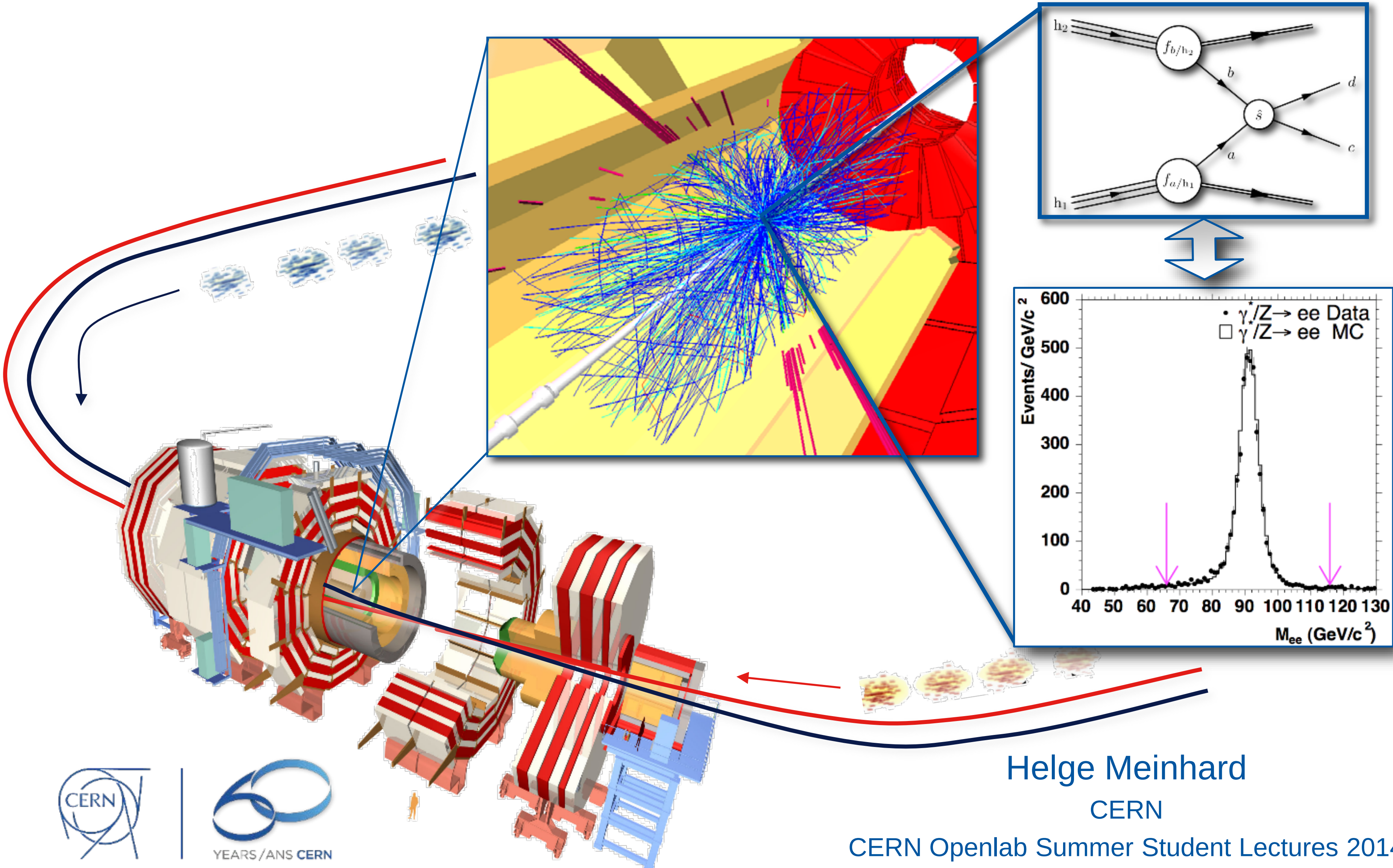
From Detectors to Publications



Helge Meinhard
CERN

CERN Openlab Summer Student Lectures 2014

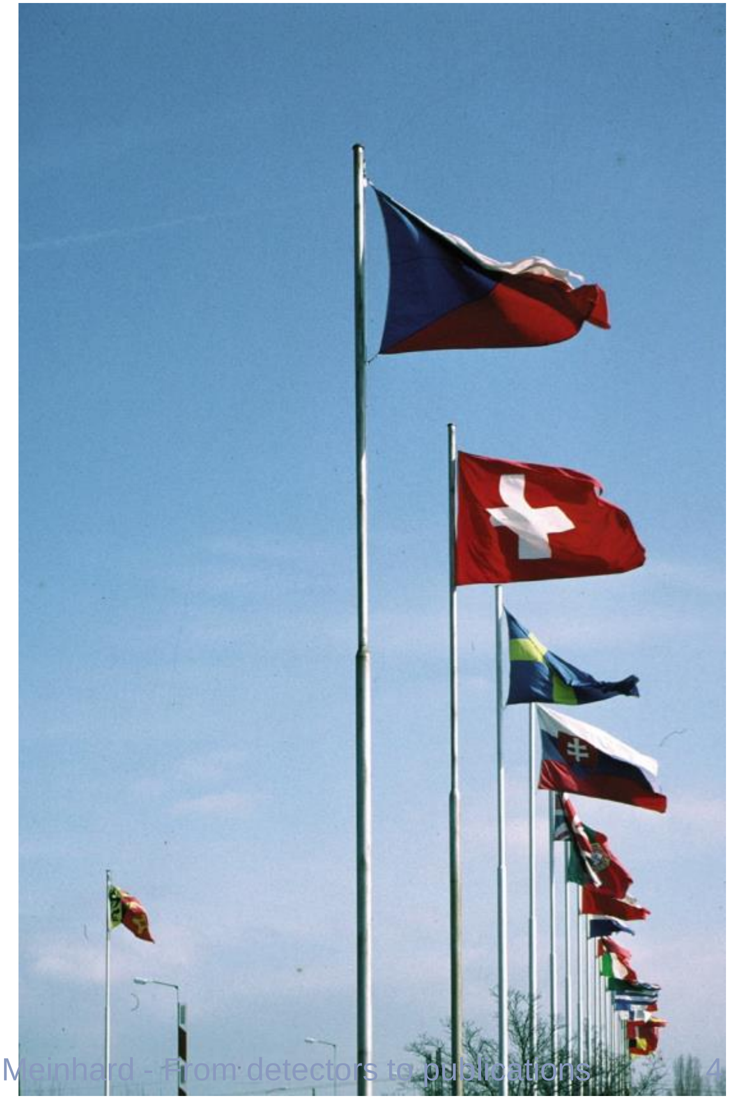
Outline

1. Brief introduction to CERN
2. Overview of process from detectors to results
3. Reconstruction
4. Simulation
5. Physics Analysis
6. Computing
7. Summary

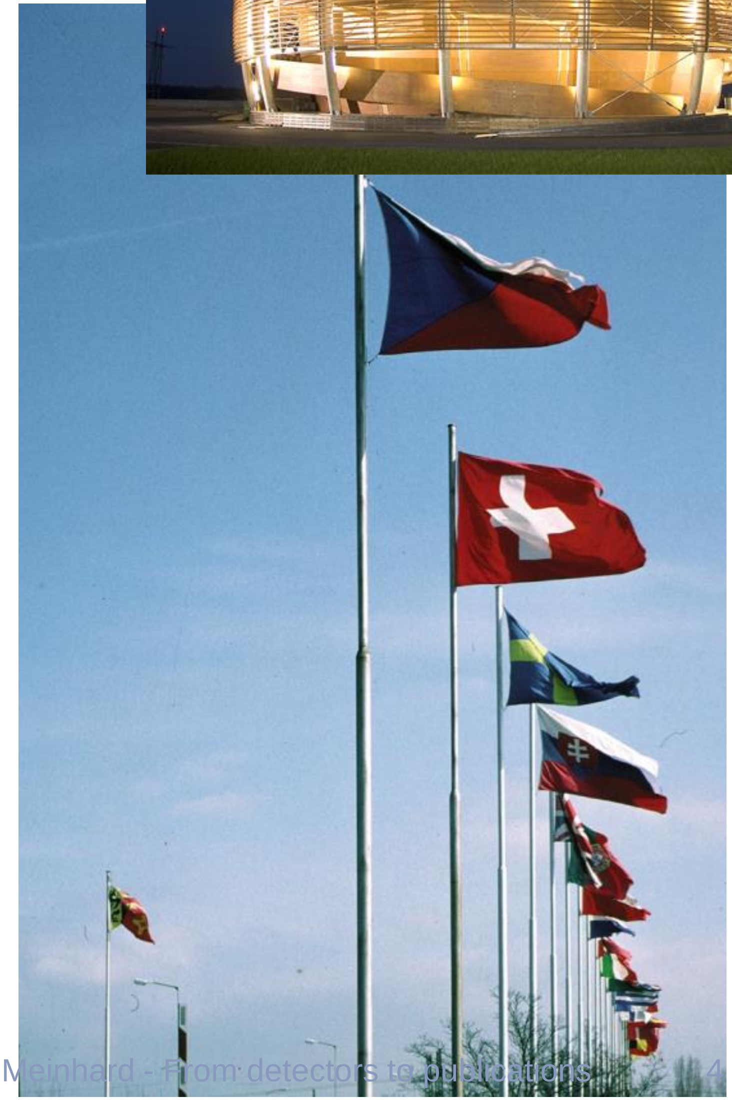
Acknowledgement: Contents partly based on 3-part lecture given in 2013 summer student lecture programme
Thanks to Jamie Boyd et al.

1. Introduction to CERN

CERN

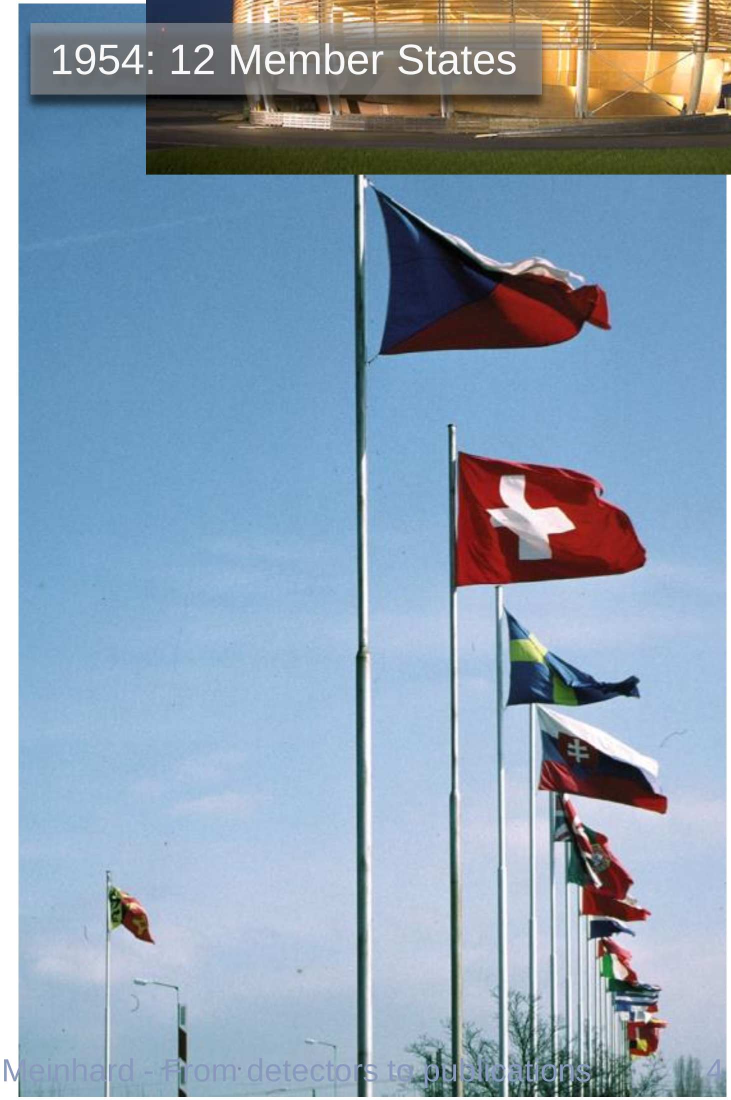


CERN



CERN

- International organisation close to Geneva, straddling Swiss-French border, founded 1954

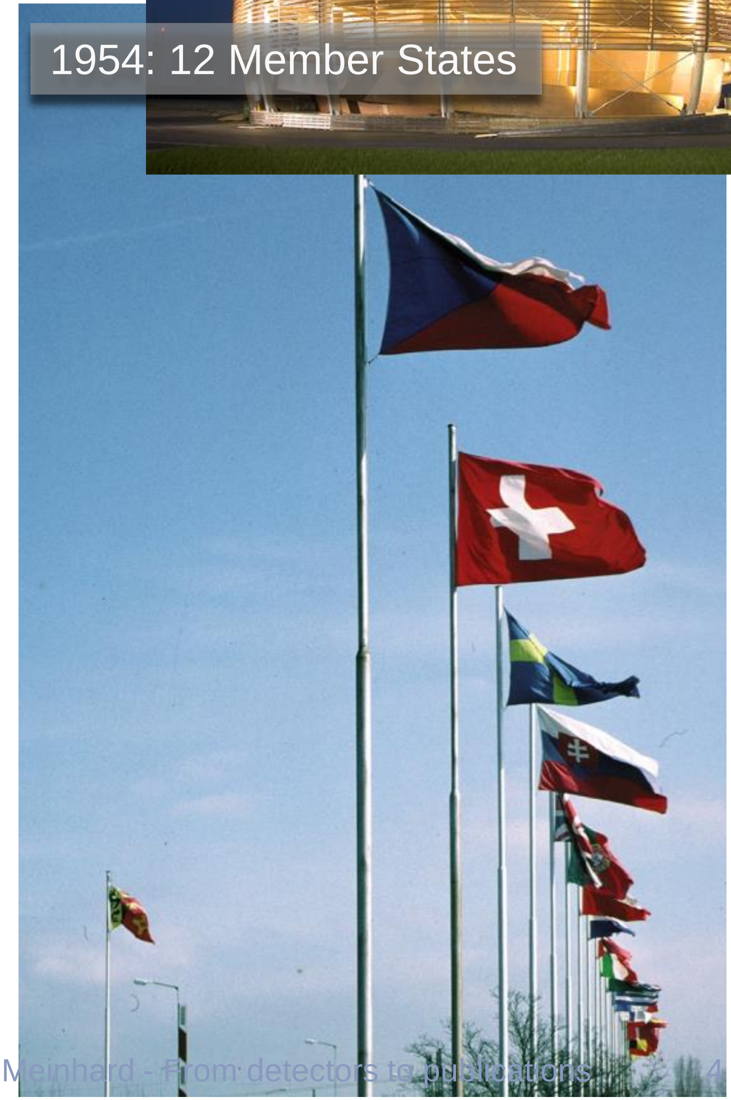


1954: 12 Member States

CERN

- International organisation close to Geneva, straddling Swiss-French border, founded 1954

  Facilities for fundamental research in particle physics



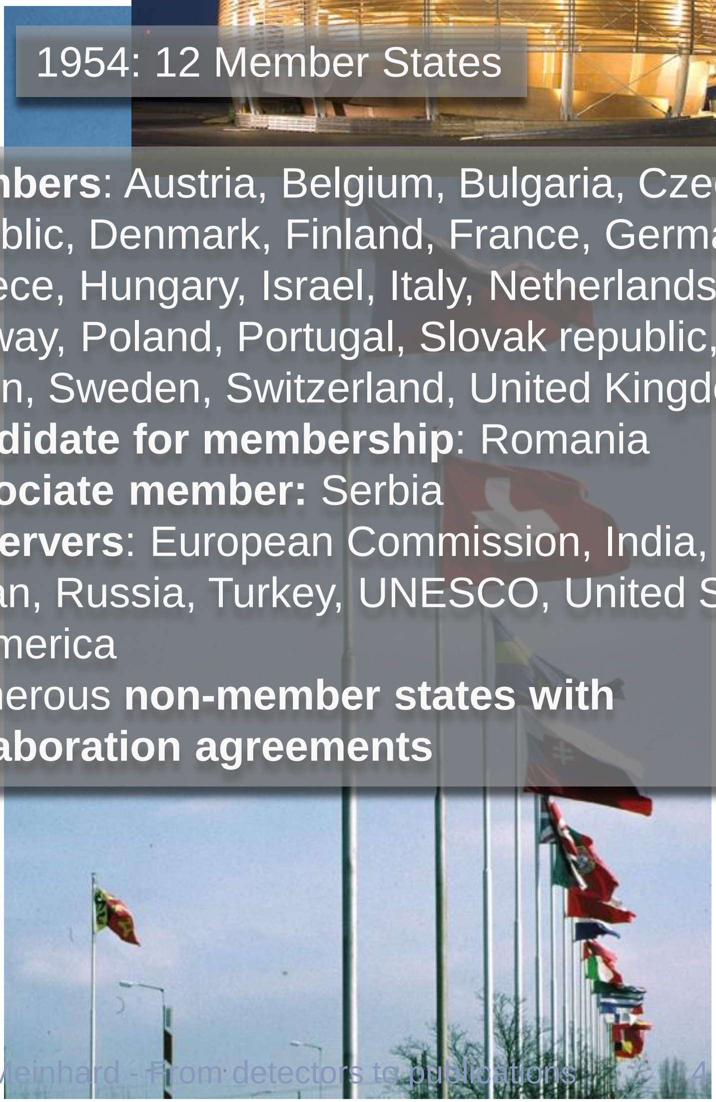
1954: 12 Member States

CERN

- International organisation close to Geneva, straddling Swiss-French border, founded 1954

  Facilities for fundamental research in particle physics
- 21 member states,
  1 B CHF budget

1954: 12 Member States



Members: Austria, Belgium, Bulgaria, Czech republic, Denmark, Finland, France, Germany, Greece, Hungary, Israel, Italy, Netherlands, Norway, Poland, Portugal, Slovak republic, Spain, Sweden, Switzerland, United Kingdom

Candidate for membership: Romania

Associate member: Serbia

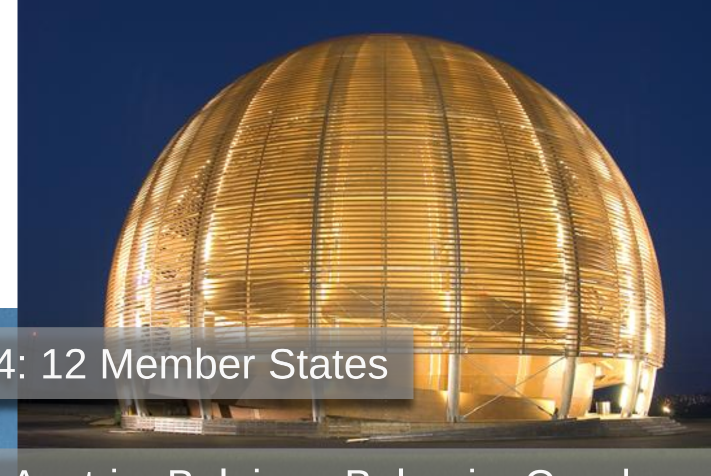
Observers: European Commission, India, Japan, Russia, Turkey, UNESCO, United States of America

Numerous **non-member states with collaboration agreements**

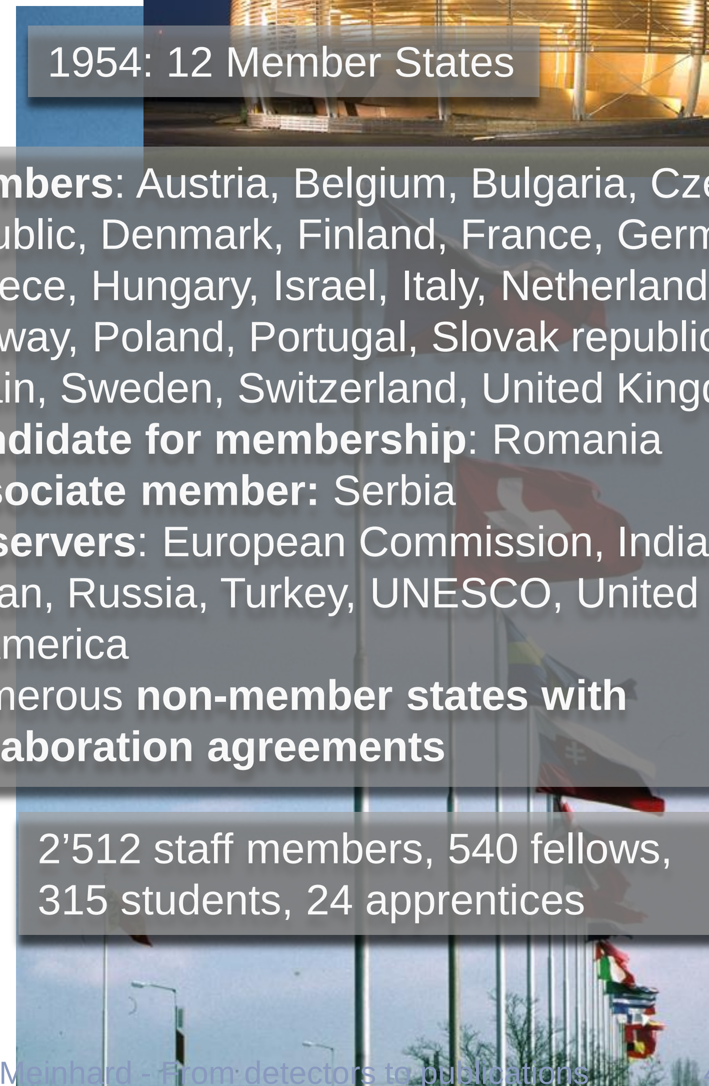
CERN

- International organisation close to Geneva, straddling Swiss-French border, founded 1954

  Facilities for fundamental research in particle physics
- 21 member states,
  1 B CHF budget
- 3’360 staff, fellows, students, apprentices, …



1954: 12 Member States



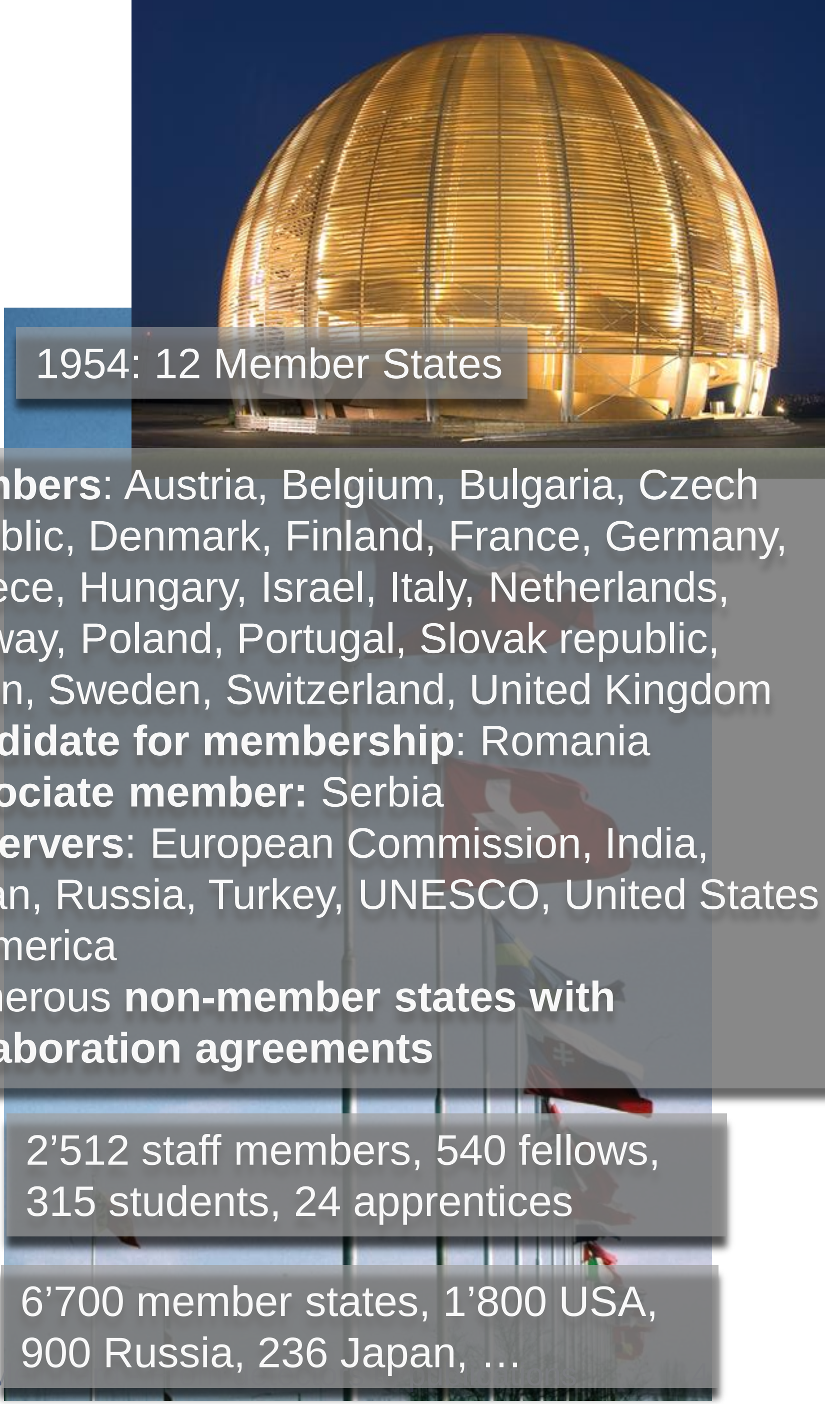
Members: Austria, Belgium, Bulgaria, Czech republic, Denmark, Finland, France, Germany, Greece, Hungary, Israel, Italy, Netherlands, Norway, Poland, Portugal, Slovak republic, Spain, Sweden, Switzerland, United Kingdom
Candidate for membership: Romania
Associate member: Serbia
Observers: European Commission, India, Japan, Russia, Turkey, UNESCO, United States of America
Numerous **non-member states with collaboration agreements**

2’512 staff members, 540 fellows, 315 students, 24 apprentices

CERN

- International organisation close to Geneva, straddling Swiss-French border, founded 1954

  Facilities for fundamental research in particle physics
- 21 member states, 1 B CHF budget
- 3’360 staff, fellows, students, apprentices, …
- 11’000 users



1954: 12 Member States

Members: Austria, Belgium, Bulgaria, Czech republic, Denmark, Finland, France, Germany, Greece, Hungary, Israel, Italy, Netherlands, Norway, Poland, Portugal, Slovak republic, Spain, Sweden, Switzerland, United Kingdom
Candidate for membership: Romania
Associate member: Serbia
Observers: European Commission, India, Japan, Russia, Turkey, UNESCO, United States of America
Numerous **non-member states with collaboration agreements**

2’512 staff members, 540 fellows, 315 students, 24 apprentices

6’700 member states, 1’800 USA, 900 Russia, 236 Japan, …

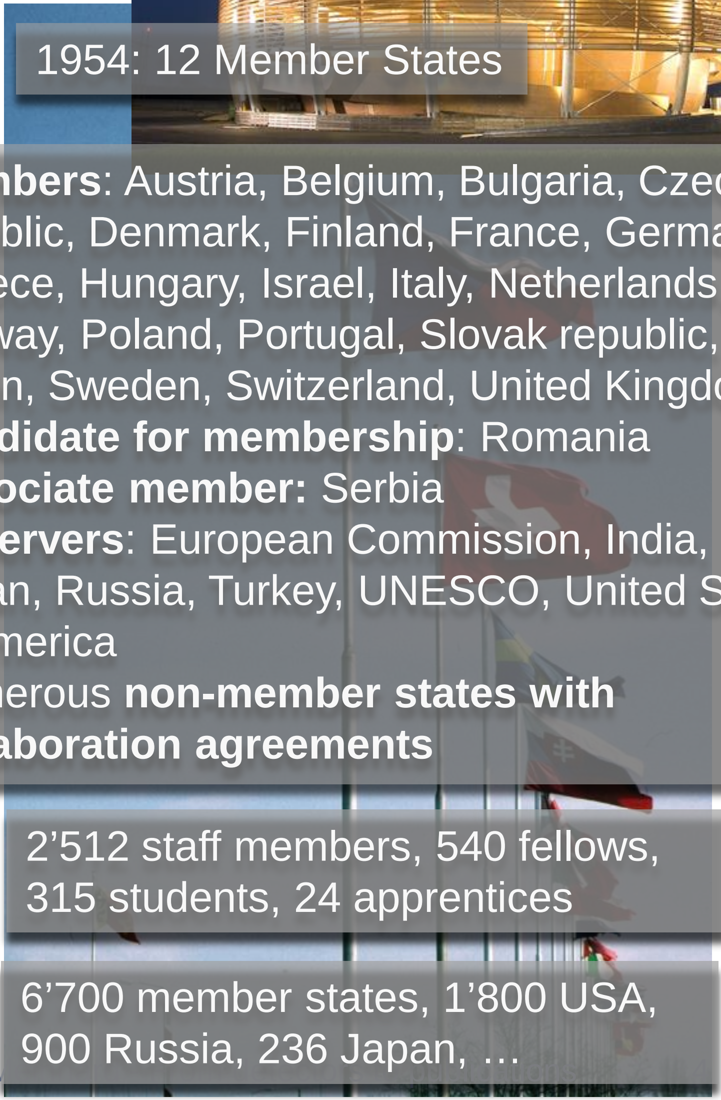
CERN

"Science for peace"

1954: 12 Member States

- International organisation close to Geneva, straddling Swiss-French border, founded 1954

  Facilities for fundamental research in particle physics
- 21 member states, 1 B CHF budget
- 3'360 staff, fellows, students, apprentices, …
- 11'000 users



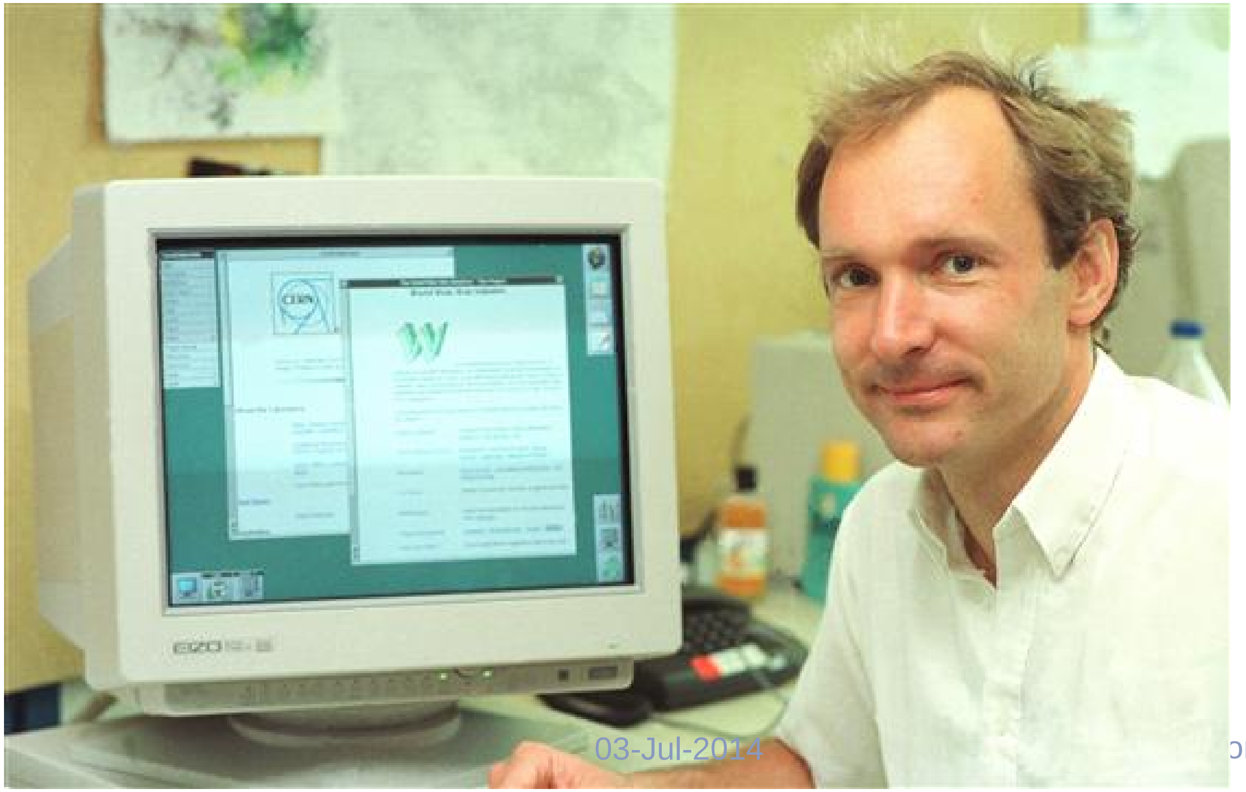
Members: Austria, Belgium, Bulgaria, Czech republic, Denmark, Finland, France, Germany, Greece, Hungary, Israel, Italy, Netherlands, Norway, Poland, Portugal, Slovak republic, Spain, Sweden, Switzerland, United Kingdom
Candidate for membership: Romania
Associate member: Serbia
Observers: European Commission, India, Japan, Russia, Turkey, UNESCO, United States of America
Numerous **non-member states with collaboration agreements**

2'512 staff members, 540 fellows, 315 students, 24 apprentices

6'700 member states, 1'800 USA, 900 Russia, 236 Japan, …

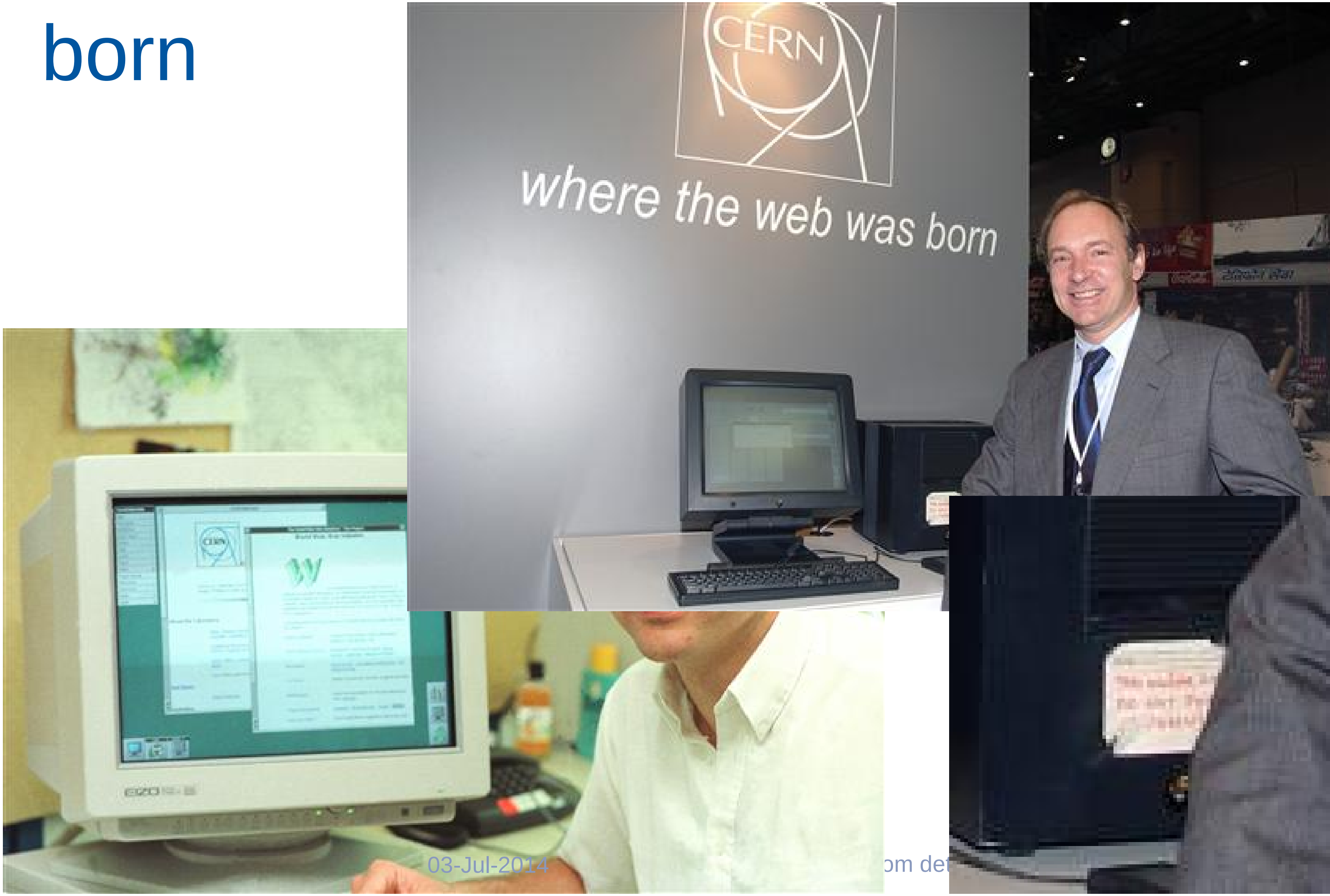
CERN | 60 YEARS/ANS CERN

CERN – where the Web was born



CERN – where the Web was born

CERN – where the Web was born




LHCb
ATLAS
CERN Meyrin
CERN Prévessin
SPS 7 km
SUISSE
FRANCE
ALICE
CMS
LHC 27 km

LHC – Large Hadron Collider

- Proton-proton collider

LHC – Large Hadron Collider

- Proton-proton collider
- 27 km circumference

LHC – Large Hadron Collider

- Proton-proton collider
- 27 km circumference
- Started operation in 2010 with 3.5 + 3.5 TeV, continued in 2011, 4 + 4 TeV in 2012
    - World's most powerful particle accelerator

LHC – Large Hadron Collider

- Proton-proton collider
- 27 km circumference
- Started operation in 2010 with 3.5 + 3.5 TeV, continued in 2011, 4 + 4 TeV in 2012
  - World's most powerful particle accelerator

  Run 1 until early 2013

LHC – Large Hadron Collider

- Proton-proton collider
- 27 km circumference
- Started operation in 2010 with 3.5 + 3.5 TeV, continued in 2011, 4 + 4 TeV in 2012
  - World's most powerful particle accelerator

  Run 1 until early 2013
- Now in Long Shutdown 1 – machine upgrade

LHC – Large Hadron Collider

- Proton-proton collider
- 27 km circumference
- Started operation in 2010 with 3.5 + 3.5 TeV, continued in 2011, 4 + 4 TeV in 2012
  - World's most powerful particle accelerator

  Run 1 until early 2013
- Now in Long Shutdown 1 – machine upgrade
- Restart early 2015 at 6.5 + 6.5 TeV

LHC – Large Hadron Collider

- Proton-proton collider
- 27 km circumference
- Started operation in 2010 with 3.5 + 3.5 TeV, continued in 2011, 4 + 4 TeV in 2012
  - World's most powerful particle accelerator

  Run 1 until early 2013
- Now in Long Shutdown 1 – machine upgrade
- Restart early 2015 at 6.5 + 6.5 TeV

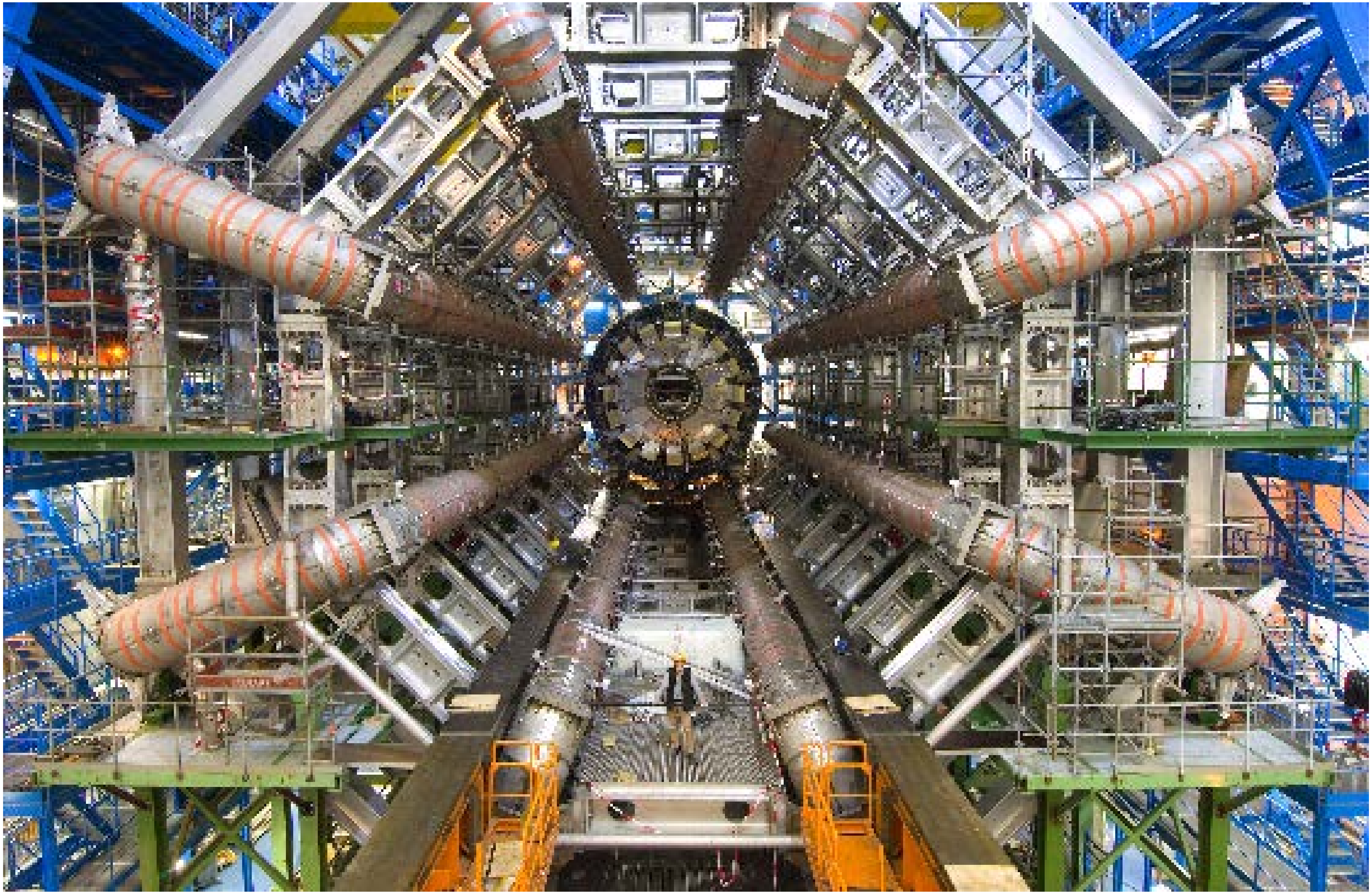
LHC – Large Hadron Collider

- Proton-proton collider
- 27 km circumference
- Started operation in 2010 with 3.5 + 3.5 TeV, continued in 2011, 4 + 4 TeV in 2012
  - World’s most powerful particle accelerator

  Run 1 until early 2013
- Now in Long Shutdown 1 – machine upgrade
- Restart early 2015 at 6.5 + 6.5 TeV

Four Large Detectors

- ATLAS, CMS, ALICE, LHCb
- Some ATLAS facts:
  - 100 million channels
  - 25 m diameter, 46 m length, 7’000 tons
  - 3’000 scientists (including 1’000 grad students)
  - 40 MHz collision rate
  - Run 1: 300 Hz event rate after filtering
- All LHC experiments: 30’000 TB in 2012, 100’000 TB in total

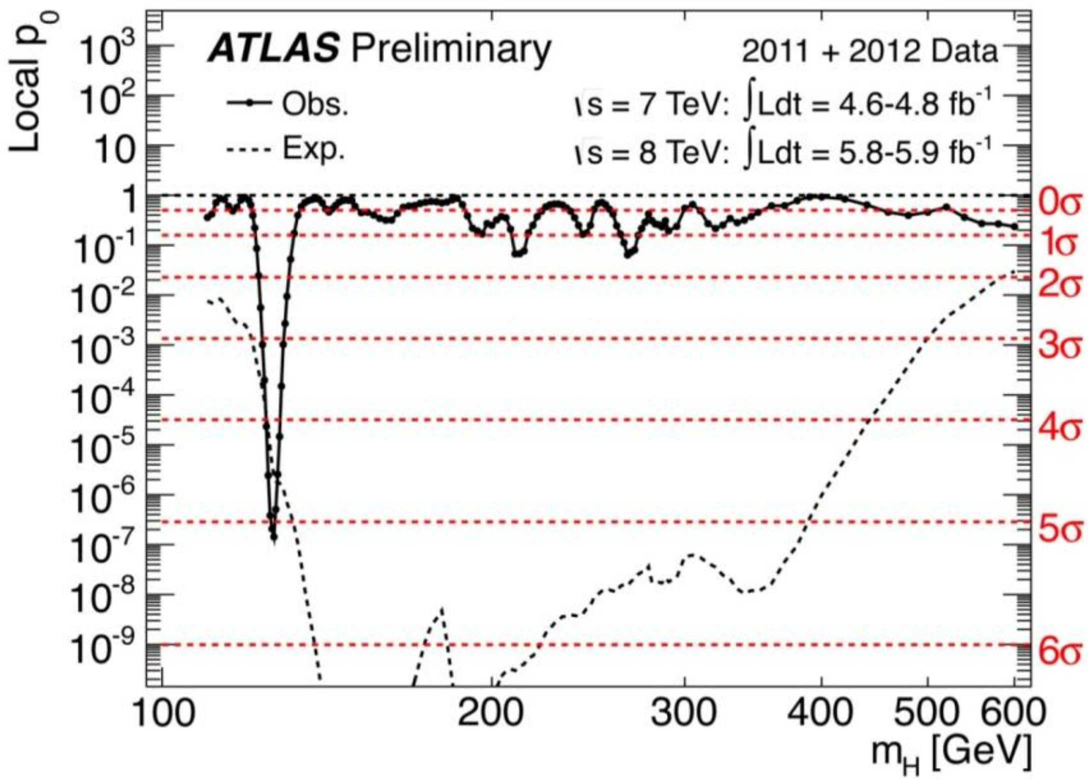


Results so far

- Many… the most spectacular one being

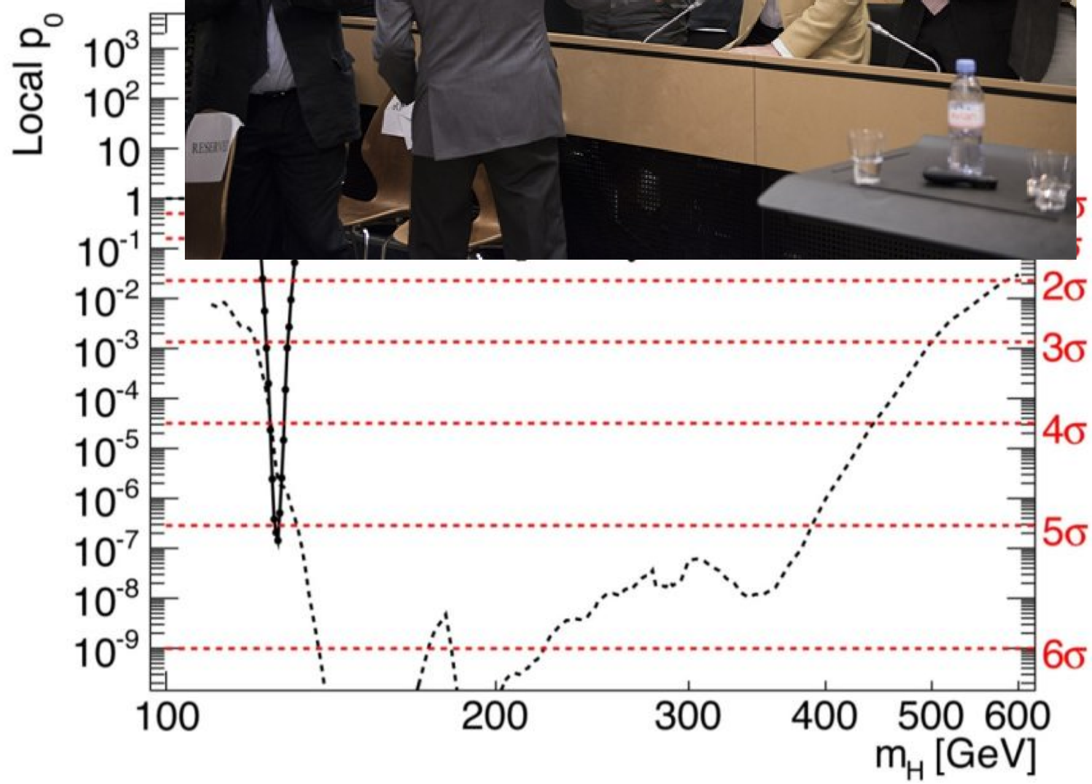
Results so far

- Many… the most spectacular one being
- 04 July 2012: Discovery of a “Higgs-like particle”



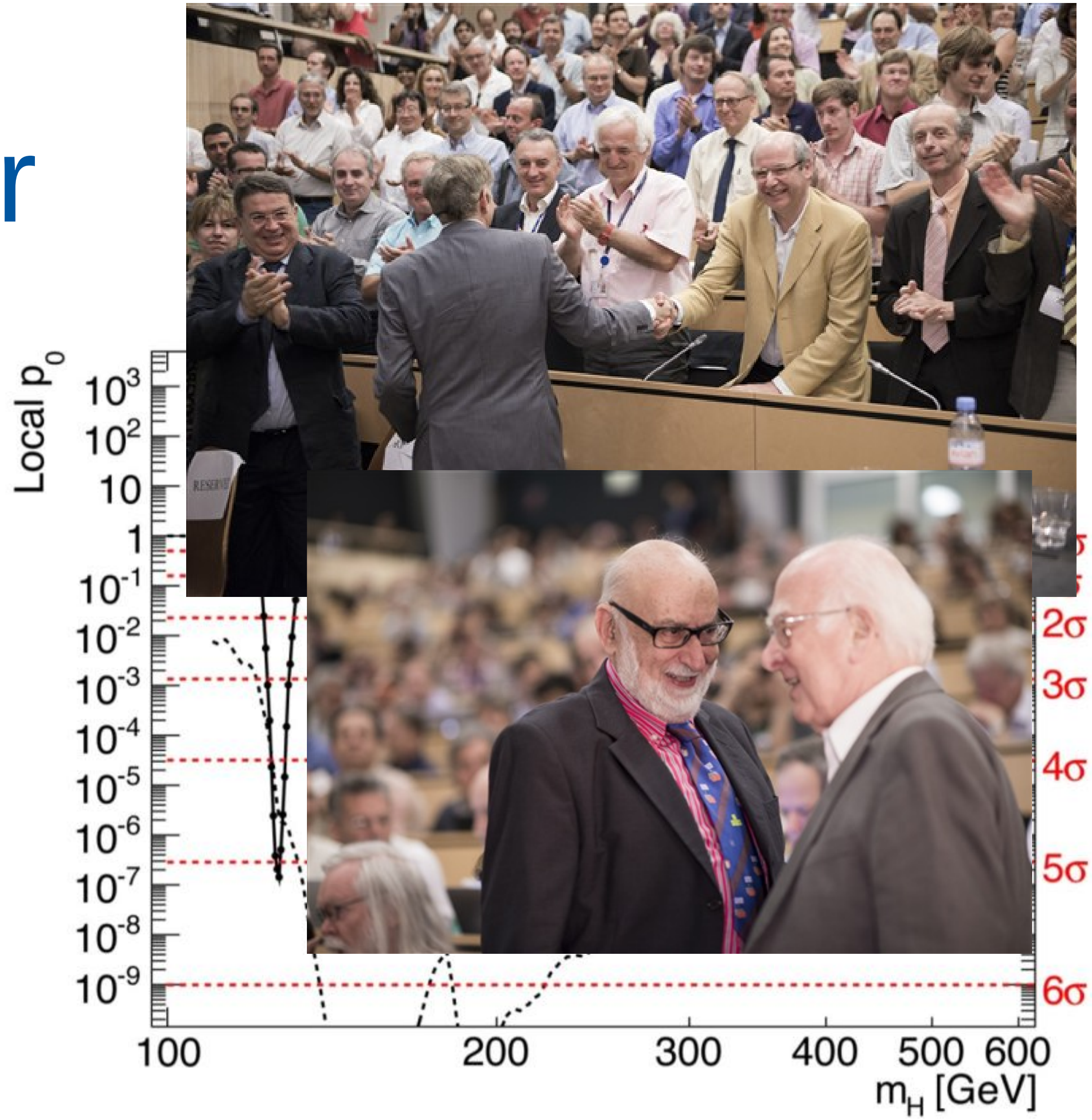
Results so far

- Many… the most spectacular one being
- 04 July 2012: Discovery of a “Higgs-like particle”



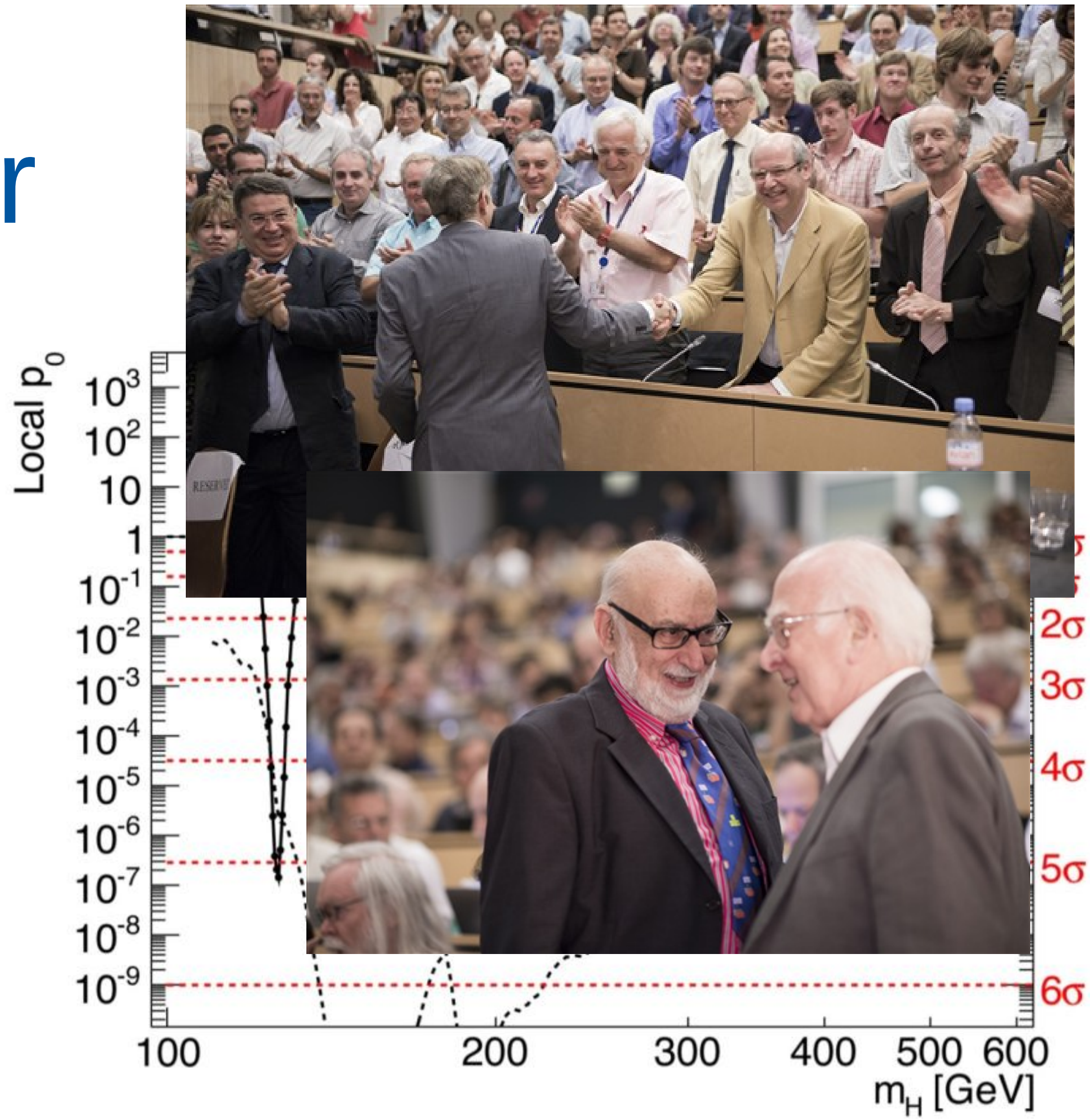
Results so far

- Many… the most spectacular one being
- 04 July 2012: Discovery of a “Higgs-like particle”



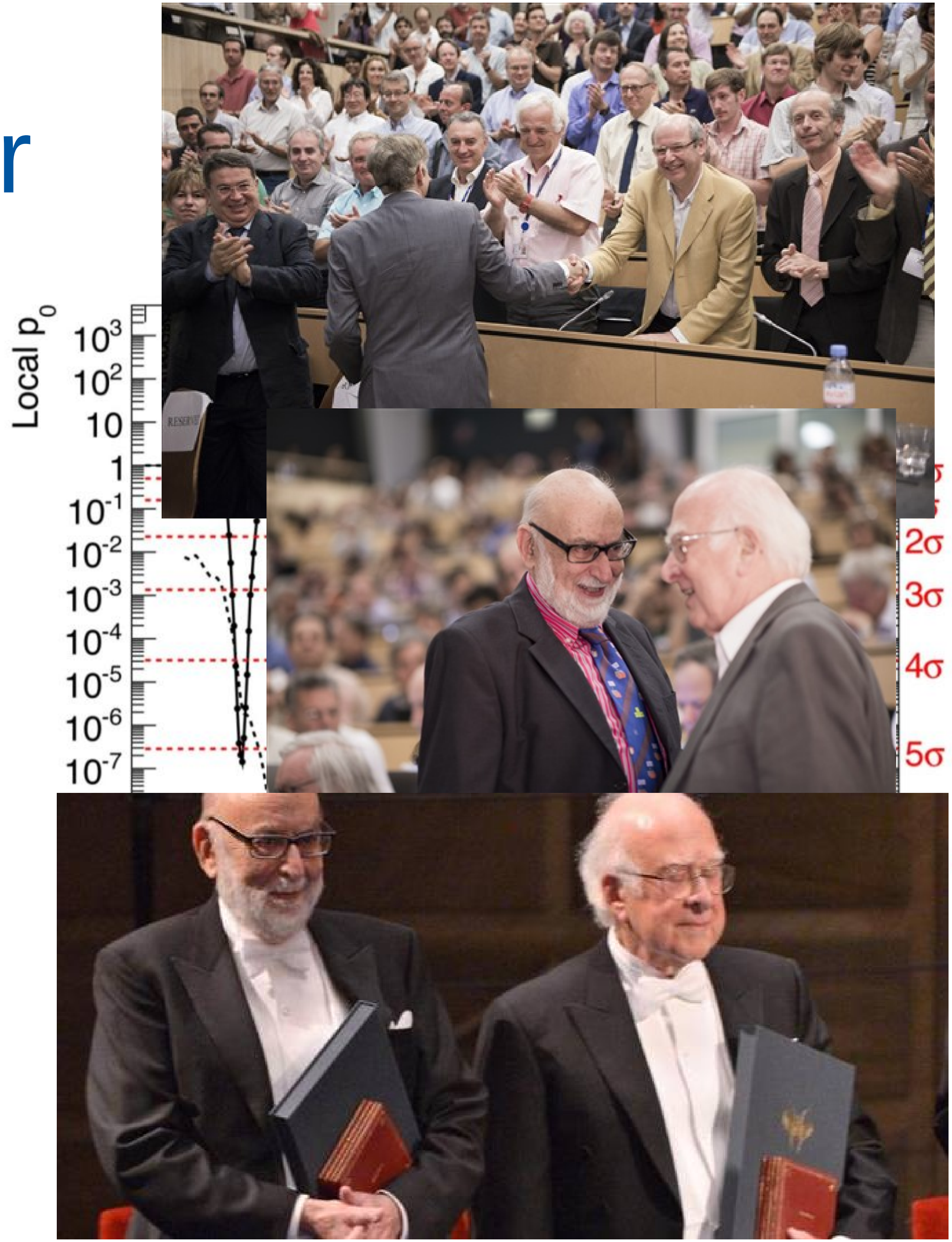
Results so far

- Many… the most spectacular one being
- 04 July 2012: Discovery of a “Higgs-like particle”
- March 2013: The particle is indeed a Higgs boson



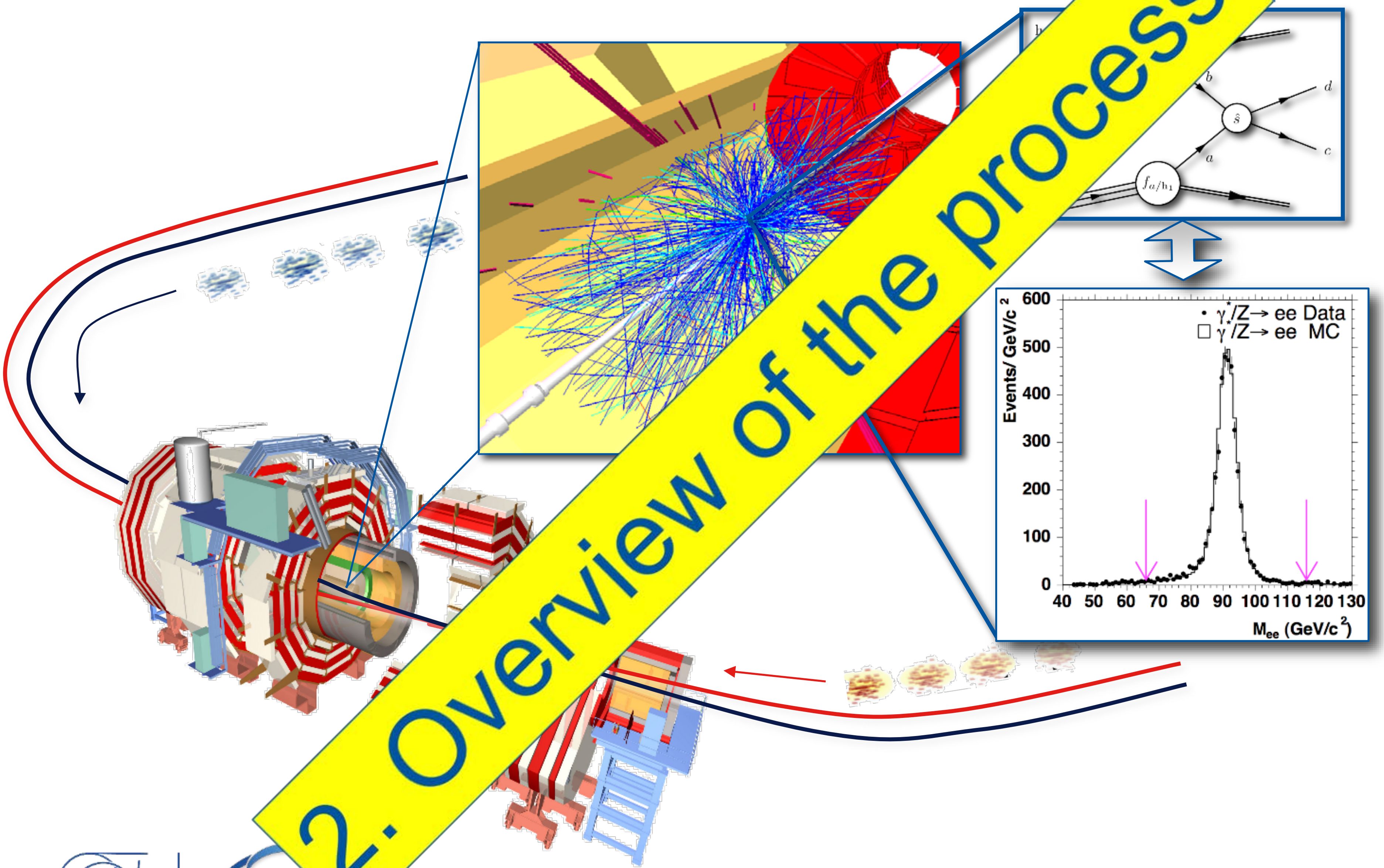
Results so far

- Many… the most spectacular one being
- 04 July 2012: Discovery of a “Higgs-like particle”
- March 2013: The particle is indeed a Higgs boson
- 08 Oct 2013 / 10 Dec 2013: Nobel price to Peter Higgs and François Englert
    - CERN, ATLAS and CMS explicitly mentioned



YEARS/ANS CERN

2. Overview of the process

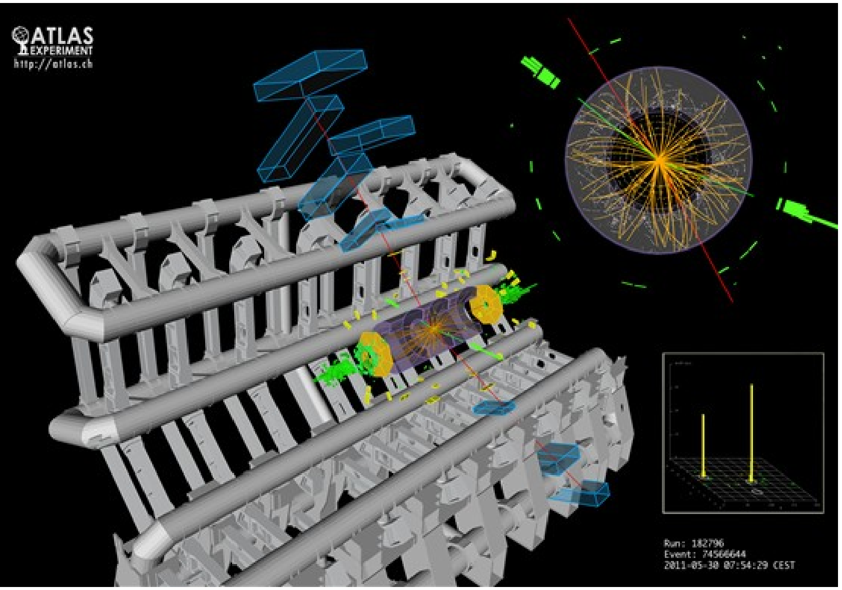


Physics Results ARE NOT…

- Finding a single ‘golden’ event (or a few of them)

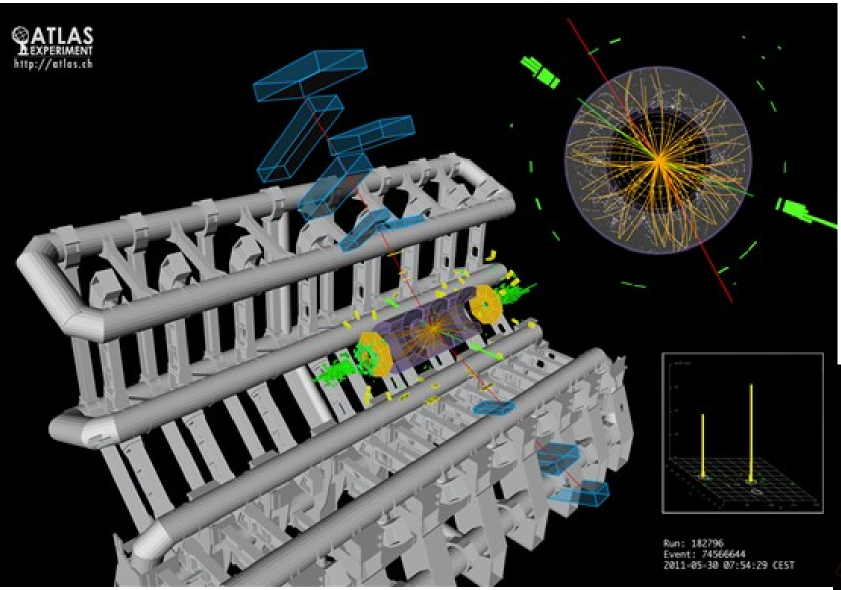
Physics Results ARE NOT…

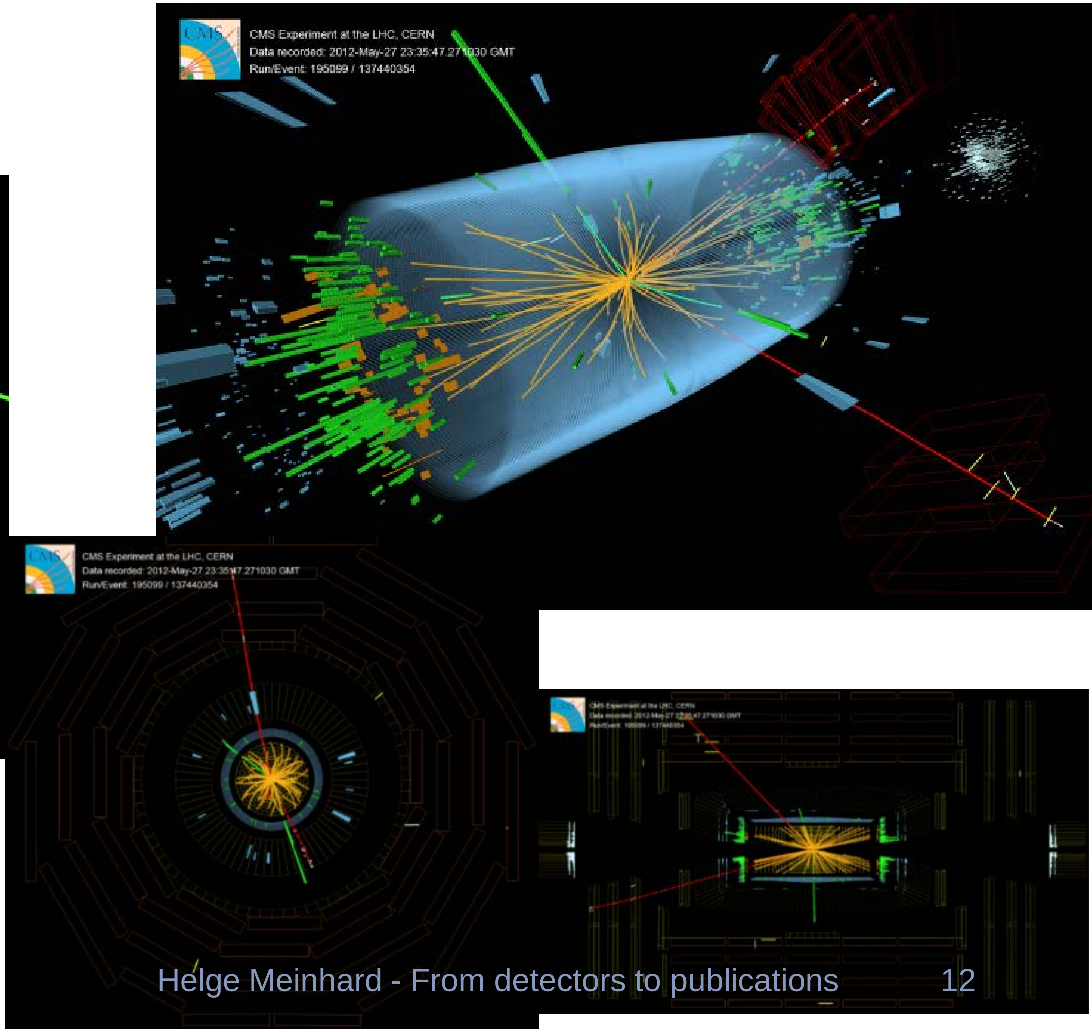
- Finding a single ‘golden’ event (or a few of them)



Physics Results ARE NOT…

- Finding a single ‘golden’ event (or a few of them)



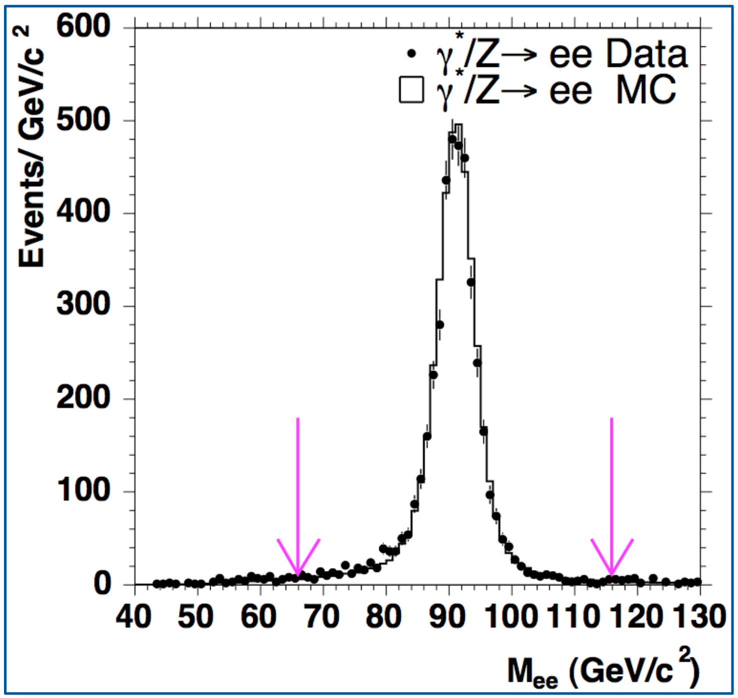


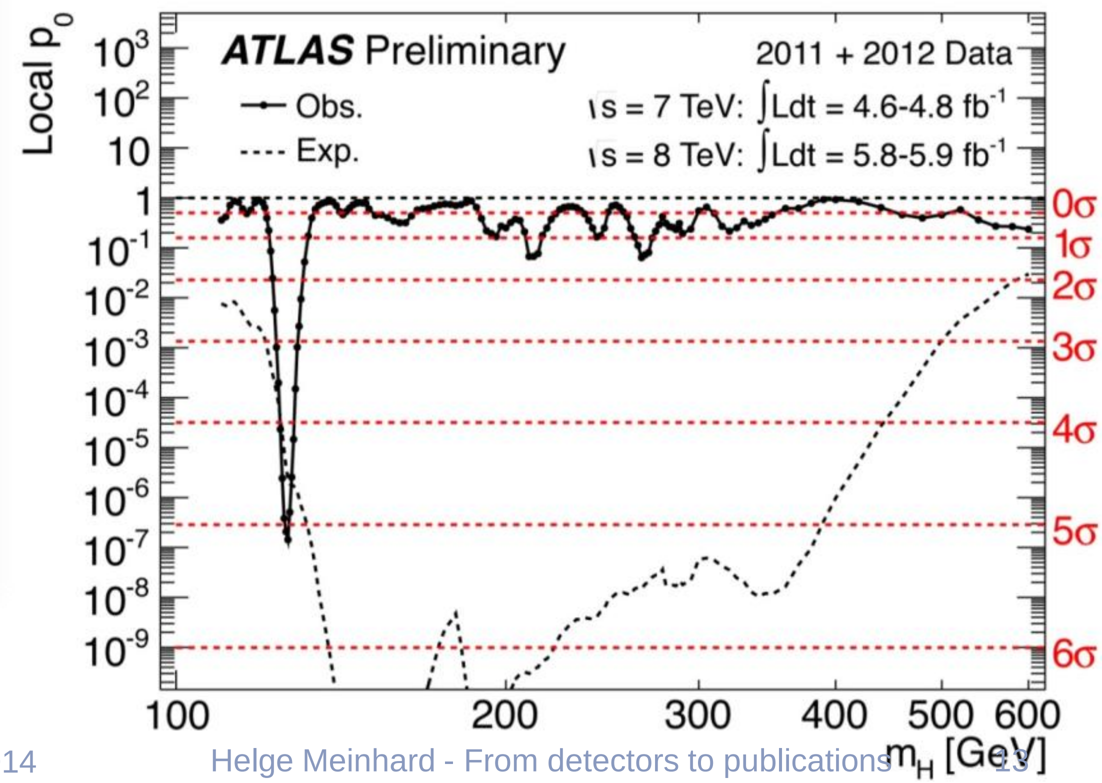
Physics Results ARE…

… statistical comparisons of experimental data with theoretical predictions

Physics Results ARE…

… statistical comparisons of experimental data with theoretical predictions





Theory...

$$\mathcal{L} = -\tfrac{1}{4}\mathbf{W}_{\mu\nu}\cdot\mathbf{W}^{\mu\nu} - \tfrac{1}{4}B_{\mu\nu}B^{\mu\nu}$$

{ $W^{\pm}, Z, \gamma$ kinetic energies and self-interactions

$$+\overline{L}\gamma^{\mu}\left(i\partial_{\mu} - g\tfrac{1}{2}\tau\cdot\mathbf{W}_{\mu} - g'\tfrac{Y}{2}B_{\mu}\right)L$$
$$+\overline{R}\gamma^{\mu}\left(i\partial_{\mu} - g'\tfrac{Y}{2}B_{\mu}\right)R$$

{ lepton and quark kinetic energies and their interactions with $W^{\pm}, Z, \gamma$

$$+\left|\left(i\partial_{\mu} - g\tfrac{1}{2}\tau\cdot\mathbf{W}_{\mu} - g'\tfrac{Y}{2}B_{\mu}\right)\phi\right|^{2}$$
$$-V(\phi)$$

{ $W^{\pm}, Z, \gamma$ and Higgs masses and couplings

$$-(G_1\overline{L}\phi R + G_2\overline{L}\phi_c R + h.c.)$$

{ lepton and quark masses and coupling to Higgs

L ... left–handed fermion (l or q) doublet
R ... right–handed fermion singlet

$\mathcal{L}$ from QCD:

$$\mathcal{L} = \bar{q}\underbrace{(i\gamma^{\mu}\partial_{\mu} - m)}_{E_{kin}(q)}q - g\underbrace{(\bar{q}\gamma^{\mu}T_a q)G^a_{\mu}}_{\text{Interaction } q,g} - \underbrace{\frac{1}{4}G^a_{\mu\nu}G^{\mu\nu}_a}_{E_{kin}(g)}$$

$E_{kin}(q)$

Interaction q, g

$E_{kin}(g)$ includes self-interaction between gluons

eg.

the Standard Model

Theory...

$$\mathcal{L} = -\tfrac{1}{4}\mathbf{W}_{\mu\nu}\cdot\mathbf{W}^{\mu\nu} - \tfrac{1}{4}B_{\mu\nu}B^{\mu\nu}$$

$W^{\pm}, Z, \gamma$ kinetic energies and self-interactions

$$+\overline{L}\gamma^{\mu}\left(i\partial_{\mu} - g\tfrac{1}{2}\tau\cdot\mathbf{W}_{\mu} - g'\tfrac{Y}{2}B_{\mu}\right)L$$
$$+\overline{R}\gamma^{\mu}\left(i\partial_{\mu} - g'\tfrac{Y}{2}B_{\mu}\right)R$$

lepton and quark kinetic energies and their interactions with $W^{\pm}, Z, \gamma$

$$+\left|\left(i\partial_{\mu} - g\tfrac{1}{2}\tau\cdot\mathbf{W}_{\mu} - g'\tfrac{Y}{2}B_{\mu}\right)\phi\right|^{2}$$
$$-V(\phi)$$

$W^{\pm}, Z, \gamma$ and Higgs masses and couplings

$$-(G_1\overline{L}\phi R + G_2\overline{L}\phi_c R + h.c.)$$

lepton and quark masses and coupling to Higgs

L ... left–handed fermion (l or q) doublet
R ... right–handed fermion singlet

$\mathcal{L}$ from QCD:

$$\mathcal{L} = \bar{q}\underbrace{(i\gamma^{\mu}\partial_{\mu} - m)}_{E_{kin}(q)}q - g\underbrace{(\bar{q}\gamma^{\mu}T_a q)G^a_{\mu}}_{\text{Interaction } q,g} - \underbrace{\frac{1}{4}G^a_{\mu\nu}G^{\mu\nu}_a}_{E_{kin}(g)}$$

$E_{kin}(q)$

Interaction q, g

$E_{kin}(g)$ includes self-interaction between gluons

eg.
the Standard Model

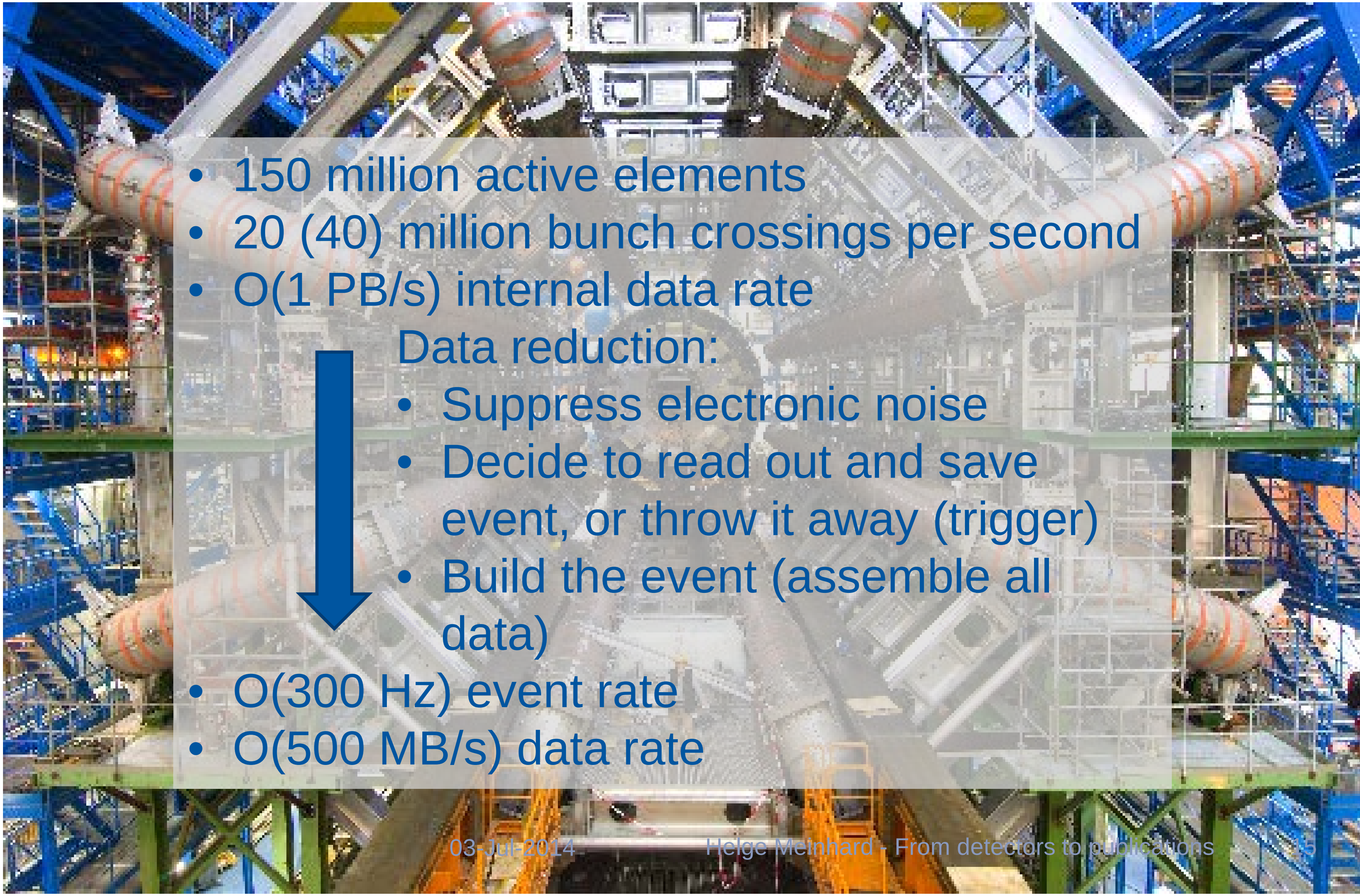
has parameters

coupling constants

masses

predicts:
cross sections,
branching ratios, lifetimes,
...

Experiment…



- 150 million active elements
- 20 (40) million bunch crossings per second
- O(1 PB/s) internal data rate

Data reduction:

- Suppress electronic noise
- Decide to read out and save event, or throw it away (trigger)
- Build the event (assemble all data)

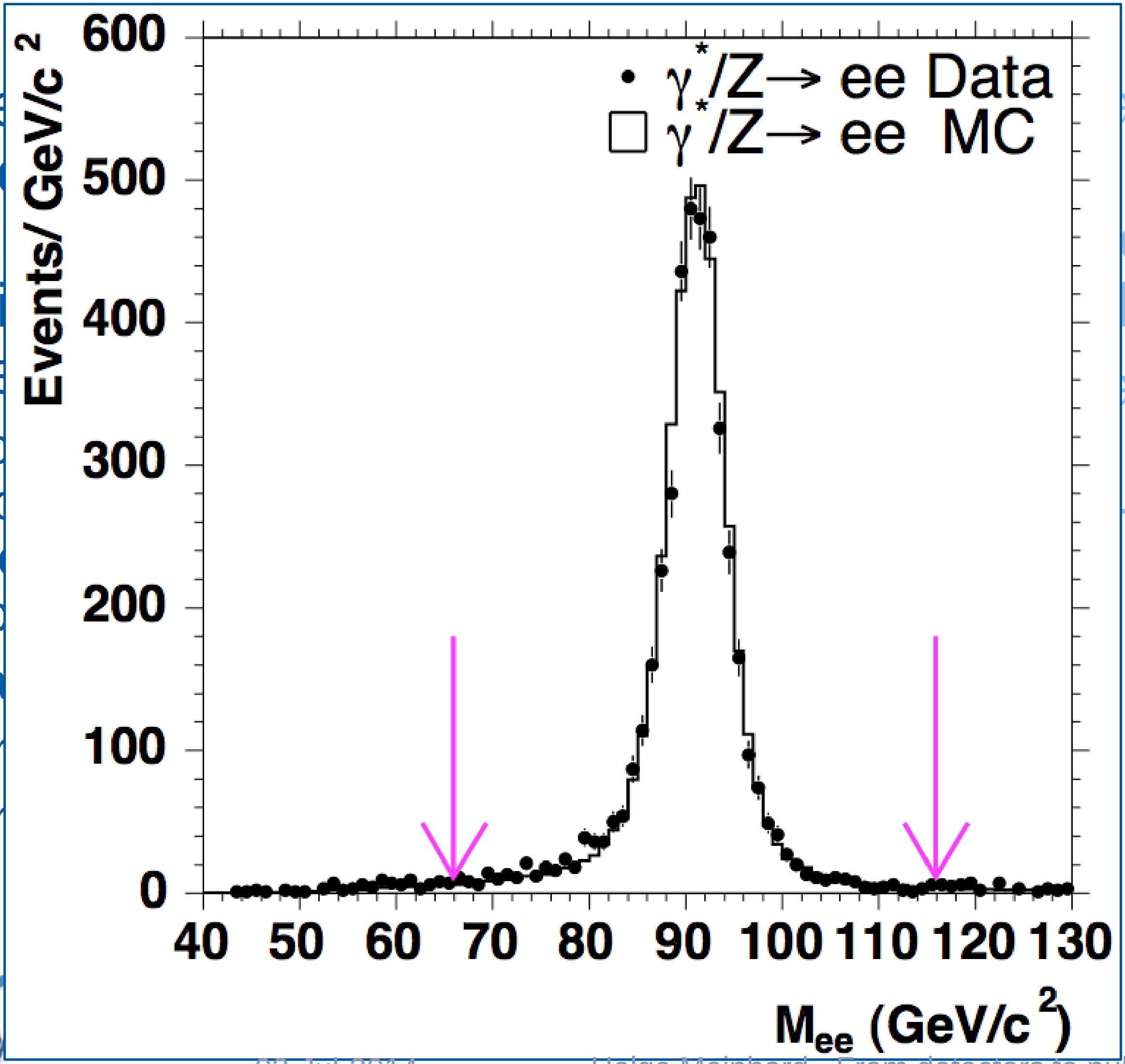
- O(300 Hz) event rate
- O(500 MB/s) data rate

Compare Theory with Experiment

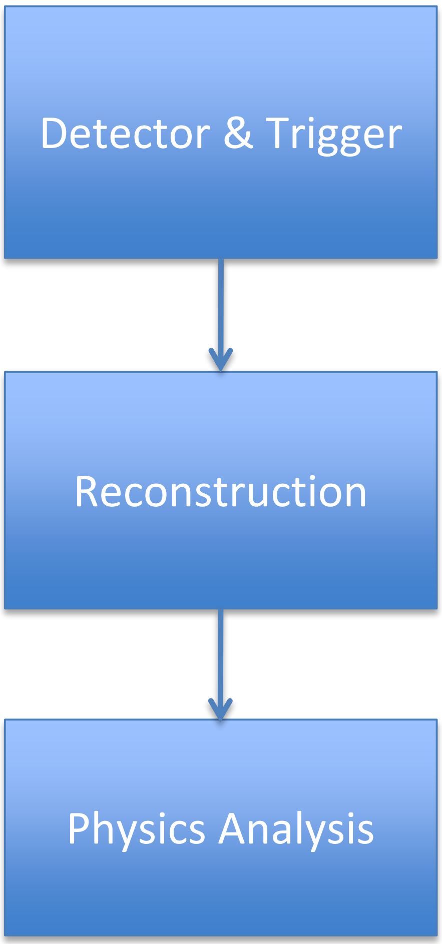
- Define "observables": Physics quantities that are sensitive to variations in theoretical predictions
  - Cross sections (probability of a certain type of interaction to occur)
  - Branching ratios
  - Particle masses
  - Particle lifetimes
  - …
- Calculate statistically these observables over a large number of events recorded
- Calculate statistically these observables as you expect them from theory
- Compare the two
  - Fit parameter(s)
  - Check fit quality

Compare Theory with Experiment

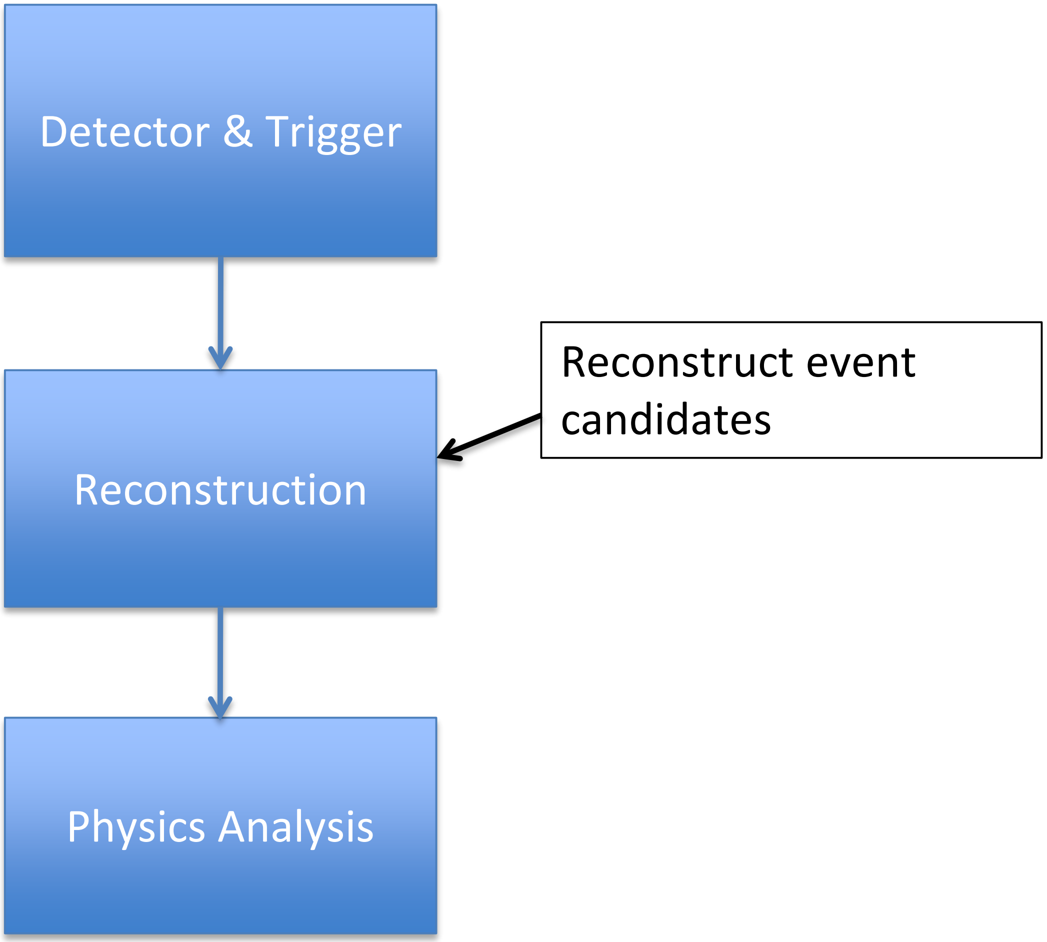
- Define ... ally Physic... that a... er variati... l theore... ally
  - Cro... as (pro... rom typ... occ...
  - Bra...
  - Par...
  - Par...
  - …



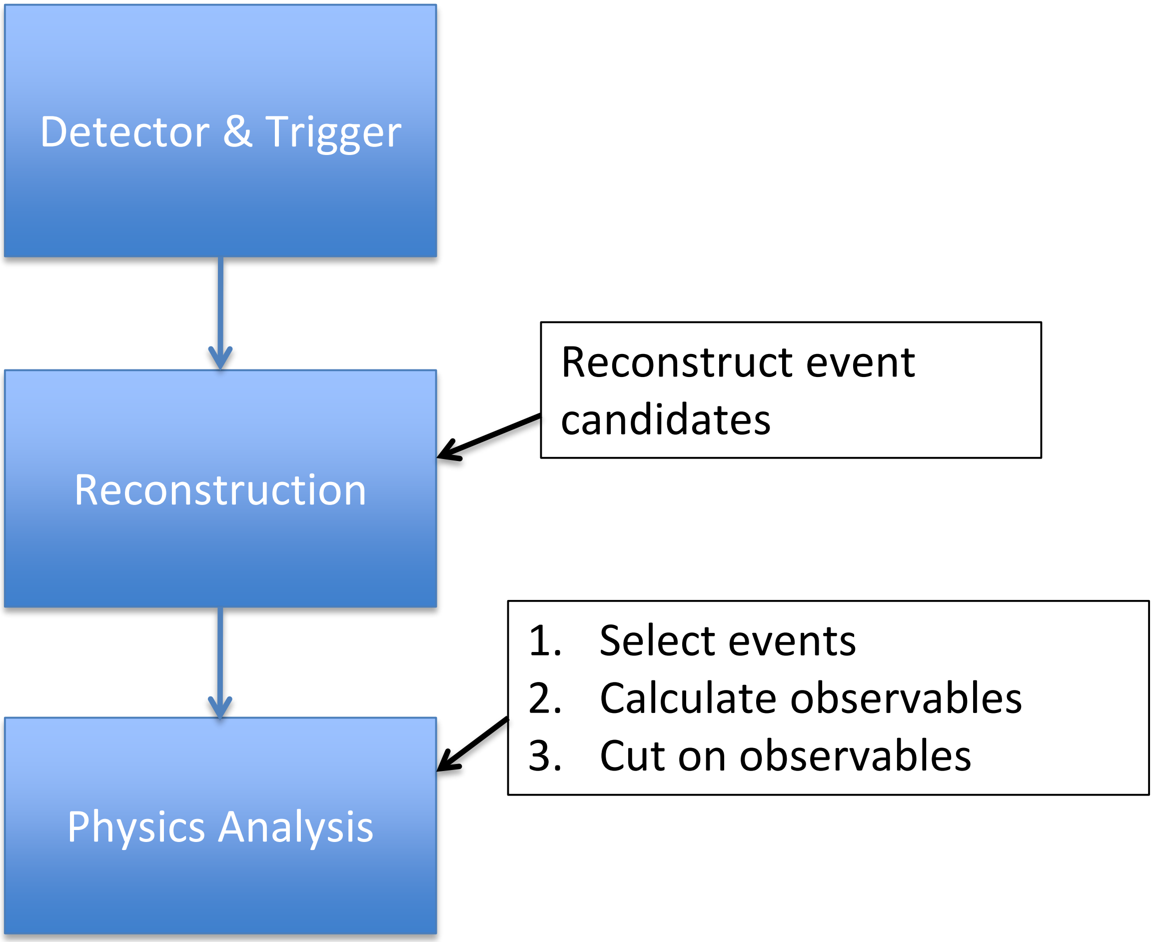
Analysis Flow



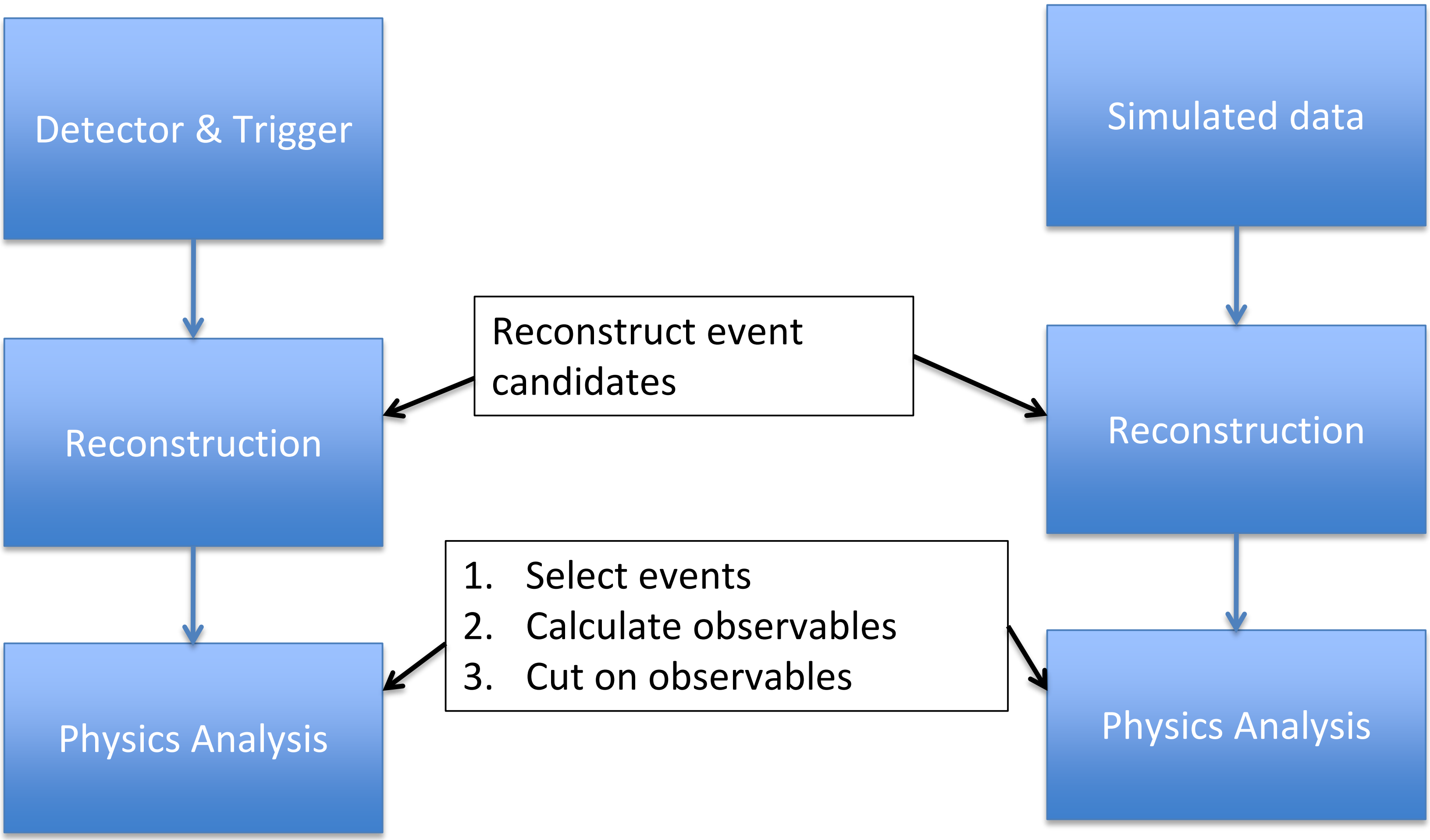
Analysis Flow



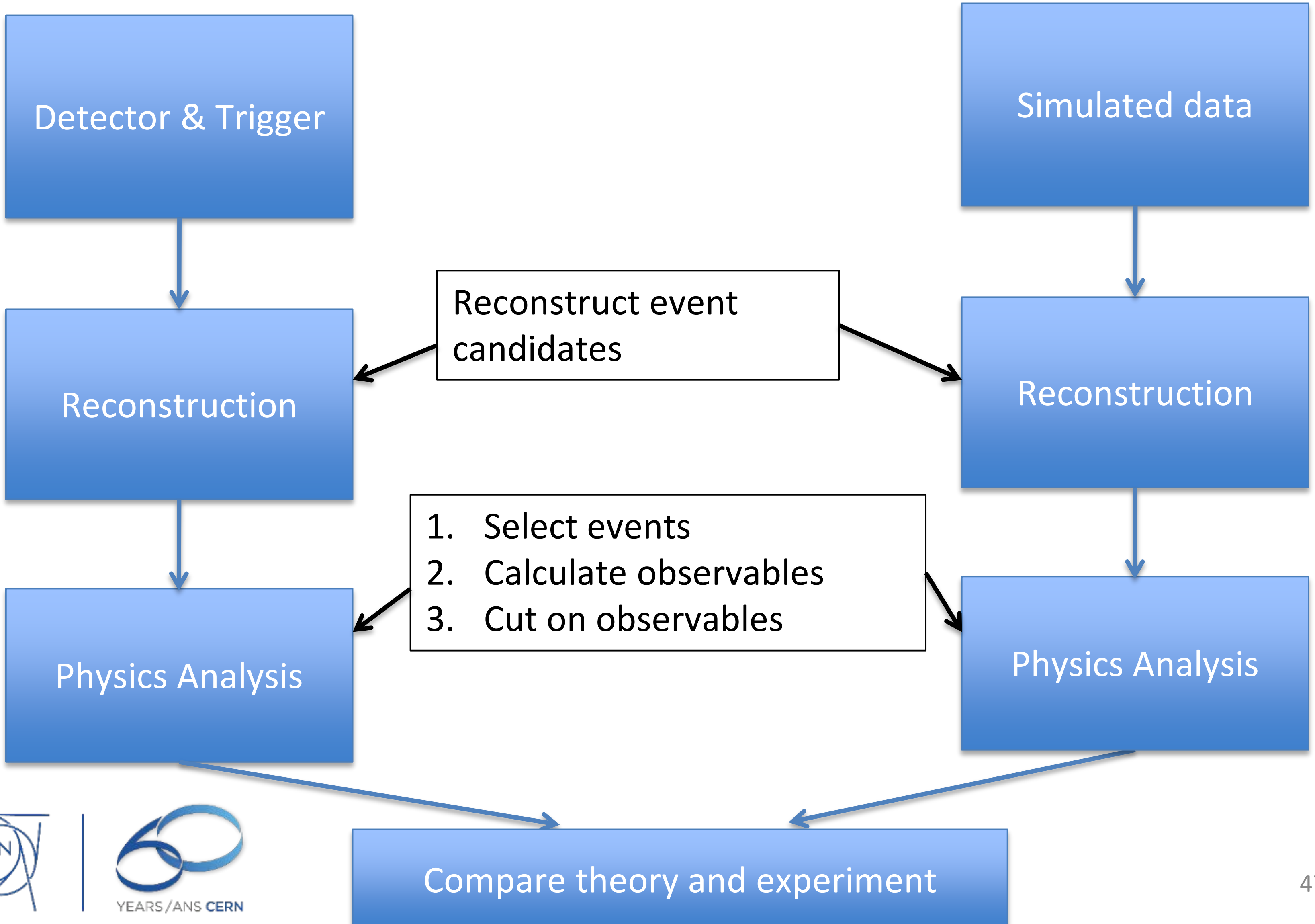
Analysis Flow



Analysis Flow



Analysis Flow



Analysis Flow

Detector & Trigger

Simulated data

Reconstruct event candidates

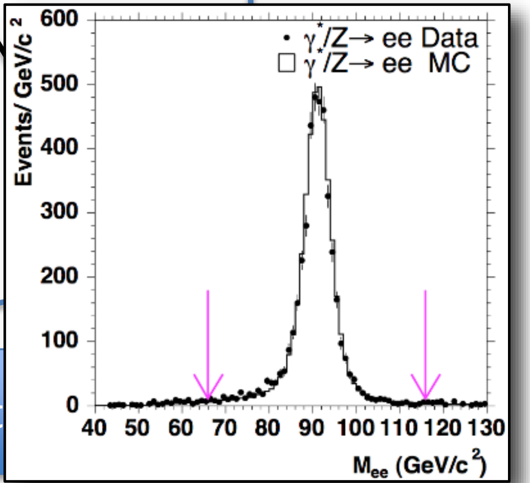
Reconstruction

Reconstruction

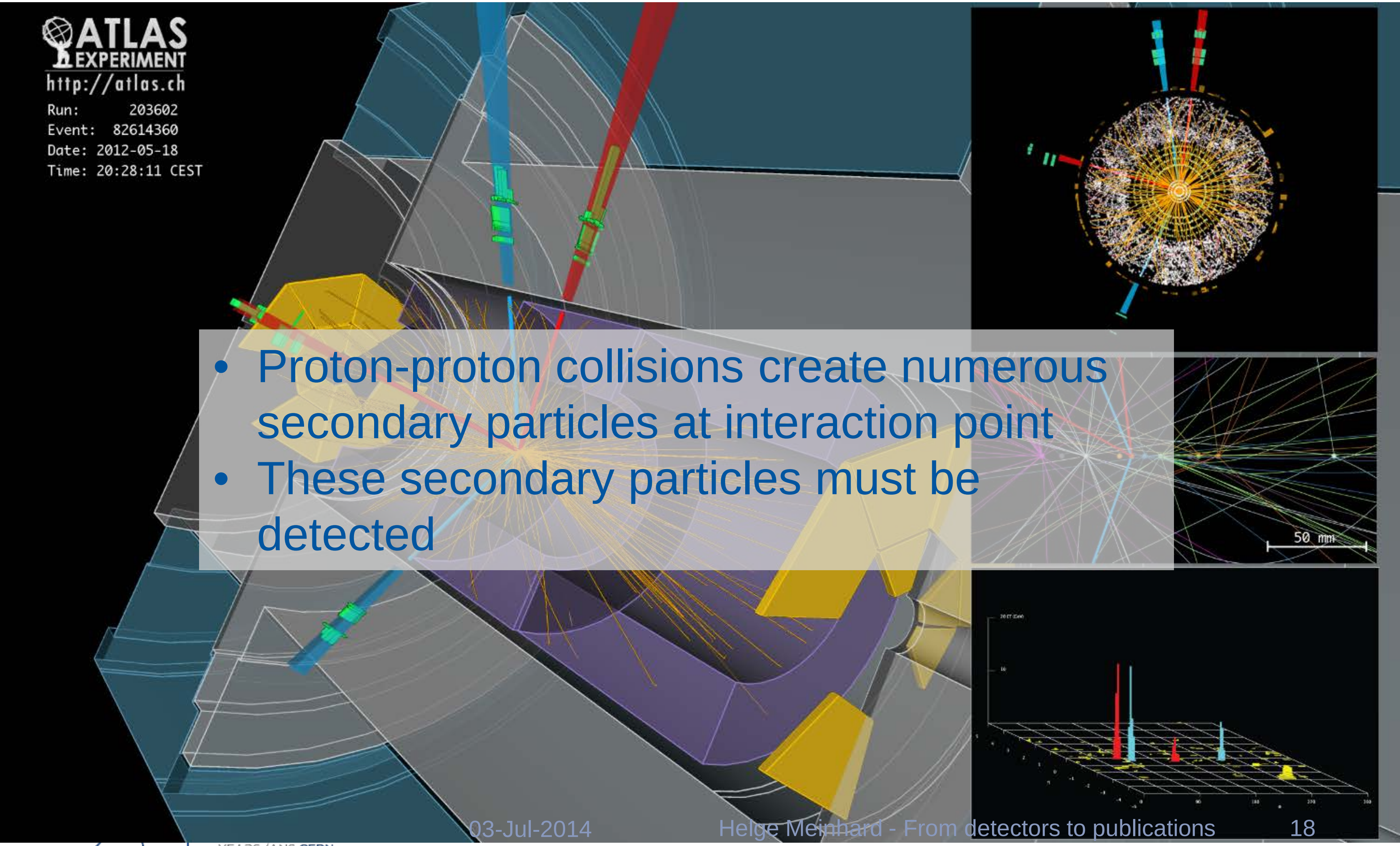
1. Select events
2. Calculate observables
3. Cut on observables

Physics Analysis

Compare theory and experiment

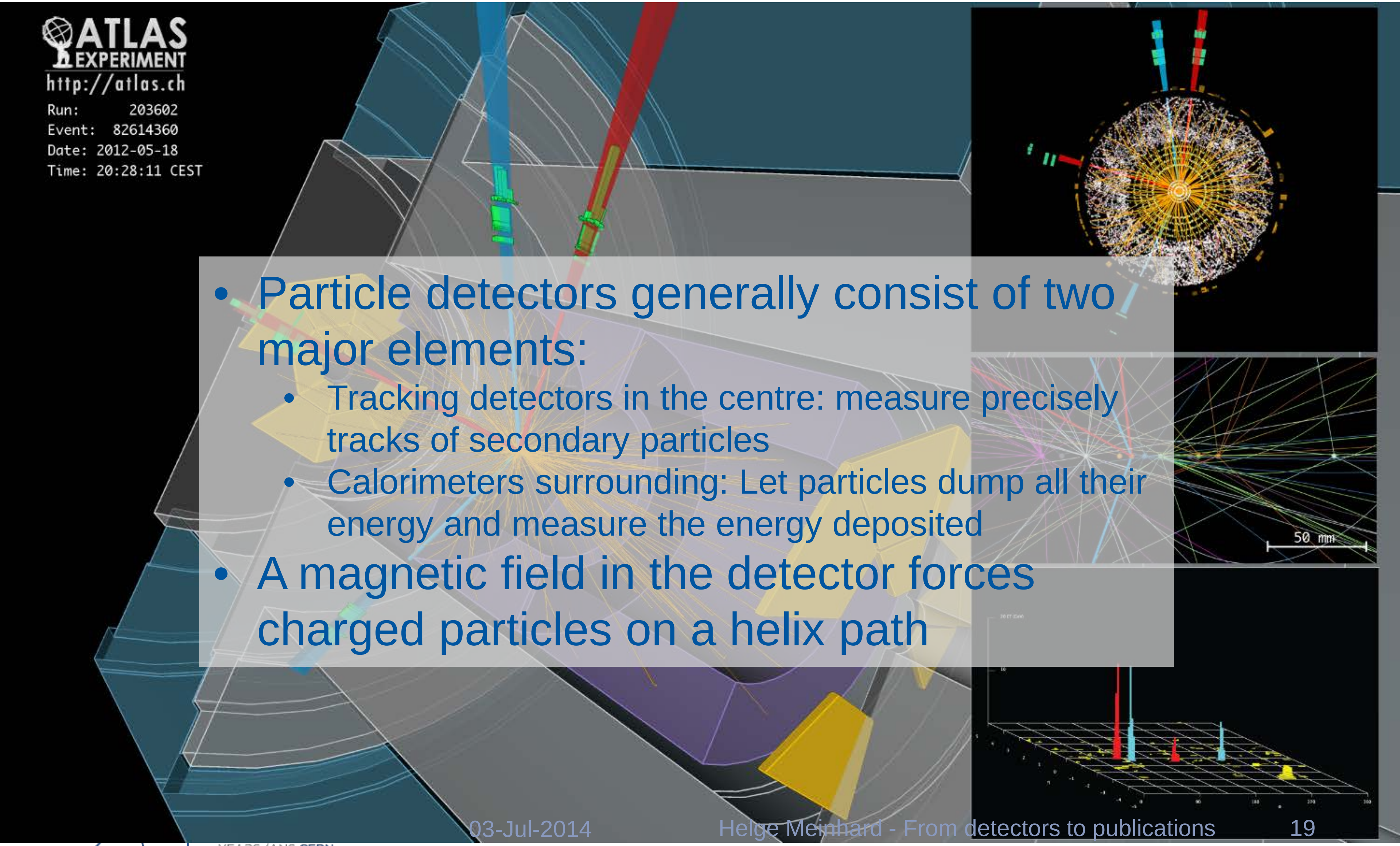


Detector (and Trigger) (1)



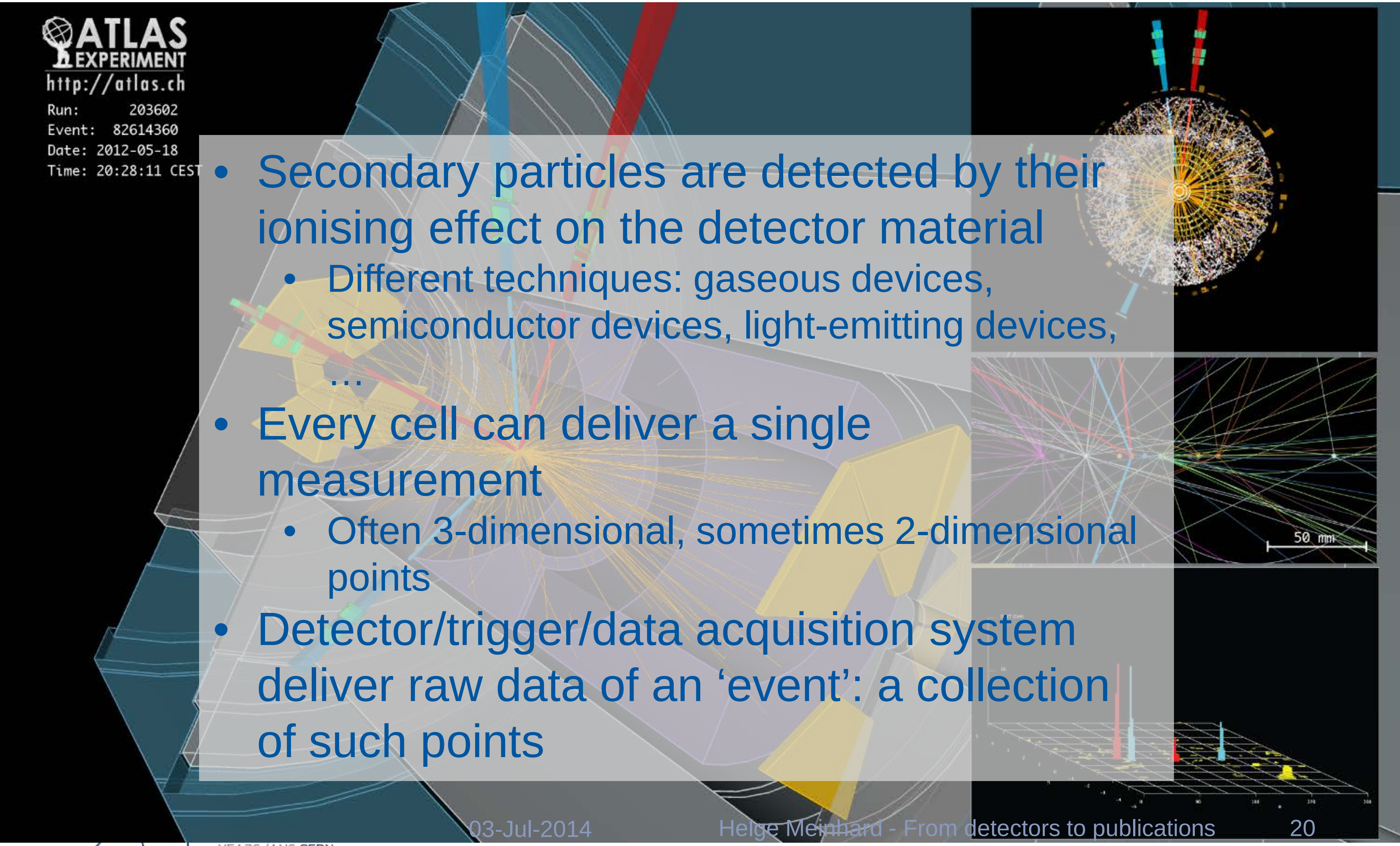
- Proton-proton collisions create numerous secondary particles at interaction point
- These secondary particles must be detected

Detector (and Trigger) (2)



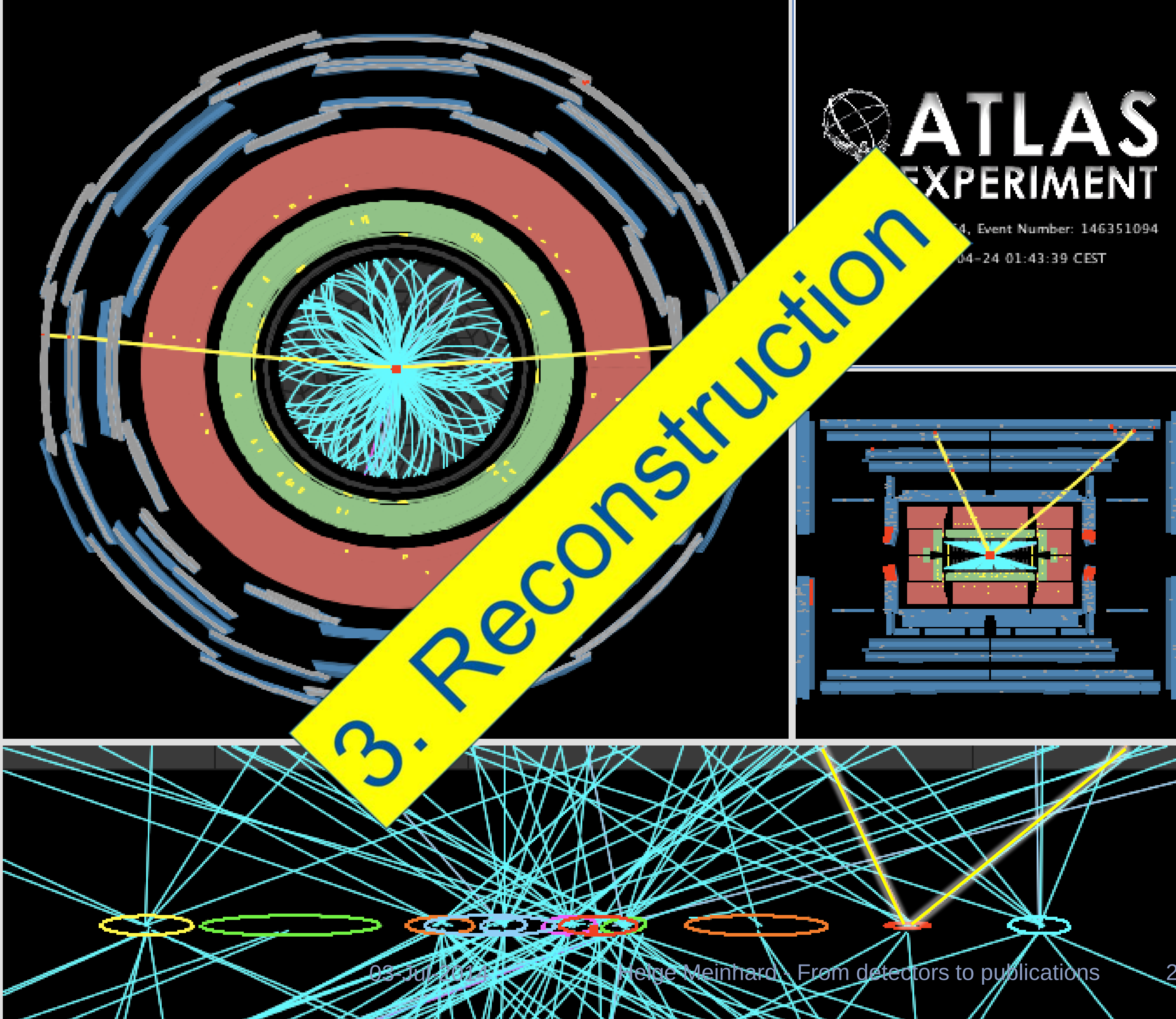
- Particle detectors generally consist of two major elements:
  - Tracking detectors in the centre: measure precisely tracks of secondary particles
  - Calorimeters surrounding: Let particles dump all their energy and measure the energy deposited
- A magnetic field in the detector forces charged particles on a helix path

Detector (and Trigger) (3)



- Secondary particles are detected by their ionising effect on the detector material
  - Different techniques: gaseous devices, semiconductor devices, light-emitting devices, ...
- Every cell can deliver a single measurement
  - Often 3-dimensional, sometimes 2-dimensional points
- Detector/trigger/data acquisition system deliver raw data of an ‘event’: a collection of such points

3. Reconstruction



Reconstruction

From points to identified particles:

- Detector reconstruction
  - Tracking
    - finding path of charged particles through the detector
    - determining momentum, charge, point of closest approach to interaction point
  - Calorimeter reconstruction
    - finding energy deposits in calorimeters from charged and neutral particles
    - determining energy deposit, location, and direction
- Combined reconstruction
  - Electron/photon identification
  - Muon identification
  - Jet finding

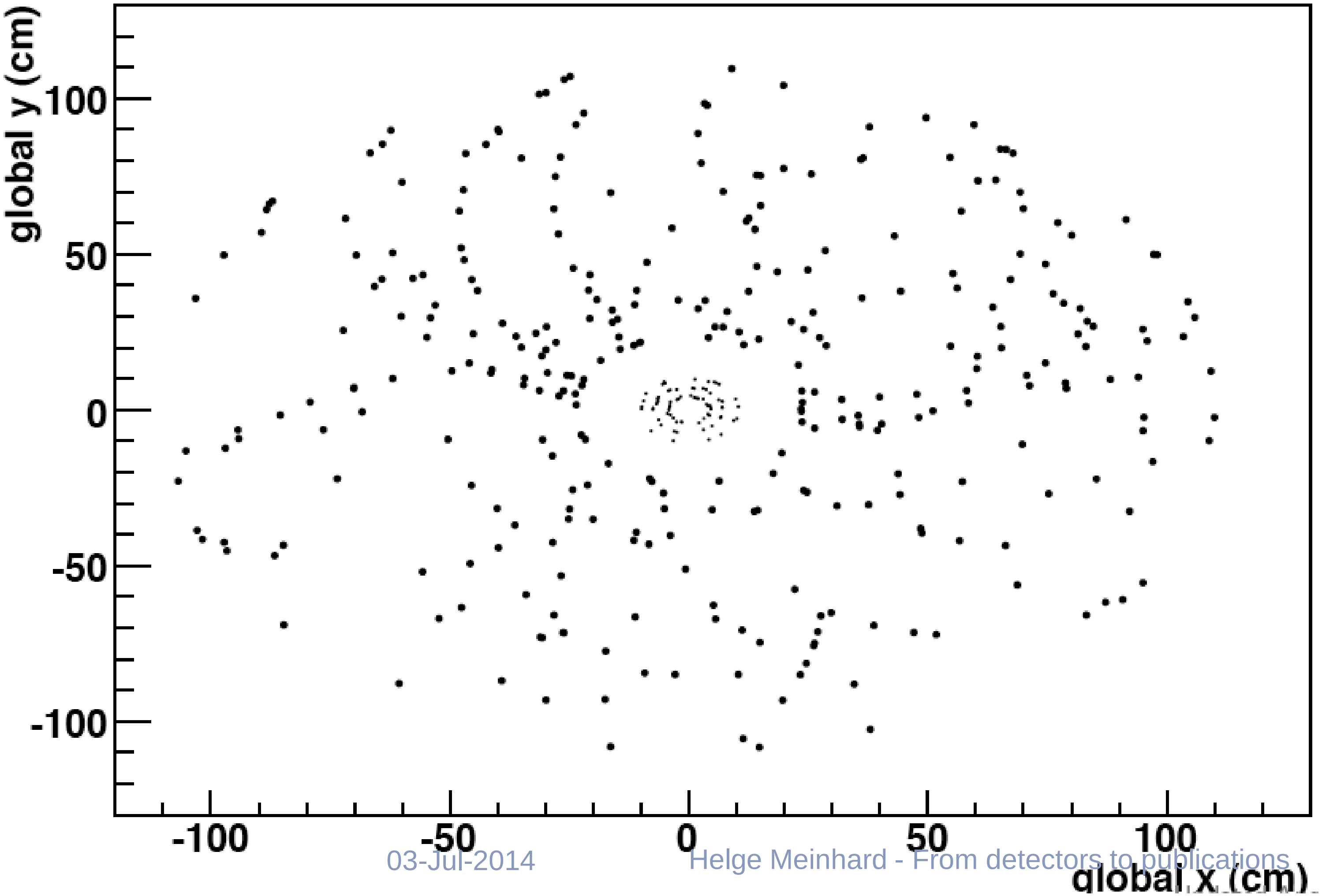
Reconstruction: Figures of Merit

- Efficiency
  - How often do we reconstruct the object correctly?
- Resolution
  - How accurately do we reconstruct it (momentum / energy, vertex position etc.)?
- Fake rate
  - How often do we reconstruct an object different from the real one?

- These figures depend on
  - Detector
  - Reconstruction algorithms, calibration, alignment
- We want/need
  - High efficiency, good resolution, low fake rate
  - … to know the efficiency, resolution, fake rate
  - Robustness against detector problems
  - Fit into computing resource limitations (CPU, memory)

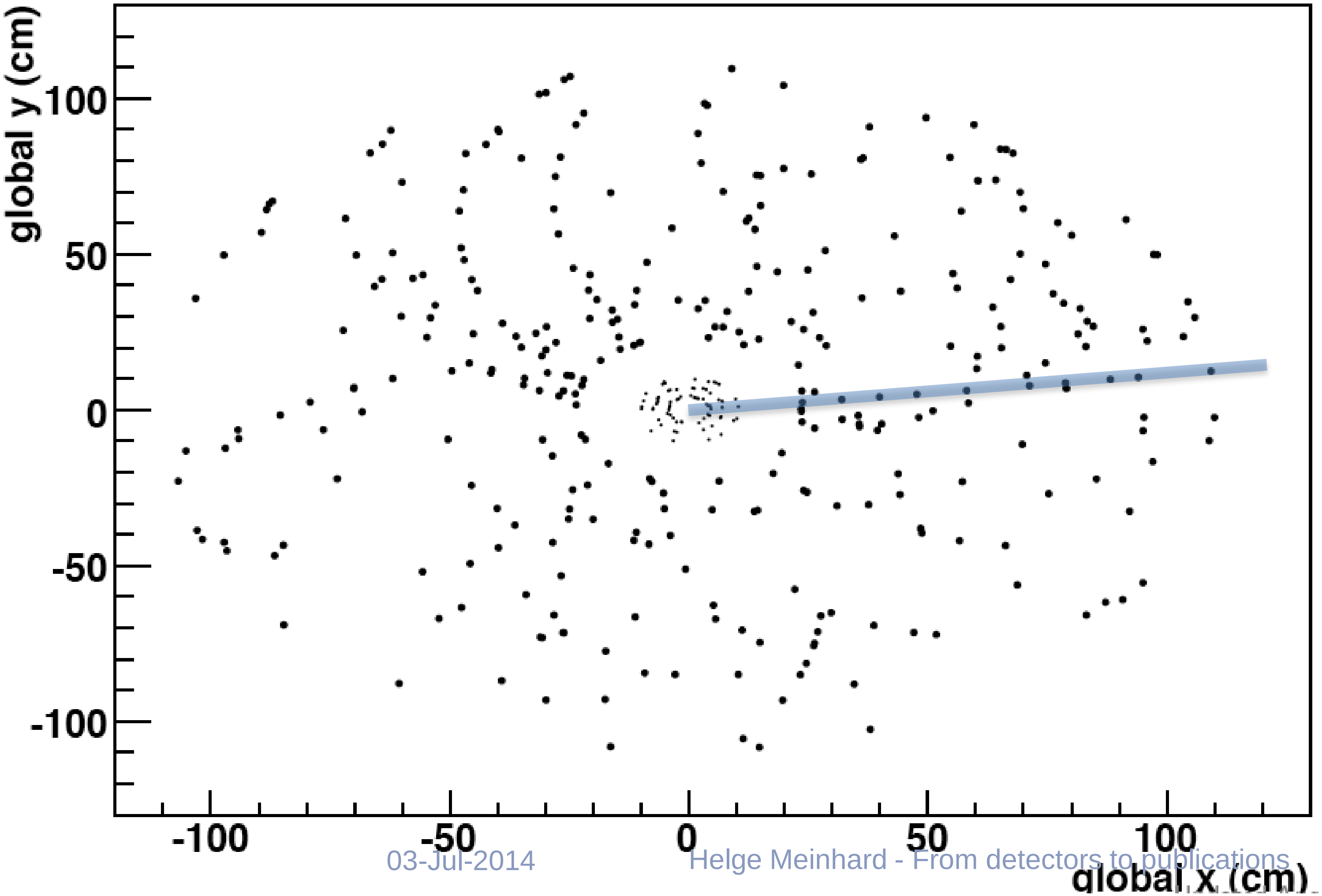
Tracking: An Example

- Where is the 50 GeV track?



Tracking: An Example

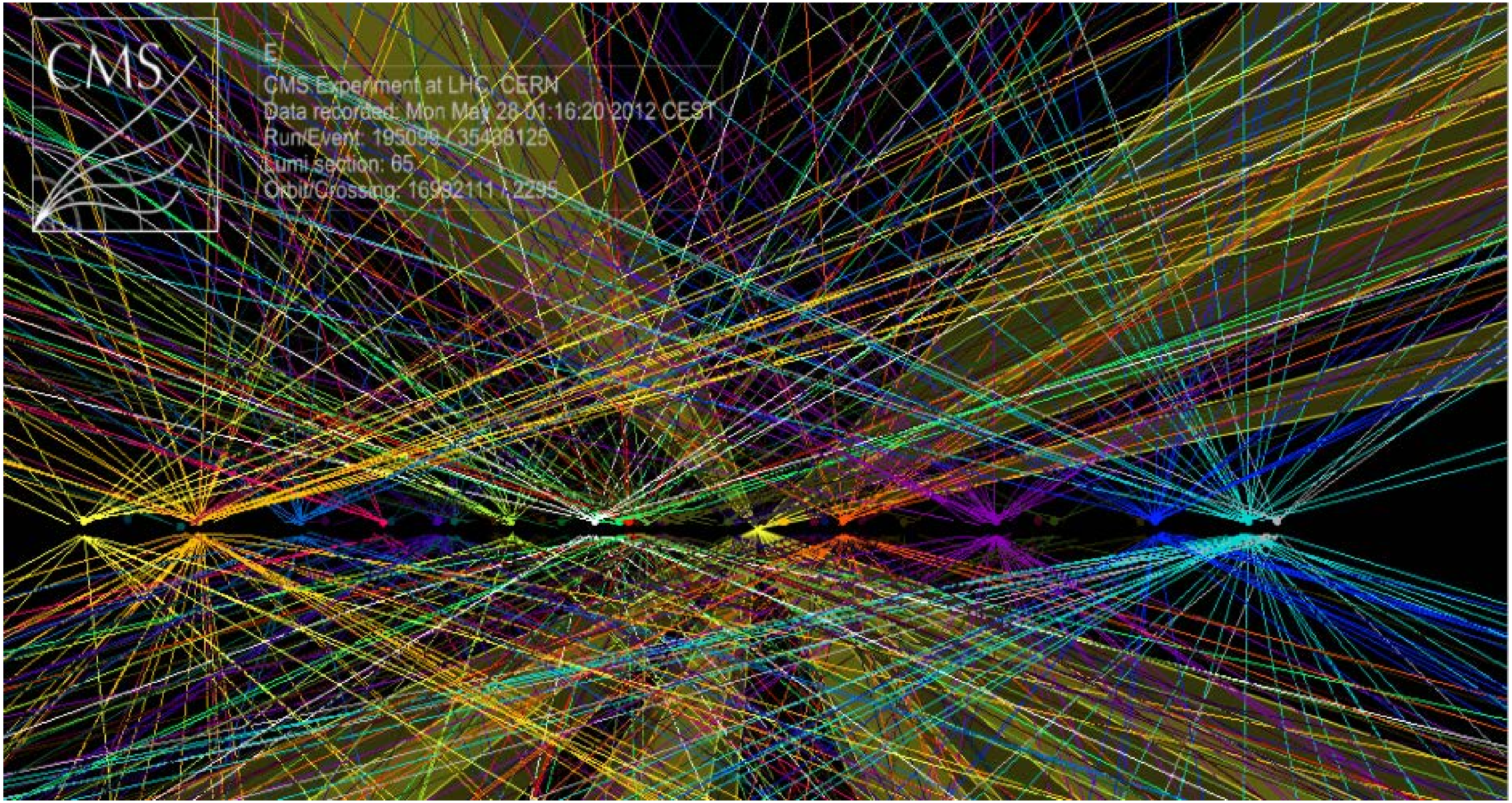
- Where is the 50 GeV track?

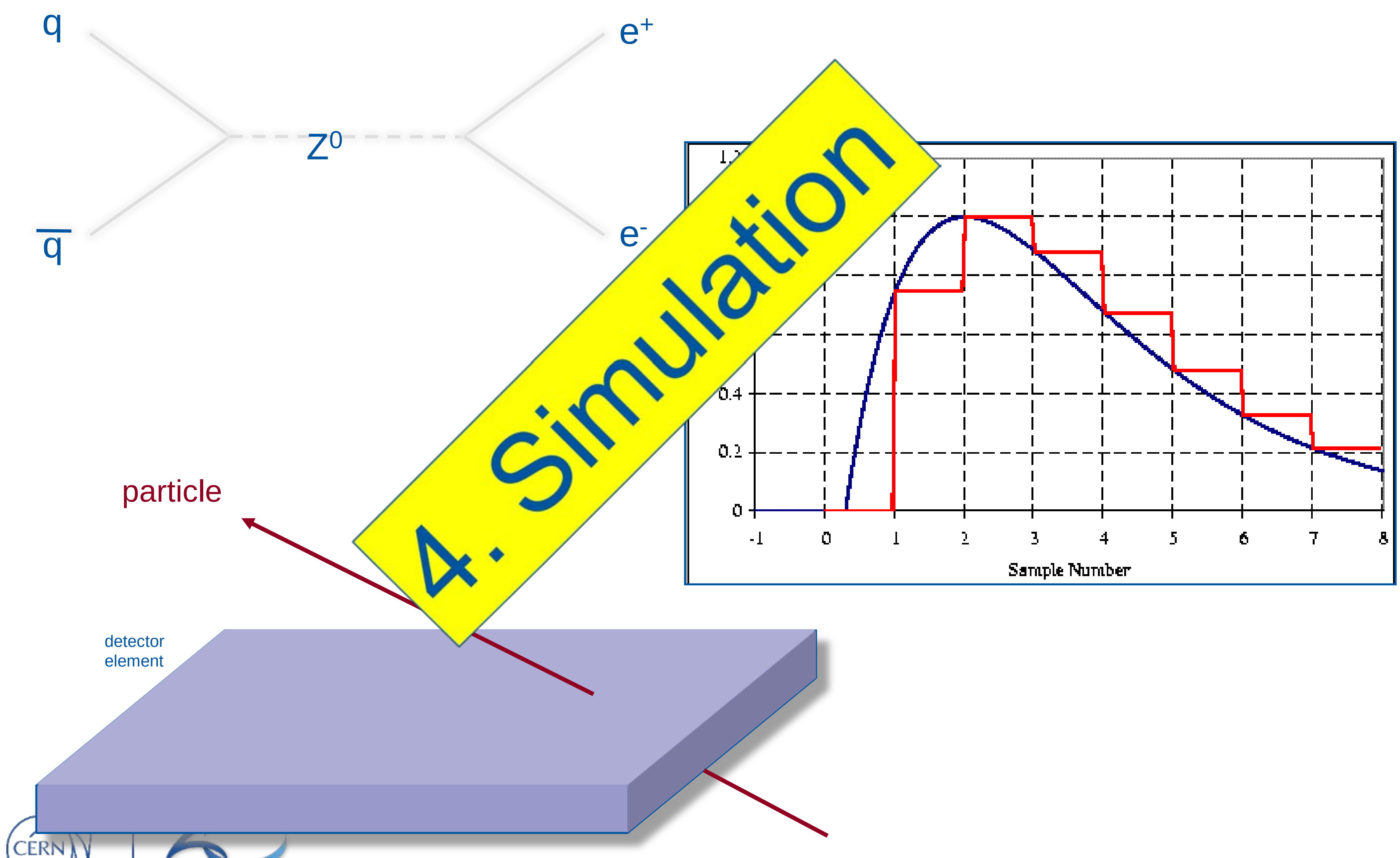


Tracking Issues

- Ionisation is a statistical process
- Ionisation means energy loss: momentum and direction of original particle get (slightly) changed
- Detector elements not perfectly aligned
- (LHC) Pile-up: a single collision of two proton bunches results in more than one proton-proton collision
  - Up to around 30 in Run 1, expected > 50 in Run 2

Pile-up Example





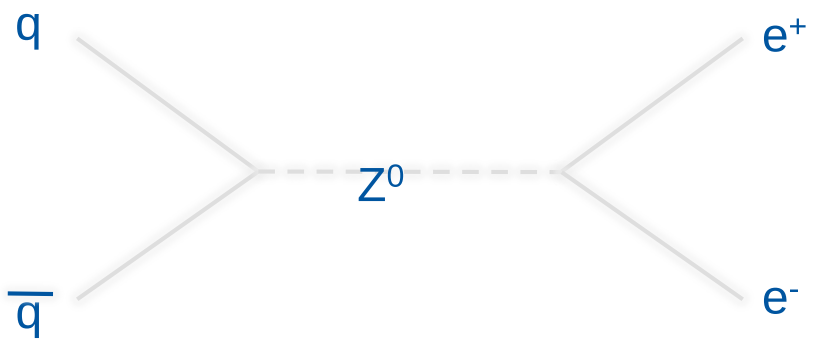
q
$\bar{q}$
Z^0
e^+
e^-
4. Simulation
particle
detector
element
0.4
0.2
0
-1
0
1
2
3
4
5
6
7
8
Sample Number

Why Simulation?

- Compare observables with expectations from theoretical models, which leads to physics results
- Design detectors
- Optimise trigger settings
- Tune analysis selections
- Estimate background and systematic errors
- Estimate efficiency, resolution and fake-rate
- …

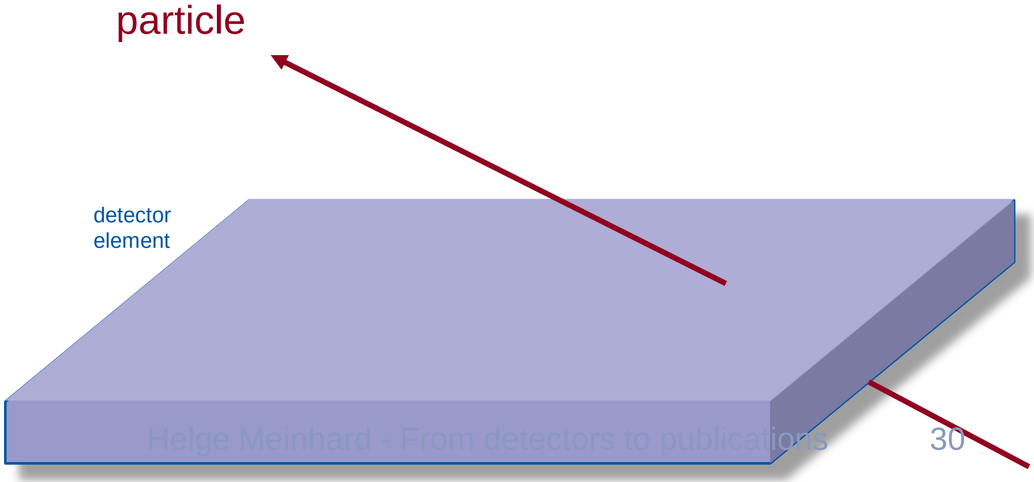
Simulation Steps: 1. Physics

- Simulate physics interaction at proton-proton collision
- Input: parameters of physics model
- Output: events with four-vectors of secondary particles



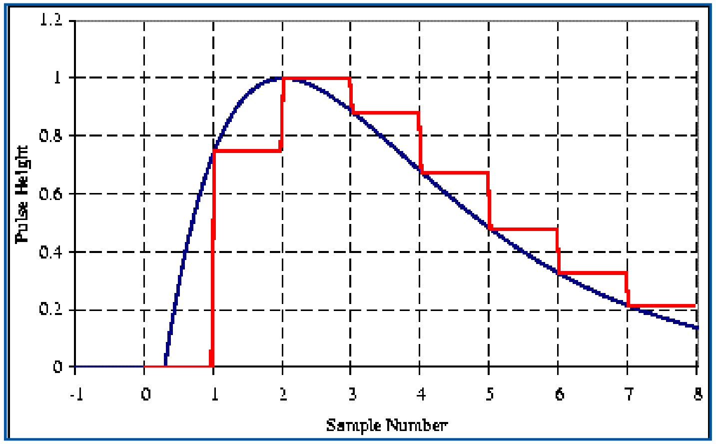
Simulation Steps: 2. Detector

- Simulate interaction of secondary particles with the detector material
  - Includes ionisation, bending of charged particles in magnetic field, …
  - Requires very complete and detailed detector description
- Based on standard toolkit GEANT 4
- Very CPU-intensive (more so than reconstruction)

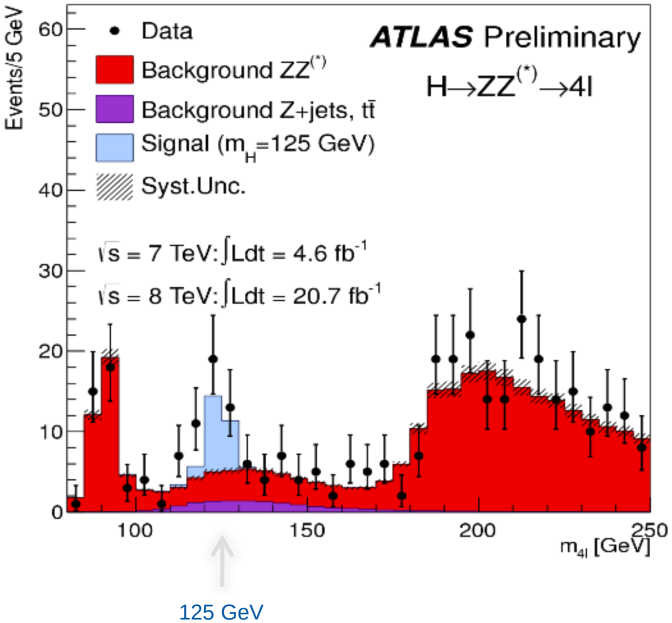


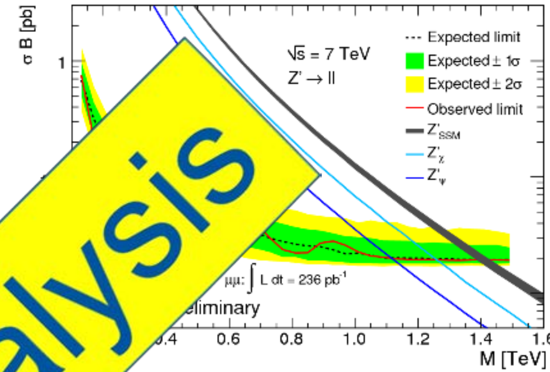
Simulation Steps: 3. Electronics

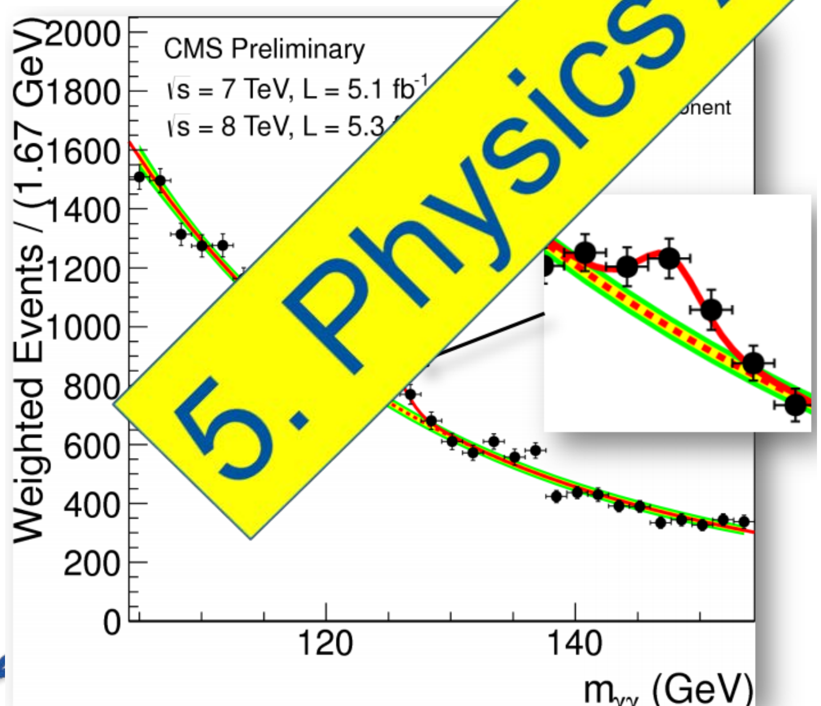
- Simulate the response of the detector elements to interactions of secondary particles passing through
  - Requires very detailed description of detector electronics
- Technically often combined with step 2 (detector simulation) – GEANT 4
- Output is very similar to detector raw data
  - Information of ‘truth’ kept all the way through

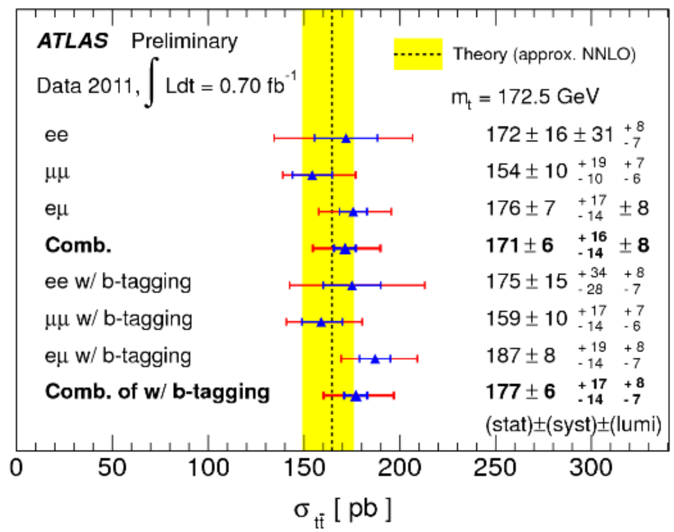


5. Physics Analysis









Types of Physics Analyses

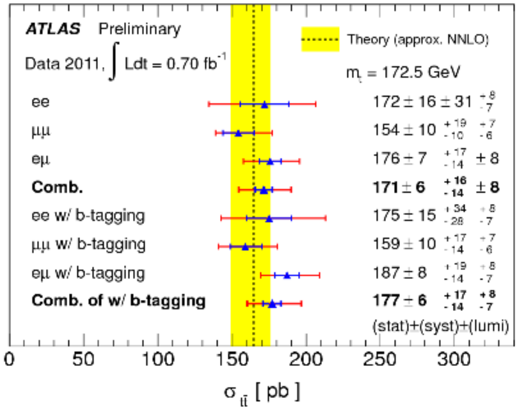
- Search for new particles / phenomena
  - Usually limited by statistics
  - Negative result (nothing found) sets new exclusion limits
- Precision measurements
  - Usually limited by systematic uncertainties
  - Serve as consistency check of the underlying theory (mostly the Standard Model)

Physics Analysis Workflow (1)

- Start with the output of reconstruction
- Apply an event selection based on the reconstructed object quantities
  - Often calculate new information, e.g. masses of combinations of particles
  - Event selection designed to improve the 'signal' to 'background' in your event sample
- Estimate
  - Efficiency of selection (& uncertainty)
  - Background after selection (& uncertainty)
  - Can use simulation for these – but have to use data-driven techniques to understand the uncertainties
- Make final plot
  - Comparing data to theory
  - Correcting for efficiency and background in data
  - Include the statistical and systematic uncertainties

Physics Analysis Workflow (1)

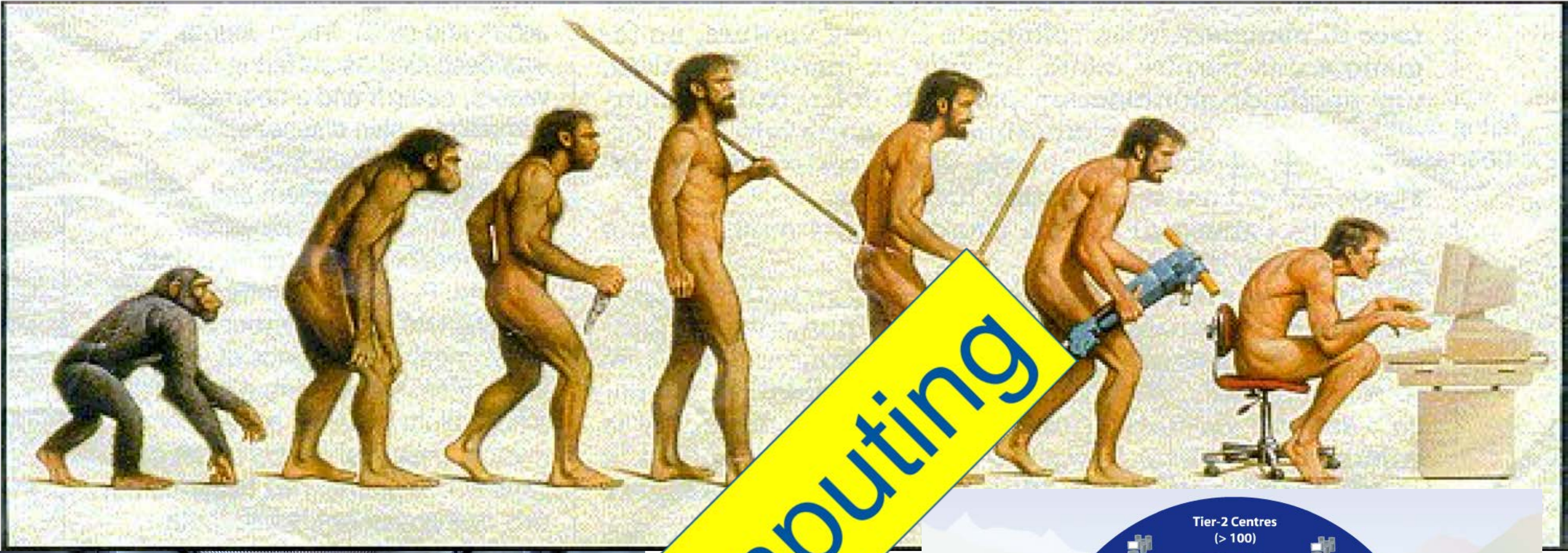
- Start with the output of reconstruction
- Apply an event selection based on the reconstructed object quantities
  - Often calculate new information, e.g. masses of combinations of particles
  - Event selection designed to improve the ‘signal’ to ‘background’ in your event sample
- Estimate
  - Efficiency of selection (& uncertainty)
  - Background after selection (& uncertainty)
  - Can use simulation for these – but have to use data-driven techniques to understand the uncertainties
- Make final plot
  - Comparing data to theory
  - Correcting for efficiency and background in data
  - Include the statistical and systematic uncertainties

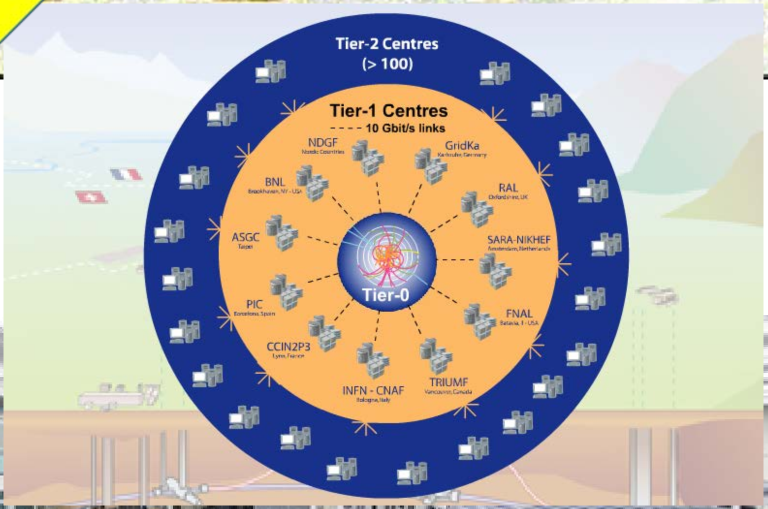


Physics Analysis Workflow (2)

- Output of reconstruction is large (order of tens of PBs); simulation similarly large
- LHC experiments have defined reduced data sets (both general and specific to certain physics channels)
- Sometimes, final steps of analysis are based on an even more reduced private data set
- Many of the analyses use standard packages for their final steps including presentation of plots and histograms, e.g. ROOT

6. Computing





Tier-2 Centres
(> 100)
Tier-1 Centres
10 Gbit/s links
NDGF
GridKa
BNL
RAL
ASGC
SARA-NIKHEF
PIC
FNAL
CCIN2P3
INFN - CNAF
TRIUMF
Tier-0

The Nature of the Problem

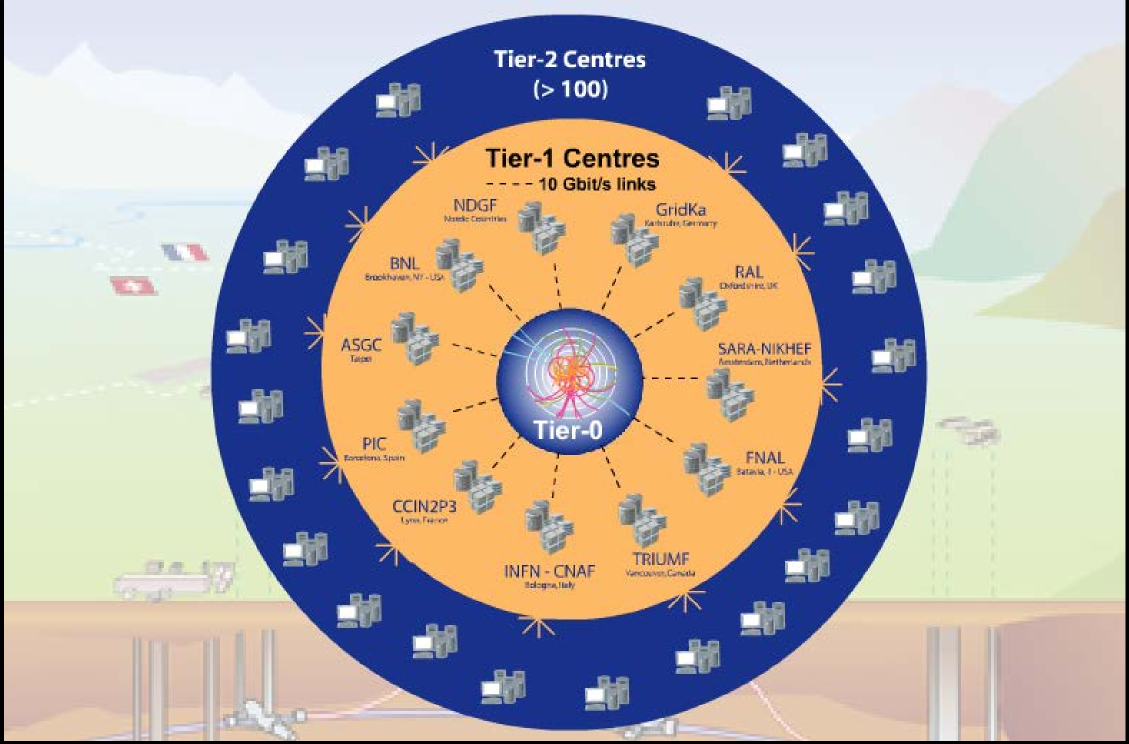
- Enormous numbers of collisions of proton bunches with each other
  - Data from each collision are small (order 1…10 MB)
  - Each collision independent of all others
- No supercomputers needed
  - Most cost-effective solution is standard PC architecture (x86) servers with 2 sockets, SATA drives, Ethernet network
  - Linux (Scientific Linux, RHEL variant) used everywhere
- Calculations are mostly combinatorics – integer (rather than floating-point) intensive

The Scale of the Problem

- 2012: 30’000 TB of new data from LHC experiments
    - CD tower of 50 km height, or 1’000 years of videos on DVD
- Requires 250’000 fast compute cores, 200’000 TB disk space, 207’000 TB tape space
- CERN able to provide some 15% of this capacity

WLCG – what and why?

- A distributed computing infrastructure to provide the production and analysis environments for the LHC experiments
- Managed and operated by a worldwide collaboration between the experiments and the participating computer centres
- The resources are distributed – for funding and sociological reasons
- Our task was to make use of the resources available to us – no matter where they are located



Tier-0 (CERN):
- Data recording
- Initial data reconstruction
- Data distribution

Tier-1 (11 centres):
- Permanent storage
- Re-processing
- Analysis

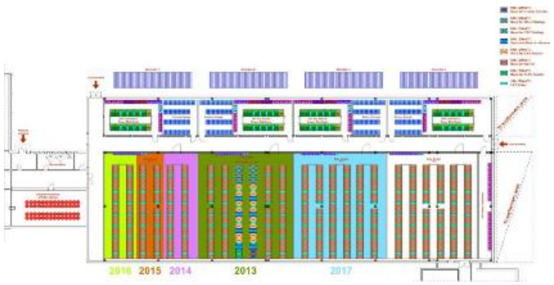
Tier-2 (~130 centres):
- Simulation
- End-user analysis

CERN Computer Centre

- 15% of LHC compute requirements
- Services for other (smaller) experiments
- Computing infrastructure for CERN and the collaborations hosted
- January 2014: 10’500 servers, 76’000 disk drives, 52’000 tape cartridges
- Total power/cooling envelope for IT equipment: 3.5 MW
  - Not sufficient for requirements of LHC Run 2

Extension: Wigner data centre in Budapest, Hungary

- Result of open tender in CERN member states
- First few 100 machines in production
- Final envelope: 2.5 MW



CERN openlab in a nutshell

- A science – industry partnership to drive R&D and innovation with over a decade of success
- Evaluate state-of-the-art technologies in a challenging environment and improve them
- Test in a research environment today what will be used in many business sectors tomorrow
- Train next generation of engineers/employees
- Disseminate results and outreach to new audiences

PARTNERS

HUAWEI

intel

ORACLE

SIEMENS

CONTRIBUTOR

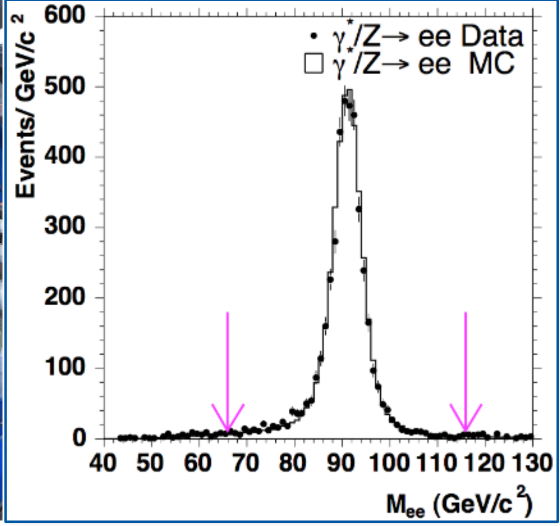
rackspace
the open cloud company

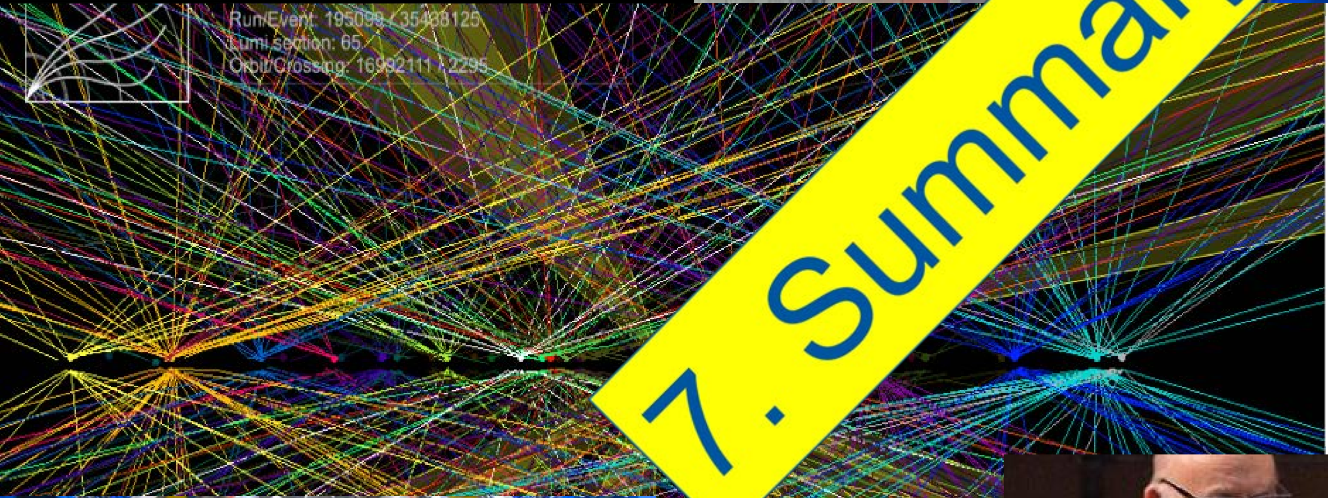
ASSOCIATE

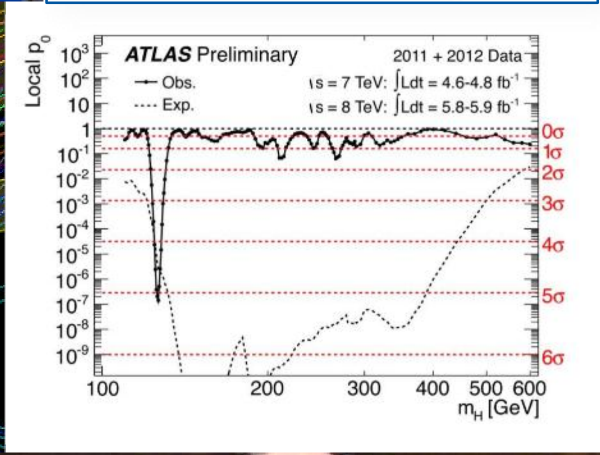
Yandex

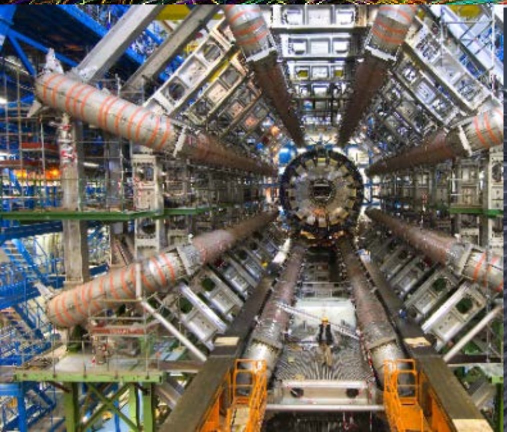
7. Summary


LHCb
ATLAS
CERN
CMS
ALICE

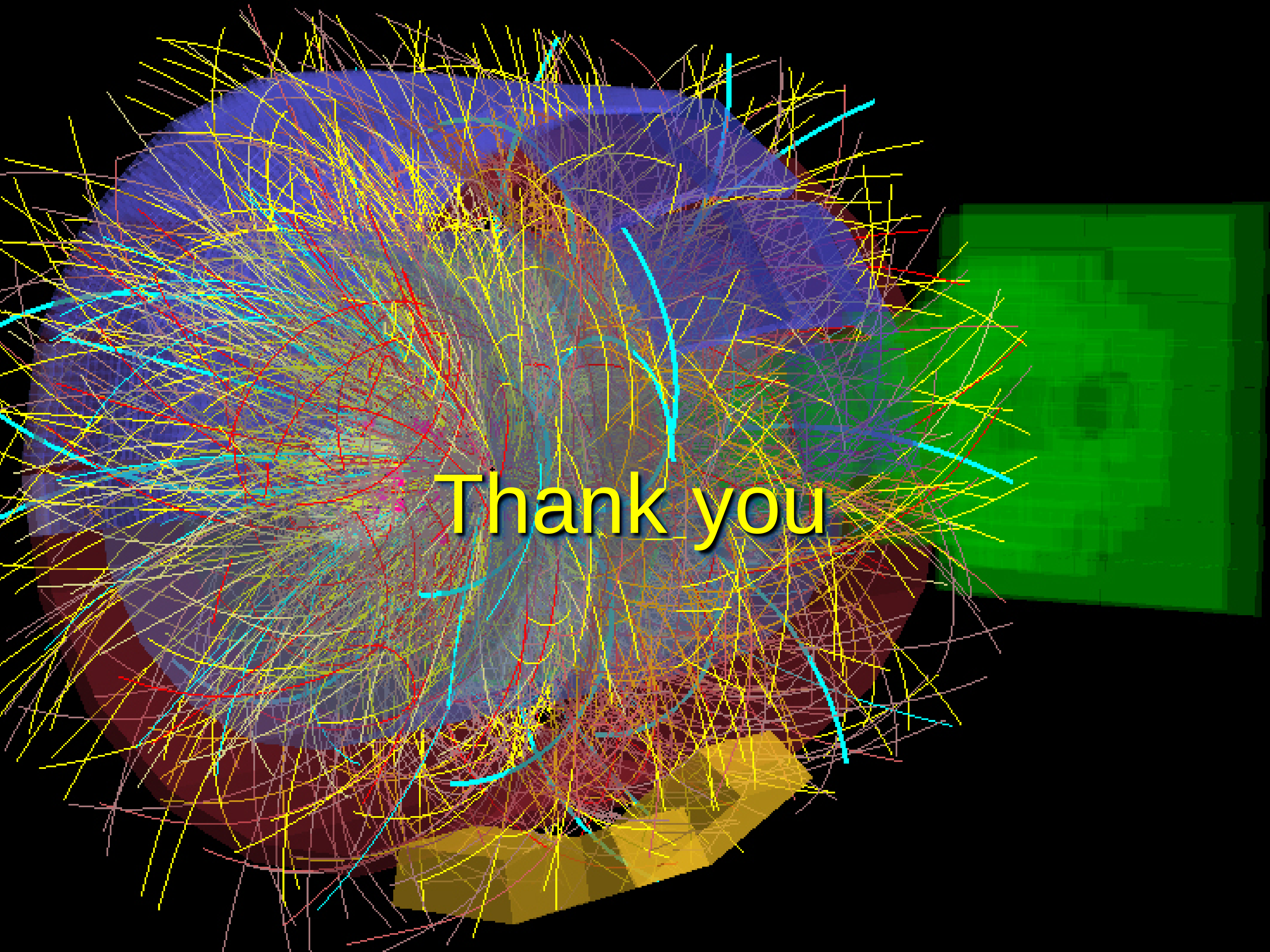

$\gamma^*/Z \to ee$ Data
$\gamma^*/Z \to ee$ MC
Events/ GeV/c^2
M_{ee} (GeV/c^2)
0
100
200
300
400
500
600
40 50 60 70 80 90 100 110 120 130


Run/Event: 195099 / 35438125
Lumi section: 65
Orbit/Crossing: 16992111 / 2295


ATLAS Preliminary
2011 + 2012 Data
Obs.
Exp.
$\sqrt{s}$ = 7 TeV: $\int Ldt$ = 4.6-4.8 fb^{-1}
$\sqrt{s}$ = 8 TeV: $\int Ldt$ = 5.8-5.9 fb^{-1}
Local p_0
0σ
1σ
2σ
3σ
4σ
5σ
6σ
100 200 300 400 500 600
m_H [GeV]



- Physics results are statistical comparisons of observables with their theoretical predictions
- Trigger/data acquisition, reconstruction, analysis, and simulation is how we get there
- Sophisticated algorithms needed for reconstruction and analysis
    - Calibration and alignment must be well controlled
- Detailed simulation is key to the success
- The single event is small and simple, but the computing scale is enormous
- Software and computing must work very well in order to achieve results with the LHC


Thank you