30.09.2020

Simulation, reconstruction and characterization of HH→4bjets

Hands On

Muon Collider framework tutorial

P. Andreetto [a], N. Bartosik [b], L. Buonincontri [a,d]
M. Casarsa [c], A. Gianelle [a], S. Jindariani [e], D. Lucchesi [a,d], S. Pagan Griso [f], L. Sestini [a]

[a] INFN Padova, [b] INFN Torino, [c] INFN Trieste, [d] University of Padova, [e] FNAL, [f] LBNL

Muon Collider software setup

Hands-on instructions can be found in:

https://confluence.infn.it/display/muoncollider/Muon+Collider+Software+tutorial

1) Log into the login.snowmass.io machine

```
ssh -XY username@login.snowmass21.io
```

2) Access the singularity container with the Muon Collider software:

```
singularity run --bind `echo $HOME` --bind
/collab/project/snowmass21/data/muonc:/data
/cvmfs/unpacked.cern.ch/registry.hub.docker.com/infnpd/mu
coll-ilc-framework:1.0-centos8
```

Muon Collider software setup

3) Check that the Muon Collider software is properly set up:

```
ddsim -h

Marlin -h
```

If error message appears, run the ILCSoft setup script:

source /opt/ilcsoft/v02-01-pre/init_ilcsoft.sh

and check again

4) Check out the configuration files and the ROOT macros:

```
git clone https://github.com/MuonColliderSoft/MuC-Tutorial.git
```

If everything worked fine, you can go to the directory for the simulation and reconstruction exercises:

```
cd ~/MuC-Tutorial/tutorial/1-mumuHHbbbb/
```

$H \rightarrow b\bar{b}$ simulation and reconstruction chain

- Generation step of the process $\mu^+\mu^- \rightarrow H \nu\bar{\nu} \rightarrow b\bar{b} \nu\bar{\nu}$ at 1.5 TeV center of mass energy ⇨ done (with PYTHIA 8):

  /data/samples/HH/mumu2H2bb750.stdhep

- Run the simulation with:

```
ddsim --steeringFile sim_steer_Hbb.py > sim.out 2>&1
```

- Output file: `mumu_H_bb.slcio`

It takes some minutes, so we can already start!

SIM steering file

Closer look at `sim_steer_Hbb.py`

```
import os

from DDSim.DD4hepSimulation import DD4hepSimulation
#from SystemOfUnits import mm, GeV, MeV, m, deg
from g4units import mm, GeV, MeV, m, deg
SIM = DD4hepSimulation()

## The compact XML file
SIM.compactFile = "/opt/ilcsoft/v02-01-pre/detector-simulation/geometries/CLIC_o3_v14_mod4/CLIC_o3_v14.xml"
## Lorentz boost for the crossing angle, in radian!
SIM.crossingAngleBoost = 0.010
SIM.enableDetailedShowerMode = True
SIM.enableG4GPS = False
SIM.enableG4Gun = False
SIM.enableGun = False
## InputFiles for simulation .stdhep, .slcio, .HEPEvt, .hepevt, .hepmc files are supported
SIM.inputFiles = ["/data/samples/HH/mumu2H2bb750.stdhep"]
## Macro file to execute for runType 'run' or 'vis'
SIM.macroFile = ""
## number of events to simulate, used in batch mode. -1 all
SIM.numberOfEvents = 10
## Outputfile from the simulation,only lcio output is supported
SIM.outputFile = "mumu_H_bb.slcio"
## Verbosity use integers from 1(most) to 7(least) verbose
## or strings: VERBOSE, DEBUG, INFO, WARNING, ERROR, FATAL, ALWAYS
SIM.printLevel = "VERBOSE"
```

Detector geometry

Input File

Number of events

Output file

SIM steering file (II)

```
################################################################################
## Configuration for the PhysicsList
################################################################################
SIM.physics.decays = False
SIM.physics.list = "QGSP_BERT_HP"

##  location of particle.tbl file containing extra particles and their lifetime information
##
SIM.physics.pdgfile = os.path.join( os.environ.get("DD4HEP"), "DDG4/examples/particle.tbl" )

##  The global geant4 rangecut for secondary production
##
##     Default is 0.7 mm as is the case in geant4 10
##
##     To disable this plugin and be absolutely sure to use the Geant4 default range cut use "None"
##
##     Set printlevel to DEBUG to see a printout of all range cuts,
##     but this only works if range cut is not "None"
##
SIM.physics.rangecut = 0.7*mm

SIM.physics.rejectPDGs = {1,2,3,4,5,6,21,23,24,25}

################################################################################
## Properties for the random number generator
################################################################################

## If True, calculate random seed for each event based on eventID and runID
## allows reproducibility even when SkippingEvents
SIM.random.enableEventSeed = True
SIM.random.file = None
SIM.random.luxury = 1
SIM.random.replace_gRandom = True
SIM.random.seed = None
SIM.random.type = None
```

GEANT4 physics list

H→b$\bar{b}$ simulation and reconstruction chain

- Generation step ⇨ done (with PYTHIA 8):

  /data/samples/HH/mumu2H2bb750.stdhep

- Run the simulation with:

```
ddsim --steeringFile sim_steer_Hbb.py > sim.out 2>&1
```

- Output file: `mumu_H_bb.slcio`

- Run the digitization/reconstruction step with:

```
Marlin reco_steer_Hbb.xml > reco.out 2>&1
```

- Output file: `Output_REC.slcio, histograms.root`

Reconstruction steering file

Three main sections in the Marlin .xml steering file

1) execute section (ordered list of processors to be executed)

```
<execute>
<processor name="MyAIDAProcessor"/>
<processor name="MyTestProcessor"/>
<processor name="MyLCIOOutputProcessor"/>
</execute>
```

2) global section (global settings)

```
<global>
<parameter name="LCIOInputFiles"> input.slcio </parameter>
<parameter name="MaxRecordNumber" value="1000" />
</global>
```

3) processor section (processor configuration)

```
<processor name="MyLCIOOutputProcessor" type="LCIOOutputProcessor">
<parameter name="LCIOOutputFile" type="string"> Output_DST.slcio </parameter>
<parameter name="DropCollectionTypes" type="StringVec">
SimCalorimeterHit
SimTrackerHit
</parameter>
<parameter name="LCIOWriteMode" type="string" value="WRITE_NEW"/>
<parameter name="SplitFileSizekB" type="int">1048576 </parameter>
<parameter name="Verbosity" type="string">WARNING </parameter>
</processor>
```

Execute section

List of active processors:

- Geometry (InitializeDD4hep)
- Digitization of vertex and tracker hits (DDPlanarDigiProcessor)
- Tracking processor (ConformalTracking)
- Calorimeter digitization (DDCaloDigi)
- Pandora PFA for particle reconstruction (DDPandoraPFANewProcessor)
- Output processors (LCIOOutProcessor)

Digitization and reconstruction algorithms

Digitization:

- Tracker: gaussian smearing of SIM hits' positions and times, time window
- Calorimeters: simple digitization with an energy calibration constant, an energy threshold and a selection time window;
- Muon detectors: simple digitization with an energy calibration constant, an energy threshold and energy saturation.

Reconstruction:

- Tracks: pattern recognition based on conformal mapping and cellular automaton (E. Brondolin et al., arXiv:1908.00256) + Kalman filter fit;
- Calorimeter clusters: Pandora Particle Flow Algorithm (J.S. Marshall and M. A. Thomson, arXiv:1308.4537) to recognize different patterns of hits released by different particle types in the high granularity calorimeters

Jet reconstruction and ntuplizer

- The jet reconstruction run in a separate steering file and ROOT trees are produced:

```
Marlin lctuple_steer.xml > ntuples.out 2>&1
```

- Output: `JetHistograms.root`
- LCTuple is a configurable flat ROOT tree, produced from the collections in the slcio files

  Link to the repository: https://github.com/iLCSoft/LCTuple

- JetHistograms.root

1-mumuHHbbbb
- PandoraSettings
- JetHistograms.root
  - MyLCTuple;1
    - evevt
    - evrun
    - evwgt
    - evtim
    - evsig
    - evene
    - evpoe
    - evpop
    - evnch
    - evpro

MC Particles

nmcp, mcori, mcpdg, mcgst, mcsst, mcvtx, mcvty, mcvtz, mcepx, mcepy, mcepz, mcmox, mcmoy, mcmoz, mcmas, mcene, mccha, mctim, mcspx, mcspy, mcspz, mccf0, mccf1, mcpa0, mcpa1, mcda0, mcda1, mcda2, mcda3, mcda4

Reconstructed Particles

nrec, rcori, rccid, rctyp, rccov, rcrpx, rcrpy, rcrpz, rcgpi, rcpiu, rcnpi, rcfpi, rcmox, rcmoy, rcmoz, rcmas, rcene, rccha, rcntr, rcncl, rcnrp, rcftr, rcvts, rcvte, rccom

Jets

nj, jmox, jmoy, jmoz, jmas, jene, jcha, jcov0, jcov1, jcov2, jcov3, jcov4, jcov5, jcov6, jcov7, jcov8, jcov9, jevis, jPxvis, jPyvis, jPzvis, jmom, jcost, jcosTheta, jTheta, jPtvis, jmvis, jmmax, jEmiss, jMmissq, jMmiss

Tracks

ntrk, trori, trtyp, trch2, trndf, tredx, trede, trrih, trshn, trnts, trfts, trsip, trsfh, trslh, trsca, ntrst, tsloc, tsdze, tsphi, tsome, tszze, tstnl, tscov, tsrpx, tsrpy, tsrpz

links to the MC truth

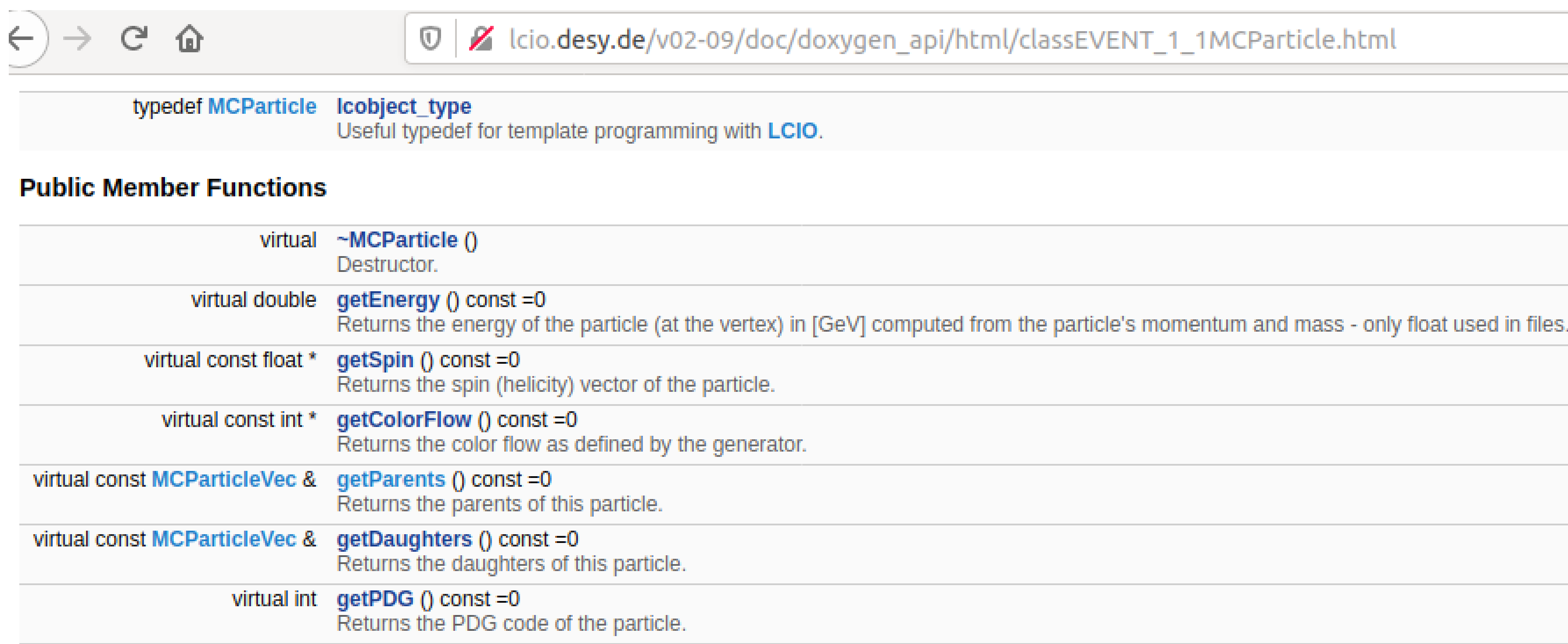
r2mnrel, r2mf, r2mt, r2mw, r2cnrel, r2cf, r2ct, r2cw, r2tnrel, r2tf, r2tt, r2tw, r2rnrel, r2rf, r2rt, r2rw

Leaf names JetHistograms.root

- To understand the full list of leaf names: see the classes https://github.com/iLCSoft/LCTuple/tree/master/src and the member function definition in http://lcio.desy.de/v02-09/doc/doxygen_api/html/annotated.html website.

```
_mcpdg[ i ] = mcp->getPDG() ;
_mcgst[ i ] = mcp->getGeneratorStatus() ;
_mcsst[ i ] = mcp->getSimulatorStatus() ;
_mcvtx[ i ] = mcp->getVertex()[0] ;
_mcvty[ i ] = mcp->getVertex()[1] ;
_mcvtz[ i ] = mcp->getVertex()[2] ;
_mcepx[ i ] = mcp->getEndpoint()[0] ;
_mcepy[ i ] = mcp->getEndpoint()[1] ;
_mcepz[ i ] = mcp->getEndpoint()[2] ;
_mcmox[ i ] = mcp->getMomentum()[0] ;
_mcmoy[ i ] = mcp->getMomentum()[1] ;
_mcmoz[ i ] = mcp->getMomentum()[2] ;
_mcmas[ i ] = mcp->getMass() ;
_mcene[ i ] = mcp->getEnergy() ;
_mccha[ i ] = mcp->getCharge() ;
_mctim[ i ] = mcp->getTime() ;
_mcspx[ i ] = mcp->getSpin()[0] ;
_mcspy[ i ] = mcp->getSpin()[1] ;
_mcspz[ i ] = mcp->getSpin()[2] ;
_mccf0[ i ] = mcp->getColorFlow()[0] ;
_mccf1[ i ] = mcp->getColorFlow()[1] ;
```

https://github.com/iLCSoft/LCTuple/blob/master/src/MCParticleBranches.cc


lcio.desy.de/v02-09/doc/doxygen_api/html/classEVENT_1_1MCParticle.html

typedef MCParticle	lcobject_type Useful typedef for template programming with LCIO.

Public Member Functions

virtual	~MCParticle () Destructor.
virtual double	getEnergy () const =0 Returns the energy of the particle (at the vertex) in [GeV] computed from the particle's momentum and mass - only float used in files.
virtual const float *	getSpin () const =0 Returns the spin (helicity) vector of the particle.
virtual const int *	getColorFlow () const =0 Returns the color flow as defined by the generator.
virtual const MCParticleVec &	getParents () const =0 Returns the parents of this particle.
virtual const MCParticleVec &	getDaughters () const =0 Returns the daughters of this particle.
virtual int	getPDG () const =0 Returns the PDG code of the particle.

Exercise with HH events

- For exercise, 1000 events of double Higgs production are already simulated and reconstructed, and you can run a simple macro invariant_mass.C for the events analysis:

```
root -l

.L invariant_mass.C

invariant_mass()
```

- The argument is used to chose the plot you want to see:

  0: number of reconstructed jets

  1: pseudorapidity of jets in the event

  2: transverse momentum of jets in the event

  3: phi of jets in the event

  4: energy of jets in the event

  5: invariant mass of jet pair associated to Higgs with highest p_T

  6: invariant mass of jet pair associated to Higgs with lowest p_T

Macro description

- For exercise, 1000 events of double Higgs process ($\mu^+\mu^- \rightarrow H\,H\,\nu\bar{\nu} \rightarrow b\bar{b}b\bar{b}\,\nu\bar{\nu}$) are already simulated and reconstructed, and you can run a simple macro invariant_mass.C for the events analysis:

- Loop on the 1000 events

- Basic object plots of all jets in the event: N_{jets}, p_T, η, E, φ

- Selection of events HH → 4b with $N_{jets} > 3$

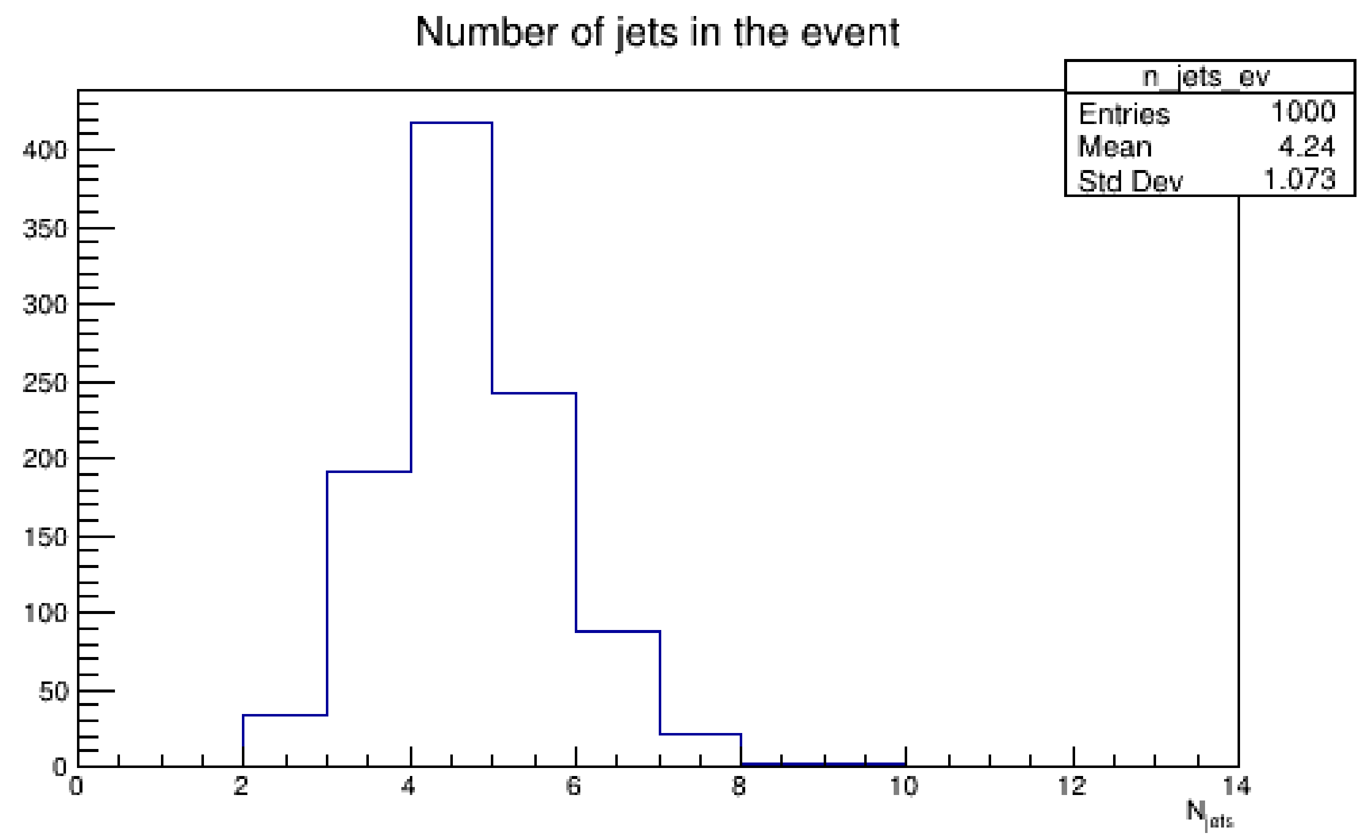
- Jets are combined in pairs and for each combination the invariant mass is calculated (m_{ij},m_{kl})

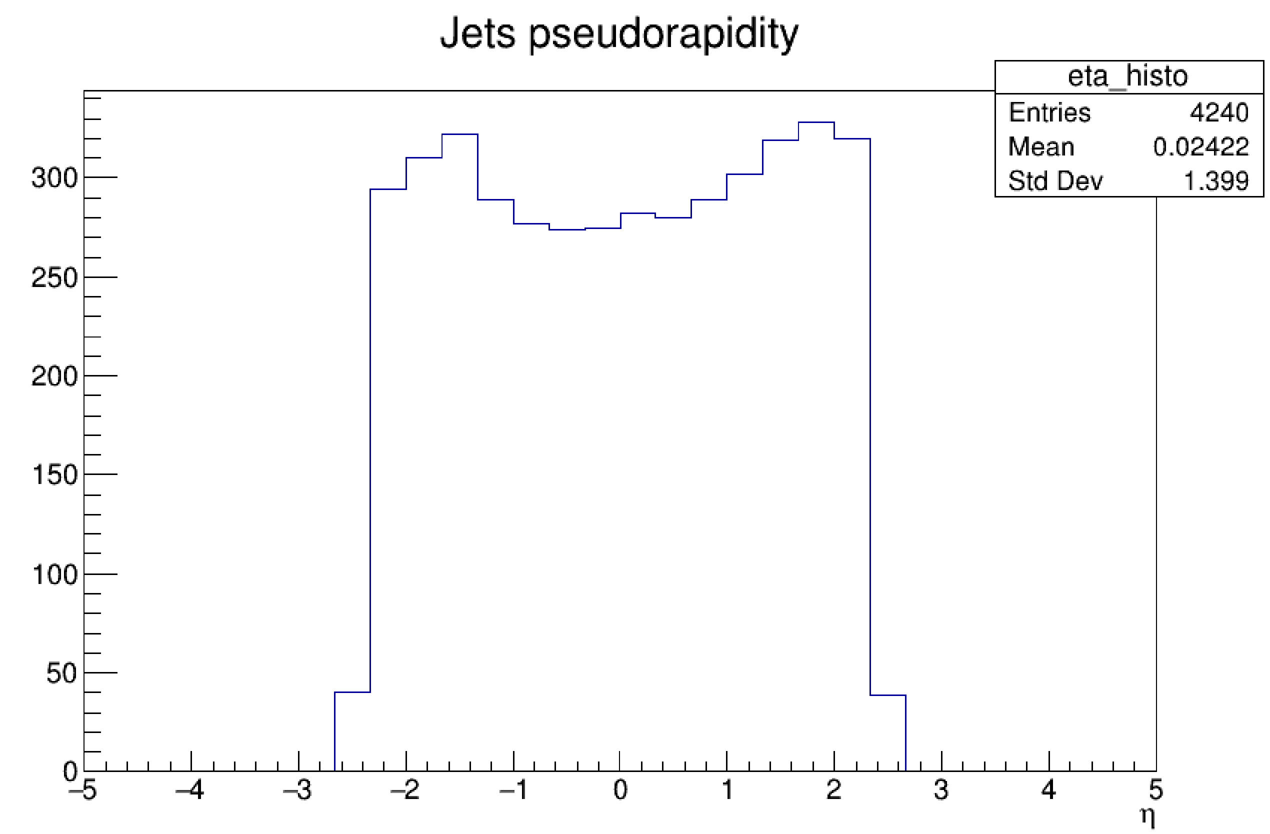
- The following relation is calculated for each pair:

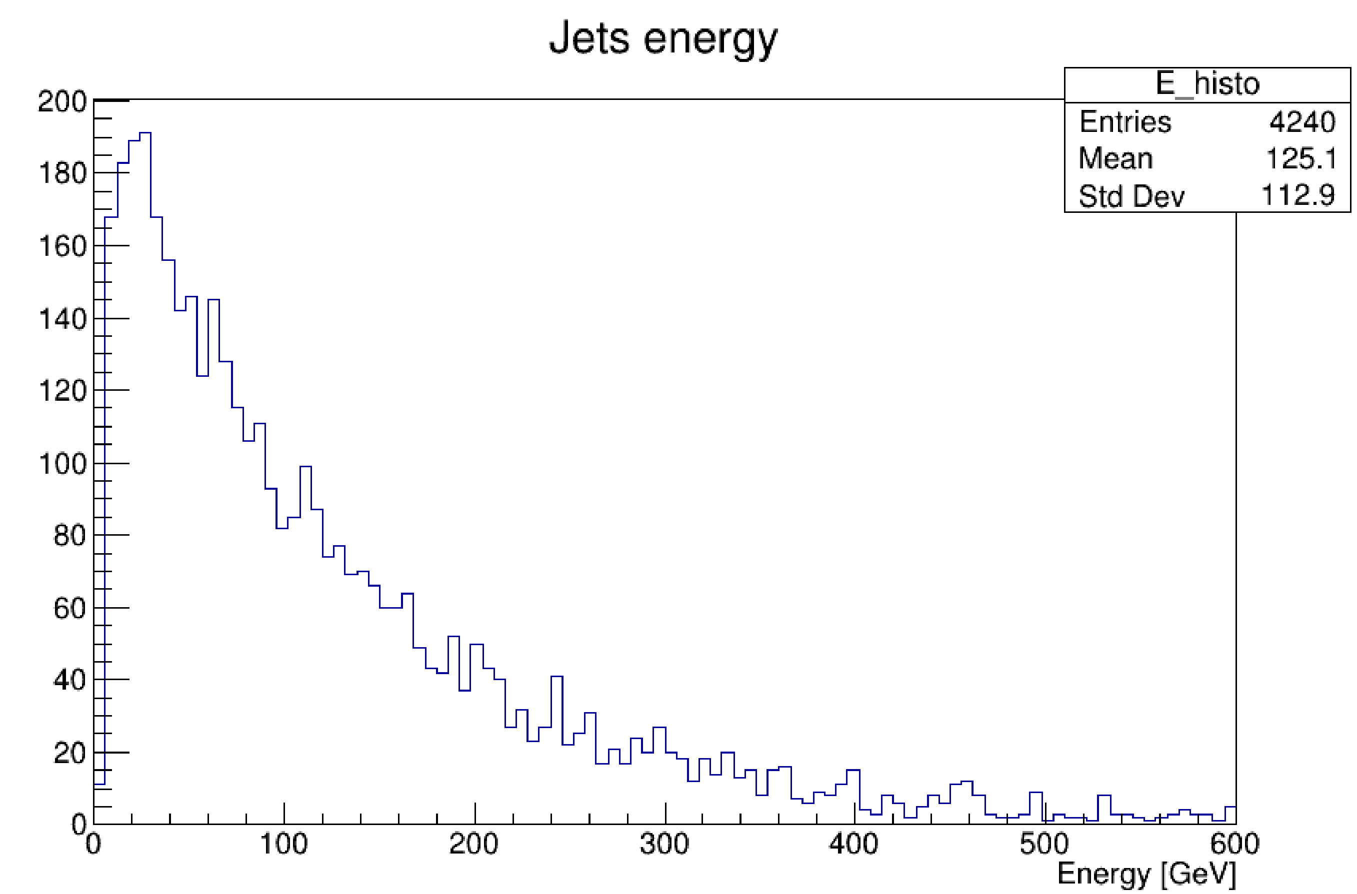
  $(m_h - m_{ij})^2 + (m_h - m_{kl})^2$

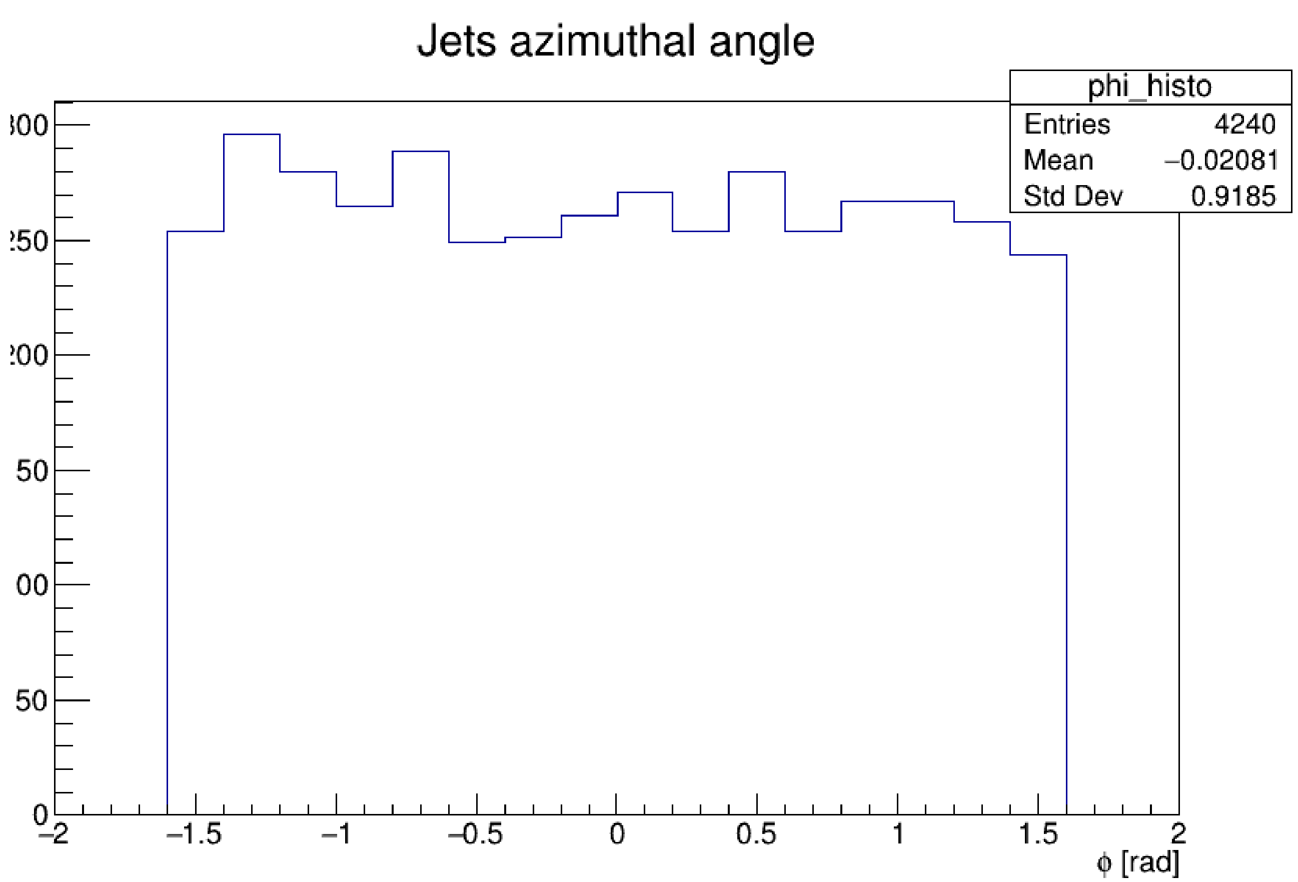
- The pair of jets which minimize the above relation is selected (m_{H1} and m_{H2} according to the Higgs p_T)

Reconstruction steering file

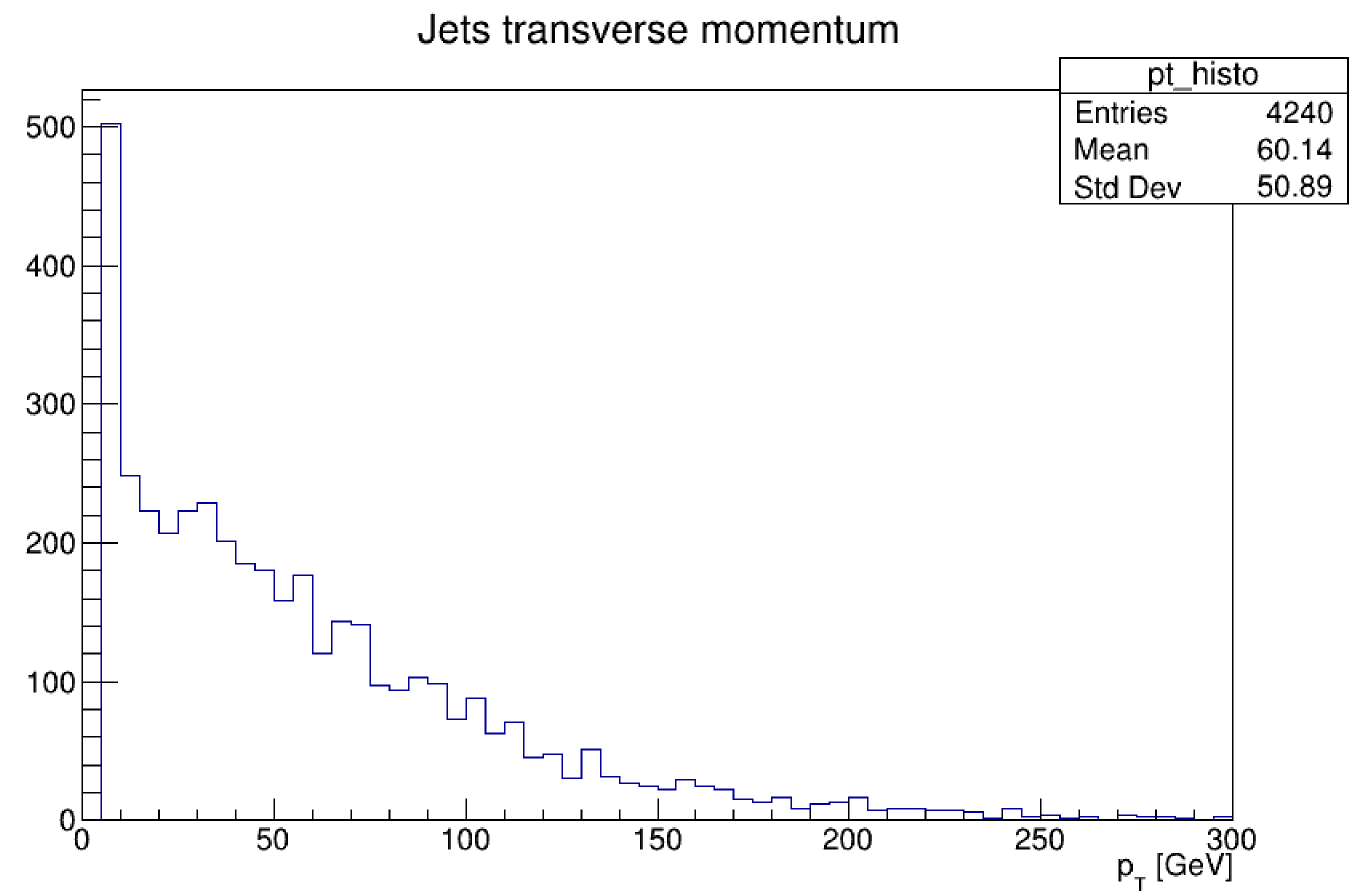


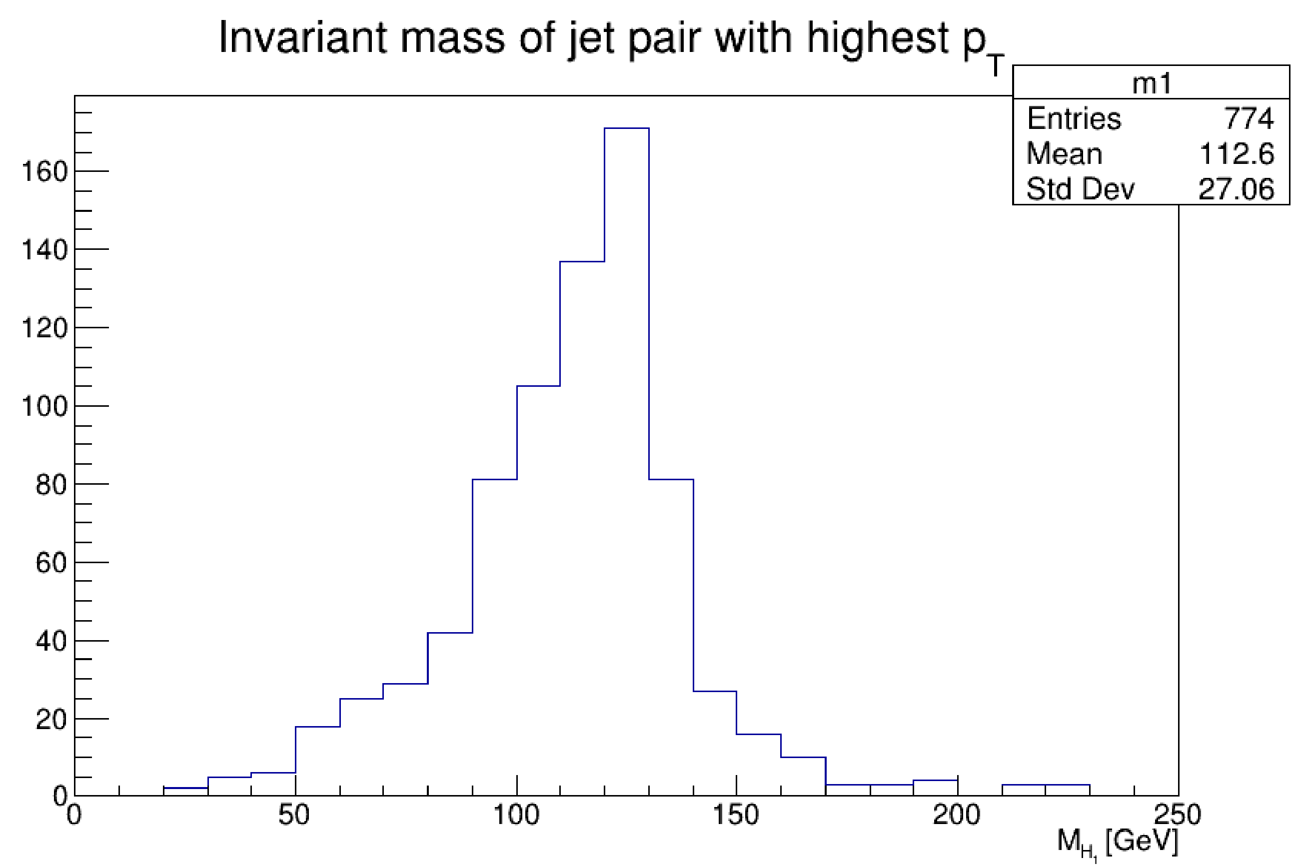


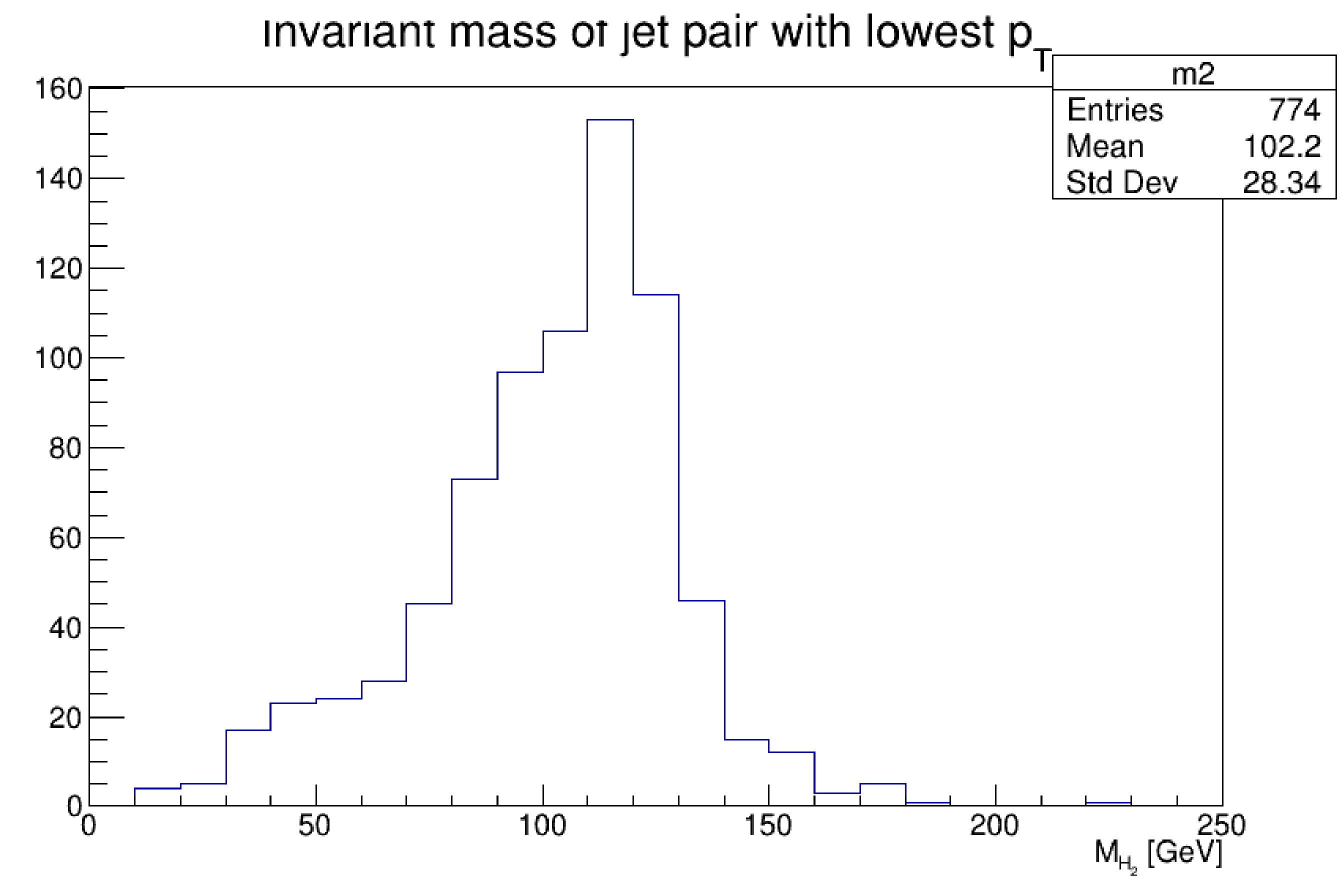




Reconstruction steering file







BACKUP

Reconstruction output files

`Output_REC.slcio` contains all the output collections

Useful commands:

- Number of events saved in an output file:

```
lcio_event_counter Output_REC.slcio
```

- List of collections saved in the slcio files:

```
anajob Output_REC.slcio
```

- Dump of collections' content:

```
dumpevent Output_REC.slcio
```

(more infos executing `dumpevent -h`).

Anajob Output_REC.slcio

COLLECTION NAME	COLLECTION TYPE	NUMBER OF ELEMENTS
BuildUpVertices	Vertex	2
BuildUpVertices_RP	ReconstructedParticle	2
BuildUpVertices_V0	Vertex	0
BuildUpVertices_V0_RP	ReconstructedParticle	0
CalohitMCTruthLink	LCRelation	3329
ClusterMCTruthLink	LCRelation	68
DebugHits	TrackerHitPlane	0
ECALBarrel	CalorimeterHit	1933
ECALEndcap	CalorimeterHit	915
ECalBarrelCollection	SimCalorimeterHit	2078
ECalEndcapCollection	SimCalorimeterHit	985
EfficientMCParticles	MCParticle	22
HCALBarrel	CalorimeterHit	149
HCALEndcap	CalorimeterHit	318
HCALOther	CalorimeterHit	0
HCalBarrelCollection	SimCalorimeterHit	309
HCalEndcapCollection	SimCalorimeterHit	659
HCalRingCollection	SimCalorimeterHit	0
ITrackerEndcapHits	TrackerHitPlane	32
ITrackerHits	TrackerHitPlane	62
InefficientMCParticles	MCParticle	0
InnerTrackerBarrelCollection	SimTrackerHit	67
InnerTrackerBarrelHitsRelations	LCRelation	62
InnerTrackerEndcapCollection	SimTrackerHit	34
InnerTrackerEndcapHitsRelations	LCRelation	32
LE_LooseSelectedPandoraPFOs	ReconstructedParticle	41
LE_SelectedPandoraPFOs	ReconstructedParticle	37
LE_TightSelectedPandoraPFOs	ReconstructedParticle	35
LooseSelectedPandoraPFOs	ReconstructedParticle	39
MCParticle	MCParticle	758
MCParticlesSkimmed	MCParticle	139
MCPhysicsParticles	MCParticle	758
MUON	CalorimeterHit	0
MergedClusters	Cluster	36
MergedRecoParticles	ReconstructedParticle	43
OTrackerEndcapHits	TrackerHitPlane	23
OTrackerHits	TrackerHitPlane	42
OuterTrackerBarrelCollection	SimTrackerHit	44
OuterTrackerBarrelHitsRelations	LCRelation	42
OuterTrackerEndcapCollection	SimTrackerHit	24
OuterTrackerEndcapHitsRelations	LCRelation	23
PFOsFromJets	ReconstructedParticle	43
PandoraClusters	Cluster	36
PandoraPFOs	ReconstructedParticle	43

Execute section

```
<execute>

  <!-- ==========  Setup  ========== -->
  <processor name="MyAIDAProcessor"/>
  <processor name="EventNumber" />
  <processor name="Config" />

  <!-- ==========  Geometry initialization  ========== -->
  <processor name="InitDD4hep_mod4"/>

  <!-- ==========  Overlay  ========== -->
  <if condition="Config.OverlayBIB">
    <processor name="OverlayBIB"/>
  </if>

  <if condition="Config.OverlayFalse">
    <processor name="OverlayFalse"/>
  </if>

  <!-- ==========  Tracker Digitization  ========== -->
  <processor name="VXDBarrelDigitiser"/>
  <processor name="VXDEndcapDigitiser"/>
  <processor name="InnerPlanarDigiProcessor"/>
  <processor name="InnerEndcapPlanarDigiProcessor"/>
  <processor name="OuterPlanarDigiProcessor"/>
  <processor name="OuterEndcapPlanarDigiProcessor"/>
```

List of active processors:

- I/O root files (LCIOOutProcessor)
- Geometry (InitDD4hep_mod4)
- Overlay (OverlayBIB and OverlayFalse)
- Digitization of vertex and tracker hits (es. VXDBarrelDigitizer)

You are running also:

- Tracking processor
- Processors for checks on tracking performances (see later sections)
- Pandora PFA for particle reconstruction
- Jets will be reconstruction will be shown later
- Output processors for .slcio files

2) global section (global settings)

```
<global>
  <parameter name="LCIOInputFiles">
  mumu_H_bb.slcio
  </parameter>
  <!-- Limit the number of processed records (run+evt): -->
  <parameter name="MaxRecordNumber" value="10" />
  <parameter name="SkipNEvents" value="0" />
  <parameter name="SupressCheck" value="false" />
  <parameter name="Verbosity" options="DEBUG0-4,MESSAGE0-4,WARNING0-4,ERROR0-4,SILENT">WARNING  </parameter>
  <parameter name="RandomSeed" value="1234567890" />
</global>
```

Input file

Number of events

Processor section

3) processor section (processor configuration)

```
<processor name="InitDD4hep_mod4" type="InitializeDD4hep">
  <!--InitializeDD4hep reads a compact xml file and initializes the DD4hep::LCDD object-->
  <!--Name of the DD4hep compact xml file to load-->
  <parameter name="EncodingStringParameter"> GlobalTrackerReadoutID </parameter>
  <!-- ALE - Use the modified geometry -->
  <parameter name="DD4hepXMLFile" type="string">
   /opt/ilcsoft/v02-01-pre/detector-simulation/geometries/CLIC_o3_v14_mod4/CLIC_o3_v14.xml
  </parameter>
</processor>
```

Detector description

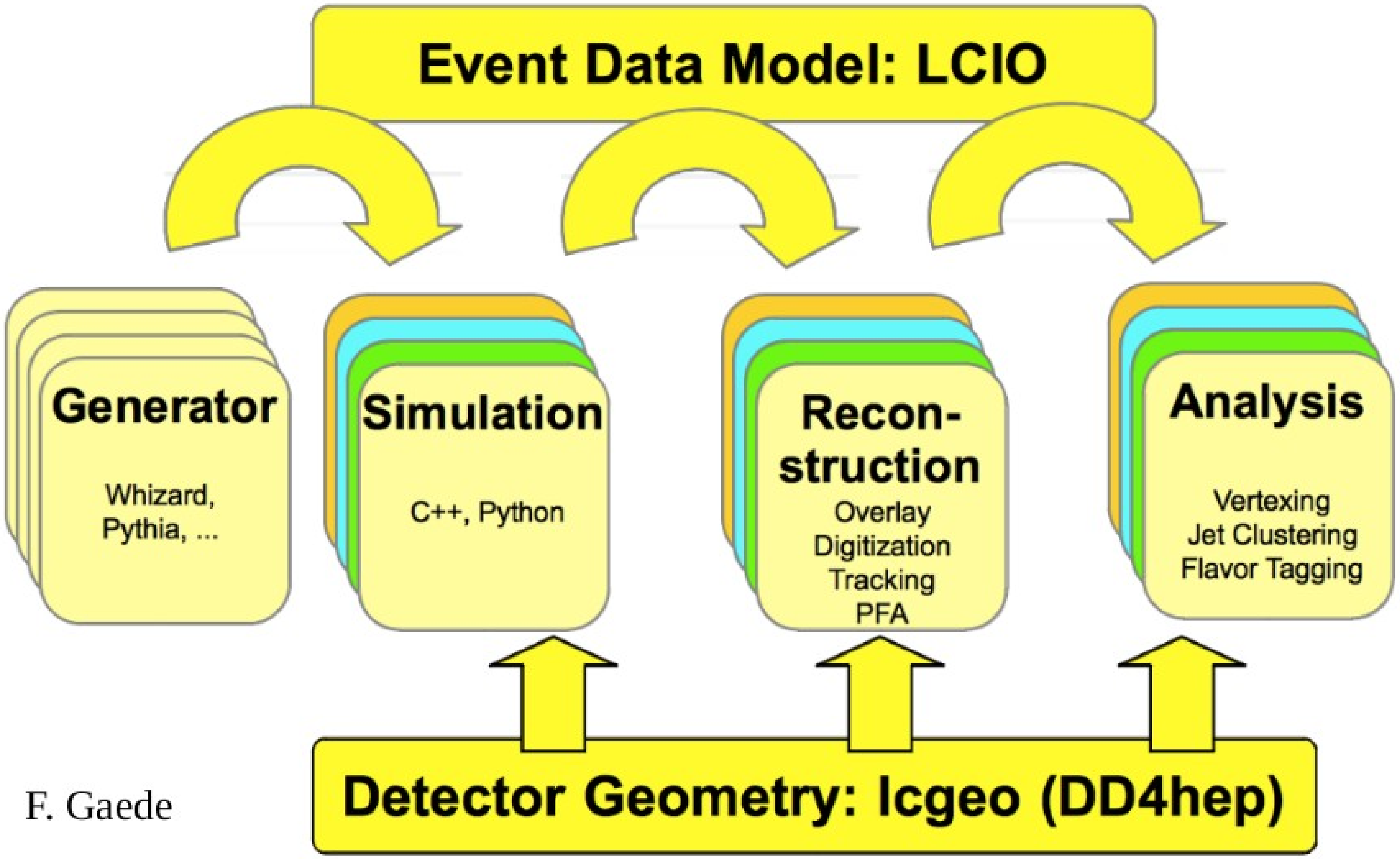
```
<processor name="Output_REC" type="LCIOOutputProcessor">
  <!--   standard output: full reconstruction keep all collections -->
  <parameter name="LCIOOutputFile" type="string"> Output_REC.slcio </parameter>
  <parameter name="FullSubsetCollections" type="StringVec"> EfficientMCParticles InefficientMCParticles </parameter>
  <parameter name="LCIOWriteMode" type="string" value="WRITE_NEW"/>
  <!--parameter name="SplitFileSizekB" type="int">996147 </parameter-->
  <parameter name="Verbosity" type="string">WARNING </parameter>
  <parameter name="DropCollectionNames" type="StringVec">  </parameter>
  <parameter name="DropCollectionTypes" type="StringVec">  </parameter>
  <parameter name="KeepCollectionNames" type="StringVec">  </parameter>
</processor>
```

Output file

ILCSoft workflow

ILCSoft workflow has three steps:

- Generation (not part of ILCSoftware framework and the tutorial)
- Simulation
- Digitization + reconstruction



Simulation step

- Based on GEANT4: The most common generator formats are supported in input: stdhep, HEPEvt, hepevt, hepmc, slcio.
- It implements a detector model