Introduction





Particle physics – Why do we care?



• What particles ?

 These particles we know, and they are the building blocks of the 5%



Collisions at LHC



Physics at the LHC: the environment

What do we mean by particle reconstruction and identification at LHC? Elementary constituents interact as such in "hard processes", namely:

Quarks and leptons as matter particles, and

	e (0.0005)	μ (0.105)	τ (1.777)
Leptons	ν _e	\mathbf{v}_{μ}	ν,
Quarks	u (< 0.005)	c (~ 1.25)	t (~ 175)
	d (< 0.005)	s (~ 0.1)	b (~ 4.2)

Gluons and EW bosons as gauge particles

Gluon(0)	Photon	W⁺,W⁻	Z
Colour octet	(0)	(80.42)	(91.188)

masses in GeV

All

- e, ν, γ : only rigorously stable particles
- μ: at collider energies, do not decay in the detector
- $\tau,c,b:\qquad \ \ Live \ long \ enough \ to \ travel \ a \ short \ distance, \ can \ see \ separated \ vertexes.$

All other particles can only be seen through their stable decay products

Energy Frontier: Jets





Physics at the LHC: the environment Which type of particles does one actually see in the final state? LHC physics processes are dominated by strong interactions (QCD) :

- Jets : hard processes: quarks and gluons materialise as hadronic jets, which consist mostly of charged and neutral hadrons (pions, kaons, and to a lesser extent protons and neutrons, which at these energies can be all considered as stable).
- "soup": soft processes: non-perturbative QCD processes with soft gluons materialising as almost uniform soup of charged and neutral π, K, etc.
- Heavy quarks with "long" lifetime are produced abundantly also
 High-p_T (above ~ 10 GeV) leptons are produced mostly in c, b decays.
 High-p_T isolated leptons may be found in fraction of J/ψ and Y decays
 For p_T > 25 GeV, dominant source of high-p_T leptons: W/Z/tt decays

<u>Main challenge at LHC</u> : find e, γ , μ , τ , b amidst q/g soup



Introduction (A)

From an outside perspective on high energy physics, people are mostly impressed by the enormous size and complexity of the accelerators and detectors.



25m





Introduction (B)

However, roughly the same effort goes into the *data acquisition*, *calibration*, *reconstruction*, and *physics analysis* of the recorded data!



.. our final product are publications in scientific journals!



Introduction (C)

- This lecture was originally made by Jan-Fiete Grosse-Oetringhaus for the German teachers programme (with input from Jamie Boyd, Ian Fisk, and Andreas Hirstius).
- Some numbers might partially already be outdated (they evolve very fast!).
- Please feel free to interrupt me at any time in case there are questions or comments!

From a collision to a physics result



Step-by-step

- Trigger: record only the interesting events
- Data acquisition
- Long term storage of the data:
 - for further processing
 - long term data preservation projects (our data is 'expensive') and legal requirements
- Reconstruction: from electric signals to particles (tracks)
- Interpretation of the collision -> physics result ("physics analysis")



Cross section, Luminosity, Reaction Rate



Luminosity

In scattering theory and accelerator physics, luminosity is the number of particles per unit area per unit time times the opacity of the target, usually expressed in cm-2 s-1. The integrated luminosity is the integral of the luminosity with respect to time. The luminosity is an important value to characterize the performance of an accelerator.



Rapidity, pseudo-rapidity



In the limit where the particle is travelling close to the speed of light, or in the approximation that the mass of the particle is nearly zero, pseudorapidity is numerically close to the experimental particle physicist's definition of rapidity,

$$y = rac{1}{2} \ln \left(rac{E + p_L}{E - p_L}
ight)$$

In hadron collider physics, the rapidity (or pseudorapidity) is preferred over the polar angle θ because, loosely speaking, particle production is constant as a function of rapidity. One speaks of the forward direction in a hadron collider experiment, which refers to regions of the detector that are close to the beam axis, at high $|\eta|$.

Physics at the LHC: the environment





rङ dơ/dp_Tdy is Lorentz-invariant

 $\Im \eta = y$ for $m \approx 0$

☞Physics is ~ constant versus η at fixed p_T ATLAS/CMS: from design to reality One word about neutrinos in hadron colliders: ✓ since most of the energy of the colliding protons escapes down the beam pipe, one can only use the energy-momentum balance in the transverse plane

→ concepts such as E_T^{miss}, missing transverse momentum and mass are often used (only missing component is E_z^{miss})

→ reconstruct "fully" certain topologies with neutrinos, e.g. W → Iv and even better H → $\tau\tau$ → Iv₁v_τ hv_τ

- ✓ the detector must therefore be quite hermetic
 - \rightarrow transverse energy flow fully measured with reasonable accuracy
 - → no neutrino escapes undetected
 - → no human enters without major effort (fast access to some parts of ATLAS/CMS quite difficult)

An example from the ATLAS detector Reconstruction of a $2e2\mu$ candidate for the Higgs boson - with $m_{2e2\mu}$ = 123.9 GeV

We need to understand the interaction of particles with matter in order to understand the design and operation of this detector, and the analysis of the data.





Data volume

1 Byte = 1 character 100 Byte = 1 SMS text message 1.024 Byte = 1 Kilobyte = 1 EMail 1.048.576 Byte = 1 Megabyte = 10 Min. phone call / 1 Foto / 1 LHC collision 734.003.200 Byte = 700 Megabyte = 1 CD-ROM 1.073.741.824 Byte = 1 Gigabyte = Data which one LHC experiment produces per second 5.046.586.572 Byte = 4,7 Gigabyte = 1 DVD 1.099.511.627.776 Byte = 1 Terabyte = Library with one million books 219.902.325.555.200 Byte = 200 Terabyte = 10 billion websites 1.801.439.850.948.198 Byte = 1,6 Petabyte = worldwide production of information in paper form 11.258.999.068.426.240 Byte = 10 Petabyte = yearly LHC data production 5.764.607.523.034.234.880 Byte = 5 Exabyte = yearly worldwide production of data in form of radio, TV etc.

19.599.665.578.316.398.592 Byte = 17 Exabyte = data volume of all phone calls worldwide per year



- The trigger system decides when an interesting collision (event) should be recorded => needs fast detectors to start the readout of slow detectors!
- Example: random versus triggered photo camera

- Goal: Recording of interesting collisions
 - Minimum bias: all collision, but at least one collision in the bunch crossing



- The trigger system decides when an interesting collision (event) should be recorded => needs fast detectors to start the readout of slow detectors!
- Example: random versus triggered photo camera



- Goal: Recording of interesting collisions
 - Minimum bias: all collision, but at least one collision in the bunch crossing



- The trigger system decides when an interesting collision (event) should be recorded => needs fast detectors to start the readout of slow detectors!
- Example: random versus triggered photo camera



- Goal: Recording of interesting collisions
 - Minimum bias: all collision, but at least one collision in the bunch crossing



- The trigger system decides when an interesting collision (event) should be recorded => needs fast detectors to start the readout of slow detectors!
- Example: random versus triggered photo camera





- Goal: Recording of interesting collisions
 - Minimum bias: all collision, but at least one collision in the bunch crossing
 - Dedicated selection of interesting events



- The most interesting events are occurring only rarely!
- E.g.: high energetic (high momentum particles) particles.





- The most interesting events are occurring only rarely!
- E.g.: high energetic (high momentum particles) particles.

 ≈ 98% of all particles are produced with
 *p*_T < 2 GeV/*c* even at LHC energies.





- 1 billion collisions happen in the LHC every second
- Only 100 1000 can be recorded per second.
- Interesting signatures have to be recognised during the data taking (= "online")
- Example Z $\rightarrow \mu\mu$
 - Look for collisions with energetic muons
 - Die The tracks of 10^9 collisions have to be investigated in one second! μ
- Combination of hardware-(dedicated hardware) and software triggers (computing farms)
 -> High Level Triggers (HLT)





Trigger – ATLAS as an example

- collision rate 40 MHz
- Level 1 Trigger ~ 100 kHz
 –first selection
 - -interesting regions in the detector are selected
- Level 2 Trigger ~ 5 kHz

 analysis of the interesting regions
- Level 3 Trigger ("event filter") ~ 400 Hz
 - Fast analysis of the entire event





Data recording

- As soon as a collision is identified as interesting, it is recorded
- Putting together the information from many sub detectors ("Event building")
 - Complex process, because many detector elements read out several collisions at the same time, but only parts should be saved.
 Event size ranges from 35 kB (LHCb) to 1-2 MB (ALICE, ATLAS, CMS) for pp collisions but up to 50 MB for PbPb collisions
- The data of several events are stored in a single file and temporarily saved on hard disks (buffer) –up to 1 GB/s; > 1 CD per second
- Data is then shipped to permanent storage on tapes –total amount of data per experiment and year: 1-2 PB



Event size



Event size depends on the number of particles produced in the collision!

Reconstruction

Small part of a read-out of an event:

audiess -		
	le84ce0: 0x0000 0x0019 0x0000 0x0000 0x0000 0x0000 0x01e8 0x5b7e	0x01e84ce0:
	le84cf0: 0x01e8 0x87ec 0x01e8 0x85d8 0x7363 0x616e 0x0000 0x0000	0x01e84cf0:
	le84d00: 0x0000 0x0019 0x0000 0x0000 0x0000 0x0000 0x01e8 0x5b7c	0x01e84d00:
	le84d10: 0x01e8 0x87e8 0x01e8 0x8618 0x7365 0x7400 0x0000 0x0000	0x01e84d10:
data = wh	le84d20: 0x0000 0x0019 0x0000 0x0000 0x0000 0x0000 0x01e8 0x5b7c	0x01e84d20:
	le84d30: 0x01e8 0x87a8 0x01e8 0x8658 0x7370 0x6c69 0x7400 0x0000	0x01e84d30:
	le84d40: + 0x0000 0x0019 0x0000 0x0000 0x0000 0x0000 0x01e8 0x5b7c	0x01e84d40:
	le84d50: 0x01e8 0x8854 0x01e8 0x8698 0x7374 0x7269 0x6e67 0x0000	0x01e84d50:
	le84d60: 0x0000 0x0019 0x0000 0x0000 0x0000 0x0000 0x01e8 0x5b7c	0x01e84d60:
	le84d70: 0x01e8 0x875c 0x01e8 0x86d8 0x7375 0x6273 0x7400 0x0000	0x01e84d70:
	le84d80: 0x0000 0x0019 0x0000 0x0000 0x0000 0x0000 0x01e8 0x5b7c	0x01e84d80:
	le84d90: 0x01e8 0x87c0 0x01e8 0x8718 0x7377 0x6974 0x6368 0x0000	0x01e84d90:
t f		
VA		

reconstruction = transform the electric signals into tracks (particles)

Track: ϕ = 0.23, η = 0.75, p_T = 2.3 GeV/c

-address = which detector?

data = what was read out?







Reconstruction steps

- Turn electric signals of the detector into physical quantities
 - -detector element 1244 has measured at time stamp 1333096259.344245 a signal of 120 ADC counts
 - \rightarrow signal at position x = 1.2 cm, y = 4.5 cm, z = 3.2 cm, deposited energy of 100 keV
 - –Needs information about the exact position of the detector element (*Alignment*) and calibration
- Particles often leave traces in neighbouring detector elements

-Clusterization of signals

- Make tracks out of the clusters (pattern recognition)
- Combination of several tracks (collision vertex, particle decays)



Pattern recognition

- Finding tracks (tracking)
 - -Linear in direction of the magnetic field
 - Curvature perpendicular to the magnetic field allows momentum determination

collision vertex



secondary vertices
 –allows to identify *weak* decays



Alignment







Commissioning using cosmic events



Haw HULL Dones HINLE, LE 161, Orbit LUMPHILL BE 201









Secondary vertices





Secondary vertices





Historical remarks

Bubble chamber photo



The 80-inch Bubble Chamber

BNL, First Pictures 1963, 0.03s cycle

Discovery of the $\Omega^{\text{-}}$ in 1964

Momentum Measurement in Tracking Device

- Charged particles deflection in magnetic field:
 - Lorentz force ⊥ to B-field and to particle direction
 - Particle trajectory projected onto plane \perp to *B*-field is *circle* with radius: $r[m] = \frac{p_r[\text{GeV}]}{0.3 \cdot B[\text{T}]}$
 - For $p_T = 10 \dots 1000$ GeV and $B = 2 \text{ T} \rightarrow R = 17 \dots 1700 \text{ m} (cf, R_{\text{ID}} \sim 1 \text{ m})$
 - manual if p_T < 0.5 GeV, the particle is trapped in solenoid → "loopers"</p>
- Obtain r and p_T from measurement of sagitta:

$$s = r(1 - \cos\frac{\alpha}{2}) \approx r\frac{\alpha^2}{8} \qquad s \approx \frac{1}{8}\frac{L^2B}{p_T}$$

Track fitting in LHC environment challenging

8

- Must handle ambiguities, hit overlaps, multiple scattering, bremsstrahlung, multiple vertices,
- Track fitters take Gaussian noise (Kalman) and non-Gaussian noise (GSF) into account
- Fitter must be fast, used in high-level trigger



B-field map accuracy ~10⁻⁴

Particle ID and Kinematics



Neutrinos are only detected indirectly via 'missing energy' not recorded in the calorimeters

Tracking detector

-Measure charge and momentum of charged particles in magnetic field

Electro-magnetic calorimeter -Measure energy of electrons, positrons and photons

Hadronic calorimeter

-Measure energy of hadrons (particles containing quarks), such as protons, neutrons, pions, etc.

Muon detector

-Measure charge and momentum of muons

ATLAS detector Wedge





20/07/2016

Roland Jansky - Searching for new physics at ATLAS: Rough guide to data analysis

Analysis objects



- List of objects needed for analysis
 - Leptons
 - muons, electrons, taus
 - Photons
 - Hadronic jets
 - Energy sums
 - Missing transverse energy
- Measure p_T , E_T , η and φ of objects

Electron identification: Track/Cluster matching



colorimeter module Lovers and granularity (3/(a,3/i) in berral; Presempler (0.025 x 0.1) Ships (0.003 x 0.1) Middle (0.025 x 0.025) Bock (0.05 x 0.025) one cell In the middle Layer 0.025x0.025 Cluster 0.075±0.175 LAr energy reconstructions Call shorpy: $E(\hat{x}d\hat{x}') = f_{\hat{x}\hat{x}\hat{x}-\alpha\hat{x}\hat{x}} \times f_{\hat{x}\hat{x}-\alpha\hat{x}\hat{x}} \times \frac{M_{\alpha\hat{x}\hat{x}}}{M_{A\alpha\hat{x}}} \times g_{\hat{x}\hat{x}\hat{x}-\alpha\hat{x}\hat{x}} \sum_{i=1}^{n} \alpha_i (\hat{s}_i - P)$ Electronic collibration constants: p = pedectol a = optimal filtering $f.g = ADC \rightarrow GeV$

alaria eter madale

Electron = track in the Inner Detector (direction measurement) Matched to an EM cluster in the calorimeter (energy measurement). Need to know:

- material distribution in the Inner detector in the front of the EM calorimeter
- calibration of the energy response of the calorimeter
- rejection against jet faking electrons

Cluster: sum of cell energies over all layers

The EM colorimeter half haved and ergostat

Photon identification



Photon = no track in the Inner Detector and an EM cluster in the calorimeter However: because of materials in the Inner Detector and in front of the calorimeter, Photon may convert into e-e+ pair. \rightarrow photon may be reconstructed as single or double Track conversion

Muon identification

Muon: identified as tracks in the Inner Detector and the Muon Spectrometer. For muon with enough energy to pass through the calorimeter, then energy loss in the calorimeter must be corrected for.

Internal alignment of Muon chambers important and relative alignment of the inner Detector and Muon Spectrometer also important to match track segments From both detectors.

Magnetic field mapping also important



Figure 1.3: Cut-away view of the ATLAS calorimeter system.





all stars to show the last

Figure 6.1: Cross-section of the barrel muon system perpendicular to the beam axis (non-bending plane), showing three concentric cylindrical layers of eight large and eight small chambers. The outer diameter is about 20 m.



Figure 6.2: Cross-section of the muon system in a plane containing the beam axis (bending plane). Infinite-momentum muons would propagate along straight trajectories which are illustrated by the dashed lines and typically traverse three muon stations.

Hadronic tau identification

The transverse momentum range of interest spans from below 10 GeV up to 500 GeV. τ leptons decay hadronically in 64.8% of all cases, while in $\sim 17.8\%$ (17.4%) of the cases they decay to an electron (muon) [1]. From the detection point of view, hadronic modes are divided by the number of charged π s among the decay products into single-prong (one charged π) and three-prong (three charged π s) decays. The small fraction (0.1%) of five-prong decays is usually too hard to detect in a jet environment. The $\tau \to \pi^{\pm} v$ mode contributes 22.4% to single-prong hadronic decays and the $\tau \to n\pi^0 \pi^{\pm} v$ modes 73.5%. For three-prong decays, the $\tau \to 3\pi^{\pm}\nu$ decay contributes 61.6%, and the $\tau \to n\pi^0 3\pi^{\pm}\nu$ mode only 33.7%. In general, one- and three-prong modes are dominated by final states consisting of π^{\pm} and π^{0} .

Properties of hadronically decaying τ -leptons:

- Collimated energy deposition in calorimeter
- I or 3 charged decay products (π[±])
- Isolated EM clusters corresponding to π⁰ in τ-decay
- Modest but significant proper lifetime
- BR (τ → hadrons) = 64.8%



Thus also needs the tracks (Inner Detector) associated to a narrow Cluster. Need a strong rejection against jets, electrons, while maintaining high tau-jet reconstruction efficiency

Jet and Missing ET identification ATLAS calorimeters

Main features for jet and E_T^{Miss} reconstruction and calibration:

- Non compensating (e/h >1) :
 - Response to hadrons jpeinternamenetic lower than that to electrons and photons
 - Developed specific calibrations
- Dead material:
 - Energy loss before EM calorimeter and between EM and HAD barrel calorimeters:
 - dead material corrections
- Different technologies and many transition regions:
 - "Crack" regions: η ≈ 1.4, 3.2
- Magnetic field bending



 $\eta = -\log(\tan(\theta/2))$

Jet and Missing ET identification

- Topo-Clusters: group of calorimeter cells topologically connected
 - Noise suppression via noise-driven clustering thresholds:
 - Seed, Neighbour, Perimeter cells (S,N,P) = (4,2,0)
 - seed cells with |E_{cell}| > So_{noise} (S = 4)
 - expand in 3D; add neighbours with |E_{cel}|>Nσ_{noise} (N = 2)
 - merge clusters with common neighbours (N < S)
 - add perimeter cells with |E_{cel}|>Po_{noise} (P = 0)
 - Attempt to reconstruct single particles in calorimeter
- Towers: thin radial slice of calorimeters of fixed size
- Topo-Tower: selecting only the cells in the tower with a significant signal



Jet Reconstruction

Sequential process:

- Input signal selection:
 - TopoClusters, Towers, TopoTowers
- Jet finding:
 - The jet finding algorithm groups the collection of clusters(towers) according to geometrical and/or kinematic criteria.
 - Many algorithms studied in ATLAS:
 ⇒ recently concentrated on AntiKt algorithm
- Jet calibration:
 - depending on detector input signal definition, jet finder choices...
- Jet selection:
 - apply cuts on kinematics to select jets of interest



Track jets use tracks as input to the jet finding and reconstruction. This would Miss the neutral component of the jet. However track jets are useful in a Number of applications

Missing ET Reconstruction



Fake Missing ET

- Fake muons can be caused by jet punch-through detected as excess activity in Muon Chambers.
- Cleaning criteria: count of muon hits and of muon segments within a cone around jet axes.
- Missing muons due to detector features
 - η=0: holes in Muon Spectrometer for cables, services to Inner Detector & Calorimeter.
 - |n| ~1.2: middle muon station missing for initial data taking
 - |n|>2.7: no muon coverage
- use calorimeter and track information to recover missing muons used in ${\sf E}_{{\sf T}^{{\sf miss}}}$ calculation





Fake Missing ET

Fake E_T^{miss} in calorimeter can also be produced by mis-measurements of jets due to cracks, gaps, transition regions used for services.

- Leakage of jets entering 'crack' region 1.3<|n|<1.6 can be detected:
 - looking for large deposits in the outermost layers of the calorimeter
 - checking the E_T^{miss} calculated from tracks found in the Inner Detector that can provide a complementary information
 - checking if E_T^{miss} is closely associated with one of the leading jets in the transverse (φ) plane

 Cleaning cuts based on those criteria could be applied⇒ analysis dependent





Calibration





CMS Experiment at the LHC, CERN Data recorded: 2012-May-13 20:06:14.621490 GMT Run/Event: 194108 / 564224000

H **→**γγ candidate









Simulation (1)

- Detailed simulations are needed to understand detector effects for the measurements
- A detector simulation reproduces the behaviour of the detector as well as possible (Monte-Carlo technique).

• Steps of a simulation:

- -Collision generator (physics): simulates which particles are created in a collision (e.g. Pythia)
- -Geometry: Exact description of all active (= sensitive and working) and inactive (= frames, cooling, cables, ...) elements in the detector
- -*Transport* software: simulates interaction of particles with detector material (e.g. GEANT4)
 - •Particles can decay, be deflected (scattering) or get stuck
 - •Energy deposit into the active detector elements is recorded
- Afterwards, the simulated signals are reconstructed with the same reconstruction software as the real data!

CERN



- Results of the simulation allow to correct for the distortions caused by the detector on the measurement.
- Important examples
 - –Track reconstruction efficiency: How probable is it to measure a particle? How many decay? How many get stuck?
 - Resolution studies: how precise do we measure a certain quantity.
 - -"fake rate" how often do we combined noise into a track (or by combinatorics) even though there was none?







Physics analysis

- Interpretation of the physics objects (tracks, vertices, calorimeter information) which were extracted during the reconstruction.
 - -Processing of many collisions in order to obtain a statistically significant result.
 - -Most common task of (young) high energy physicists
- Step-by-step
 - -Selection of interesting collisions
 - -Combine the physics objects from the reconstruction to the observable under study (e.g. decay pattern etc.)
 - -Summarize the relevant physics quantities (mainly in histograms)
 - -Correction for detector effects
 - -Extraction of the final physics measurement
 - -Estimation of the systematic uncertainties (<- often most of the work!!)
 - -Comparison of measurement with theoretical predictions.
- Reconstruction is a central task done by a few groups guided by experienced experts. Physics analysis is done by essentially all groups.



LHC Detectors are well-described in simulation e.g. Tracking -Material

TOB1

R (cm)



TIR

20

2 0.008

0.006

0.004

0 ***

10







Commissioning with Beam : Tracking- Resonances



Commissioning with Beam : Tracking-ECALs



Commissioning with Beam: Understanding Muons, b-tagging, Triggers



Muons remarkably well understood, b-tagging progressing well, newer triggers being rapidly understood as luminosity increases

Conclusions

- We need to understand the interactions of particles with matter in order to understand the design and operation of the detector, and the analysis of the data
- See latter detailed lectures on
 - Geant4 detector simulation
 - Data analysis with ROOT
 - Statistical analysis
 - Extracting physics results from LHC data