This review article introduces the design of the general purpose experiments ATLAS and CMS, which independently discovered the Higgs boson, showing how generic features are motivated by the characteristics needed to explore the physics landscape made accessible by the Large Hadron Collider accelerator, whose high collision rate creates an extremely challenging operating environment for instrumentation. Examples of the very different component designs chosen by the two experiment collaborations are highlighted, as an introduction to briefly describing techniques used in the construction of some of these elements and, subsequently, in the assembly of both detection systems in their respective underground caverns.

1. Introduction

Formidable technical challenges had to be overcome during the design and construction of the general-purpose experiments ATLAS [1–3] and CMS [4–6], used to discover the Higgs boson [7–11] in proton–proton (p–p) collisions at CERN’s Large Hadron Collider (LHC) [12,13]. These challenges arise directly from the practicalities of searching for very rare processes at high energies, the motivation for the construction of the LHC. The general aims of both experiments are to test how matter constituents and force carriers acquired, in part, distinctive masses; to search for candidate particles to explain dark matter; to search for pointers to how gravity could be consolidated into a unified theory; and to search for unexpected phenomena, such as previously unobserved matter constituents, subconstituents or
forces. Specifically, the experiments are designed to capture the outcome (produced particles and energy) of all p–p collisions occurring during encounters (‘crossings’) between circulating bunches of protons, rejecting those which are not interesting and disentangling individual, interesting p–p collisions (‘events’) from uninteresting ones, reconstructing the outcome of each and identifying distinctive signatures (of known and new phenomena) in the set of interesting events. Because the lifetime of the Higgs boson after production is unobservably short (around $10^{-22}$ s for the observed mass), only its decay products (familiar particles) remain to be detected. Among many possible decay modes into familiar particles of the standard model [14–16], some are distinctive footprints or ‘signatures’. These signatures typically involve leptons with high momentum transverse to the beam direction, jets of hadrons originating from a quark or gluon, ‘missing energy’ (an apparent topological imbalance signifying a particle invisible to the detector, such as a neutrino) or short-lived heavy particles that fly a short, but detectable, distance in the detector before decaying.

2. Generic detector design

Detectors to identify the variety of signatures resulting from production of a Higgs boson or other novel heavy particle are hermetic, multi-layered (each layer optimized for a particular job) and highly granular (many detecting elements per layer). The concentric layers, usually engineered as a cylindrical barrel closed at both ends by endcaps, must identify and precisely measure the energies and directions of all the particles originating from a specific p–p collision. Working outwards from the p–p collision point, the generic detector consists of a tracking detector to reconstruct charged particle tracks bending in a magnetic field and thus measure their charge, momentum and point of origin; next, an electromagnetic calorimeter that absorbs and measures the energy and direction of electrons, positrons and photons; then, a hadronic calorimeter that absorbs and measures the energy and direction of hadrons (i.e. particles containing constituent quarks and gluons); finally, a muon spectrometer, measuring via magnetic bending the charge and momentum of charged particles emerging from the calorimeters.

3. The Large Hadron Collider environment

Because the mass of new particles is only weakly constrained, a very large sensitive dynamic range is needed and thus a proton collider, which produces a large range of constituent centre of mass energies, gives the maximum discovery potential. However, constituent collisions at high centre of mass energies are a rarity; the probability of a collision producing, for example, a Higgs boson is very low (a few hundred every trillion p–p collisions) and distinctive decays of the Higgs are also rare (e.g. 1 in approx. 500 leads to the distinctive two photon signature). The very high rate of p–p collisions needed to search for the Higgs on a reasonable time scale is achieved in the LHC by having bunches of 130 billion protons colliding 40 million times per second in each experiment. About 25 p–p collisions result each time bunches cross, resulting in an overall p–p interaction rate of $10^9$ per second. The huge track density (about 2000 per 25 ns) creates a very harsh radiation environment, and the detector builders face a challenging trade-off between occupancy versus granularity as well as the general effect of very high particle rates. Additionally, as the vast majority of bunch collisions contain no interesting events, a fast filtering system (‘trigger’) is needed to reject these, after which even those containing an interesting p–p collision event are usually superimposed on the result of the 25 uninteresting p–p collisions occurring during the same bunch crossing (a feature known as ‘pile-up’). Storing the 500 most promising events per second still leads (for an event record size averaging 1 MB) to a total data volume of 10 million gigabytes per year and a formidable data management and processing challenge.
Figure 1. Structure of the (a) ATLAS and (b) CMS detectors. (Online version in colour.)

Table 1. Principal systems and technologies of ATLAS and CMS.

<table>
<thead>
<tr>
<th>system</th>
<th>ATLAS</th>
<th>CMS</th>
</tr>
</thead>
<tbody>
<tr>
<td>tracking</td>
<td>silicon (pixel + microstrip)</td>
<td>silicon (pixel + microstrip)</td>
</tr>
<tr>
<td></td>
<td>gas (transition radiation)</td>
<td></td>
</tr>
<tr>
<td>electromagnetic calorimeter</td>
<td>sampling (lead/liquid argon)</td>
<td>PbWO₄ crystals</td>
</tr>
<tr>
<td>hadron calorimeter</td>
<td>sampling (steel/plastic scintillator)</td>
<td>sampling (brass/plastic scintillator)</td>
</tr>
<tr>
<td></td>
<td>(copper/liquid argon)</td>
<td></td>
</tr>
<tr>
<td>muon spectrometer (tracking + gas)</td>
<td>gas</td>
<td>gas</td>
</tr>
<tr>
<td>fast trigger</td>
<td></td>
<td></td>
</tr>
<tr>
<td>magnet</td>
<td>solenoid (tracking) + toroid (muon)</td>
<td>solenoid (around tracking and calorimetry) + iron return yoke</td>
</tr>
<tr>
<td>magnetic field</td>
<td>approximately 2 T in solenoid,</td>
<td>approximately 4 T in solenoid</td>
</tr>
<tr>
<td></td>
<td>approximately 0.5–1 T in toroid</td>
<td></td>
</tr>
</tbody>
</table>

4. Design of ATLAS and CMS

Prior to 1990, it was not obvious that detectors suitable for LHC conditions of rate, pile-up and radiation could be developed; however, by 1996, two solutions, ATLAS [1,2] and CMS [4,5], were approved for construction. Although many of the foundation detection principles are a century old, the implementations in these experiments involve unprecedented scale and precision, combined with rigorous construction techniques and testing procedures in order to operate well in the aggressive environment. The overall structure of ATLAS and CMS is illustrated in figure 1, and the technologies used in the different layers are shown in table 1.

Polar coordinates are used to specify the structures, except that the polar angle $\theta$ (relative to the beam axis) is replaced by pseudo-rapidity $\eta = -\ln[\tan(\theta/2)]$, because the flux of particles produced in $p$–$p$ collisions is approximately uniform in this variable.

In this review, only the magnet systems and electromagnetic calorimeters will be considered in any detail, as these are where the most radical differences between the two designs appear.
5. Magnet systems

The very different relative sizes and masses of ATLAS and CMS result from the different magnet solutions for momentum and charge analysis of charged particles through their bending in a magnetic field. Both these superconducting magnet systems required pushing the limits of technology and making the largest magnets of their type so far. The magnetic deflection of a charged particle reduces with increasing momentum and increases with magnetic field and with the square of the path-length within the field. The minimum detectable bending is limited by the precision of the sensor system and by perturbations of the particle trajectory owing to interactions with material along the flight-path.

In the case of CMS, a single, very high field (4 T), long, large bore solenoid surrounds all the tracking and calorimetry and an instrumented iron flux return yoke constitutes the muon measurement system. A transverse view of this solenoid, with part of the longitudinally segmented return yoke, is shown in figure 2, just before testing in the surface assembly building. This solution is relatively compact (hence the ‘Compact Muon Solenoid’, CMS), leading to a cylindrical experiment 15 m in diameter and 24 m long, having a mass of 12 500 t.

In the case of ATLAS, a thin solenoid (2 T) surrounds the tracking system and an arrangement of massive barrel and endcap toroids (hence ‘A Toroidal LHC Apparatus’, ATLAS) provide a field for muon tracking outside the calorimeters, of approximately 0.5 T in the barrel and 1 T in the endcaps. Figure 3 shows the barrel toroid system just after completion in the underground cavern. The ATLAS solution leads to cylindrical experiment of diameter 25 m and length 44 m, substantially bigger than CMS, but with a mass of 7000 t, substantially smaller than CMS.

The challenge of the ATLAS and CMS magnet systems is illustrated in figure 4 by comparing the magnets of some particle physics experiments, built or proposed over the past two decades, in terms of stored energy and stored energy per unit volume.
The CMS solenoid [17], with a stored energy of 2.6 GJ and a stored energy per unit cold mass of 11.6 kJ kg\(^{-1}\) was a step into unexplored territory compared with previous such magnets in particle physics. The key to its construction was the reinforced conductor, able to resist the radial magnetic pressure of 60 atm encountered during operation. The conductor was formed by continuous electron beam welding, in 2.6 km lengths, of aluminium alloy reinforcement on either side of a pure aluminium stabilizer containing the ‘Rutherford’ superconducting Nb–Ti cable. The four-layer coil made from this conductor was wound in five elements, which were assembled vertically and then rotated on a special jig for insertion into the steel vacuum tank, visible in figure 2, that houses the coil and provides thermal insulation.

The ATLAS toroids [18] were also the largest ever produced, each being wound in a ‘double pancake’ structure using a conductor whose pure aluminium stabilizer is reinforced by microalloying with Ni/Zn. As is evident in figure 3, the eight independent barrel toroids, each with its own cryostat, were assembled into a support structure embedded in the experimental cavern, whereas the two endcap torus systems involved incorporating a further eight, smaller coils into a common cryostat.

6. Electromagnetic calorimeters

The challenges facing those designing electromagnetic calorimeters (the second detector layer around the collision point) are to provide adequate energy and position resolution, a large dynamic range, fast and uniform response, hermetic coverage, radiation tolerance, low noise and good rejection power against hadrons. For instance, the energies of photons from the Higgs boson decay, \(H^0 \rightarrow \gamma\gamma\), must be measured to a precision of less than or equal to 0.5%. Calorimeters
are limited in their energy measurement by the primary statistics of energy deposit (improving as energy increases), by noise and pile-up effects (which are typically energy independent) and by non-uniformities in construction, which, because their effects scale with energy, can easily become the limiting factors on high energy performance. CMS chose a homogeneous calorimeter composed of lead tungstate scintillating crystals, each coupled to a pair of large-area silicon avalanche photodiodes (APDs). By contrast, ATLAS chose a sampling calorimeter with liquid argon as the active medium and interleaved absorber layers made of lead, formed into an ‘accordion-like’ structure. Although based on well-known principles, both technologies had to be developed from scratch and were the subject of years of prototyping, beam-tests and optimization, before the final designs were committed to manufacture.

The ATLAS ‘accordion’ calorimeter [3,19], illustrated in figure 5, was first proposed in 1990 and took 15 years to bring to fruition. It has the advantages of linear response, being very stable and intrinsically radiation hard (owing to the noble gas active medium), having a Moliere radius (approx. 48 mm) consistent with an affordable size, being easy to calibrate, having an electrode structure allowing freedom to select the spatial granularity used to create longitudinal as well as transverse samplings and, finally, being simple to make hermetic, as the signals are picked out at front and back. The complications of the design are the difficult-to-construct, high-precision accordion structure and the reliance on a cryogenic medium (operating temperature 89 K, monitored to 10 mK), requiring three separate cryostats (one for barrel combined with tracker solenoid plus two in the endcaps). The purity of the liquid argon is crucial to uniform response, any electronegative impurities are fatal, the choice of materials is crucial, there are very stringent cleanliness requirements during assembly and impurities must be closely monitored. For instance, radioactive source monitoring of the oxygen contamination is done with a precision less than 0.3 ppm. The readout electronics must respond to the initial, impulse ionization signal, because electron drift over the 2 mm liquid argon gap takes 400 ns, much longer than the interval between bunch crossings. The full barrel detector is constructed of two half-barrels, each comprising 15 modules. Each barrel module is 3.2 m long, spans a radius between 1.5 and 2.0 m and has 1024 accordion absorbers, with the width of the gap between electrode and absorber maintained to better than 0.1 mm over the full length. To keep uniform response with depth the lead thickness varies with $\eta$ and the accordion angle increases with radius. A separate, first radial compartment acts as a ‘pre-shower’ to help compensate energy loss in the material of the tracker and the solenoid. Crucial to the uniformity of response was the consistency of tooling, process, cleanliness, quality assurance and testing procedures between the manufacturing sites.

In 2004, the two half barrels were assembled inside the cryostat shared with the central solenoid, 14 years after the original accordion calorimeter concept was described and the first prototype module constructed.
The development of lead tungstate scintillating crystals for the CMS electromagnetic calorimeter \cite{6,20} started with a few small, yellowish samples in 1993. Over the next 5 years, dramatic advances in materials purity, stoichiometry, quality control and radiation hardness paved the way for production of around 76 000 crystals (weighing 91 tonnes in total). Figure 6 illustrates individual crystal quality control at CERN. Large matrices of crystals were also tested and irradiated in particle beams to confirm the required properties and to establish calibration. Lead tungstate crystals provide a homogeneous medium with high density, very small Molière radius (approx. 22 mm) and radiation length (approx. 9 mm), leading to a very compact calorimeter, well adapted to the CMS design, which encloses the calorimeters within the solenoid. The light output is fast with no slow component, suitable for the high bunch-crossing rate at LHC. Complications arise, because the crystal is brittle, thus making a mechanical structure with minimal gap between crystals challenging. The light-yield is low, and highly temperature-dependent, as is the response of the APDs used to detect the scintillation light. Thus the temperature must be stabilized to 0.1°C, through a feedback mechanism in the cooling system. Mass-production quality control, using dedicated, custom-built, measuring machines, proved crucial for radiation hardness and light collection uniformity. Establishing intercrystal calibration required lengthy cosmic ray and particle beam tests of many modules, and the evolution of the light output with irradiation had to be continuously followed using sophisticated (laser-based) light output monitoring having a precision of 0.1%.

In 1993, a few APDs also existed, but only as small area devices with poor radiation tolerance. The radiation-tolerant, large-area APD phototransducers (two per crystal) used in the transverse magnetic field of the CMS barrel had to be custom-developed, passing through 100 prototypes before commitment to mass-production of 140 000 units. In the endcaps, radiation dose is too severe for the APDs, but vacuum phototriodes, which can tolerate the axial field, proved suitable for the task, again after several years of prototyping. The readout electronics was completely redesigned in 2003–2004 in the emerging 0.25 μm feature size, to help meet the imposed cost ceiling and performance criteria. A total of 512 000 radiation-hard, on-board ASICs were constructed with an equivalent noise level of just 40 MeV per channel. The 36th and final barrel supermodule was installed in CMS in July 2007, 14 years after the original conception and tests, an identical development period as for the equivalent ATLAS electromagnetic calorimetry system.

7. Tracking systems, silicon sensors

No review of the instrumentation challenges faced by the builders of ATLAS and CMS would be complete without at least a brief mention of the tracking systems, which provide efficient detection of charged particle trajectories within the respective solenoids over the range \(|\eta|\) less
than 2.5. These rely partly or wholly on custom-developed silicon sensors, which, along with their specialized on-board electronics, have been qualified to face the LHC challenges of particle rate and radiation level, typically 10–100 kGy yr\(^{-1}\) (but up to 1000 kGy for the innermost layers) and \(3 \times 10^{14}\) neutron-equivalent cm\(^{-2}\) yr\(^{-1}\). The sensors are p–n junction diodes, operated at reverse bias to create a sensitive region depleted of all mobile charge carriers except those created by ionizing particles, which are swept by the electric field towards detection electrodes [21–23]. The pixel detectors, not much bigger than a shoebox, which surround the beampipe, typically have around 70 million silicon sensor elements (each element typically 100–200 \(\mu\)m across) arranged in layers starting just 4–5 cm from the collision point. Their information is crucial to the algorithms used to reconstruct charged particle tracks for momentum measurement and to identify and disentangle multiple p–p collisions and the tracks associated with each collision. They are also critical for picking out the decay points of short-lived particles, slightly displaced from the p–p collision point.

Moving outwards in the tracking system, the pixels are replaced by strips a few times 10 cm long and of the order of 100 \(\mu\)m wide. The CMS microstrip tracker, for instance, has an area of 210 m\(^2\), with 10 million strips arranged in 15 000 modules, with a total of 25 million wire bonds connecting sensors and electronics. This represents two orders of magnitude increase in scale compared with the largest silicon sensor systems in use in the early 1990s. Thanks to automated assembly techniques and the development of radiation-tolerant 0.25 \(\mu\)m electronics, this extraordinary instrument was completed in just 7 years.

8. Assembly of ATLAS and CMS

The assembly techniques of ATLAS and CMS differ dramatically, resulting from the different magnet system solutions; however, the same challenges—integration, logistics, alignment, schedule and safety, along with daunting cable, fibre and pipework installation tasks—were faced in both cases.

In the ATLAS case, pre-assembly of most subdetectors was done on the surface, but the major structures, particularly the barrel toroid system and its supporting structures, were assembled underground once the experimental cavern was ready. In the CMS case, highly modular construction allowed assembly and system-test of the magnet, yoke and parts of all subsystems on the surface while cavern engineering progressed underground. The experiment was then lowered in 15 large pieces (weighing 300 to 2000 tonnes) using the ‘strand-jacking’ technique and a custom-designed gantry crane spanning the roof of the surface assembly building.

9. Selecting interesting collisions

Although this review has focused on detector design and construction, the problem of selecting the few hundred potentially interesting p–p interactions per second, from the thousand million per second which occur within ATLAS or CMS, is one of the fundamental challenges for detector systems at LHC. The process amounts to making successively more complex decisions on successively lower data rates. Custom-integrated circuits, as well as field-programmable gate arrays, at the limits of commercially available technology, are used to make the initial selection, whereas the data of each bunch crossing are stored for a few millionths of a second in memories built into the detector. ATLAS and CMS differ in the number of levels of pre-selection, but both ultimately send between 1 and 100 kHz of events to processor farms in the surface installation above the experiments, using very high-speed telecommunications switches. The processor farms complete an online selection process where 99.999% of p–p collisions are rejected and discarded forever. The retained data (flowing at a rate of about 2 GB s\(^{-1}\) from CERN) are stored and analysed using the worldwide LHC computing GRID [24], comprising 150 computing centres on five continents, which can deploy 300 000 processor cores, more than 200 petabytes of disk storage and a similar amount of tape storage. This completes the achievement of a remarkable 1 in \(10^{13}\)
selection, requiring a processing power of teraflops, designed to extract observable Higgs decays occurring at millihertz rates from p–p bunch collisions happening at a 40 MHz rate.

10. Conclusion

ATLAS and CMS both met their detector performance targets, as witnessed by their reconstruction of the bulk of the previous 40 years of painstaking discoveries in particle physics in just a few months. This was a crucial step in the verification of the apparatus before searching for new phenomena. It should not, however, be overlooked that completion, commissioning and pre-calibration were significantly assisted by the unexpected, additional year before collisions started, resulting from an LHC cryogenic incident in autumn 2008. Ultimately, the LHC delivered the required collisions (about 2000 trillion p–p interactions per experiment), and both ATLAS and CMS were tuned up and ready for discovery physics. Building the instruments that revealed the Higgs boson \[25–28\] set unprecedented challenges to the worldwide particle physics community. It was necessary to design and construct the fast, radiation-tolerant, highly granular detectors needed to make precise and consistent measurements over a large dynamic range. This required the development of several specific new detector technologies (with a typical timeframe from concept to installation of about 15 years), deployment of new and existing technologies in apparatus of unprecedented scale, plus the exploitation of cutting edge developments in microelectronics and computing. Importantly, the particle physics collaboration model had to be extended to global teams of more than 2000 scientists. The common motivation of scientific curiosity was the key factor in making this work. Thanks to these teams, to those that built and operated the LHC accelerator and to the key role of CERN as host laboratory, the whole enterprise started with extraordinary success. These are still early days; new challenges and hopefully new discoveries lie ahead!

Acknowledgements. Most of the content of this review is based upon the existing and extensive documentation (\[26,28–32\], for example) but I thank numerous co-workers who provided advice and material. I am particularly grateful to David Barney, Benoit Cure, Patrick Fassnacht, Ian Fisk, Frans Meijers, Steve Myers, Tejinder Virdee and Pippa Wells.

References


