UNIVERSITY OF SUSSEX

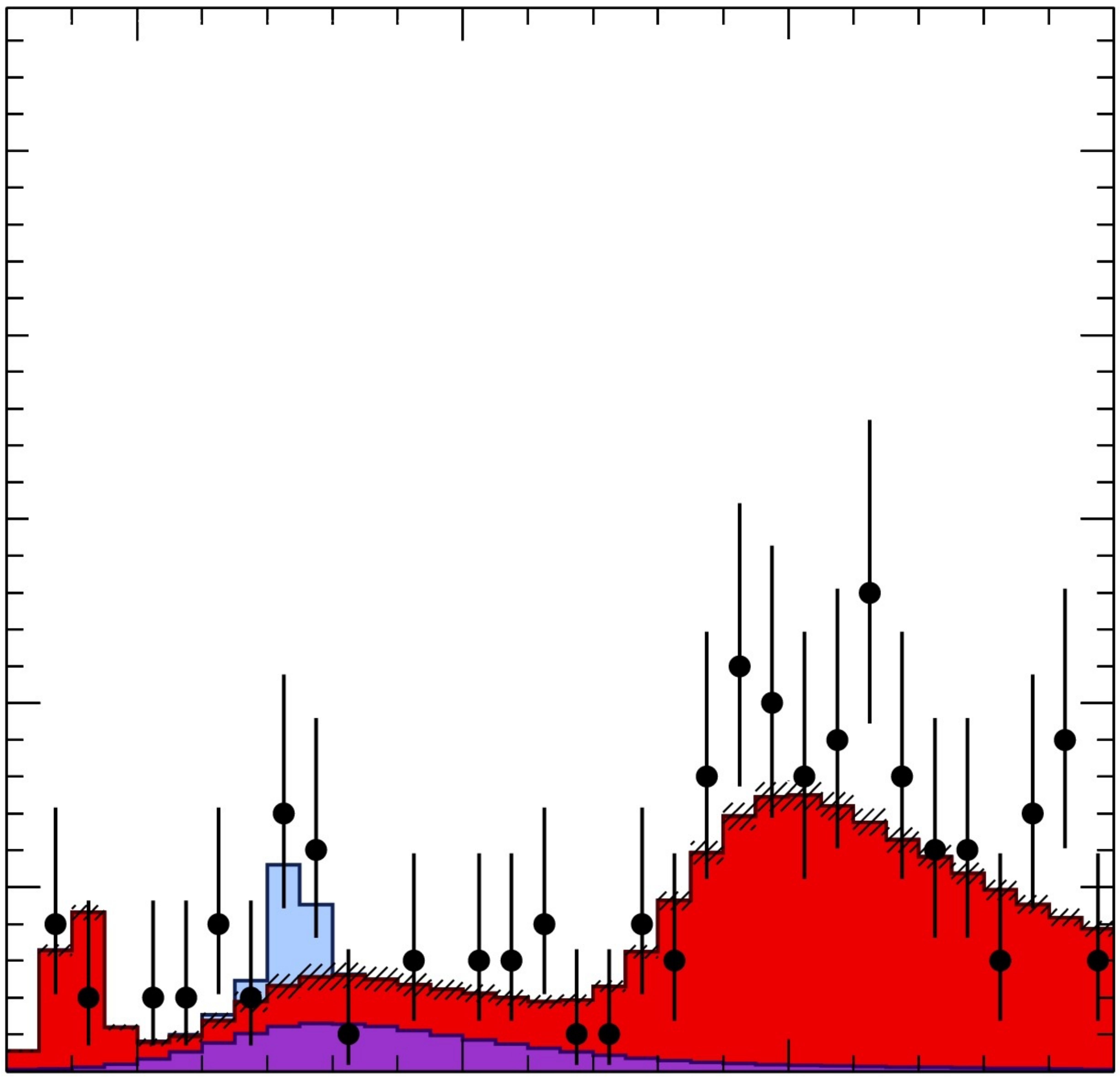
ATLAS EXPERIMENT

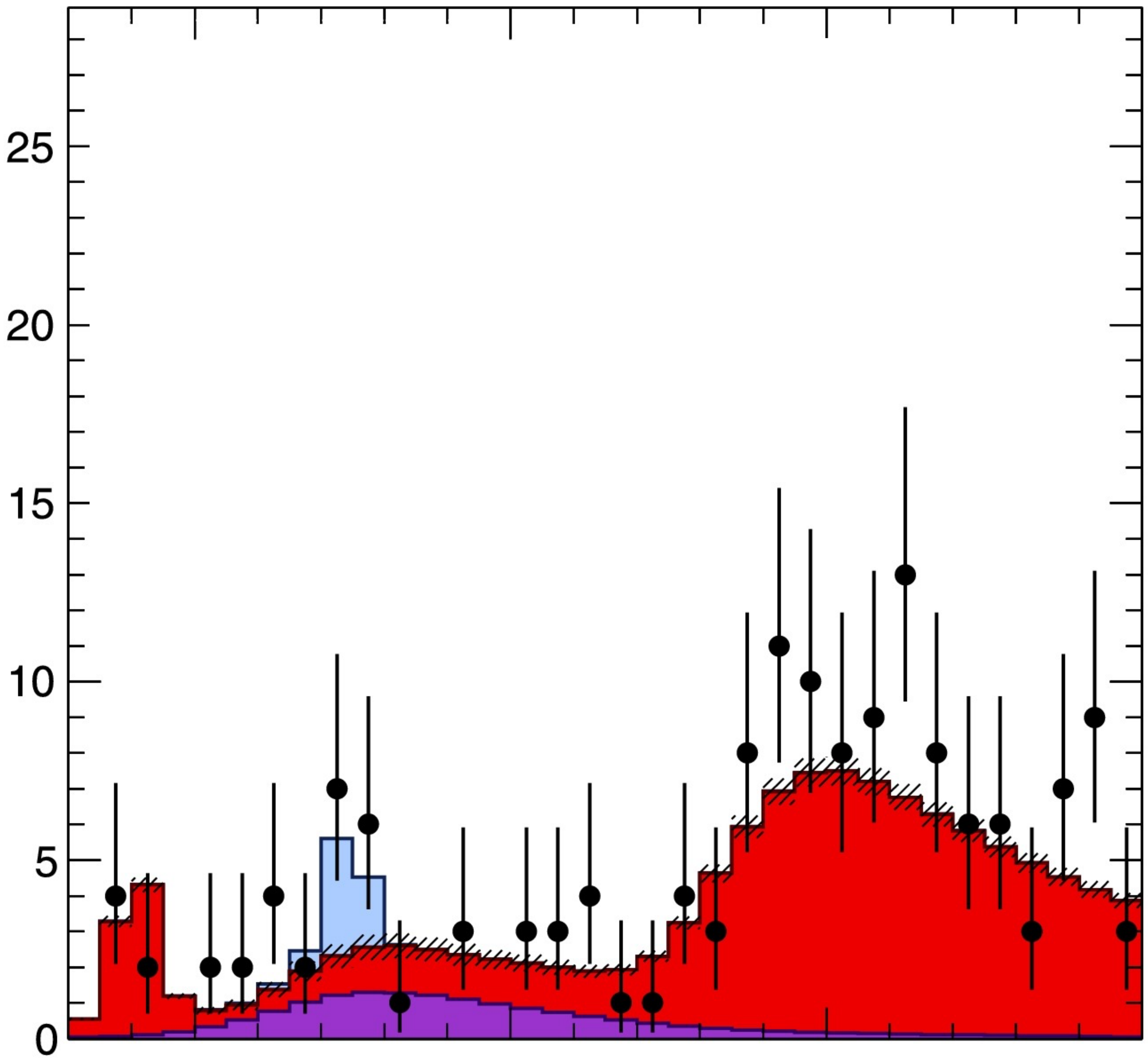
Meirin Oan Evans, Adam Powell (+ help of many people)

CAPS 2019 Abertawe

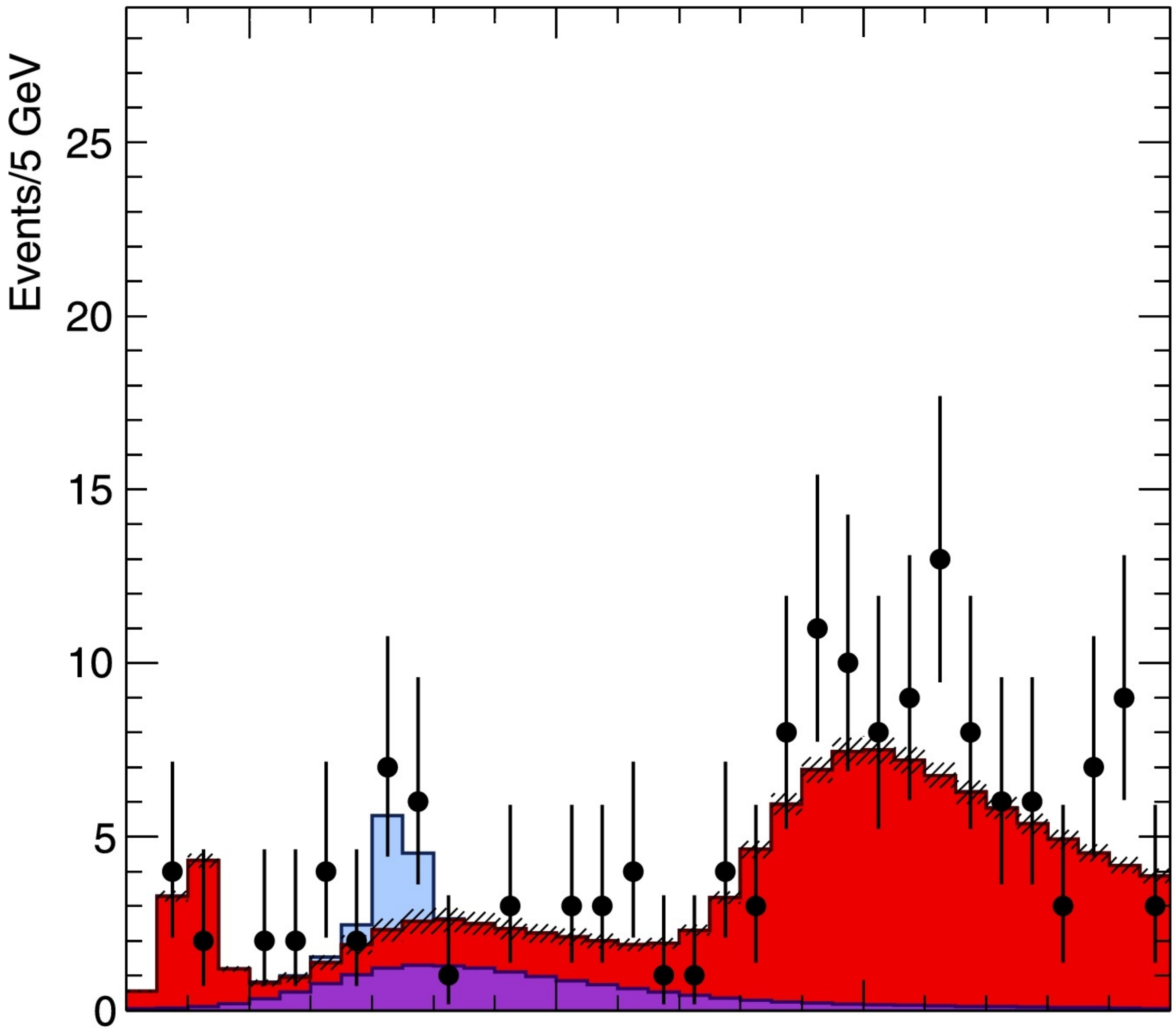
9th July 2019

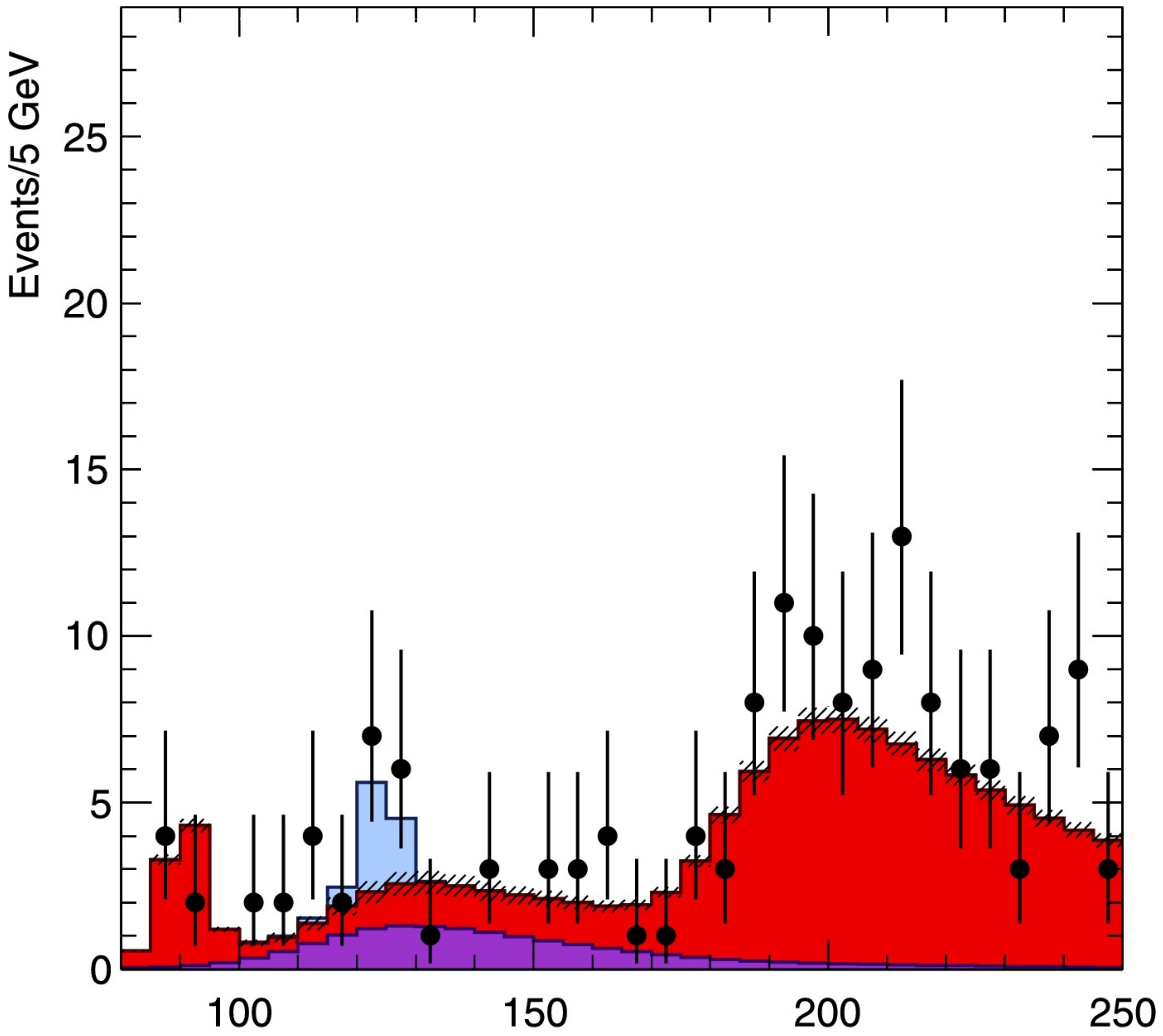
How to rediscover the Higgs

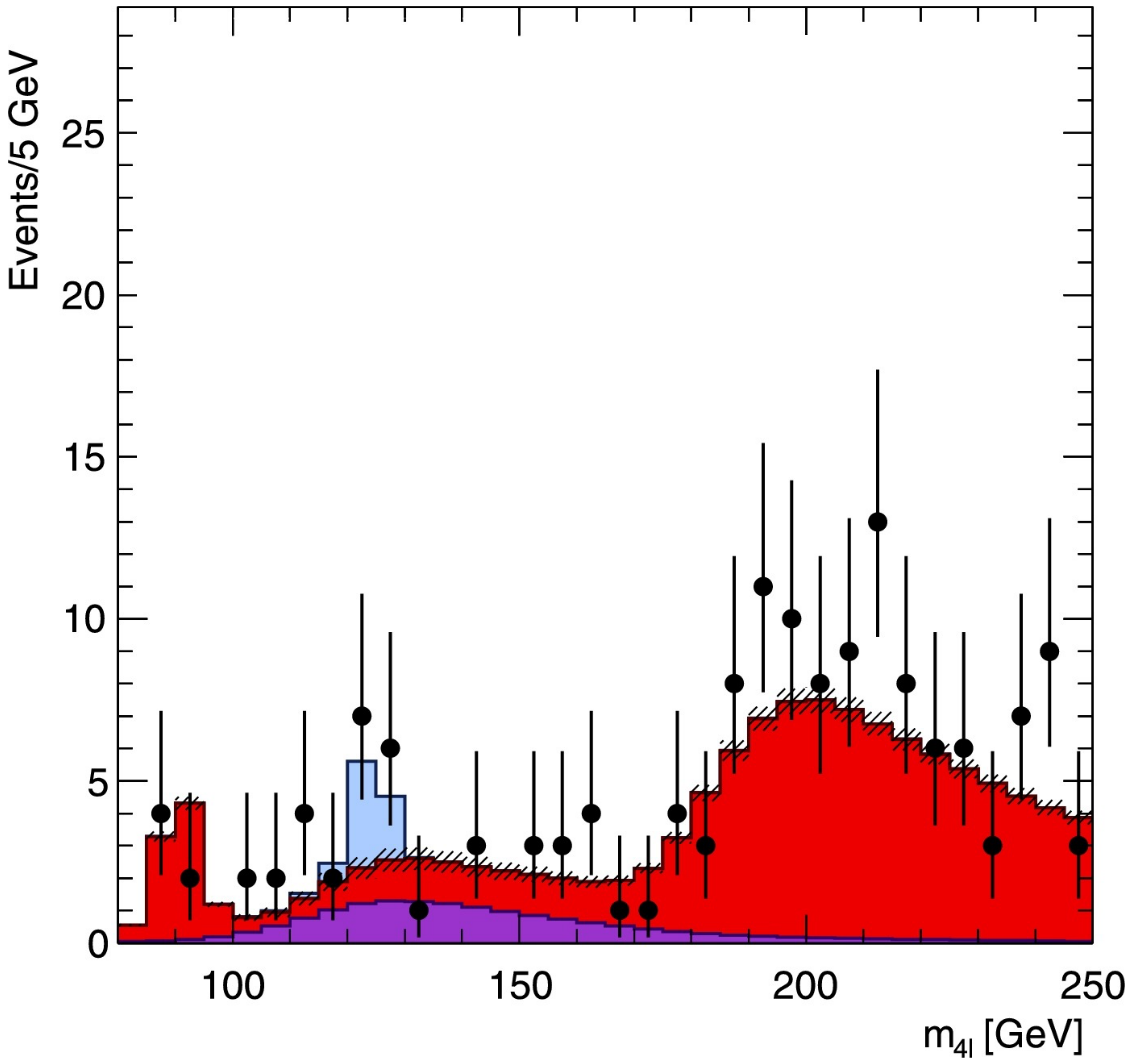


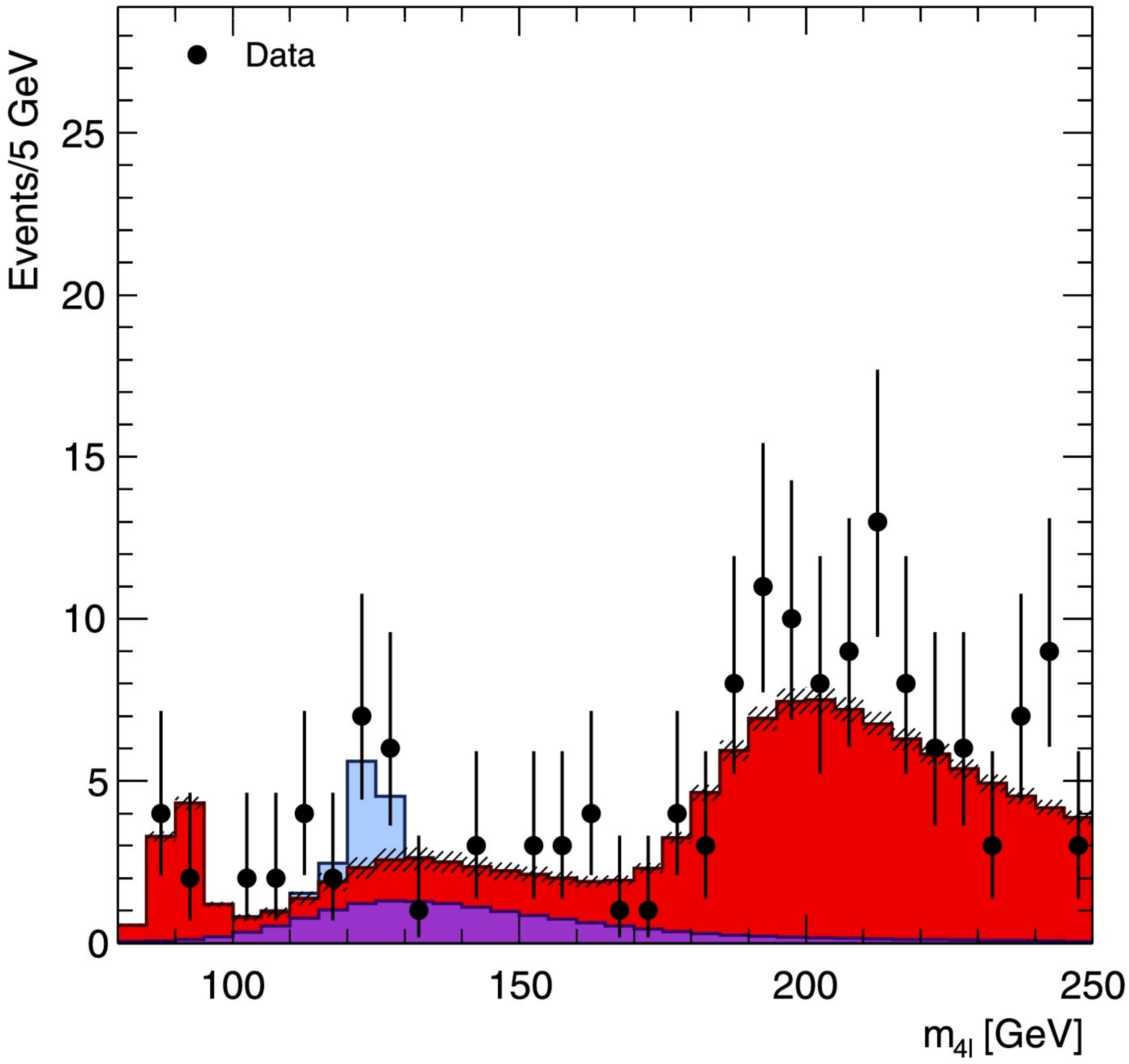


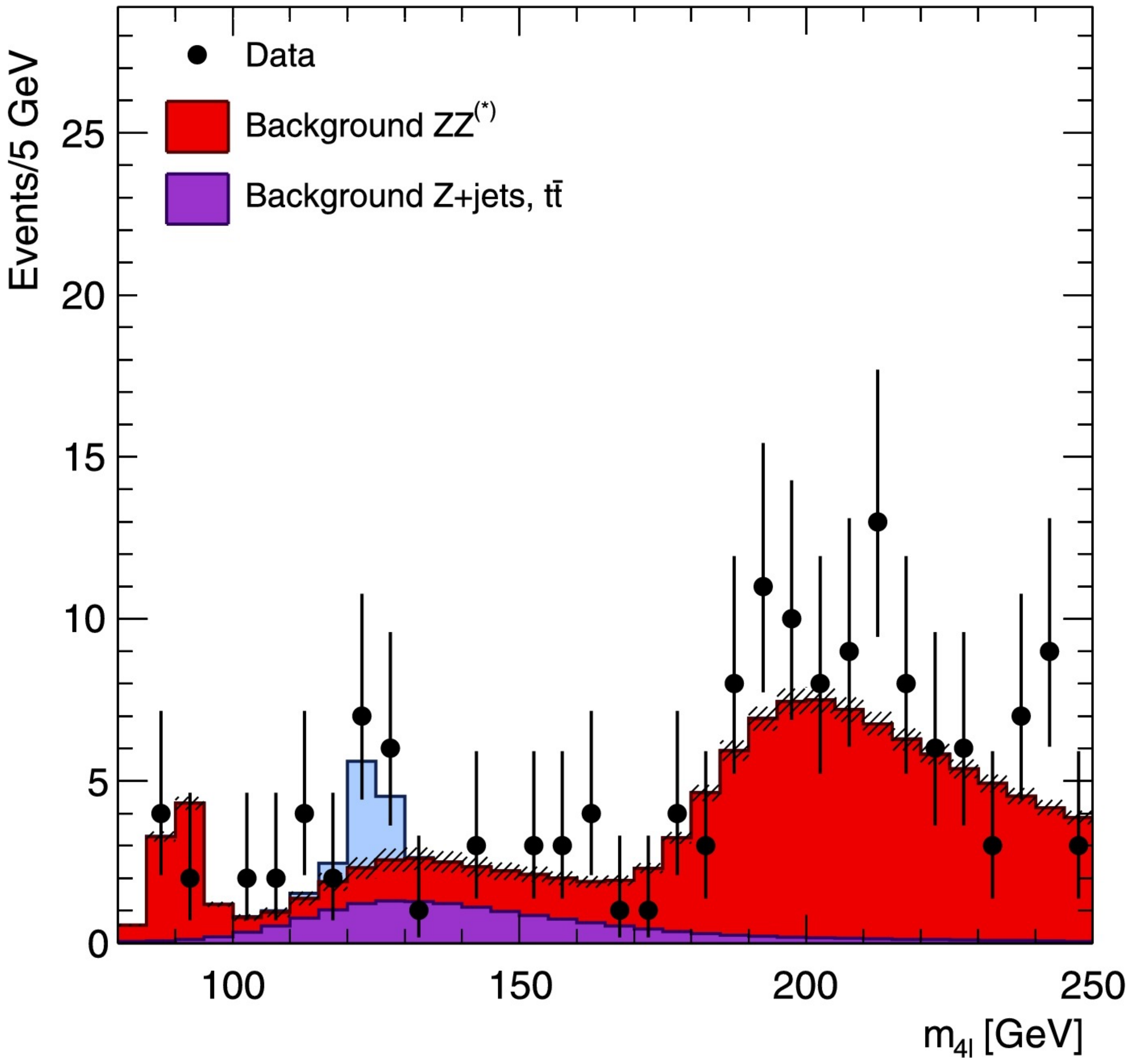
25
20
15
10
5
0


Events/5 GeV
25
20
15
10
5
0


Events/5 GeV
25
20
15
10
5
0
100
150
200
250

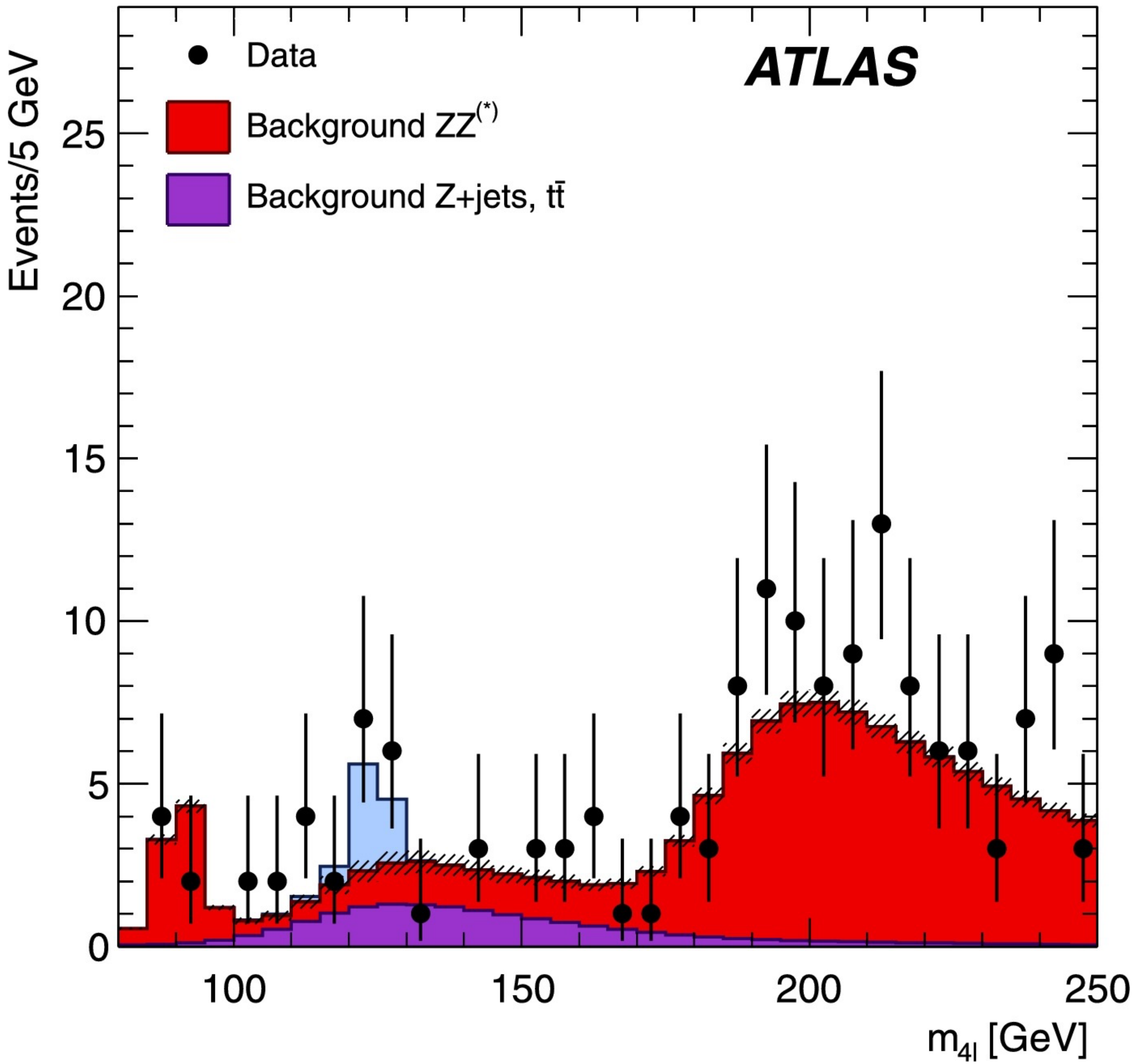

Events/5 GeV
25
20
15
10
5
0
100
150
200
250
m_{4l} [GeV]


Data
Events/5 GeV
25
20
15
10
5
0
100
150
200
250
m_{4l} [GeV]

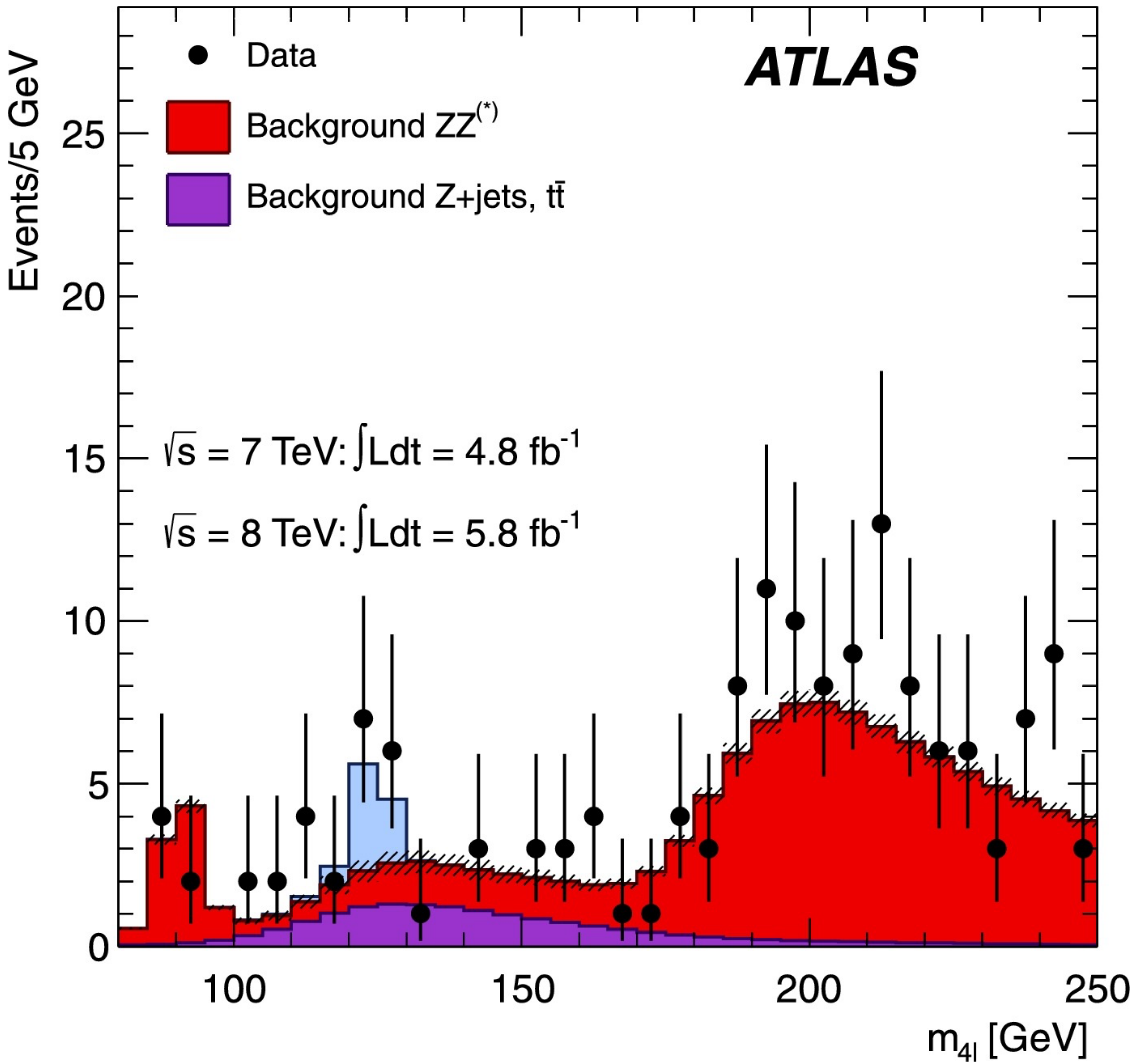


UNIVERSITY OF SUSSEX

ATLAS EXPERIMENT

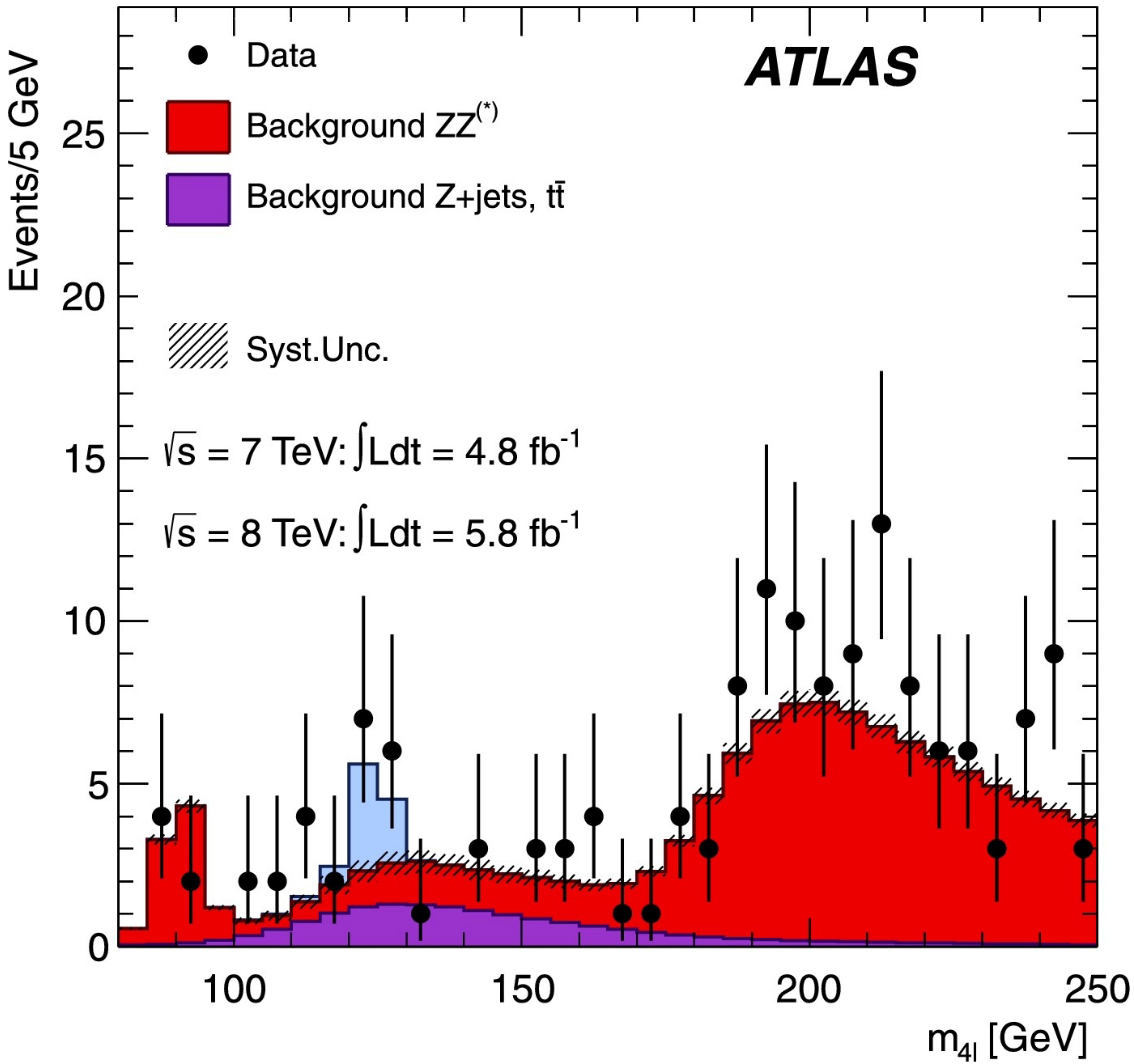


Data
Background $ZZ^{(*)}$
Background Z+jets, $t\bar{t}$
ATLAS
Events/5 GeV
25
20
15
10
5
0
100
150
200
250
m_{4l} [GeV]

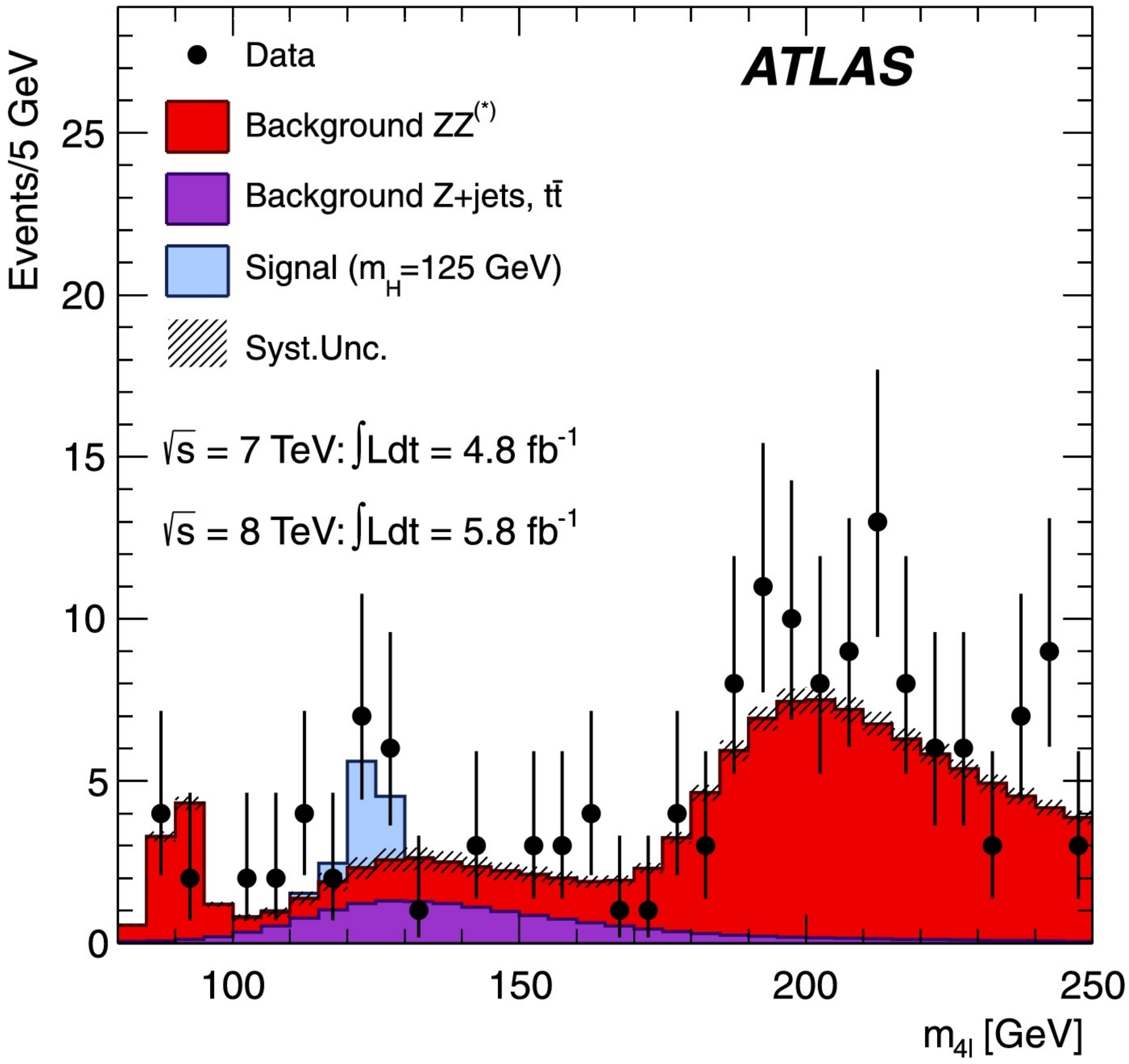


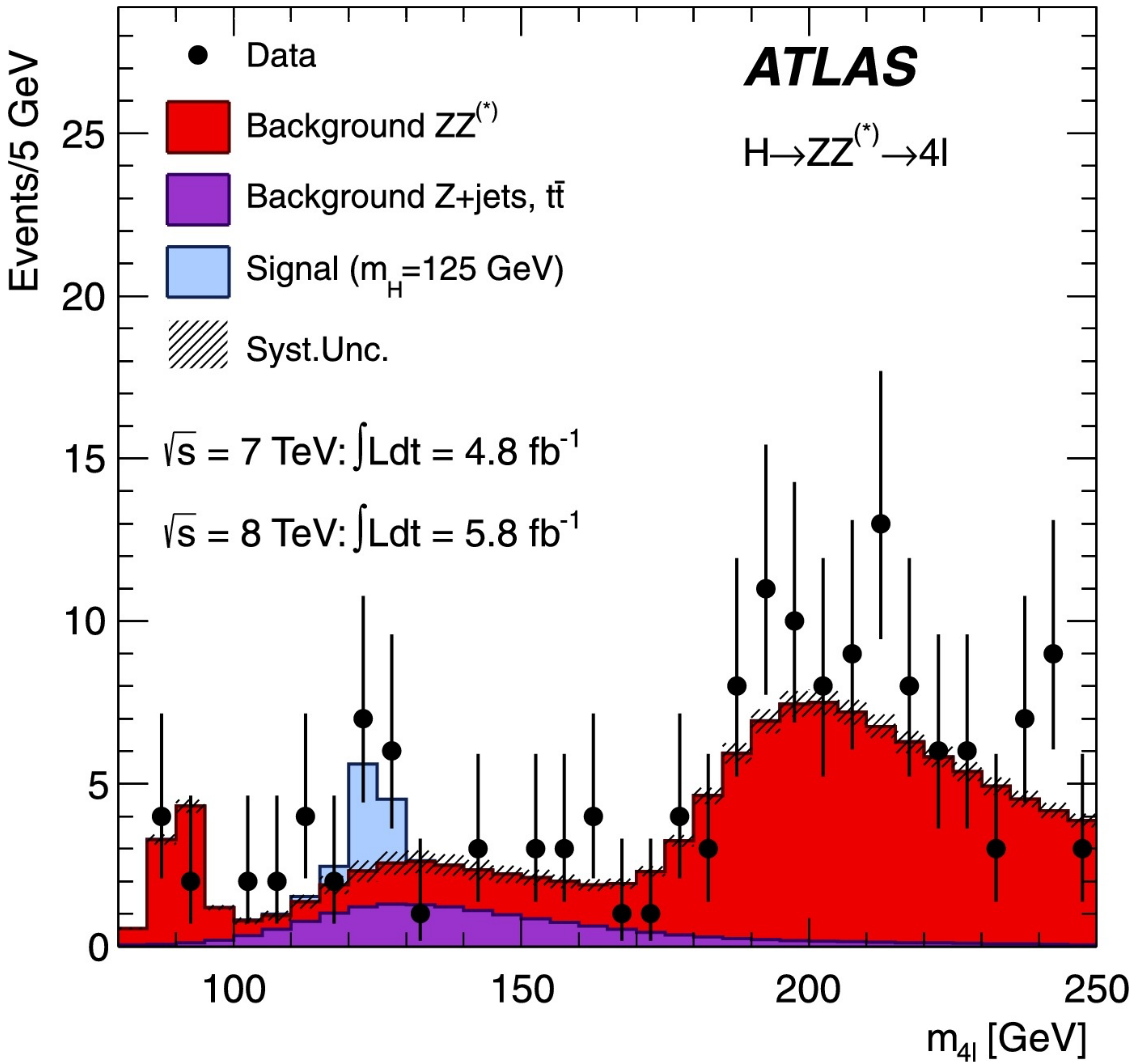
UNIVERSITY OF SUSSEX

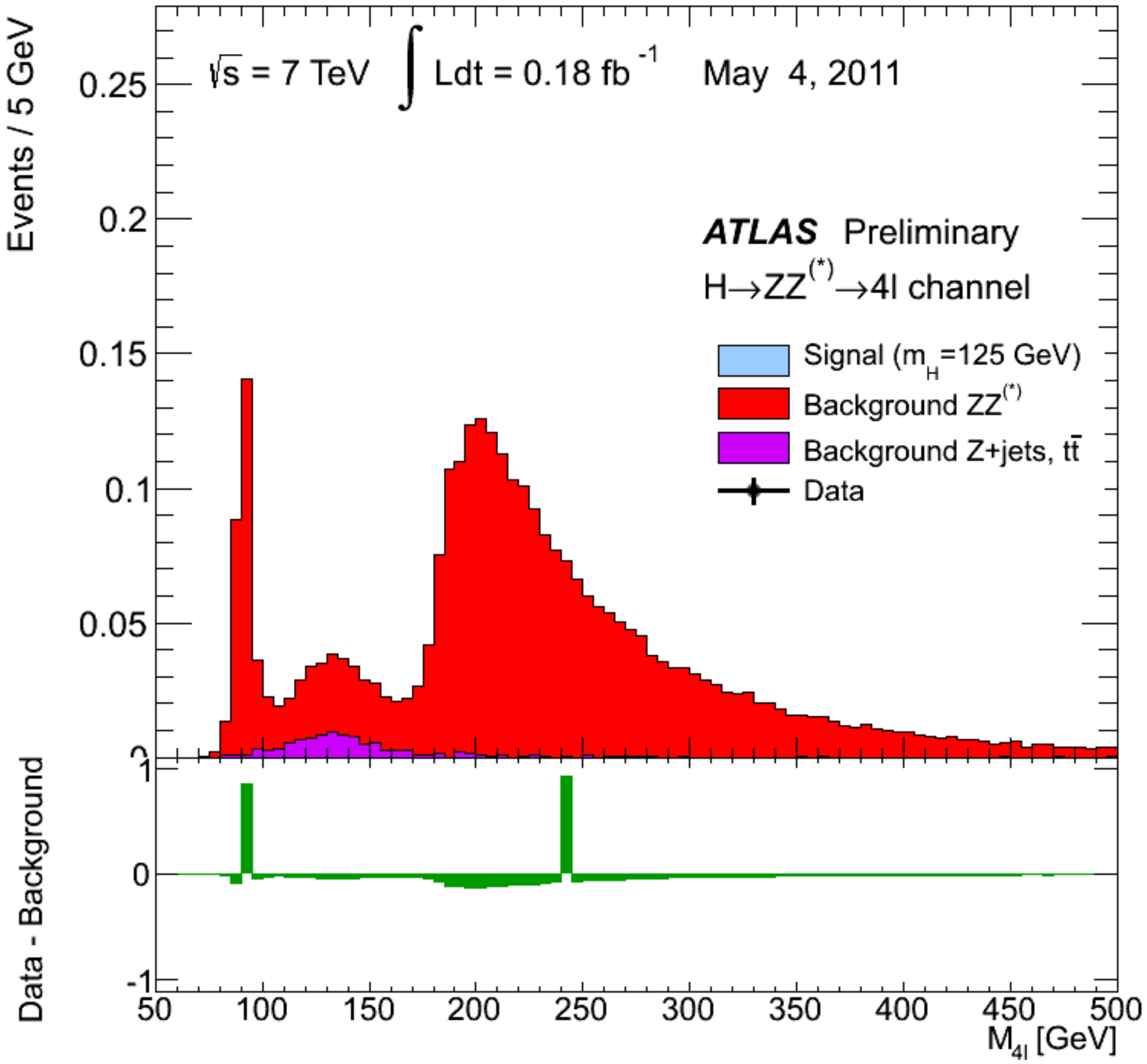
ATLAS EXPERIMENT



ATLAS
Data
Background ZZ(*)
Background Z+jets, tt̄
Syst.Unc.
√s = 7 TeV: ∫Ldt = 4.8 fb-1
√s = 8 TeV: ∫Ldt = 5.8 fb-1
Events/5 GeV
m4l [GeV]
0
5
10
15
20
25
100
150
200
250


Data
ATLAS
Background ZZ(*)
Background Z+jets, tt̄
Signal (m_H=125 GeV)
Syst.Unc.
√s = 7 TeV: ∫Ldt = 4.8 fb-1
√s = 8 TeV: ∫Ldt = 5.8 fb-1
Events/5 GeV
0
5
10
15
20
25
100
150
200
250
m4l [GeV]

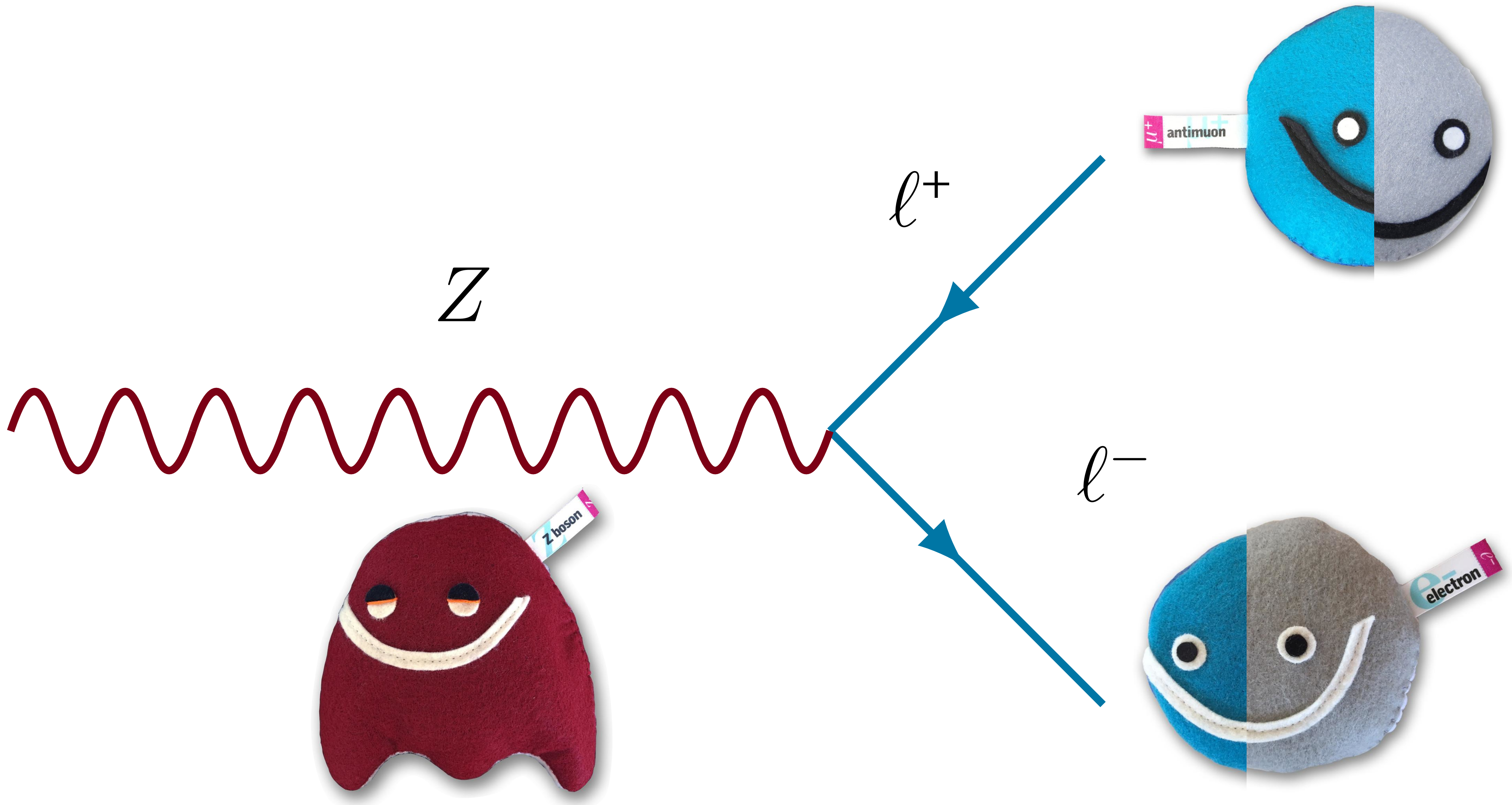

Events/5 GeV
Data
Background $ZZ^{(*)}$
Background Z+jets, $t\bar{t}$
Signal (m_H=125 GeV)
Syst.Unc.
ATLAS
$H \rightarrow ZZ^{(*)} \rightarrow 4l$
$\sqrt{s}$ = 7 TeV: $\int Ldt$ = 4.8 fb^{-1}
$\sqrt{s}$ = 8 TeV: $\int Ldt$ = 5.8 fb^{-1}
0
5
10
15
20
25
100
150
200
250
m_{4l} [GeV]



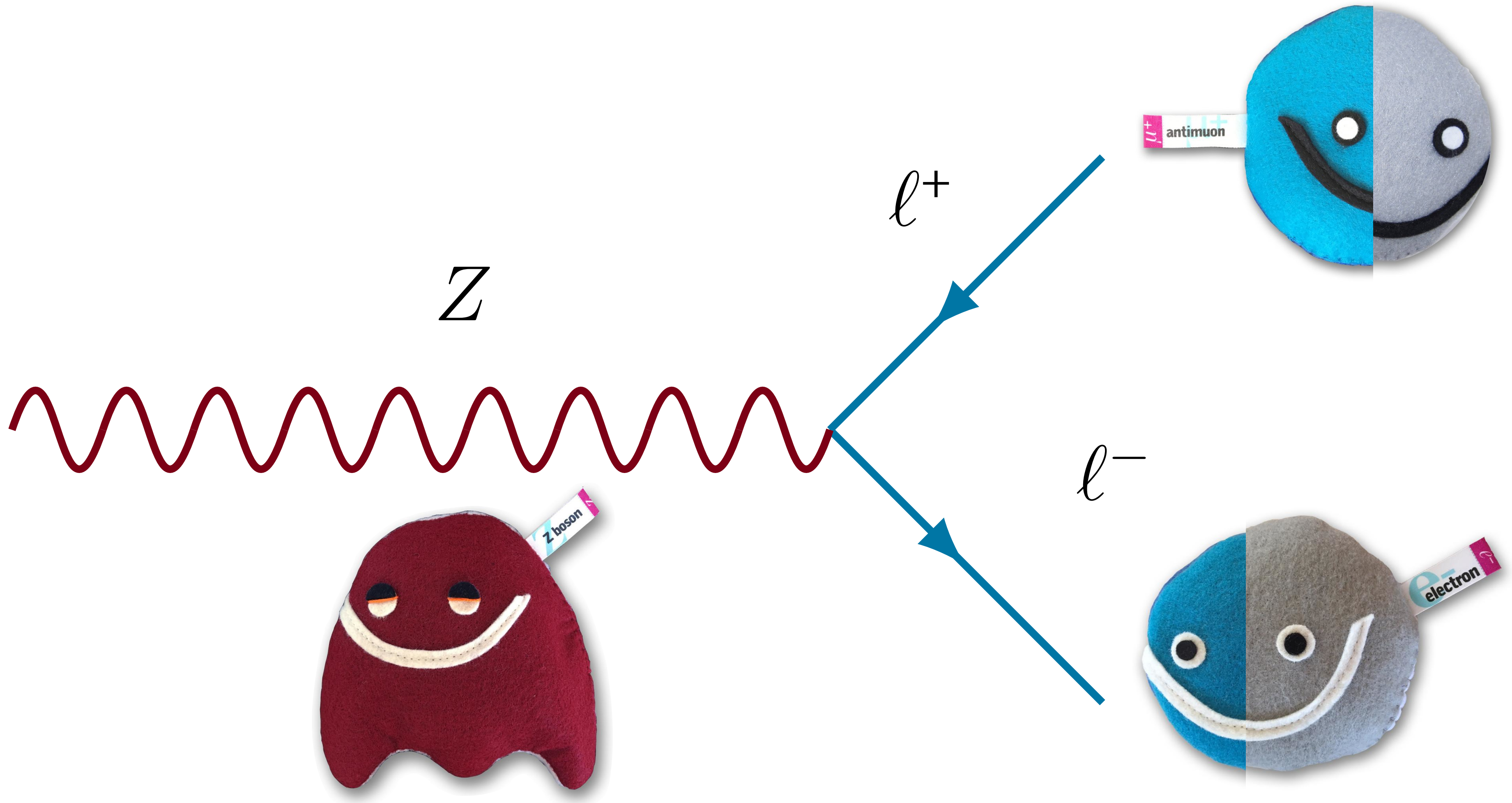
UNIVERSITY OF SUSSEX

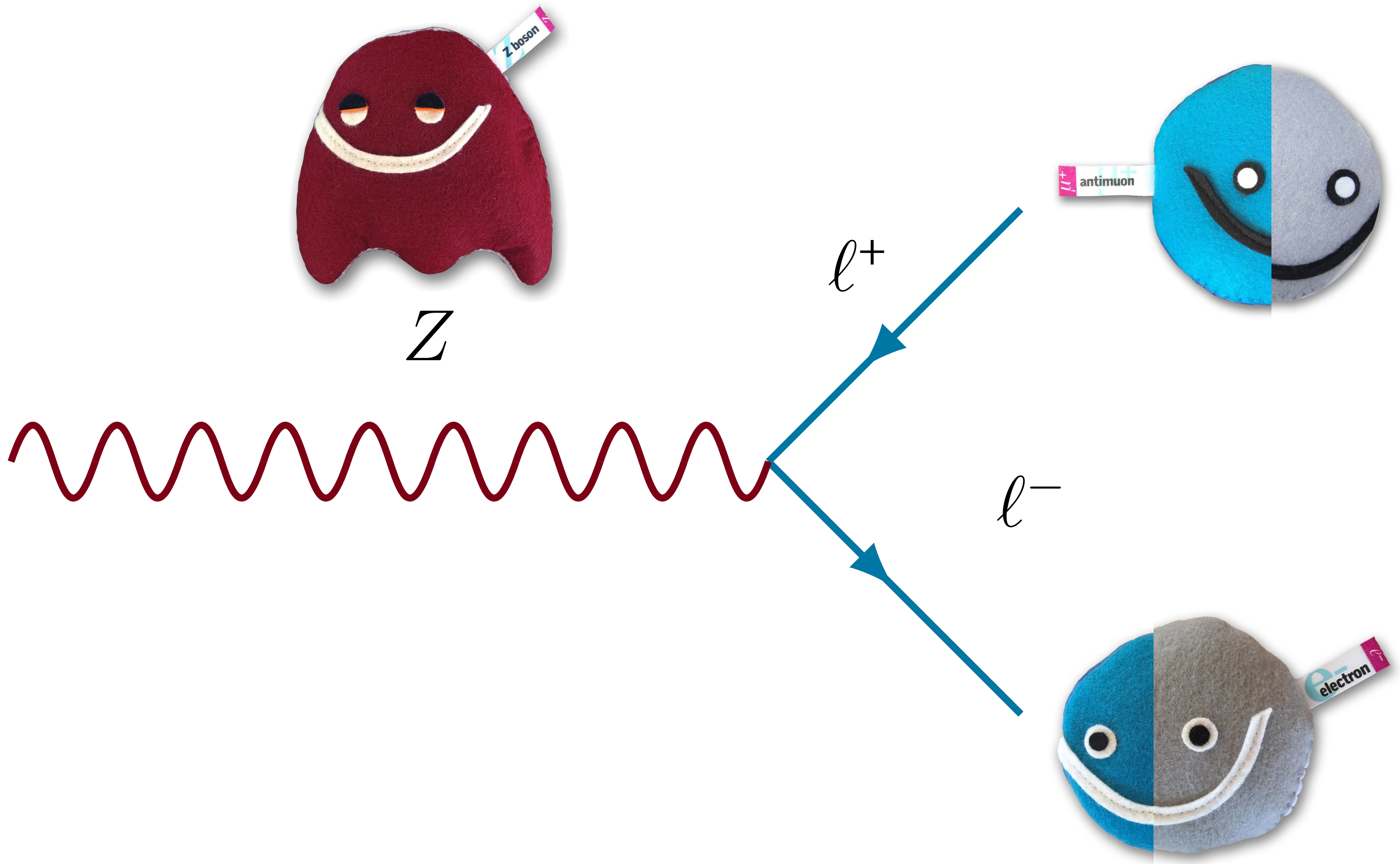
ATLAS EXPERIMENT

Z

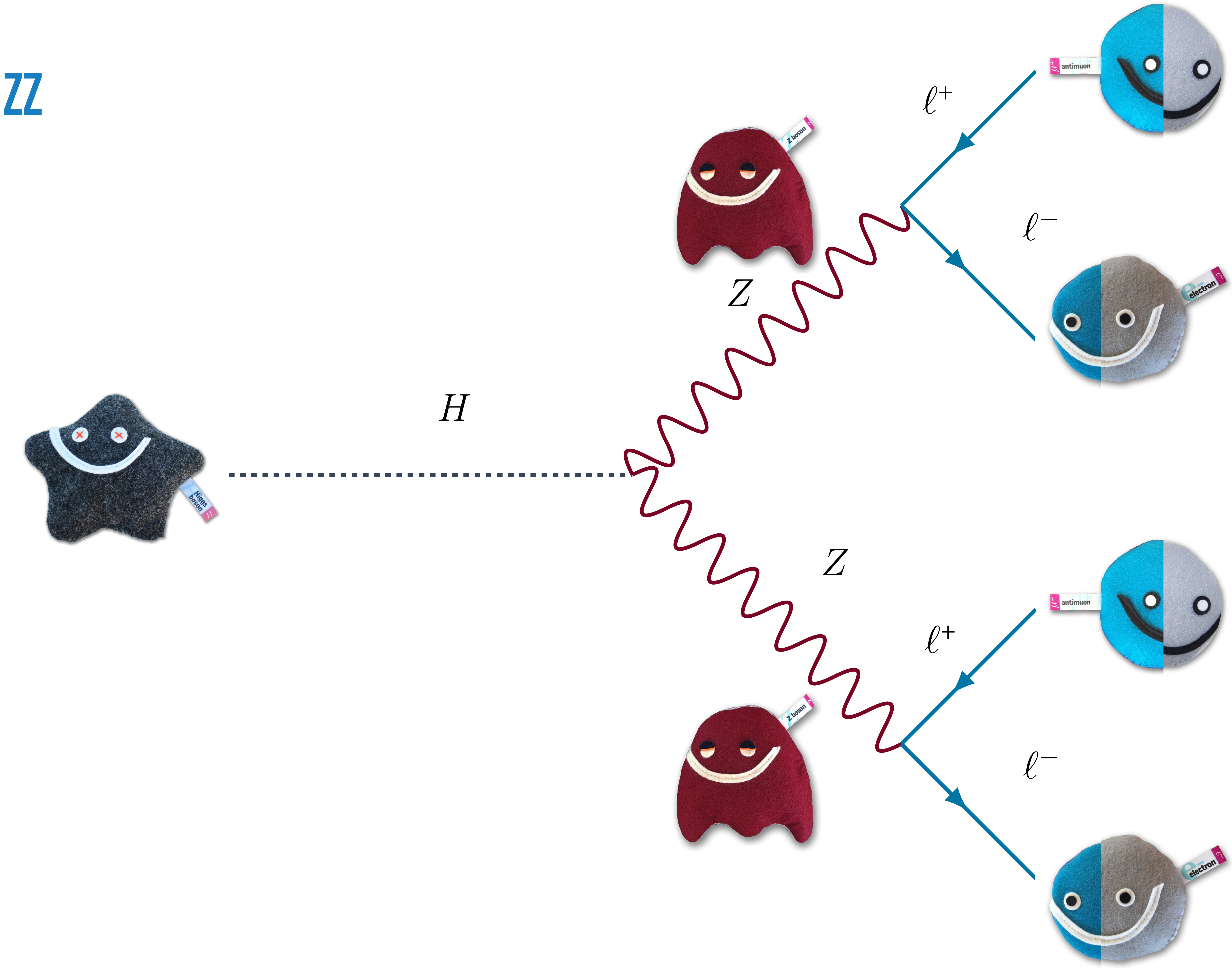


ZZ

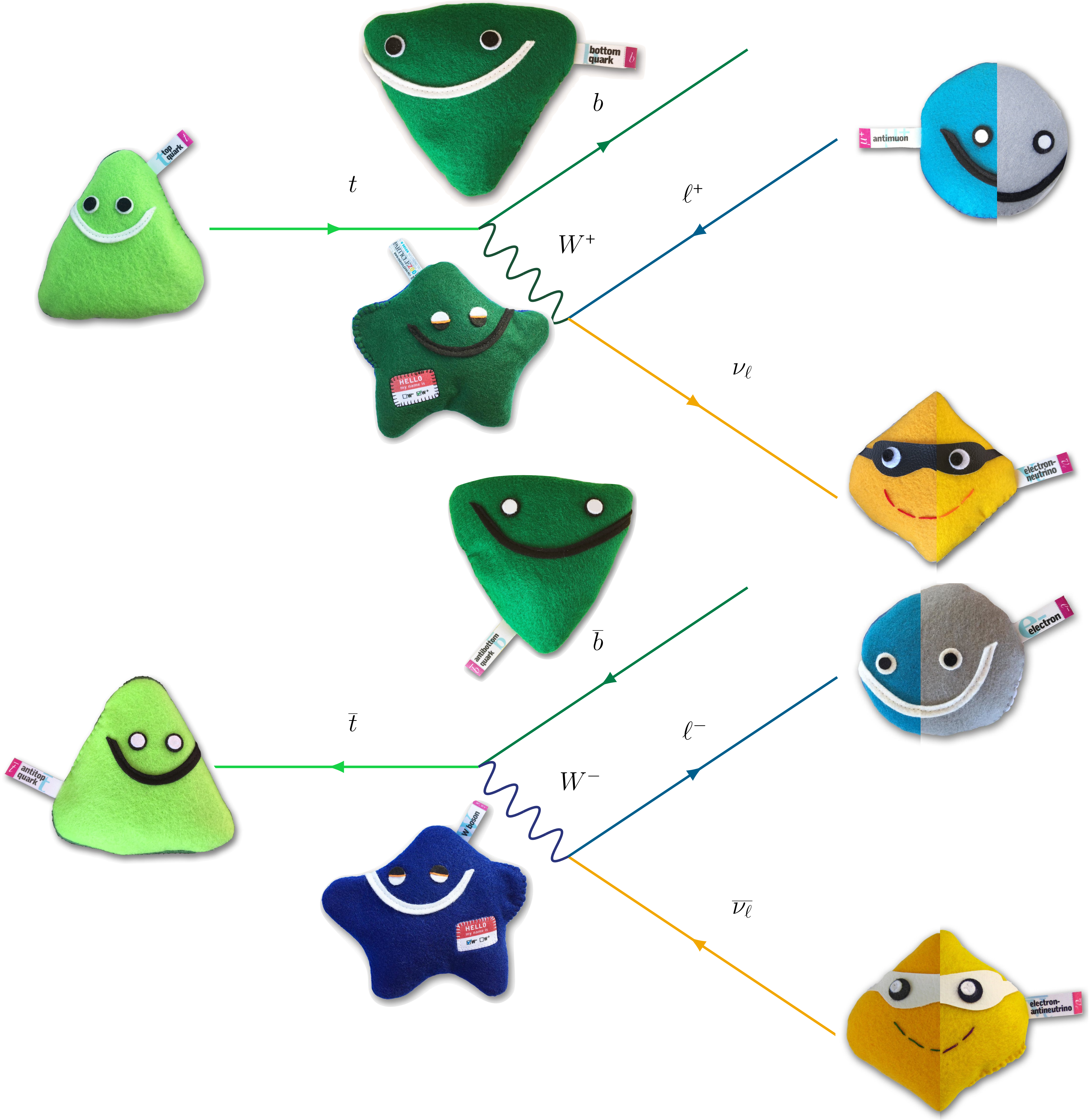

Z
ℓ^+
ℓ^-
Z boson
antimuon
electron


Z boson
Z
ℓ^+
ℓ^-
antimuon
electron

H → ZZ



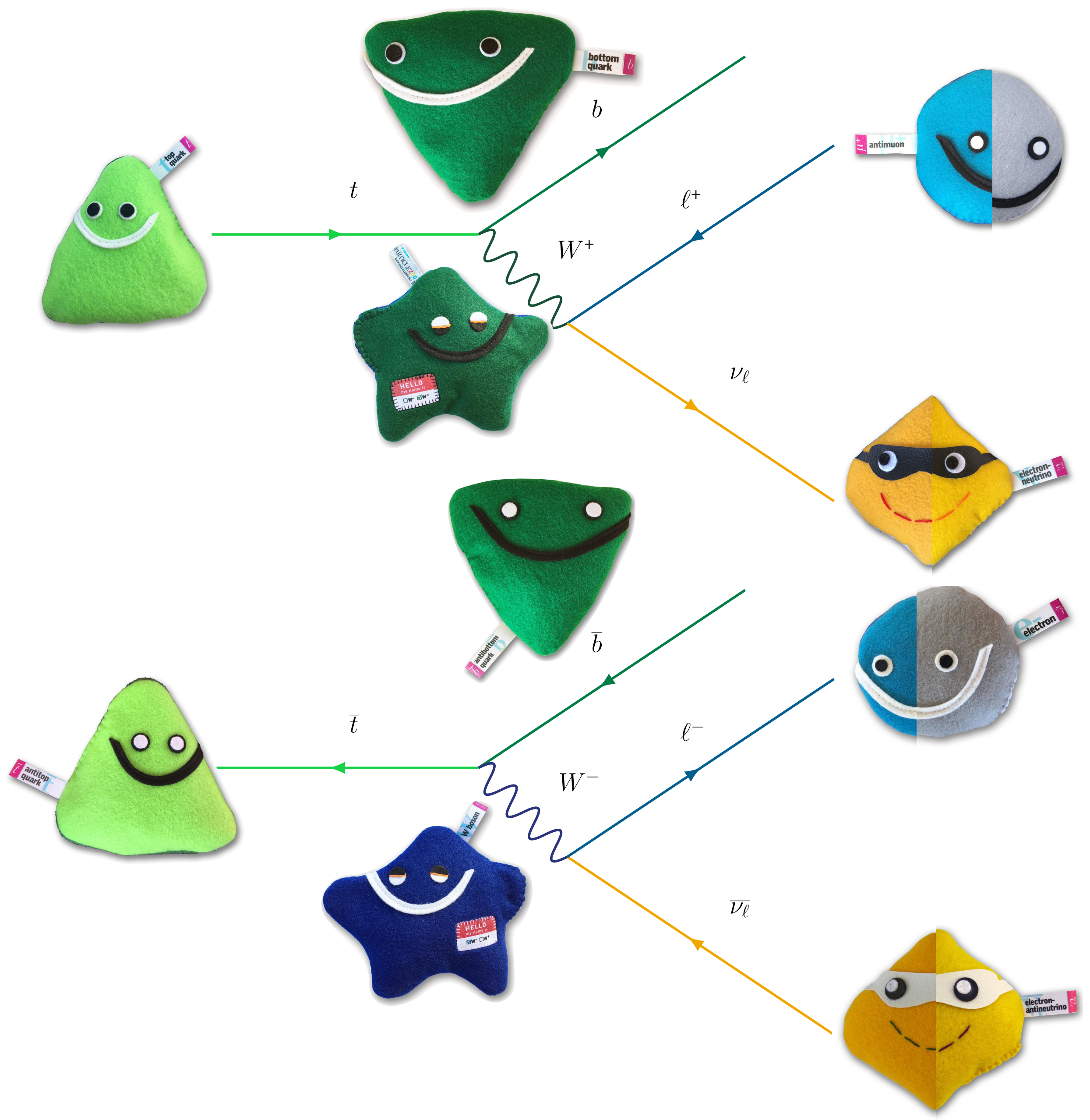
$t\bar{t}$



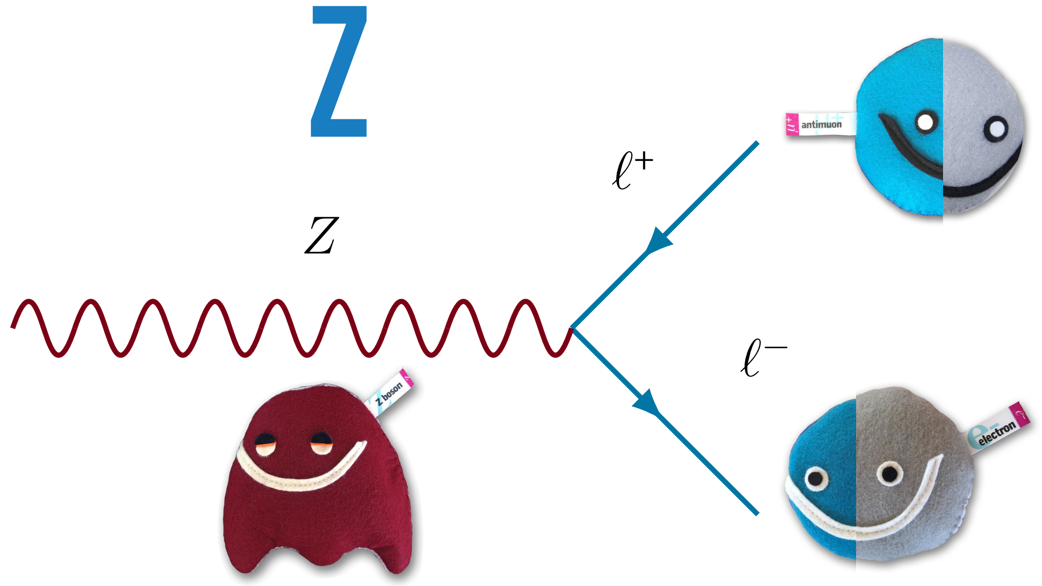
top quark
bottom quark
t
b
W^+
ℓ^+
antimuon
ν_ℓ
electron-neutrino
antibottom quark
$\bar{b}$
electron
antitop quark
$\bar{t}$
W^-
ℓ^-
W boson
$\overline{\nu_\ell}$
electron-antineutrino

SIGNAL VS BACKGROUND

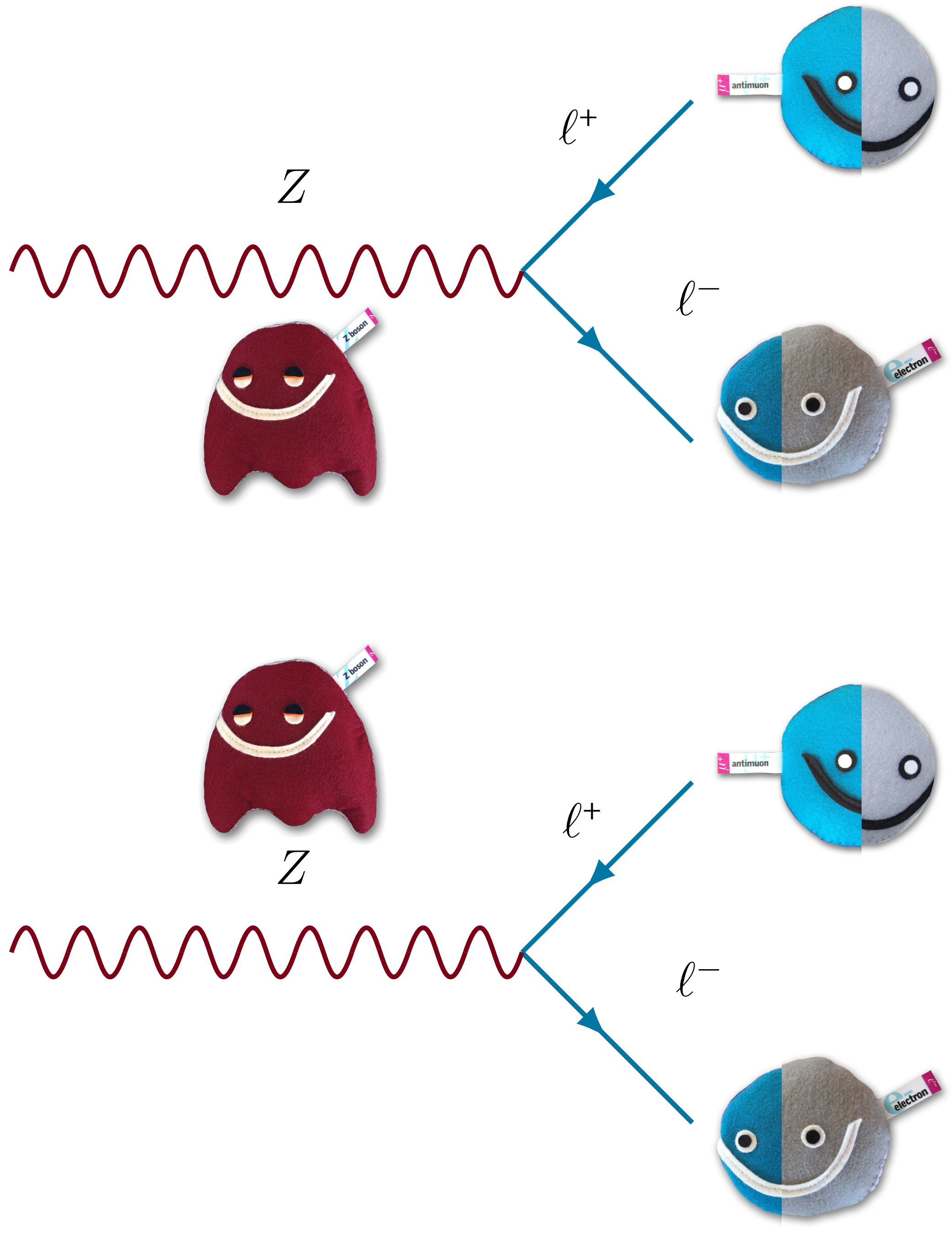
$t\bar{t}$



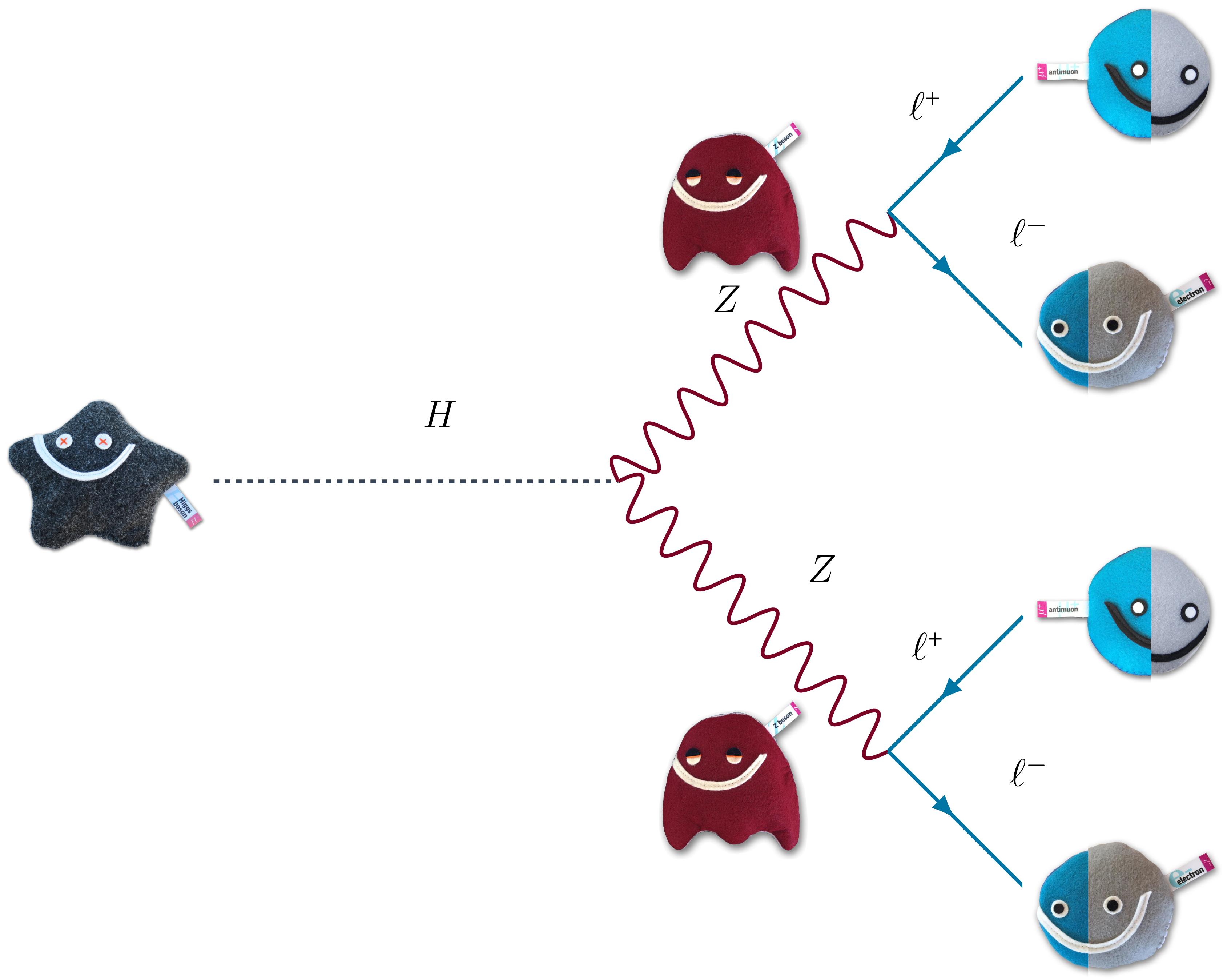
Z



ZZ



H → ZZ

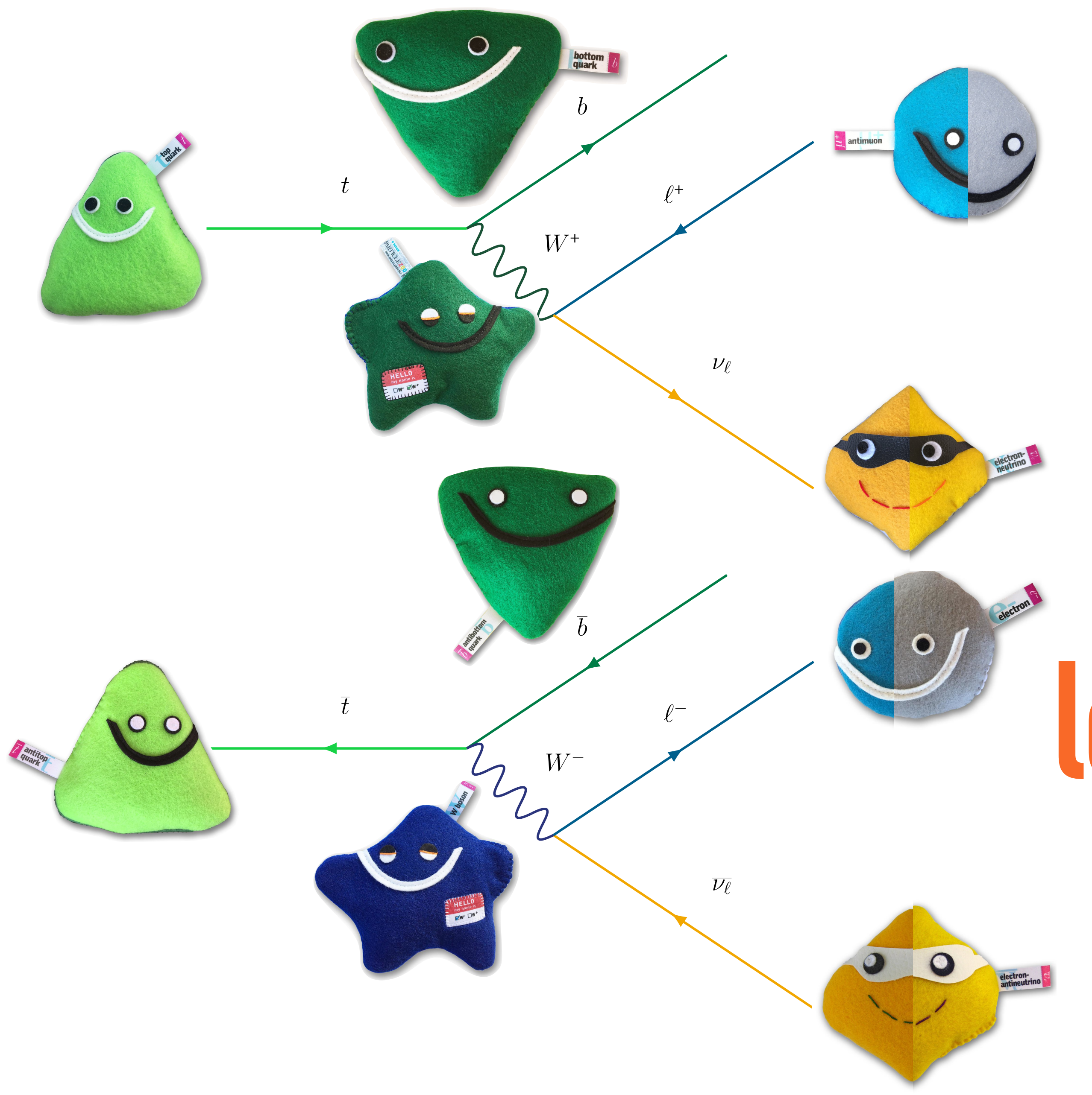


UNIVERSITY OF SUSSEX

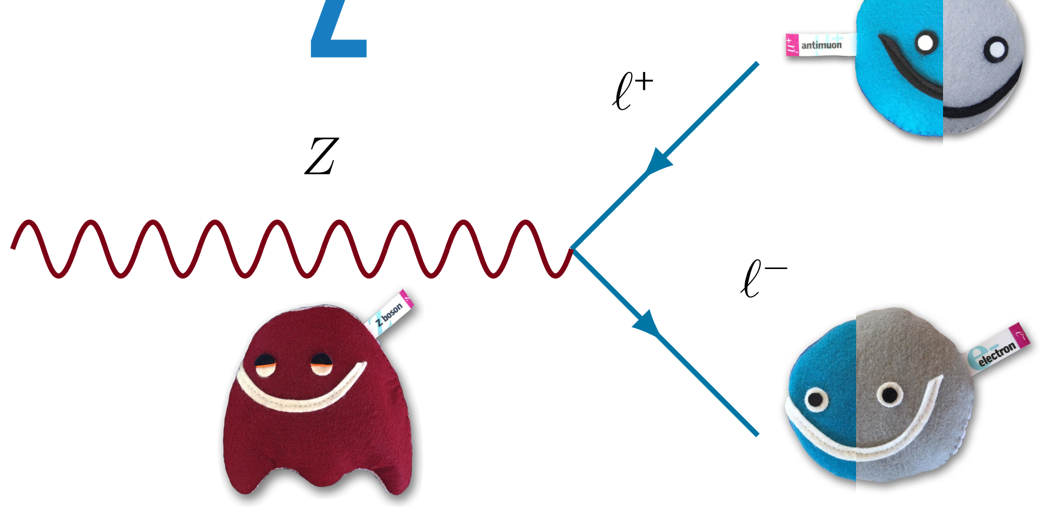
ATLAS EXPERIMENT

SIGNAL VS BACKGROUND

$t\bar{t}$

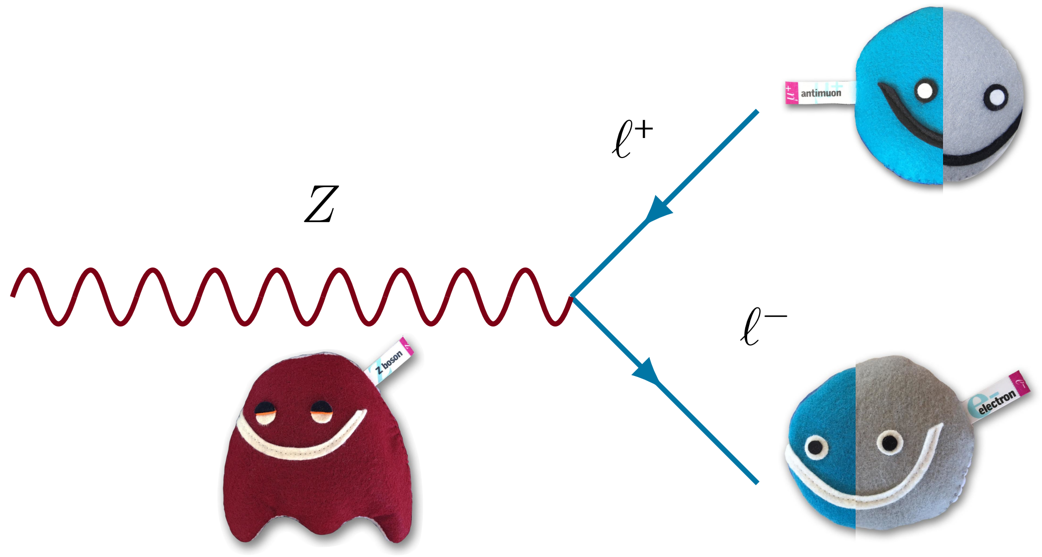


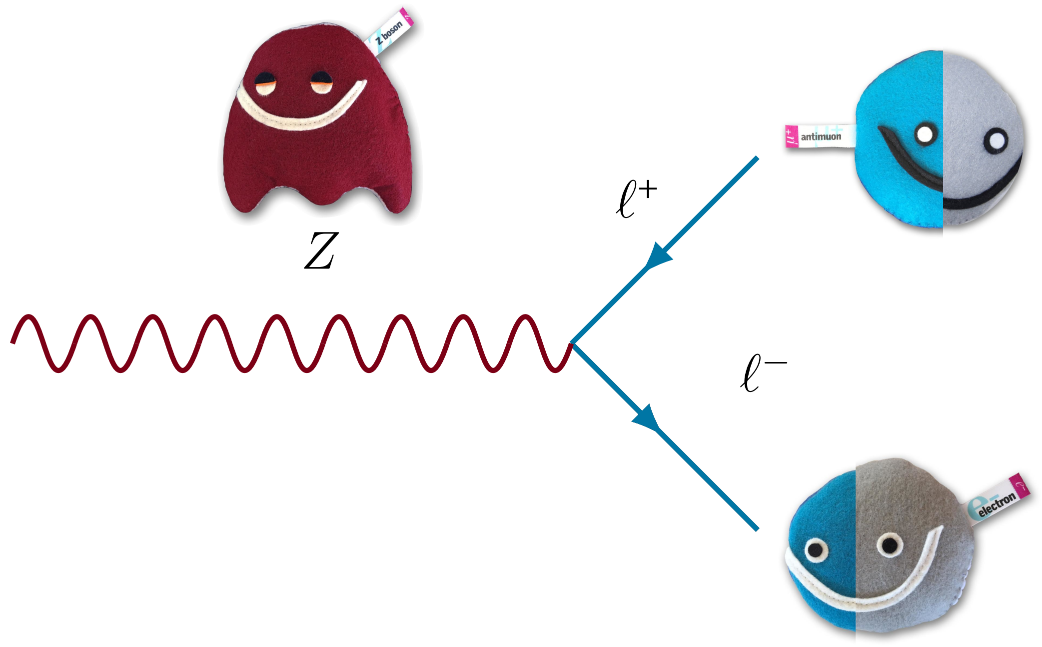
Z



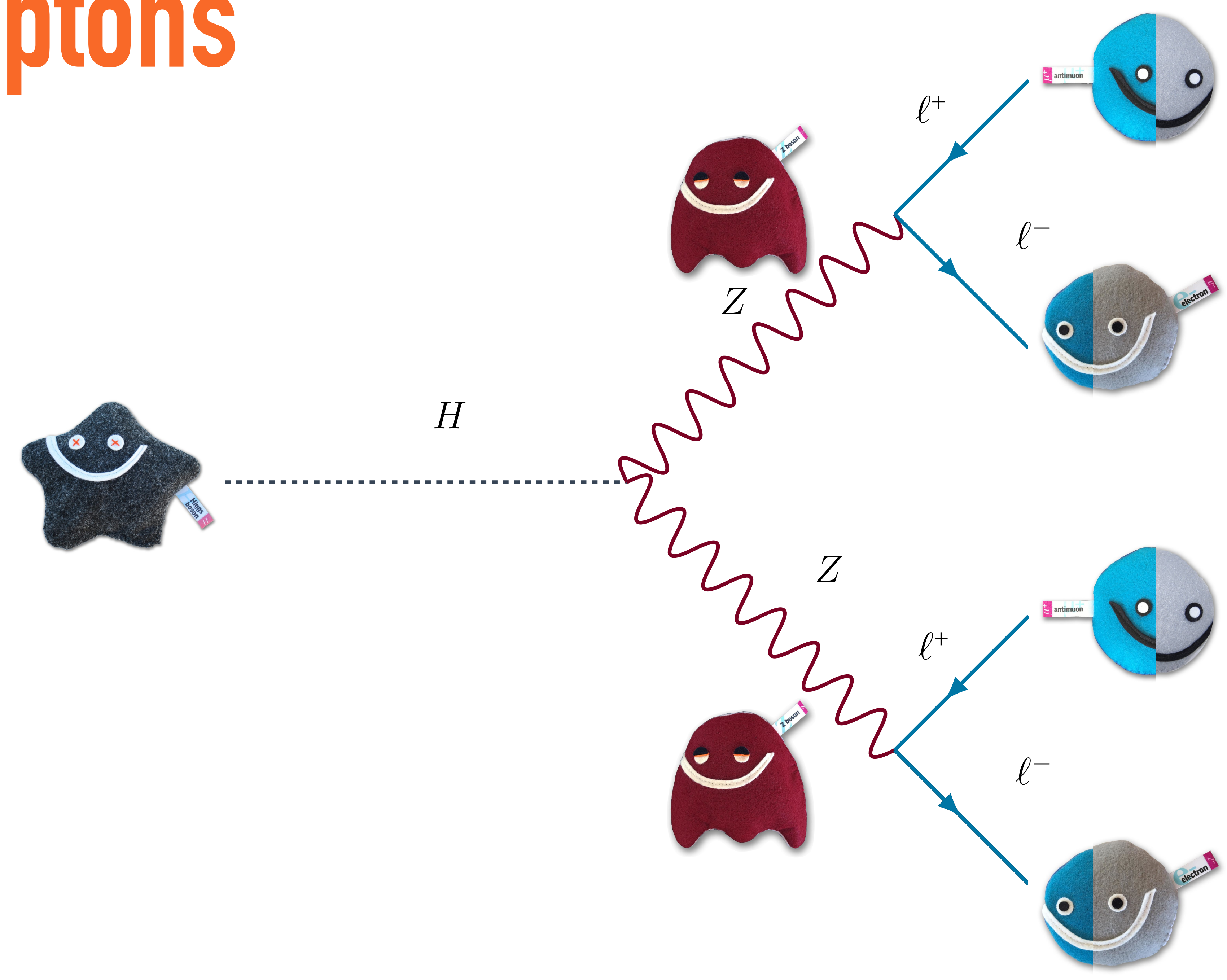
2 leptons

4 leptons





ZZ



H → ZZ

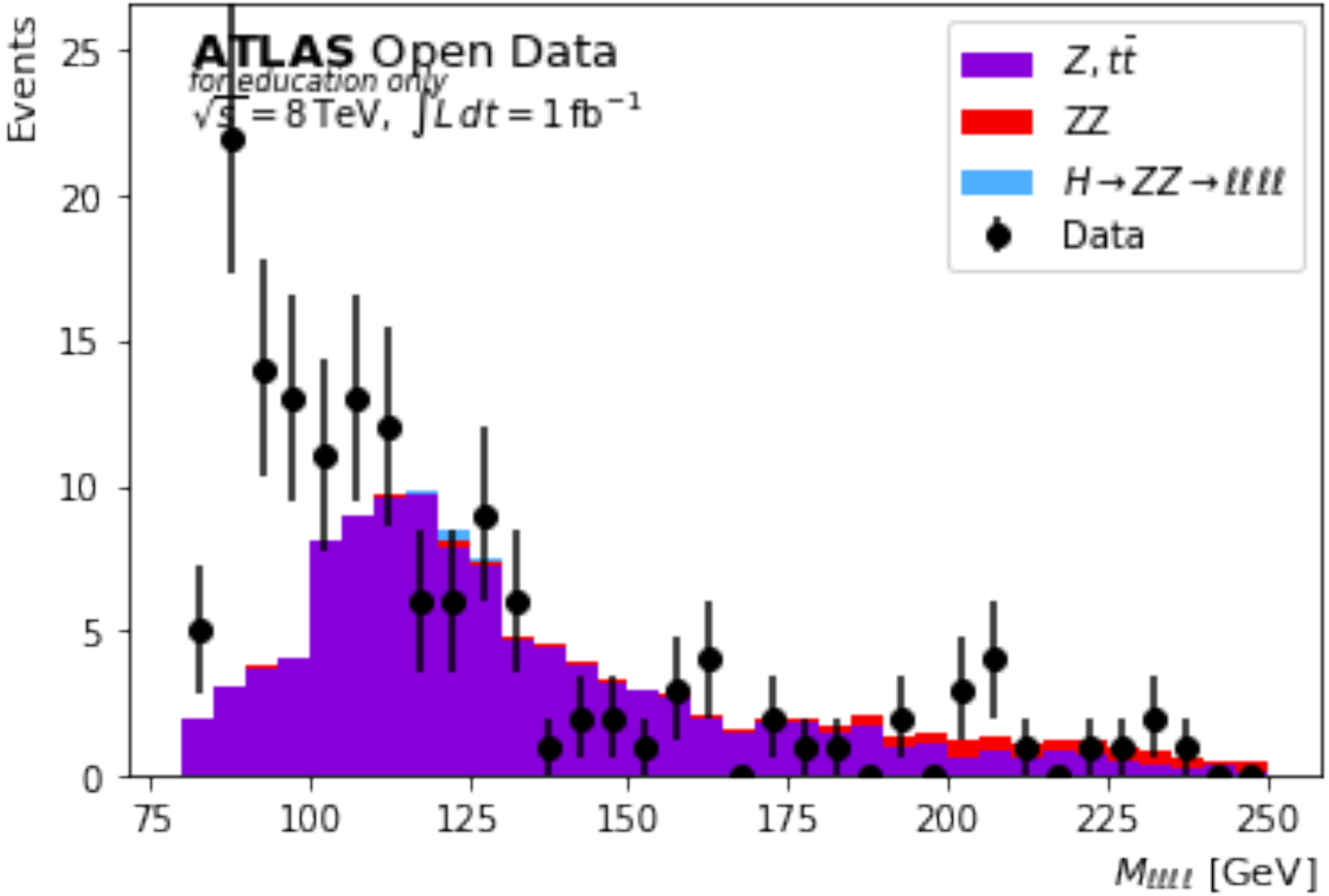
UNIVERSITY OF SUSSEX

ATLAS EXPERIMENT

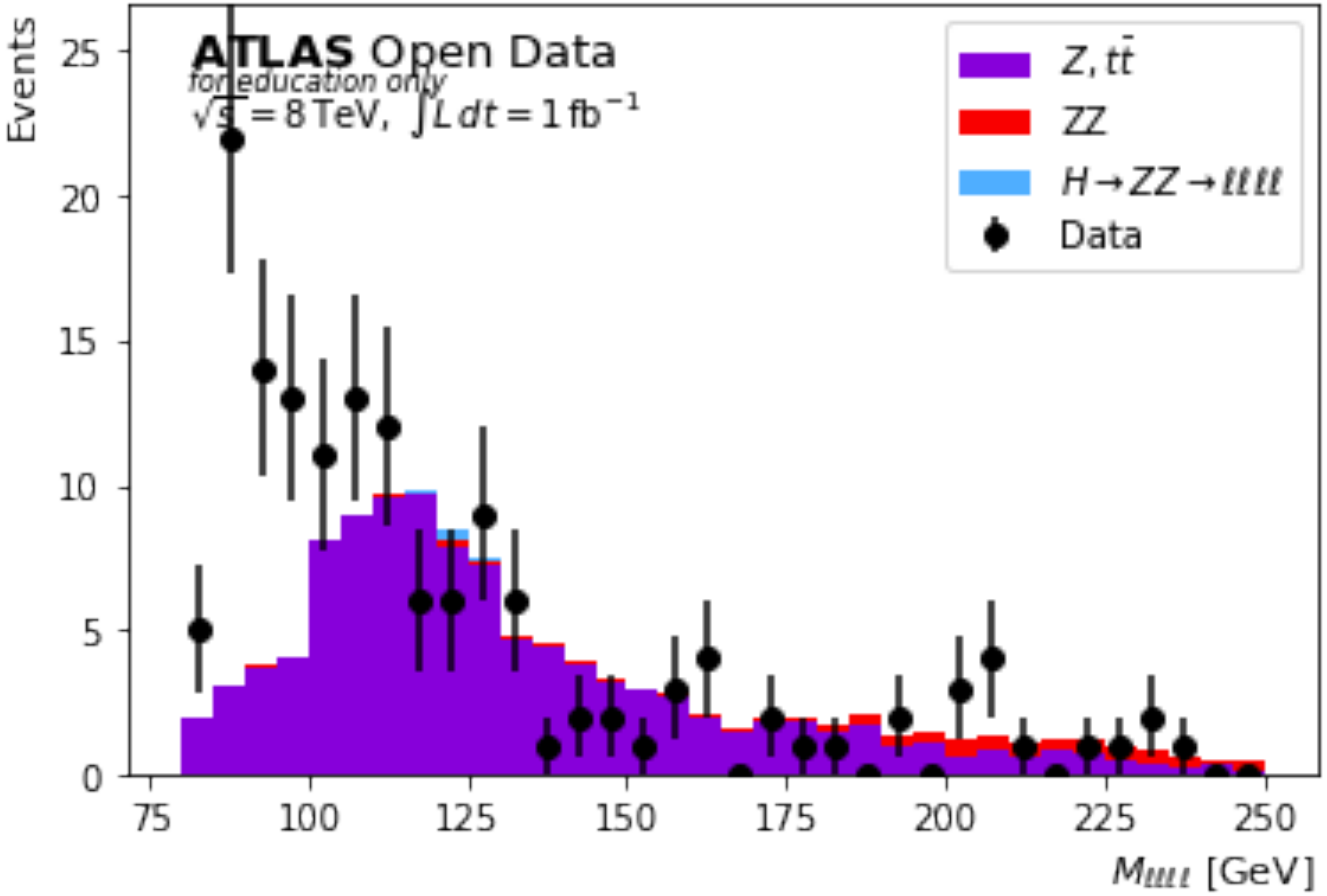
Run the template analysis

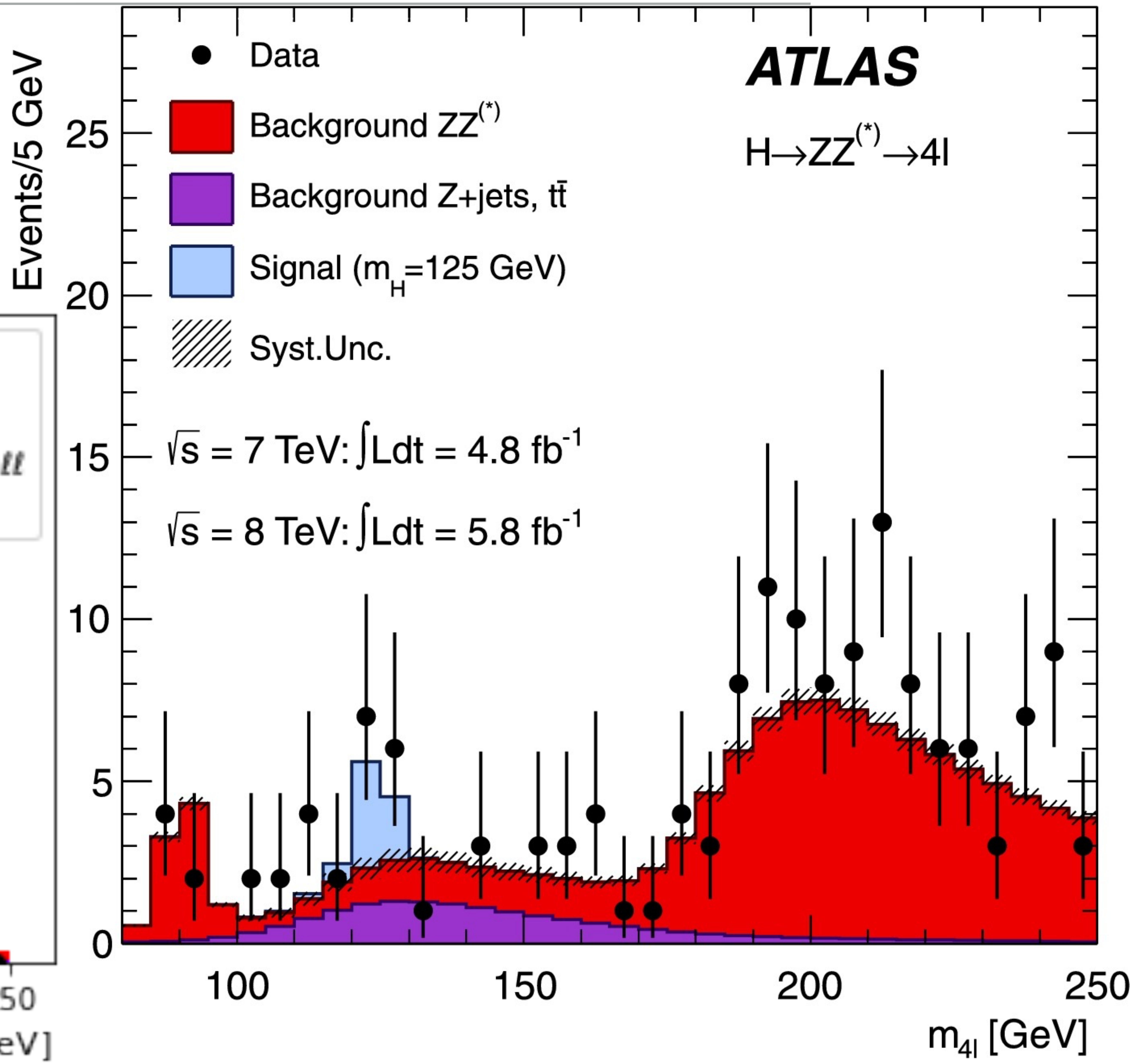
- Go to github.com/meevans1/How-to-rediscover-the-Higgs
- Follow the instructions to open the Jupyter notebook
- Cell -> Run All

Output from template analysis



Output from template analysis





Cut on more variables

- Variables described on the ATLAS Open Data website
- Try add your own cuts to improve the signal / background ratio
- I'll be updating the live google doc with hints

Competition time!

- Submit your team's notebook + its output for entry into our competition
- Winners' code will be modified to use for the upcoming release of new 13 TeV data
  - (+ a prize at the dinner Thursday)
- Instructions on the GitHub code page on how to enter

UNIVERSITY OF SUSSEX

ATLAS EXPERIMENT

Judgement criteria

- In this order:
  - Something crazy like Machine Learning
  - Code elegance
  - Number of cuts correctly implemented
  - Total elapsed time
  - Plot
  - $\text{Signal}/\sqrt{\text{Background}}$

UNIVERSITY OF SUSSEX

ATLAS EXPERIMENT

Diolch! (Thanks!)

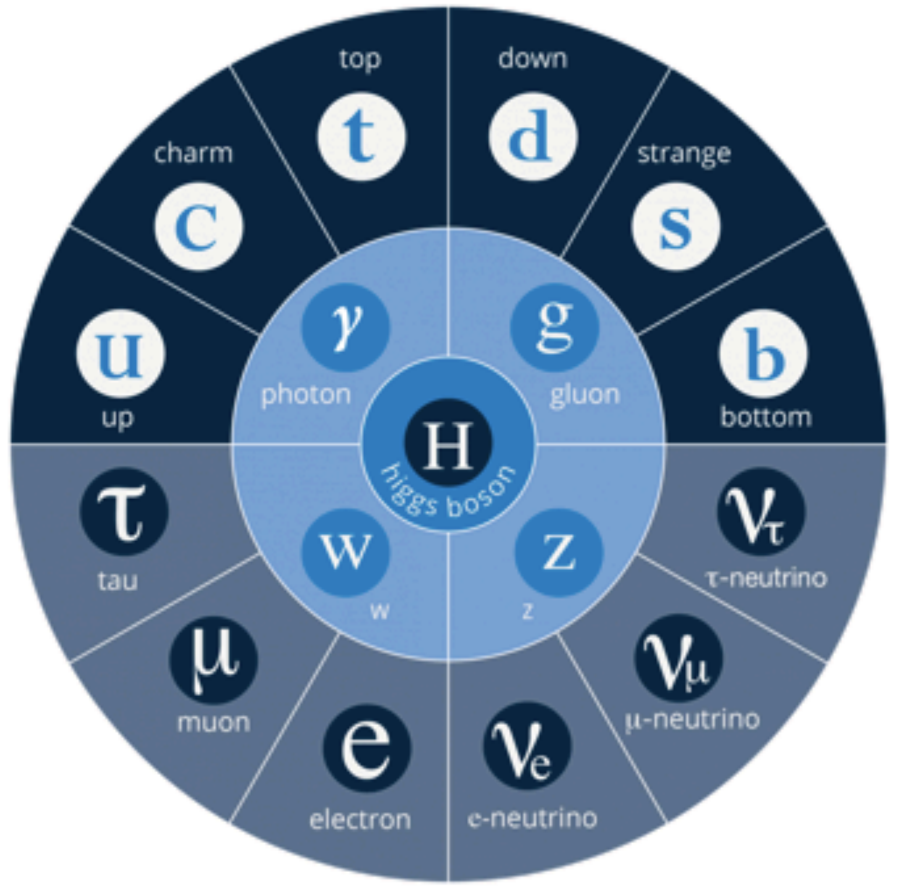
- From this hackathon I hope you've got an insight into:
  - How Python, Pandas, Numpy & Matplotlib can be used for (really) big data analysis
  - How interesting (& challenging) particle physics can be
  - How the Higgs boson was discovered
  - How you can go onto rediscover the Higgs boson
  - How to (re-)win a Nobel Prize

meirin.oan.evans@cern.ch
@meirinoanevans
@meirinoanevans
Meirin Oan Evans

DIOLCH! (THANKS!)

ATLAS EXPERIMENT

Experimental Particle Physics



What are the building blocks of matter?

What are the forces between them?

What happened to antimatter?

What's dark matter?

What was the early universe like & how did it evolve?

What about gravity?

Anything else?

UNIVERSITY OF SUSSEX

open data ATLAS

ATLAS data

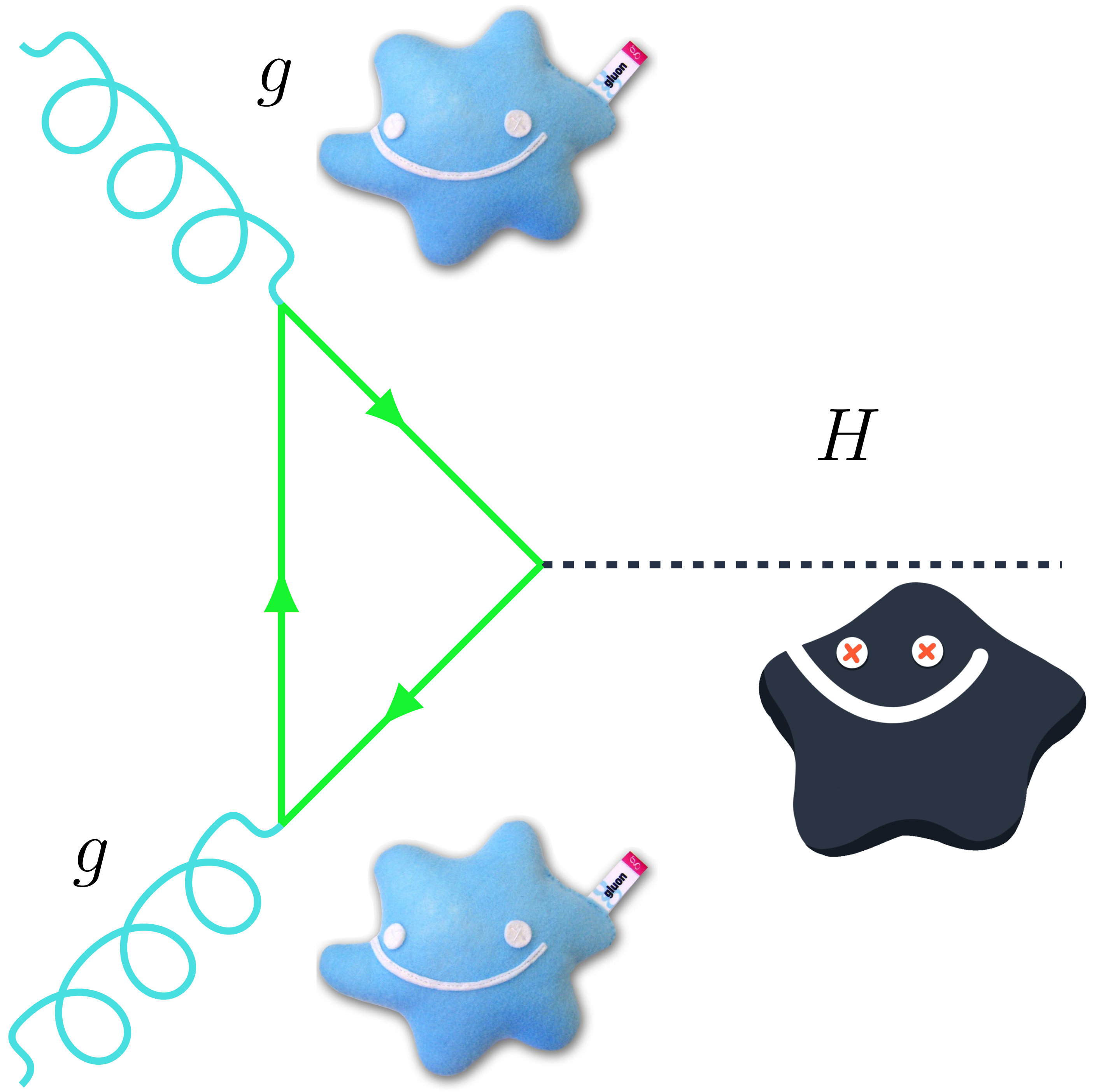
- ATLAS is designed to observe up to 1.7 billion collisions a second
  - (> 60 million MB/s)
  - On a stack of standard 120mm 700MB CDs, it'd stretch from the Earth to the Sun every year!
- So a trigger system selects ~1000 of these collisions s^{-1}
  - Our CD stack now only stretches from Brighton to Oxford every year…

*not to scale

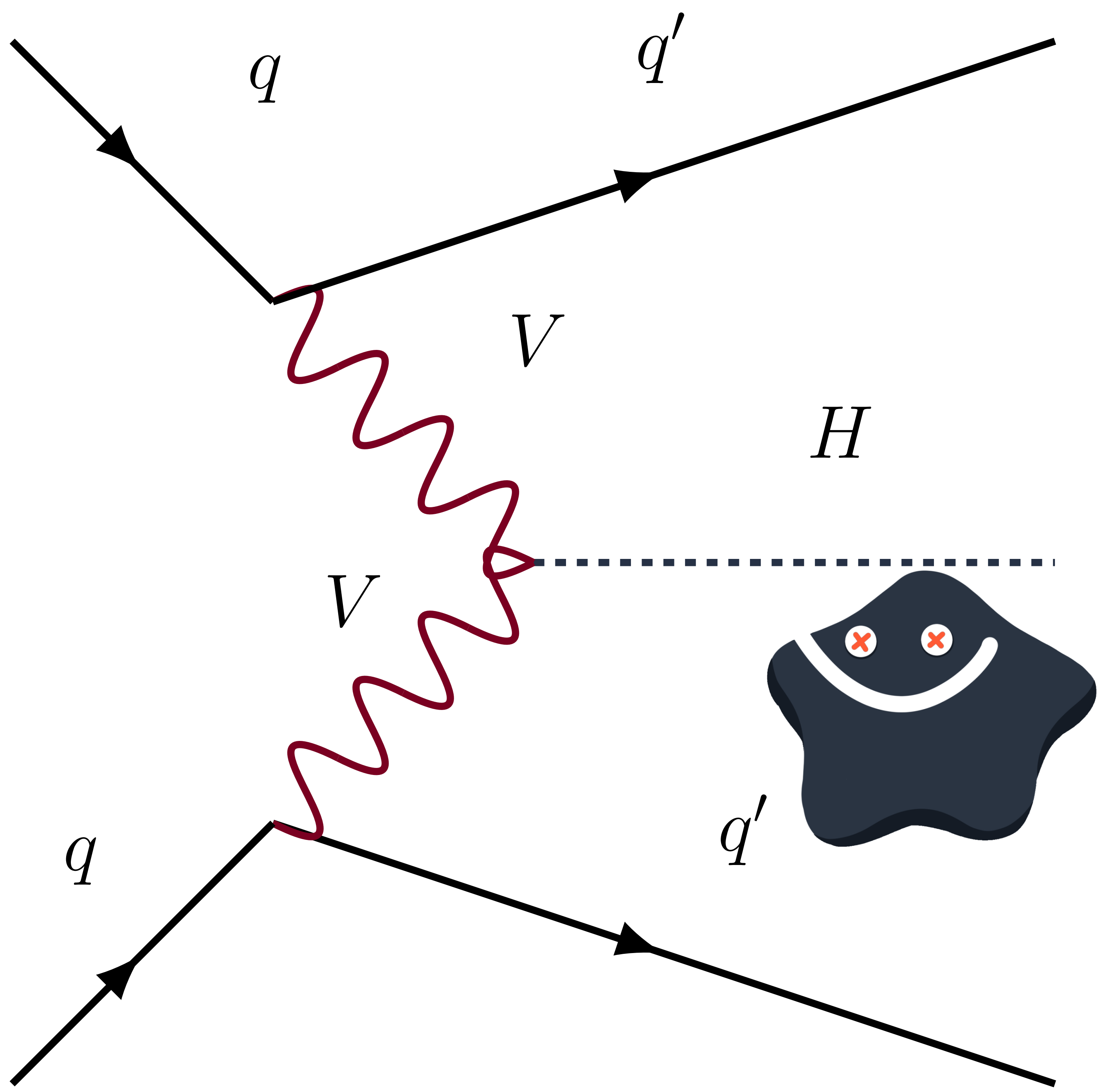
UNIVERSITY OF SUSSEX

open data ATLAS

Higgs production

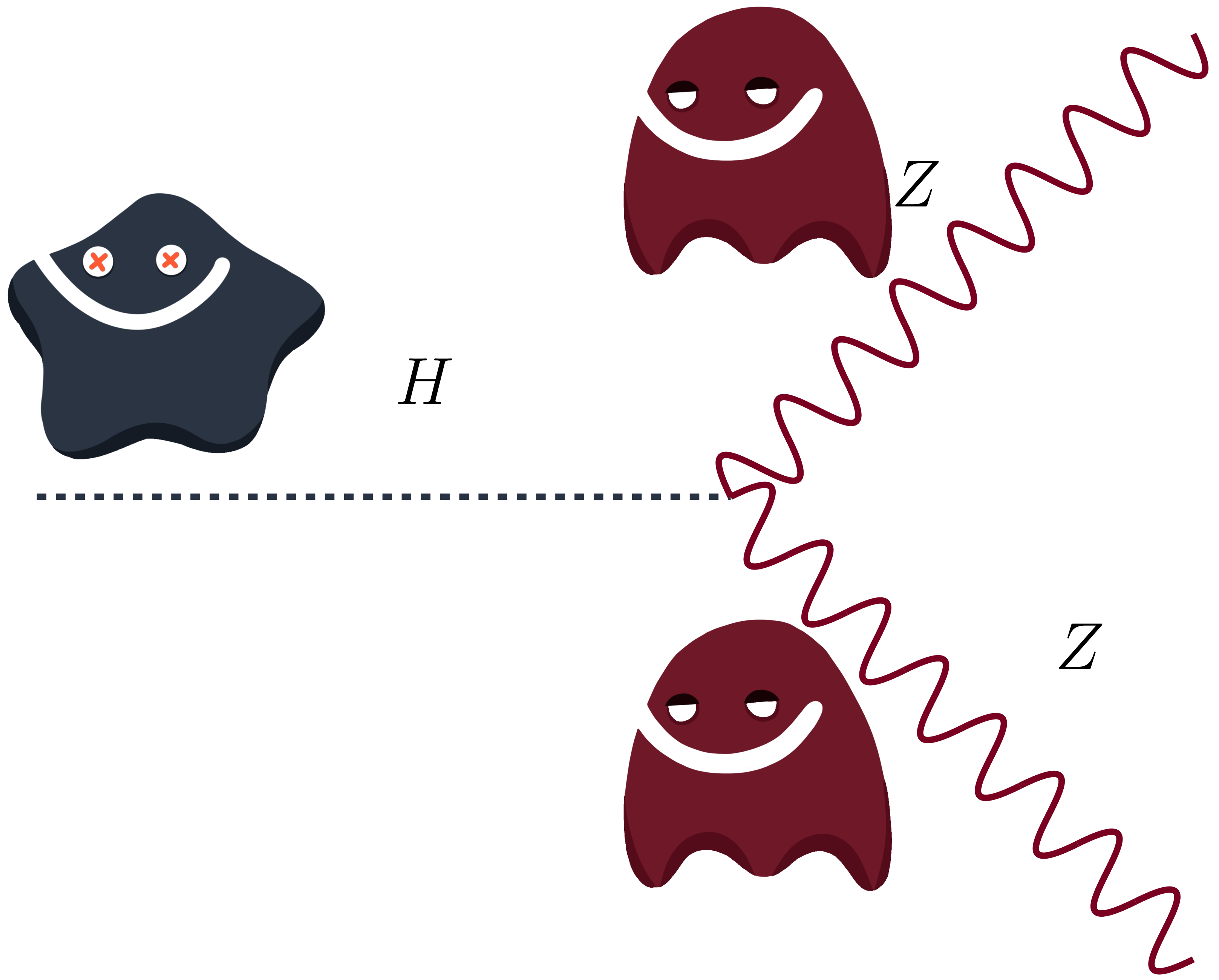


gluon-gluon (gg)

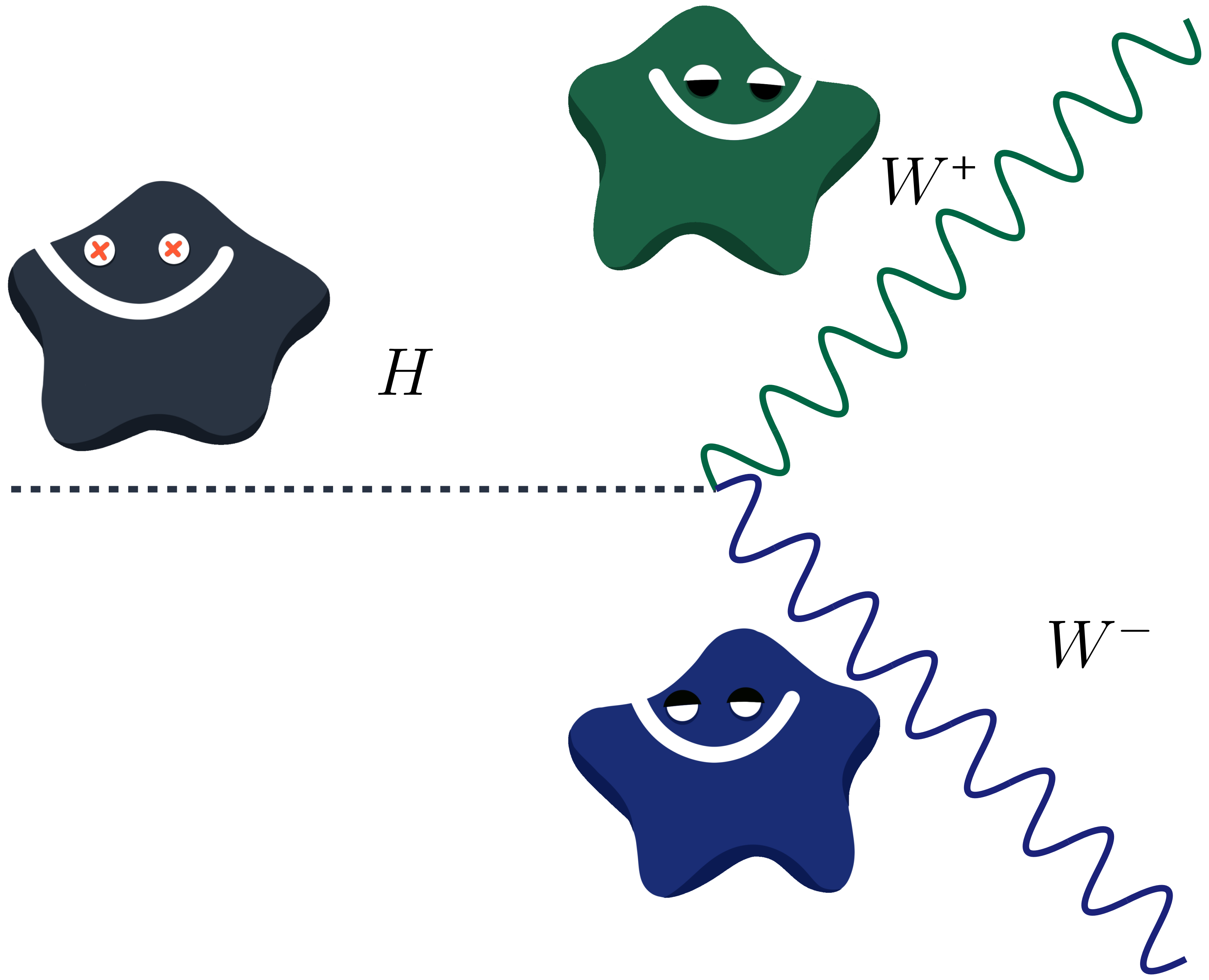


Vector Boson Fusion (VBF)

Higgs decay

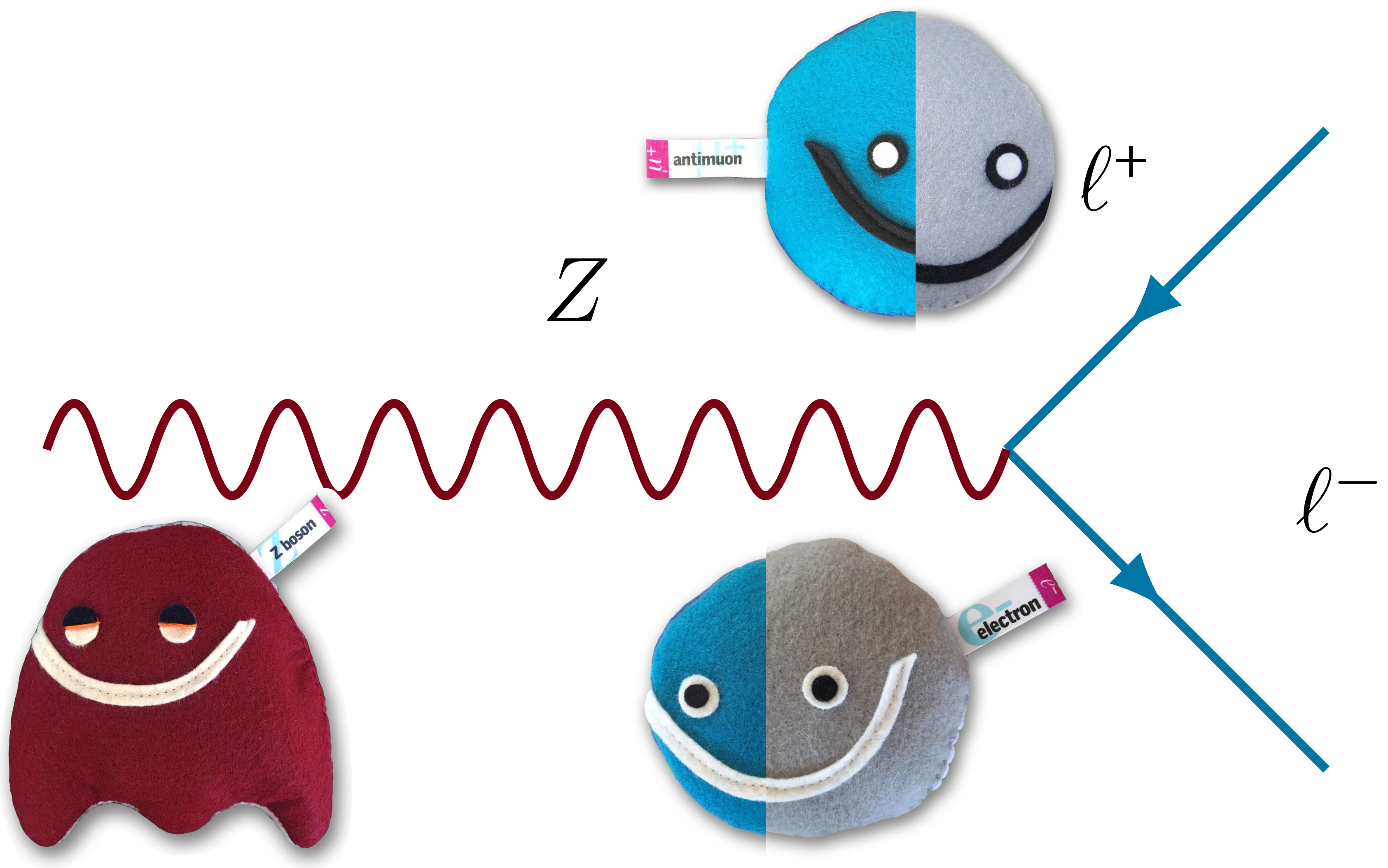


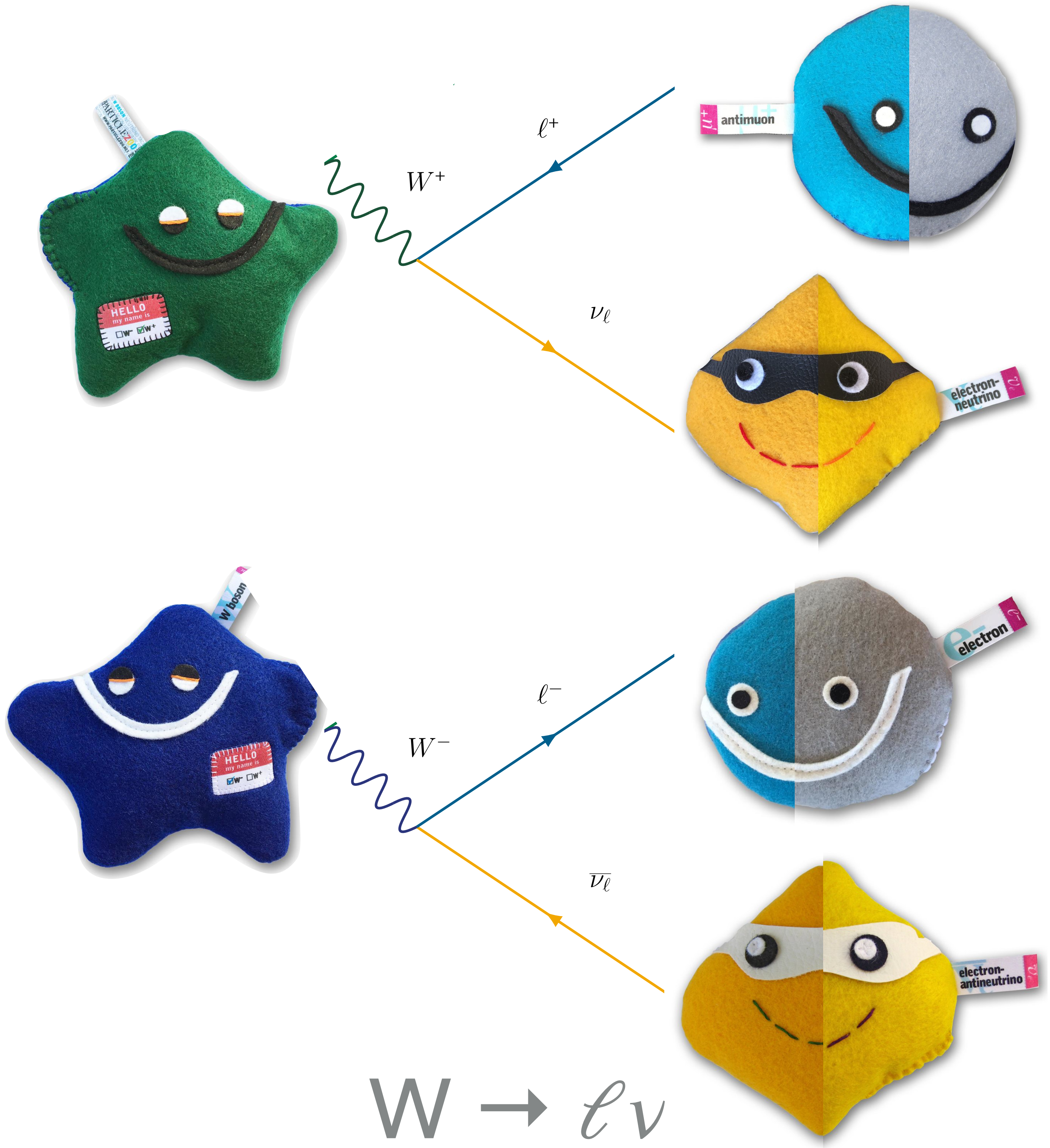
H → ZZ



H → WW

Z & W decay





$Z \rightarrow \ell\ell$

$W \rightarrow \ell\nu$

UNIVERSITY OF SUSSEX

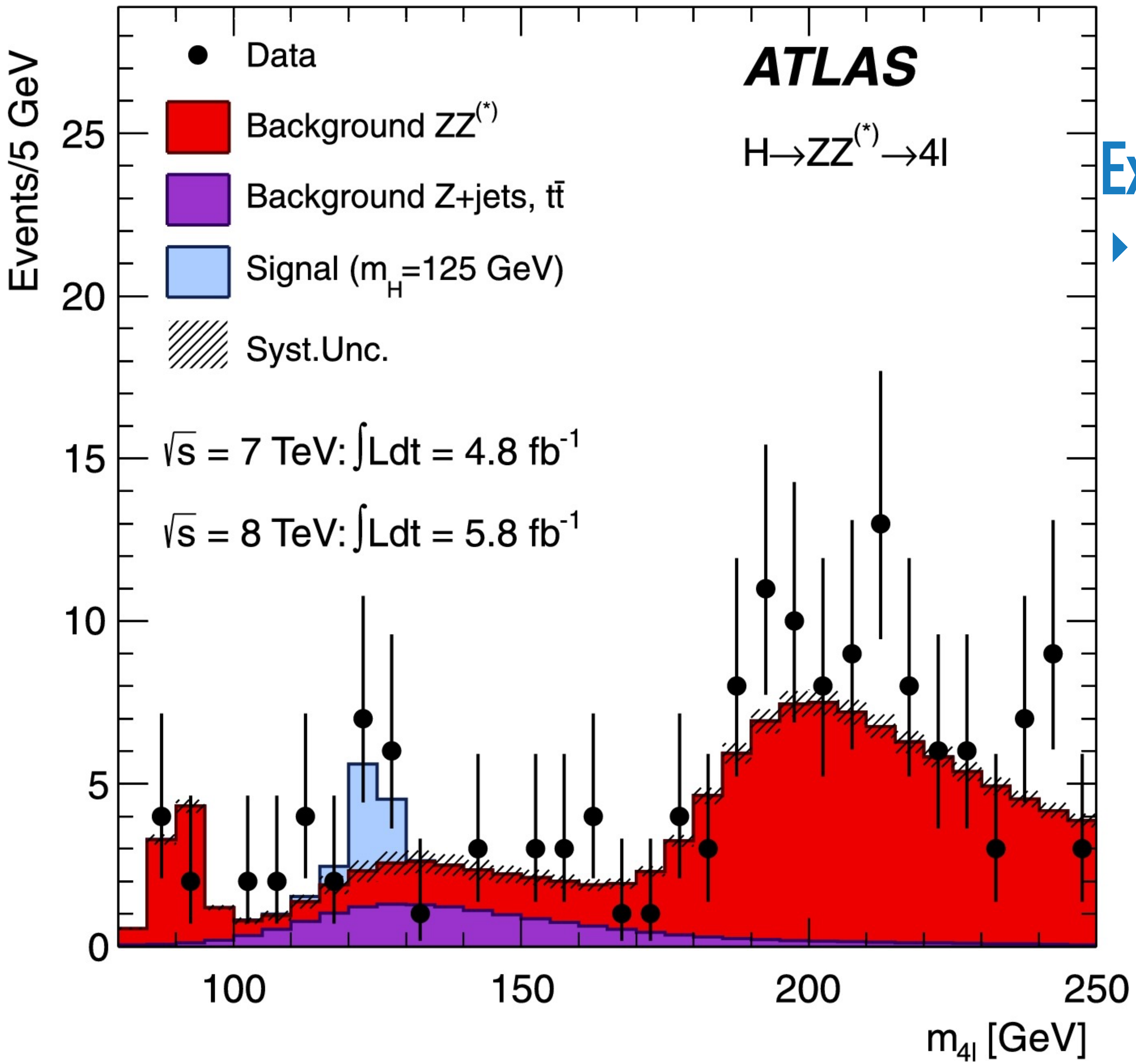
open data ATLAS

Final state particles

- Of the particles we've talked about, only charged leptons ($\ell^{\pm}$) are directly measured by ATLAS
- Higgs, Z, W bosons have lifetimes < 10^{-22} s
  - They decay before being measured
  - They're measured indirectly by their final state particles

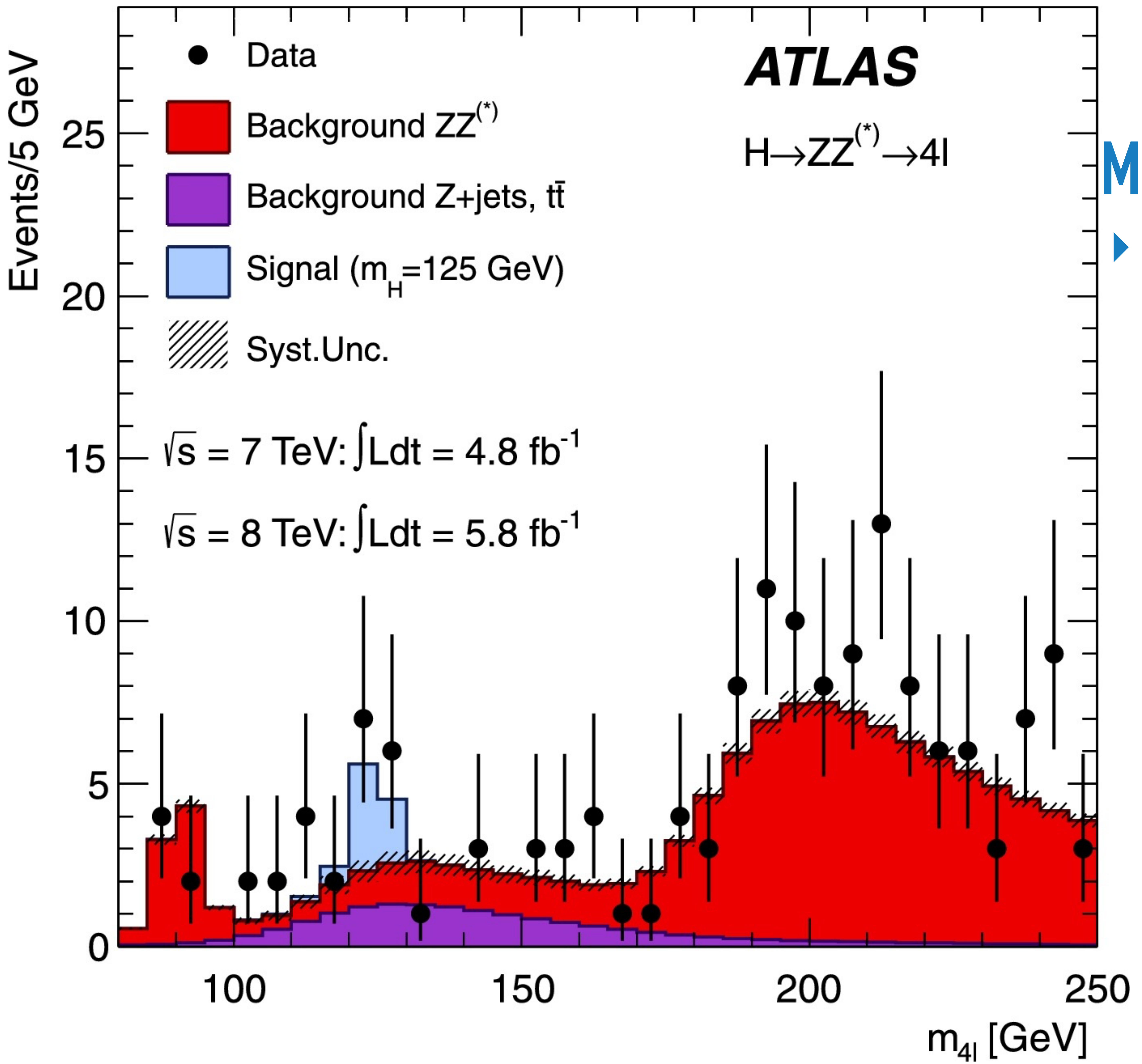
UNIVERSITY OF SUSSEX

open data ATLAS



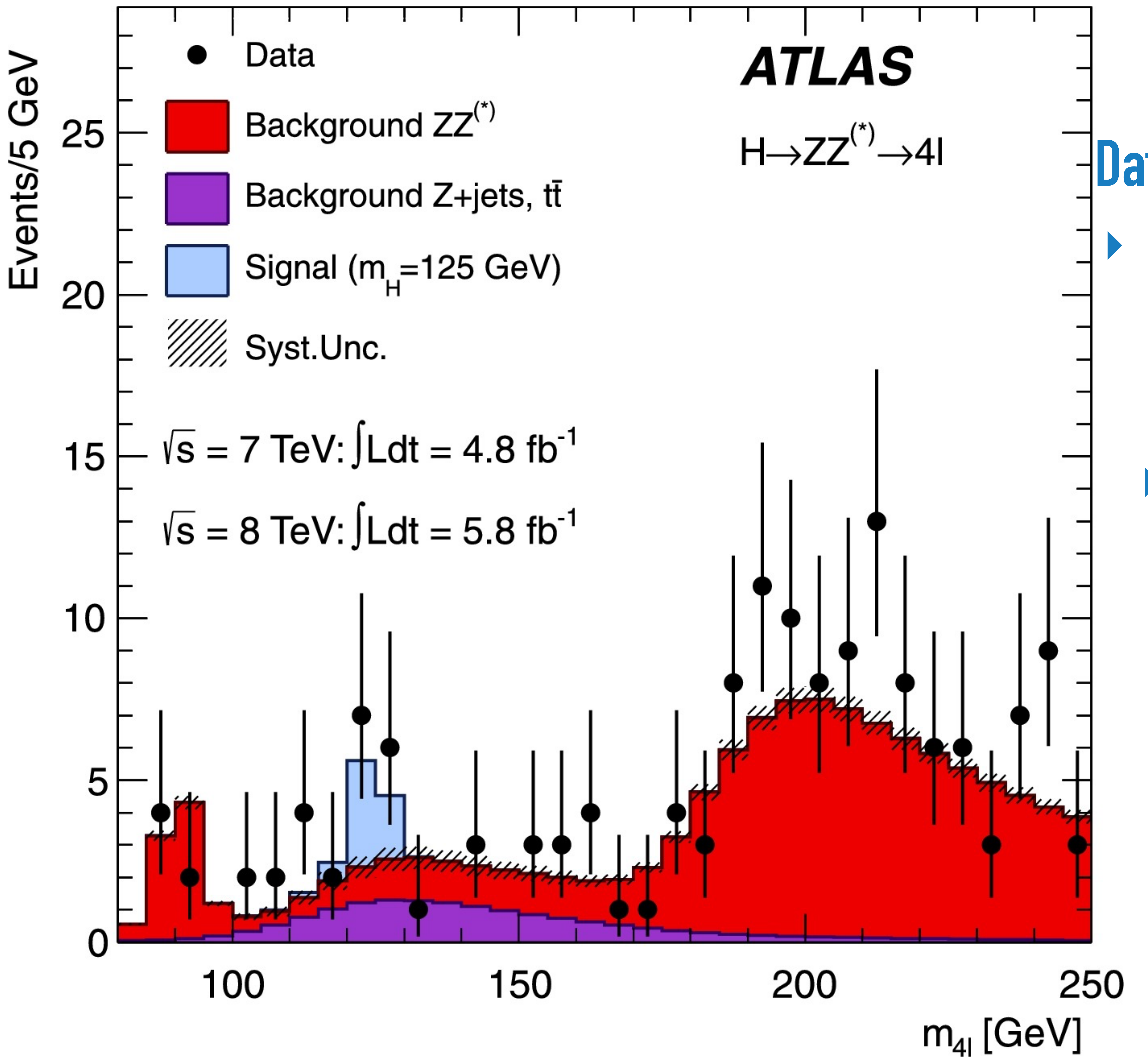
Experimental data

- Black points represent experimental data
  - Statistical uncertainty represented by error bars on the data points



MC simulation

- Filled histograms show the prediction from different MC simulations
  - The contributions are stacked

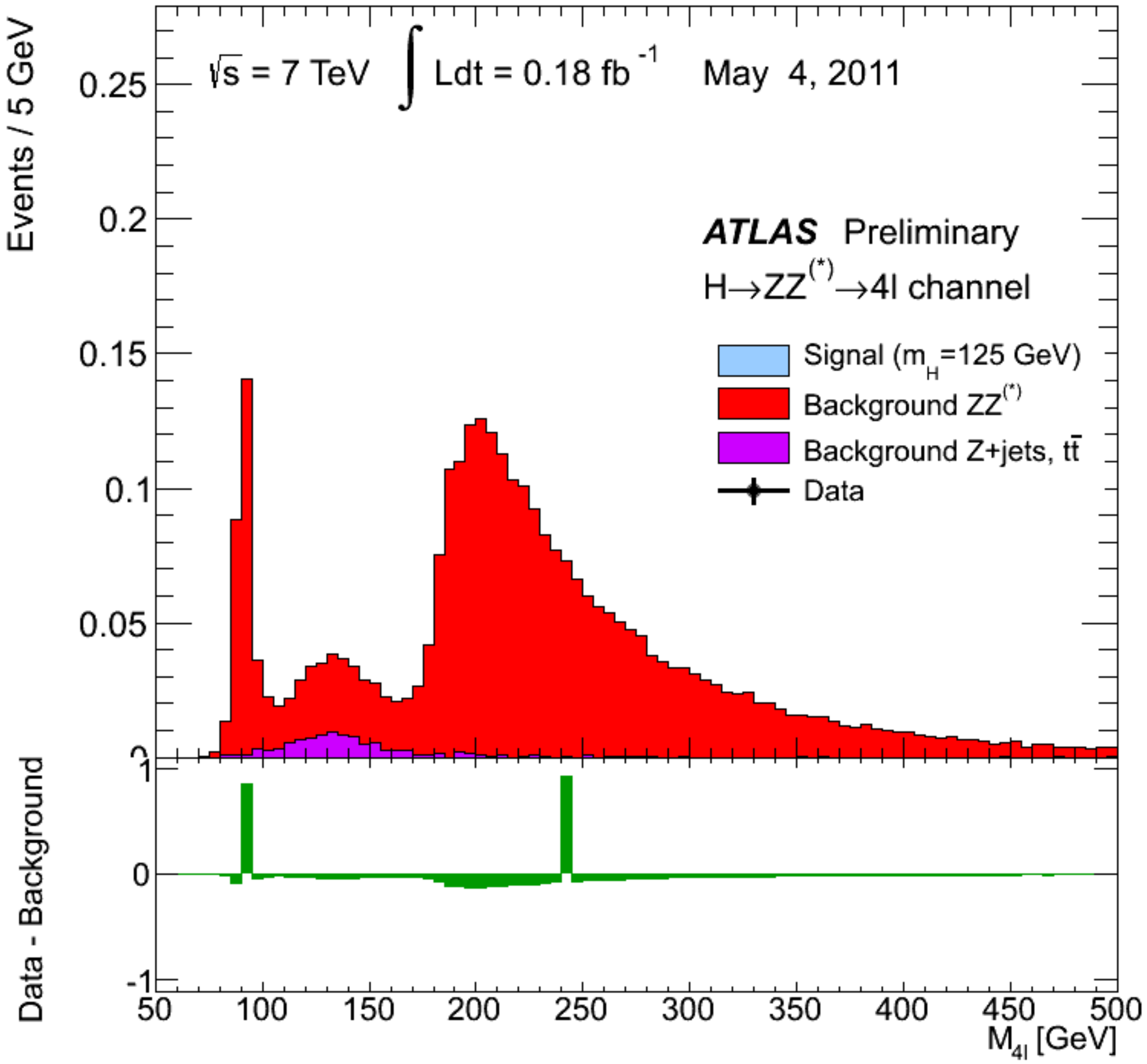


Data/MC discrepancy

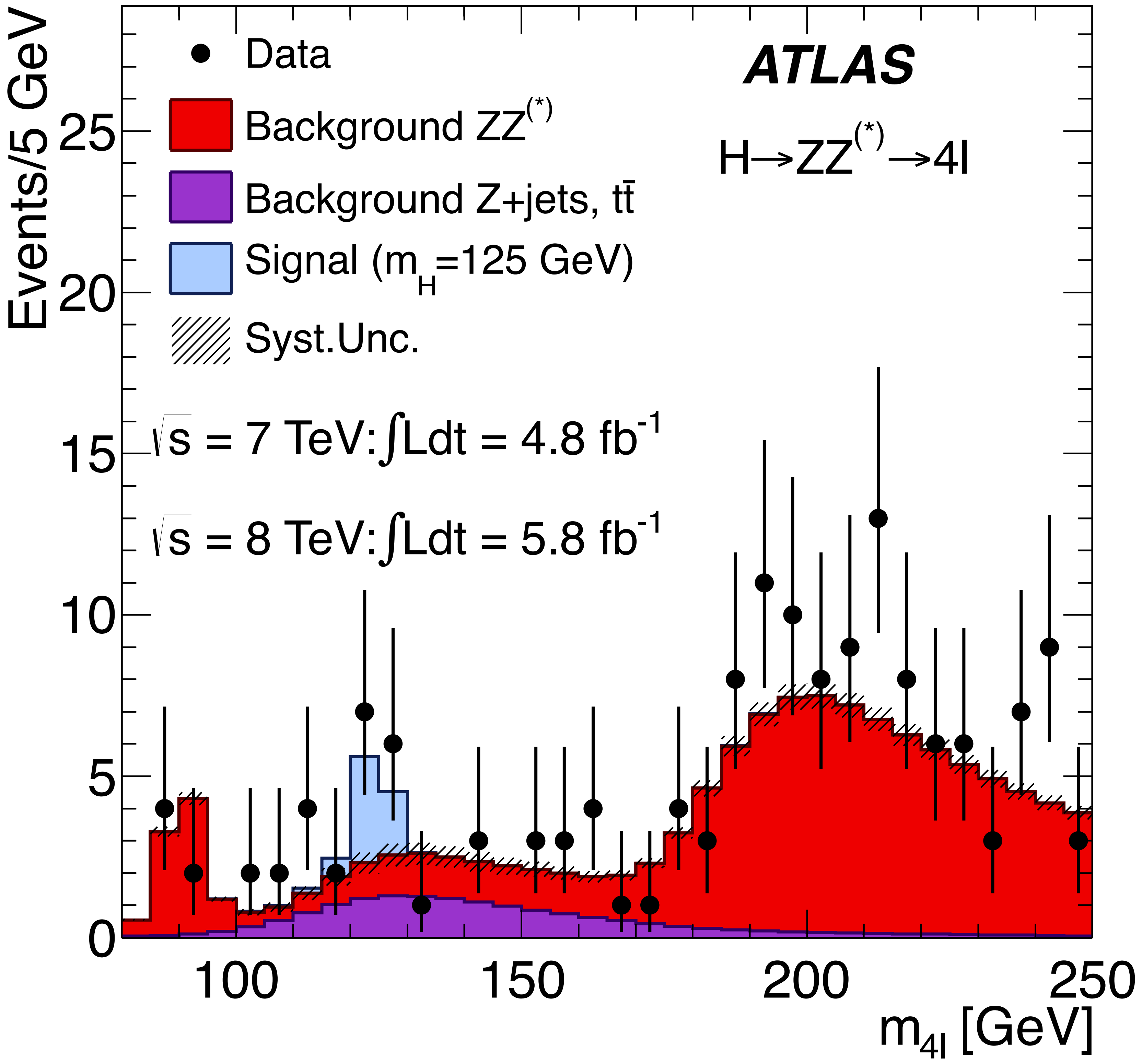
- Any discrepancy between Data & MC is statistical
  - Like line-of-best-fit only passing through 67% of 1σ error bars

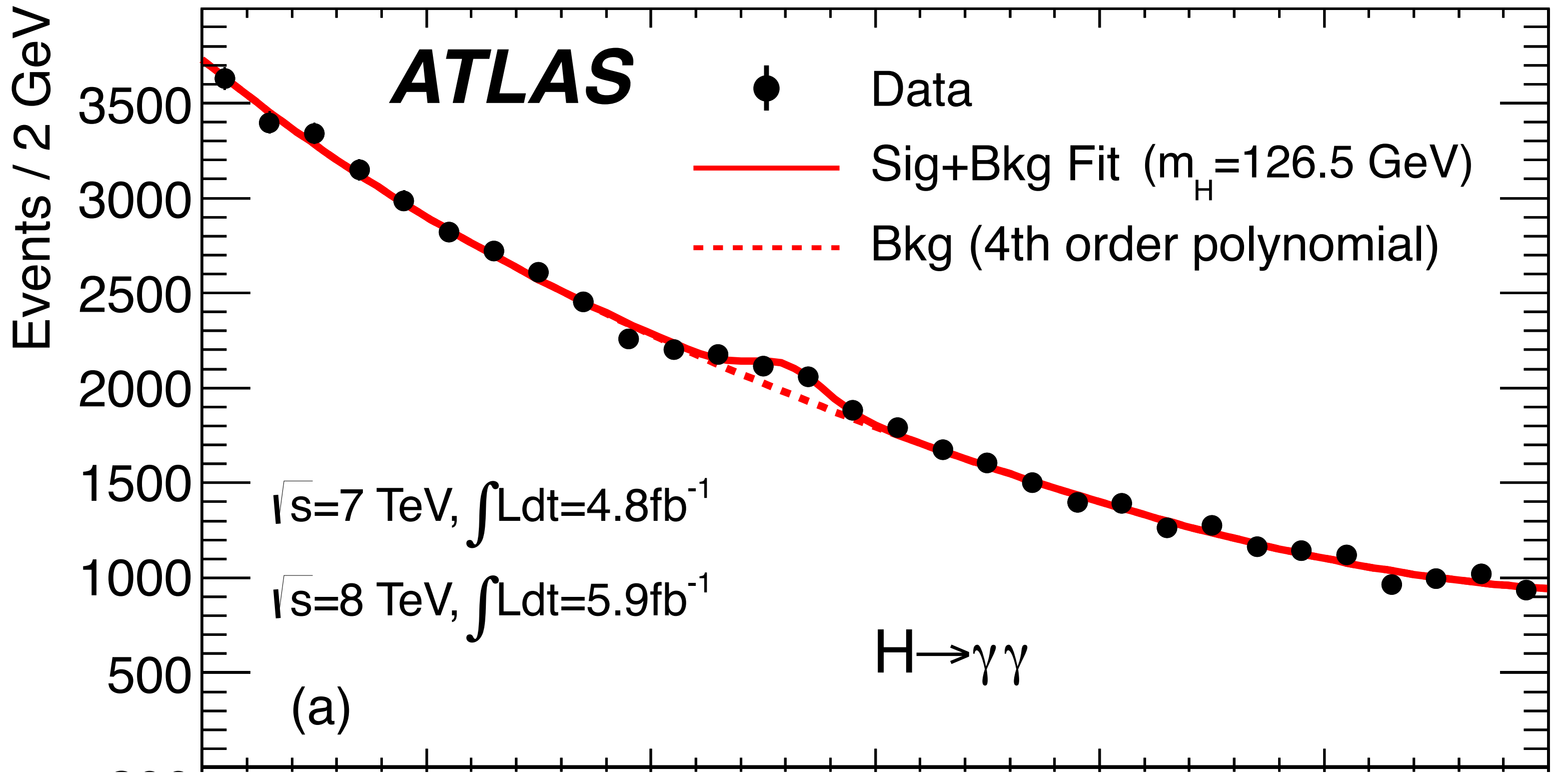
UNIVERSITY OF SUSSEX

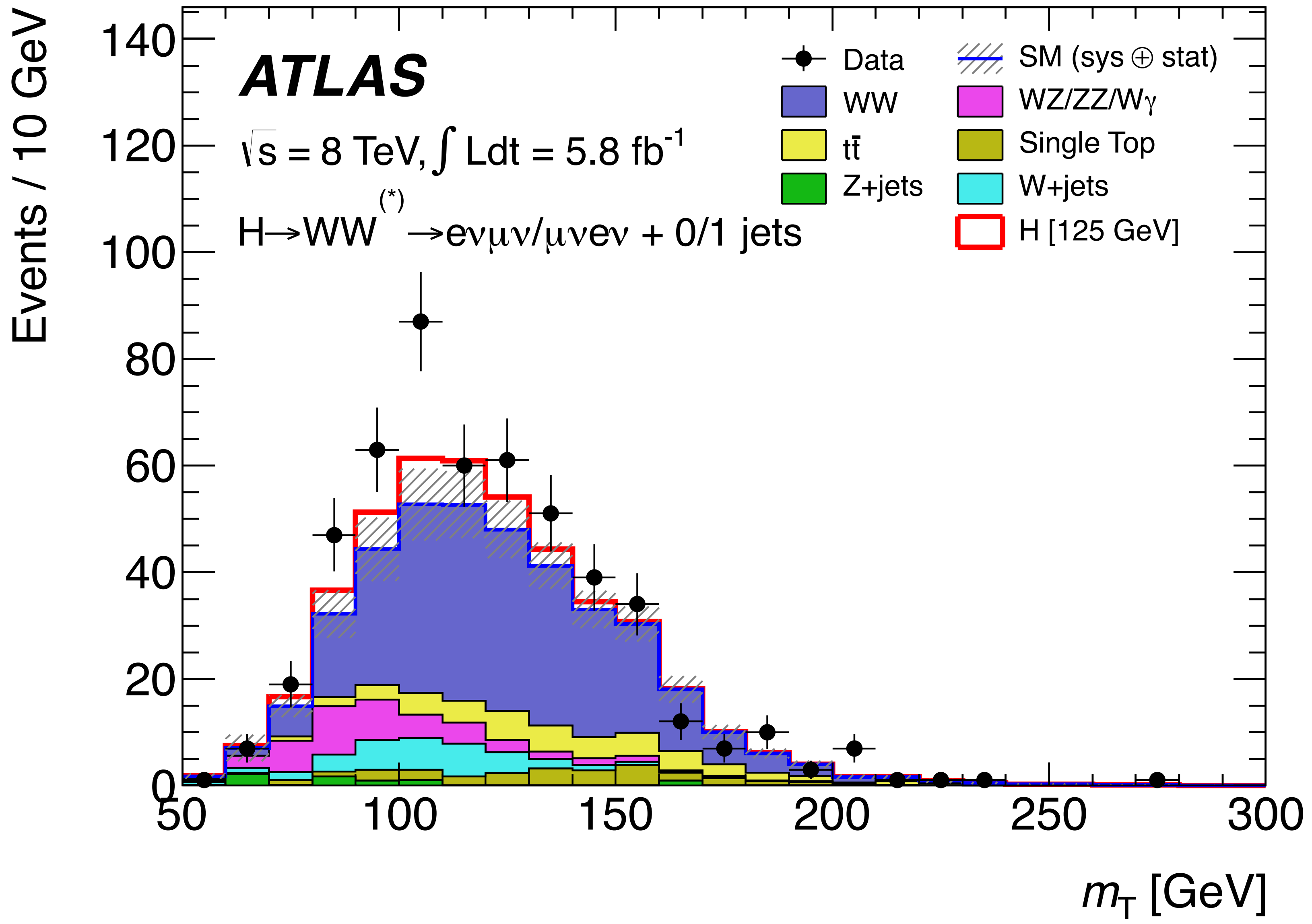
open data ATLAS


Events / 5 GeV
√s = 7 TeV
∫ Ldt = 0.18 fb^-1
May 4, 2011
ATLAS Preliminary
H→ZZ(*)→4l channel
Signal (m_H=125 GeV)
Background ZZ(*)
Background Z+jets, tt̄
Data
Data - Background
M_4l [GeV]
0.25
0.2
0.15
0.1
0.05
1
0
-1
50
100
150
200
250
300
350
400
450
500

How did this lead to a discovery?







- Combining $H \to ZZ$, $H \to \gamma\gamma$, $H \to WW$, gave total significance of 5σ in data over background-only

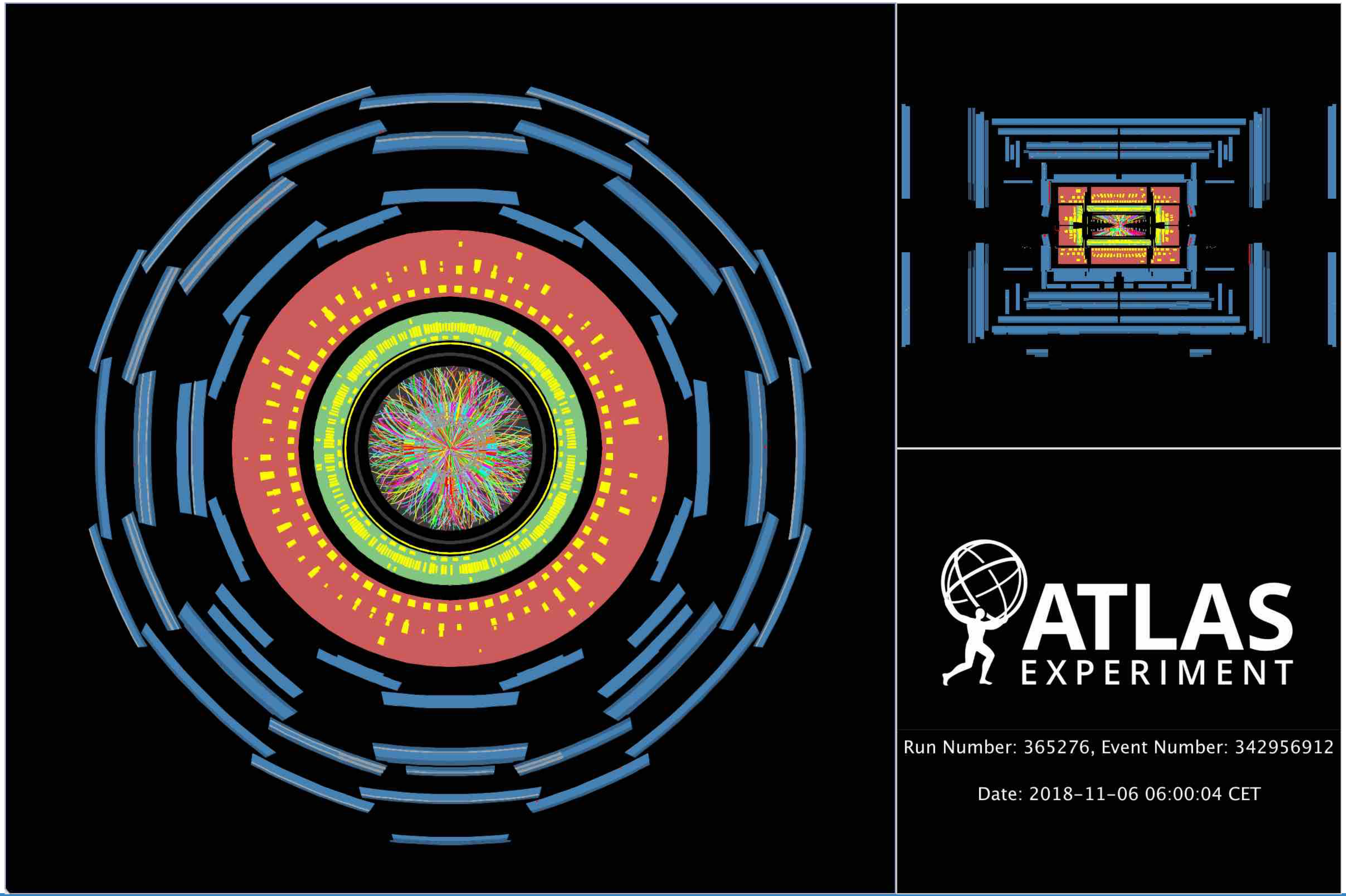
Why do Z & $t\bar{t}$ sometimes have > 2 leptons?

- From Higgs discovery paper: “Non-prompt leptons from heavy flavour decays, electrons from photon conversions and jets-misidentified as electrons…”

➡ Can have more than just the leptons from the leading order Feynman diagrams

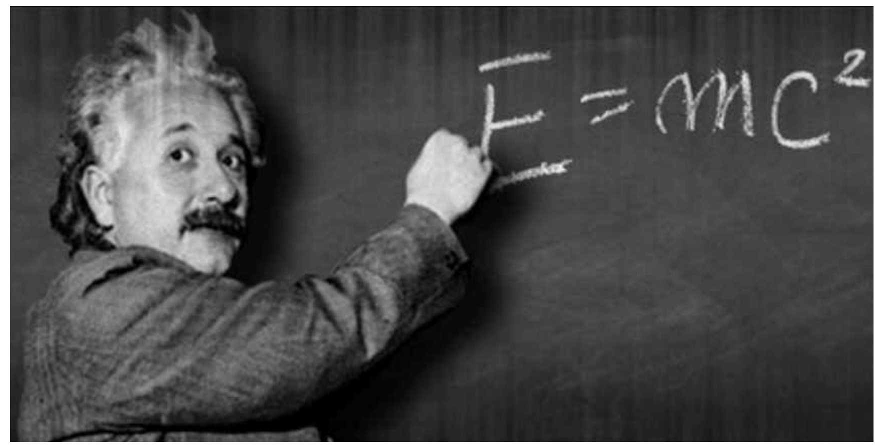
➡ e.g. any charged particle can radiate a photon, which decays into an electron-positron pair (2 extra leptons)

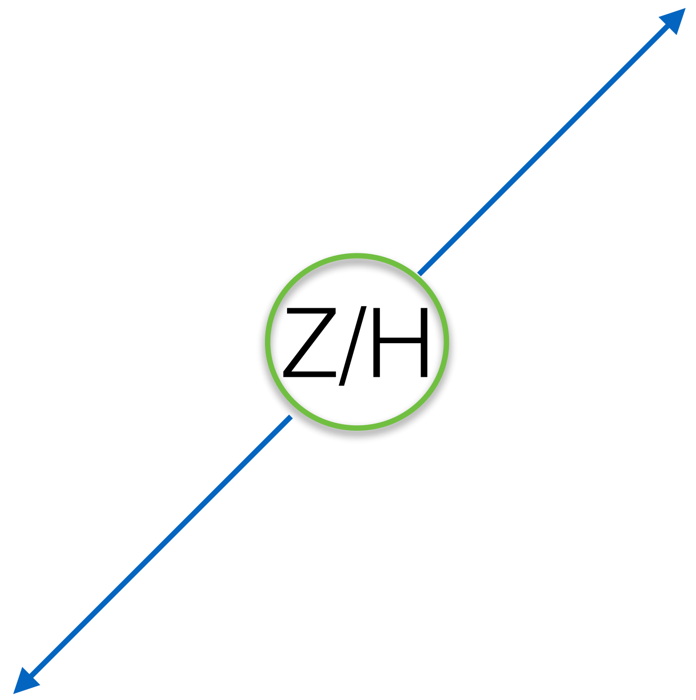
Transverse momentum? Why not momentum?



Transverse momentum? Why not momentum?

- Z and Higgs bosons are ~100x heavier than protons
  - ➡ ~all proton momentum goes to making Z/Higgs rest mass
  - ➡ Z/Higgs produced almost at rest
  - ➡ decay products move ~back-to-back in random direction
  - ➡ products often oriented in transverse plane
- Leptons are ~1000x lighter than Z/Higgs





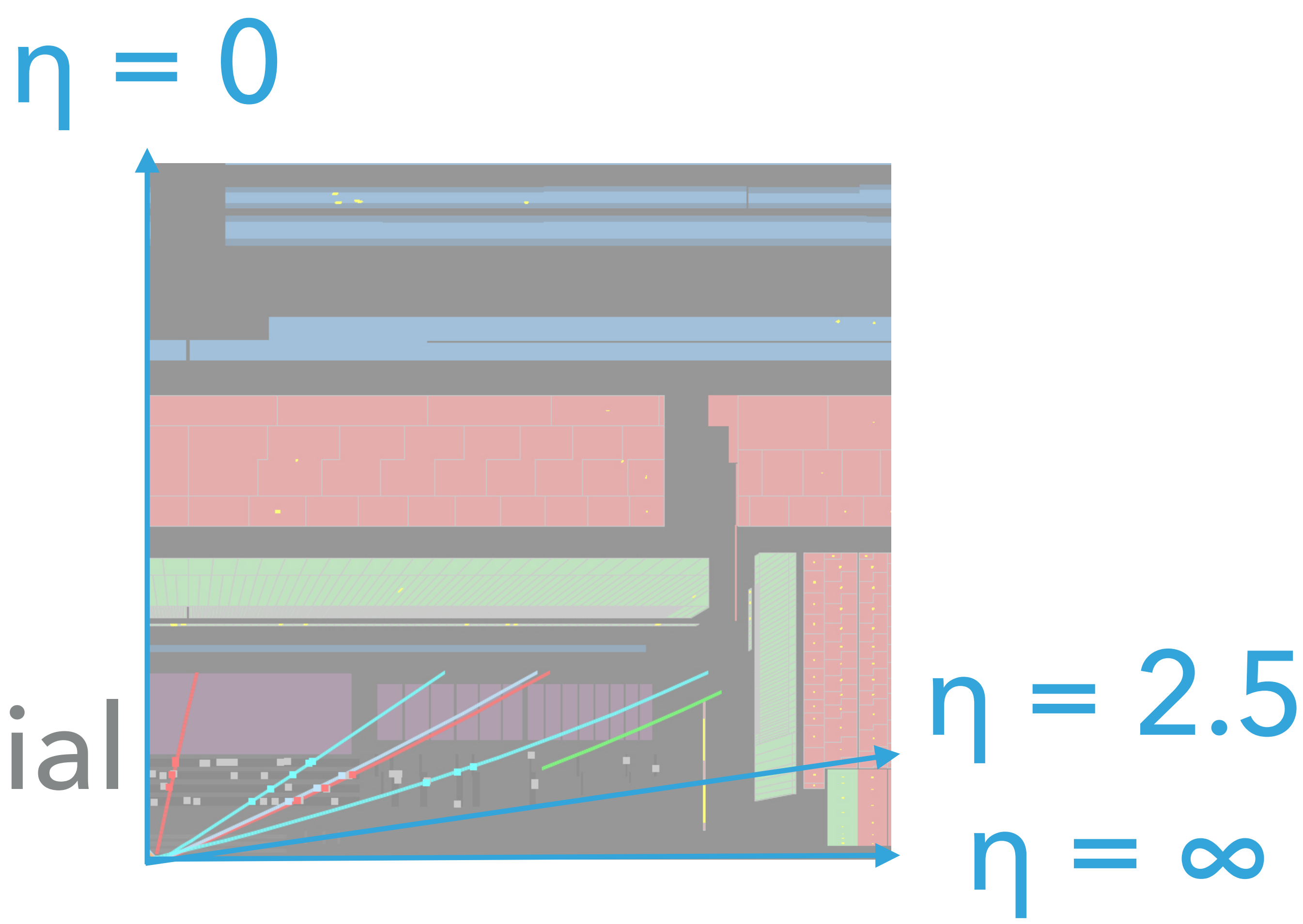
UNIVERSITY OF SUSSEX

ATLAS EXPERIMENT

Pseudorapidity? Why not θ?

- $lep\_eta = -\ln\tan(\frac{\theta}{2})$.
- Differences in pseudorapidity are Lorentz invariant
- Differences in θ aren't

  ➡ difference in pseudorapidity gives handle on spatial separation

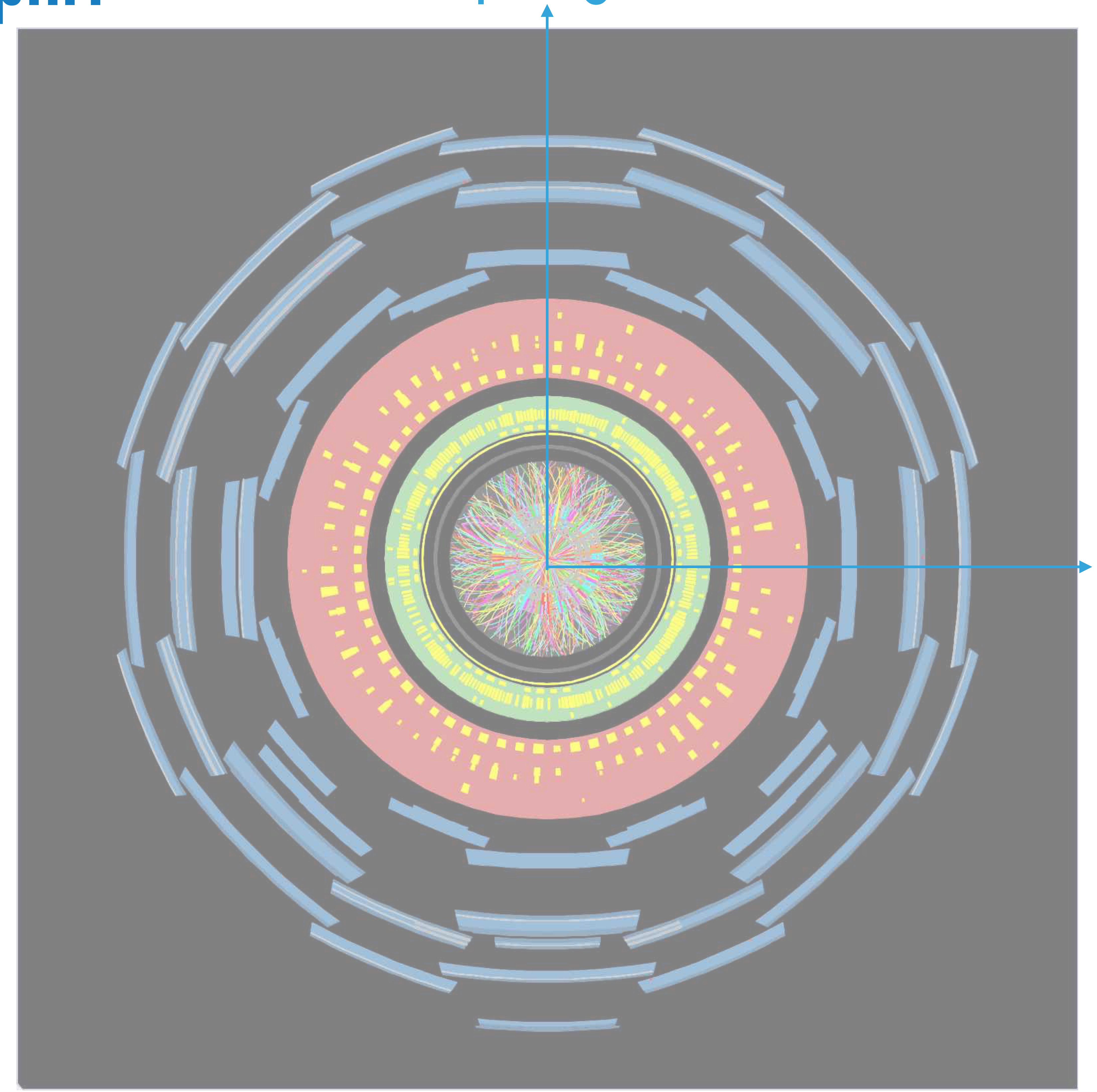


UNIVERSITY OF SUSSEX

ATLAS EXPERIMENT

What's lep_phi?

$\Phi = 0^0$



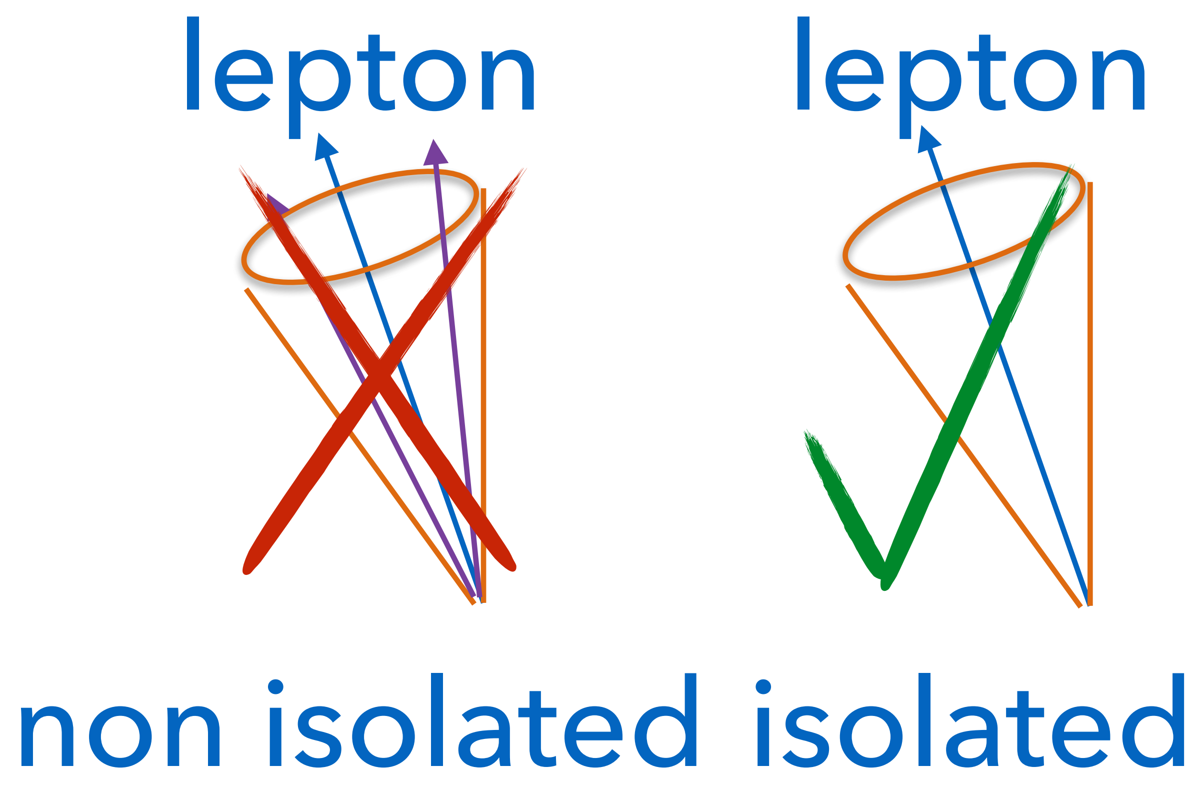
$\Phi = 90^0$

UNIVERSITY OF SUSSEX

ATLAS EXPERIMENT

What's lep_etcone20?

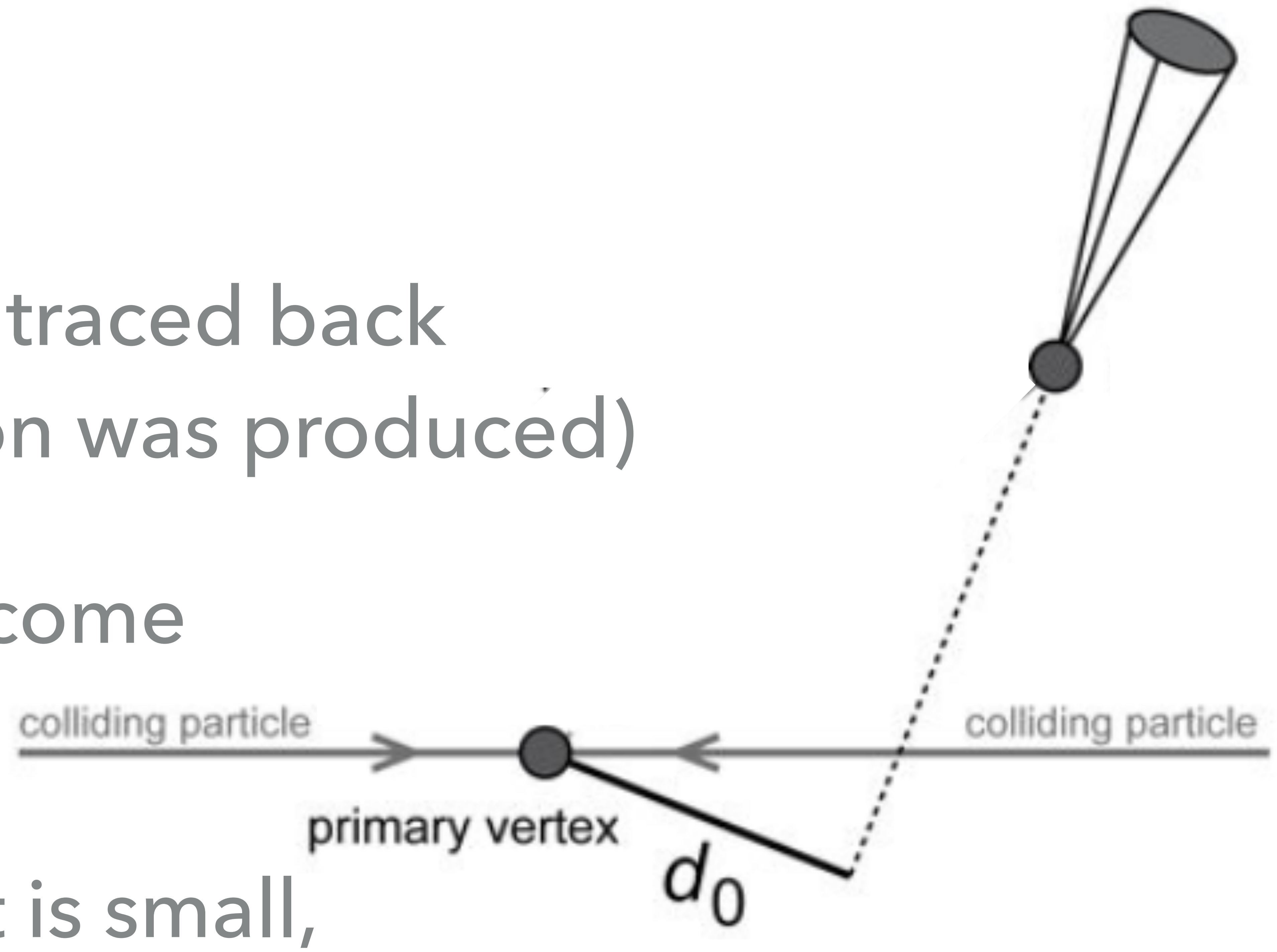
- A measure of the energy in a cone around the lepton, not including the lepton



- Leptons from boson decays are more likely to be isolated
- But we don't want to throw away leptons with really high pt, even if they aren't completely isolated
- So throw away leptons with etcone20/pt > 0.3 (e.g.)

What's lep_d0?

- Trace back the track of the lepton
- Take the minimum distance between the traced back track and primary vertex (where the boson was produced)
- Want to throw away leptons that haven't come from the primary vertex (large d0)
- If the significance of the d0 measurement is small, we can be confident of our d0 measurement
- So cut away leptons with d0/sigd0 > 6.5 (e.g.)



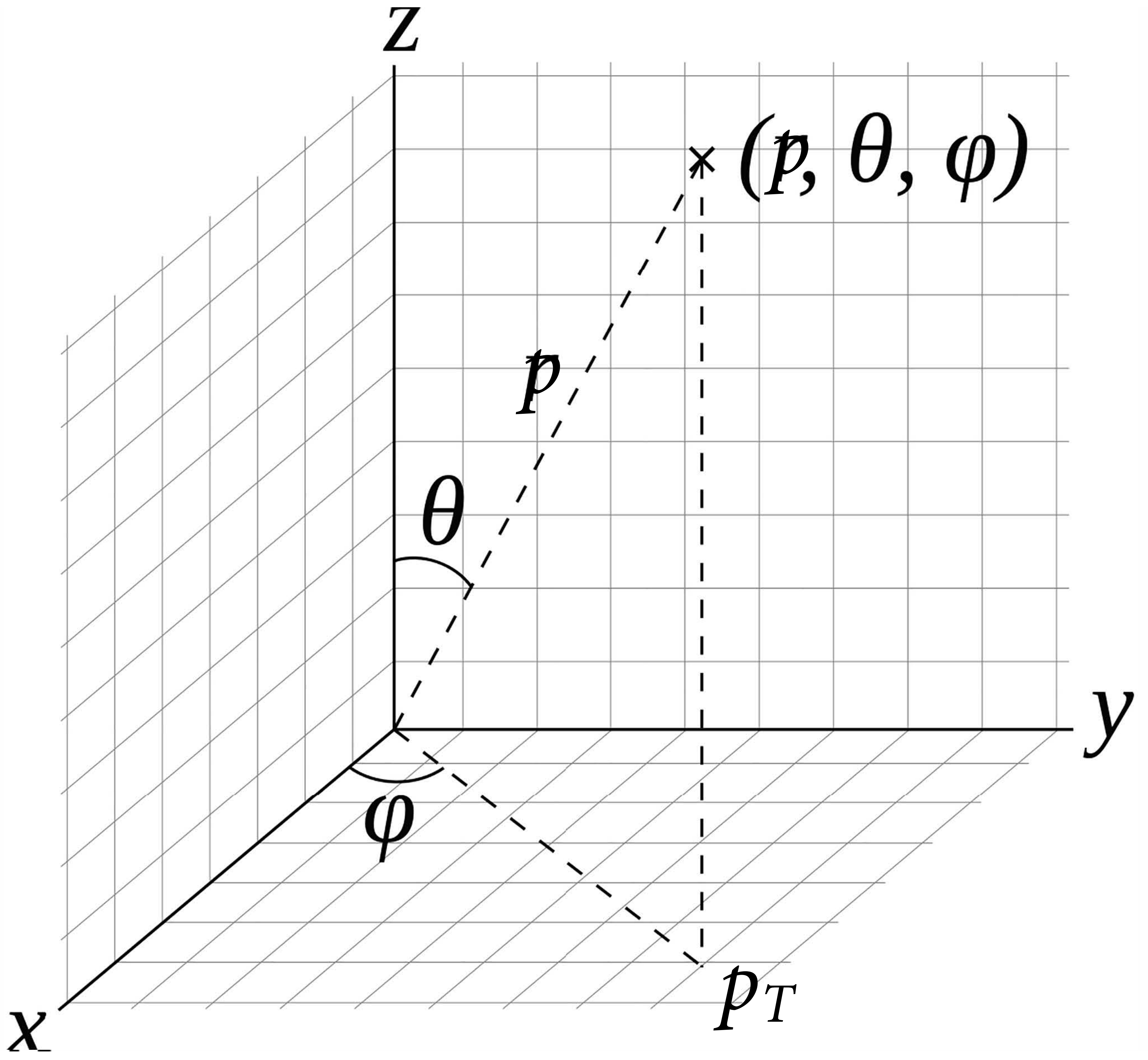
UNIVERSITY OF SUSSEX

ATLAS EXPERIMENT

Momentum components

- $p_x = p_T \cos \phi$
- $p_y = p_T \sin \phi$
- $p_z = p \cos \theta, p_T = p \sin \theta$

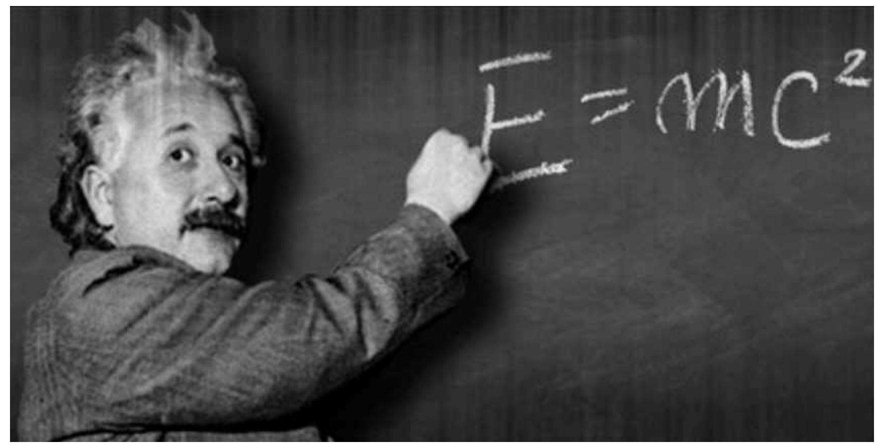
➡ $p_z = \dfrac{p_T}{\sin \theta} \cos \theta$



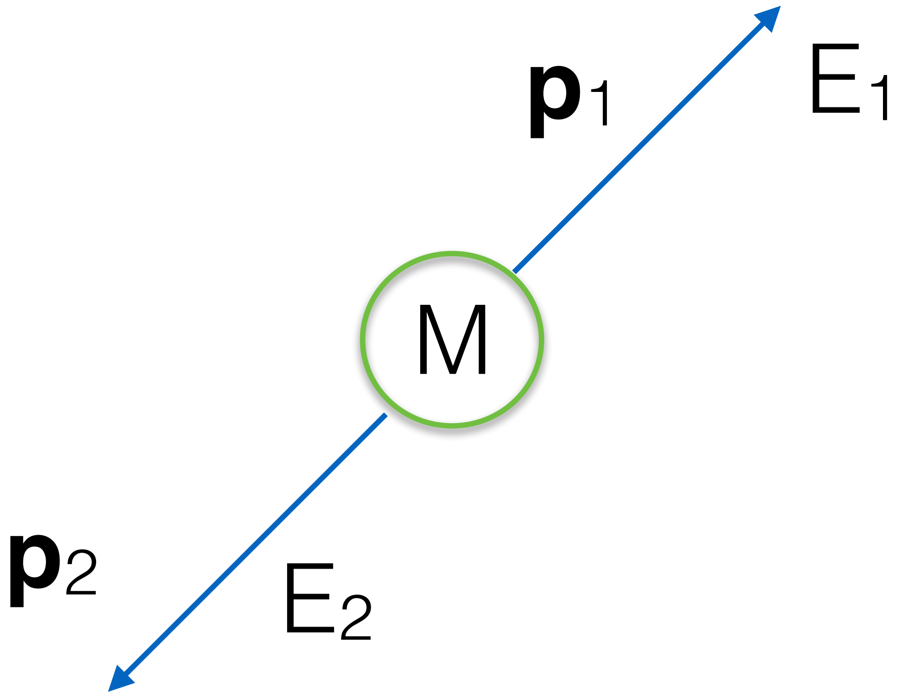
UNIVERSITY OF SUSSEX

ATLAS EXPERIMENT

Invariant mass



- $E^2 = p^2 + M^2$
- $M = \sqrt{E^2 - p^2}$
- $M = \sqrt{(E_1 + E_2 + \ldots)^2 - (\vec{p_1} + \vec{p_2} + \ldots)^2}$



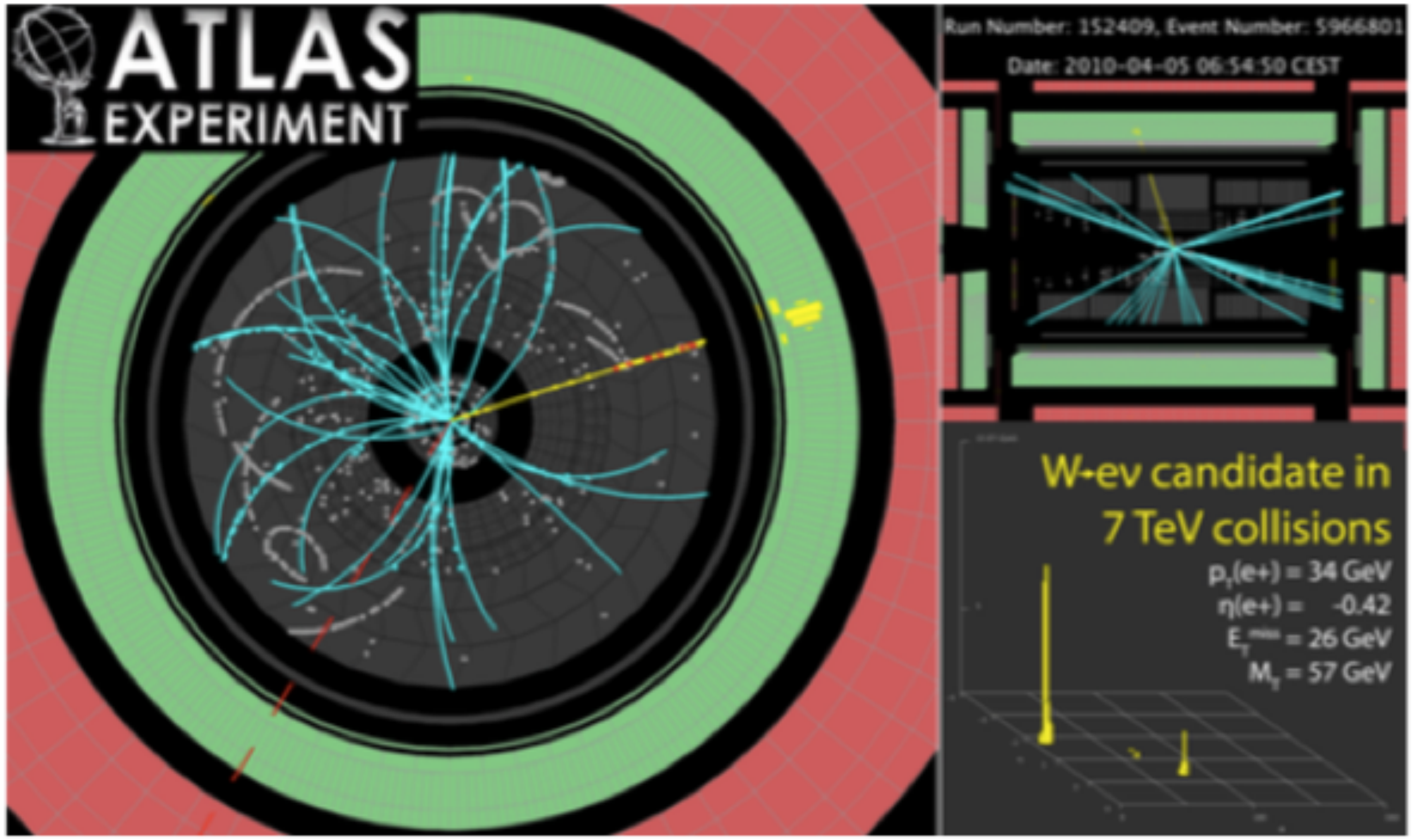
- Use final state leptons to find invariant mass of the boson from which they decayed
- Throw away if invariant mass very different to prediction

UNIVERSITY OF SUSSEX

ATLAS EXPERIMENT

$$\Delta R = \sqrt{(\Delta\eta)^2 + (\Delta\phi)^2}$$

- Don't want the tracks of leptons to overlap
- If they did overlap, it'd be hard to say whether it was 1 or 2 tracks
- So throw away leptons that are very close together in ΔR

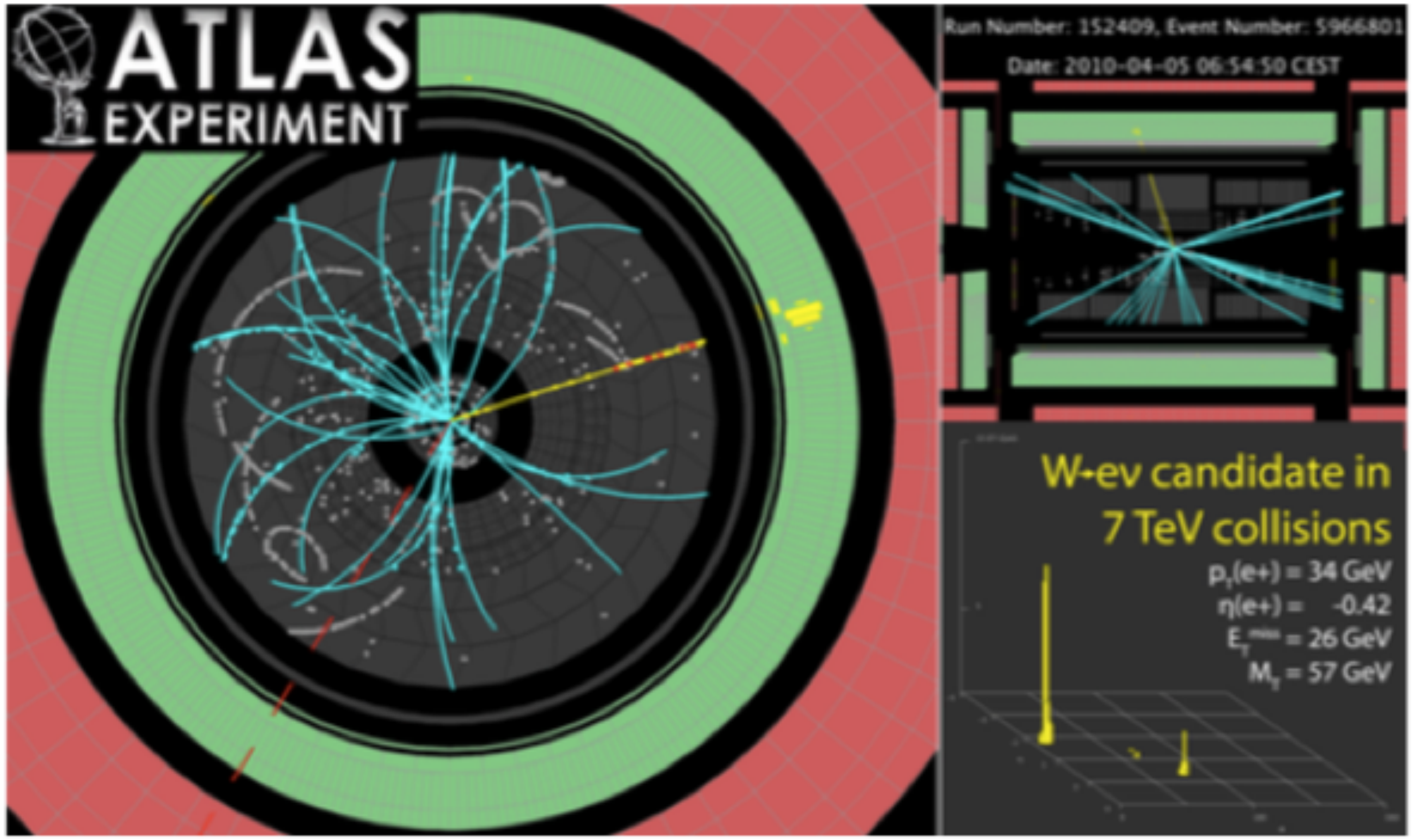


UNIVERSITY OF SUSSEX

ATLAS EXPERIMENT

Why a stricter ΔR for leptons of same type?

- Electrons are detected by their energy deposited
- If electrons are really close together, you're not sure whether you're detecting 2 electrons, or a single high-energy electron
- Similar for muons

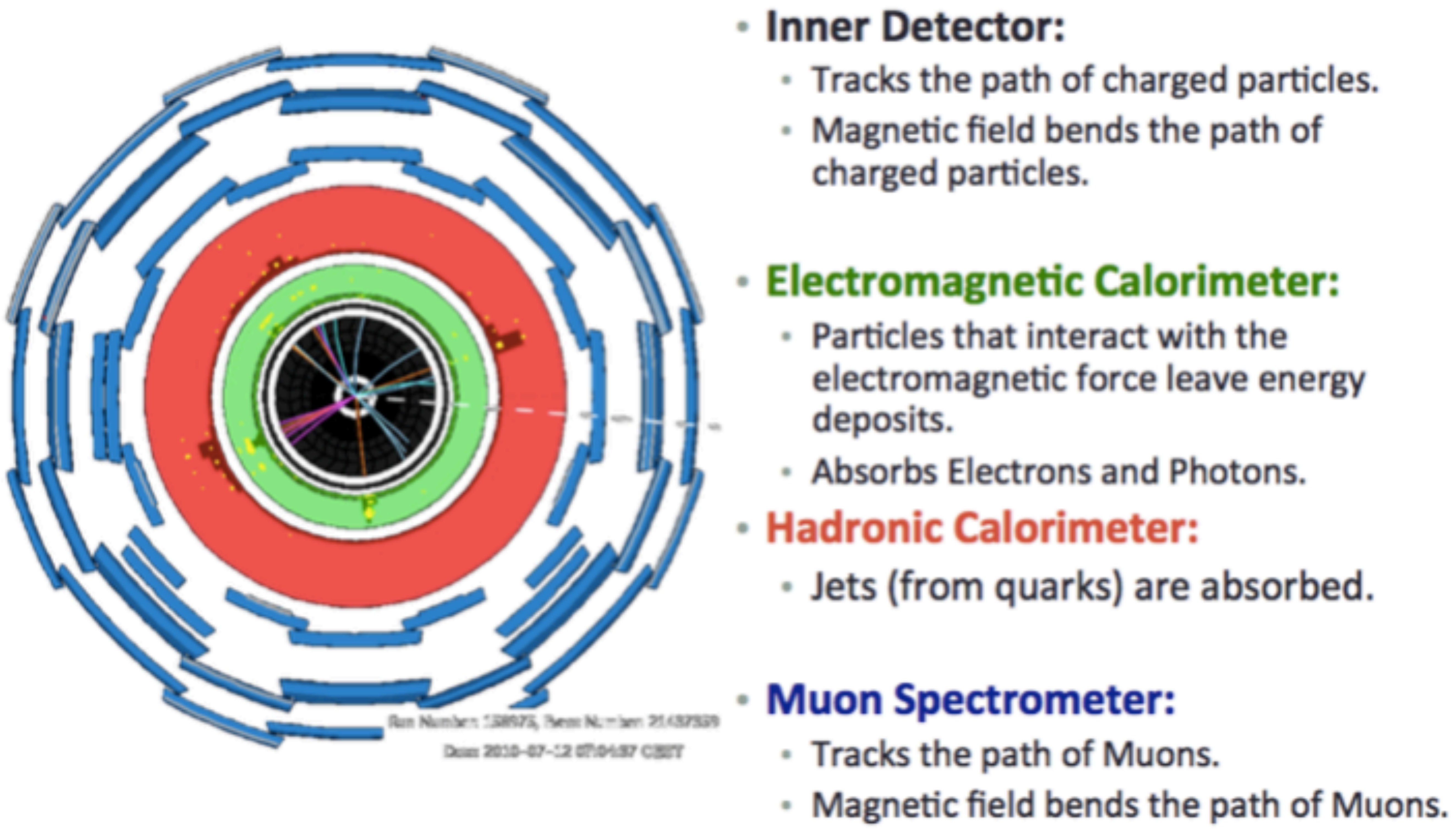


UNIVERSITY OF SUSSEX

ATLAS EXPERIMENT

Why a lower p_T threshold for muons?

- Muons are the only particles detected in the outer part of ATLAS
- So if something is detected here, high chance it's a muon

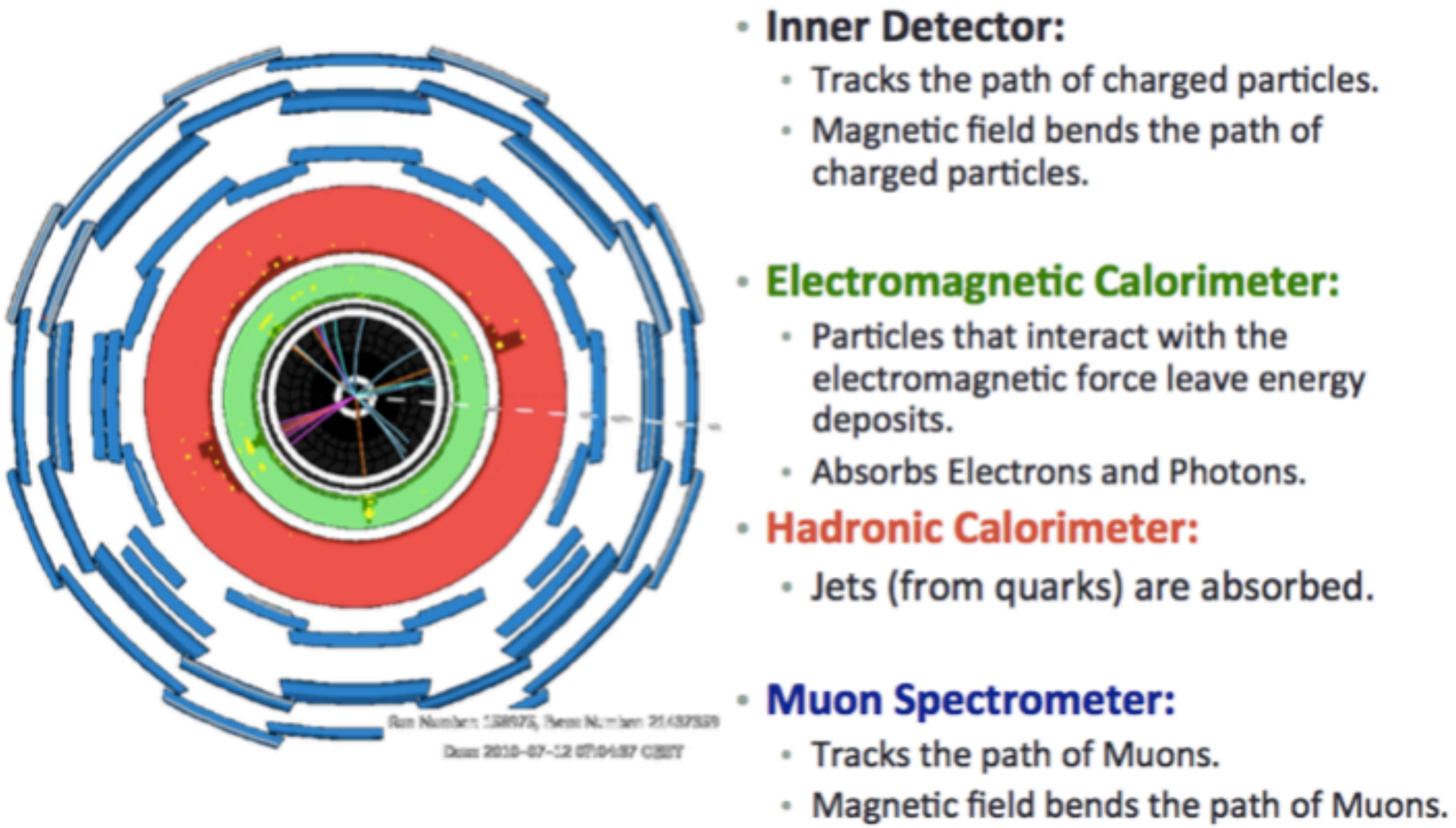


UNIVERSITY OF SUSSEX

ATLAS EXPERIMENT

Why a lower etcone20 threshold for electrons?

- Electrons are detected in a busier part of ATLAS than muons
- Need to account for other energy in Electromagnetic Calorimeter when measuring electrons



UNIVERSITY OF SUSSEX

ATLAS EXPERIMENT

$\text{Signal}/\sqrt{\text{Background}}$

- $\text{Signal}/\sqrt{\text{Background}}$ is a statistical measure of how well you're separating signal and background
- It's an example of "significance"
- This is taking the statistical uncertainty on the number of background events as a Poisson distribution
- Only background in denominator because when searching for a new particle, you have to reject the background-only hypothesis

UNIVERSITY OF SUSSEX

ATLAS EXPERIMENT

Lepton p_T

- "the second (third) lepton in p_T order must satisfy $p_T > 15$ GeV ($p_T > 10$ GeV)"

Minimum opposite-charge-same-type lepton pair invariant mass

- "All possible lepton pairs in the quadruplet that have the same flavour and opposite charge must satisfy $m_{\ell\ell} > 5$ GeV in order to reject backgrounds involving the production and decay of J/ψ mesons"

UNIVERSITY OF SUSSEX

ATLAS EXPERIMENT

Invariant mass of Z boson candidate 1

- "The same-flavour and opposite-charge lepton pair with an invariant mass closest to the Z boson mass (m_Z) in the quadruplet is referred to as the leading lepton pair. Its invariant mass, denoted by m_{12}, is required to be between 50 GeV and 106 GeV"

UNIVERSITY OF SUSSEX

ATLAS EXPERIMENT

Invariant mass of Z boson candidate 2

- "The same-flavour and opposite-charge lepton pair with an invariant mass closest to the Z boson mass (m_Z) in the quadruplet is referred to as the leading lepton pair. The remaining same-flavour, opposite-charge lepton pair is the sub-leading lepton pair. Its invariant mass, m_{34} ,is required to be in the range $m_{min} < m_{34} < 115$ GeV, where the value of m_{min} depends on the reconstructed four-lepton invariant mass, $m_{4\ell}$. The value of m_{min} varies monotonically from 17.5 GeV at $m_{4\ell}$ = 120 GeV to 50 GeV at $m_{4\ell}$ = 190 GeV and is constant above this value."

UNIVERSITY OF SUSSEX

ATLAS EXPERIMENT

ΔR

- "leptons are required to be separated from each other by $\Delta R = \sqrt{(\Delta\eta)^2 + (\Delta\phi)^2} > 0.1$ if they are of the same flavour and by $\Delta R = \sqrt{(\Delta\eta)^2 + (\Delta\phi)^2} > 0.2$ otherwise."

UNIVERSITY OF SUSSEX

ATLAS EXPERIMENT

Minimum lepton p_T

- "Each electron (muon) must satisfy $p_T > 7$ GeV ($p_T > 6$ GeV)"

UNIVERSITY OF SUSSEX

ATLAS EXPERIMENT

Lepton etcone20

- "The … isolation for electrons is computed as the sum of the E_T of …energy … clusters … within a cone of size ΔR = 0.2 around the candidate electron …, divided by the electron E_T. … The … energy of the cells assigned to the electron … is excluded,…The … isolation for electrons is required to be less than 0.20. The … isolation …for muons is defined by the ratio to the p_T of the muon of the E_T sum of the calorimeter cells inside a cone of size ΔR = 0.2 around the muon direction minus the energy deposited by the muon. Muons are required to have a … isolation less than 0.30 "

UNIVERSITY OF SUSSEX

ATLAS EXPERIMENT

Lepton d0

- "Non-prompt leptons from heavy flavour decays, electrons from photon conversions and jets mis-identified as electrons have broader … impact parameter distributions than prompt leptons from Z boson decays and/or are non-isolated. Thus, the Z and $t\bar{t}$ background contributions are reduced by applying a cut on the … impact parameter significance, defined as the transverse impact parameter divided by its uncertainty, d_0/σ_{d_0}. This is required to be less than 3.5 (6.5) for muons (electrons). The electron impact parameter is affected by bremsstrahlung and thus has a broader distribution."

Significance

- $$\frac{\text{measurement}}{\text{error}}$$
- Quote this as a number of "σ"
  - (Not to be confused with a 1σ error bar!)
- Error can usually be estimated using Poisson (square root)
- What measures of significance do you use in your work?

UNIVERSITY OF SUSSEX

open data ATLAS

Significance measures

- $$\frac{N_{\text{signal}}}{\sqrt{N_{\text{background}}}}$$
- $$\frac{N_{\text{data}} - N_{\text{background}}}{\sqrt{N_{\text{background}}}}$$
- $$\frac{N_{\text{signal}}}{\sqrt{N_{\text{signal}} + N_{\text{background}}}}$$
- $$\frac{N_{\text{data}} - N_{\text{background}}}{\sqrt{N_{\text{signal}} + N_{\text{background}}}}$$
- $$\frac{N_{\text{data}} - N_{\text{background}}}{\sqrt{N_{\text{data}}}}$$
- $$\sqrt{2((N_{\text{signal}} + N_{\text{background}})\ln(1 + \frac{N_{\text{signal}}}{N_{\text{background}}}) - N_{\text{signal}})}$$

UNIVERSITY OF SUSSEX

open data ATLAS

If $N_{\text{background}}$ known

- $\dfrac{N_{\text{signal}}}{\sqrt{N_{\text{background}}}}$
- Profile likelihood ratio test

$$\sqrt{2((N_{\text{signal}} + N_{\text{background}})\ln(1 + \frac{N_{\text{signal}}}{N_{\text{background}}}) - N_{\text{signal}})}$$

- Useful resource: statistics presentation by Glen Cowan

Implement more cuts

- What results did you get?
- How does the new result compare with previous results?
- How would you measure the success of these cuts?
- How might you improve?

UNIVERSITY OF SUSSEX

open data ATLAS

Optimisation

- Cuts used in Higgs discovery paper will have been optimised for that dataset
- We only have a subset (~10%) of that dataset
  - So the optimised cuts for our dataset might be different
- Modify some cuts and see what happens
- How might you optimise your cuts?

UNIVERSITY OF SUSSEX

open data ATLAS

Machine Learning

- Neural Network (NN)
- Boosted Decision Tree (BDT)
- Support Vector Machine (SVM)
- Random Forest

UNIVERSITY OF SUSSEX

open data ATLAS