

Introduction to particle physics: experimental part

❖ **RAW data to Physics**

- The road from collisions to physics publications

❖ **From RAW data to Standard Model particles**

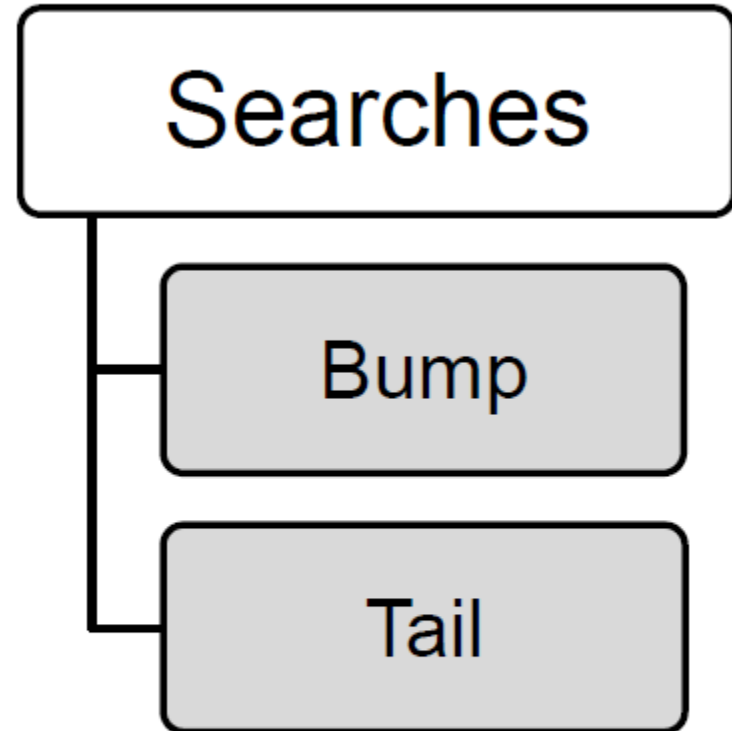
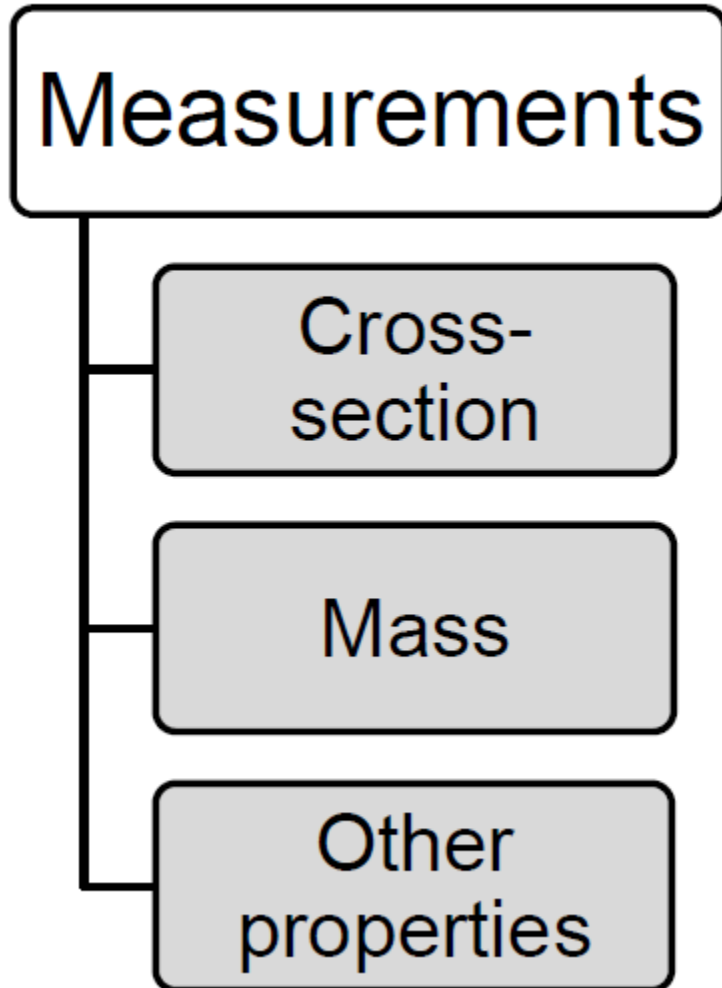
- About measuring properties of the final particles created from proton-proton collisions

❖ **From Standard Model particles to measurements and searches**

- About how we analyse data using ingredients we have constructed

Large fraction of slides from A. Sfyrla lectures at CERN Summer School 2018

Physics analyses



Phycis analyses

Measurements

- ◎ Allow important tests of the consistency of the theory.
- ◎ Typically limited by systematic uncertainties.

Searches

- ◎ ... For new particles.
- ◎ If no signal, set limits on some model.
- ◎ If signal, a potential discovery!
- ◎ More data typically improve a search.

Physics analyses

Measurements

- © Allow important tests of the consistency of the theory.
- © Typically limited by systematic uncertainties.

Searches

- © More data typically improve a search.

“Systematic” uncertainties are introduced by inaccuracies in the methods used to perform the measurement.

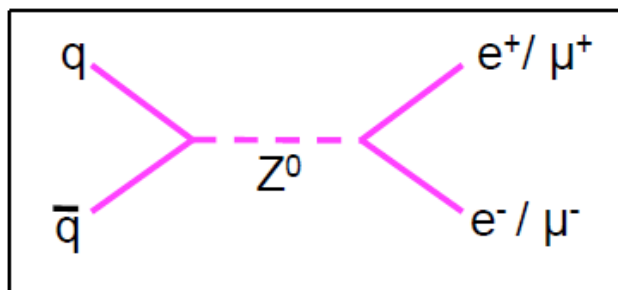
SIMPLE EXAMPLE:

MEASURING Z^0 CROSS-SECTION AT LHC

Measuring Z^0 cross-section at LHC

⊙ Z^0 boson decays to lepton or quark pairs

⊙ We can reconstruct it in the e^+e^- or $\mu^+\mu^-$ decay modes



⊙ Discovery and study of the Z^0 boson was a critical part understanding the electroweak force.



⊙ And now, at the LHC?

⊙ **Important test of theory:** does the measurement agree with the theoretical prediction at LHC collision energy?

⊙ **A standard candle** for studying reconstruction and deriving calibrations.

⊙ Can be used for luminosity determination!

Physics analyses

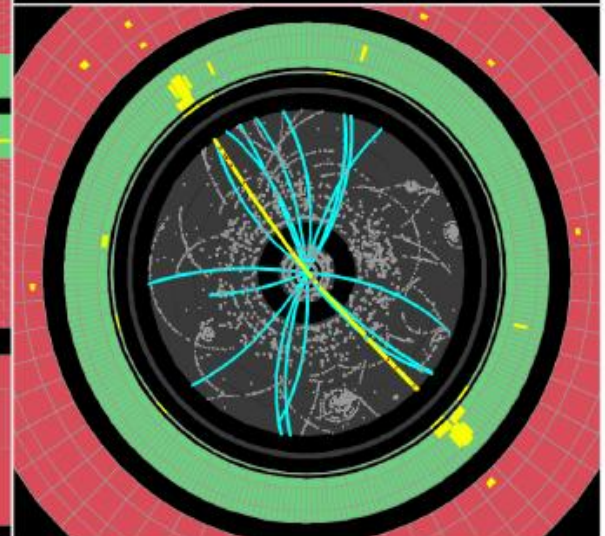
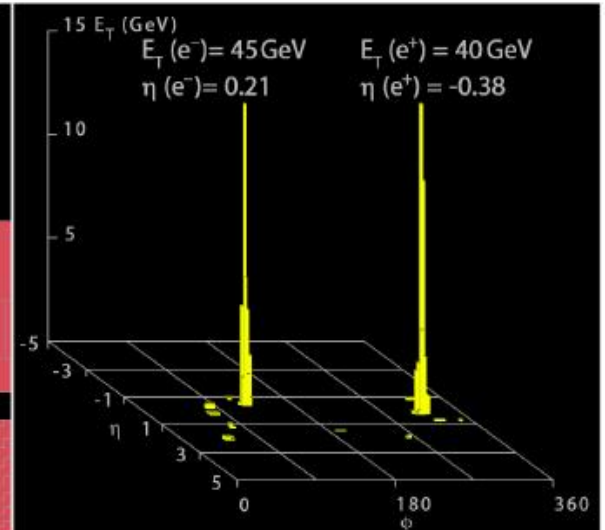
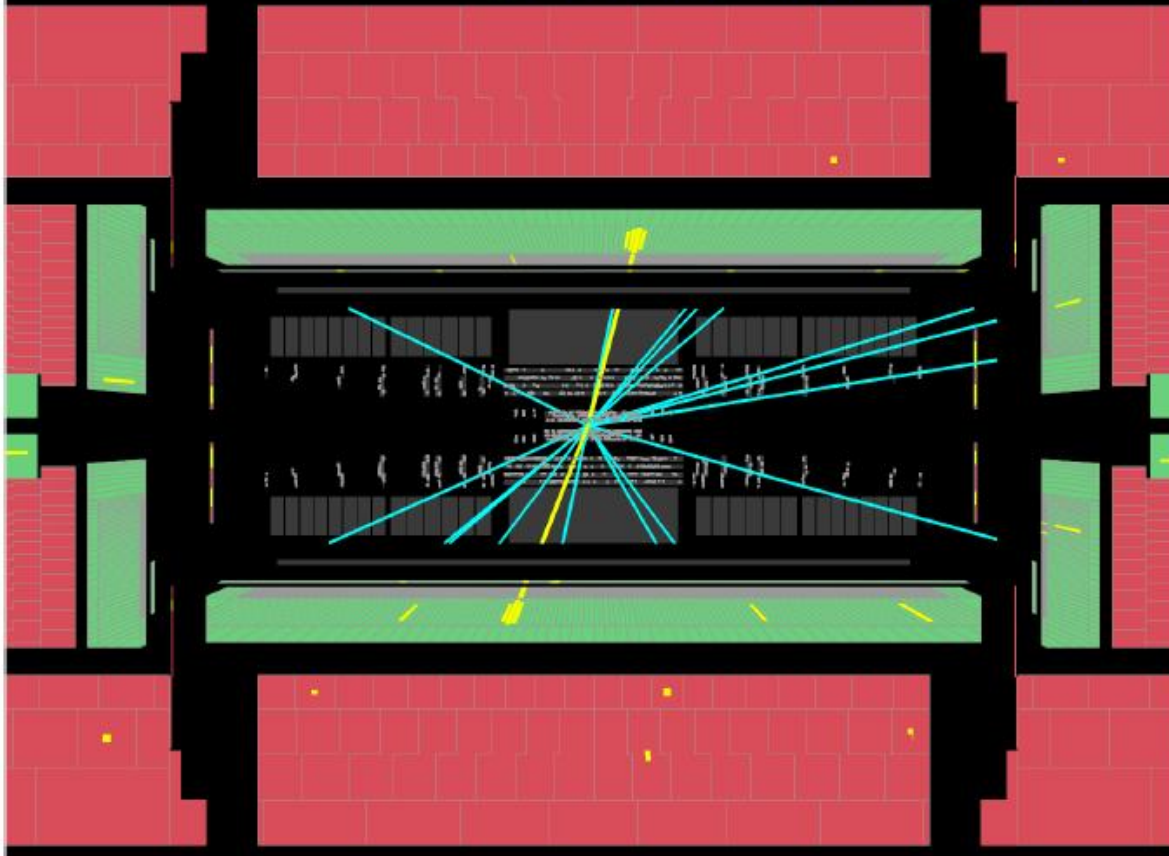


Run Number: 154817, Event Number: 968871

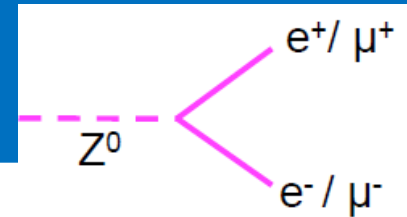
Date: 2010-05-09 09:41:40 CEST

$M_{ee} = 89 \text{ GeV}$

$Z \rightarrow ee$ candidate in 7 TeV collisions



Reconstructing Z^0 's



How do we know it's a Z^0 ?

Identify Z decays using the invariant mass of the 2 leptons

$$M^2 = (L_1 + L_2)^2 \quad \text{where } L_i = (E_i, \mathbf{p}_i) = 4\text{-vector for lepton } i$$

Under assumption that lepton is massless compared to mass of Z^0

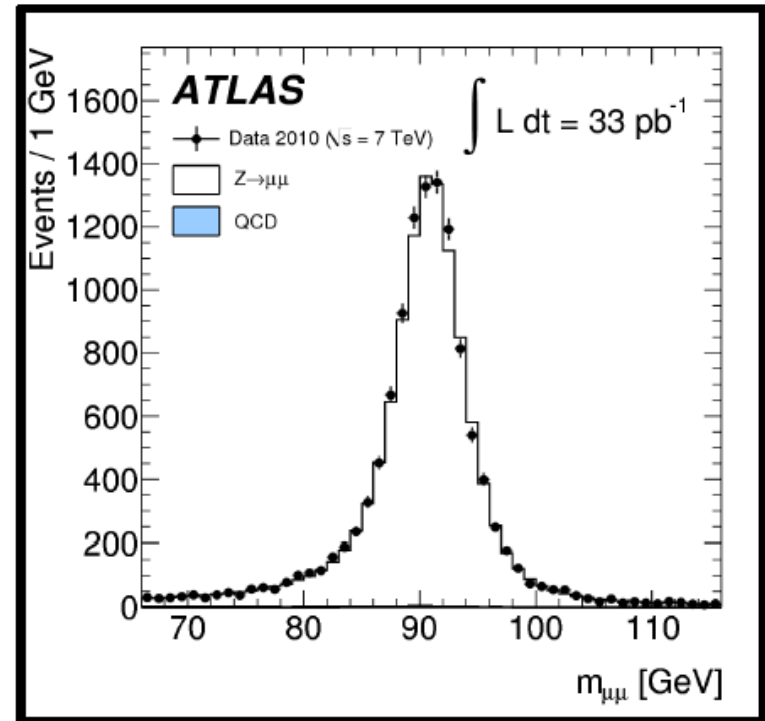
$$\Rightarrow M^2 = 2 E_1 E_2 (1 - \cos\vartheta_{12}) \quad \text{where } \vartheta_{12} = \text{angle between the leptons}$$

⊙ So need to reconstruct the electron and muon energy and direction. Then can calculate the mass.

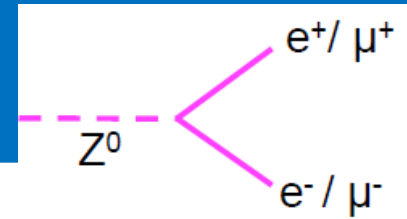
Select Z^0 events with 'analysis cuts':

- ⊙ Events with 2 high momentum electrons or muons
- ⊙ Require the electrons or muons are of opposite charge
- ⊙ With di-lepton mass close to the Z^0 mass (e.g. $70 < m_{l+l-} < 110$ GeV)

Very little background in Z^0 mass region!



Reconstructing Z^0 's



How do we know it's a Z^0 ?

Identify Z decays using the invariant mass of the 2 leptons

$M^2 = (L_1 + L_2)^2$ where $L_i = (E_i, \mathbf{p}_i)$ = 4-vector for lepton i

Under assumption that lepton is massless compared to mass of Z^0

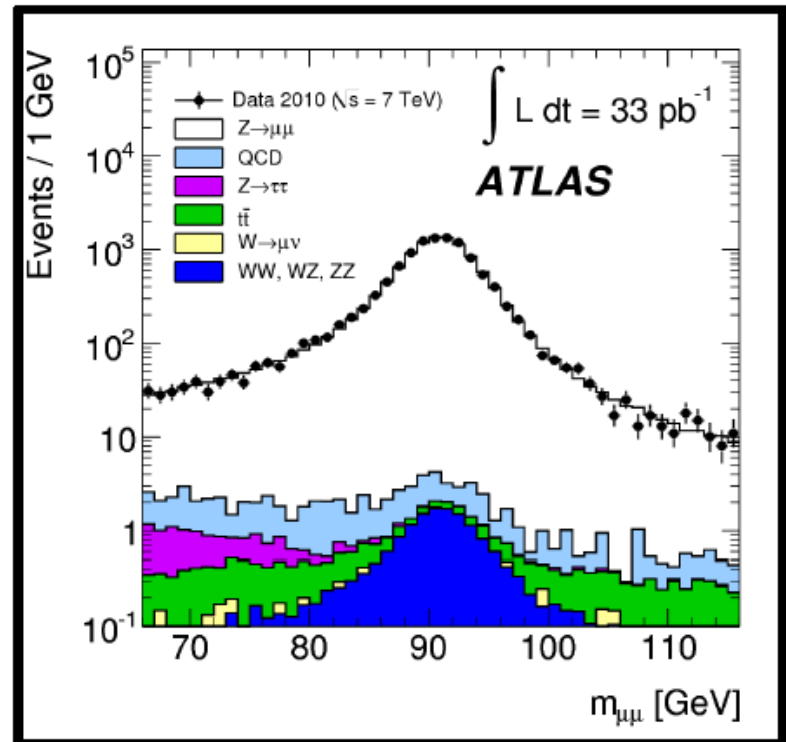
$\Rightarrow M^2 = 2 E_1 E_2 (1 - \cos\vartheta_{12})$ where ϑ_{12} = angle between the leptons

⊙ So need to reconstruct the electron and muon energy and direction. Then can calculate the mass.

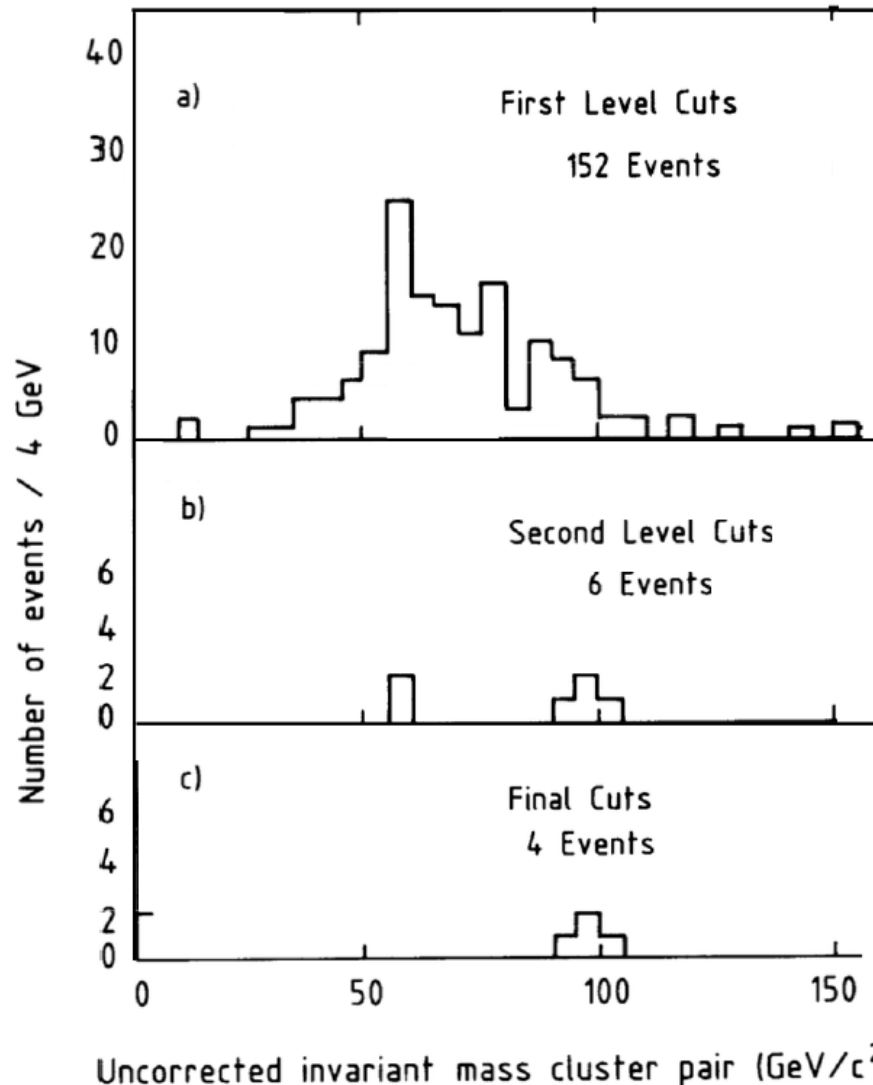
Select Z^0 events with 'analysis cuts':

- ⊙ Events with 2 high momentum electrons or muons
- ⊙ Require the electrons or muons are of opposite charge
- ⊙ With di-lepton mass close to the Z^0 mass (e.g. $70 < m_{l+l-} < 110$ GeV)

Very little background in Z^0 mass region!



A step back in time ...



Z → ee in UA1

Two EM clusters with $E_T > 25 \text{ GeV}$.

As above plus a track with $p_T > 7 \text{ GeV}$ pointing to the cluster. Hadronic and track isolation requirements applied.

A second cluster has also an isolated track.

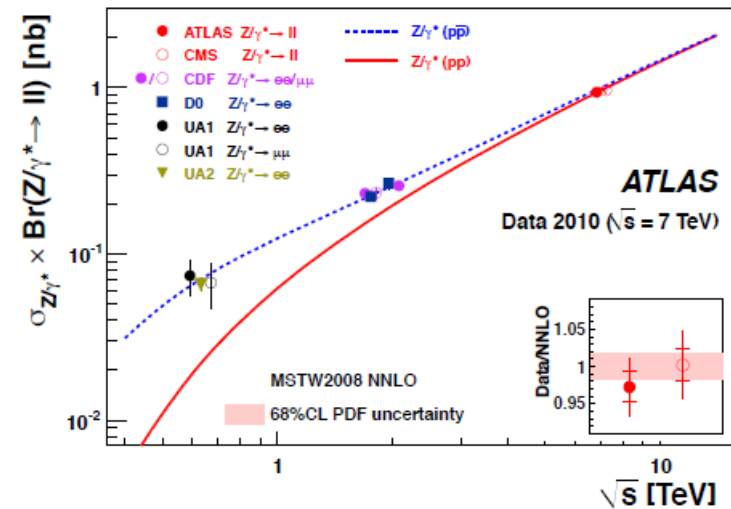


Measuring the Z^0 cross-section

Theoretically

Cross-section calculated for:

- ⊙ Specific production mechanism (pp, $p\bar{p}$, e^+e^-)
- ⊙ Centre-of-Mass of the collisions (7, 8, 13 TeV at LHC)



Experimentally

$$\sigma \cdot \text{BR} = \frac{\text{Number of events}}{\alpha \cdot \epsilon \cdot L}$$

N of events:

N of events on data – N of expected background events

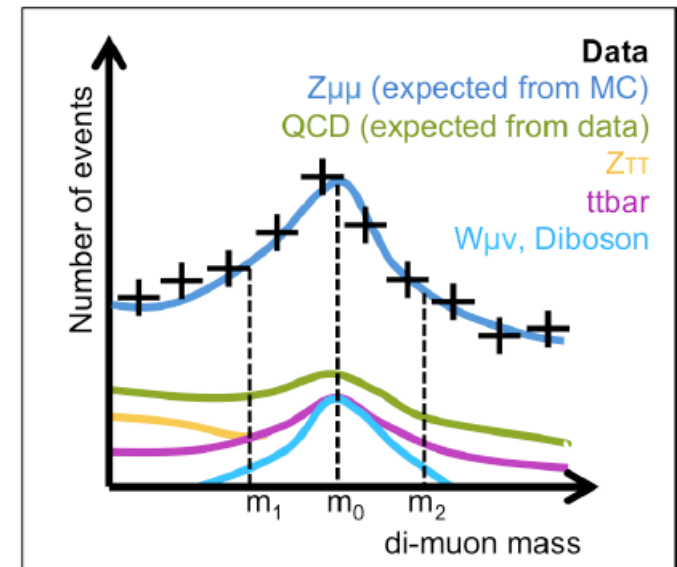
α – acceptance:

fraction of events passing selection requirements

ϵ – efficiency:

reconstruction efficiency of relevant objects

L – luminosity

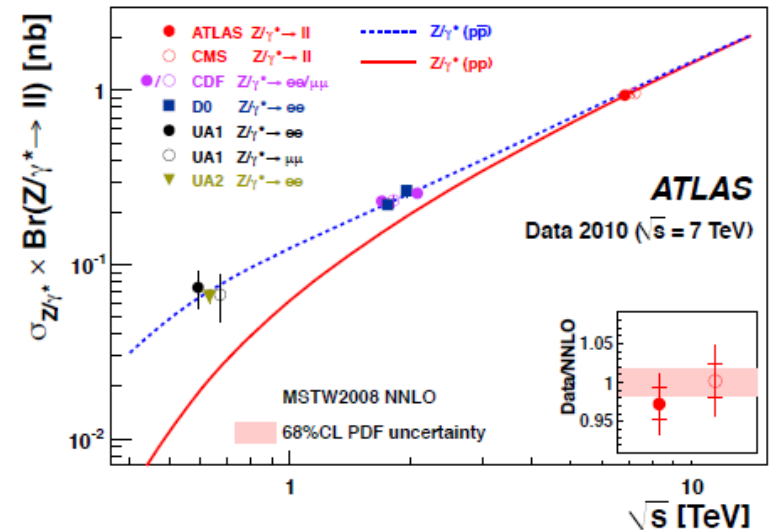


Measuring the Z^0 cross-section

Theoretically

Cross-section calculated for:

- ⊙ Specific production mechanism (pp, $p\bar{p}$, e^+e^-)
- ⊙ Centre-of-Mass of the collisions (7, 8, 13 TeV at LHC)



Experimentally

$$\sigma \cdot BR = \frac{\text{Number of events}}{\alpha \cdot \epsilon \cdot L}$$

N of events:

N of events on data – N

α – acceptance:

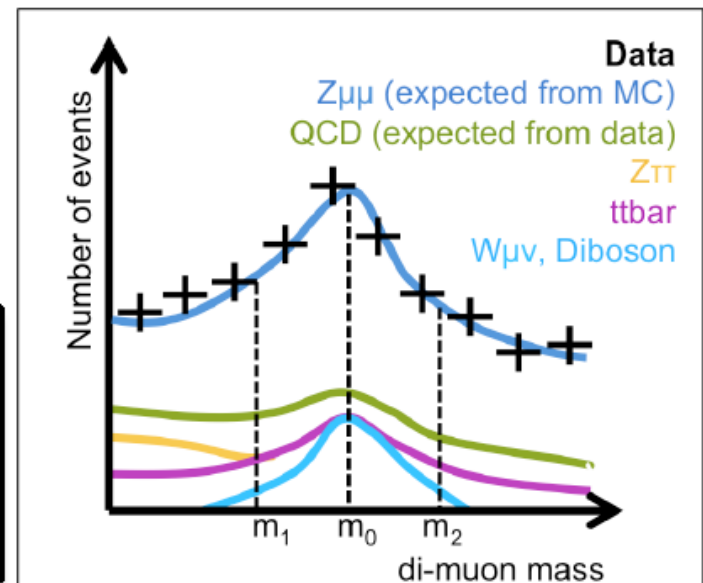
fraction of events passing selection

ϵ – efficiency:

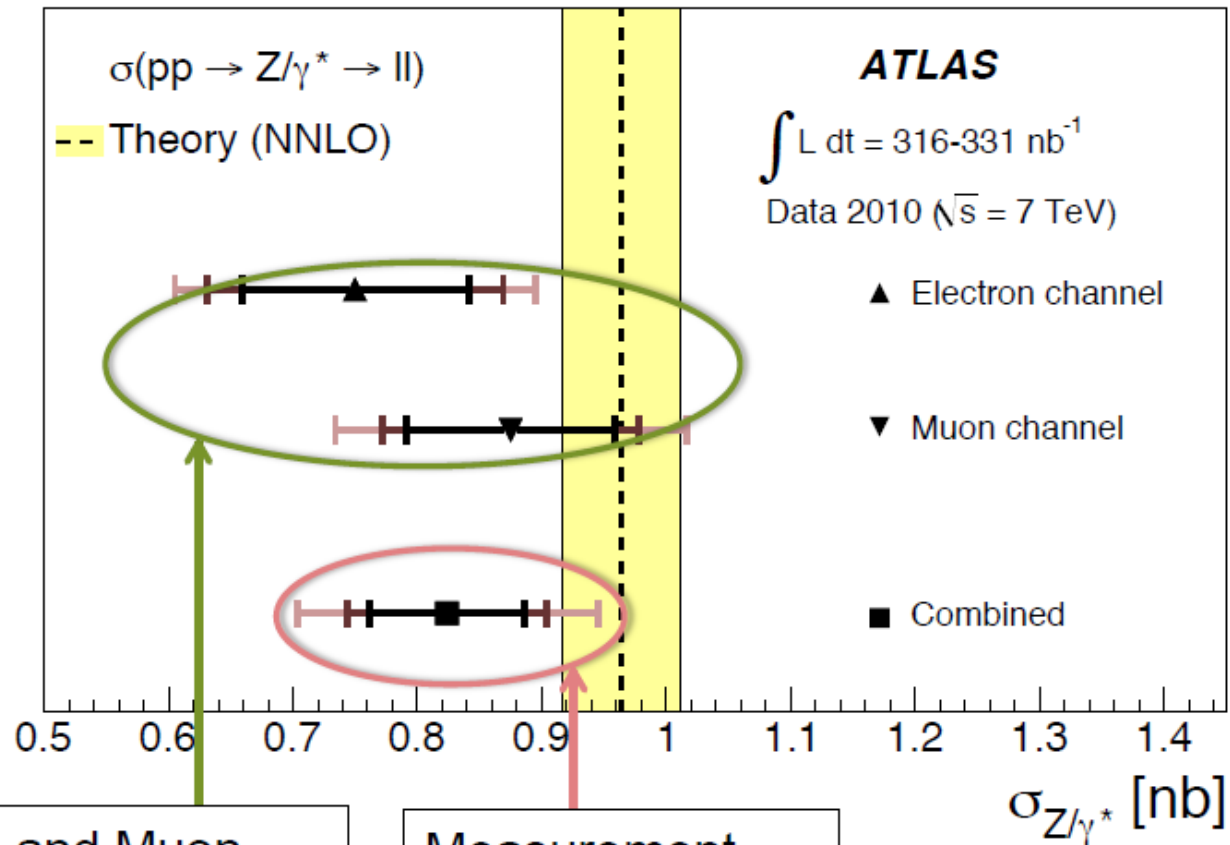
reconstruction efficiency

L – luminosity

All numbers carry uncertainties – both “statistical” and “systematic”!



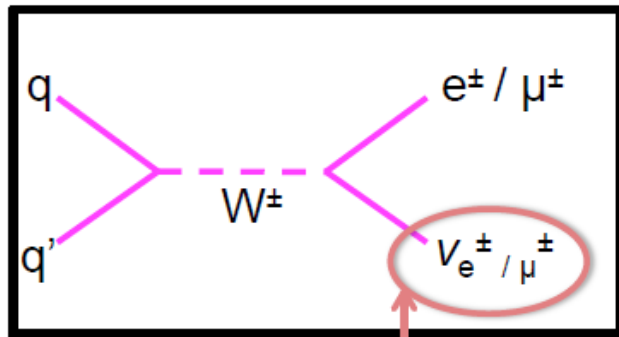
Measuring the Z^0 cross-section



Electron and Muon channel agree within uncertainties

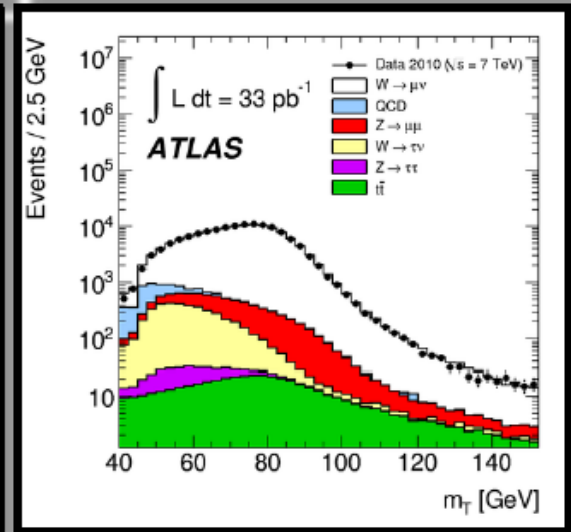
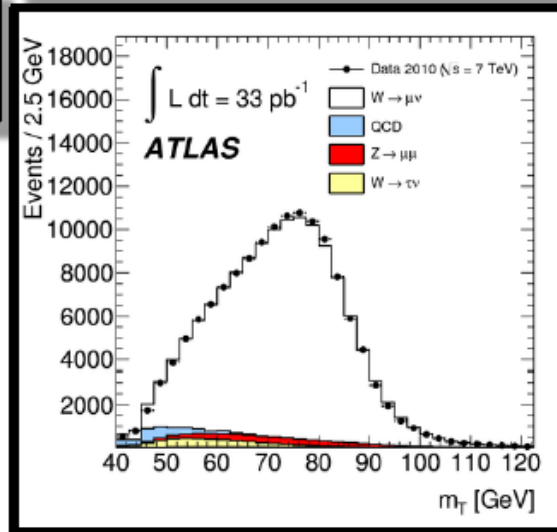
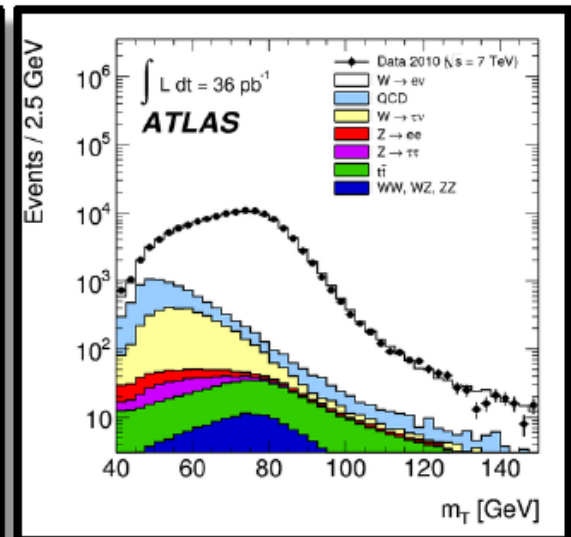
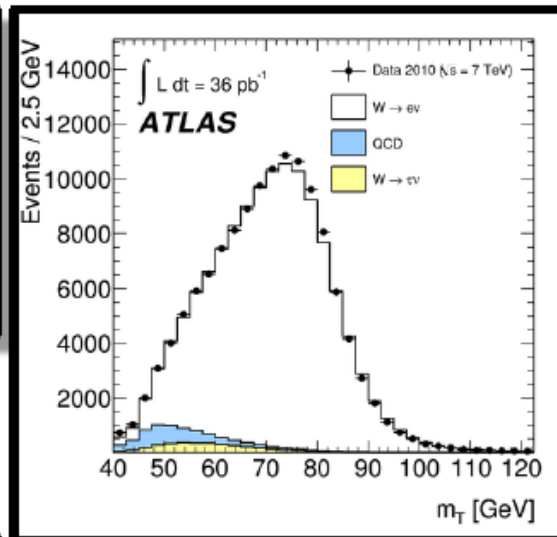
Measurement consistent with prediction within uncertainties

Measuring the W cross-section

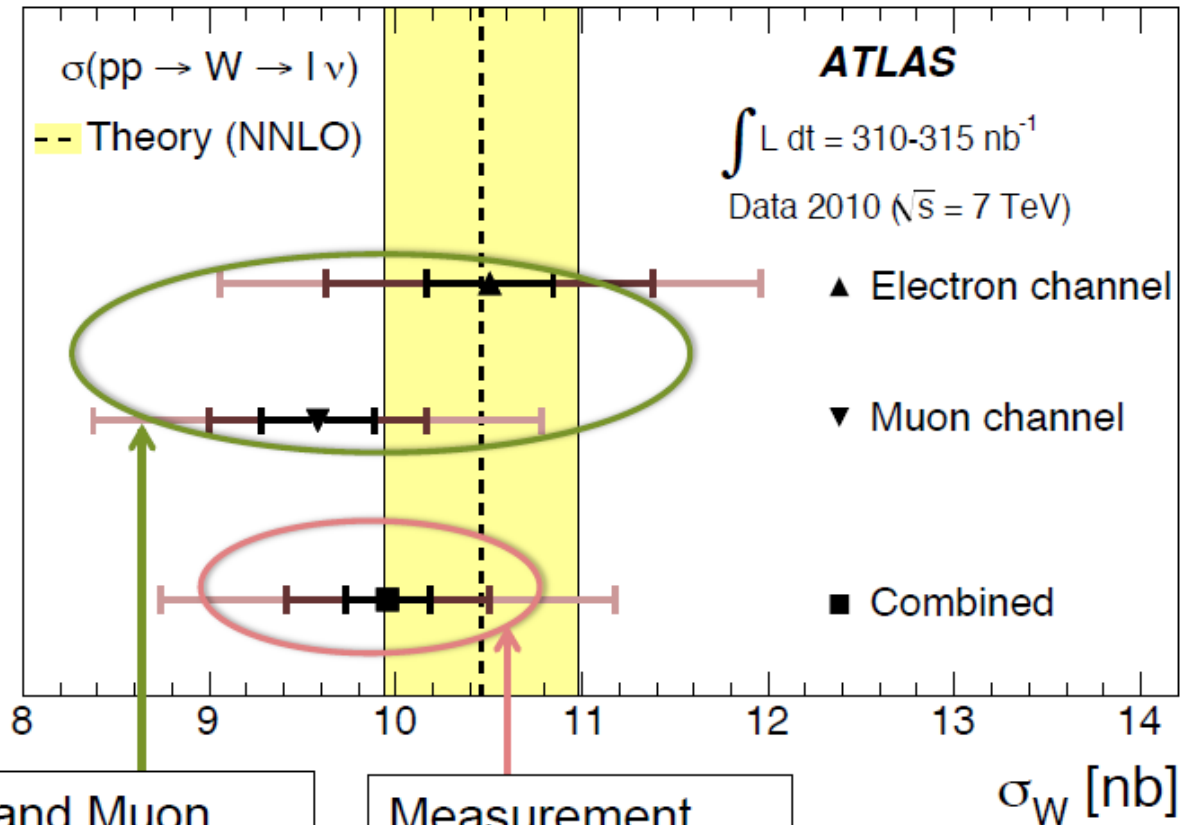


Available in the transverse plane only!

$$M_T^2 = 2 E_{T1} E_{T2} (1 - \cos\theta_{12})$$



Measuring the W cross-section



Electron and Muon channel agree within uncertainties

Measurement consistent with prediction within uncertainties

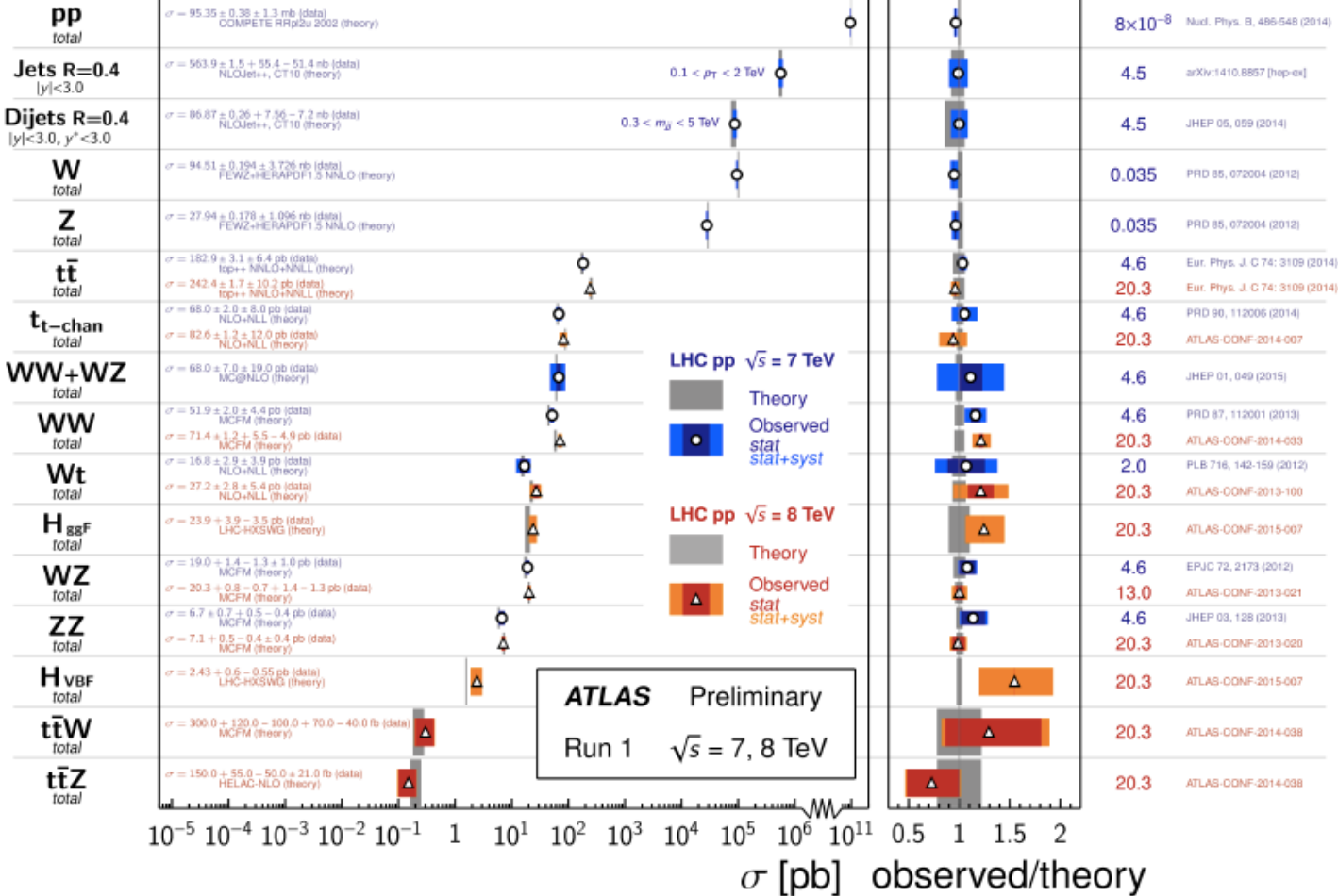
„Final” calibration

Standard Model Total Production Cross Section Measurements

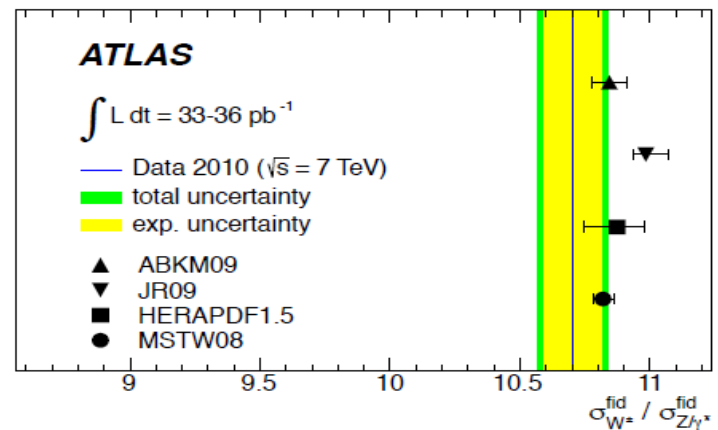
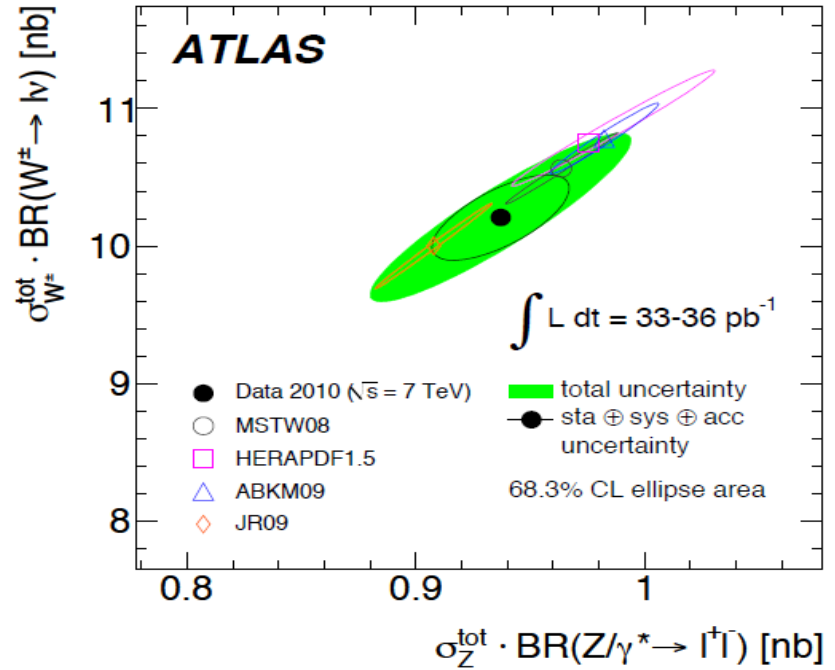
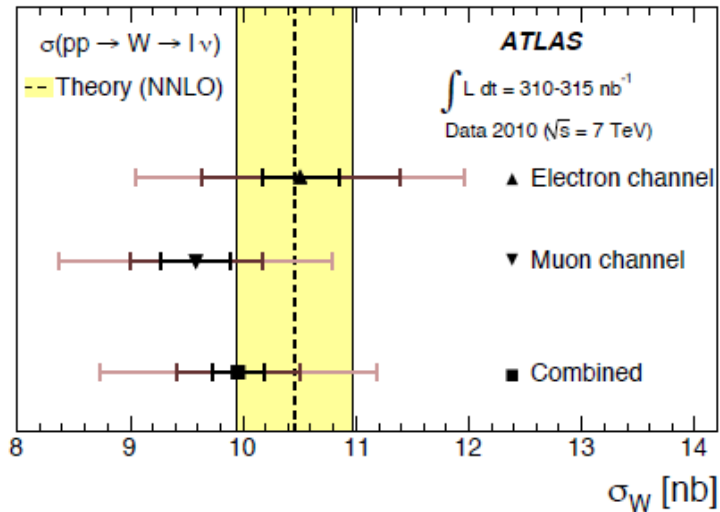
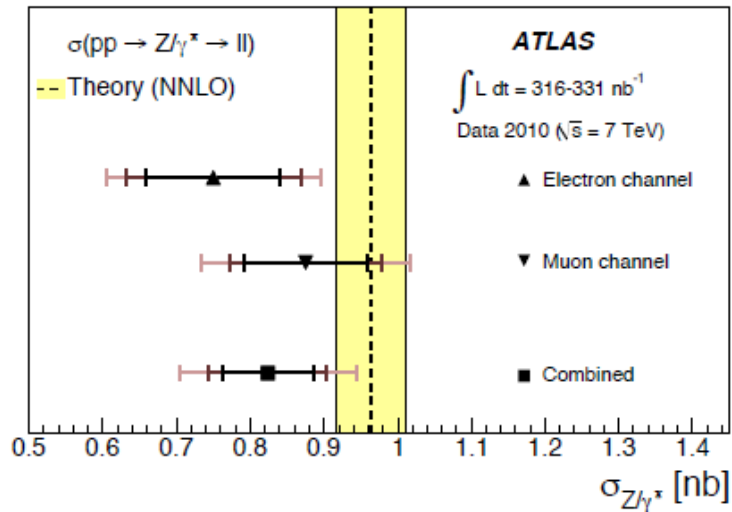
Status:
March 2015

$\int \mathcal{L} dt$
[fb⁻¹]

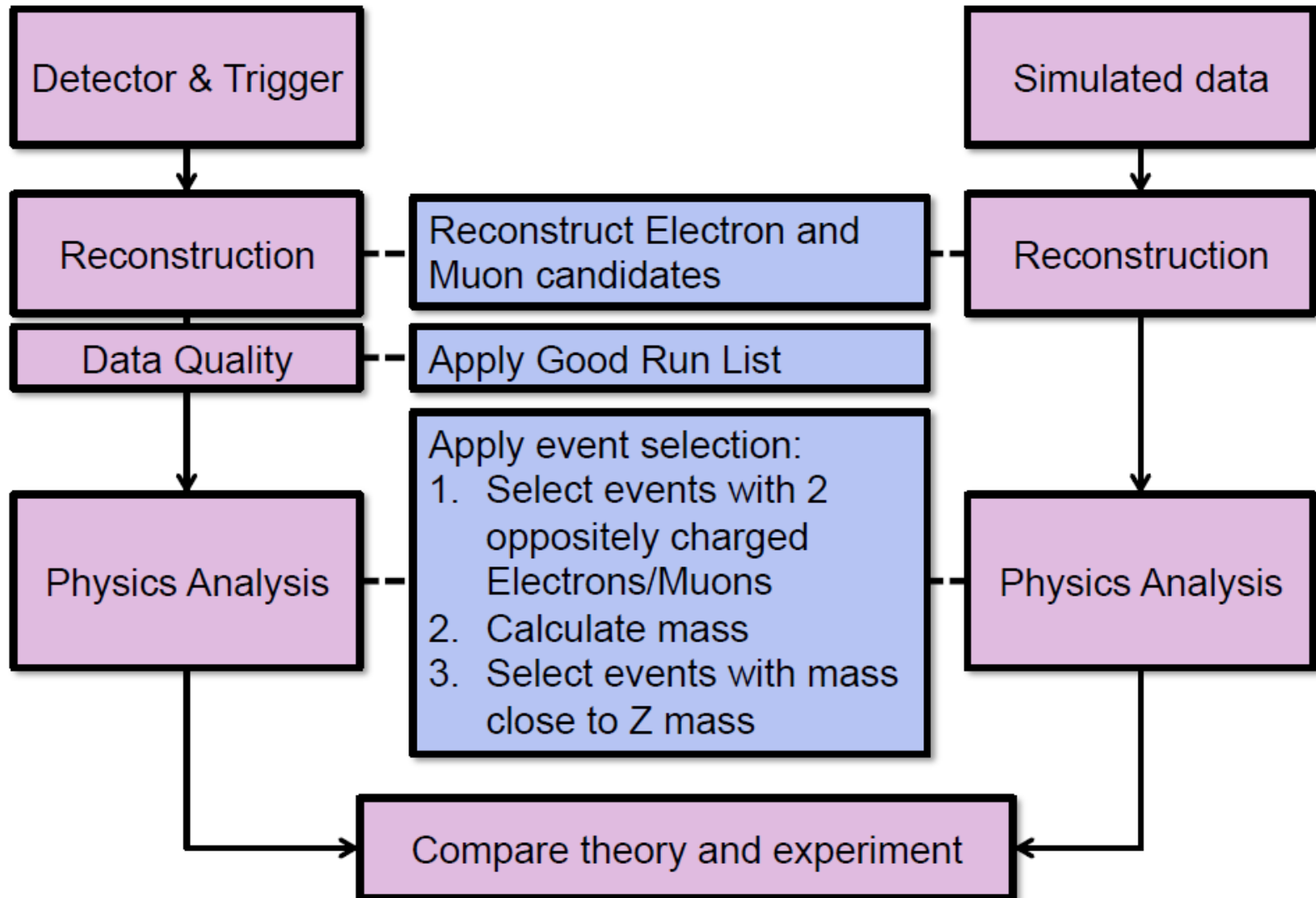
Reference



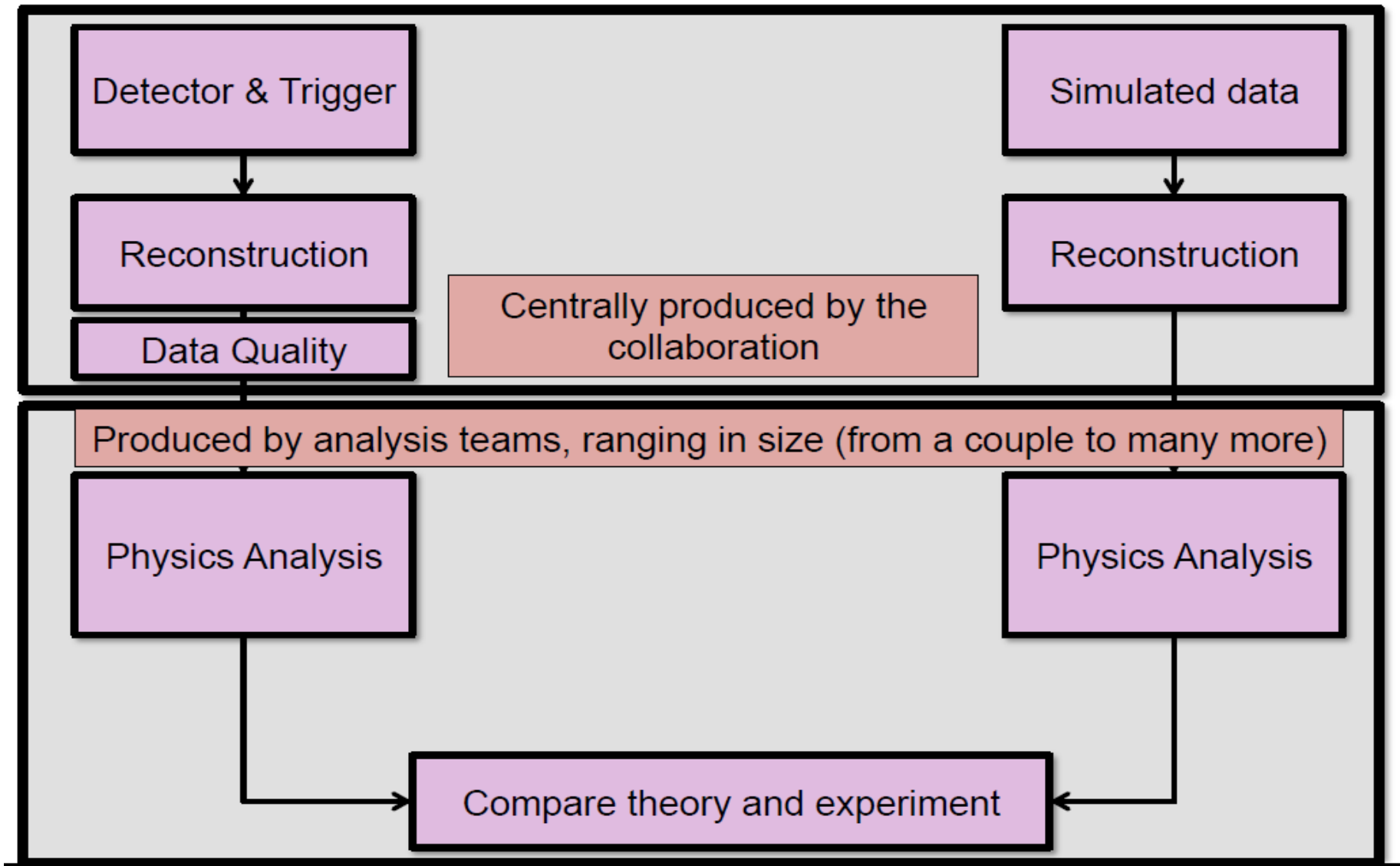
Measuring cross-sections ratio

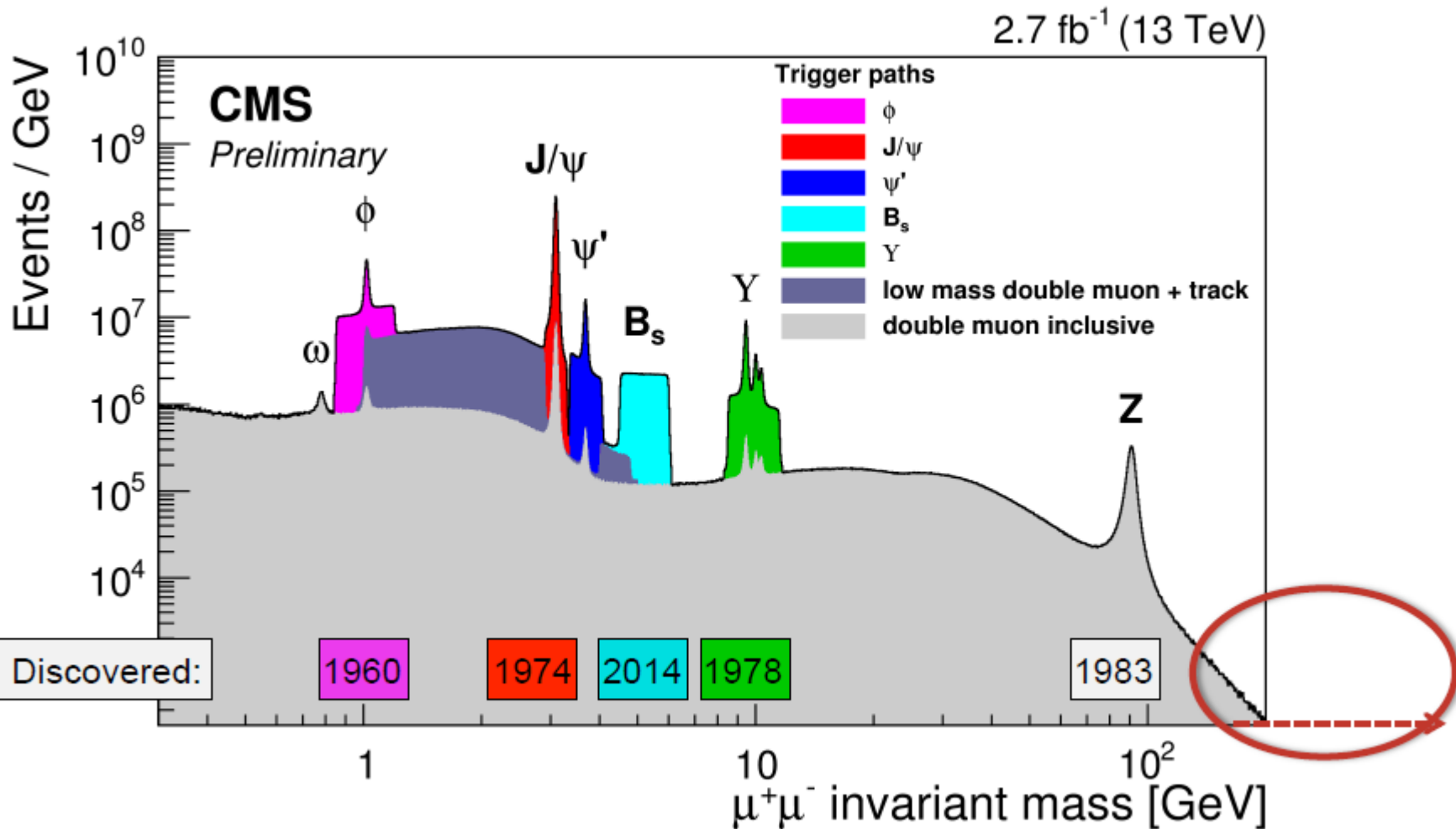


Analysis flow in Z^0 cross-section measurement



Analysis flow in Z^0 cross-section measurement





What's out there?

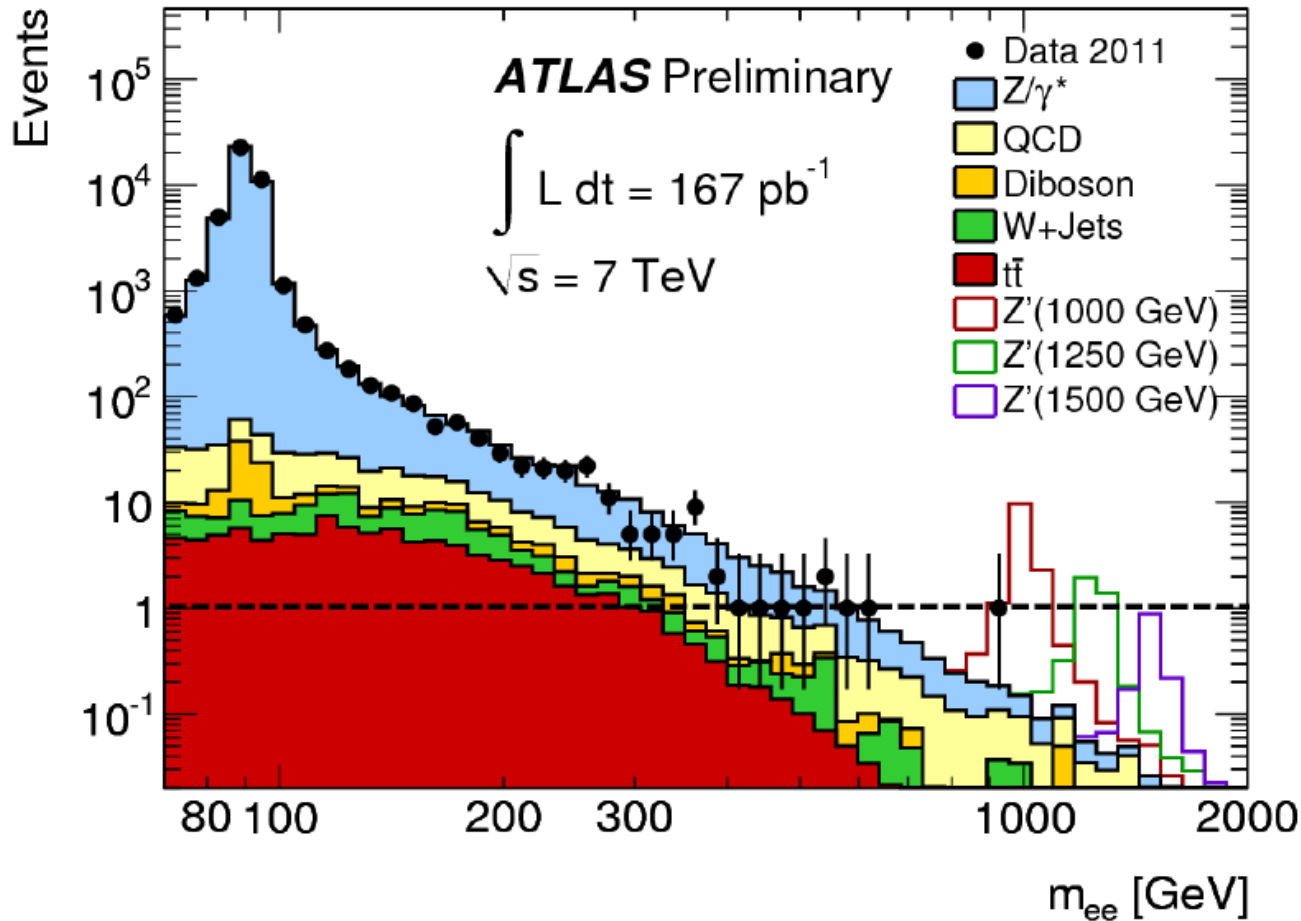
Simple search example

SIMPLE SEARCH EXAMPLE:

SEARCH FOR A HEAVY Z'

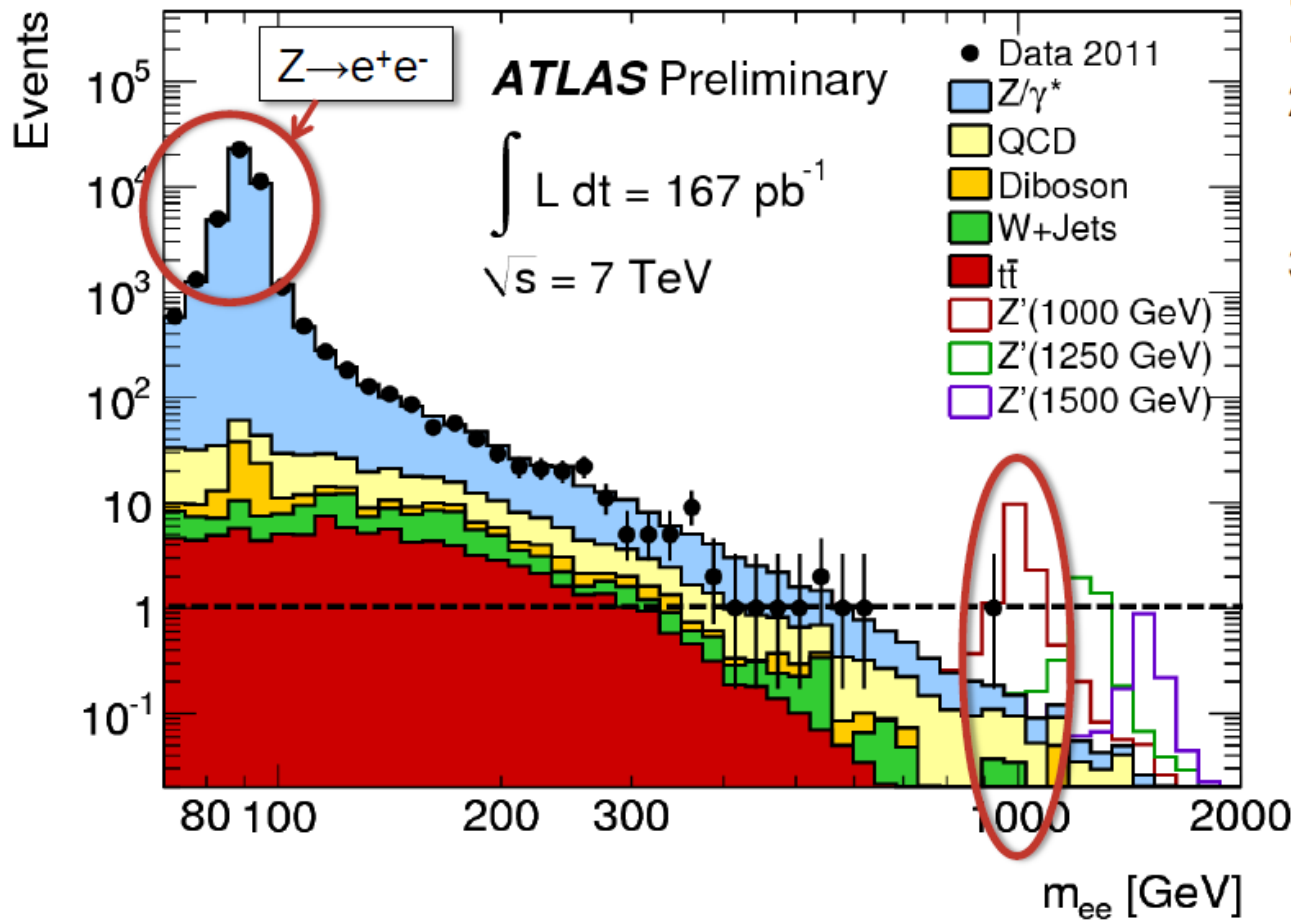
Search for a new heavy Z'

© Like $Z \rightarrow ee$ but at higher mass.



Search for a new heavy Z'

© Like $Z \rightarrow ee$ but at higher mass.



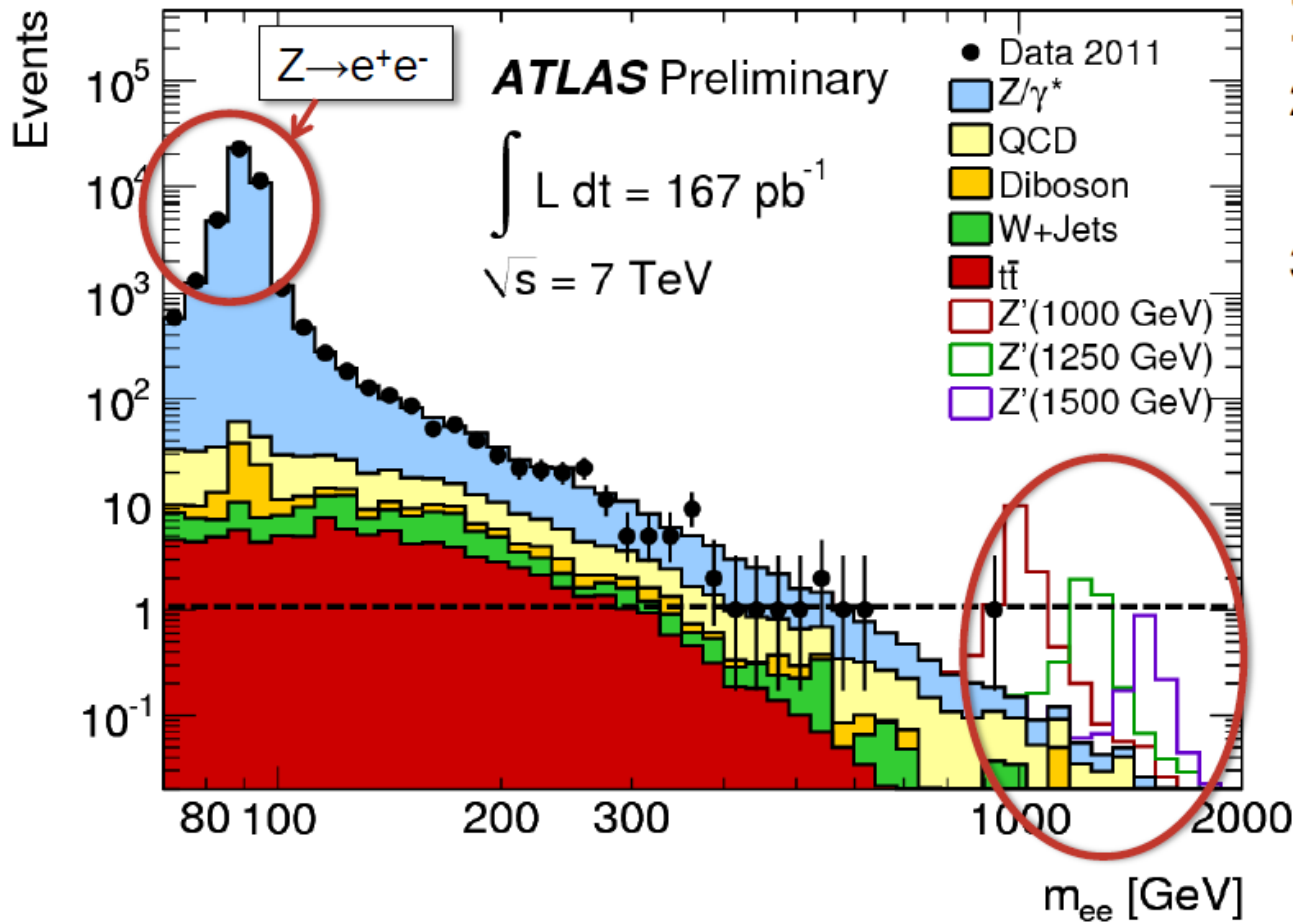
Select 2 electron candidates and plot their invariant mass for:

1. **Data**
2. **Simulated background events**
3. **Simulated signal with different masses**

Data inconsistent with a 1 TeV Z'

Search for a new heavy Z'

© Like $Z \rightarrow ee$ but at higher mass.



Select 2 electron candidates and plot their invariant mass for:

1. **Data**
2. **Simulated background events**
3. **Simulated signal with different masses**

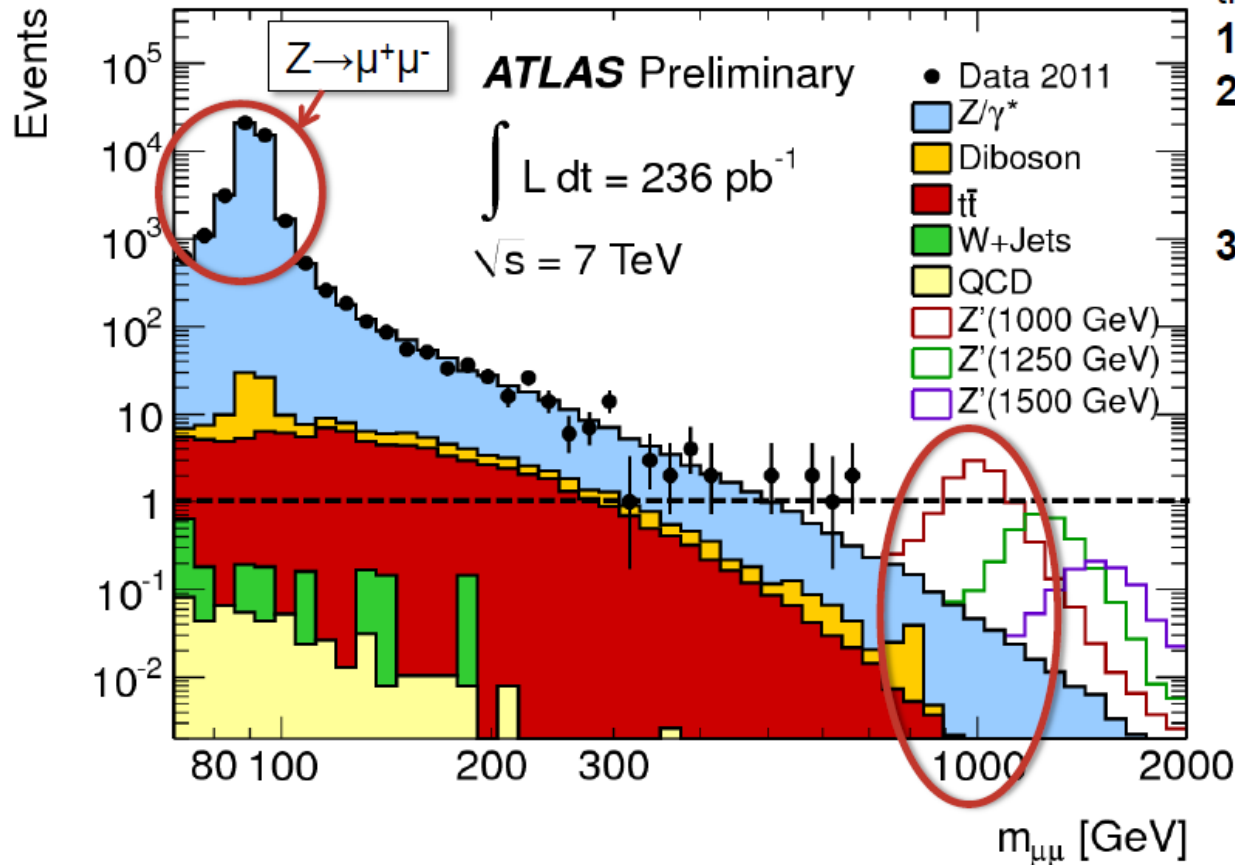
Cross-section decreases with mass (higher the mass of the Z' , the more data needed to discover it)

Search for a new heavy Z'

© **And similar for muons**

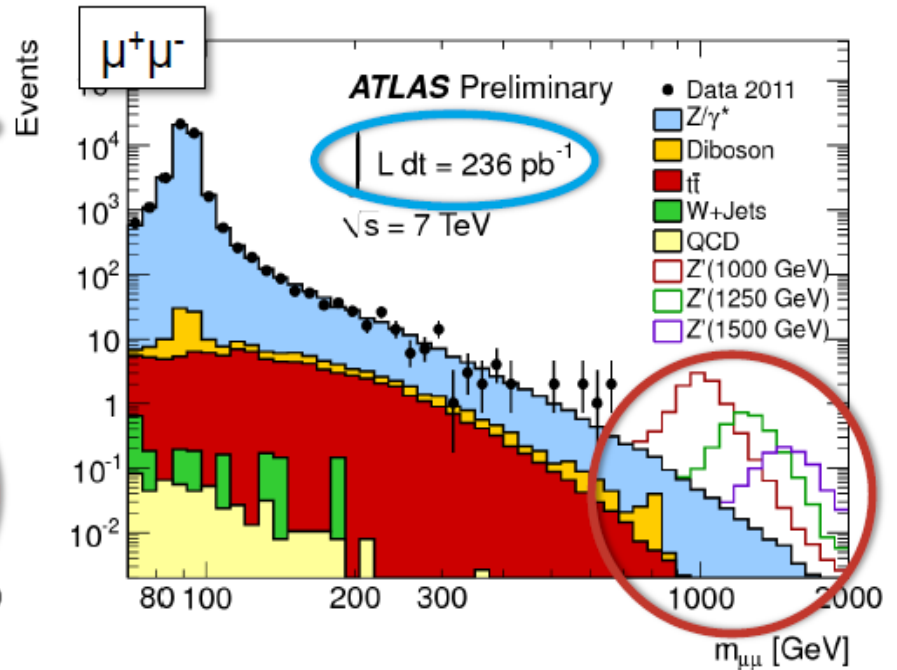
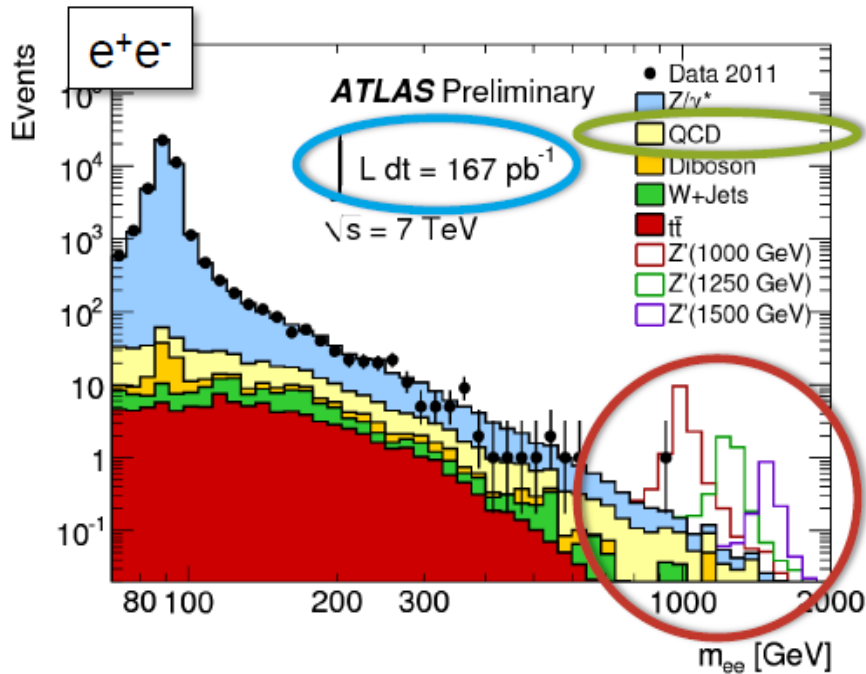
Select 2 muon candidates and plot their invariant mass for:

1. **Data**
2. **Simulated background events**
3. **Simulated signal with different masses**



Data inconsistent with a 1TeV Z'

A small comparison



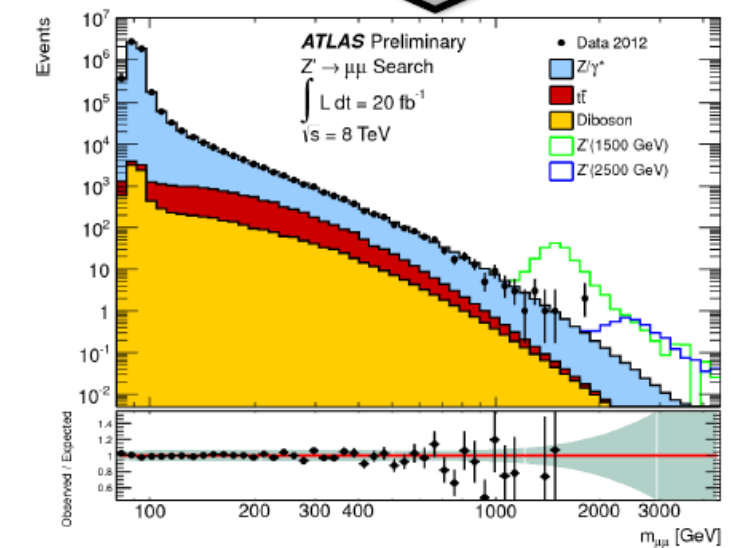
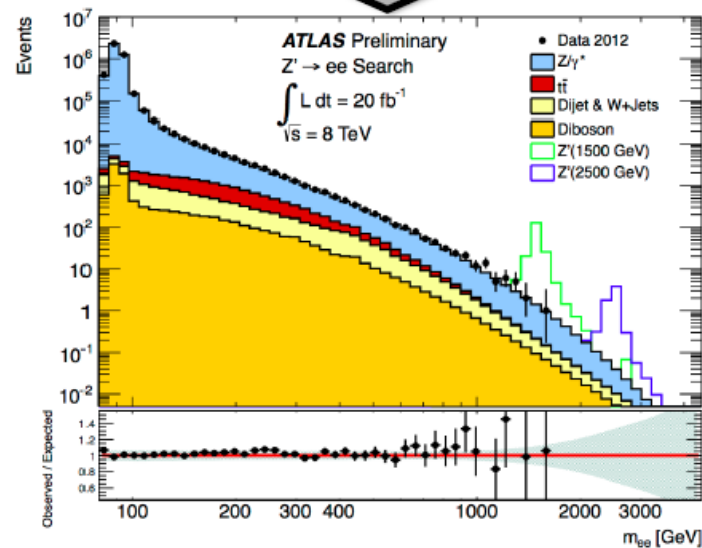
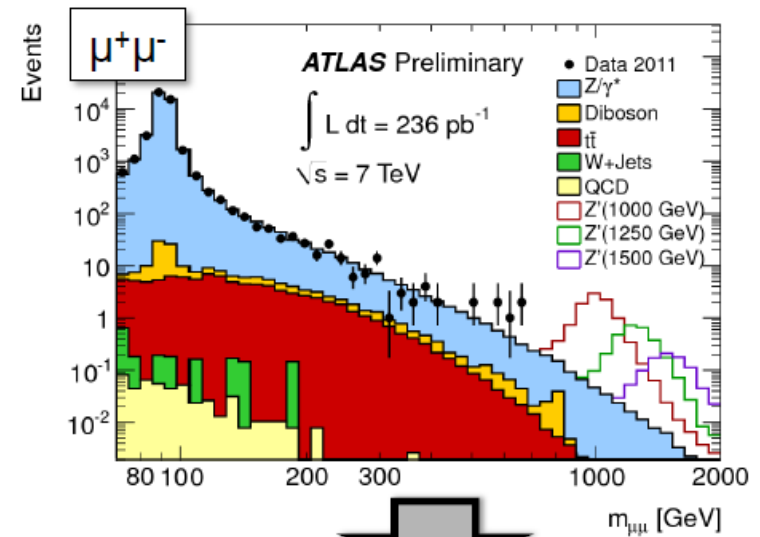
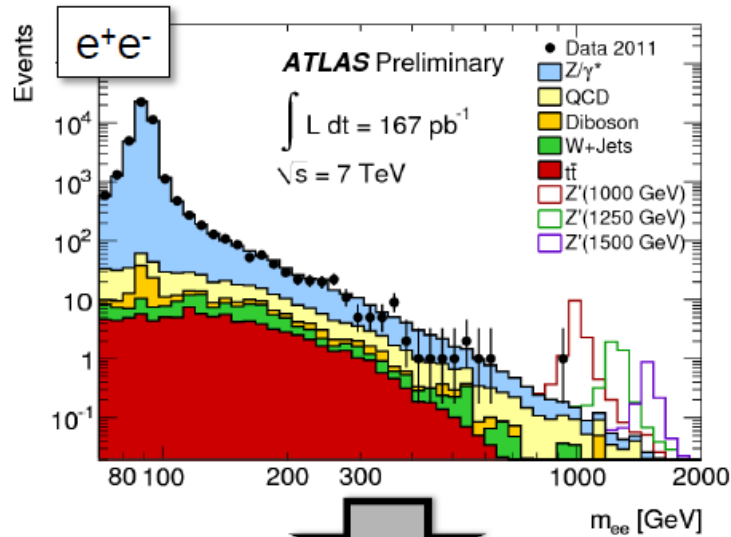
Differences in:

⊙ **Resolution**

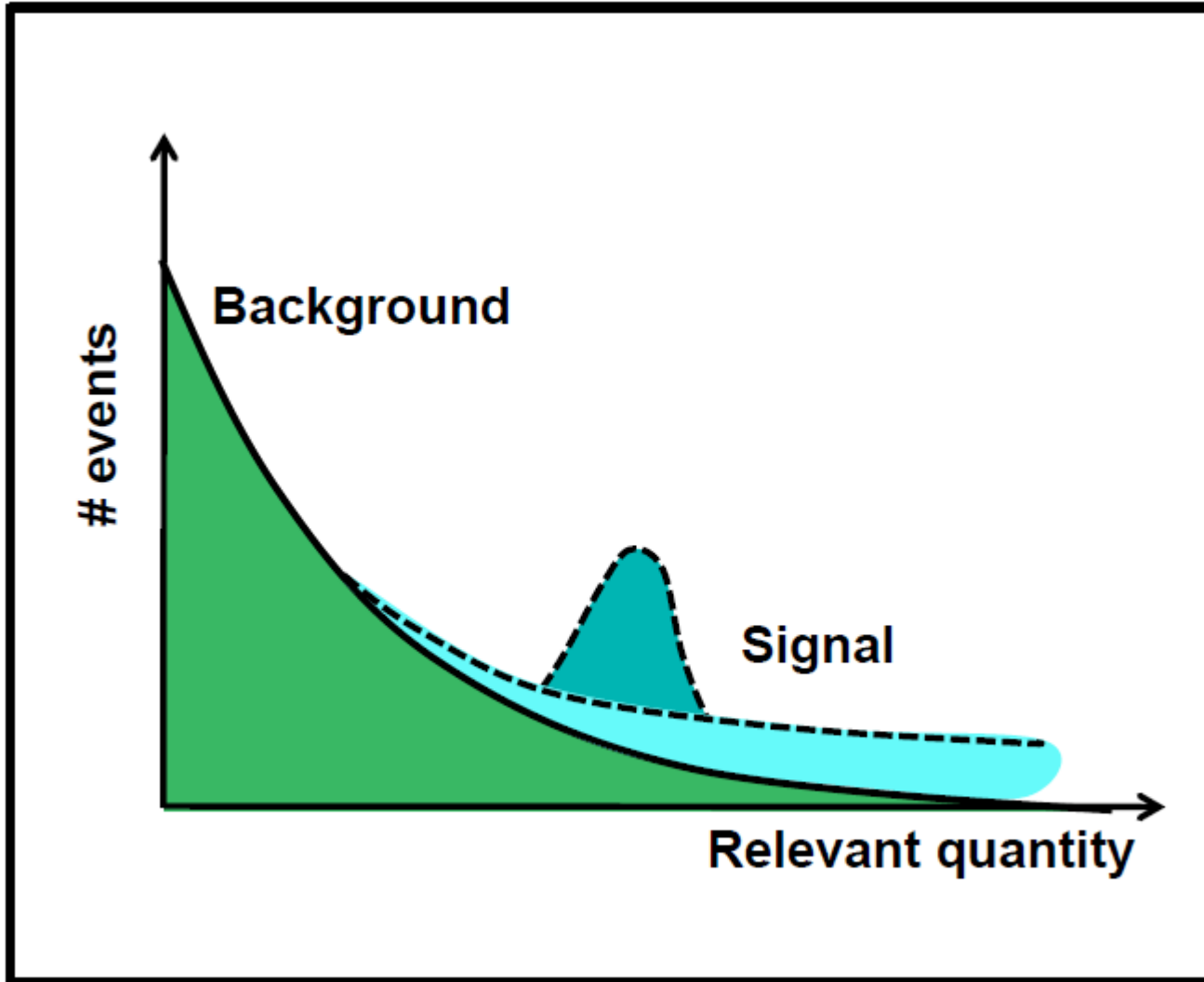
⊙ **Background composition**

⊙ **Dataset**

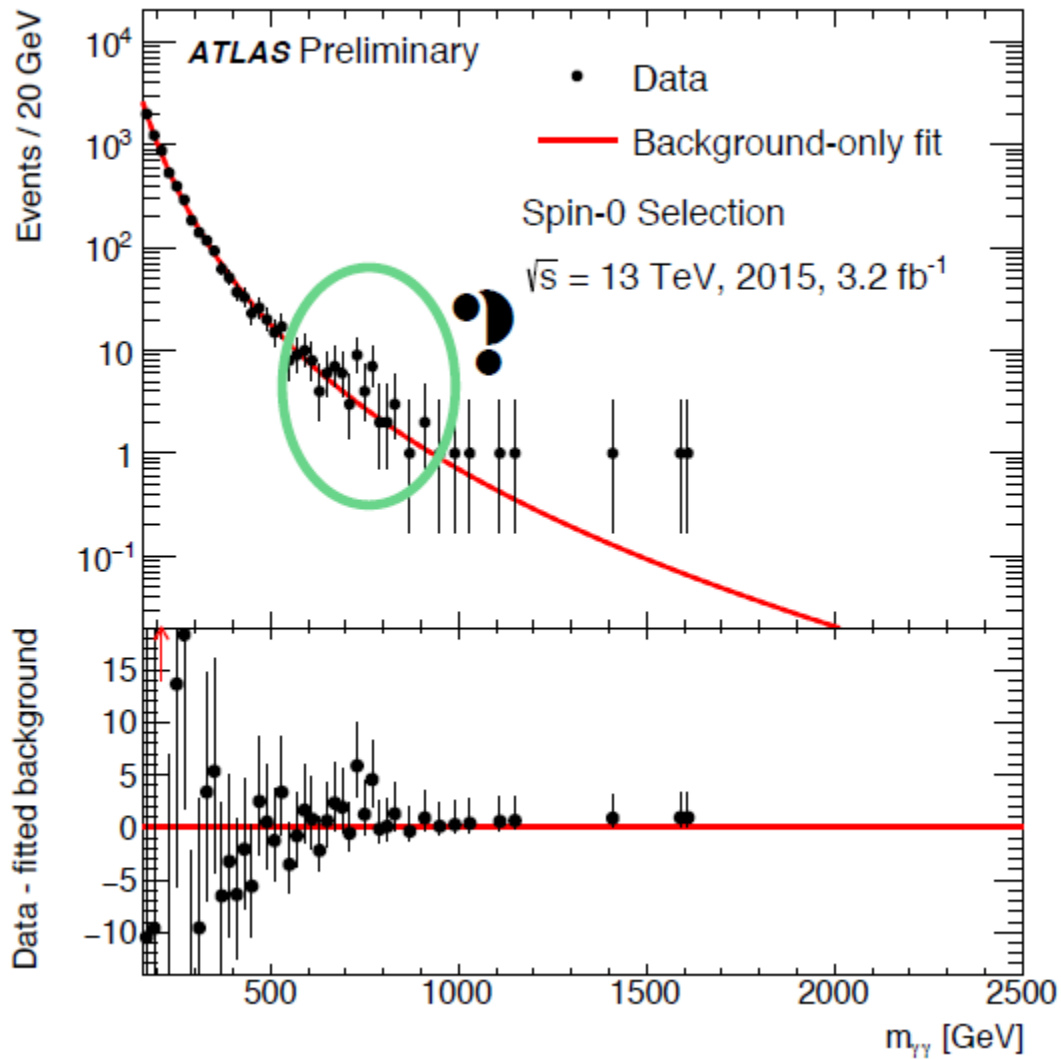
Evolution...



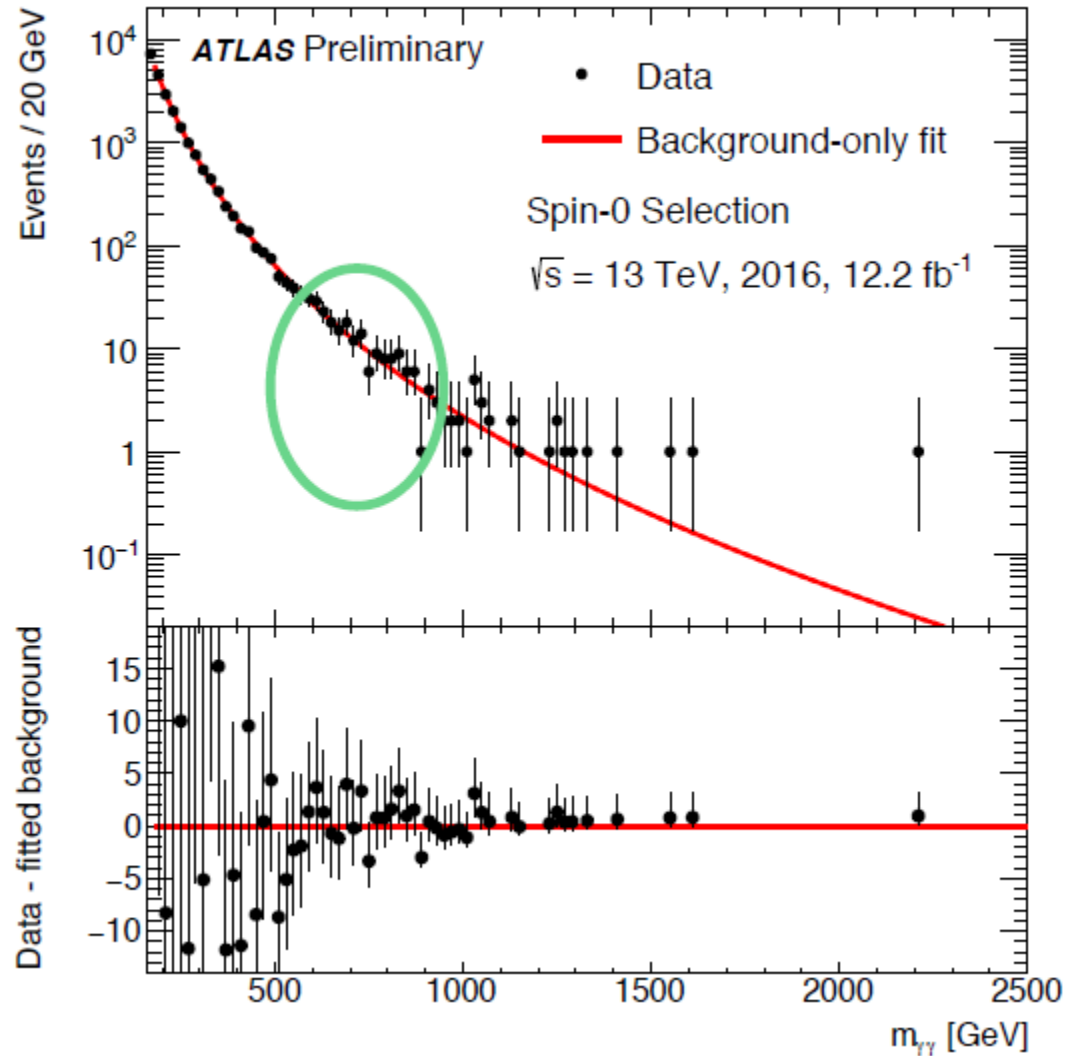
Searches



A well known bump search

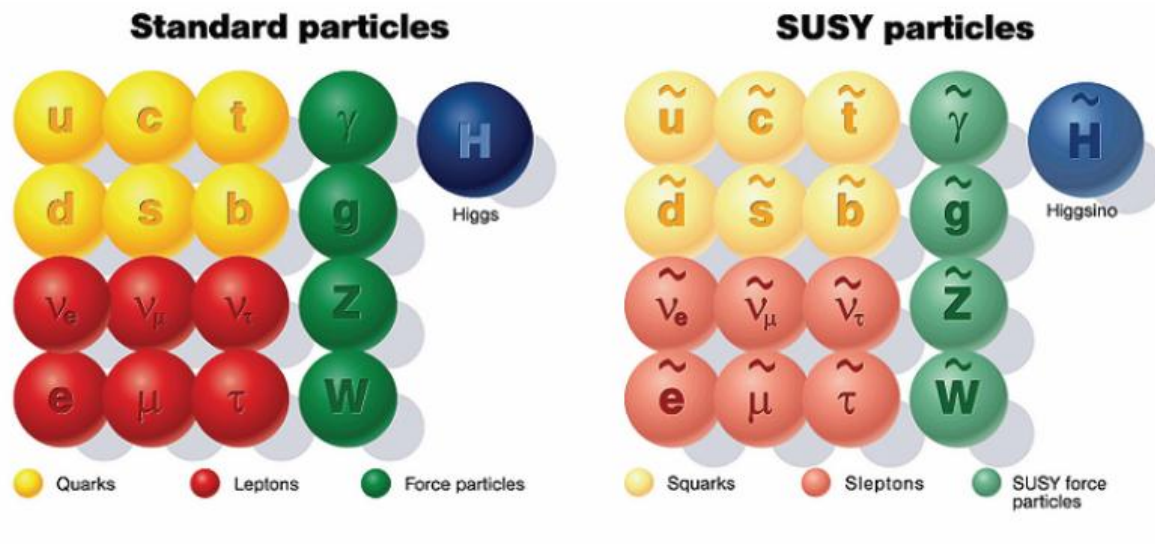


A well known bump search



Typical SUSY searches

- Super-symmetry?

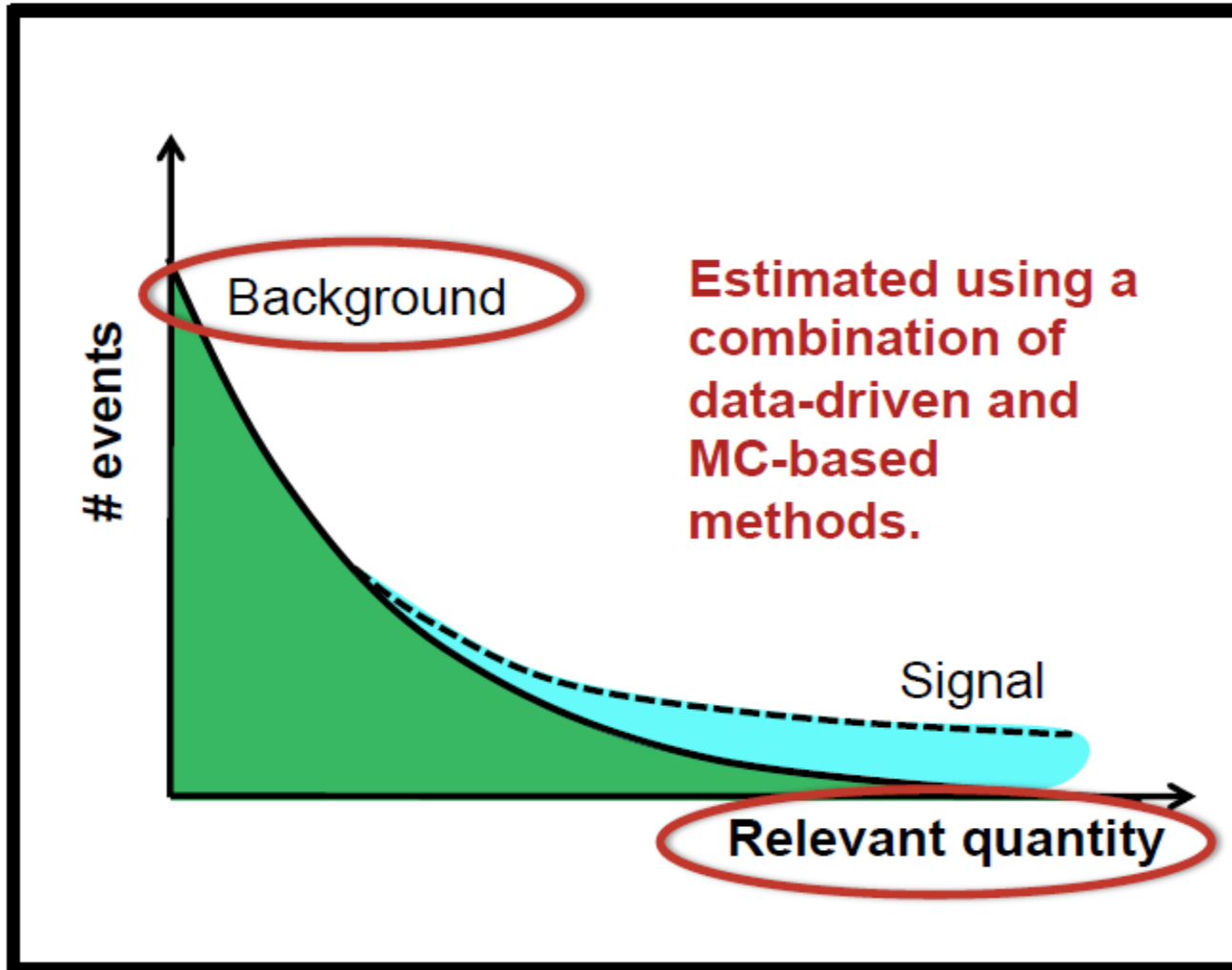


- Composite quark and/or leptons?
- New Heavy bosons?
- Gravitons?
- Dark Matter particles?
- ...

u	c	t	g
d	s	b	γ
ν_e	ν_μ	ν_τ	W
e	μ	τ	Z

Any new theory need to agree with the SM!

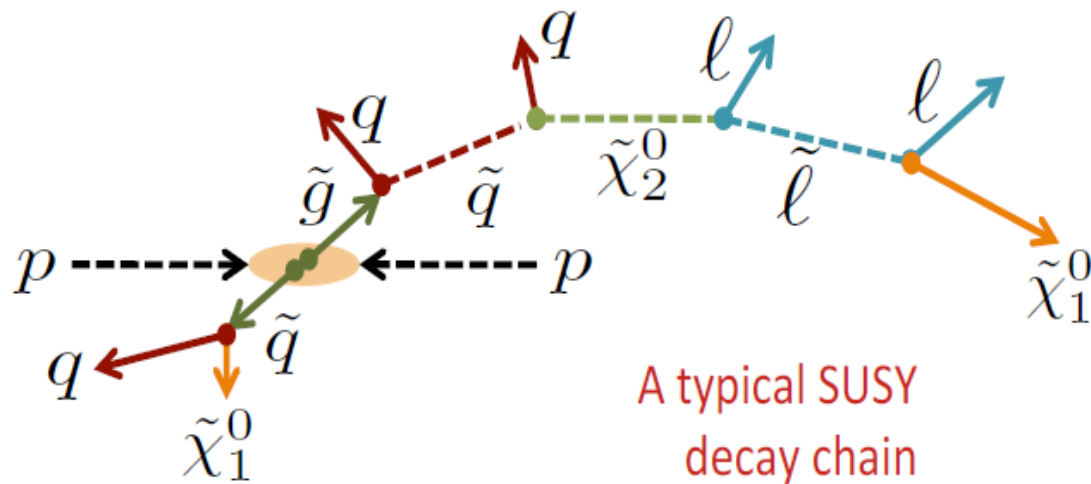
Typical SUSY searches



E.g. MET

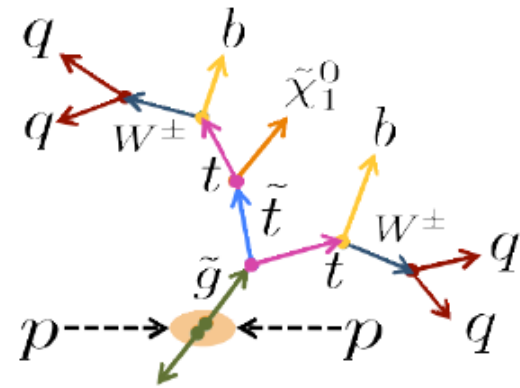
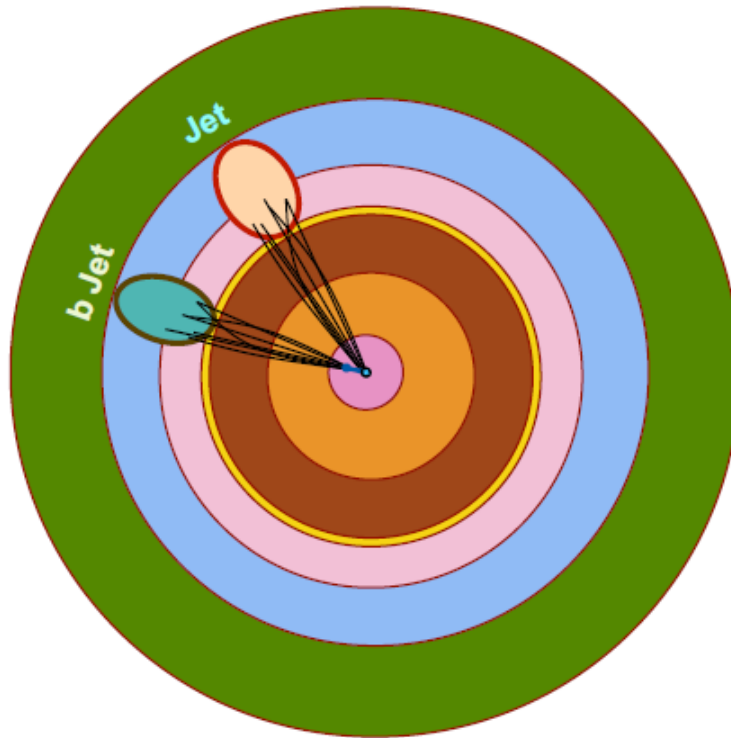
Another search example

SEARCH FOR SUSY IN EVENTS WITH LARGE JET MULTIPLICITIES



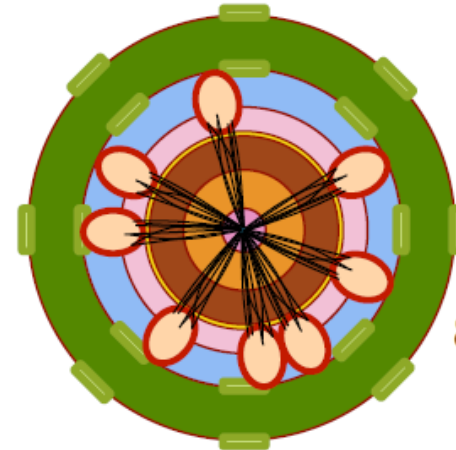
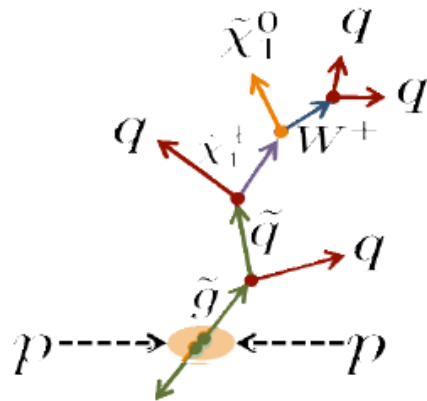
Event selection

b-jets

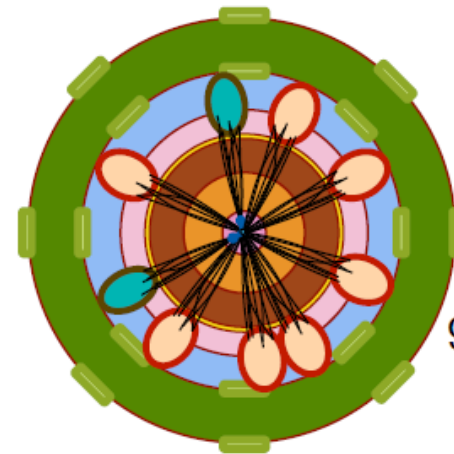
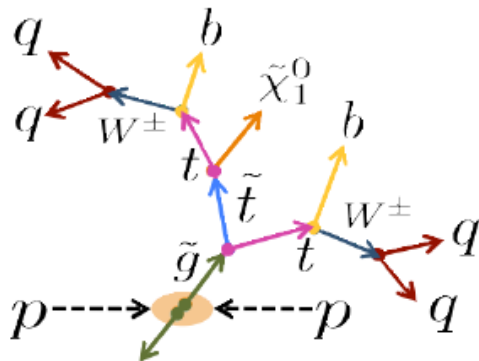


Event selection

b-jets



8 jets, 0 b-jets

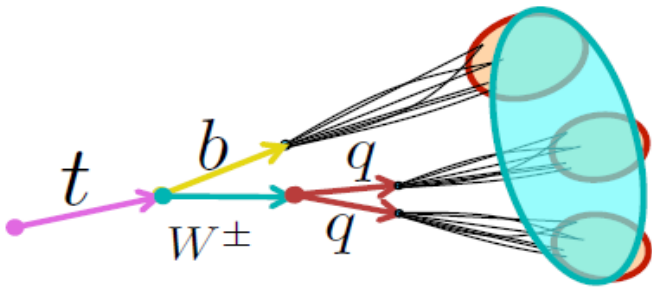
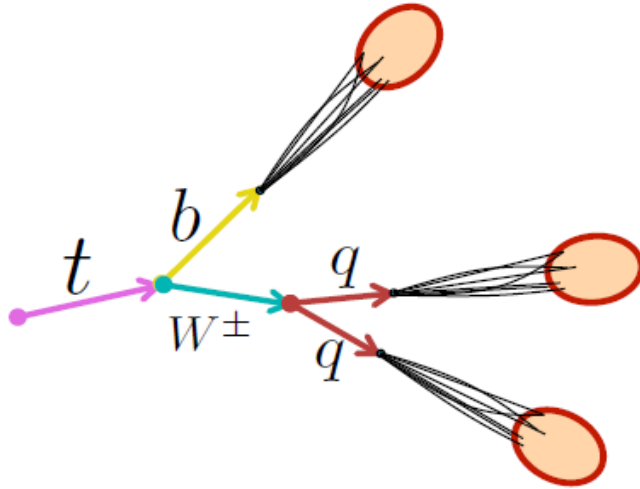


9 jets, ≥ 2 b-jets

Signal regions can range in jet p_T and jet & b-jet multiplicity.

Event selection

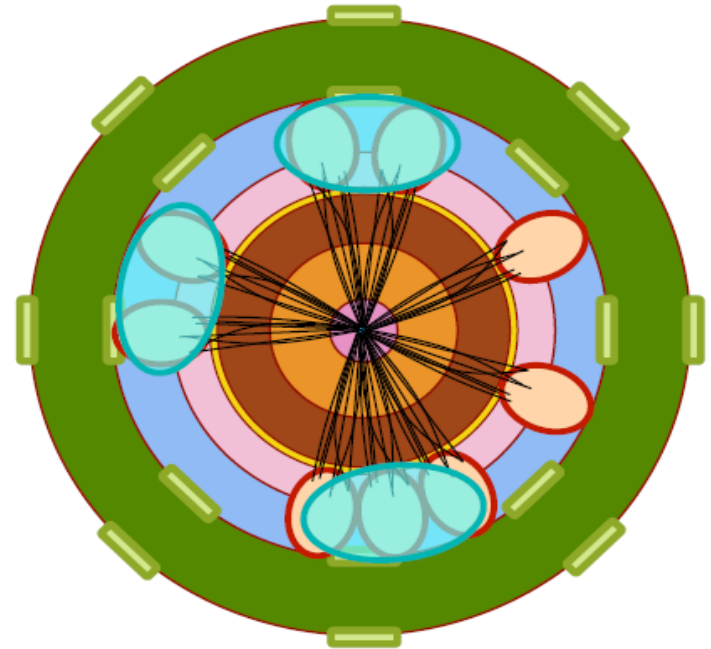
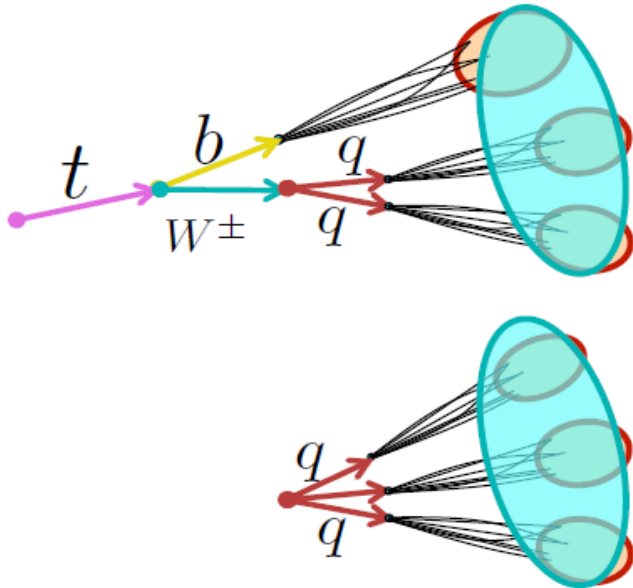
“fat-jets”



Fat-jets are a key signature in searches for boosted objects, e.g. boosted tops.

Event selection

“fat-jets”



$$m_j \text{ (QCD)} < m_j \text{ (SUSY)}$$

Proposed in arXiv:1202.0558

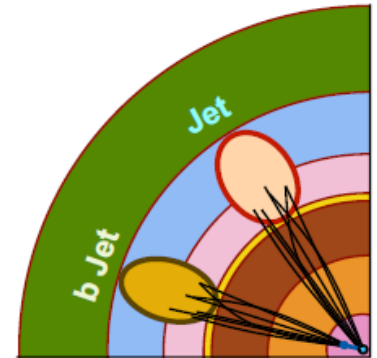
$$M_J^\Sigma = \sum_{i=1}^{nJ} m_{j_i}$$

Signal regions can range in jet multiplicity and M_J^Σ cuts.

Example of search

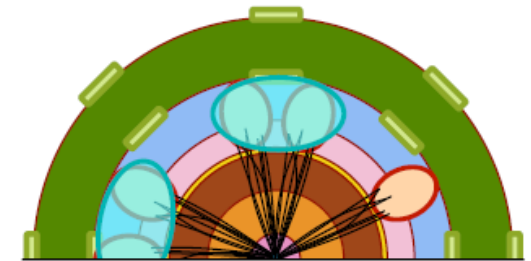
“b-jet stream”

ID	8j50			9j50			≥10j50			7j80			≥8j80		
Jet $ \eta $	< 2.0														
Jet p_T	50 GeV						80 GeV								
Jet count	=8			=9			≥10			=7			≥8		
b-jets	0	1	≥2	0	1	≥2	-			0	1	≥2	0	1	≥2
$ME_T/\sqrt{H_T}$	> 4 GeV ^{1/2}														



“fat-jet stream”

ID	≥8j50		≥9j50		≥10j50	
Jet $ \eta $	< 2.8					
Jet p_T	50 GeV					
Jet count	≥8		≥9		≥10	
M_J^Σ (GeV)	>340	>420	>340	>420	>340	>420
$ME_T/\sqrt{H_T}$	> 4 GeV ^{1/2}					



Proposed in arXiv:1202.0558

$$M_J^\Sigma = \sum_{i=1}^{n_J} m_{j_i}$$

Results

b-jet stream

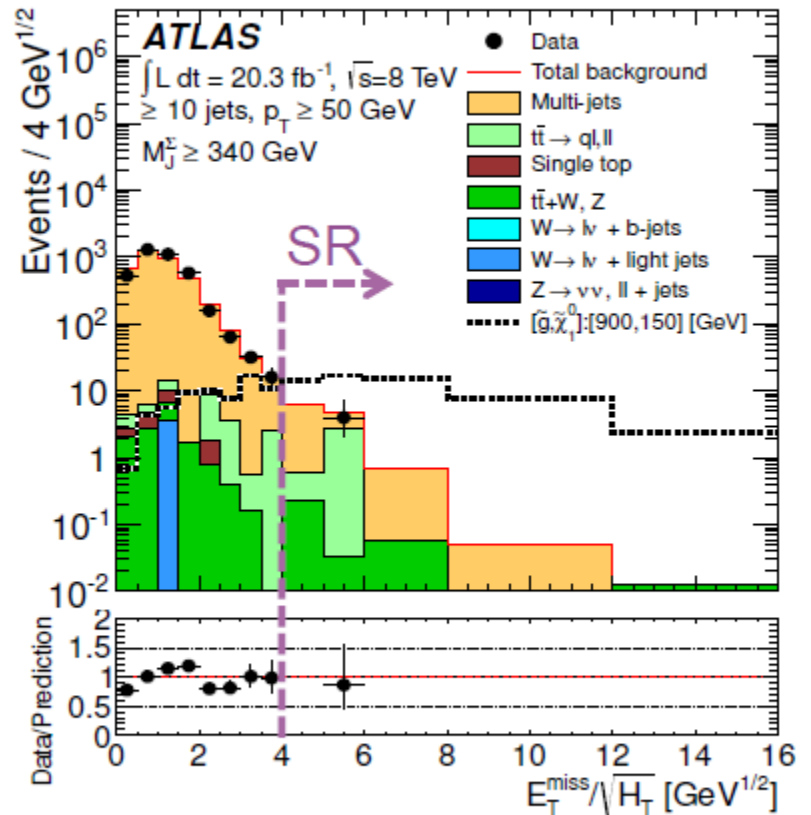
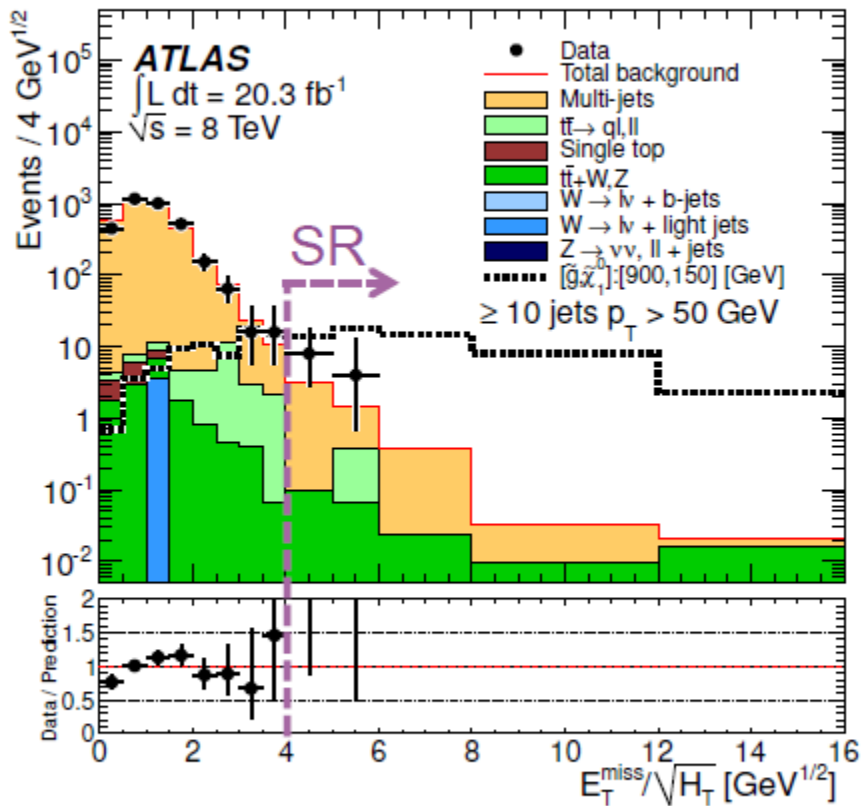
ID	8j50			9j50			≥10j50
b-jets	0	1	≥2	0	1	≥2	0
Expected evts	35±4	40±10	50±10	3.3±0.7	6.1±1.7	8.0±2.7	1.37±0.35
Observed evts	40	44	44	5	8	7	3
Significance (σ)	0.7	-0.02	-0.6	0.8	0.6	-0.28	1.11

ID	7j80			≥8j80		
b-jets	0	1	≥2	0	1	≥2
Expected evts	11.0±2.2	17±6	25±10	0.9±0.6	1.5±0.9	3.3±2.2
Observed evts	12	17	13	2	1	3
Significance (σ)	0.05	-0.14	-1.0	0.9	-0.28	-0.06

fat-jet stream

ID	≥8j50		≥9j50		≥10j50	
M_J^Σ (GeV)	340	420	340	420	340	420
Expected evts	75±19	45±14	17±7	11±5	3.2±3.7	2.2±2.0
Observed evts	69	37	13	9	1	1
Significance (σ)	-0.27	-0.6	-0.6	-0.34	-0.8	-0.6

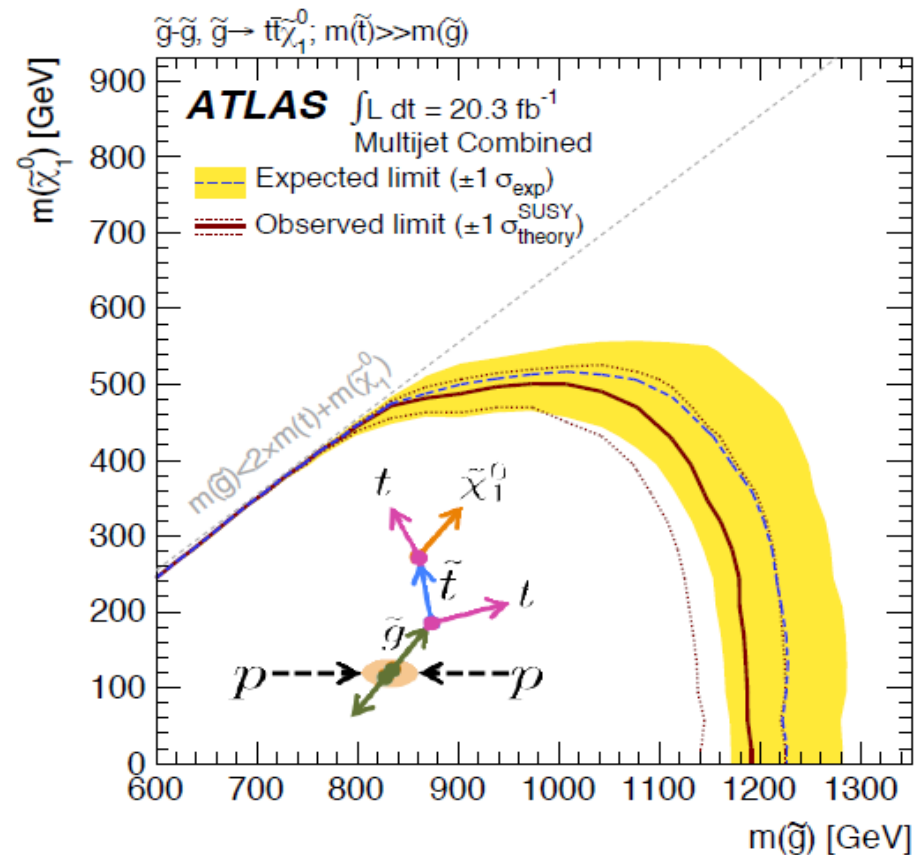
Results



Interpretations

Real or Simplified models

- ⊙ Simplified topologies include typically one production and one decay process. Provide useful information for theorists.



CONTENTS OF A RANDOM SEARCH INTERNAL DOCUMENTATION

26	Contents	
27	1 Introduction	3
28	2 Differences compared to the previous multijet analysis	3
29	3 Data-set and Monte Carlo samples	5
30	3.1 Data Sample	5
31	3.2 Standard Model Monte-Carlo	5
32	3.3 Signal Models	5
33	4 Trigger	10
34	5 Object selection and event cleaning	
35	5.1 Definition of Primary Objects	
36	5.2 General Analysis Cuts	
37	6 Effect of Pile-up on Jets	
38	7 Event selection	
39	7.1 Signal region optimization	
40	7.2 Jet mass selection	
41	7.3 Signal region definitions	
42	8 Standard Model background estimation	
43	8.1 Multi-Jet Production	
44	8.1.1 Central jets, split by b -tag	
45	8.1.2 Total Jet Mass, M_J^Σ	
46	8.2 Leptonic backgrounds	
47	8.2.1 $t\bar{t}$ and W +jets Background	
48	8.2.2 Z +jets Background	
49	8.3 Other SM Processes	
50	8.4 Summary	
51	9 Systematic uncertainties	
52	9.1 Experimental Uncertainties	
53	9.2 Theoretical Uncertainties	
54	9.2.1 $t\bar{t}$ Systematics	
55	9.2.2 W +jets Systematics	
56	9.2.3 $W + b/\bar{b}/b\bar{b}$ +jets Systematics	
57	9.2.4 $Z \rightarrow \nu\nu$ +jets Systematics	
58	9.2.5 Single Top Systematics	
59	9.2.6 $t\bar{t}$ + W/Z Systematics	
60	10 Statistical Methods	
61	11 Results and Interpretation	
62	11.1 Fit results	
63	11.1.1 b -jet analysis stream	
64	11.1.2 M_J^Σ analysis stream	

65	11.2 Interpretation	156
66	11.2.1 Model independent limits	156
67	11.2.2 Limits	157
68	12 Conclusions	161
69	A B-tagging in Different Monte-Carlo Samples	165
70	B Optimization studies	166
71	B.1 Multijet plus b -jets signal regions	166
72	B.2 Multijet plus M_J^Σ signal regions	173
73	B.3 Multijet plus M_J^Σ control regions - systematic uncertainties	177
74	B.3.1 Reweighting distributions using the leading fat jet p_T	182
75	C Trigger	197
76	D Theory Systematic Variations	203
77	D.1 $t\bar{t}$ Systematic variations	203
78	D.2 Single Top Systematic variations	204
79	D.3 Vector Boson + Jets Systematic variations	204
80	D.4 T bar+ V Systematic variations	204
81	E Sensitivity to SUSY models	208
82	F Signal contamination	229
83	G Heavy Flavour systematics	237
84	H Signal region distributions	242
85	H.1 b -jet SRs	242
86	H.2 M_J^Σ SRs	244
87	I Minor backgrounds	247
88	J Fit tests	250
89	K Stream overlap	250
90	L HEP data	256
91	M Event displays	285

Components of a physics analysis

- ⊙ Data-set and Monte Carlo samples
- ⊙ Trigger
- ⊙ Object definitions and event selections
- ⊙ Background determination
- ⊙ Systematic uncertainties
- ⊙ Statistical methods
- ⊙ Results
- ⊙ [Interpretations]

Components of a physics analysis

- ⊙ Data-set and Monte Carlo samples
- ⊙ Trigger
- ⊙ Object definitions
- ⊙ Background detection
- ⊙ Systematic uncertainties
- ⊙ Statistical methods
- ⊙ Results
- ⊙ [Interpretations]

The data and simulation samples used in the analysis. Data for the measurement / search, simulation to compare data to predictions.

Data-set specifics:

- ⊙ Data quality \Rightarrow Good run list.
- ⊙ Luminosity.

Monte carlo sample specifics:

- ⊙ Generator, tunes.
- ⊙ Statistics.

Components of a physics analysis

◎ Data-set and Monte Carlo samples

◎ Trigger

◎ Object definition

◎ Background determination

◎ Systematic uncertainties

◎ Statistical methods

◎ Results

◎ [Interpretations]

The trigger used to collect the data with.

Trigger specifics:

◎ Prescales; typically unprescaled triggers are used, prescaled triggers for QCD / high stat measurements.

◎ Trigger (in)efficiencies.

Components of a physics analysis

◎ Data-set and Monte Carlo samples

◎ Trigger

◎ Object definitions and event selections

◎ Backgrounds The exact definition of objects (electrons, muon, jets, ...) and how these are combined in selecting events to be analyzed.

◎ Systematics

◎ Statistics

◎ Results

◎ [Interpretation]

Object definition specifics:

- ◎ “Flavor” of the identification (loose, medium, tight).
- ◎ Calibrations.

Event selection specifics:

- ◎ Event cleaning (e.g. from noise and cosmics).
- ◎ Momentum, geom. acceptance and multiplicity of objects.
- ◎ Higher level cuts, such as invariant mass.
- ◎ “**Signal regions**”.

Components of a physics analysis

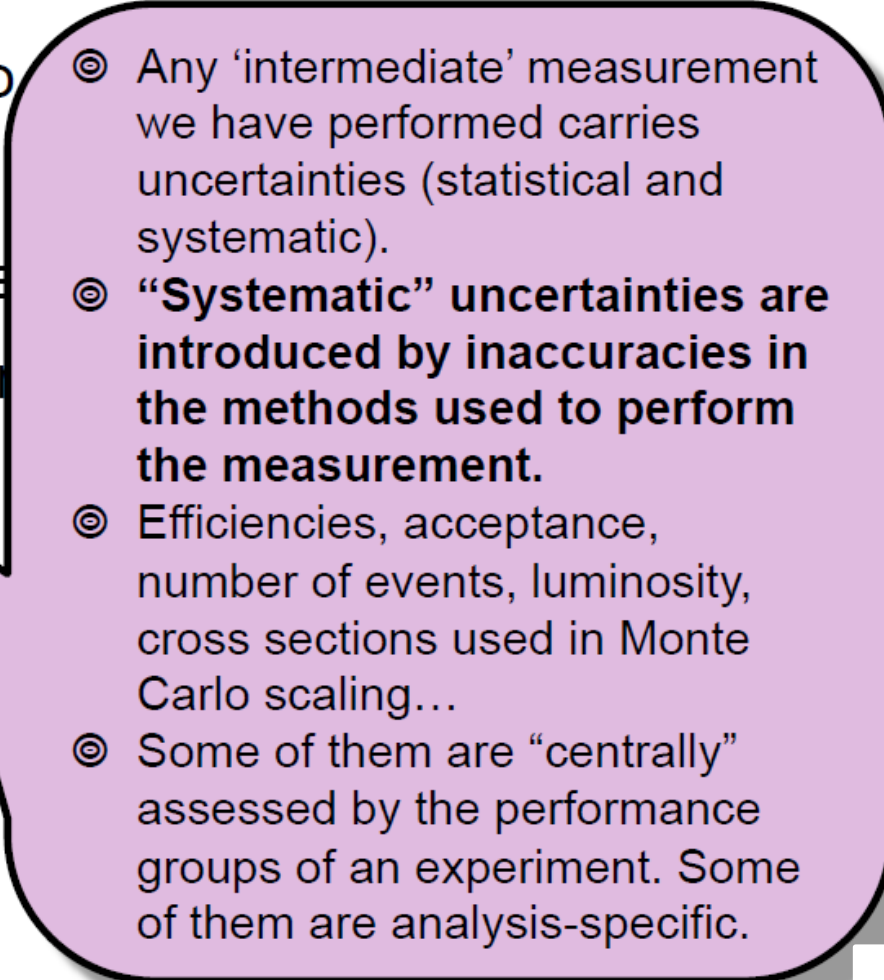
- ⊙ Data-set and Monte Carlo samples
- ⊙ Trigger
- ⊙ Object definitions and event selection
- ⊙ Background determination
- ⊙ Systematic uncertainties
- ⊙ Statistical methods
- ⊙ Results
- ⊙ [Interpretations]

Events that are imitating the signal we are searching for or measuring.

Background determination specifics:

- ⊙ Can/must be **data-driven** or **simulation-based**.
- ⊙ “**Validation regions**” and “**control regions**” required. These can use different triggers wrt signal regions.

Components of a physics analysis

- ◎ Data-set and Monte Carlo
 - ◎ Trigger
 - ◎ Object definitions and event selection
 - ◎ Background determination
 - ◎ Systematic uncertainties
 - ◎ Statistical methods
 - ◎ Results
 - ◎ [Interpretations]
- 
- ◎ Any 'intermediate' measurement we have performed carries uncertainties (statistical and systematic).
 - ◎ **“Systematic” uncertainties are introduced by inaccuracies in the methods used to perform the measurement.**
 - ◎ Efficiencies, acceptance, number of events, luminosity, cross sections used in Monte Carlo scaling...
 - ◎ Some of them are “centrally” assessed by the performance groups of an experiment. Some of them are analysis-specific.

Components of a physics analysis

- ⊙ Data-set and Monte Carlo
- ⊙ Trigger
- ⊙ Object definitions and event selection
- ⊙ Background determination
- ⊙ Systematic uncertainties
- ⊙ Statistical methods
- ⊙ Results
- ⊙ [Interpretations]

Dealing with large data-sets, we use statistical methods to make sense of the numbers we measure.

Typical method:

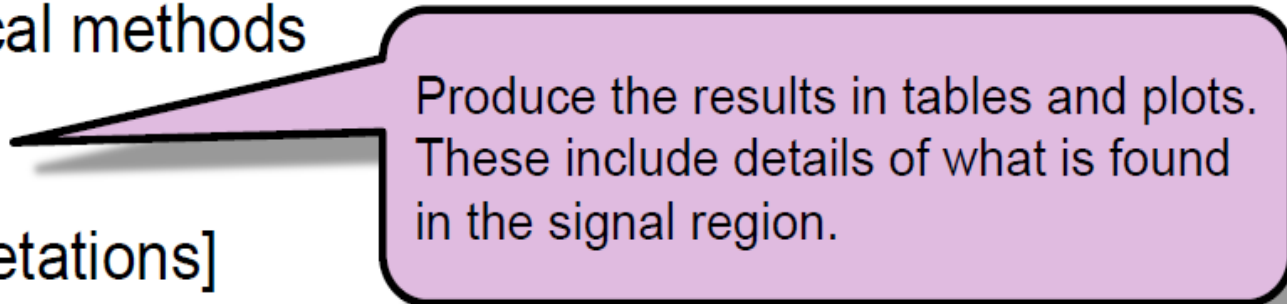
- ⊙ Do a fit to extract signal from background.

Methodologies can vary a lot, but nowadays they are pretty unified within and across experiments.

Neural nets and other machine learning methods are broadly used, primarily to improve signal over background discrimination!

Components of a physics analysis

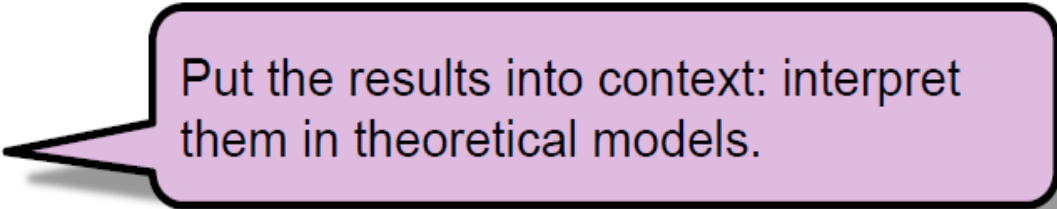
- ◎ Data-set and Monte Carlo samples
- ◎ Trigger
- ◎ Object definitions and event selections
- ◎ Background determination
- ◎ Systematic uncertainties
- ◎ Statistical methods
- ◎ Results
- ◎ [Interpretations]



Produce the results in tables and plots. These include details of what is found in the signal region.

Components of a physics analysis

- ⊙ Data-set and Monte Carlo samples
- ⊙ Trigger
- ⊙ Object definitions and event selections
- ⊙ Background determination
- ⊙ Systematic uncertainties
- ⊙ Statistical methods
- ⊙ Results
- ⊙ [Interpretations]



Put the results into context: interpret them in theoretical models.

(Instead of) conclusions

