Search for the Higgs boson in the H→ZZ*→ 4-lepton decay channel, with the ATLAS experiment at LHC

Part 2: data analysis

Stefano Rosati
INFN Roma

Introduction

- In the previous seminar by Ludovico, you have seen how a detector (let's pick a random one: ATLAS @ LHC) has been:
  - Designed
  - Built
  - Calibrated
- These are of course the first fundamental steps

- Today we'll see the following steps that lead to the Higgs boson discovery:
  - Design the analysis, collect the data and look for a signal
  - Did we really do a discovery ?
- And, we'll briefly see how some of the properties of the new resonance have been measured

Designing an analysis

- What signal am I looking for ?
  - Define the signature in terms of the final state properties
- Is my trigger able to store the results ?
  - Particular important at high-luminosity colliders like LHC
- What are the backgrounds ?
  - Are there already known processes, that produce a signature exactly like the one I am looking for
  - What are the handles I can use to reject the backgrounds
  - Can I use Monte-Carlo (MC) or do I need to estimate the backgrounds from the data
- How much data do I need, to be able to say something ?
- How can I measure the properties of a possible discovery ?

Outline

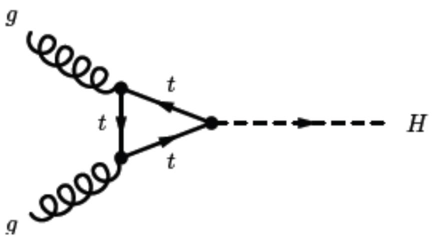
- A brief introduction on the SM Higgs search at LHC
  - Production mechanisms and decay channels
  - Description of main search channels and their characteristics
  - Assuming you already know the basics of Higgs Physics and its motivations
- The $H \rightarrow ZZ^* \rightarrow$4-lepton channel (lepton = electron or muon)
  - Leptons reconstruction in brief
  - Signal signature and backgrounds
  - Event selection
  - Exclusion limits and signal significance
  - Measurement of the signal properties
- In the end, you should be able to see how the main results of the Higgs search at LHC Run 1 were derived at ATLAS

Section 1

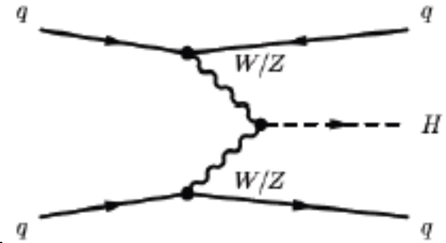
INTRODUCTION

Standard Model Higgs Boson @ LHC

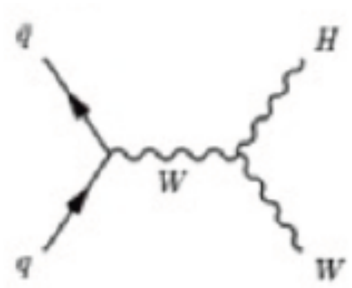
Production Feynman diagrams for a Higgs Boson at a p-p collider

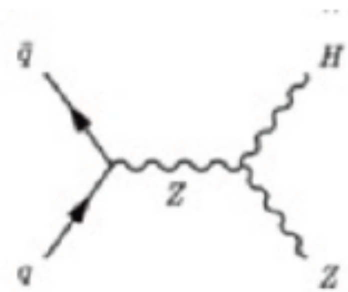


Gluon-gluon fusion:
~87%
Calculated with ~10% uncertainty
(PDF and QCD scales)

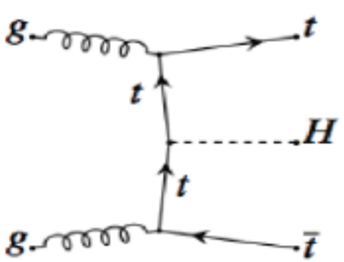


Vector Boson Fusion (VBF):
~7%
Calculated with ~3% uncertainty



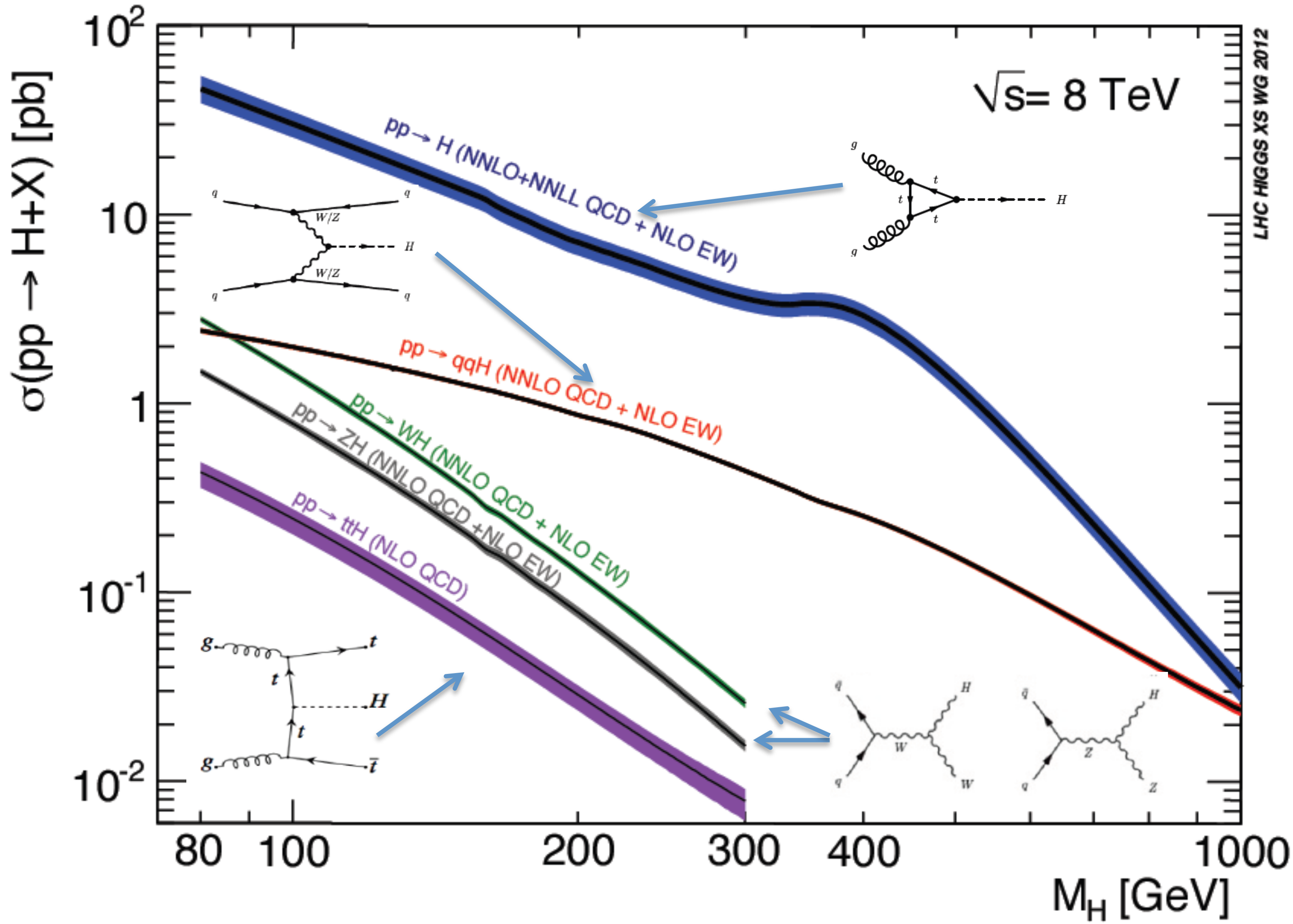


Associated production (WH, ZH):
~5% Calculated with ~3% uncertainty



tt-Higgs:
~1% Calculated with ~12% uncertainty

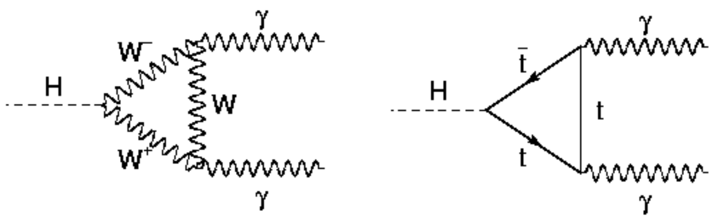
Mass dependence of the cross sections



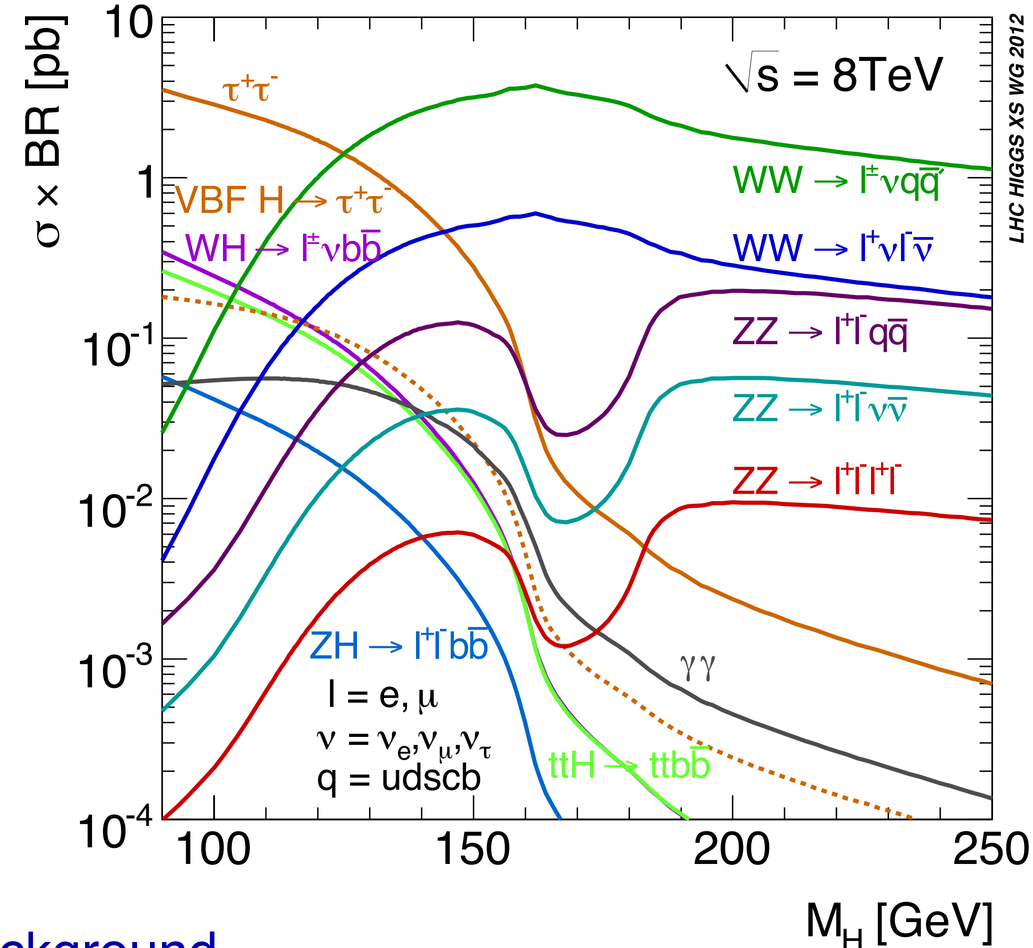
Decay branching ratios

- Higgs coupling proportional to M_f, M_W, M_Z

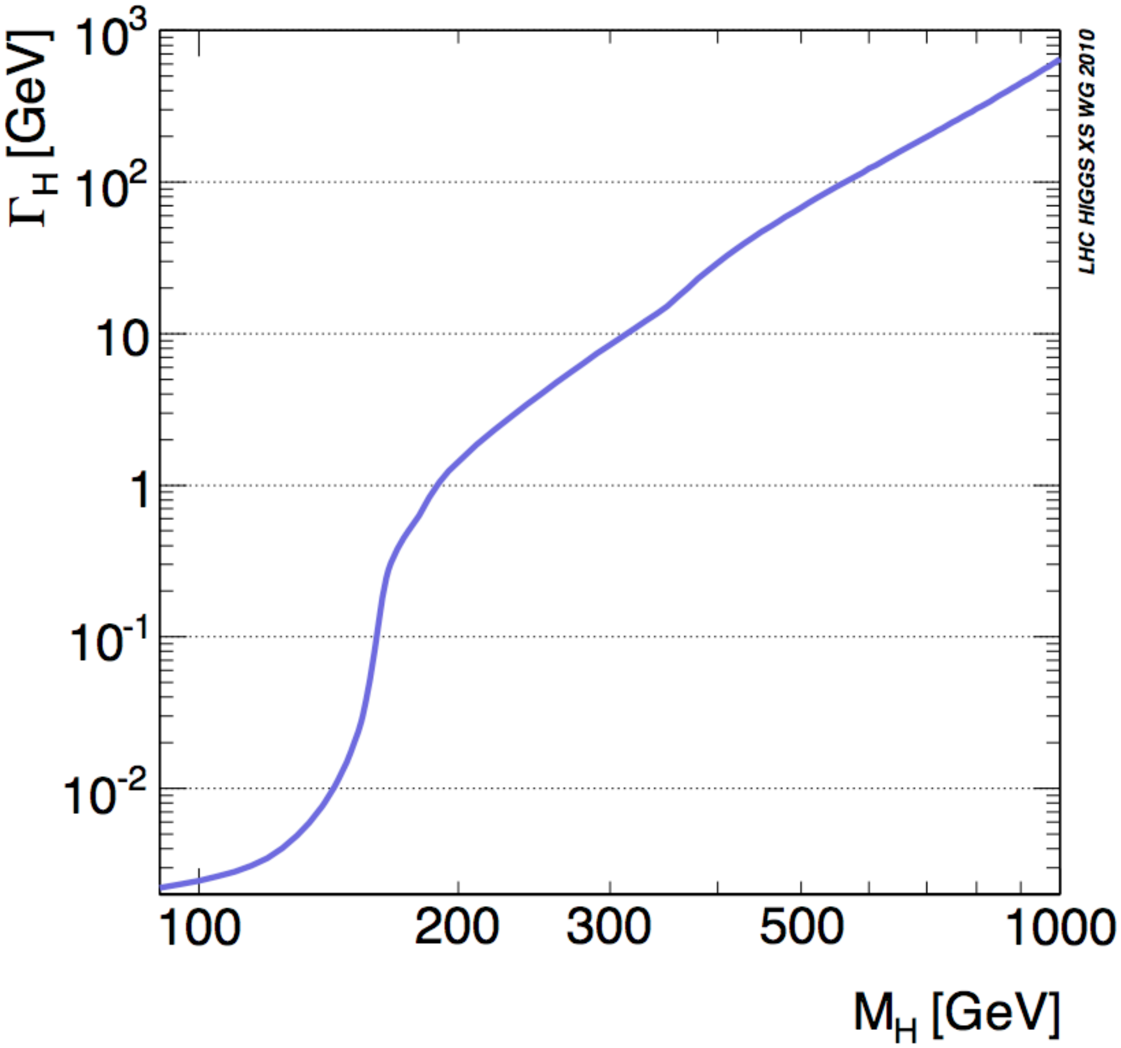
- The decay to a pair of photons Is possible through W and top loops:



- In spite of the low BR, the 4-lepton decay channel is a fundamental one:
  - Clean signature with low background
  - Final state is fully reconstructed allowing the best measurement of the properties (mass, spin/parity, cross sections)



Higgs width



In the low-mass region, i.e. below ~200 GeV:

→ natural width of the resonance negligible w.r.t. the typical resolutions of the "best" channels

→ $H \rightarrow \gamma\gamma$ and $H \rightarrow ZZ^* \rightarrow 4l$ have mass resolutions between ~1 and ~2 GeV

Event rates

Number of signal events in 20 fb^{-1} at 8 TeV center of mass energy
Standard Model Higgs with M_H = 125 GeV
Expected events before any selection

Channel	Events before selection
$H \rightarrow \gamma\gamma$	1000
$H \rightarrow WW \rightarrow lnln$	1200
$H \rightarrow ZZ^* \rightarrow 4l$ (Inclusive)	60
$H \rightarrow ZZ^* \rightarrow 4l$ (VBF)	3
$H \rightarrow ZZ^* \rightarrow 4l$ (VH)	2
$H \rightarrow \tau\tau$	30000

But, not only the signal rate is important of course.
What are the backgrounds ? What's their size with respect to the signal ?

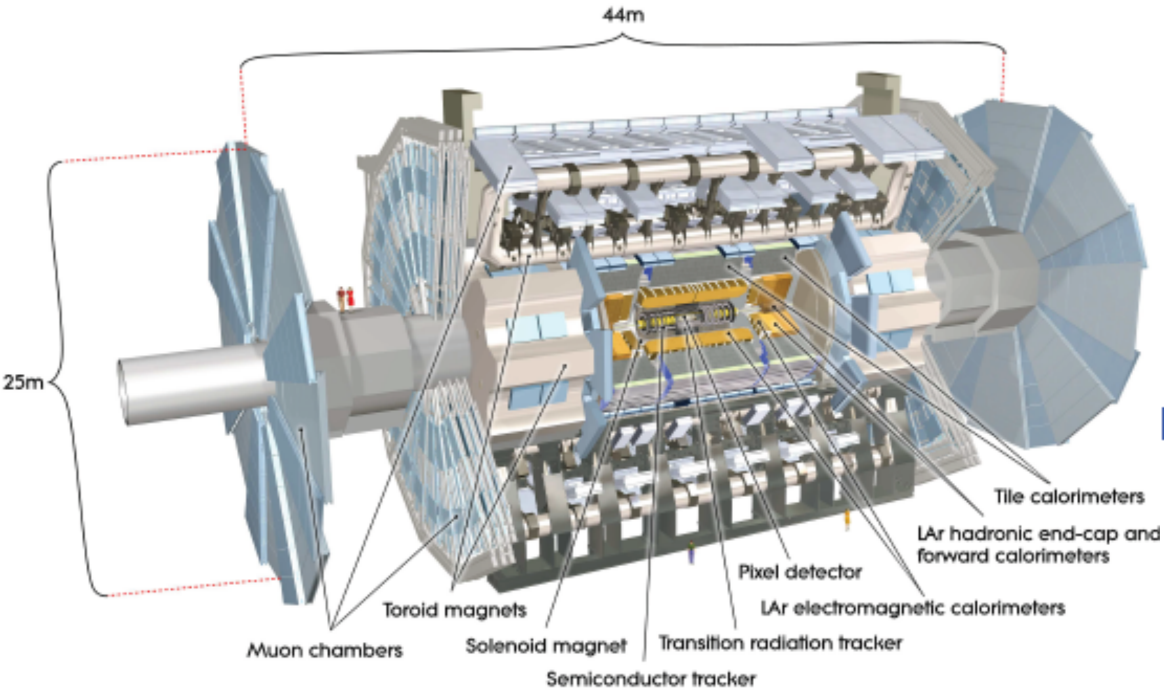
Some remarks on the plots

- When looking at histograms of any distribution for different physics processes, one can usually use two approaches:
- Normalize each histo to given integral (usually 1) in order to look in detail at shape differences
  - I.e. to normalize to 1, weight each event with the factor:
    - $w=1/N_i$
      where N_i is the total number of events you have in your MC for the process i
  - You will not see the actual number of expected events in the histo, but will be able to compare the shapes
- Normalize each histogram based on the cross section and a given integrated luminosity
  - This means to weight each event of the process i with a factor:
    - $w= \mathcal{L}\sigma_i / N_i$
      where $\mathcal{L}$ is the integrated luminosity you want to normalize to, σ_i is the cross section of the process, and N is the tot number of MC events
- In the following you will see plots normalized in both ways

Section 2

THE ATLAS EXPERIMENT AND THE LHC RUN-1 DATASET

The ATLAS detector



Magnets	1 Central Solenoid (2 T) + 3 air-core toroids
Tracking	Silicon+Transition radiation tracker
EM calo	Sampling LAr calo
HAD calo	Plastic scintillator (barrel) LAr technology (endcap)
Muon	Reco and trigger Standalone reco capabilities

ATLAS p-p run: April-Sept. 2012

Inner Tracker			Calorimeters		Muon Spectrometer				Magnets	
Pixel	SCT	TRT	LAr	Tile	MDT	RPC	CSC	TGC	Solenoid	Toroid
100	99.3	99.5	97.0	99.6	99.9	99.8	99.9	99.9	99.7	99.2

All good for physics: 93.7%

Luminosity weighted relative detector uptime and good quality data delivery during 2012 stable beams in pp collisions at √s=8 TeV between April 4th and September 17th (in %) – corresponding to 14.0 fb^{-1} of recorded data. The inefficiencies in the LAr calorimeter will partially be recovered in the future.

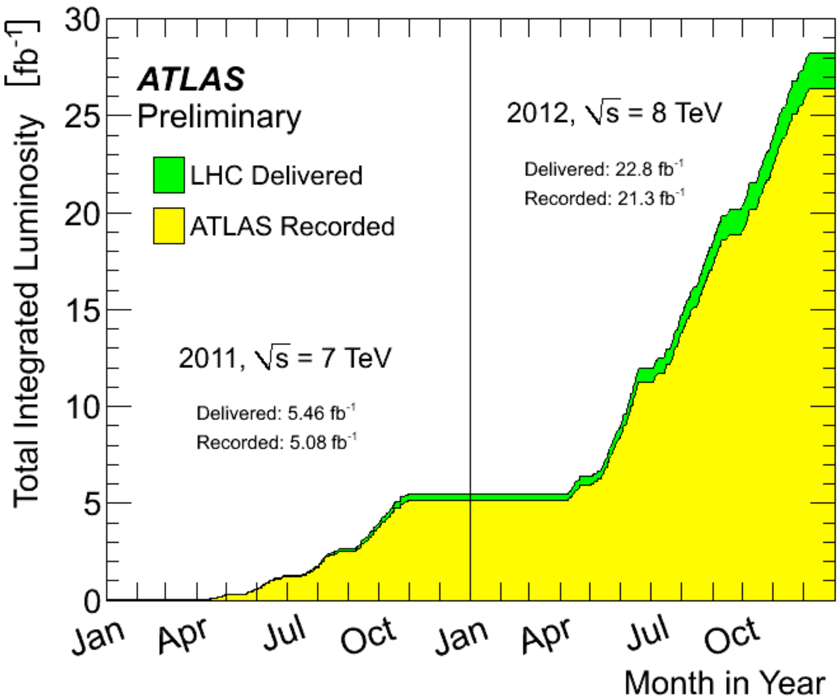
You have seen all the details in the seminar by Ludovico

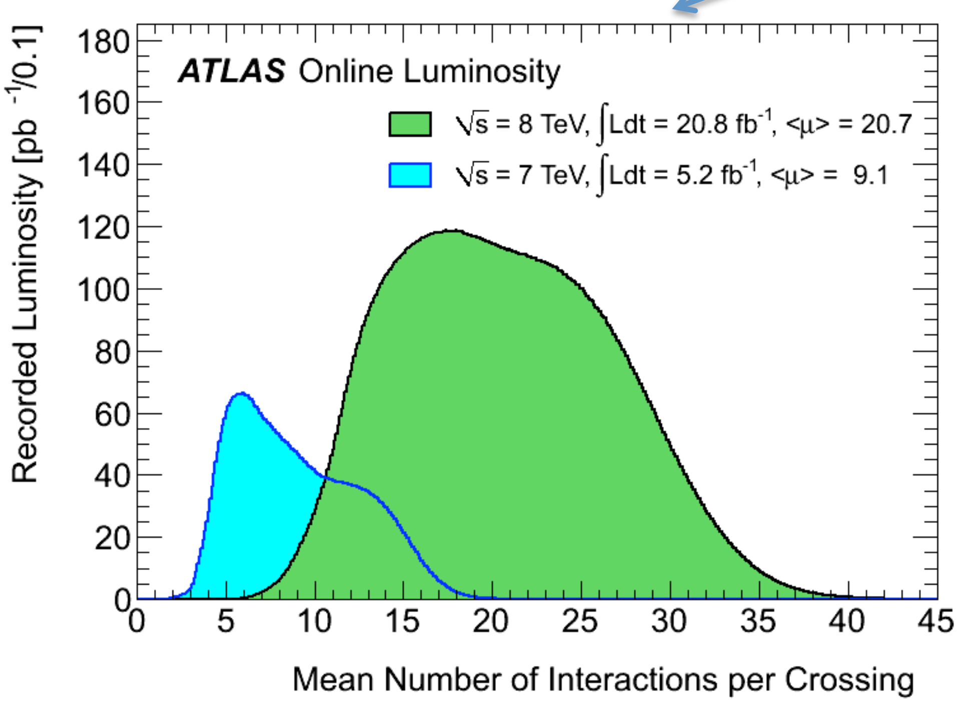
2012 data
→ 10 PB of data !

The Run-1 LHC dataset

p-p center of mass energy at 7 TeV in 2011 (integrated luminosity $\mathcal{L}$=5.08 fb^{-1}) and 8 TeV in 2012 ($\mathcal{L}$=20.8 fb^{-1})

Pileup !

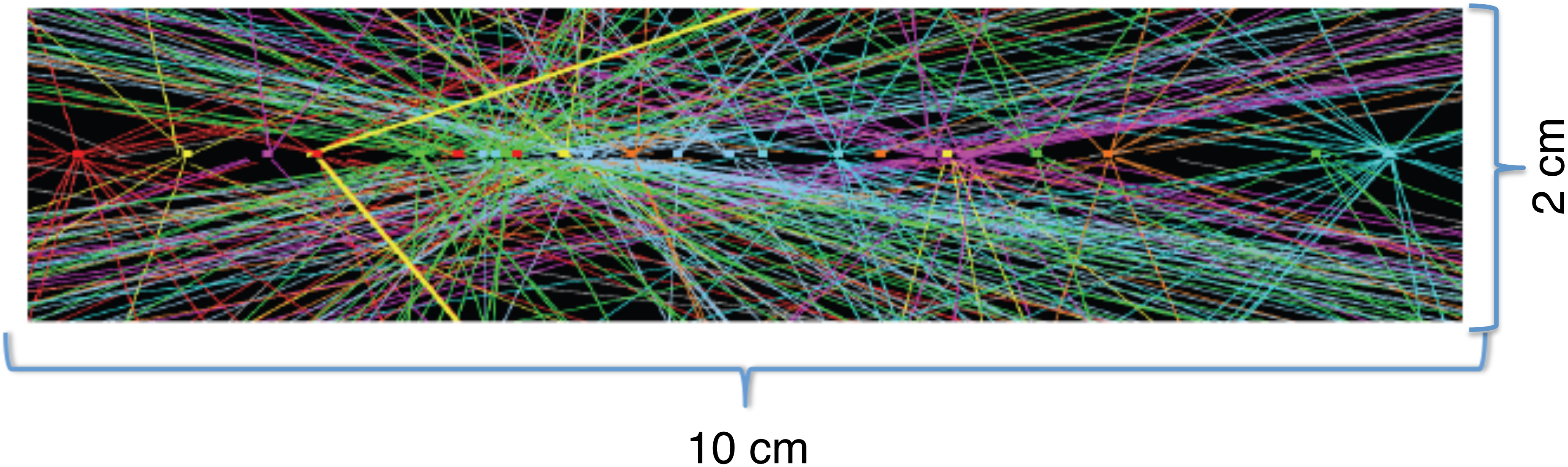




Peak instantaneous lumi was 7.7•10^{33} cm^{-2}s^{-1} in 2012

Pileup

That's how a typical event at 8 TeV looks like, in the region around the primary vertex



The σ of the interaction point along the beam axis (z axis) is ~5 cm, while the reconstruction resolution is ~90 μm
→ The z IP is used to associate the tracks to each primary vertex

While on the x and y planes (transversal to the beam) the resolution is comparable to the spread of the interaction point (~15 μm)

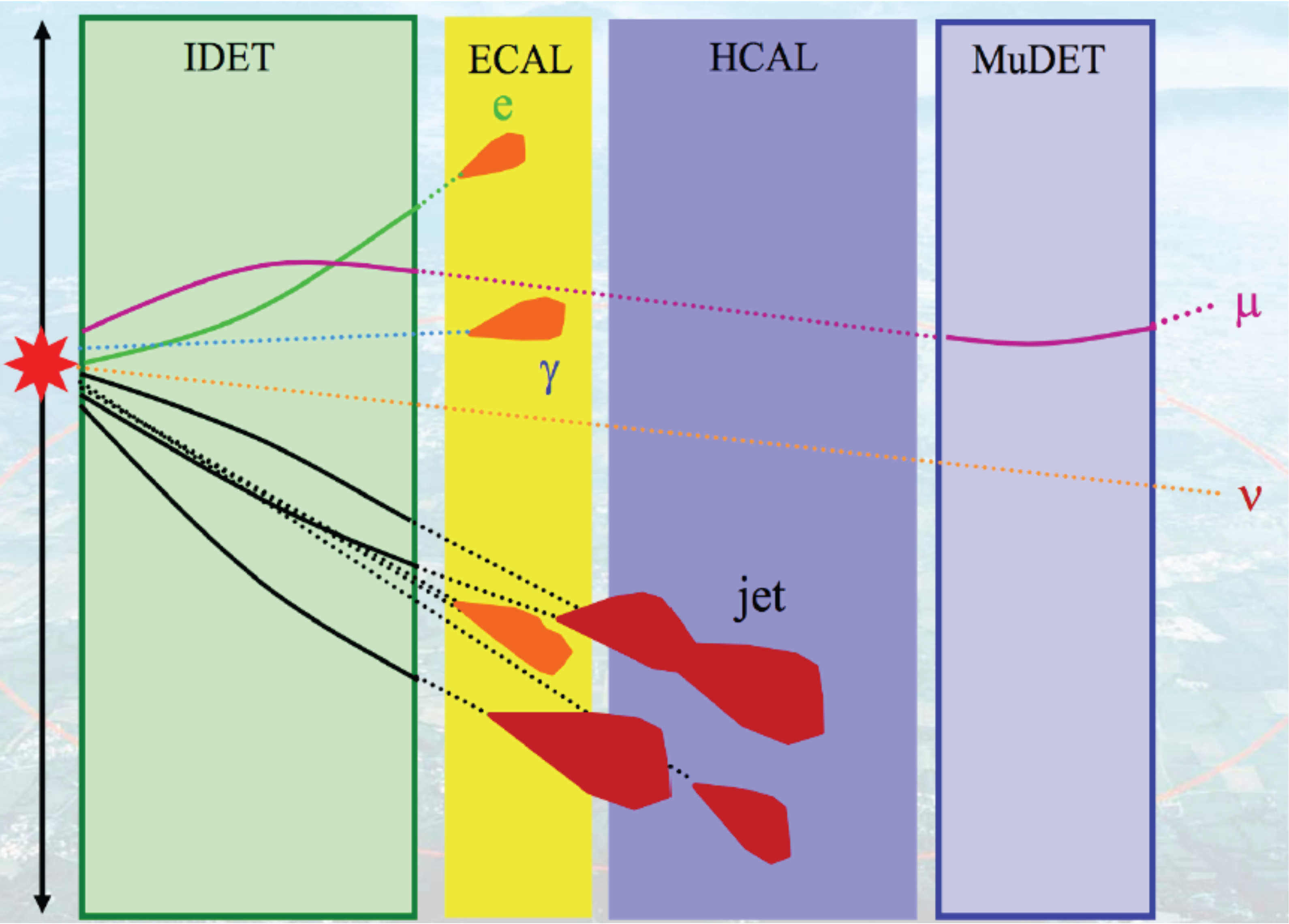
Section 3

SELECTION OF H→ZZ* EVENTS, IN THE 4-LEPTON CHANNEL

Steps for building an event selection - 1

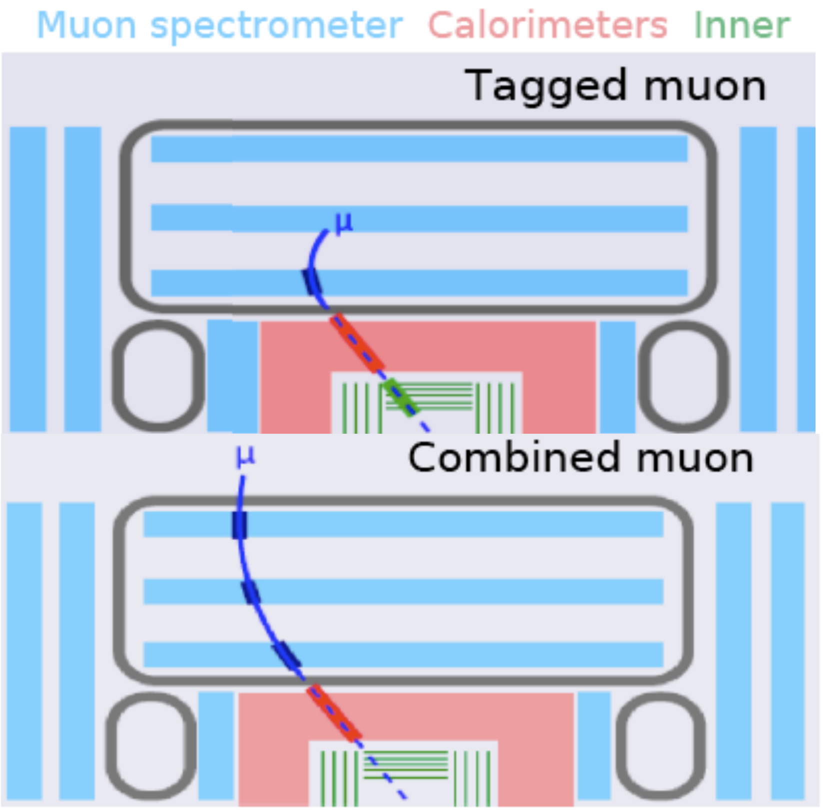
- When you think about an event selection, the first fundamental step is:
- Define the reconstructed "objects" that you will use in the analysis:
  - E.g. in the case of the $H \rightarrow ZZ^* \rightarrow 4l$ search the relevant objects are muons and electrons
- Check the reconstruction performance
  - The MC models a perfect detector, in the best possible ways, some corrections might be needed
    - Which methods can be used to determine them

Identification and reconstruction



Muon identification and reconstruction

- Muons are identified by:
  - Tracks in the Muon Spectrometer
  - Segments of tracks in the inne station of the muon spectrometer (low-pT muons not reaching all 3 stations)
    → tagged muon
- Tracks in the MS are back-extrapolated to the ID, correcting for energy losses in the calo
  - Look for a matching ID track
    →combined muon
- Tracks in the ID are out-extrapolated in the case of tagged mons



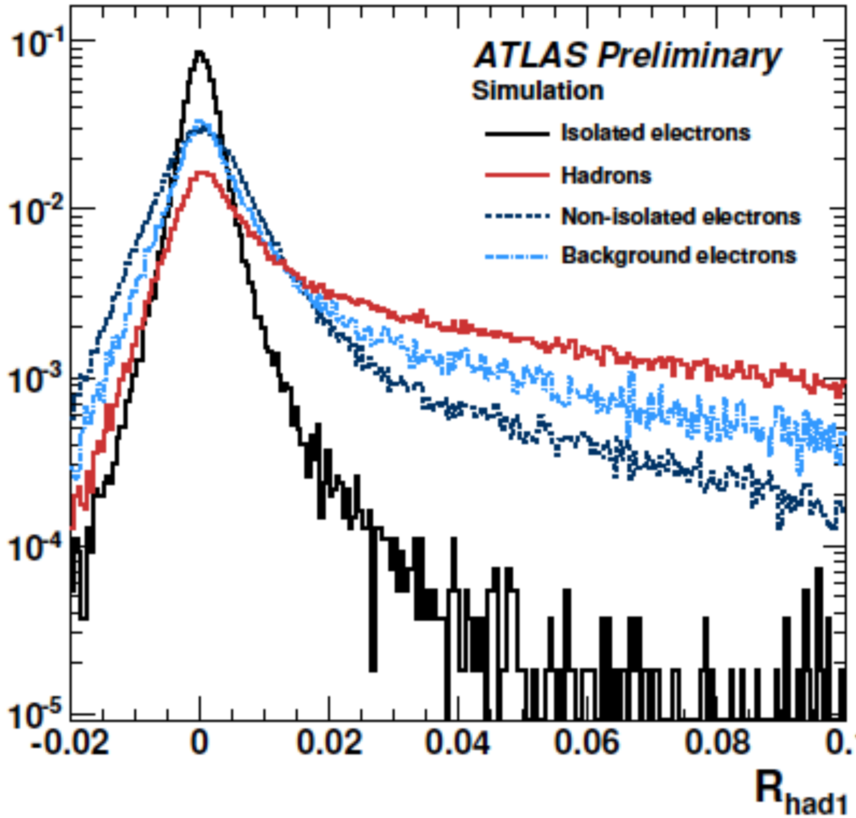
See the last seminar by Ludovico for all the details on the performance !

Electron identification and reconstruction

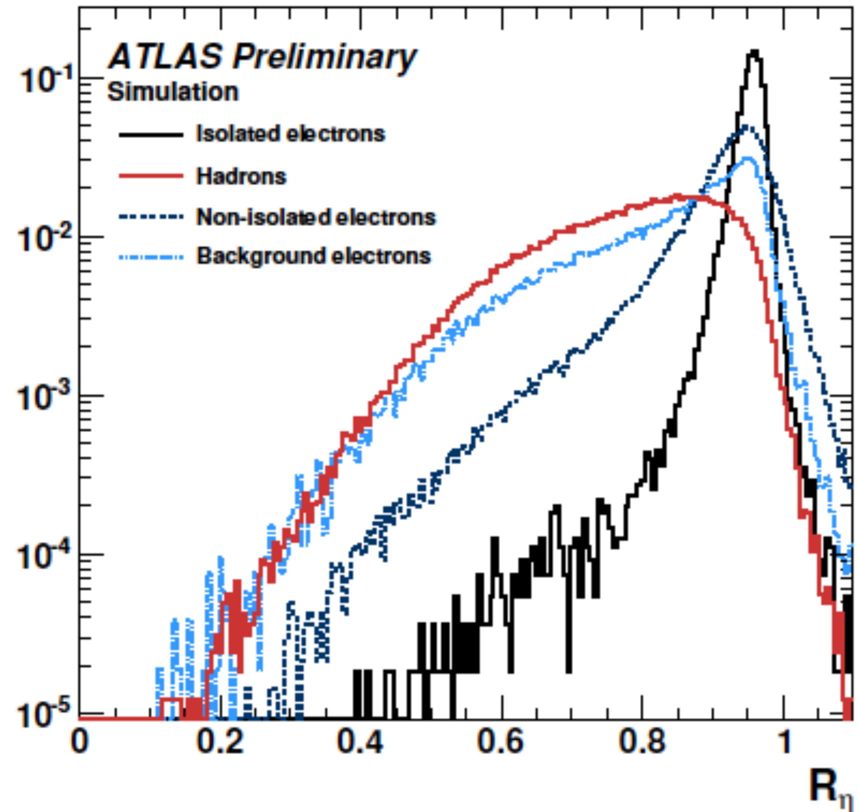
- Look for an ID track pointing to a cluster in the electromagnetic calorimeter
- Identification criteria must provide good separation with respect to jets faking electrons
- Some examples of general discriminating variables are:
  - Hadronic leakage:
    - ratio of energy in hadronic calorimeter / EM cluster energy (Rhad)
  - Shower shape variables
    - Ratio of inner cluster cells/total cluster (R_η)
    - Ratio of last sampling / first samplings
  - Track / cluster matching:
    - $\Delta\eta$, $\Delta\phi$ btw track and cluster
    - Ratio of cluster energy / track momentum
  - A few more ATLAS-specific variables (like hits in the TRT etc.)
- Either cut-based selection, or multi-variate
  - We'll discuss the difference in more detail later

Electron discriminating variables – 1

Hadronic leakage

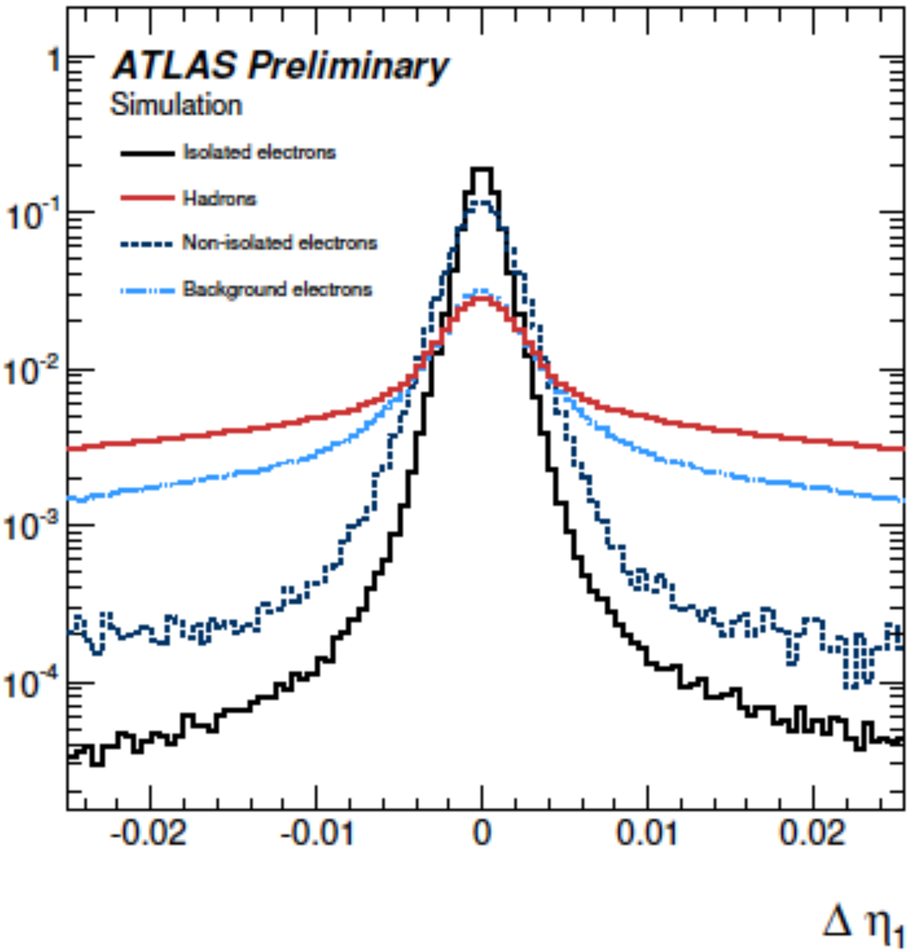


Inner cells / total cluster energy

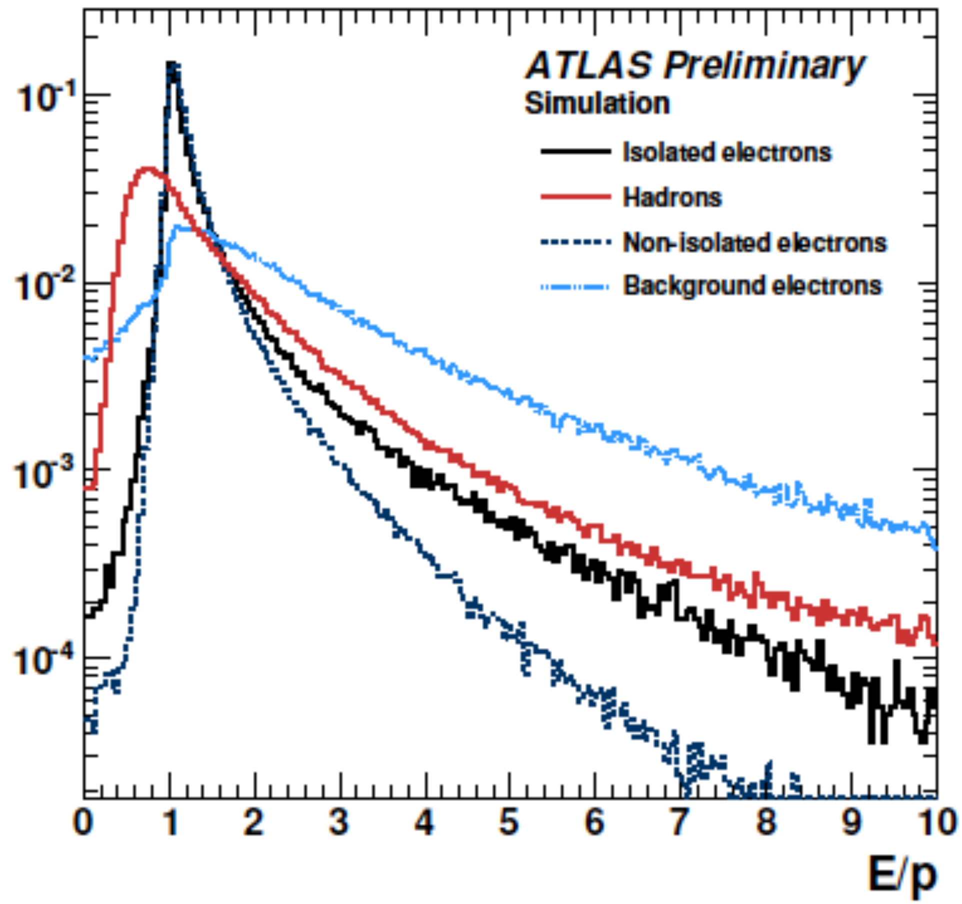


Electron discriminating variables - 2

$\Delta\eta$ between track and cluster



Cluster energy / track momentum

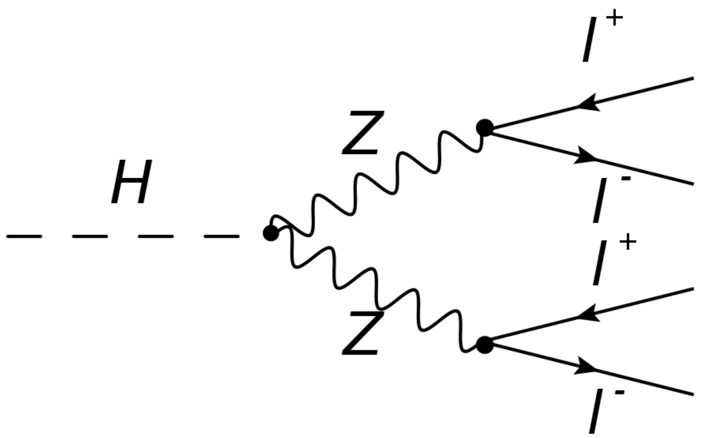


Steps for building an event selection - 2

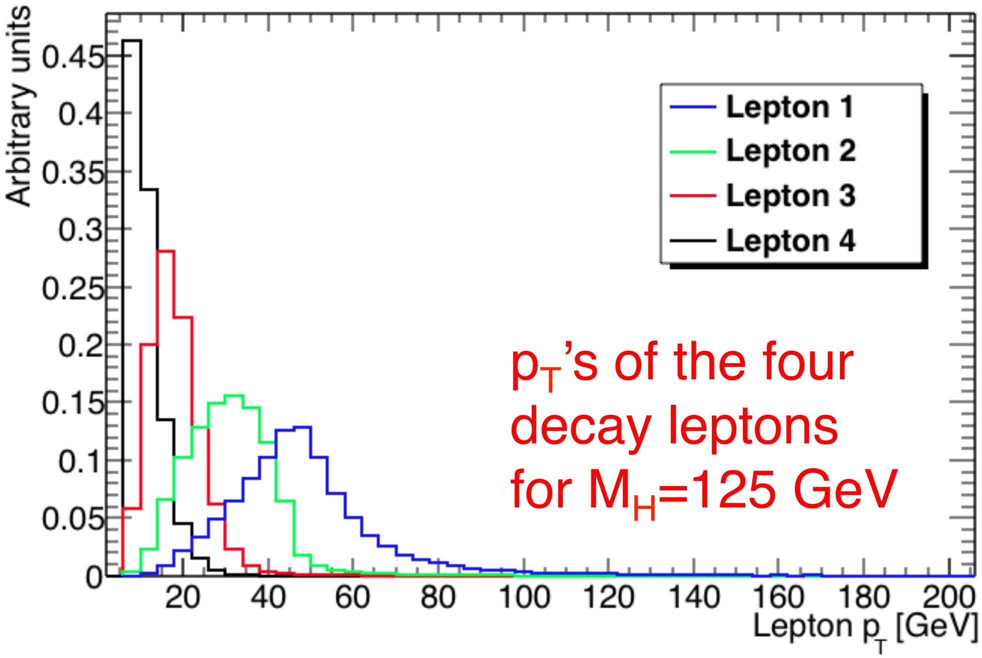
- In particular for searches, it is important to define the event selection without looking at the data (blind analysis)
  - Avoid biases in the definition of cuts: looking at the data one can pick excess regions and artificially enhance a significance
    - This should be avoided by all means
- So in general all the selection steps are defined using MC only
  - Data can also be used for some purposes, but in regions where the signal is not expected, the so-called control regions, or sidebands
- Once the selection is fully defined on MC:
  - Look at the data applying the cuts that you have defined

Basic signal kinematics

- The Higgs decays to a pair of Z bosons
- For M_H<~180 GeV (smaller than twice the Z mass)
  - One Z is on-shell, the other is off-shell at lower masses
- For MH>~180 GeV
  - Both Z's are on-shell
- Here we'll focus on the search in the low-mass region
  - Where the Higgs was actually found !

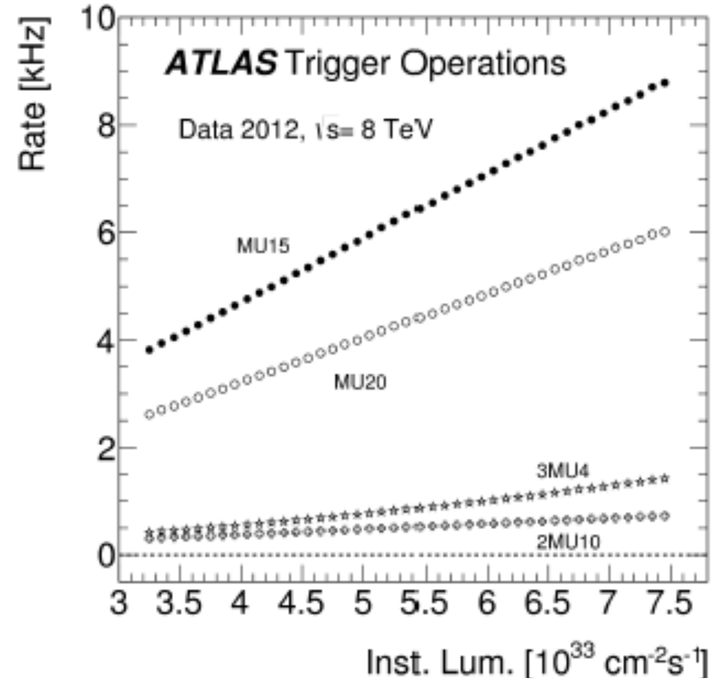


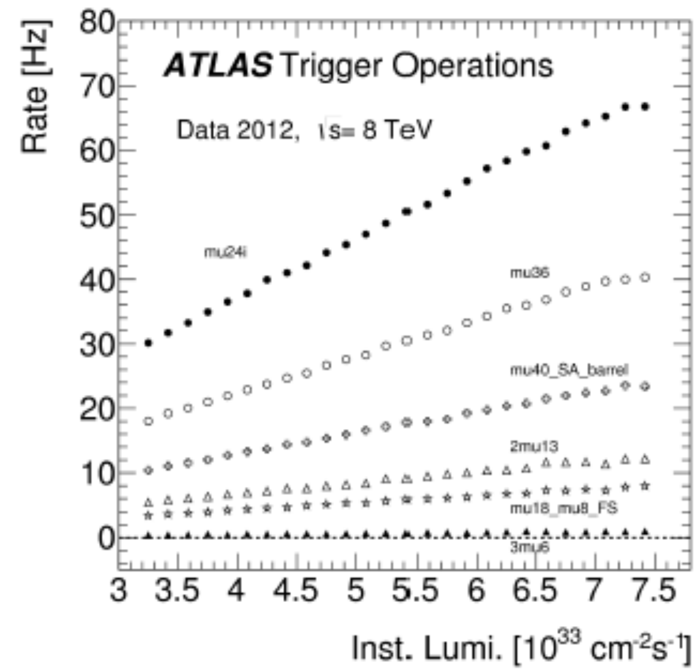
p_T's from ~5 to 100 GeV



The Trigger

- The trigger setup is the result of a compromise:
  - Keep the rate of accepted events at a level that can be sustained by the system:
    - ➢ Raise the thresholds
  - Keep high efficiencies for relevant Physics signals
    - ➢ Lower the thresholds
- Data Acquisition (DAQ) limits:
  - L1: 65-70 Hz
  - L2: 5 kHz
  - EF: 400 Hz
- Many signatures have to be combined in a menu, keeping the total rate within the DAQ limits
- Lepton thresholds always below 25 GeV during Run-1 (or 12-15 GeV for di-lepton)
  - Not a problem for the 4l-channel
- Efficiency is important → more later



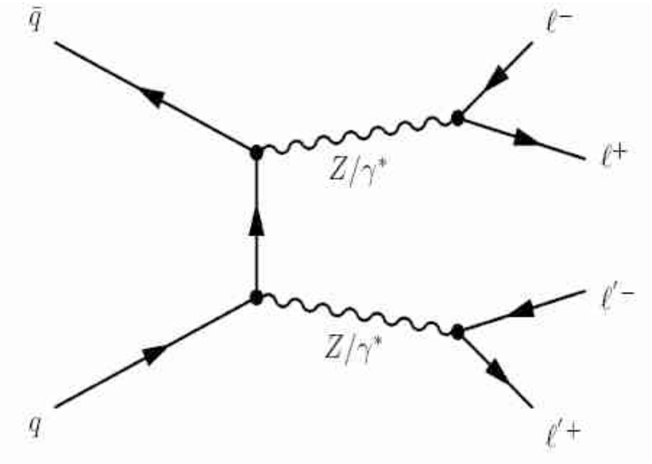


Backgrounds

- When looking for a signal, most important thing is… the background
- Think of all possible processes that can give the same signature of your signal
- Can these events be rejected (fully or partially) ?
  - Identify any property of the background events, different from those of the signal → "reducible" background
  - If it can't be fully rejected, it will have to be taken into account in all following measures
- If instead the background signature is in everything identical to the signal, its events can't be rejected and will have to be taken into account later in the analysis
  - In this case → "irreducible" background

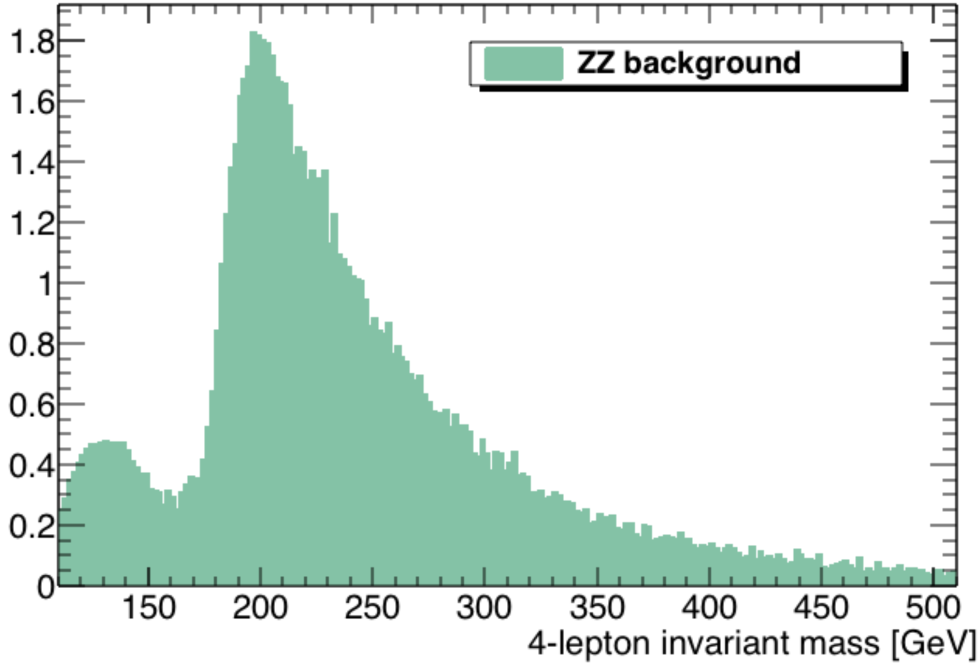
Irreducible background

- The process qq→ZZ$^{(*)}$→4l has the same final state of the signal
- Only difference is the mass distribution
  - There are a couple of more subtle ones to which we'll come back later



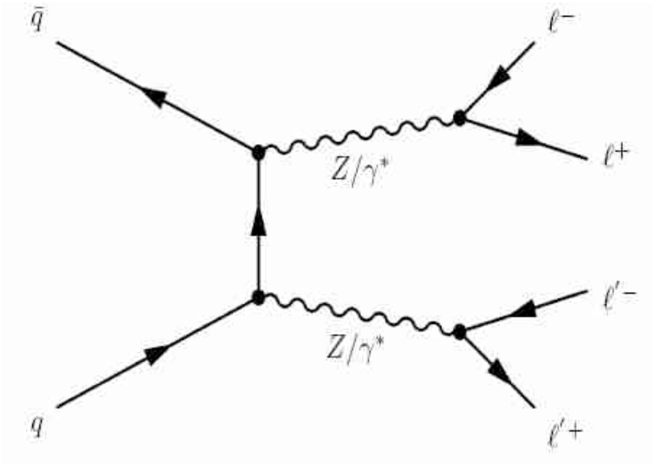
Onset of two on-shell Z's production from 2 M_Z, i.e. from about 180 GeV

Events are normalized to 20 fb^{-1} and based on each process cross section



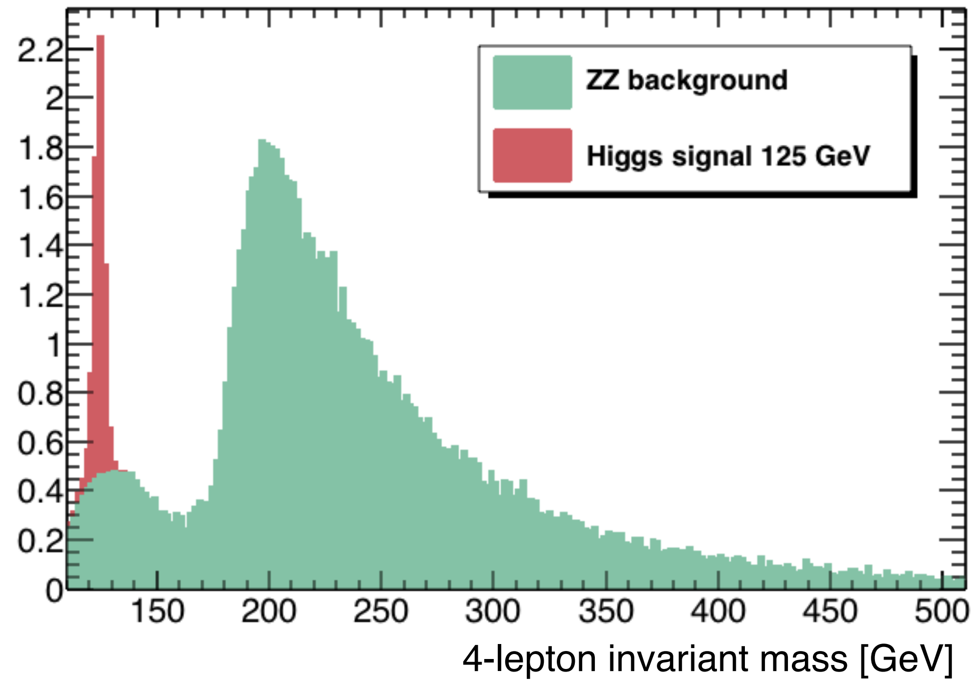
Irreducible background

- The process qq→ZZ(*)→4l has the same final state of the signal
- Only difference is the mass distribution
  - There are a couple of more subtle ones to which we'll come back later



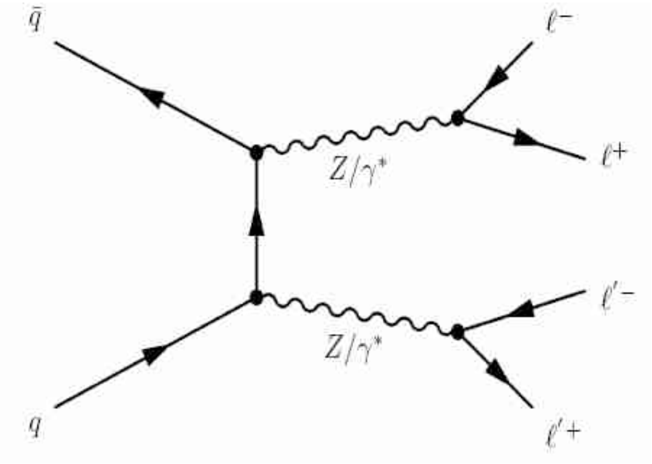
With the signal (@125 GeV) superimposed

Events are normalized to 20 fb^{-1} and based on each process cross section



Irreducible background

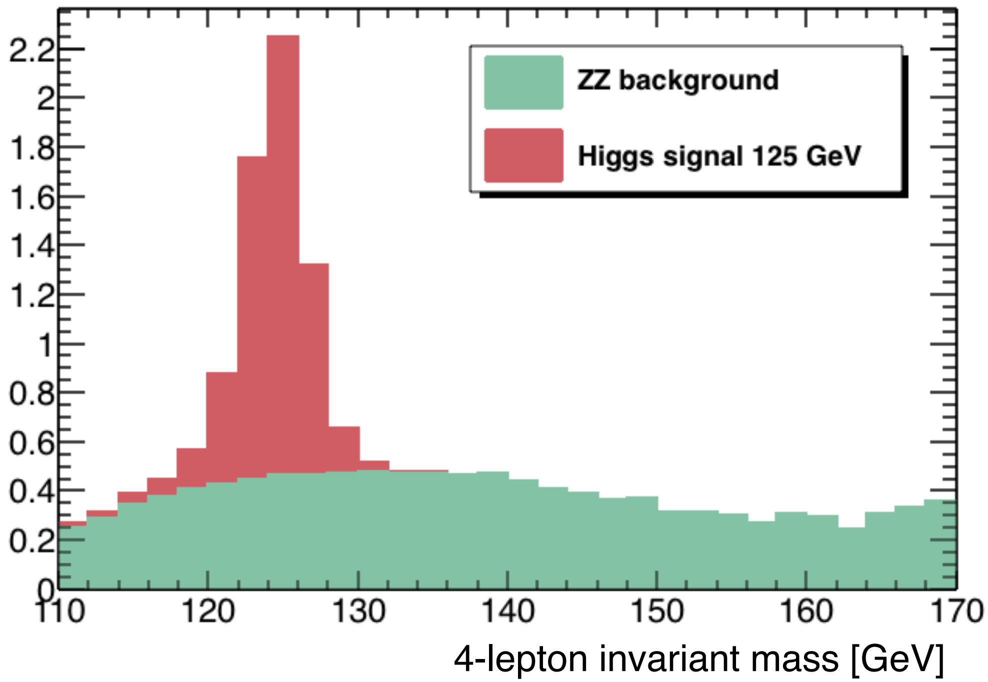
- The process qq→ZZ$^{(*)}$→4l has the same final state of the signal
- Only difference is the mass distribution
  - There are a couple of more subtle ones to which we'll come back later



With the signal (@125 GeV) superimposed

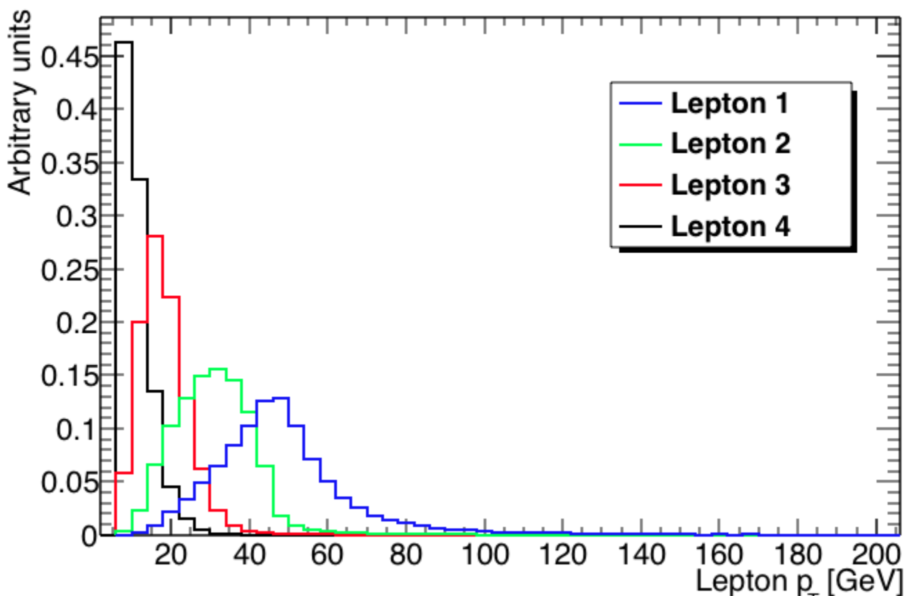
Events are normalized to 20 fb^{-1} and based on each process cross section

Zoom in the low-mass region →



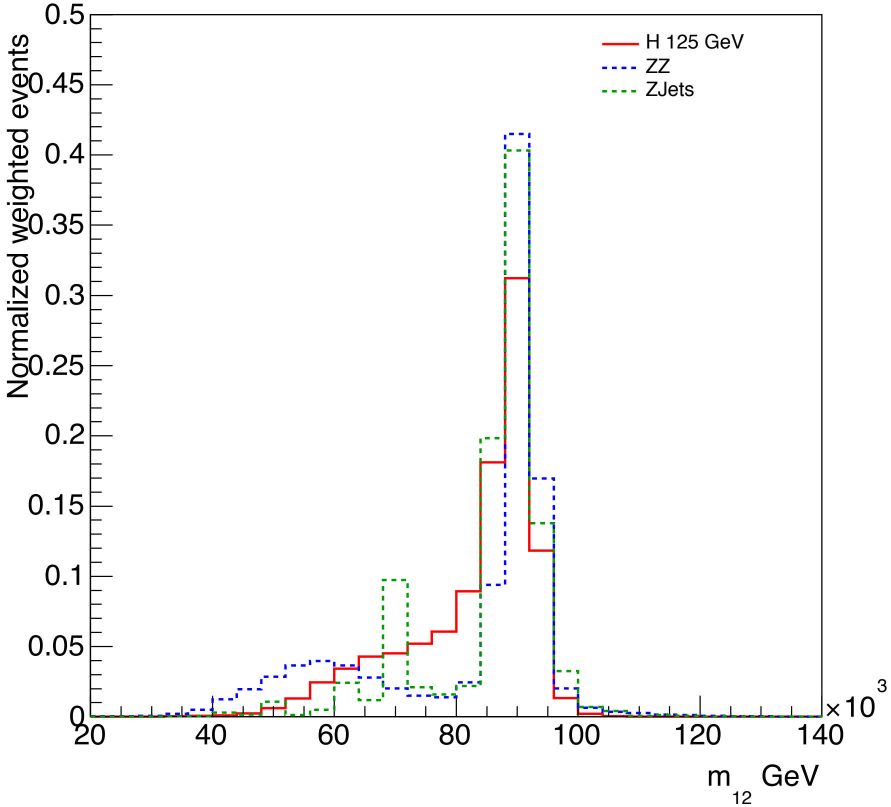
First steps of the selection

- Require 4 reconstructed leptons (e or μ) coming from the same primary vertex
  - Some basic quality cuts are applied besides the standard identification:
    - ➢ Number of hits used for the reconstruction in each detector
- The composition of the quartet defines the decay channel
  - 4μ, 4e, $2e2\mu$
- Apply the first cuts on the lepton p_T's:
  - p_T>6 (7 for e), 10, 15, 20 GeV
- Cut a little higher on electrons due to performance corrections (more details later)

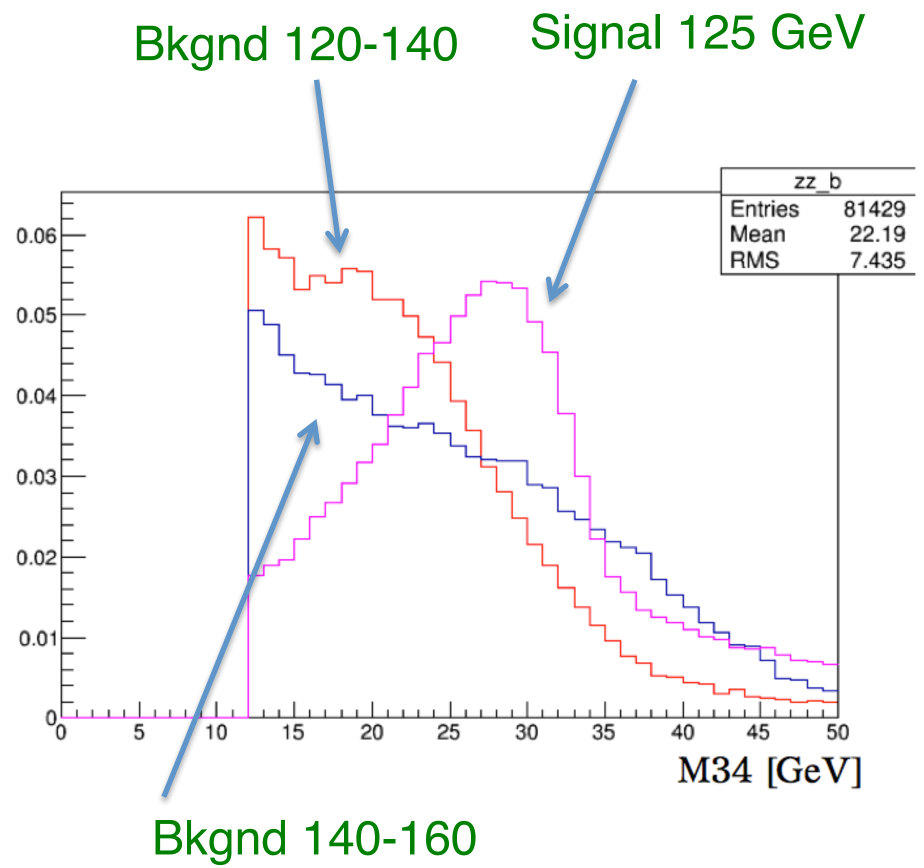


Other kinematic cuts

On-shell Z



Off-shell Z



Leptons are paired according to type and charge, then for low-mass H search, the paired closer to the Z mass is called Z1 (on-shell Z), the other Z2 (off

Significance

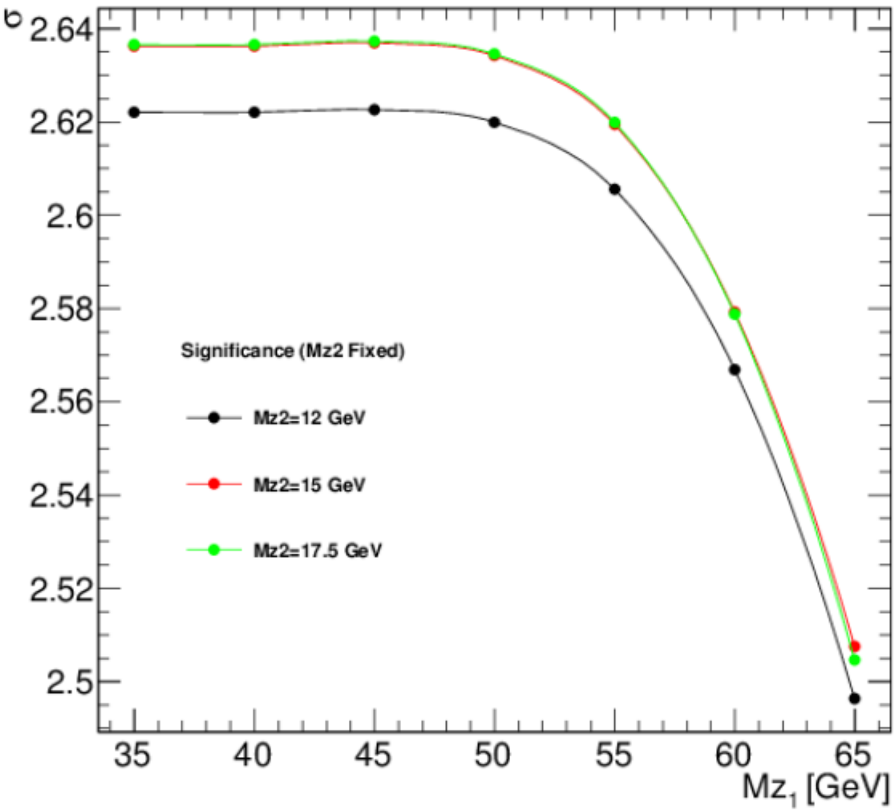
- When doing a search, the optimization of the cuts is driven by the maximization of the signal significance
- During the optimization, if the systematics are small one can normally use the simple expression you have seen in the lectures for the significance (n. of std. deviations):

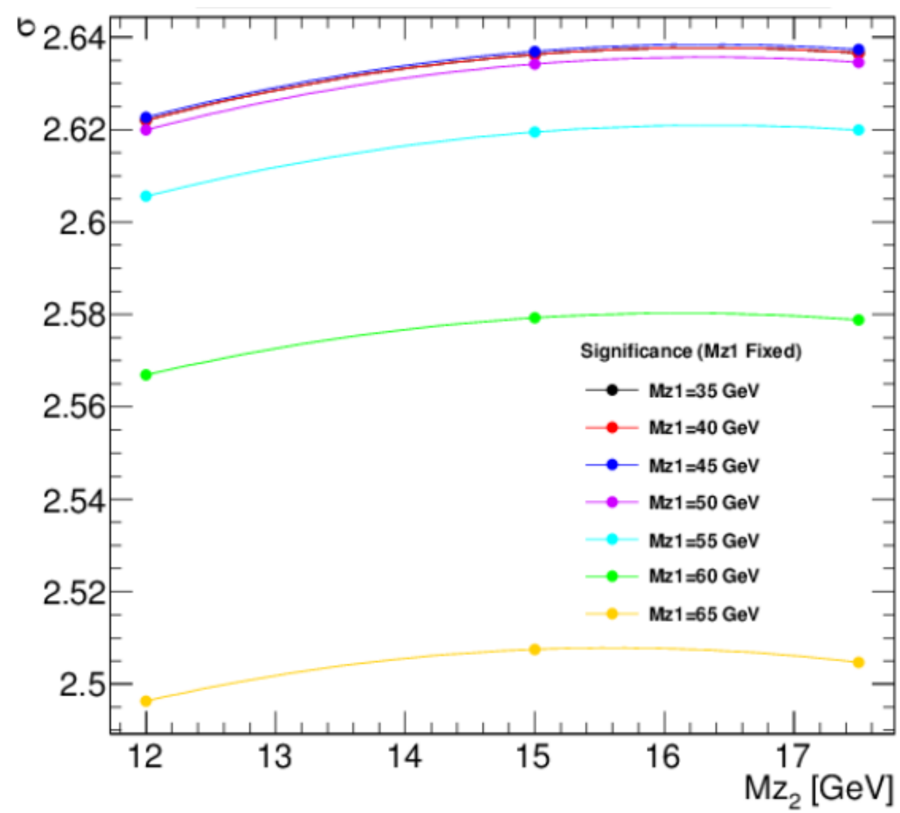
$$\frac{S}{\sqrt{S+B}}$$

- Or, in case of low statistics, using the likelihood ratio as test statistics:

$$\sqrt{2(S+B)\ln\left(1+\frac{S}{B}\right)-2S}$$

Optimizing the cuts vs the significance

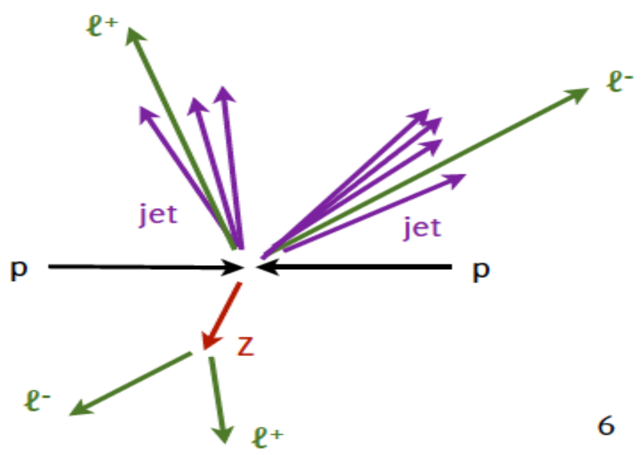




Significance vs the cuts on MZ1 and MZ2

Reducible backgrounds

- In this case, the final state is not exactly as your signal, but it has characteristics that can fake the signal final state
- In the 4-lepton example the main example of such background is the Z +jets process
- Cuts on the 3rd and 4th leading leptons are used for the rejection of these backgrounds
  - Z+light jets: the additional jets can fake electrons
    - Main handle to reject is an optimal electron identification
  - Z+bb: b leptonic decays produce leptons in the final state
    - Main rejection handle are the characteristics of leptons in jets from heavy quarks with long lifetime
      - Isolation
      - High impact parameter
- Processes with very large σ w.r.t. the signal, but easier to discriminate
  - E.g. Z inclusive σ is ~1 nb, i.e. ~2*10^5 times the signal cross section !

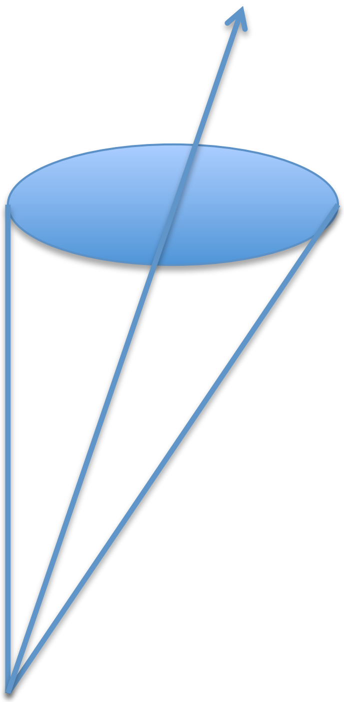


Lepton isolation

- Two possible ways of calculating it:
  - Sum of the tracks pT in a cone around the lepton track (track isolation)
  - Sum of the calorimeter cells energy in a cone around the lepton track
- The two variables are correlated but can be used in a complementary way
- The cone is defined as:

$$\Delta R = \sqrt{(\Delta\eta)^2 + (\Delta\varphi)^2} < 0.2$$

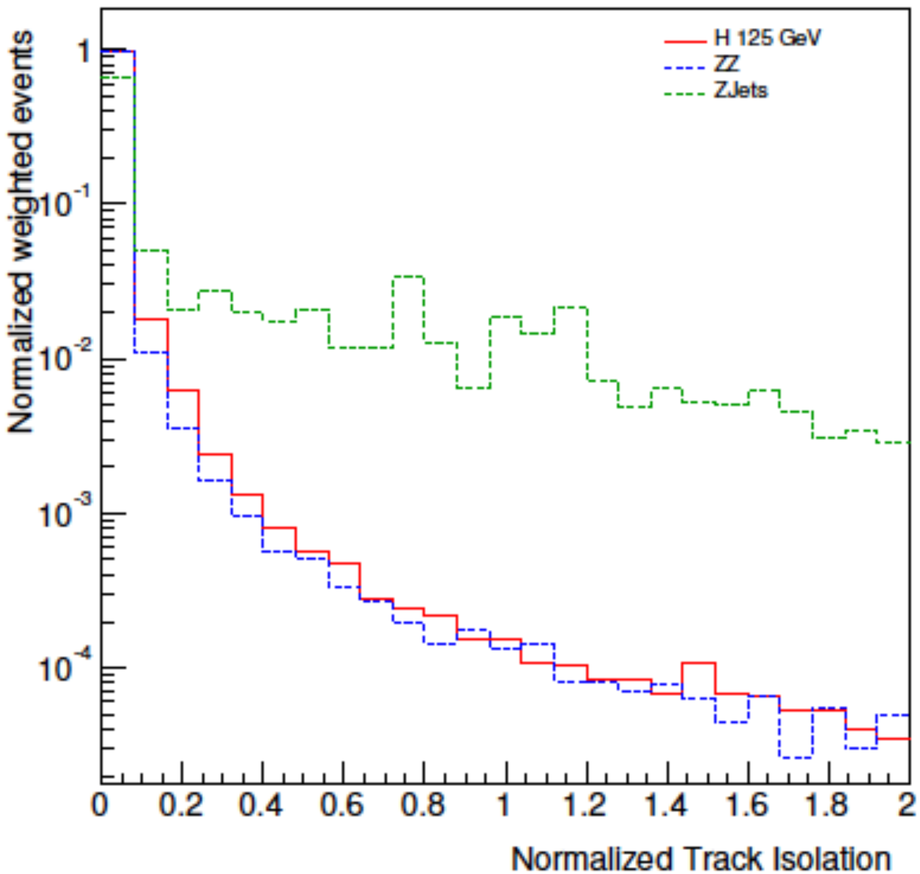
- The size is optimized on the basis of efficiency and rejection
  - An important component is the impact of pileup events



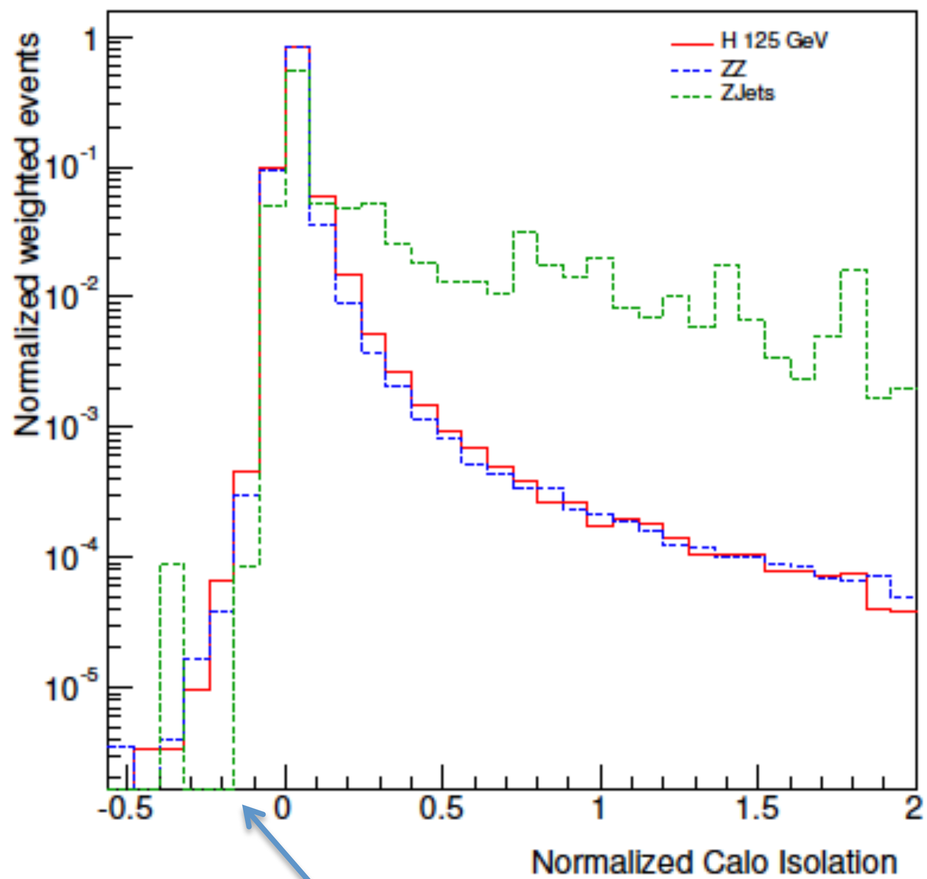
Lepton isolation

- In both cases the isolation is normalized to the lepton p_T

Track isolation



Calorimeter isolation

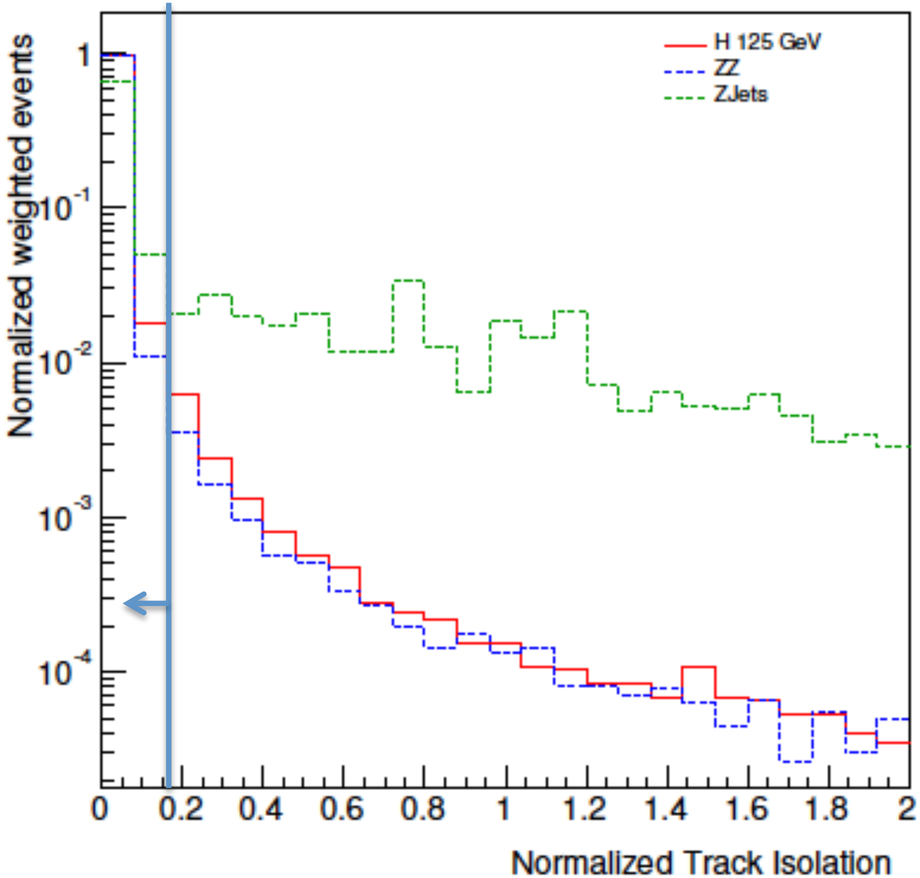


Noise subtraction

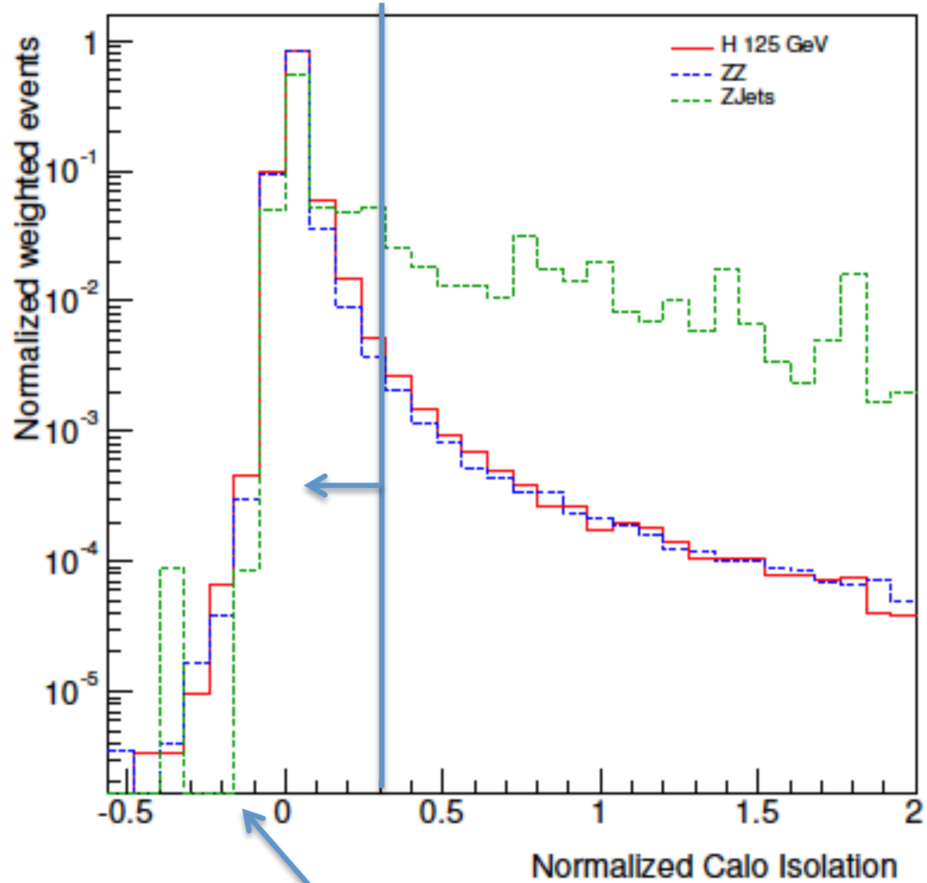
Lepton isolation

- In both cases the isolation is normalized to the lepton p_T

Track isolation

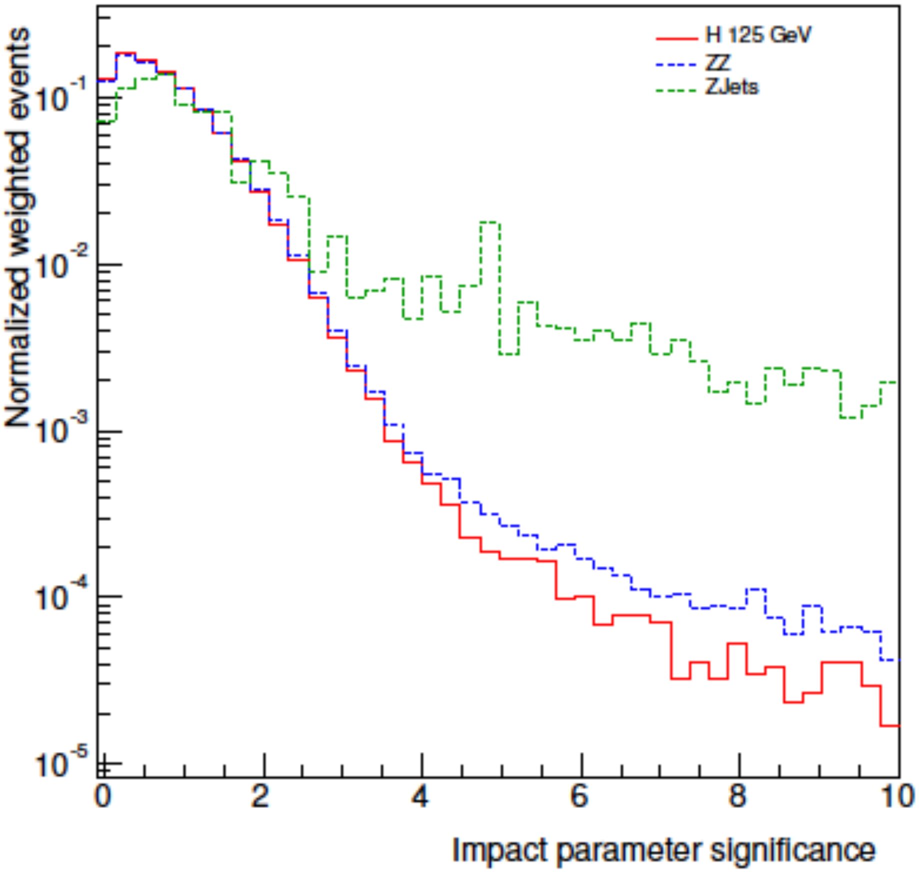


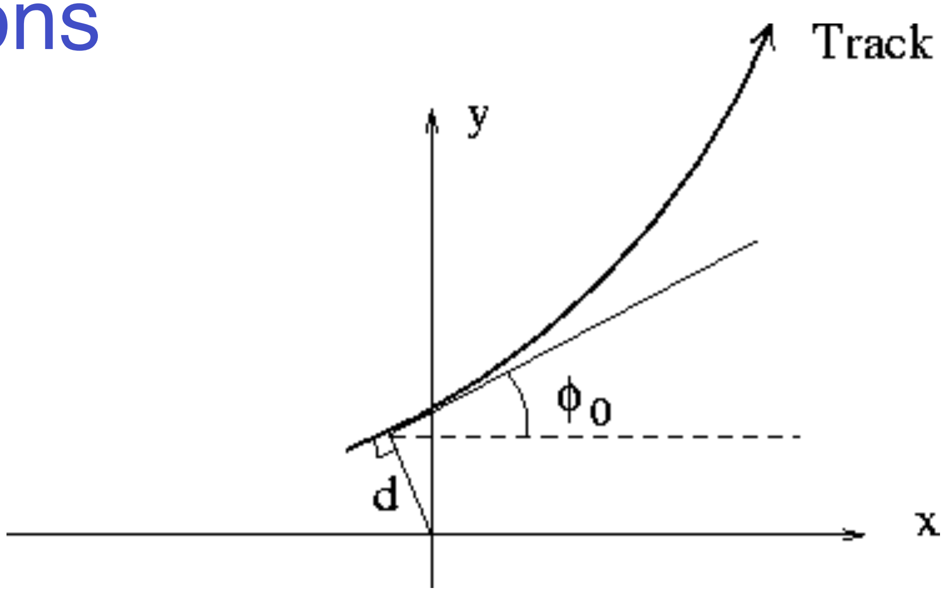
Calorimeter isolation



Noise subtraction

Impact Parameter: muons





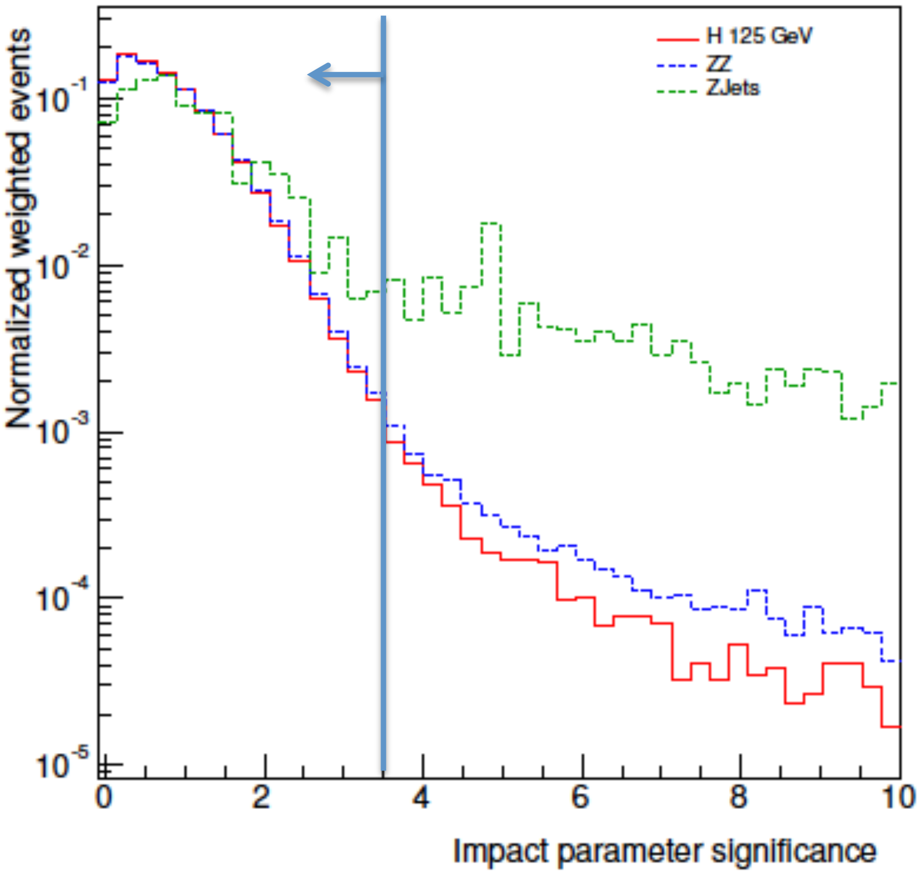
Long B mesons lifetime t~1.5 ps, i.e. ~750 μm at 10 GeV p_T

Transverse IP is used because of the much better resolution

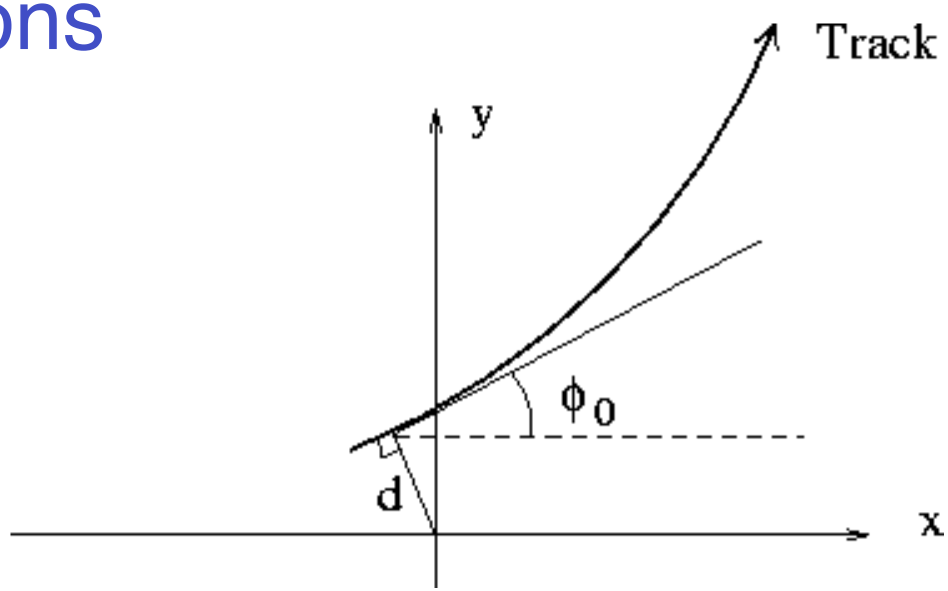
Large IP of the sub-leading leptons can be used as a discriminating variable

IP significance = d/σ_d

Impact Parameter: muons



IP significance = d/σ_d

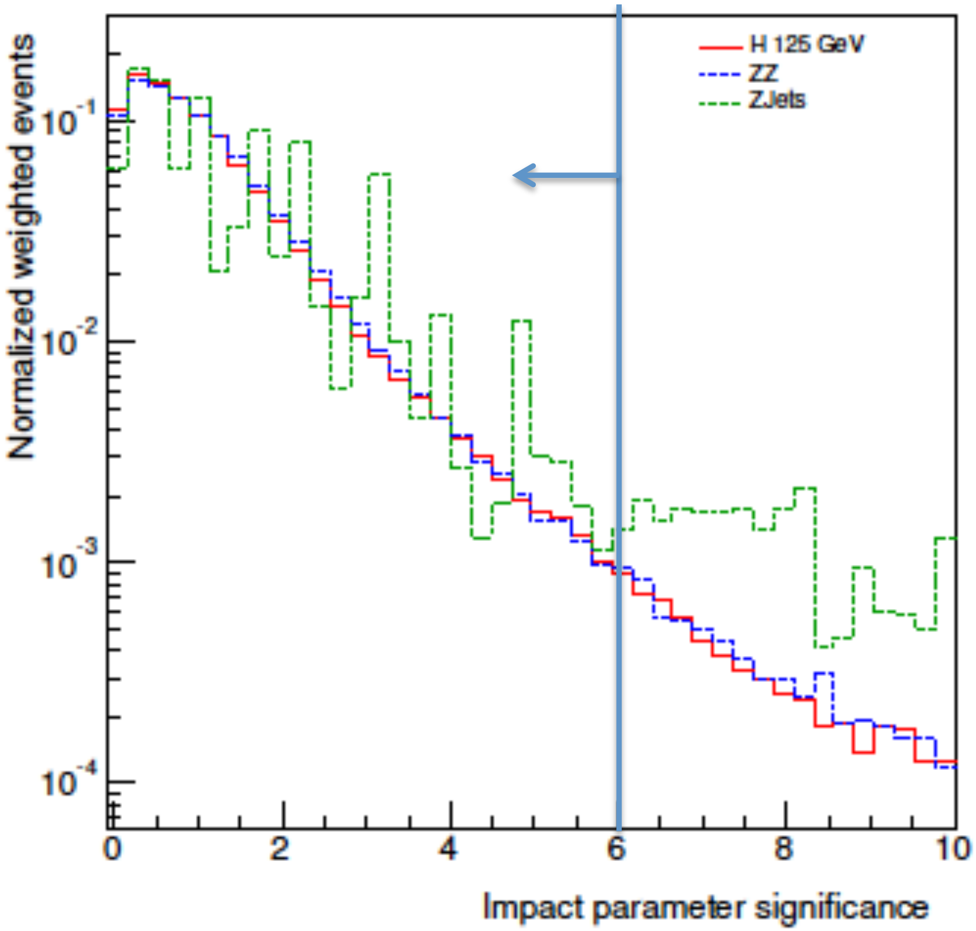


Long B mesons lifetime t~1.5 ps, i.e. ~750 μm at 10 GeV p_T

Transverse IP is used because of the much better resolution

Large IP of the sub-leading leptons can be used as a discriminating variable

Impact parameter: electrons



Because of the bremsstrahlung, the performance on IP reconstruction is much worse for electrons

It can still be used, but the cut has to be re-tuned to a looser value
→ keep a high enough efficiency

Derive backgrounds from data

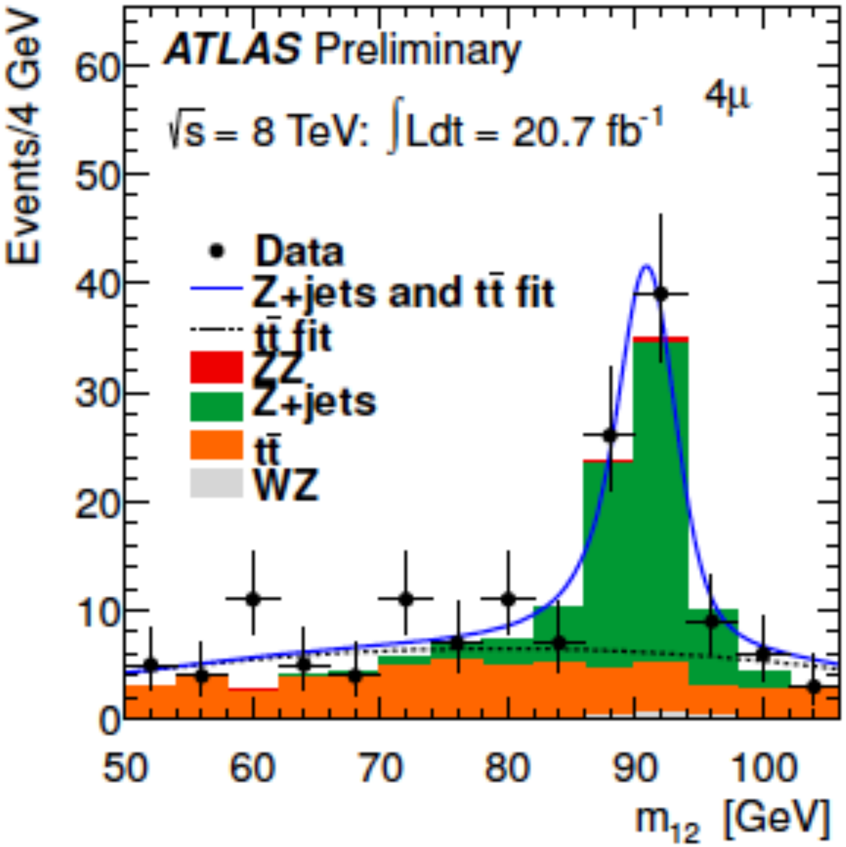
- In some cases, the cross section of a background is so large, that is not possible to generate enough MC events to determine its characteristics precisely

- E.g. for Z+jets one would need ~$2*10^5$ the signal statistics

- To derive signal normalization and/or shapes, and in general to cross check all backgrounds, the so-called Control Regions (CR) are built, by removing, or reverting some of the selection cuts
  - Create background-enhanced CRs
- The MC can then be used to extrapolate from the CR to the Signal Region (SR) via some transfer factors or functions
  - But in some cases just the data are used, and also the transfer functions are derived from the data
  - Various methods that we'll not have time to cover in detail

Example of a background CR

- Obtained by removing/reverting the cuts on isolation and IP
- Signal and irreducible background are in this way completely removed
- An almost pure Z+jets + ttbar (another irreducible background that we didn't cover here) sample is obtained



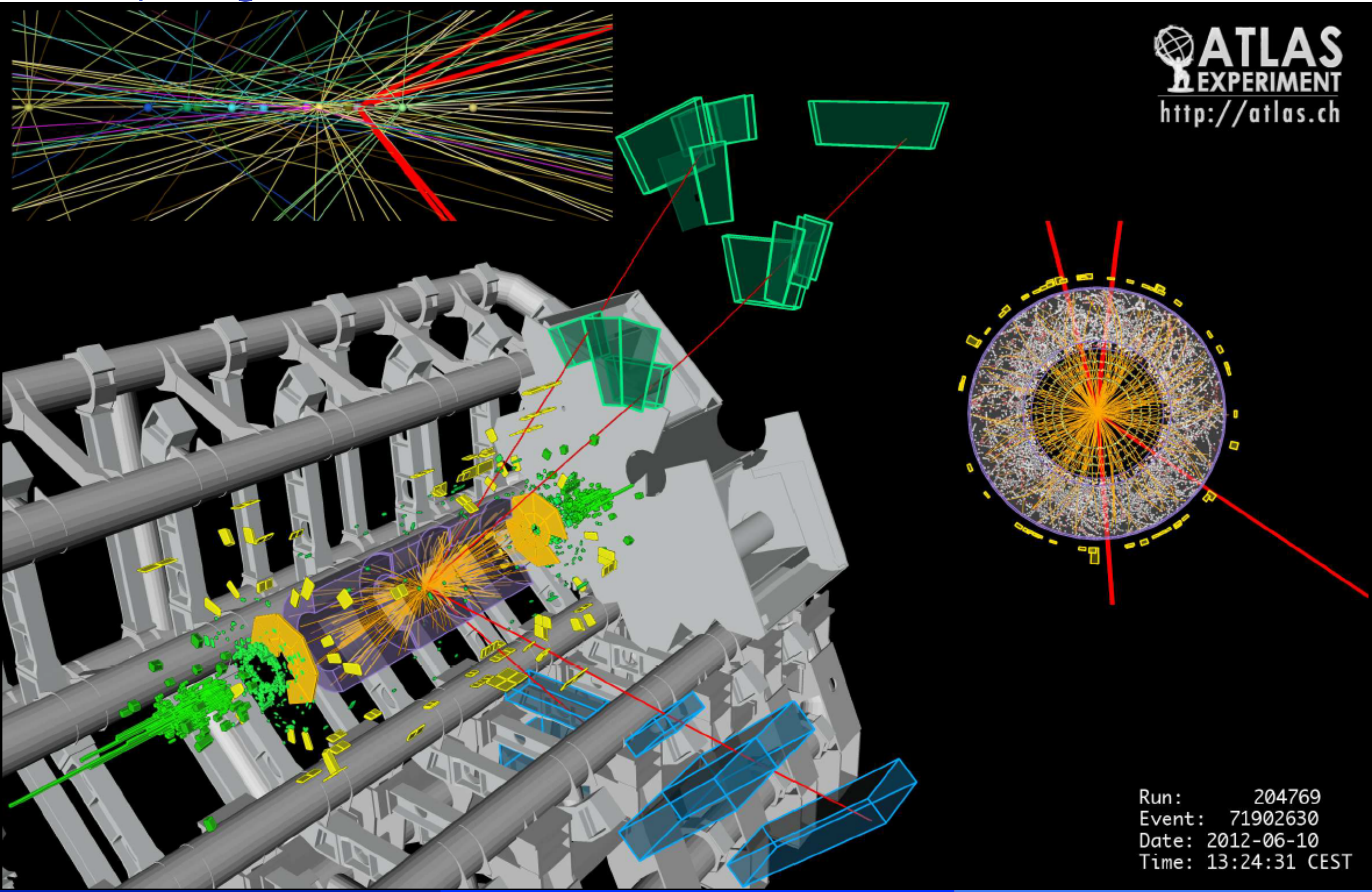
By fitting the data with two template functions:

→Breit-Wigner + Gaussian (Z-peak from the Z+jets)

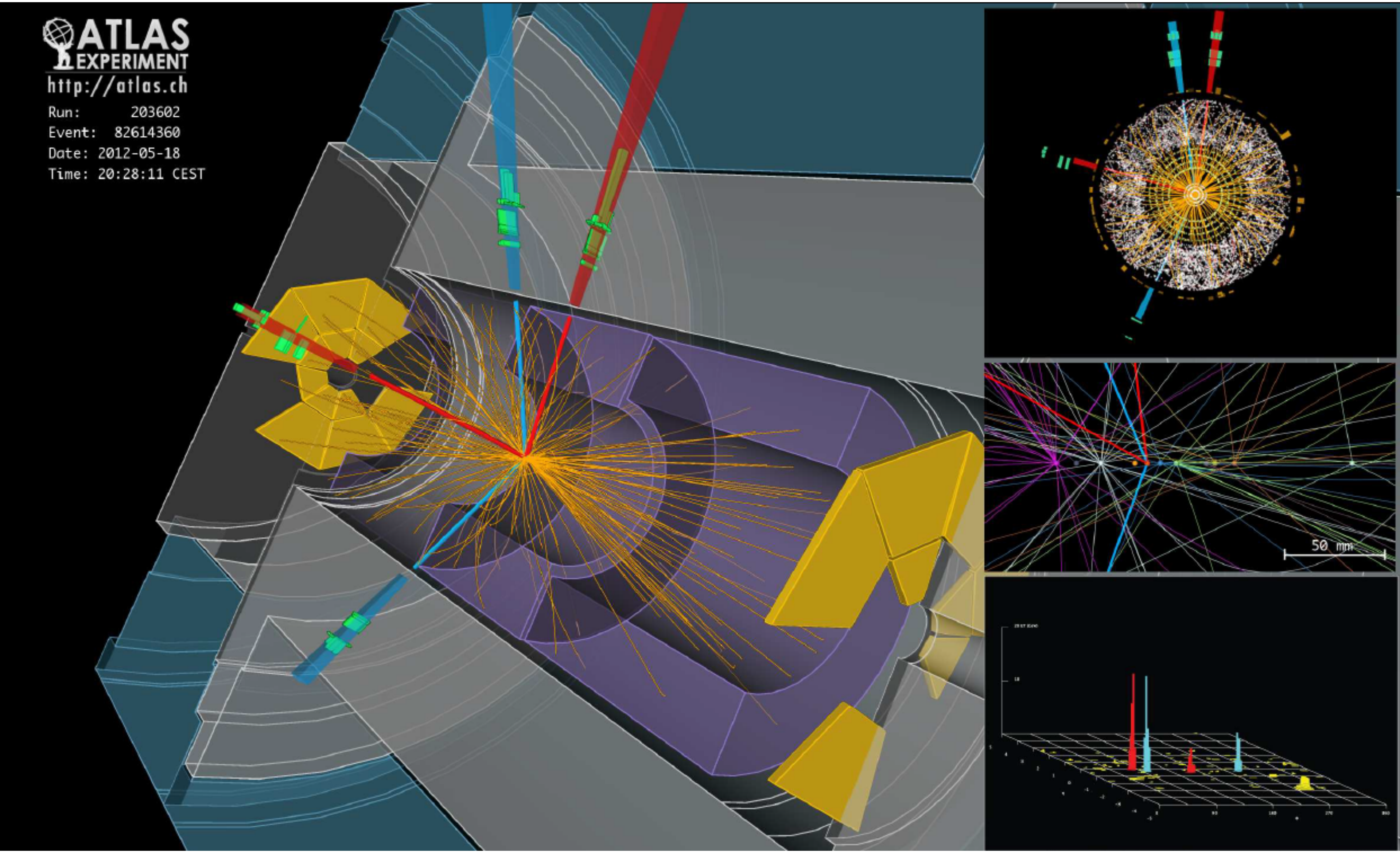
→Polynomial (~flat ttbar component)

The two contributions can be disentangled and each MC separately rescaled to the data

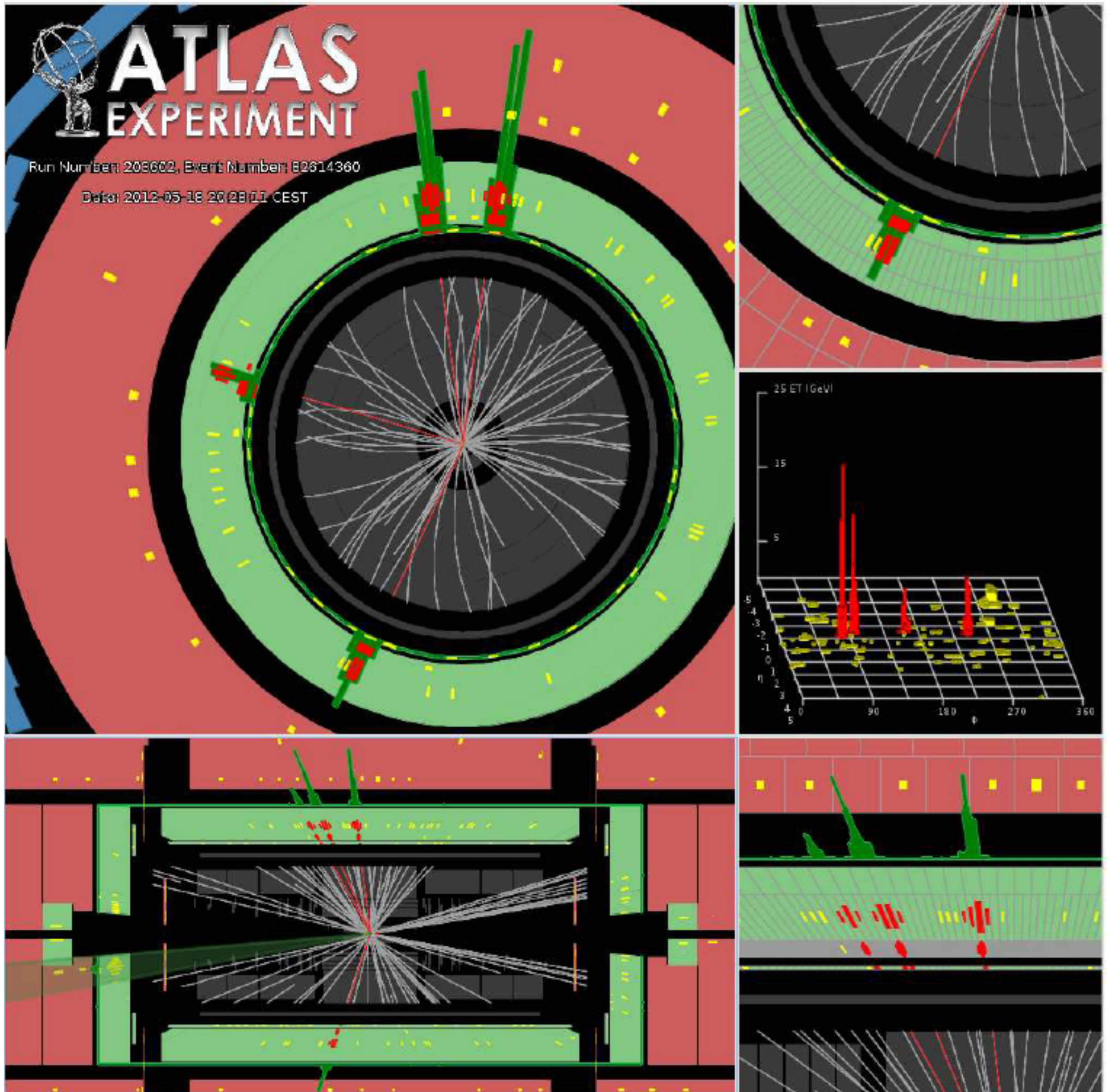
A 4μ signal candidate



A 4e signal candidate



Another view

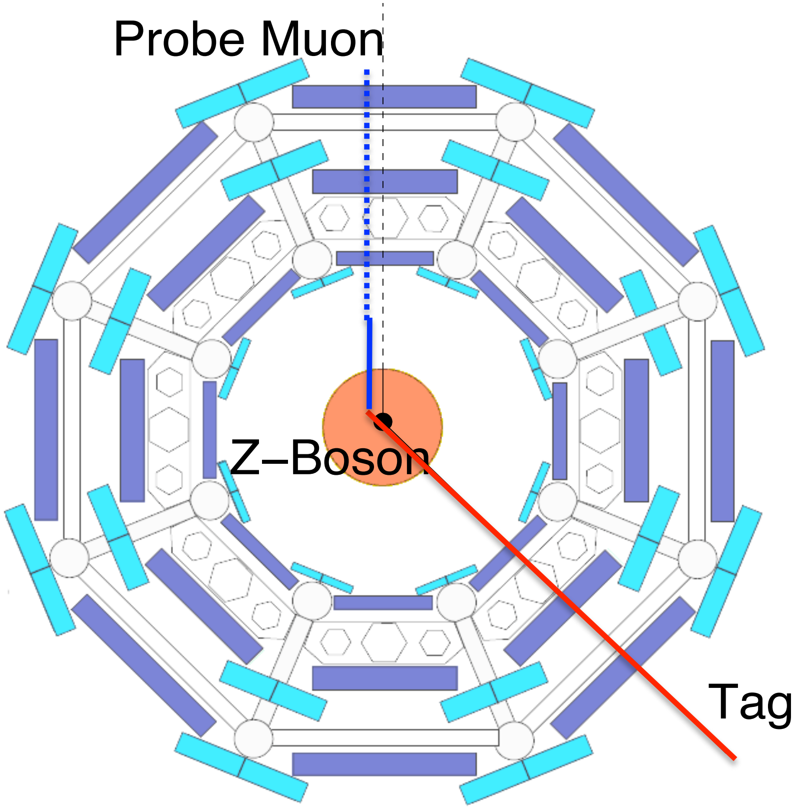


Modeling of the detector performance

- The MC models the detector performance with great accuracy
- Still some effects might not be perfectly modeled and will need some data-driven tuning
  - Need to make sure that the MC models data correctly, before claiming that any difference comes from new physics
- Fine-tuning of small local effects, or time-dependent detector issues:
  - E.g. if one part of the detector becomes inefficient for a limited data taking interval
- Typical effects to take into account are:
  - Trigger and reconstruction efficiencies
  - Momentum and energy resolutions
  - Momentum and energy scales
- This is a very important ingredient in all Physics analyses and implies:
  - Develop methods to measure data-driven performances
  - Apply corrections to the MC or to the data

Efficiencies

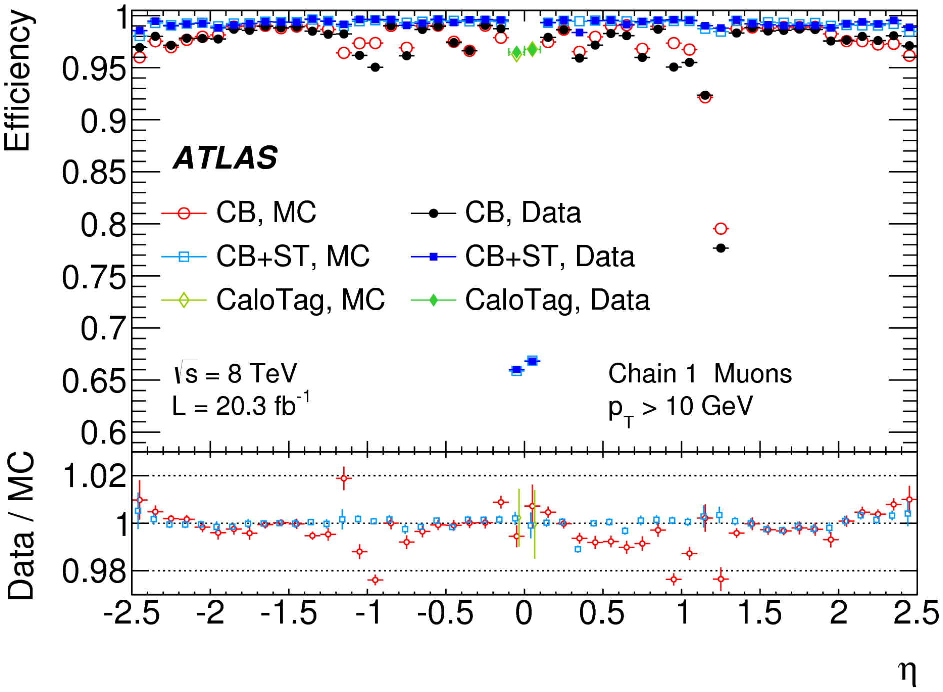
- Use known process/resonances decaying to electrons or muons
  - $Z \rightarrow ll$ (pT>~10 GeV) and $J/\psi \rightarrow ll$ (pT<~10 GeV)
- Look for a reconstructed muon, which also provided the trigger (tag muon)
- Look for a track in the Inner Detector (probe):
  - Same vertex as the tag
  - Isolated (reject tracks in jets)
  - M(tag,probe) close to M_Z or $M_{J/\psi}$



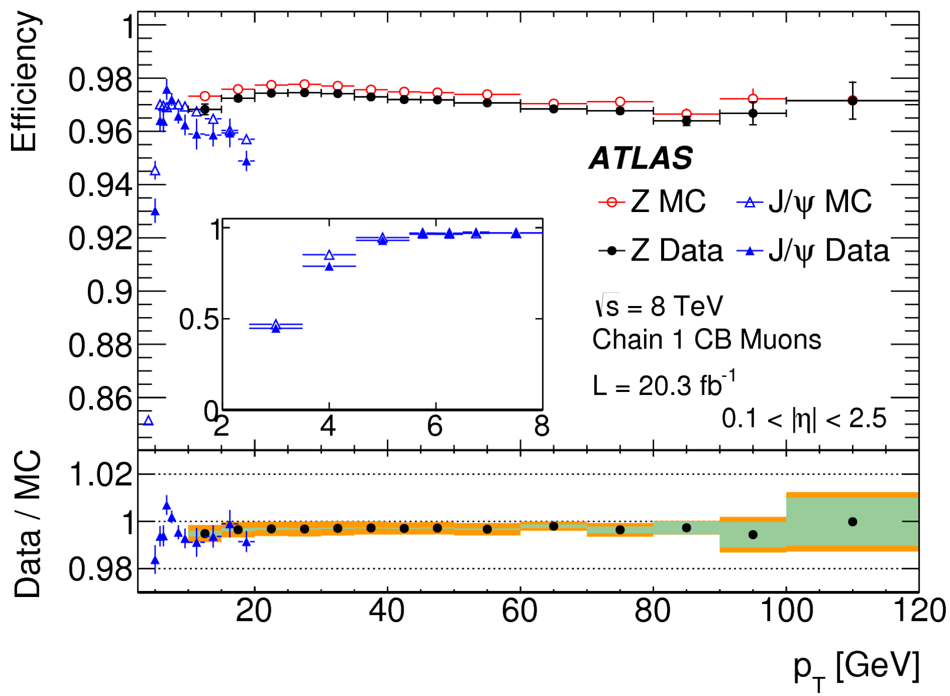
- Check if the probe ID track matches a reconstructed muon (or a muon trigger element)
  - The efficiency is given by $N_{matching}/N_{total}$ probes
- Same method used also for electrons (look for matching tracks/calo clusters)

Some example results for muons

Efficiencies vs η for various types of reconstructed muons

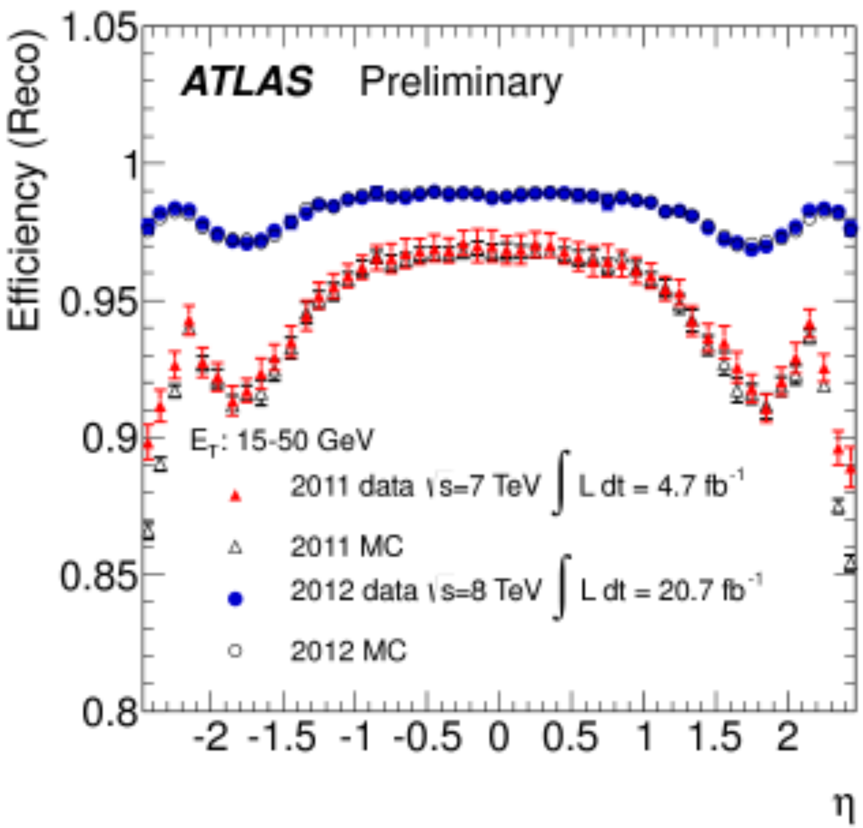


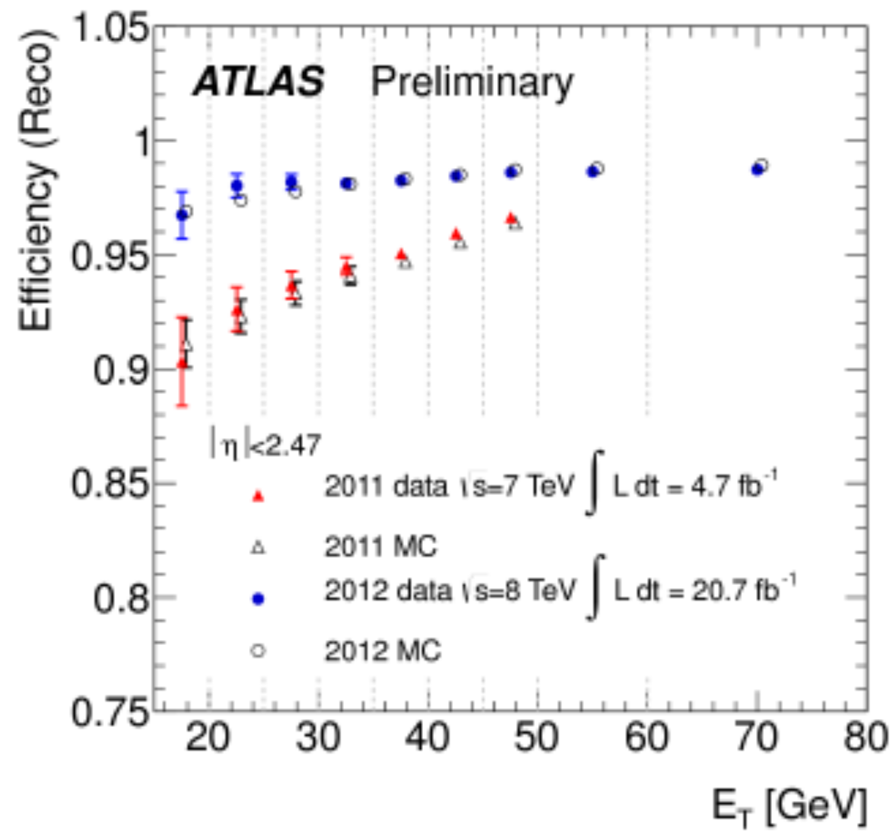
Efficiencies vs p_T for combined muons



Local Efficiencies are calculated for single muons in bins of p_T, η, ϕ, chosen with the proper granularity, depending on detector structure, or known inefficiency regions

Efficiencies corrections for electrons





Also in this case, the efficiencies are calculated in bins of the phase-space
Large improvement in 2012 reconstruction

Efficiency reweighting

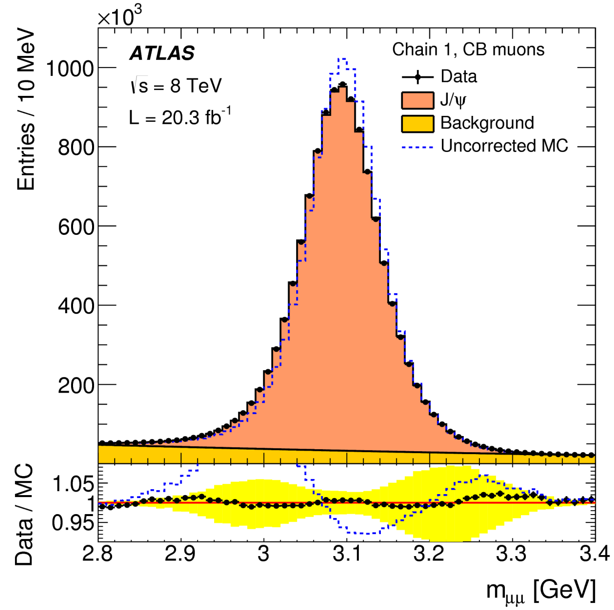
- Efficiency is corrected by giving to each event a weight
- Get for each bin:
  - ε_i^{data}: efficiency in data
  - ε_i^{MC}: efficiency in MC
- In 4-lepton events:
- Reconstruction efficiency
  - Reweight each MC event according to: $$w = \prod_{i=1,n} \frac{\varepsilon_i^{DATA}}{\varepsilon_i^{MC}}$$
  - Where the product runs over all reconstructed leptons entering in the analysis
- Trigger efficiency: just need at least one lepton triggering
  - Reweight each MC event according to:

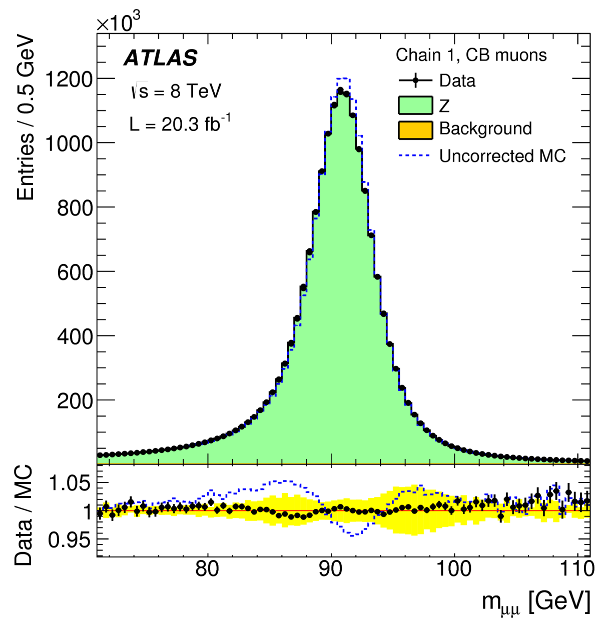
$$w = \frac{1 - \prod_{i=1,N} (1 - \varepsilon_i^{DATA})}{1 - \prod (1 - \varepsilon_i^{MC})}$$

Ratio of the probabilities that at least one lepton is passing the trigger

Momentum scales and resolutions

- Scales of momentum and energy are determined by comparing to the MC distributions of known resonances
  - Z, J/ψ , Υ
- The very well known masses and widths allow the precise determination of momentum and energy scales
  - Same method used for muons and electrons
  - Just the level of backgrounds is different
- Scales are determined in bins of p_T, η, ϕ
- Systematic uncertainties on scales are a fundamental ingredient in the mass measurement





Section 4

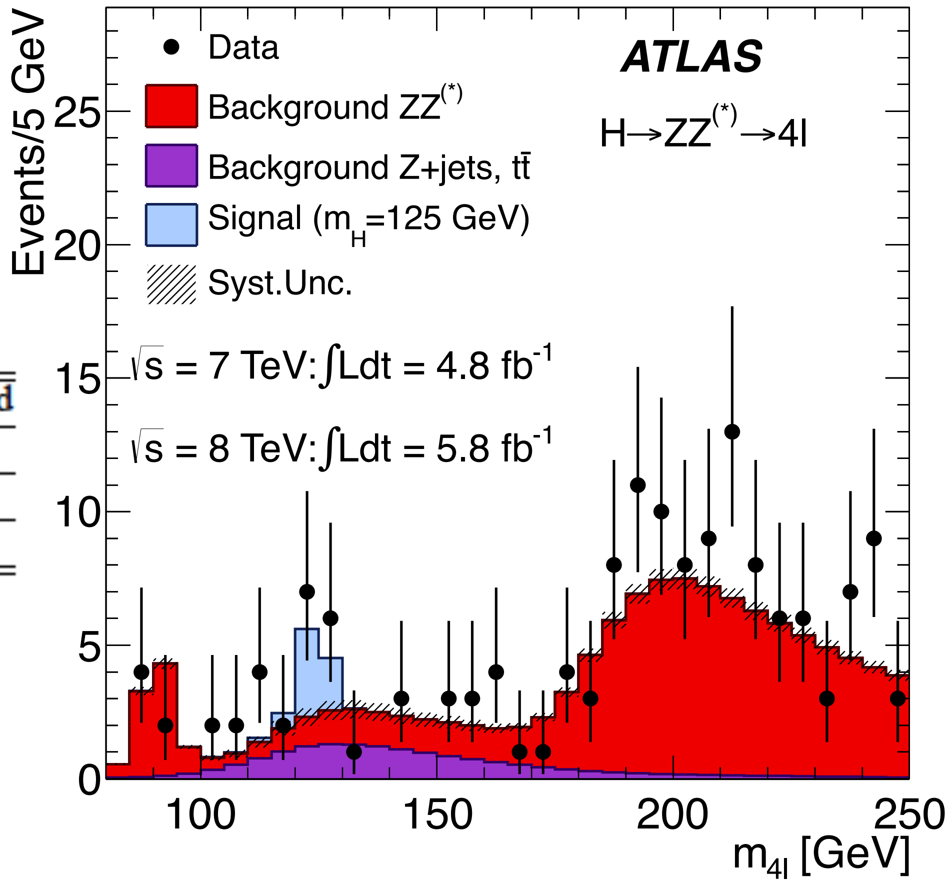
ANALYSIS RESULTS

Results of July 2012 – the discovery !

Number of events in a mass window 120-130 GeV around the signal peak

Theory error on the σ

	Signal	$ZZ^{(*)}$	$Z+\text{jets}, t\bar{t}$	Observed
4μ	2.09±0.30	1.12±0.05	0.13±0.04	6
$2e2\mu/2\mu2e$	2.29± 0.33	0.80±0.05	1.27±0.19	5
$4e$	0.90±0.14	0.44±0.04	1.09±0.20	2

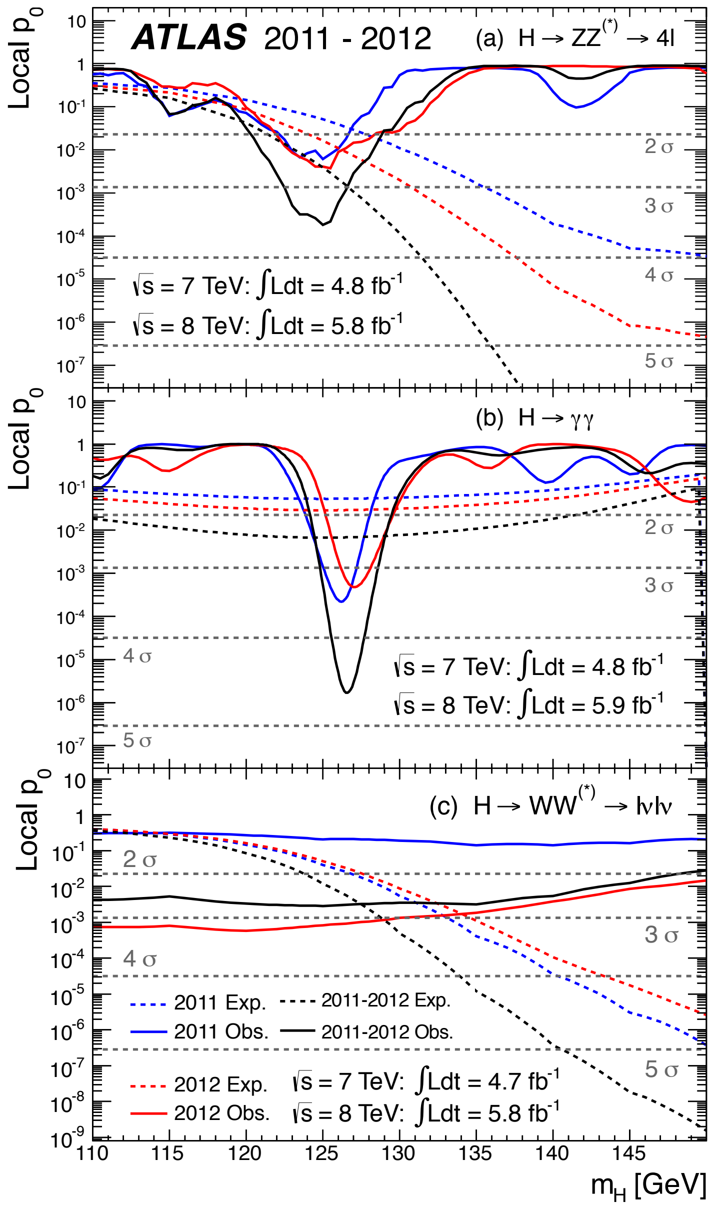


Total expected signal: 5.3 events
Total expected background: 4.9 events

Total observed events in data: 13 !

Already a simple calculation shows you something: probability to observe 13 events when you expect 4.9 is 0.17 % (of course no errors no syst here)

Significance of the first observation



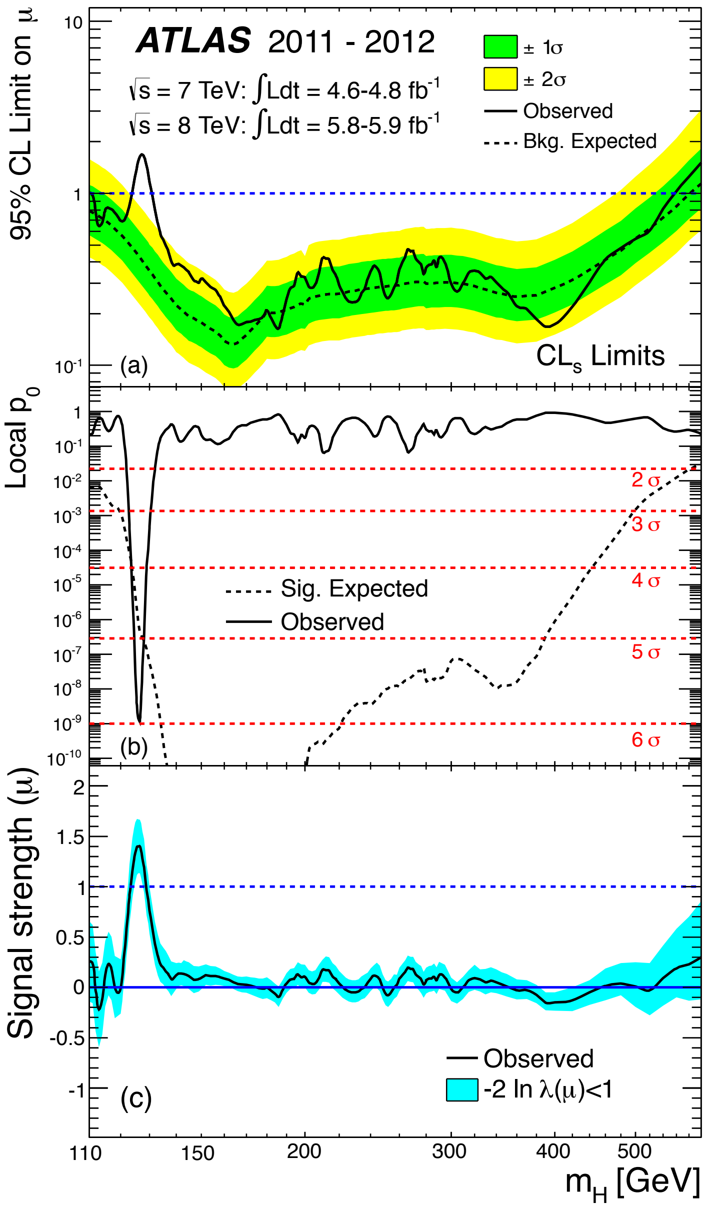
- Test of the background only hypothesis using the test statistics that you've seen the lectures:

$$q_0 = -2 \ln \frac{L(0, \hat{\hat{\theta}})}{L(\hat{\mu}, \hat{\theta})}$$

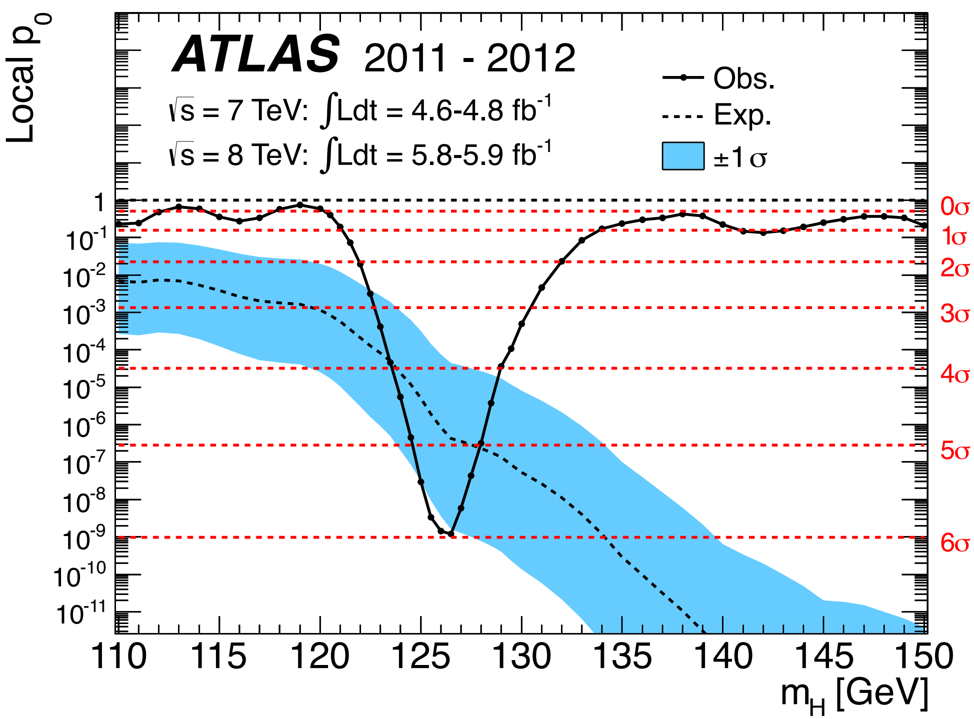
- The p0 is the probability that a background fluctuation is more signal-like than expected for the signal (or than the data for the observed)

- The expected curves correspond to the p0 vs mass in case of a SM signal

Combined limits and significance

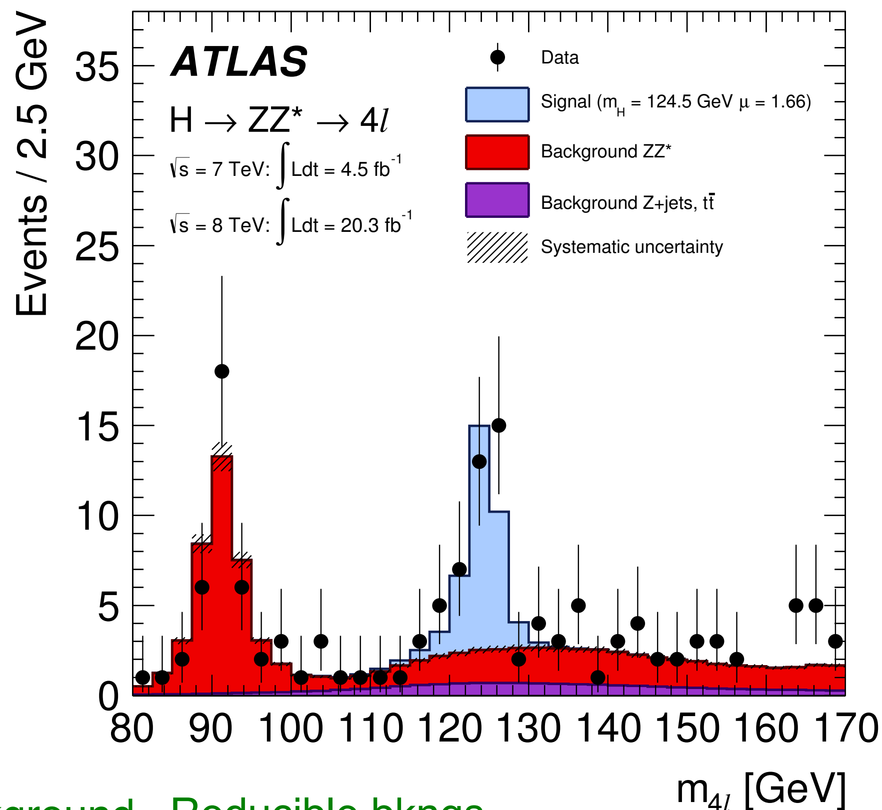


Combined p_0 in the low-mass region



The new resonance, with the full run1 dataset

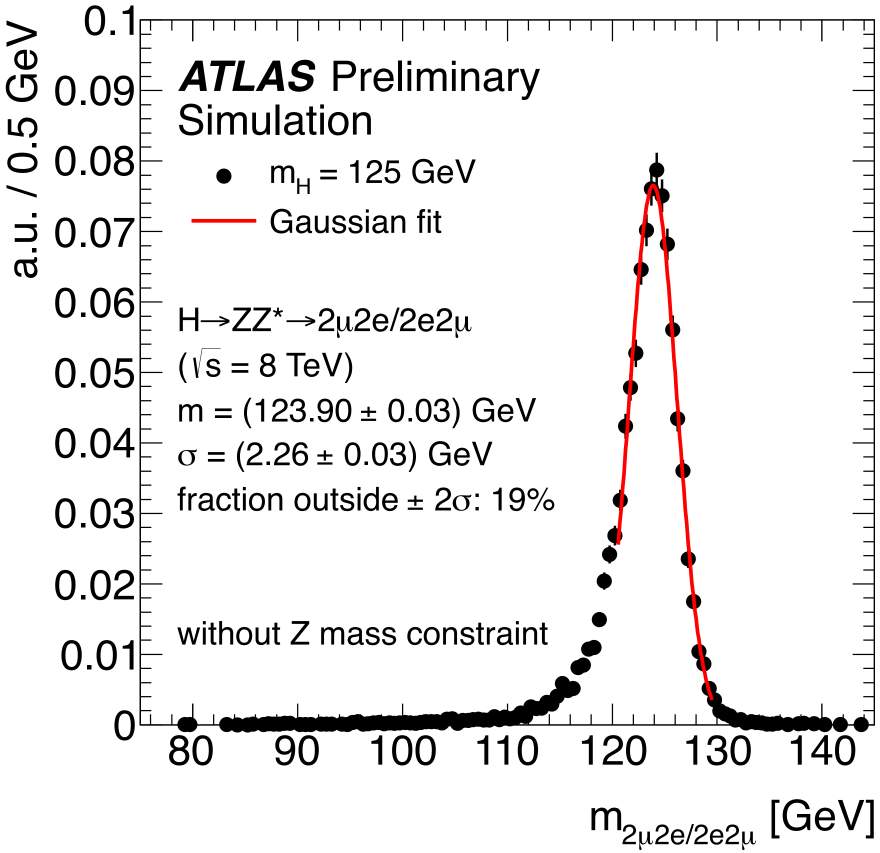
- At the end of the run-1 data taking (end of 2012)
  - $4.5\ fb^{-1}$ at 7 TeV and $20.3\ fb^{-1}$ at 8 TeV
- In the 120-130 GeV mass window:
  37 observed events with
  10.4 expected from background only (well above
  5-sigma significance)
  - Light excess w.r.t. expected SM signal

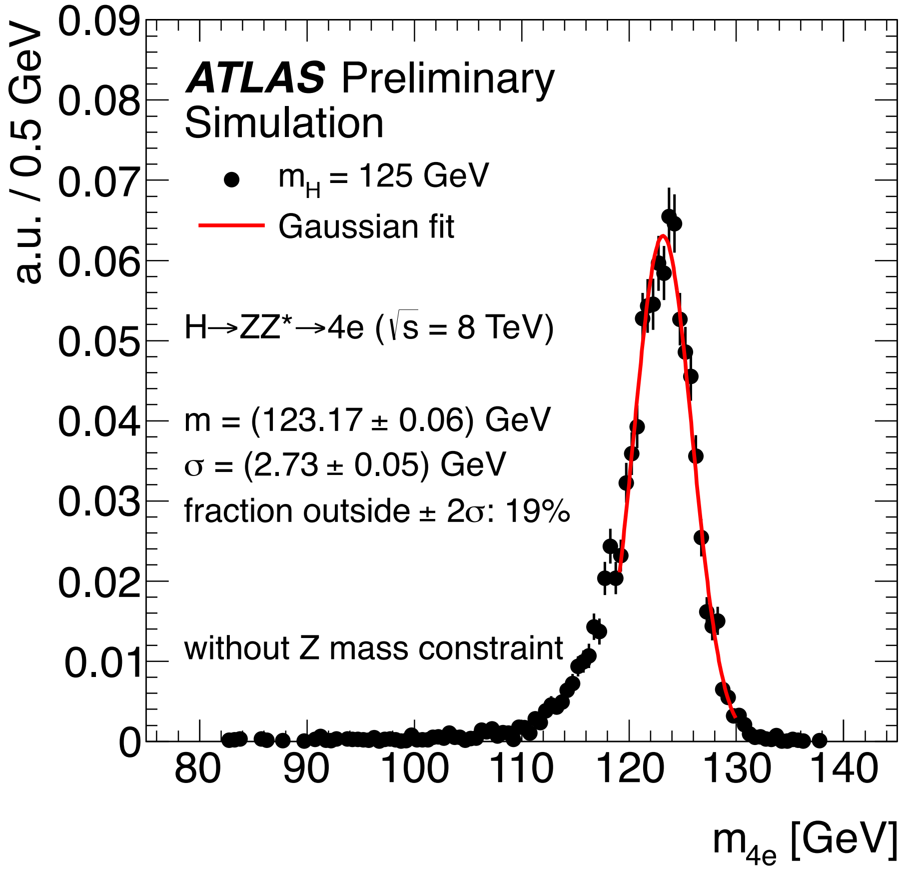


$\sqrt{s} = 7$ TeV and $\sqrt{s} = 8$ TeV

	Total signal	Signal 120-130	ZZ background	Reducible bkngs			
4μ	6.80 ± 0.67	6.20 ± 0.61	2.82 ± 0.14	0.79 ± 0.13	1.7	9.81 ± 0.64	14
$2e2\mu$	4.58 ± 0.45	4.04 ± 0.40	1.99 ± 0.10	0.69 ± 0.11	1.5	6.72 ± 0.42	9
$2\mu 2e$	3.56 ± 0.36	3.15 ± 0.32	1.38 ± 0.08	0.72 ± 0.12	1.5	5.24 ± 0.35	6
$4e$	3.25 ± 0.34	2.77 ± 0.29	1.22 ± 0.08	0.76 ± 0.11	1.4	4.75 ± 0.32	8
Total	18.2 ± 1.8	16.2 ± 1.6	7.41 ± 0.40	2.95 ± 0.33	1.6	26.5 ± 1.7	37

Mass resolution





Natural width @ 125 GeV is 4 MeV

→ The reconstructed peak width is completely determined by the experimental resolution

→ It's important to reduce it as much as possible (mass mesurement)

Z mass constraint

- The resolution can be improved by applying the so-called Z mass constraint
  - The signal decay must have the two leading leptons having an invariant mass close to an on-shell Z
  - Rescale the lepton momenta so that the invariant mass of the leading lepton pair corresponds to the Z mass
- Simplest way is to rescale the momenta minimizing a chi-sq and imposing a constraint to the Z mass

$$M_{ll}^2 = p1 \cdot p2 \cdot (1 - \cos\theta_{12})$$

$$M_{ll}^{2'} = k1 p1 \cdot k2 p2 \cdot (1 - \cos\theta_{12}) = M_Z^2$$

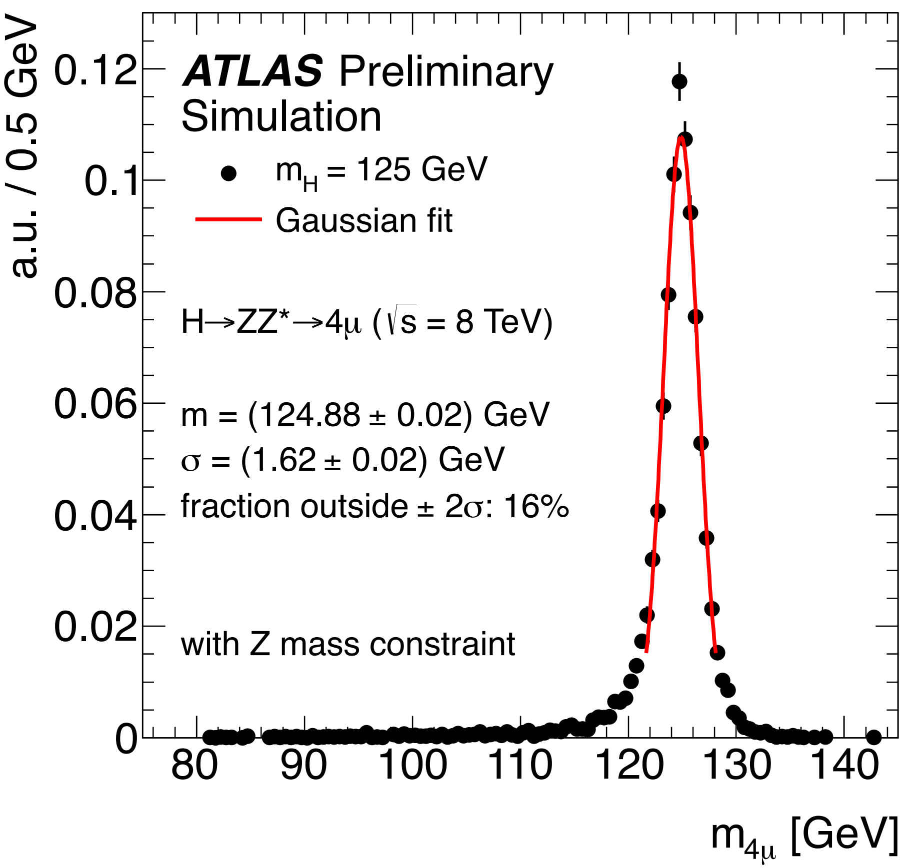
$$k1 \cdot k2 = M_Z^2 / M_{ll}^2$$

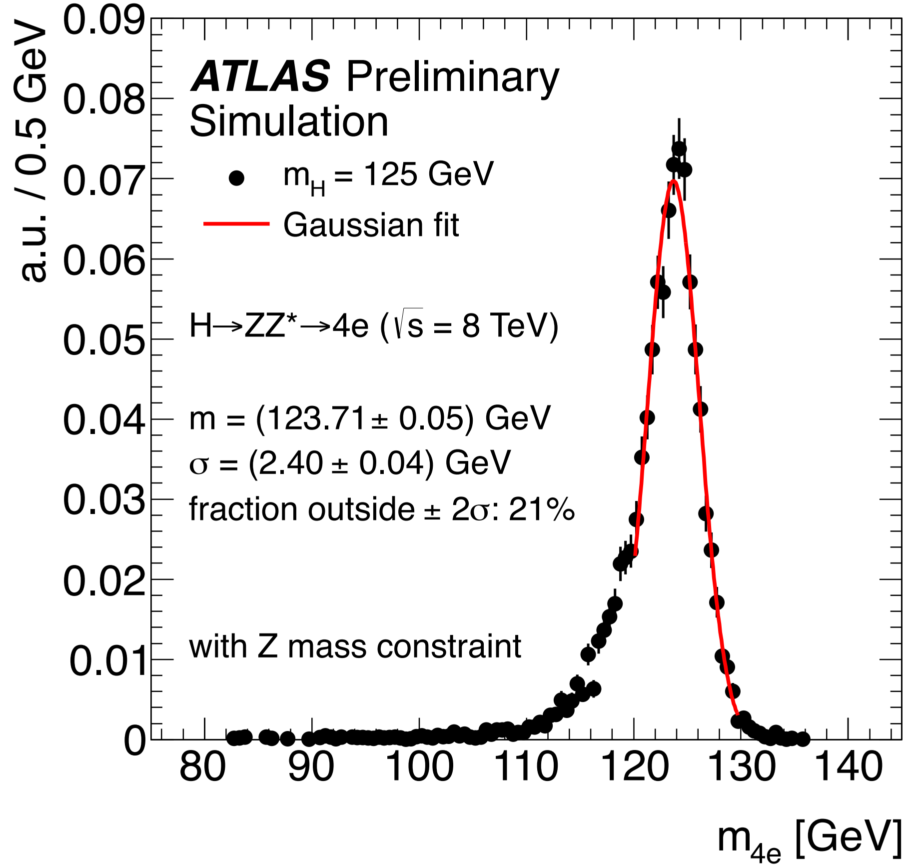
$$\chi^2 = \frac{(k1p1 - p1)^2}{\sigma_{p1}^2} + \frac{(k2p2 - p2)^2}{\sigma_{p2}^2}$$

Or, a more complex Constrained kinematic fit, taking into account also the Z natural width

Signal mass resolution with mass constraint

For M_H=125 GeV

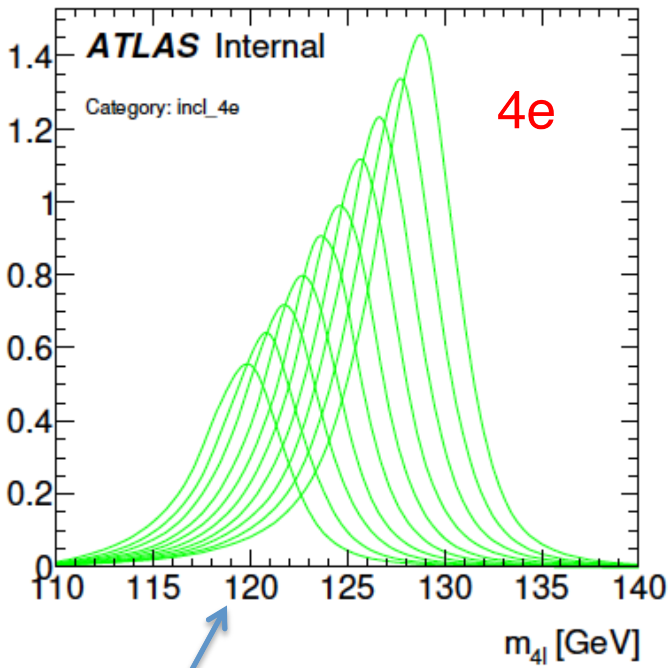


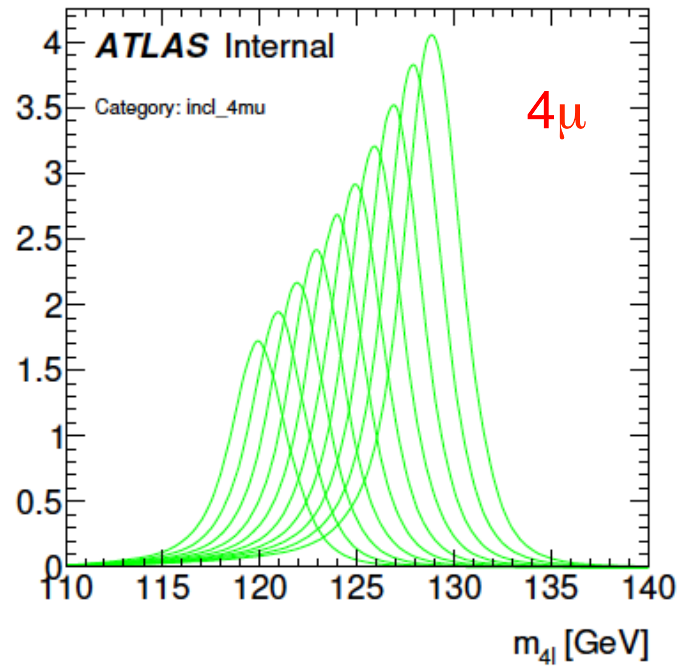


The Higgs mass is a free parameter of the theory → its precise measurement it's fundamental

Mass measurement

Template fit: buld a MC mode of the signal mass distribution
Continuos variation of the distribution vs M_H and σ_H via a morphing function (basically an interpolation among the fixed points)





Of course, only the signal depends on M_H

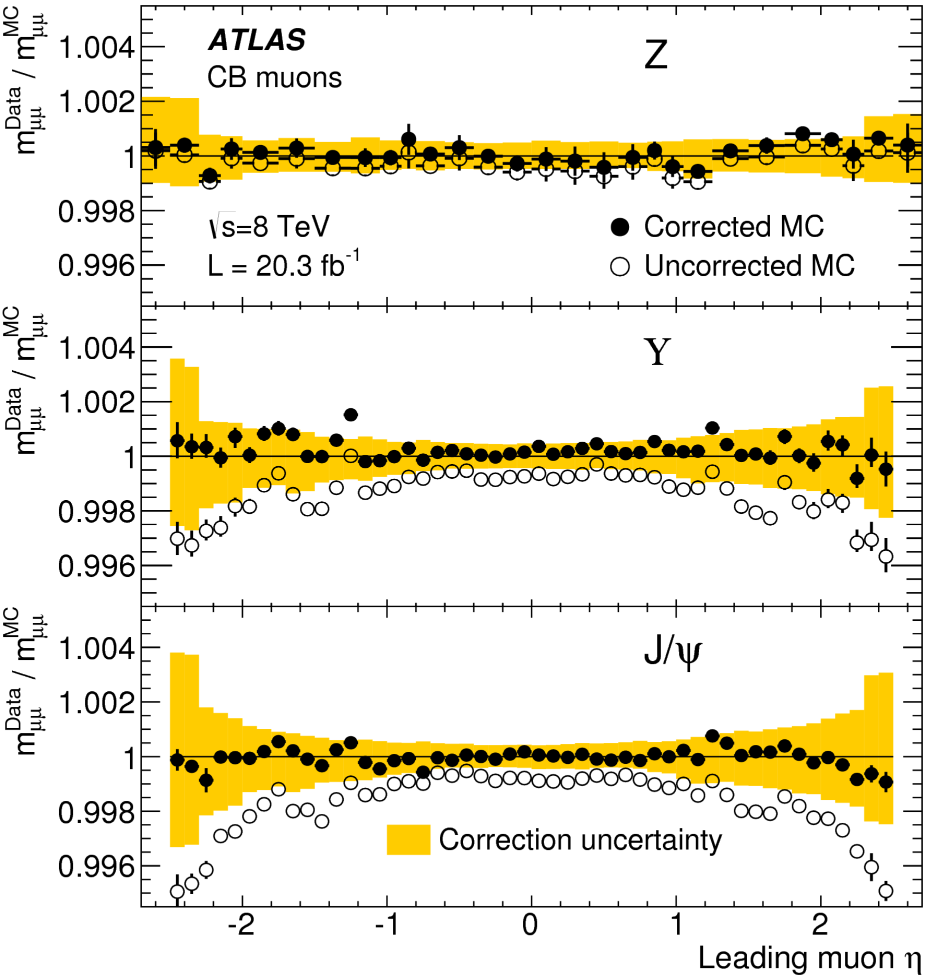
In the same fit, both mass and "signal strength" extracted: $$\mu = \frac{\sigma^{Signal}_{Observed}}{\sigma^{Signal}_{SM}} = \frac{N^{Signal}_{Observed}}{N^{Signal}_{SM}}$$

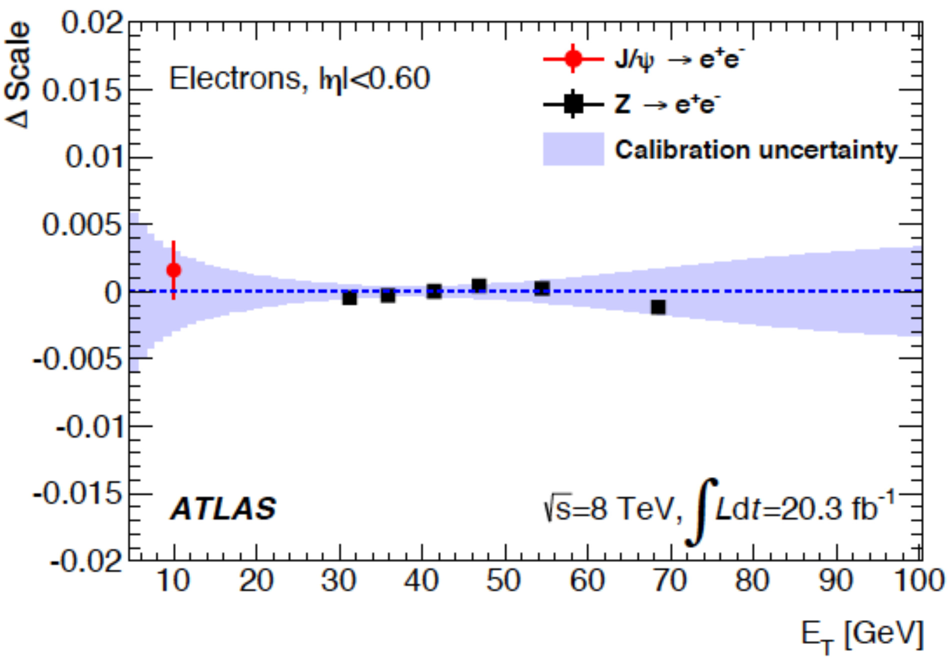
Systematics

Systematics on the energy scales are:

→ 0.04% - 1%   for electrons
→ 0.07% - 0.2% for muons

Thanks to the large stat of the resonance and the extensive work on calibrations

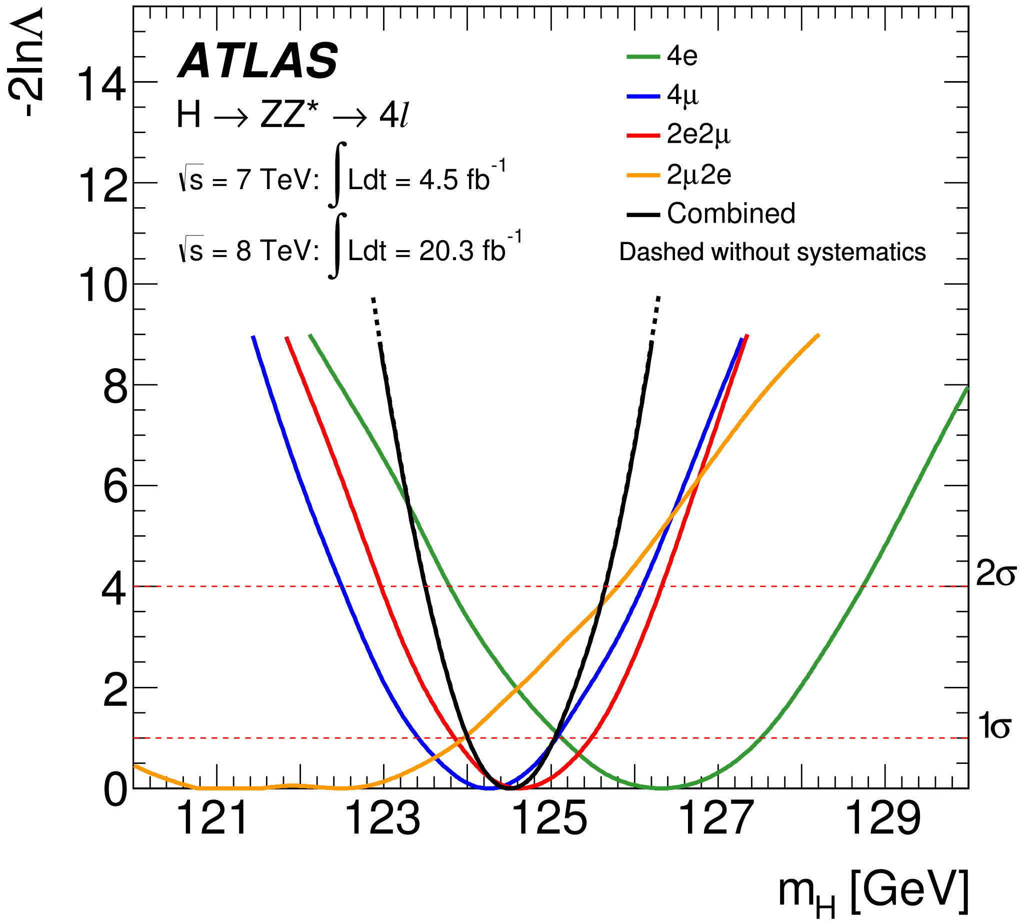




Statistical and systematic errors

- The error on a measurement has:
  - A statistical component, i.e. depending only on the number of events entering in the measurement
  - A systematic component, i.e. connected to the methods and assumptions used in the measurement (e.g. the energy scales, the errors on the background knowlegde)
- Due to the limited statistics the $H \rightarrow ZZ^* \rightarrow 4l$ channel is not strongly affected by systematics
- Mass measurementmain errors:
  - Momentum and energy scales
  - Background shapes
- Signal strenght measurement:
  - Backgrounds normalizations (and shape)
  - Theoretical error on the SM signal cross section

Mass fit results



M_H=124.51±0.52 (stat)±0.6 (syst)

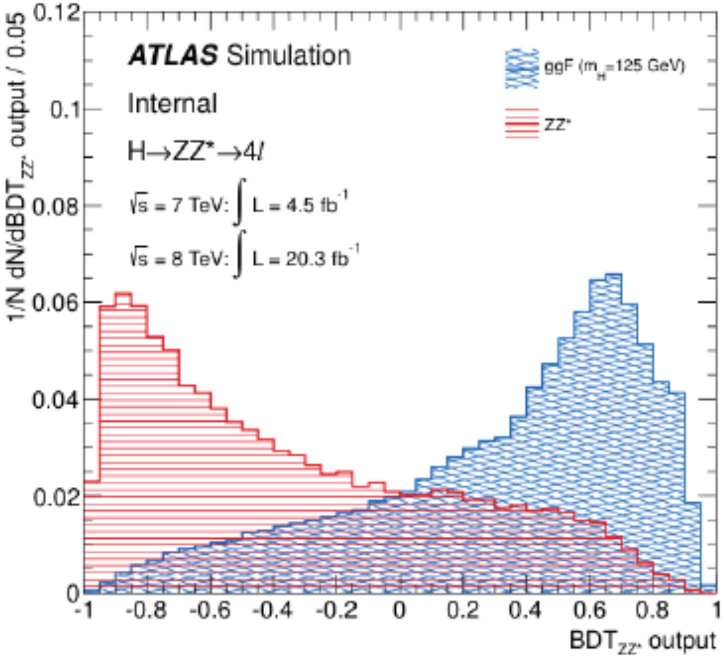
The impact of the scale syst. on the mass is negligible

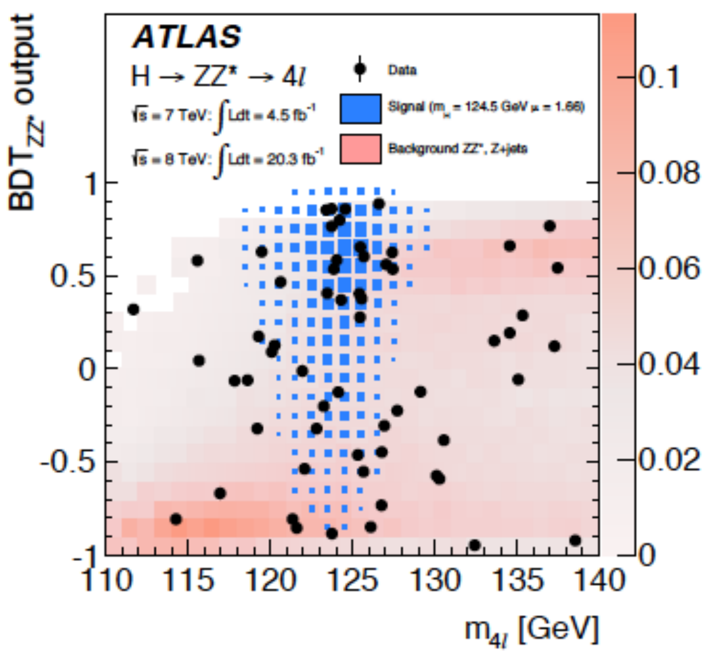
Signal strength compatible with the SM:
μ=1.44 +0.40/-0.33

In the case of m the impact of the systematics is larger mainly due to the theory uncertainties on the signal cross section

Improving the mass fit

- The fit of the mass peak that we just saw, is model-independent, i.e. does not make assumptions on the spin/parity quantum numbers of the resonance
  - The fit can be improved by assuming SM hypothesis $J^P=0^+$ for the decaying resonance
  - This hypothesis has been verified on data (see in the following)
- While, the ZZ background has a different composition of Z's polarization states (total J not forced to be 0)
  - This feature can be used to build a discriminant variable between the signal and the irreducible background, and include it as additional dimension in the fit





Section 5

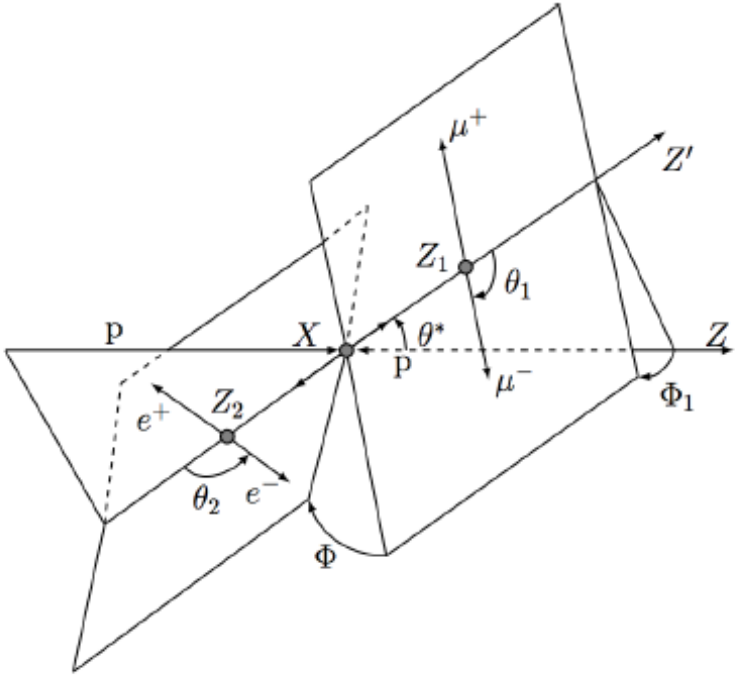
MULTI-VARIATE ANALYSES

Multivariate Analyses

- It is not always possible to place a cut on a discriminating variable
  - The level of discrimination doesn’t always allow to remove the background keeping a high efficiency
- In some cases, the discrimination is just in tiny shape differences, or even in the correlation among various variables
  - More sophisticated methods have to be applied
  - Multi-Variate Analyses are very powerful in this cases
  - We’ll see here just an example (the Boosted Decision Tree) of many different methods available

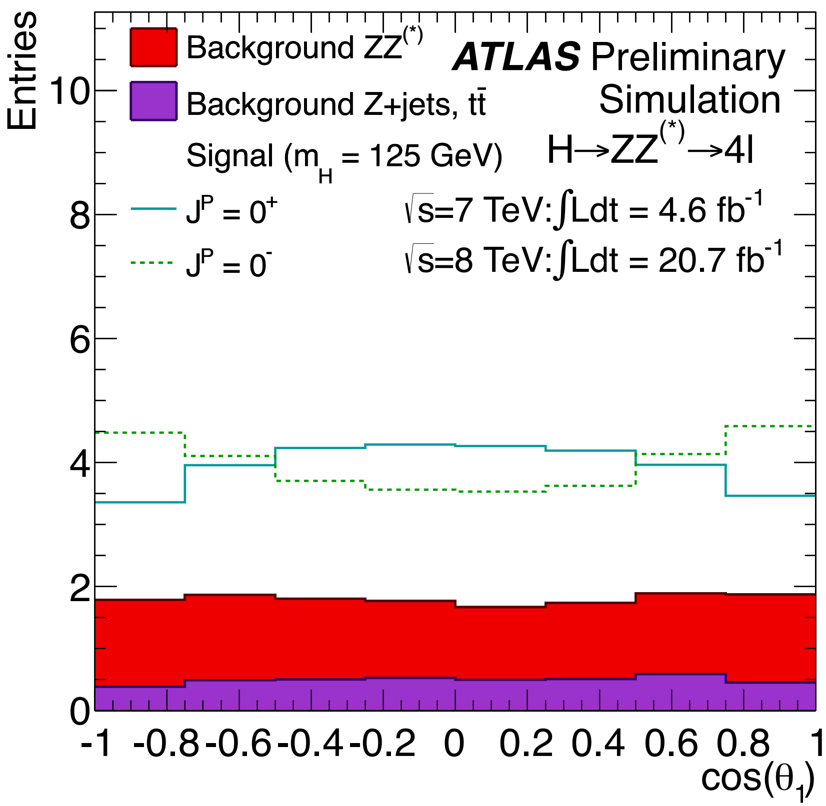
Measurement the spin and parity quantum numbers

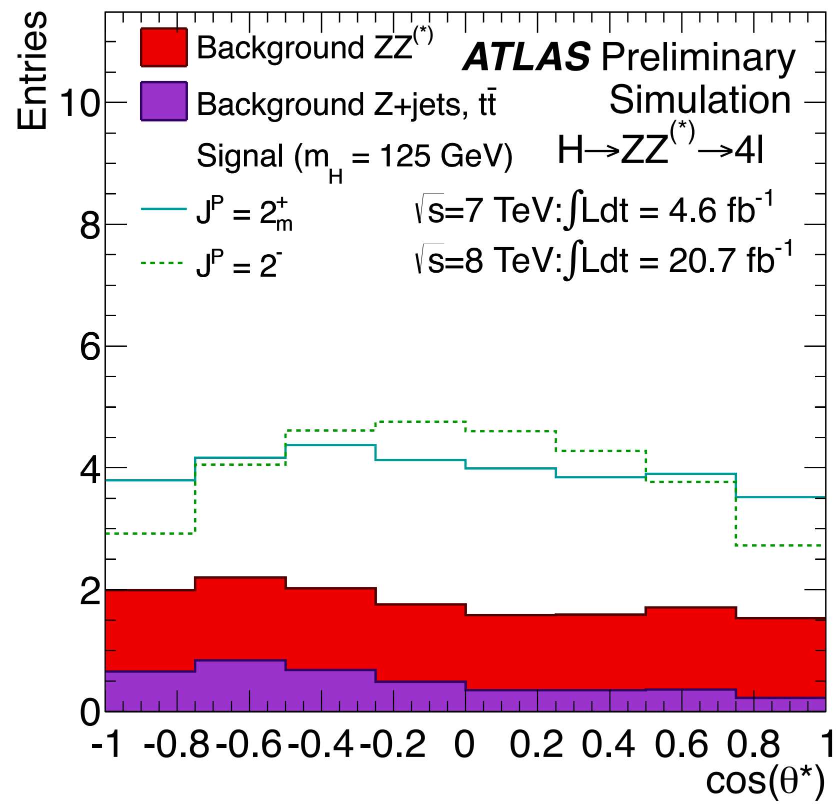
- The H→ZZ*→4l channel allows the full reconstruction of the final state
  - It's one of the most important channels for the measurement of the resonance properties
- Spin and parity of the decaying resonance affect the polarization of the two Z's in the final state
  - Angular distributions can be used to test spin and parity quantum numbers of the decaying resonance



Example of sensitive variables

Decay angle of the first l^+l^- pair



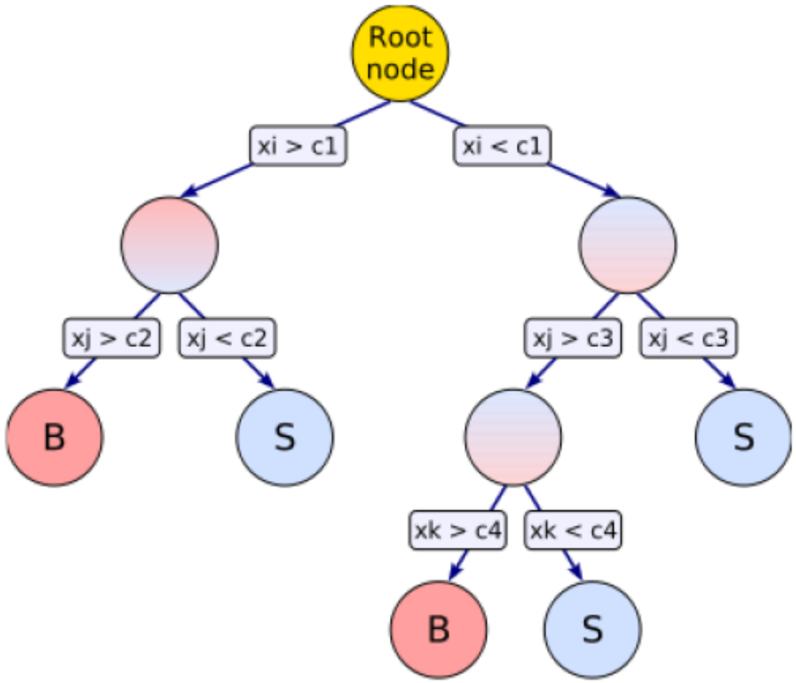


- The variables are discriminant, but not enough to be able to just place a cut.
- Any cut would not remove much more background than signal

→ Need a so-called multi-variate analysis (MVA)

Building a Boosted Decision Tree (BDT)

- A BDT is a sequence of binary cuts on one the selection variables
- At each step, the variable providing the best S / B separation, and the optimal cut is chosen
- At each S/B split the procedure is repeated for each of the two subsets obtained



- The procedure stops when the subsets become so small, that the statistical fluctuations are larger than any improvement in separation
- The nodes of the final level (“leaves”) are classified as S or B according to the class the majority of the events belongs to.

Examples of MVA: BDT

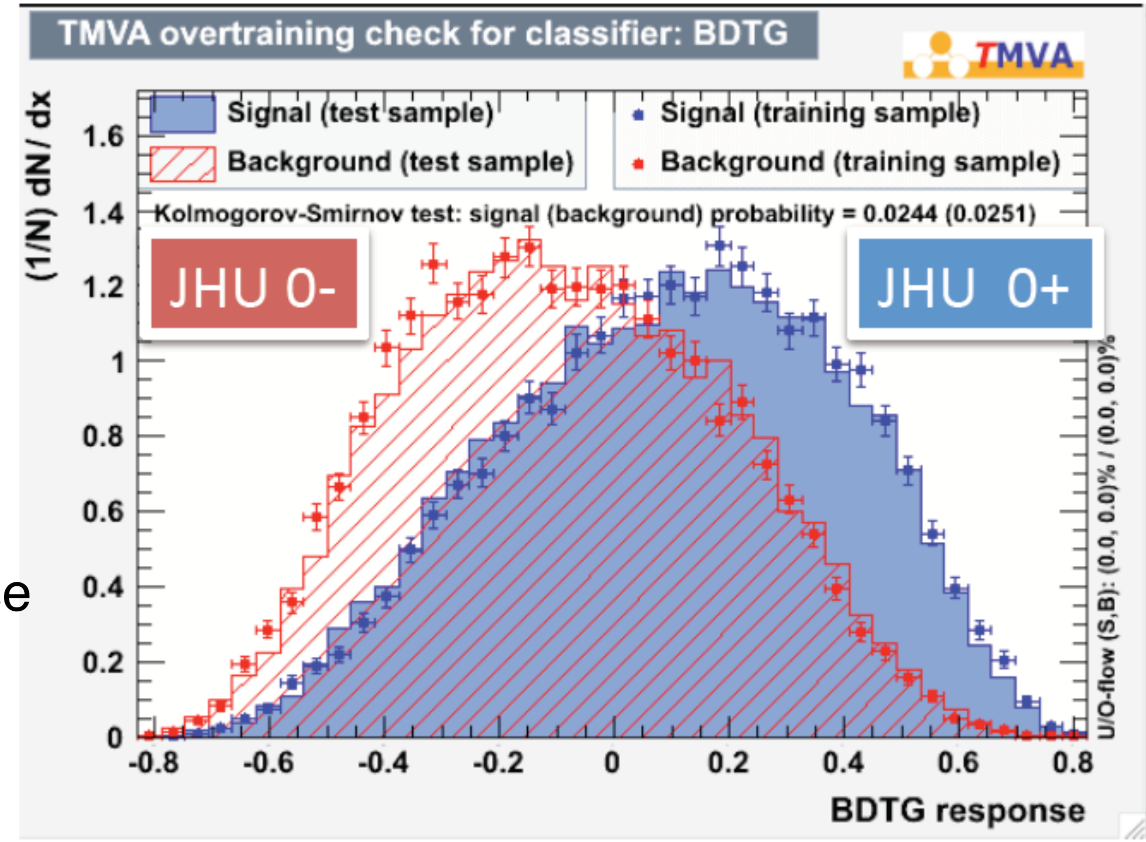
- Boost: optimize the sensitivity
- Use multiple trees:
  - Each can be trained with the same dataset, but reweighted in order to optimize the sensitivity
- An example (Adaptive Boost)
  - Train the first tree with the original weights (e.g. with the cross sections)
  - The subsequent tree is trained reweighting the previously misclassified events with a weight $\alpha=(1\text{-err})/\text{err}$, where err is the mis-classification rate
- The number of trees and their depth, must be chosen according to the available MC training statistics

More on MVA analyses

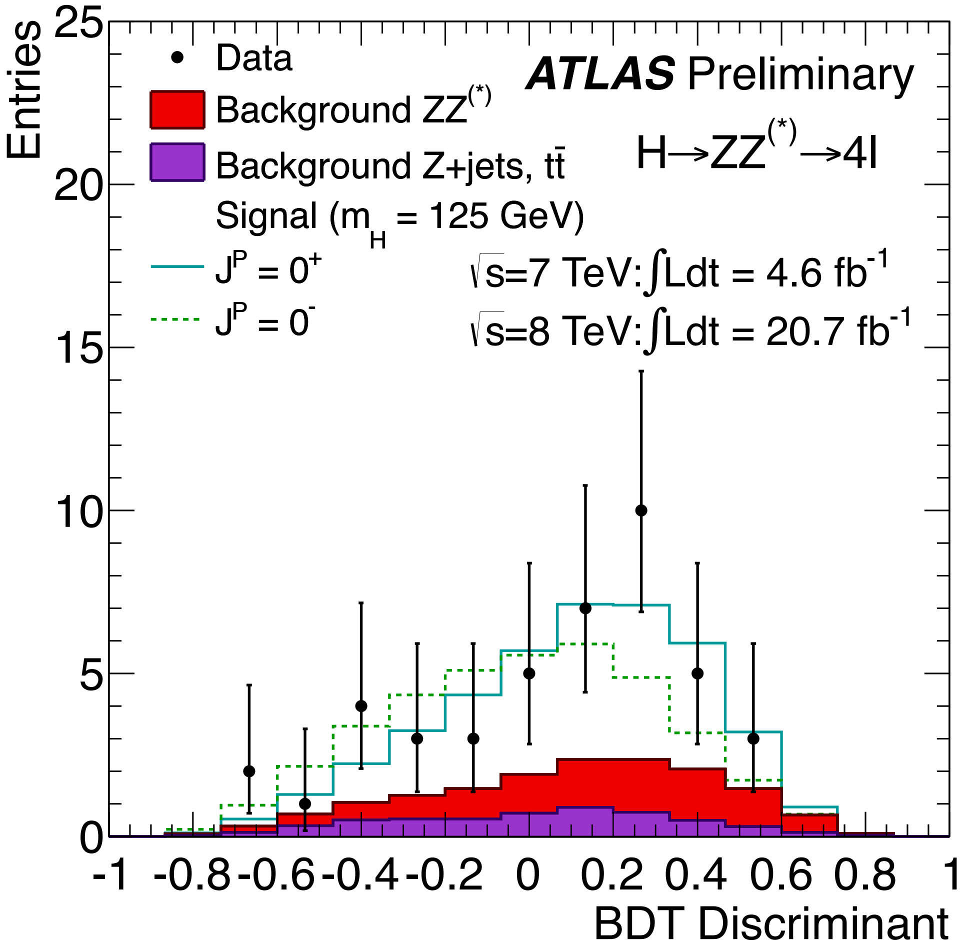
- The “training” of the method is usually performed on MC events
- To correctly evaluate the separation, divide the sample in two sub-sets: training and test samples
- This is also done in order to avoid the over-training
  - The MVA might “learn” the fluctuations of your training sample

This is an example of a light overtraining
→the separation between the training samples is slightly better than for the test samples

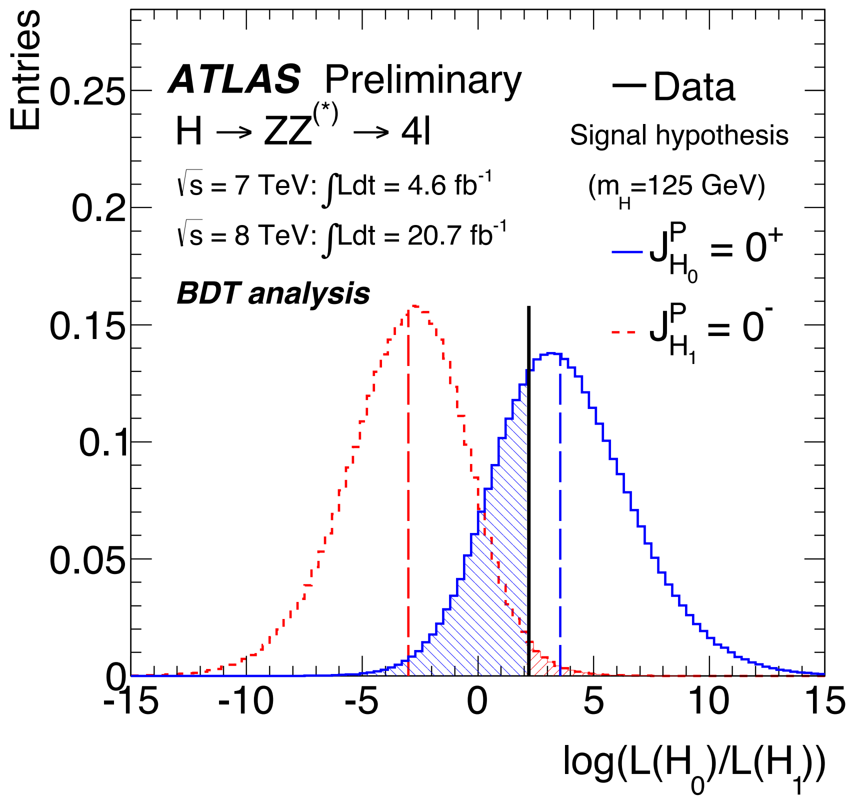
In this case one should increase the MC stat, or decrease the number of trees



Discriminant distribution and hypothesis test

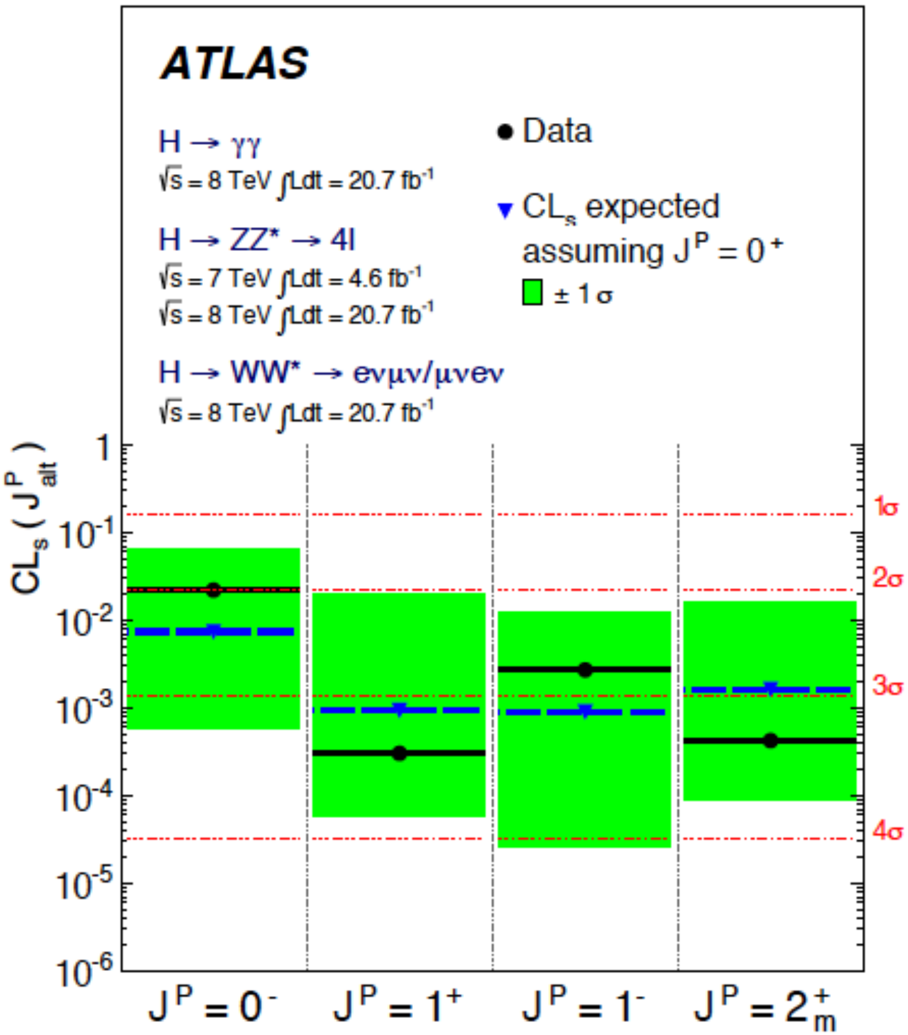


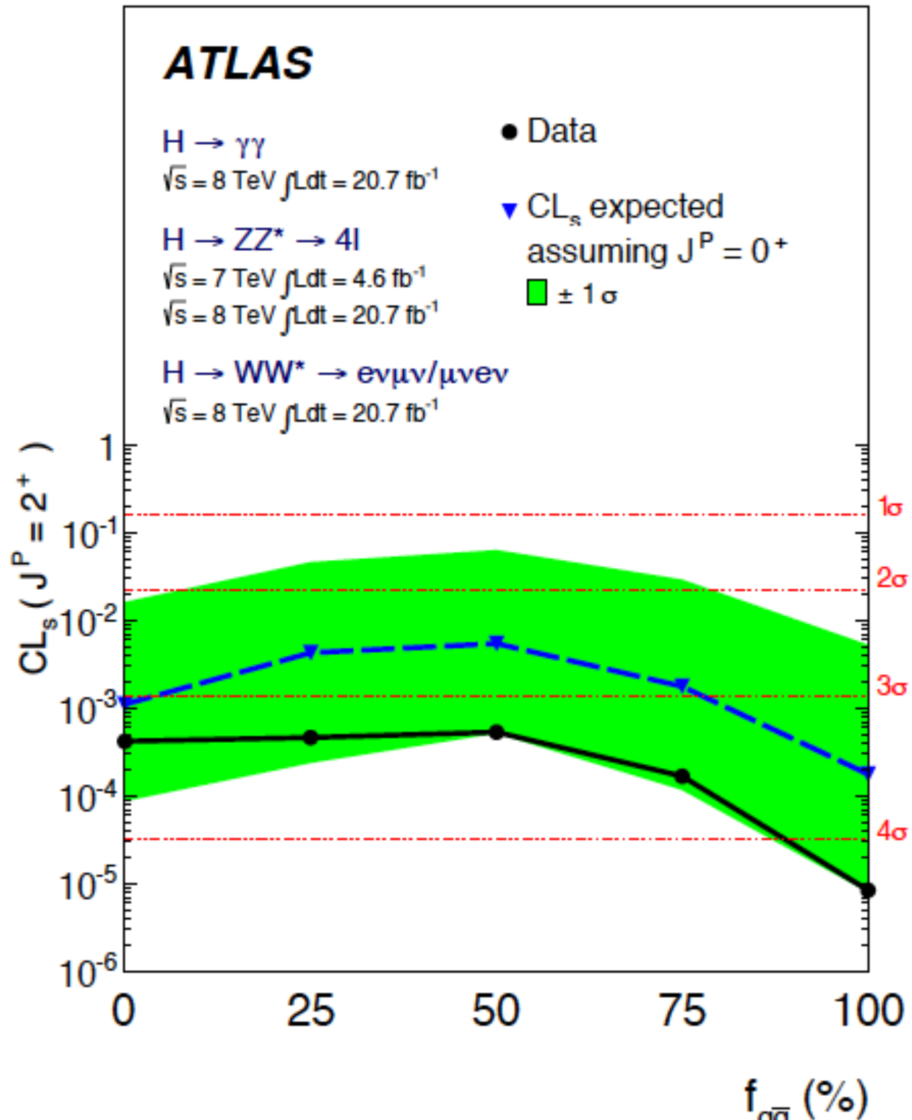
- Log-likelihood ratio as test statistics
- Generate pseudo-experiments assuming each of the two hypotheses
- Median of 0^+ → SM expected p0
- Data value → Observed p0



Final results on spin and parity

$$CL_s = \frac{p_0(J^P_{alt})}{1 - p_0\ (J^P = 0^+)}$$





CONCLUSIONS

Conclusions

- This was an attempt to consider one real example of an analysis and follow all the aspects of its development
  - Initial thoughts
  - Designing the analysis
  - Applying it to the data → the discovery
  - Measurements of the new resonance properties
- Due to lack of time, I couldn't give you many details, and a real summary of the results, but in case you would like to discuss more please do not hesitate to contact me:
  - Mail: stefano.rosati@cern.ch
  - Office: Building Marconi, 2nd floor, 229-b
- And in general if you are interested in having a thesis with the ATLAS group, many topics available both on data analysis and detector development