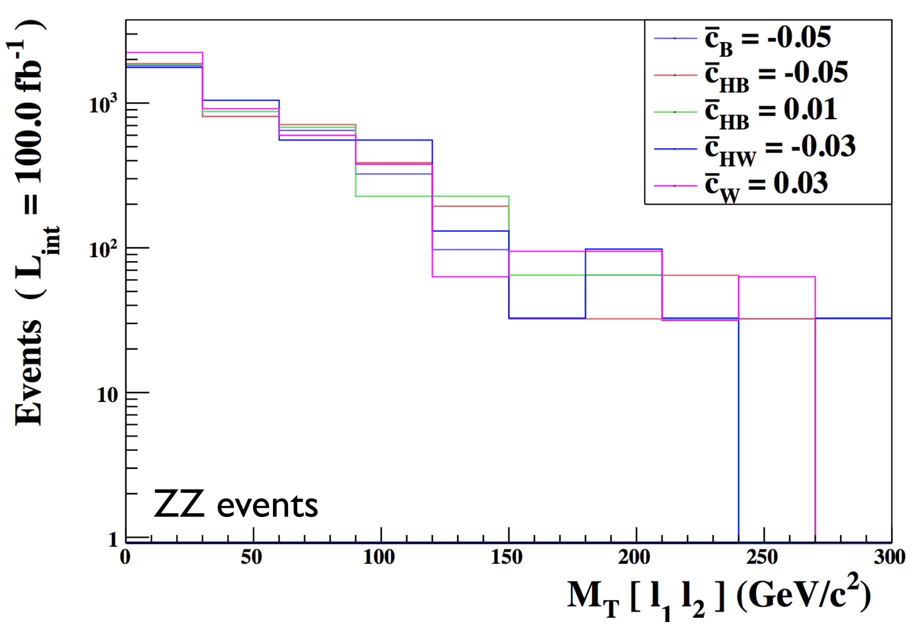
```
ma5>plot MT(l[1] l[2]) 10 0 300 [logY]
ma5>plot DPHI_0_PI(l[1], l[2]) 10 0 3.14 [logY]
ma5>plot DELTAR(l[1], l[2]) 10 0 5 [logY]
ma5>plot MT_MET(l[1]) 10 0 300 [logY]
```

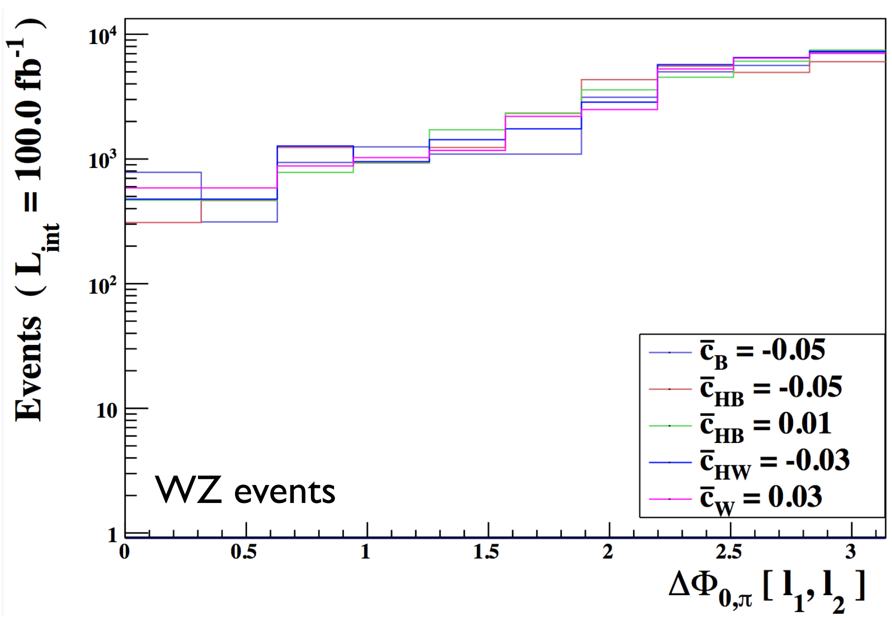
✦ Combining particles

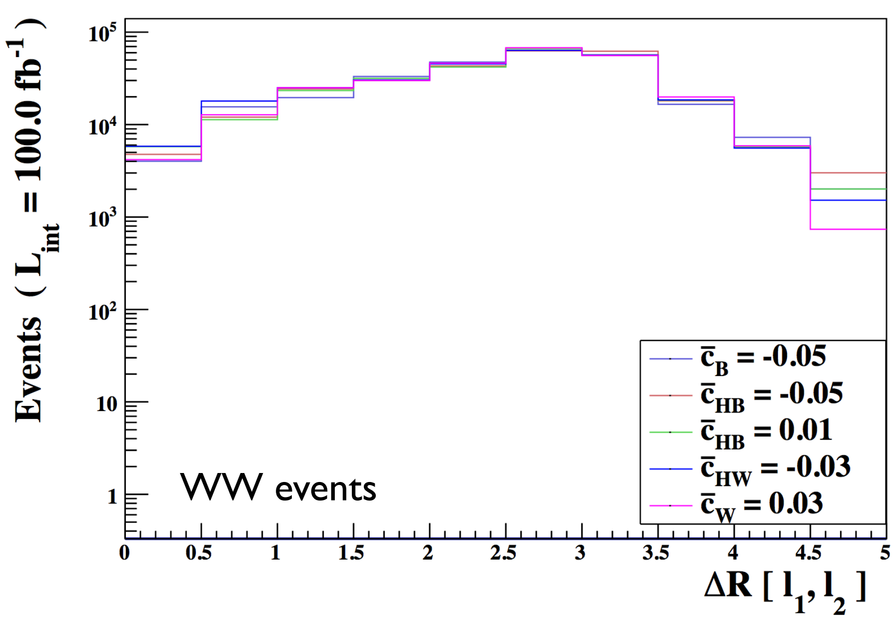
- Several particles passed as arguments
- Sum of the four-momenta; the observable is then computed
- Vectorial and scalar sums/differences as well as ratios are available (*s*, *v*, *ds*, *dv*, *r* prefixes)

✦ Four special functions

- *DELTAR, DPHI_0_PI, DPHI_0_2PI*: take two arguments
- *MT_MET*: transverse mass when combining a particle with the missing momentum

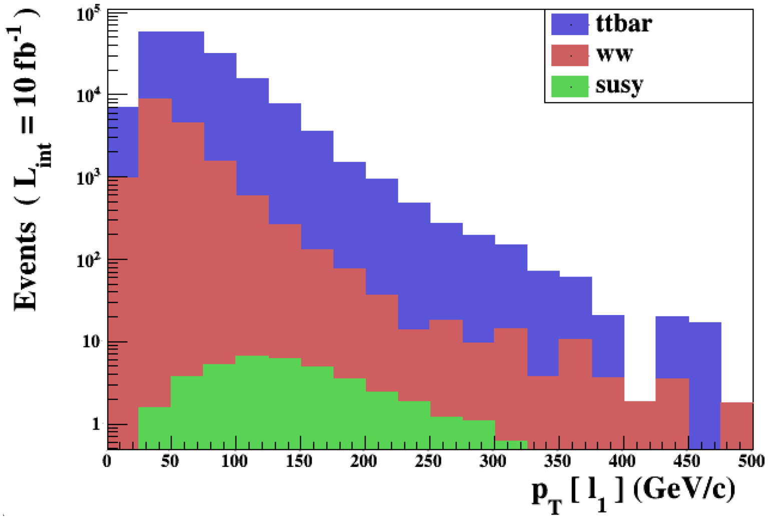






Signal over background ratios

- Analysis of multiple samples: back to the slepton example
  - SUSY signal: smuon pair production
  - Background: ttbar and ww
  - Calculating S/B ratio with MADANALYSIS 5 (by typing in formulas)
    - S = number of signal events
    - B = number of background events
    - ΔS = error on the number of signal events
    - ΔB = error on the number of background events
  - Here: $\frac{S}{\sqrt{S+B}} \Rightarrow \frac{\sqrt{(S+2B)^2(\Delta S)^2+S^2(\Delta B)^2}}{(S+B)^{3/2}}$



[see O. Mattelaer's talk]

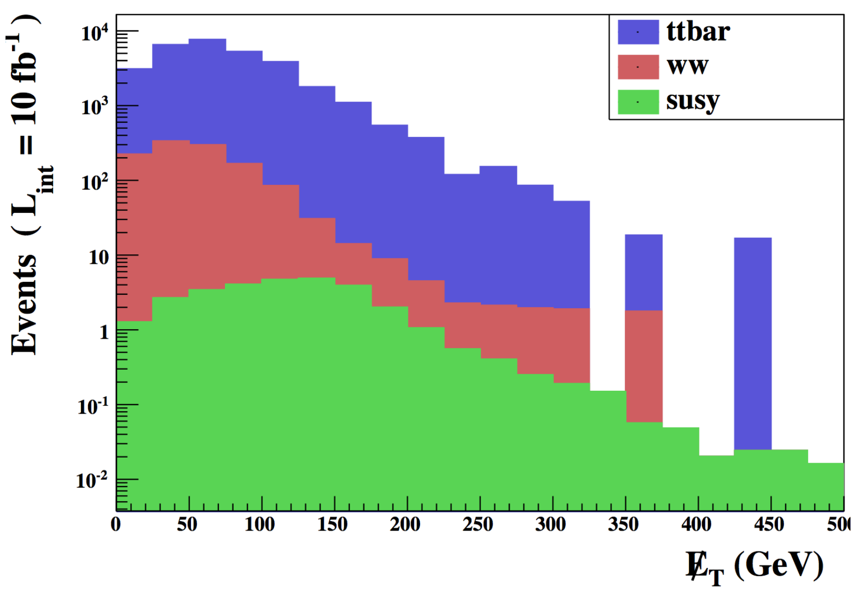
```
ma5>set ww.type = background
ma5>set ttbar.type = background
ma5>define l = l+ l-
ma5>select PT(l[1]) > 100
ma5>set main.SBratio = 'S/sqrt(S+B)'
Checking the formula ...
Formula corresponding to the uncertainty calculation has been found and set to the variable main.SBerror :
  1./pow(S+B,3./2.)*sqrt((S+2*B)**2*ES**2+S**2*EB**2)
ma5>plot MET 20 0 500 [logY]
ma5>submit
```

Cut-flow chart

- How to compare signal (S) and background (B): S/sqrt (S+B) .
- Associated uncertainty: 1./pow(S+B,3./2.)*sqrt((S+2*B)**2*ES**2+S**2*EB**2).

Cuts	Signal (S)	Background (B)	S vs B
Initial (no cut)	40.528 +/- 0.139	184995 +/- 596	0.094217 +/- 0.000714
Cut 1	29.9 +/- 2.8	30806 +/- 190	0.1701 +/- 0.0319

Signal and Background comparison



Outline

Motivation for the expert mode of MADANALYSIS 5

[Conte, Dumont, BenjFuks, Wymant (1405.3982)]

✦ MADANALYSIS 5 is used without its PYTHON interface

- ✤ More freedom in the observables (only some of them can be called from the PYTHON console)
- ✤ Complicated cuts can be implemented
- ✤ More suitable for large numbers of events (using several cores)

✦ The expert mode is developer-friendly

The analysis is a C++ class

The SAMPLEANALYZER internal data format

- ★ Readers for LHE, STDHEP, HEPMC, LHCO and DELPHES
- ★ Many classes and methods for particle and object properties
- ★ Specific methods for histograms and cuts
- ★ *etc.* (see 1405.3982)

Services

- ★ Physics observables (transverse variables, object identification, isolation)
- ★ Streamers
- ★ Exceptions
- ★ *etc.*

Interfaces

- ★ FASTJET
- ★ DELPHES 3
- ★ New DELPHES modules [LH 2013 proceedings]

The new DELPHES modules (linked to SAMPLEANALYZER):

- ★ isolation
- ★ tracks
- ★ output files

Scripts

- ★ Compilation
- ★ Linking
- ★ Analysis skeleton generator

The new extension of the expert mode

[Conte, Dumont, BenjFuks, Wymant (1405.3982)]

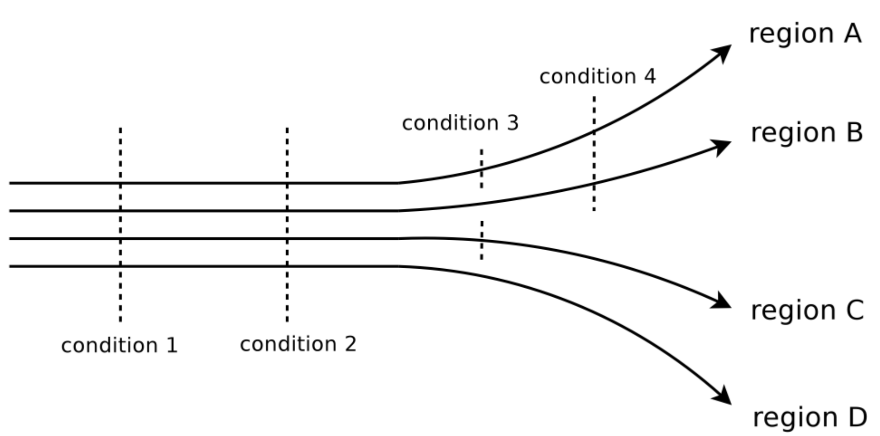
✦ Main features (enable a recast of most of the cut-based LHC analyses on the market)

- Support for multiple sub-analyses (signal and control regions)
- New ready-to-use observables (M_{T2}, M_{T2W}, *etc.*)
- New optimized handling of cuts and histograms

✦ Handling cuts and histograms

- Naive approach not efficient (see cut #4 for instance)

```
count the event in region D
if (condition 3)
{
  count the event in region C
  if (condition 4)
  {
    count the event in region A
  }
}
if (condition 4)
{
  count the event in region B
}
```



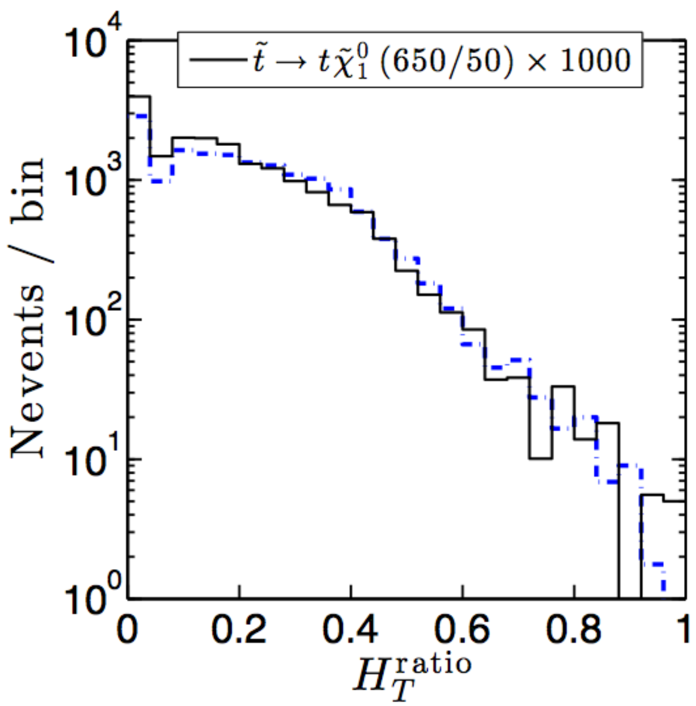
- A more efficient algorithm has been implemented
  - ★ Each cut condition is only evaluated once
  - ★ It is applied to all surviving regions simultaneously
- Similar treatment for histograms

Example (proof of concept): CMS-SUS-13-011

- CMS search for stops in the single lepton channel
  - SUSY benchmark: stop of 650 GeV and neutralino of 50 GeV
  - Normalization: NLO+NLL total cross sections (14 fb here)
  - Simulation chain:
    SLHA ➢ FeynRules ➢ UFO ➢ MadGraph 5 ➢ Pythia 6 ➢ modified Delphes ➢ MadAnalysis 5

- Cross check with publicly available material from CMS

Cut	MadAnalysis 5	CMS
At least one lepton, four jets and 100 GeV of missing transverse energy	31.4	29.7
At least one b-tagged jet	27.1	25.2
No extra loosely-isolated lepton or track	22.5	21.0
No hadronic tau	22.0	20.6
Angular separation between the missing momentum and the two hardest jets	18.9	17.8
Hadronic top quark reconstruction	12.7	11.9
The transverse mass M_T (defined in the text) is larger than 120 GeV	10.4	9.6
At least 300 GeV of missing transverse energy and $M_{T2}^W > 200$ GeV	5.1	4.2



CMS results (for this analysis) can be reproduced with a pretty good accuracy: at the 20%-30% level

Outline

Summary

✦ MADANALYSIS 5 in a nutshell

- ✤ A unique framework for collider phenomenology (parton, hadron, detector and reconstructed levels)
- ✤ User-friendly by means of its PYTHON interface (normal mode)
- ✤ Flexible thanks to its C++ kernel (expert mode)
- ✤ Interfaced to several other HEP packages

✦ Presented examples for the normal mode of MADANALYSIS 5: Higgs EFT

- ✤ Two equivalent (FEYNRULES) implementations exist
- ✤ Can be used in many different purposes (diboson production investigated in this presentation)
  - ★ Total rates
  - ★ Differential distributions

✦ MADANALYSIS 5 and LHC analyses

- ✤ The expert mode has been extended to facilitate the implementation of LHC analyses
- ✤ Proof of concept: CMS-SUS-13-011 (good agreement has been found)
- ✤ More analyses are coming…

✦ Website: http://launchpad.net/madanalysis5