CentraleSupélec

université PARIS-SACLAY

Machine Learning Introduction
Hands On

David Rousseau

IJCLab-Orsay

david.rousseau@in2p3.fr @dhpmrou

SOS 2022, Carry-le-Rouet

cnrs IN2P3 Les deux infinis

IJCLab Irène Joliot-Curie Laboratoire de Physique des 2 Infinis

ML for Higgs physics

Using ML to see the Higgs Boson
Using Boosted Decision Tree first

Seeing the Higgs

Proton collisions

Before observation, all was known about the Higgs boson, except its mass

Probabilités de désintégration prédites pour une masse de 125 GeV

$H \rightarrow bb$	58%
$H \rightarrow WW^*$	21%
$H \rightarrow \tau^+\tau^-$	6.4%
$H \rightarrow ZZ^*$	2.7%
$H \rightarrow \gamma\gamma$	0.2%

$E=mc^2$

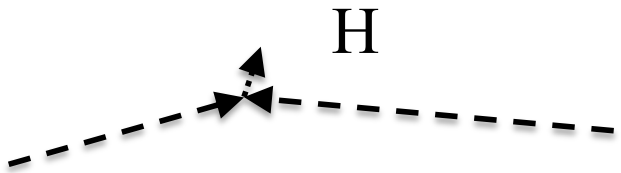
Einstein en 1905

$E^2=p^2+m^2$

gamma 1

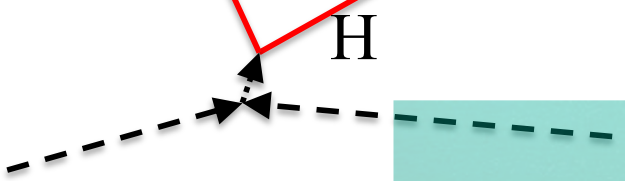
Measured (et m=0)!

gamma 2



H, before decay

$m_H^2=E_H^2-p_H^2$



H,Just after decay

Energy Momentum conservation

$E_H=E_{g1}+E_{g2}$

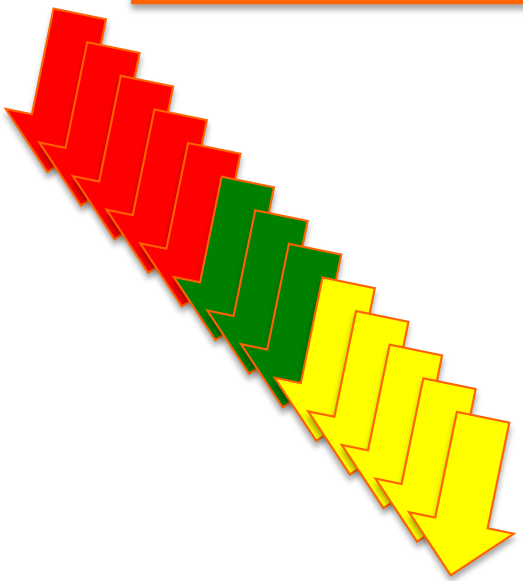
$\vec{p}_H=\vec{p}_{g1}+\vec{p}_{g2}$

⇨we get m_H!

10^{14} collisions / year

Trigger: fast rough selection

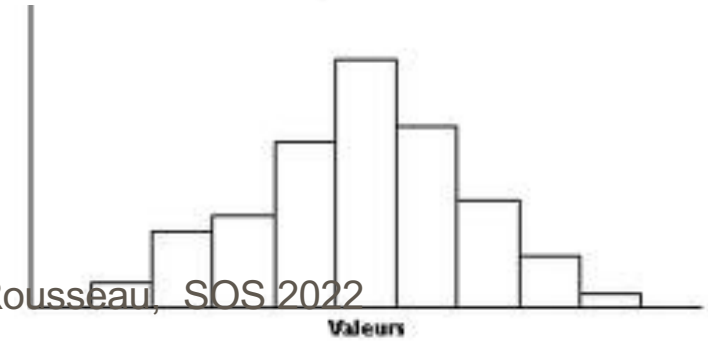
10^{9} events on disk

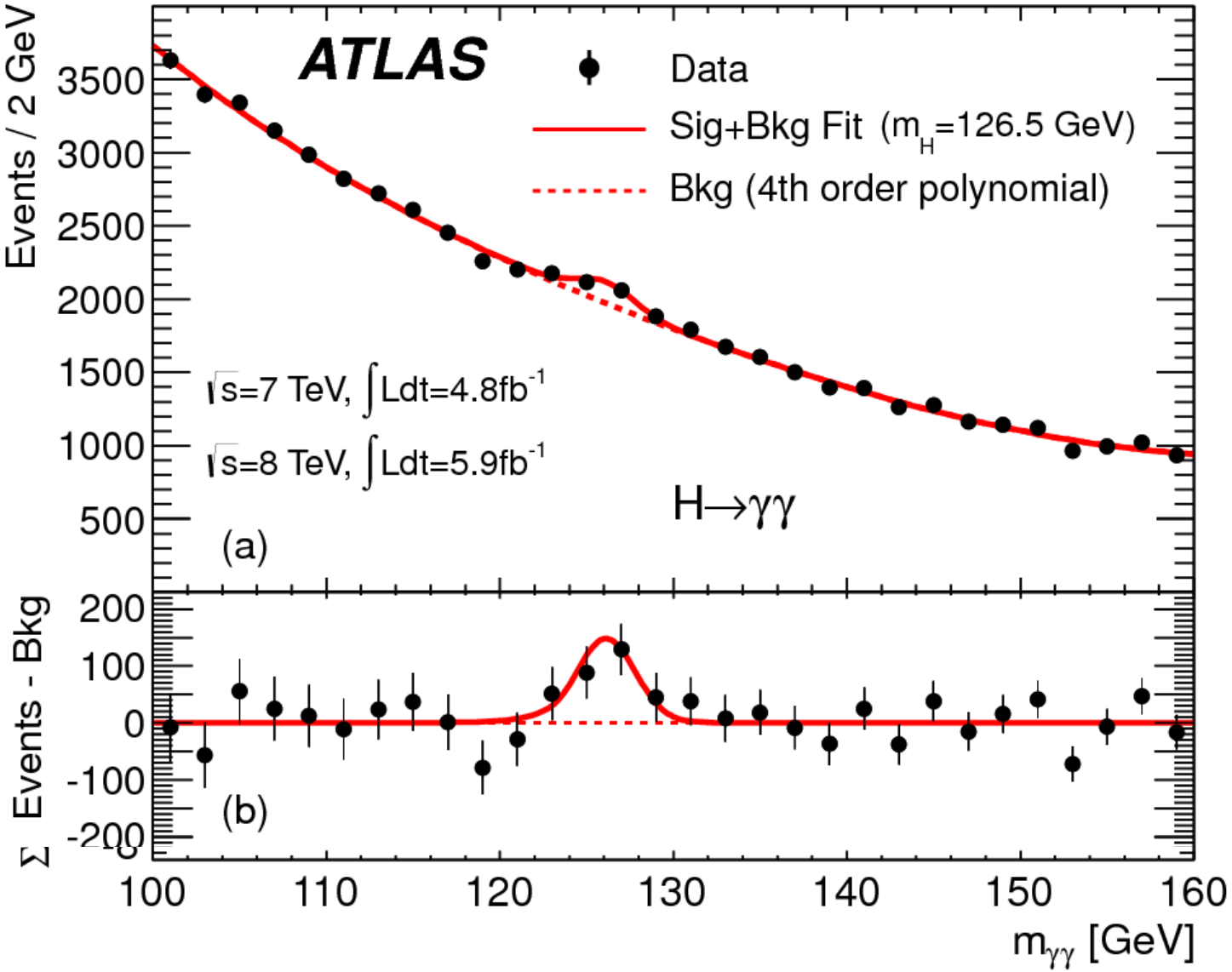


Tri précis

10^{5} events with 2 photons

Mass calculation
→histogramme

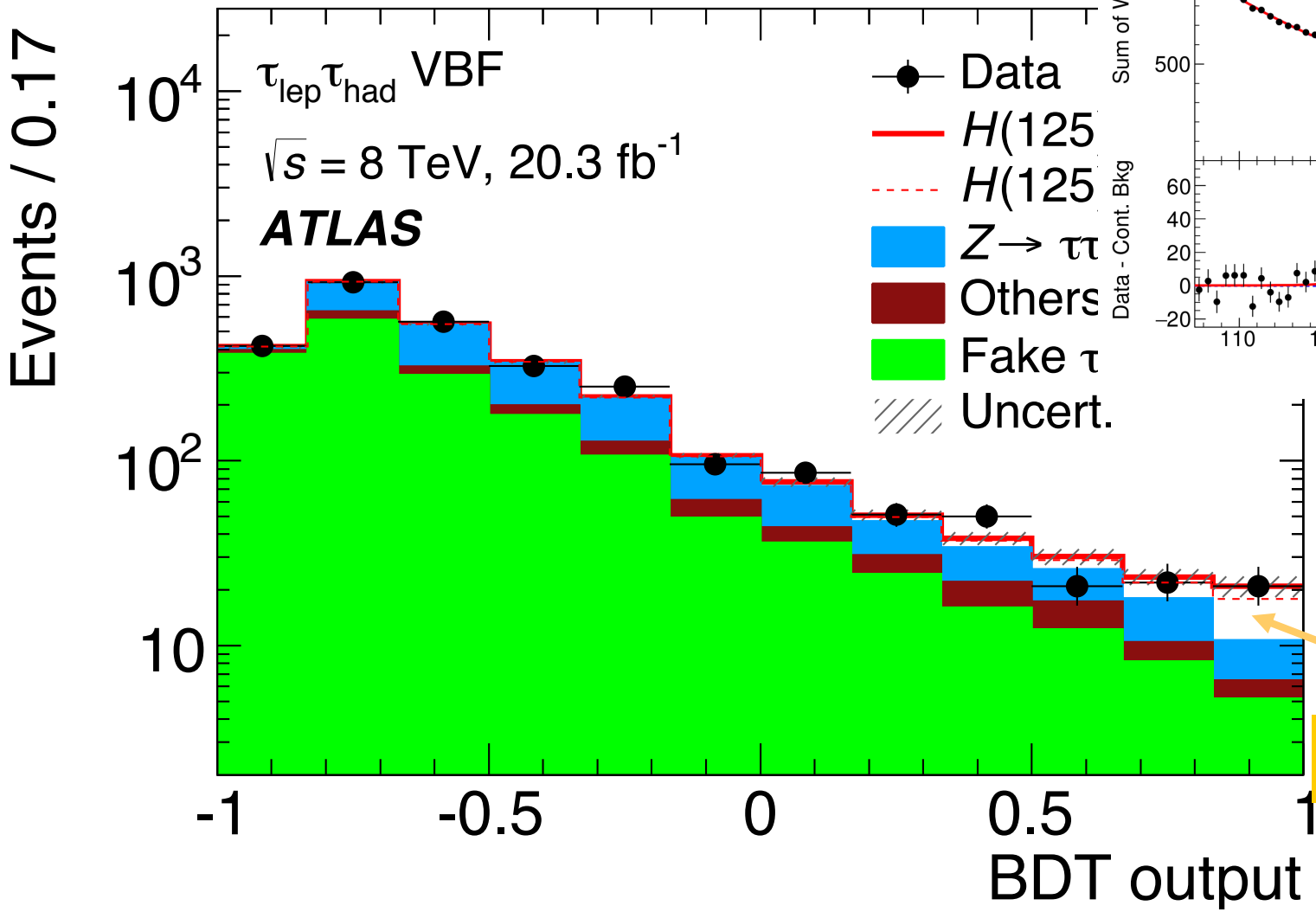


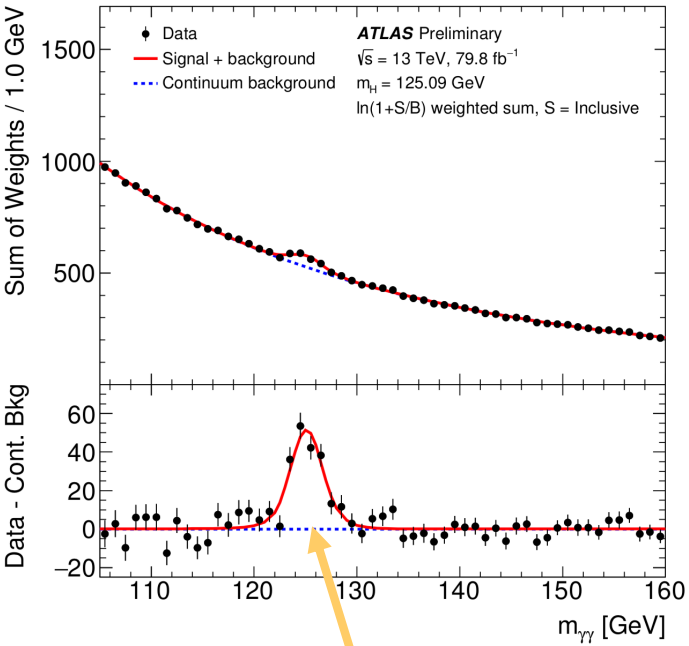

ATLAS
Data
Sig+Bkg Fit (m_H=126.5 GeV)
Bkg (4th order polynomial)
√s=7 TeV, ∫Ldt=4.8fb^-1
√s=8 TeV, ∫Ldt=5.9fb^-1
H→γγ
(a)
(b)
Events / 2 GeV
Σ Events - Bkg
m_γγ [GeV]

Classifier in Higgs Physics

JHEP 04, 117 (2015) 1501.04943

BDT using ~dozen of high level variables



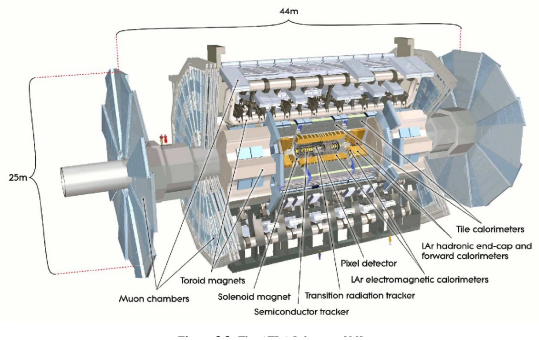


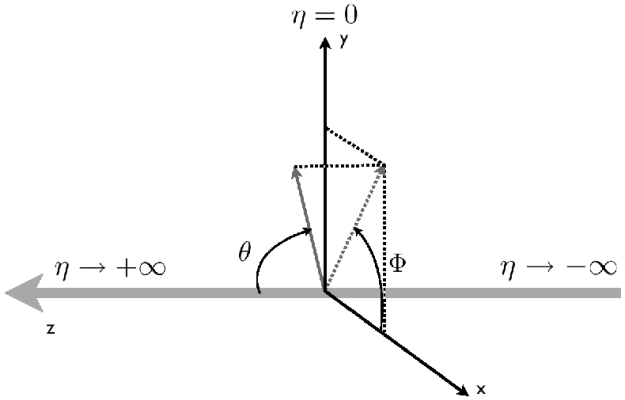
Higgs evidence

Two fundamental entities

- « Events » :
  - All measurements from one proton collision
  - List of particles with their properties
  - Derived quantities
  - =>ML to help select interesting events « Signal » with respect to « Background »
- « Particles »:
  - Extracted from an event
  - Jet, lepton, photon Missing ET
- ➔tabular data, one line per event, columns for particle properties
- (also possible to deal with variable number of particles, not today)

Coordinates





- P : momentum
- E : energy =sqrt(P^2+M^2)~P because P>>M
- Angles (cylindrical)
  - ϕ : azimuth angle]-π, +π]
  - θ : dip angle [0, +π]
  - η : eta, pseudo-rapidity = log(tan(θ/2)), ~[-5,5]
- P_T : =P sin(θ) : transverse momentum
- ME_T : Missing Transverse Energy = $-\Sigma_{\text{all particles}} \overrightarrow{P_T}$: estimator of transverse momentum of neutrinos

H→WW

- One of the Higgs Discovery channel
- $H \rightarrow W^+(\rightarrow l^+ \nu)\ W^+(\rightarrow l^- \nu)$
  - ➔2 leptons of opposite charge
  - Neutrinos undetected ! => Missing Transverse Energy
  - No invariant mass peak!
- Background :
  - Other processes leading to $W^+(\rightarrow l^+ \nu)\ W^+(\rightarrow l^- \nu)$

Before we start

- IN2P3 ML activities
  - IN2P3/CEA ML workshop 26-27 sep 2022
- Enquiry about IN2P3 ML project sent on Tuesday 17th may (deadline this Sunday…)
  - ➔collect material for IN2P3 Scientific Council 23rd June
- machine-learning-l@in2p3.fr
  - send an email to listserv@in2p3.fr with body of email:
    - subscribe machine-learning-l firstname lastname
- Other advertisements:
  - CERN IML workshop 9th-13th May 2022
  - Learning To Discover 19th to 29th April 2022

Let's do it !

- We'll use this notebook
- We'll run in Google Colab : dhrou/HEPMLtutorials/HEPML_HandsOn_BDT.ipynb branch SOS2022
- You can also install anaconda and run locally