Beyond the standard model with the LHC

John Ellis[1]

Whether or not the Large Hadron Collider reveals the long-awaited Higgs particle, it is likely to lead to discoveries that add to, or challenge, the standard model of particle physics. Data produced will be pored over for any evidence of supersymmetric partners for the existing denizens of the particle 'zoo' and for the curled-up extra dimensions demanded by string theory. There might also be clues as to why matter dominates over antimatter in the Universe, and as to the nature of the Universe's dark matter.

The unparalleled high energy of the Large Hadron Collider (LHC), with its 7 TeV per beam and its enormously high collision rate that should reach a billion collisions per second, makes it a microscope able to explore the inner structure of matter on a scale that is an order of magnitude smaller than previously achieved. Results at the energies and distances explored so far led physicists to successfully describe matter using the standard model of particle physics[1–3]. But this description is incomplete, and the standard model raises, but leaves unanswered, many fundamental questions. Explanations are needed for the origin of particle masses and the small differences seen in the properties of matter and antimatter, as well as to establish whether fundamental interactions can be unified. Moreover, the standard model has no explanation for some of the basic puzzles of cosmology, such as the origin of matter and the nature of the Universe's dark matter and dark energy. There are high hopes that the LHC will help resolve at least some of these basic issues in cosmology and in physics beyond the standard model[4].

Theoretical calculations made using the standard model agree well with data collected at lower-energy accelerators, such as at CERN's Large Electron–Positron (LEP) accelerator in the 1990s and, more recently, at the Tevatron proton–antiproton collider at Fermilab (Batavia, Illinois)[5]. Data collected at LEP agreed with the standard model at the per-mille level, and recent measurements of the masses of the intermediate vector boson W (ref. 6) and the top quark[7] agree well with standard-model predictions. But the theoretical calculations are valid only with an ingredient that has not yet been observed — the notorious Higgs boson. Without this missing ingredient, the calculations yield incomprehensible, infinite results[8,9]. The agreement of the data with the calculations implies not only that the Higgs boson (or something equivalent) must exist, but also suggests that its mass should be well within the reach of the LHC[5].

In this review, I discuss the likelihood of finding the Higgs boson and what other physics beyond the standard model the accelerator might reveal.

Searching for symmetry breaking

Why should the Higgs boson exist, and are there any alternatives? In the underlying equations of the standard model, none of the elementary particles seems to have mass. In the real world, however, only the photon and gluon, the carriers of the electromagnetic and strong nuclear interactions, are massless. All the other elementary particles are massive, with the W and Z bosons, intermediaries of the weak nuclear interaction, and the top quark weighing as much as decent-sized nuclei. The underlying symmetry between the different particles of the standard model must be broken so that some may acquire masses.

There are two ways to break the symmetry of the standard model. The preferred way is to respect the symmetry of the underlying equations, in which the massless photon and the massive W and Z bosons appear in the same way, but look for an asymmetric solution, much as the reader and writer are lopsided solutions of the symmetric equations of electromagnetism. According to this approach to the standard model, symmetry is thought to be already broken in the lowest-energy state, the so-called vacuum. This 'spontaneous' symmetry breaking is ascribed to a field that permeates all space, taking a specific value that can be calculated from the underlying equations, but with a random orientation in the internal 'space' of particles that breaks the underlying symmetry. This mechanism, which was suggested by Peter Higgs[10] and independently by Robert Brout and François Englert[11], forces some particles, such as the photon, to remain massless, but gives masses to others in proportion to their coupling to this vacuum field (Fig. 1).

In the same way that the electromagnetic field has a quantum particle associated with it, the photon, this vacuum field would also have an associated quantum particle, the Higgs boson. Experiments at LEP seemed at one time to have found a hint of its existence[12]. In the end, however, these searches were unsuccessful and told us only that any Higgs boson must weigh at least 114 GeV (ref. 13). If its mass is less than about 200 GeV, researchers using the Tevatron may find some evidence for it before the LHC comes into operation[14].

The large experiments, ATLAS[15] and CMS[16], at the LHC will be looking for the Higgs boson in several ways (Fig. 2). The Higgs boson is predicted to be unstable and decay into other particles, such as photons, bottom quarks, tau leptons, W or Z bosons. It may well be necessary to combine several different decay modes to uncover a convincing signal. The LHC experiments should be able to find the Higgs boson even if it weighs as much as 1 TeV, and there are high expectations that it could be found during the first couple of years of LHC operation. Its discovery would set the seal on the success of the standard model.

Higgs or bust?

With the impending confirmation or refutation of the Higgs hypothesis, many theorists are getting cold feet. Some are beginning to support alternative scenarios that go beyond the standard model[17]. One popular suggestion is that the Higgs boson might not be an 'elementary' particle in the same sense as the quarks, leptons and the photon, but instead might be composed of simpler constituents[18]. This model would be analogous to the Bardeen–Cooper–Schrieffer (BCS) theory of superconductivity, in which a photon acquires an effective mass by interacting with 'Cooper pairs' of electrons. In this analogy, the W and Z bosons would 'eat' tightly bound pairs of novel strongly interacting fermions

[1]Theory Division, Physics Department, CERN, CH-1211 Geneva 23, Switzerland.

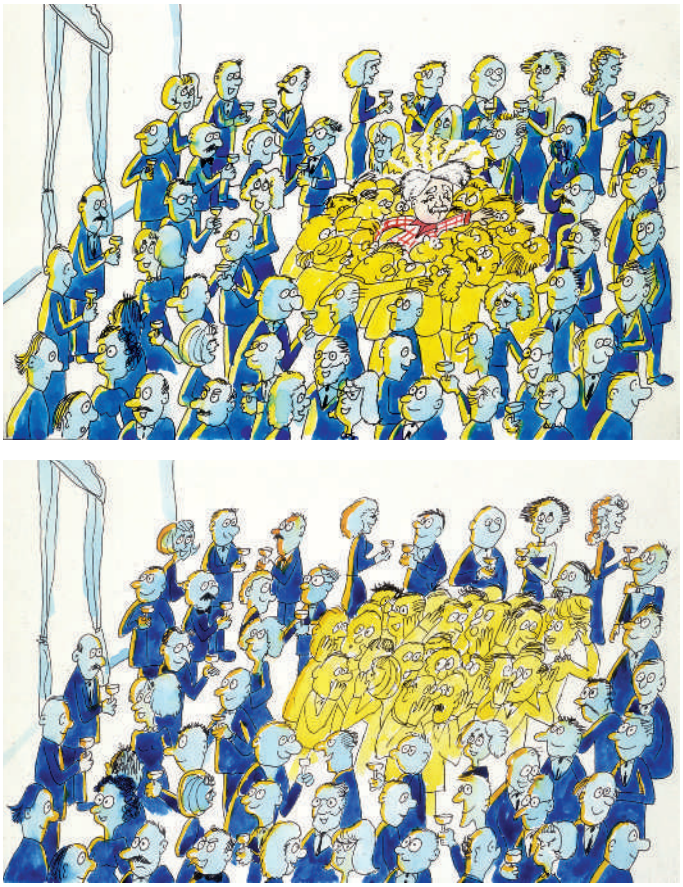


Figure 1 | Picturing the Higgs field. The behaviour of physicists in a crowded social event at a conference is an analogy for the Higgs mechanism, as proposed by David Miller (University College London). The physicists represent a non-trivial medium permeating space. In the upper panel, the physicists cluster around a famous scientist who enters the room, slowing the scientist's progress. In much the same way, a particle passing through the Higgs–Brout–Englert field slows down and acquires a mass. In the lower panel, a rumour propagates. This is an excitation of the medium — the group of physicists — itself, forming a body with a large mass; this is analogous to the formation of a Higgs boson. Figure reproduced with permission from CERN.

rather than an elementary Higgs field. It seems rather difficult to reconcile this composite alternative with the accurate low-energy data from LEP[5], but some enthusiasts are still pursuing this possibility. Alternatively, it has been suggested that the Higgs boson is indeed elementary, but is supplemented by some additional physics — for example, being supersymmetric (discussed later).

The most radical alternative to the Higgs hypothesis exploits the second way of breaking the standard model's symmetry. It postulates that, although the underlying equations are symmetric, their solution is subject to boundary conditions that break that symmetry. What boundary would that be, given that space is apparently infinite (or at least very large compared to the scale of particle physics)? The answer is that there might be additional, very small dimensions of space with edges where the symmetry may be broken[19]. Such models would have no Higgs boson, and are difficult to reconcile with the data already acquired that seem to require a relatively light Higgs boson.

Theorists are amusing themselves discussing which would be worse: to discover a Higgs boson with exactly the properties predicted in the standard model or to discover that there is no Higgs boson. The former would be a vindication of theory, but would teach us little new. The latter would upset the entire basis of the standard model. The absence of a Higgs boson would be exciting for particle physicists, but it might not be so funny to explain to the politicians who have funded the LHC mainly to discover this particle. Whichever option nature chooses, the good news is that the LHC will provide us with a clear-cut experimental answer and end the speculation.

The hierarchy problem

Resolving the Higgs question will set the seal on the standard model, but, as I mentioned at the beginning, there are plenty of reasons to expect other physics beyond the standard model to be discovered (Fig. 3). Specifically, there are good reasons to expect other discoveries at the TeV energy scale, within reach of experiments at the LHC. Many would consider this to be the primary motivation for the leap into the unknown that the LHC represents.

For example, it is generally thought that the elementary Higgs boson of the standard model cannot exist in isolation. Specifically, difficulties arise when one calculates quantum corrections to the mass of the Higgs boson owing to the exchanges of virtual particles (see, for example, ref. 20). Not only are these corrections infinite in the standard model, but, if the usual procedure of controlling them by cutting the theory off at some high energy or short distance is adopted, the net result depends on the square of the cut-off scale. This implies that, if the standard model were embedded in some more complete theory that kicks in at high energy — such as a grand unified theory of the particle interactions or a quantum theory of gravity — the mass of the Higgs boson would be sensitive to the details of this high-energy theory. This would make it difficult to understand why the Higgs boson has a (relatively) low mass. It would also, by extension, make it difficult to explain why the energy scale of the weak interactions — as reflected in the masses of the W and Z bosons — is so much smaller than that of unification or quantum gravity.

One might be tempted simply to wish away this 'hierarchy problem' by postulating that the underlying parameters of the theory are tuned finely, so that the net value of the Higgs boson mass obtained after adding in the quantum corrections is unnaturally small as the result of some sneaky cancellation. But it would surely be more satisfactory either to abolish the extreme sensitivity to the quantum corrections or to cancel them in a systematic manner. Indeed, this has been one of the reasons for believing that the Higgs boson is composite. If it is, the Higgs boson would have a finite size, which would cut the pesky quantum corrections off at some relatively low scale. In this case, the LHC might uncover a cornucopia of new particles with masses around this cut-off scale, which should be near 1 TeV. At the very least, the interactions of the W and Z vector bosons would be modified in an observable way.

The supersymmetric solution

An alternative way to get rid of these quantum corrections is provided by supersymmetry[21]. This is an elegant theory that would pair up fermions, such as the quarks and leptons that make up ordinary matter, with bosons, such as the photon, gluons, W and Z that carry forces between the matter particles or even the Higgs itself (Fig. 4). Supersymmetry also seems to be essential for making a consistent quantum theory of gravity based on string theory (of which more later). However, these elegant arguments give no clue as to what energies would be required to observe supersymmetry in nature.

The first argument that supersymmetry might appear near the TeV scale was provided by the hierarchy problem: in a supersymmetric theory, the quantum corrections owing to the pairs of virtual fermions and bosons cancel each other systematically[22], and a low-mass Higgs boson no longer seems unnatural[23]. The residual quantum corrections to the mass of the Higgs boson would be small if differences in mass between supersymmetric partner particles were less than about 1 TeV. Because the fermions and bosons of the standard model do not pair up with each other in a neat supersymmetric manner, this theory would require each of the standard-model particles to be accompanied by an as-yet unseen supersymmetric partner. It might seem profligate for there to be all these partners, but at least the hypothesis predicts a 'scornucopia' of supersymmetric particles that should weigh less than about 1 TeV and hence could be produced by the LHC[15,16].

In the wake of this hierarchy argument, at least three other reasons have surfaced for thinking that supersymmetric particles weigh about 1 TeV.

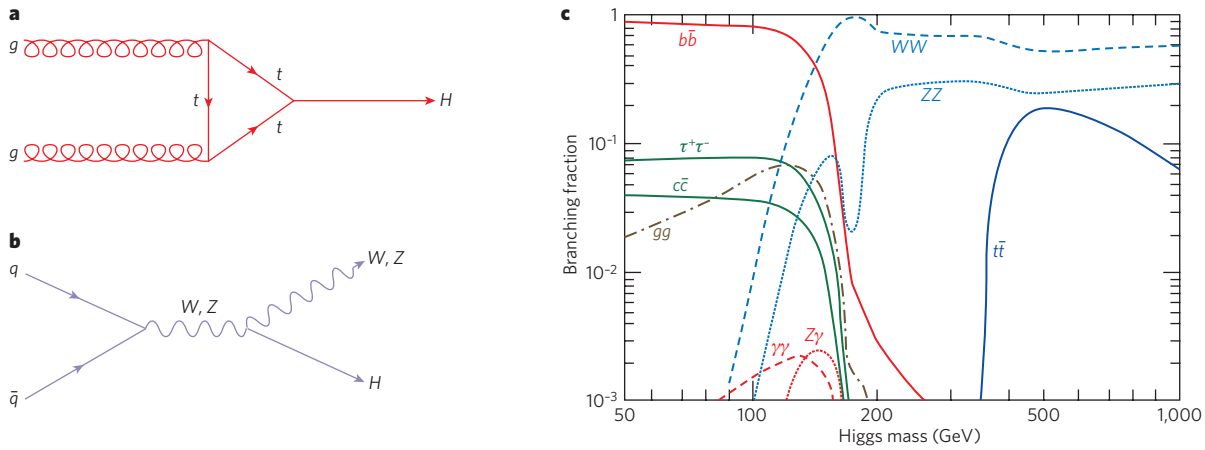


Figure 2 | The Higgs boson at the LHC. A Higgs (H) boson may be produced by a range of interactions, two examples of which are shown here. The first, **a**, is through fusion of gluons (g) from the protons in the LHC beams, through a top (t) quark loop; and the second, **b**, is through a *bremsstrahlung* process, in which a quark (q) and antiquark ($\bar{q}$) annihilate to create a W or Z boson, which may then radiate a Higgs. **c**, The Higgs itself then decays, and it is these decay products that will be caught in a detector. The 'branching fraction' or probability of decay to certain products depends on the (as-yet unknown) mass of the Higgs particle, which is dominated by decay to a bottom–antibottom quark pair at low mass, but by decay to pairs of W bosons at high mass.

The first is that these particles would facilitate the unification of the strong, weak and electromagnetic forces into a simple grand unified theory[24]. Another argument is that a theory with low-energy supersymmetry would predict that the Higgs boson weighs less than about 150 GeV (ref. 25), which is precisely the range favoured indirectly by the present data. The final one is that, in many models, the lightest supersymmetric particle (LSP) is an ideal candidate for the dark matter advocated by astrophysicists and cosmologists.

The LSP is ideal because it is stable when a suitable combination of baryon and lepton numbers is conserved[26], as happens in the minimal supersymmetric extension of the standard model, as well as in simple models of grand unification and neutrino masses. In this case, LSPs would be left over as relics from early in the Big Bang, and calculations of their abundance yield a density of dark matter in the range favoured by astrophysics and cosmology if the LSP weighs at most a few hundred GeV, probably putting it within reach of the LHC[27].

Supersymmetry could be a bonanza for the LHC, with many types of supersymmetric particle being discovered. In many models, the LHC would produce pairs of gluinos (the supersymmetric partners of the gluons) or squarks (the supersymmetric partners of the quarks) that would subsequently decay through various intermediate supersymmetric particles. Finally, each of these pairs of particles would yield a pair of LSPs that interact only weakly and hence carry energy away invisibly. In favourable cases, the masses of several intermediate particles could be reconstructed this way. It might even be possible to use these measurements to calculate what the supersymmetric dark-matter density should be, so as to compare the result with the astrophysical estimates[28].

Into extra dimensions?

Postulating a composite Higgs boson or supersymmetry are not the only strategies that have been proposed for dealing with the hierarchy problem. Another suggestion is that there are additional dimensions of space[29]. Clearly, space is three-dimensional on the scales that we know so far, but the idea that there are additional dimensions curled up so small that they are invisible has been in the air since it was first proposed by Kaluza and Klein over 80 years ago. This idea has gained ground in recent years with the realization that string theory predicts the existence of extra dimensions of space[30].

According to string theory, elementary particles are not idealized points of euclidean geometry, but are objects extended along one dimension (a string) or are membranes with more dimensions[31]. For the quantum theory of strings to be consistent, particles have to move in a space with more than the usual three dimensions. Initially, it was thought that these extra dimensions would be curled up on scales that might be as small as the Planck length of around 10^{-33} cm. But more recently, it was realized that at least some of these new dimensions might be much larger and possibly have consequences observable at the LHC.

One of the possibilities offered by these speculations is that gravity is strong when these extra dimensions appear, possibly at energies close to 1 TeV. Under this condition, according to some variants of string theory, microscopic black holes might be produced by the LHC[32]. These would be short-lived, decaying rapidly through thermal (Hawking) radiation. Measurements of this radiation would offer a unique laboratory window on the mysteries of quantum gravity. The microscopic black holes would emit energetic photons, leptons, quarks and neutrinos, providing distinctive experimental signatures. In particular, the neutrinos they emit would carry away more invisible energy than LSPs would in the supersymmetric models discussed previously[33].

Although microscopic black holes would be the most dramatic sign of large extra dimensions, they are not the only sign of such theories that might be visible at the LHC. If the extra dimensions are curled up on a sufficiently large scale, the ATLAS and CMS projects might be able to see Kaluza–Klein excitations of standard-model particles, or even of the graviton, the mediator particle of gravity. Indeed, the spectroscopy of some extra-dimensional theories might be as rich as that of supersymmetry[34]. If so, how do we tell which cornucopia the LHC is uncovering? There are significant differences in the relationship between, for example, the masses of the partners of quarks and leptons in supersymmetric theories and in theories with large extra dimensions. Moreover, the spins of the Kaluza–Klein excitations would be the same as those of their standard-model progenitors, whereas the spins of the supersymmetric partners

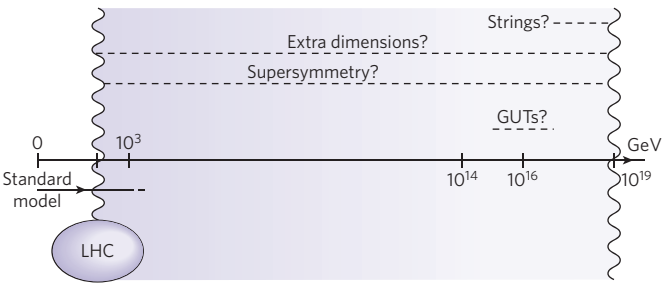


Figure 3 | Physics beyond the TeV scale. The standard model has been well tested up to around the 100-GeV mass scale. The LHC will test beyond this, to the crucial 1,000-GeV level, the TeV scale, at which hints of new physics, such as supersymmetry and extra dimensions, may emerge. String theory or grand unified theories (GUTs) inhabit much higher energy scales, approaching 10^{19} GeV, which is called the Planck scale.

	Known particles of the standard model	Postulated supersymmetric partners or 'sparticles'	
Half-integer spin	Electron	Selectron	Integer spin
	Neutrino	Sneutrino	
	Top quark	Stop	
Integer spin	Gluon	Gluino	Half-integer spin
	Photon	Photino	

Figure 4 | Examples of supersymmetric partners. Supersymmetry is a symmetry drawn between fermions (with half-integer spin) and bosons (with integer spin). It postulates that, for each fermion, there exists a bosonic partner — such as the supersymmetric electron, or 'selectron', which partners the electron. Similarly, each boson is thought to have a fermionic superpartner, which for the gluon is the 'gluino'.

would be different. These underlying differences translate into characteristic differences in the spectra of decay products in the two classes of model and into distinctive correlations between them[35].

It is amusing that, in some theories with extra dimensions, the lightest Kaluza–Klein particle (LKP) might be stable[36], rather like the LSP in supersymmetric models. In this case, the LKP would be another candidate for astrophysical dark matter. Thus, there is more than one way in which LHC physics beyond the standard model might explain the origin of dark matter: fortunately, the tools seem to be available for distinguishing between them.

The matter–antimatter conundrum

Will the LHC explain the origin of conventional matter? As was first pointed out by the Russian physicist Andrei Sakharov[37], particle physics can explain the origin of matter in the Universe in terms of small differences in the properties of matter and antimatter, such as those discovered in the decays of *K* and *B* mesons. Present experimental data accord well with the matter–antimatter differences allowed by the standard model. However, by themselves, these differences in the properties of matter and antimatter would be insufficient to generate the matter seen in the Universe. It is possible that the deficit will be explained by new physics at the TeV scale revealed by the LHC. For example, supersymmetry allows many more possibilities for differences between the properties of matter and antimatter than are possible in the standard model[38]; some of these differences might explain the amount of matter in the Universe.

This provides one of the motivations for the LHCb experiment[39], which is dedicated to probing the differences between matter and antimatter, notably looking for discrepancies with the standard model (Box 1). In particular, LHCb has unique capabilities for probing the decays of mesons containing both bottom and strange quarks, the constituents of the *B* and *K* mesons probed in other experiments investigating matter–antimatter differences. There are many other ways to explore the physics of matter and antimatter, and the ATLAS and CMS experiments will also contribute to them, in particular by searching for rare decays of mesons containing bottom quarks.

If these experiments detect any new particles beyond the standard model at the TeV scale, questions will immediately arise as to whether this new physics distinguishes between matter and antimatter, and whether or not this new physics explains the origin of matter in the Universe. For example, if the Higgs boson is discovered at the LHC, are its couplings to matter and antimatter the same? If supersymmetry is discovered at the LHC, do supersymmetric 'sparticles' and 'antisparticles' behave in the same way? There are many models in which matter–antimatter differences in the Higgs or sparticle sector are responsible for the origin of the matter in the Universe.

Into the future

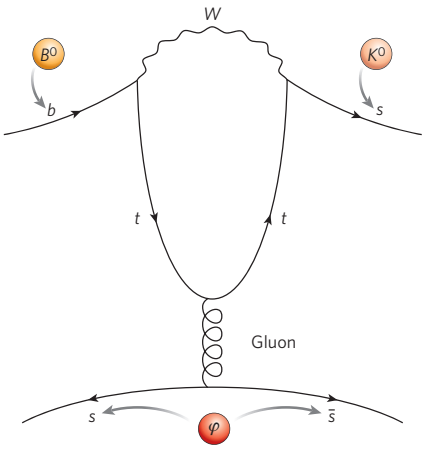
According to present plans, the first full-energy collisions of the LHC will take place in 2008, although it will take some time for the accelerator to build up to its designed nominal collision rate. There are hopes, however, that in its first couple of years of operation, it will already start to provide crucial information on physics beyond the standard model, for example by discovering the Higgs boson — or other new particles such as those

Box 1 | Penguin hunting at LHCb

Why does matter dominate over antimatter in the Universe, considering that both were thought to be created in equal quantities in the Big Bang? Part of the explanation is that some interactions between particles take place at different rates when two fundamental symmetries of the quantum field theory that underlie them are simultaneously reversed. There are two symmetries involved: charge conjugation, C; and parity symmetry, P. Charge conjugation turns particles into their antiparticles by reversing internal properties such as electric charge. By contrast, parity symmetry flips external particle properties such as spin, similar to looking at an interaction in a mirror.

CP violation was first discovered experimentally[45] in decays of *K* mesons, which contain a strange quark in addition to an up or down quark. Later theoretical work showed[46] that CP violation would occur naturally in interactions mediated by the weak nuclear force in the standard model with three quark generations. (At the time, particles from only two were known.) The degree of violation is, however, insufficient to explain the Universe's matter–antimatter imbalance.

The subsequent discovery of the third quark generation, formed of the bottom (*b*) and top (*t*) quarks, vindicated the model. A plethora of



experiments has since confirmed CP violation, indirectly and directly, in decay channels of *B* mesons (those containing bottom quarks), where the effect is expected to be particularly large. These experiments notably include two specially constructed '*B* factories', the Belle detector at KEK in Japan, and BaBar at the Stanford Linear Accelerator Center (SLAC), in California, which have delivered a series of more precise values for the parameters of CP violation since 2001.

The LHCb experiment is LHC's dedicated CP-violation detector. It is a 20-m-long spectrometer with a conical detection volume expanding in radius along the beam axis. It is attuned to detecting the distinctive signature of *B* decays — charged particles with high transverse momenta originating from a vertex significantly displaced from the interaction point of the proton beams.

To maximize the probability of only a single *B* interaction per beam crossing, the LHC beams are defocused slightly to a luminosity of around $2.5 \times 10^{32}\,cm^{-2}s^{-1}$, below the LHC's nominal value of $10^{34}\,cm^{-2}s^{-1}$. The implied collision rates and the high energy of the LHC beams should allow CP-violation parameters to be more tightly constrained and perhaps also provide a glimpse of physics beyond the standard model.

Such physics could manifest itself, in particular, in 'penguin' processes such as $B^0 \rightarrow K^0 \varphi$ (see page 270) in which the decay of a highly energetic *B* meson takes place, legitimately according to the rules of the weak interaction, through an intermediate loop of massive particles such as a top quark and a *W* boson (see figure). Does the degree of CP violation in such a process differ significantly from that found, for example, in the decay $B^0 \rightarrow K^0 J/\Psi$, which does not include a penguin loop? If so, that could be an indication of new physics participating in the penguin loop — such as the involvement of supersymmetric particles.

predicted by supersymmetry, if they are not too heavy[40]. Continued running of the LHC at its nominal luminosity would enable many properties of the Higgs boson to be verified, for example by providing measurements of its couplings to some other particles and checking whether these are proportional to the particles' masses. This period should also enable the properties of any other newly discovered particles to be checked, such as establishing whether their spins are the same as those of their standard-model counterparts or are different.

What might be possible using the LHC after these planned phases of exploitation? One possibility is to add new components to the existing ATLAS and CMS detectors that would provide new ways to study the Higgs boson. For example, new components close to the beams several hundred metres from the interaction points might be able to detect rare proton–proton collisions that produce nothing except a single isolated Higgs boson[41]. Another possibility is that supersymmetric or other new particles might show up in unexpected ways. For example, in some supersymmetric scenarios there would be a metastable charged particle that would have quite distinctive experimental signatures[42], and it might be interesting to devise new detectors to explore this possibility.

It might also be possible to increase the LHC collision rate significantly beyond the nominal value. This possibility would be particularly interesting if, for example, the initial runs of the LHC discover new physics with a very low production rate, perhaps because it has a high energy threshold. Increasing the LHC collision rate might be possible by redesigning the collision points using new magnet technologies; it would also require replacing at least some of CERN's lower-energy accelerators, such as the low-energy linear proton accelerator and the Proton Synchrotron, so as to feed more intense beams into the LHC[43]. Technical options for increasing the LHC collision rate are now being evaluated, so that they can be considered when the first experimental results from the initial LHC runs become available, some time around 2010.

Exploitation of the LHC and the study of possible upgrade options are among the highest priorities for European particle physics and were decided upon at a special meeting of the CERN Council in Lisbon in July 2006 (ref. 44). Possible future accelerators were also considered, such as a linear electron–positron collider or a neutrino factory. The priorities for these options will surely depend on the nature and energy scale of whatever new physics beyond the standard model the LHC reveals, as well as on developments in other areas such as neutrino physics. A central element in the European strategy for particle physics is the need to review advances in particle physics in the coming years, and in particular to review the implications of any LHC discoveries at the end of this decade.

Particle physics stands on the brink of a new era. Research using the LHC will make the first exploration of physics in the TeV energy range. There are good reasons to hope that the LHC will find new physics beyond the standard model, but no guarantees. The most one can say for now is that the LHC has the potential to revolutionize particle physics, and that in a few years' time we should know what course this revolution will take. Will there be a Higgs boson, or not? Will space reveal new properties at small distances, such as extra dimensions or supersymmetry? Will experiments at the LHC cast light on some fundamental cosmological questions, such as the origin of matter or the nature of dark matter? Whatever the answers to these questions might be or whatever surprises the LHC might spring, it will surely set the agenda for the next steps in particle physics.■

1. Glashow, S. Partial symmetries of weak interactions. *Nucl. Phys.* **22,** 579–588 (1961).
2. Weinberg, S. A model of leptons. *Phys. Rev. Lett.* **19,** 1264–1266 (1967).
3. Salam, A. in *Elementary Particle Physics: Relativistic Groups and Analyticity* (Nobel Symposium No. 8) (ed. Svartholm, N.) 367 (Almqvist and Wiksills, Stockholm, 1968).
4. Ellis, J. Physics at LHC. *Acta Phys. Polon.* **B38,** 1071–1092; preprint at <http://arxiv.org/abs/hep-ph/0611237> (2007).
5. LEP Electroweak Working Group. *LEP Electroweak Working Group* <http://lepewwg.web.cern.ch/LEPEWWG/> (2007).
6. Tevatron Electroweak Working Group. *Tevatron Electroweak Working Group W/Z Subgroup* <http://tevewwg.fnal.gov/wz/> (2007).
7. Tevatron Electroweak Working Group. A combination of CDF and D0 results on the mass of the top quark. Preprint at <http://arxiv.org/abs/hep-ex/0703034> (2007).
8. 't Hooft, G. Renormalizable Lagrangians for massive Yang–Mills fields. *Nucl. Phys. B* **35,** 167–188 (1971).
9. 't Hooft, G. & Veltman, M. Regularization and renormalization of gauge fields. *Nucl. Phys. B* **44,** 189–213 (1972).
10. Higgs, P. W. Broken symmetries and the masses of gauge bosons. *Phys. Rev. Lett.* **13,** 508–509 (1964).
11. Englert, F. & Brout, R. Broken symmetry and the mass of gauge vector mesons. *Phys. Rev. Lett.* **13,** 321–322 (1964).
12. Barate, R. *et al.* (ALEPH Collaboration) Observation of an excess in the search for the standard model Higgs boson at ALEPH. *Phys. Lett. B* **495,** 1–17 (2000).
13. LEP Higgs Working Group. *LEP Higgs Working Group* <http://lephiggs.web.cern.ch/LEPHIGGS/www/Welcome.html> (2007).
14. Cavalli, D. *et al.* The Higgs working group: summary report. *Proc. Workshop on Physics at TeV Colliders* (Les Houches, 2001).
15. ATLAS Collaboration. *Detector and Physics Performance Technical Design Report.* <http://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/TDR/access.html> (2007).
16. CMS Collaboration. *CMS Physics: Technical Design Report* (ed. De Roeck, A.) CERN-LHCC-2006-021 (2006).
17. Ellis, J. Summary of the International Conference on High-Energy Physics, Beijing, China, August 2004. *Int. J. Mod. Phys.* **A20**, 5297 (2005).
18. Farhi, E. & Susskind, L. Technicolor. *Phys. Reports* **74,** 277–321 (1981).
19. Csaki, C., Grojean, C. Pilo, L. & Terning, J. Towards a realistic model of Higgsless electroweak symmetry breaking. *Phys. Rev. Lett.* **92,** 101802 (2004).
20. 't Hooft, G. Naturalness, chiral symmetry, and spontaneous chiral symmetry breaking. Under the spell of the gauge principle. (eds 't Hooft, G. *et al.*) 352–374 (World Scientific, Singapore, 1994).
21. Wess, J. & Zumino, B. A Lagrangian model invariant under supergauge transformations. *Phys. Lett. B* **49,** 52–54 (1974).
22. Ferrara, S., Iliopoulos, J. & Zumino, B. Supergauge invariance and the Gell-Mann–Low eigenvalue. *Nucl. Phys. B* **77,** 413–419 (1974).
23. Witten, E. Mass hierarchies in supersymmetric theories. *Phys. Lett. B* **105,** 267–271 (1981).
24. Ellis, J., Kelley, S. & Nanopoulos, D. *Phys. Lett. B* **249,** 441–448 (1990).
25. Ellis, J., Ridolfi, G. & Zwirner, F. Higgs boson properties in the standard model and its supersymmetric extensions. Preprint at <http://arxiv.org/pdf/hep-ph/0702114> (2007).
26. Fayet, P. in *Unification of the Fundamental Particle Interactions* (eds Ferrara, S., Ellis, J.& van Nieuwenhuizen, P.) 587 (Plenum, New York, 1980).
27. Ellis,J., Hagelin, J., Nanopoulos, D., Olive, K. & Srednicki, M. Supersymmetric relics from the Big Bang. *Nucl. Phys. B* **238,** 453–476 (1984).
28. Battaglia, M. *et al.* Updated post-WMAP benchmarks for supersymmetry. *Eur. Phys. J.* **C33,** 273–296 (2004).
29. Antoniadis, I. A possible new dimension at a few TeV. *Phys. Lett. B* **246,** 377–384 (1990).
30. Green, M., Schwarz, J. & Witten, E. *Superstring Theory* (Cambridge Univ. Press, Cambridge, 1987).
31. Randall, S. & Sundrum, R. An alternative to compactification. *Phys. Rev. Lett.* **83,** 4690–4693 (1999).
32. Arkani-Hamed, N., Dimopoulos, S. & Dvali, G. The hierarchy problem and new dimensions at a millimeter. *Phys. Lett. B* **429,** 263–272 (1998).
33. Harris, C. M. *et al.* Exploring higher dimensional black holes at the Large Hadron Collider. *JHEP* **0505,** 053 (2005).
34. Cembranos, J., Feng, J. Rajaraman, A. & Takayama, F. Exotic collider signals from the complete phase diagram of minimal universal extra dimensions. *Phys. Rev. D* **75,** 036004 (2007).
35. Athanasiou, C., Lester, C. G., Smillie, J. M. & Webber, B. R. Distinguishing spins in decay chains at the Large Hadron Collider. *JHEP* **0608,** 055 (2006).
36. Servant, G. & Tait,T. M. P. Is the lightest Kaluza–Klein particle a viable dark matter candidate? *Nucl. Phys. B* **650,** 391–419 (2003).
37. Sakharov, A. D. Violation of CP invariance, C asymmetry, and baryon asymmetry of the Universe. *Pisma Zh. Eksp. Teor. Fiz.* **5,** 32–35 (1967).
38. Cline, J., Joyce, M. & Kainulainen, K. Supersymmetric electroweak baryogenesis. *JHEP* **0007,** 018 (2000).
39. LHCb Collaboration. *The Large Hadron Collider Beauty Experiment for Precise Measurements of CP Violation and Rare Decays.*<http://lhcb.web.cern.ch/lhcb/> (2007).
40. Blaising, J.-J. *et al. Potential LHC Contributions to Europe's Future Strategy at the High-Energy Frontier.* <http://council-strategygroup.web.cern.ch/council-strategygroup/BB2/contributions/Blaising2.pdf> (2006).
41. FP420 Research and Development Project. FP420 R&D Project <http://www.fp420.com/> (2007).
42. Feng, J. L. & Smith, B. T. Slepton trapping at the CERN Large Hadron Collider and the International Linear Collider. *Phys. Rev. D* **71,** 015004 (2005).
43. Blondel, A. *et al.* Physics opportunities with future proton accelerators at CERN. Preprint at <http://arxiv.org/pdf/hep-ph/0609102> (2006).
44. Strategy Group for European Particle Physics. *CERN Council Strategy Group Home Page* <http://council-strategygroup.web.cern.ch/council-strategygroup/> (2007).
45. Christenson, J. H., Cronin, J. W., Fitch, V. L. & Turley, R. Evidence for the 2π decay of the K_2^0 meson. *Phys. Rev. Lett.* **13,** 138–140 (1964).
46. Kobayashi, M. & Maskawa, T. CP violation in the renormalizable theory of weak interaction. *Prog. Theor. Phys.* **49,** 652–657 (1973).

Author information Reprints and permissions information is available at npg.nature.com/reprintsandpermissions. The author declares no competing financial interests. Correspondence should be addressed to the author (john.ellis@cern.ch).

NATURE|Vol 448|19 July 2007|doi:10.1038/nature06080

The quest for the quark–gluon plasma

Peter Braun-Munzinger[1] & Johanna Stachel[2]

High-energy collisions between heavy nuclei have in the past 20 years provided multiple indications of a deconfined phase of matter that exists at phenomenally high temperatures and pressures. This 'quark–gluon plasma' is thought to have permeated the first microseconds of the Universe. Experiments at the Large Hadron Collider should consolidate the evidence for this exotic medium's existence, and allow its properties to be characterized.

Shortly after the idea of asymptotic freedom — that the interaction between quarks, which is strong at large separations, weakens as the quarks get closer to one another — was introduced by David Gross and Frank Wilczek[1] and David Politzer[2], two groups[3,4] realized independently that it has a fascinating consequence. When temperatures or densities become very high, strongly interacting quarks and gluons become free and transform themselves into a new, deconfined phase of matter, for which the term 'quark–gluon plasma' was coined. We ourselves live at low densities and temperatures, in the normal world of hadronic matter, where quarks and gluons are confined to the size of hadrons. But at its origin, the Universe was a fireball of much higher density and temperature. At times from the electroweak phase transition — some 10 picoseconds after the Big Bang, and lasting for 10 microseconds — it is thought to have taken the form of a quark–gluon plasma. Here, we review the current state of knowledge of this peculiar phase of matter, and outline how the Large Hadron Collider (LHC) should further our understanding of it.

Transition temperature

Various simple estimates lead to a critical temperature for the transition between the familiar, confined hadronic phase of matter and the deconfined, plasma phase of the order of 100 MeV. (In this review, we use the kT unit system, in which all temperatures (T) are multiplied by Boltzmann's constant $k = 8.617 \times 10^{-5}$ eV K^{-1} to express them in more convenient energy units; for reference, a temperature of 100 MeV is somewhat more than 1 trillion kelvin, at 1.16×10^{12} K.) In detailed investigations of hadronic matter, Rolf Hagedorn[5] discovered in the 1960s a limiting temperature for hadronic systems of around the π-meson mass of 140 MeV. It turns out that this temperature is nothing other than the critical temperature for the deconfinement phase transition.

With the advance of solving quantum chromodynamics — the quantum field theory of the strong interaction — on a space-time lattice, more accurate values have become available for the transition temperature. The most readily calculable values are those for a zero net baryon density (that is, no difference between baryon and antibaryon densities). For instance, researchers obtained a temperature[6] of 173 MeV with a systematic error of about 10% for a system involving the two light quark flavours (up and down) and one heavier quark flavour (strange). Very recently, higher values in the vicinity of 190 MeV have been quoted. The reason for the variations is that the lattice energy units have been normalized differently: now, the calculations use quantities that involve heavy bottom quarks, whereas in the past they used the mass of the ρ meson, which contains only the light up and down quarks (see page 270). There is currently a lively debate (see, for example, ref. 7) about the most accurate way to calculate the transition temperature.

Extending lattice quantum chromodynamics into the regime of non-zero net baryon density has met with great technical difficulties. Results[8–10] have become available indicating that the transition temperature drops moderately with increasing density. Going a third of the way from zero net density to the density of atomic nuclei, it drops by 2–3% — not very much. The critical energy density for the phase transition is 0.7 ± 0.2 GeV fm^{-3} (ref. 6). This energy density is about five times that of nuclear matter.

Towards the nuclear fireball

Since the early 1980s, collisions of heavy atomic nuclei at as large energies as possible have been seen as the ideal way to probe these harsh conditions of extremely high temperature and density. To be able to talk about thermodynamic phases, phase transitions, temperatures and so on, the system under consideration must behave like 'matter', not like individual elementary particles or a group of elementary particles. That implies two things. First, the system must consist of a large number of particles (thousands or, better, tens of thousands). Second, it needs to reach local equilibrium, at least approximately, so that variables such as temperature, pressure, energy and entropy density can be defined, and so that thermodynamic relations between those quantities (the equation of state, the speed of sound) can be investigated. This means that the system's lifetime must be significantly larger than the inverse rate of interactions, so that at least a few (order of magnitude five) interactions occur for each constituent, driving the system towards equilibrium.

Collisions of protons (or electrons) produce too few particles to fulfil these conditions. But we know now that collisions between nuclei create enough particles that, if the energy is high enough, they do indeed create a fireball of interacting quarks and gluons above the temperature needed for the phase transition into deconfinement. This fireball quickly expands and cools, until it rehadronizes on passing the deconfinement temperature again. The hugely energetic fireball created in the aftermath of the Big Bang had cooled sufficiently for protons and neutrons (and other confined, but unstable, hadrons) to form after about 10^{-5} s. The fireball created in a nuclear collision in the laboratory contains much less energy, and so is much shorter-lived than that after the Big Bang: after only about 10^{-22} s, the quark–gluon plasma phase of the fireball transforms back to hadronic matter.

Collisions of atomic nuclei have been studied for about 20 years at sufficiently high energies to cross into the deconfined phase. Experimental programmes started simultaneously in late 1986 at the Alternating Gradient Synchrotron at the Brookhaven National Laboratory (BNL) in Upton, New York, and at the Super Proton Synchrotron (SPS)

[1]Gesellschaft für Schwerionenforschung, Planckstr. 1, D 64291, Darmstadt, Germany and Technische Universität Darmstadt. [2]Physikalisches Institut, Universität Heidelberg, Philosophenweg 12, D69120 Heidelberg, Germany.

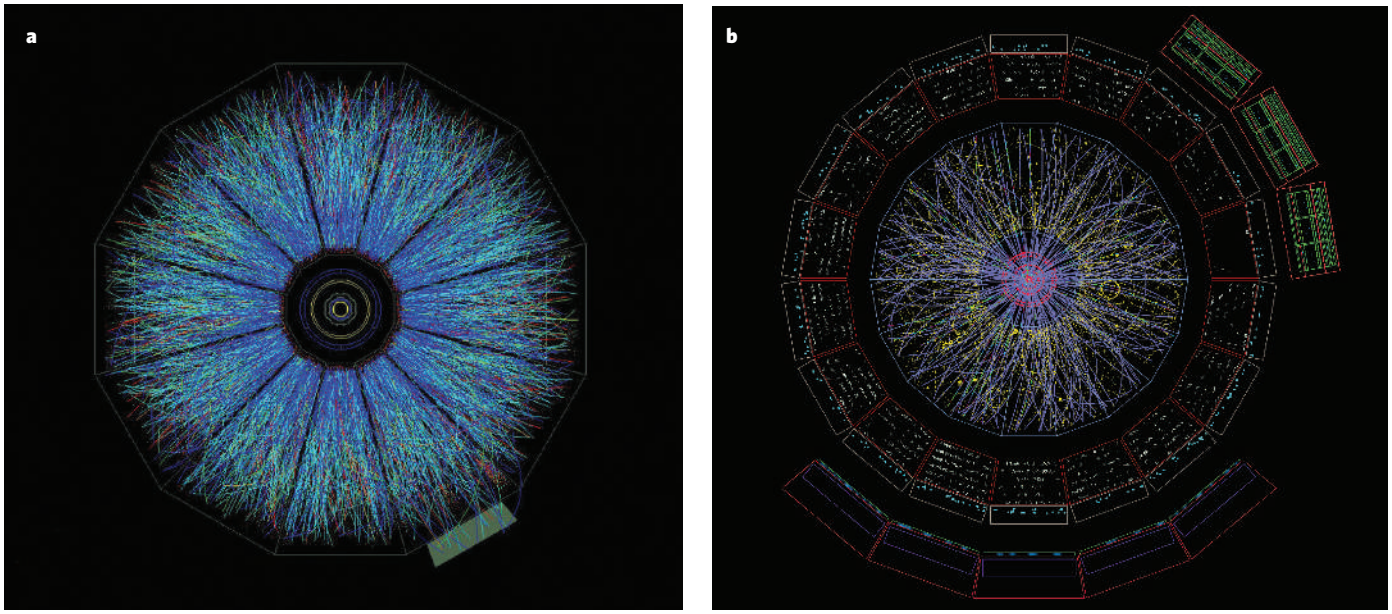


Figure 1 | Fireball remnants. a, Charged particles from a central gold–gold collision at RHIC, recorded by the time projection chamber of the STAR experiment. Colours represent the level of ionization deposited in the detector, with red equating to high values and blue to low values. **b**, A simulation of a central lead–lead collision — just a one-degree slice in polar angle is shown — in the central barrel of the ALICE experiment at the LHC. **Images courtesy of the STAR and ALICE collaborations.**

at CERN. At both facilities, collisions were studied initially with light atomic nuclei (up to silicon and sulphur, with mass numbers of 28 and 32, respectively) and, from the early 1990s, also with heavy nuclei such as gold (mass number 197) and the most abundant isotope of lead (208). For these heavy colliding nuclei, BNL has reached energies in the centre-of-mass system of close to 1,000 GeV, and CERN has reached 3,600 GeV (corresponding to a centre-of-mass energy per colliding nucleon pair, written as $\sqrt{s_{NN}}$, of 4.6 and 17.2 GeV, respectively). At least for the CERN energy regime, enough evidence was gathered to conclude that a new state of matter had been created in these collisions[11,12].

At the same time, a huge next step was being taken. At BNL, a dedicated new accelerator, the Relativistic Heavy-Ion Collider (RHIC), went into operation, servicing four experiments, called BRAHMS, PHENIX, PHOBOS and STAR. At RHIC, heavy nuclei such as gold collide at a relativistic centre-of-mass energy of 40,000 GeV ($\sqrt{s_{NN}} = 200$ GeV). This higher collision energy means a much larger and hotter fireball than had previously been possible (Fig. 1). Data from the first three years of running at RHIC are summarized in refs 13–16, and more recent data can be found in ref. 17.

An even braver new world will come about with the start of operations at the LHC, in which nuclei with masses up to that of lead will be able to collide at a centre-of-mass energy of 1,150 TeV ($\sqrt{s_{NN}} = 5.5$ TeV). This is a huge step in collision energy, about 30 times more than that of RHIC and, at about 0.18 mJ, the first really 'macroscopic' energy to be investigated. The fireball is expected to contain tens of thousands of gluons and quarks, and its temperature should exceed the critical temperature for the deconfinement phase transition several times over. This huge increase in energy should allow the unambiguous identification and characterization of the quark–gluon plasma.

A fireball in chemical equilibrium

As mentioned earlier, one of the crucial questions to be addressed in considering ultra-relativistic collisions between nuclei is the extent to which matter is formed in the fireball. There are two important sets of observations that support the idea of a matter-like fireball. The first concerns the fact that the fireball yields hadrons that are in chemical equilibrium, forming a statistical ensemble. Hadron yields have been studied with high precision in nuclear collisions at the energies used in

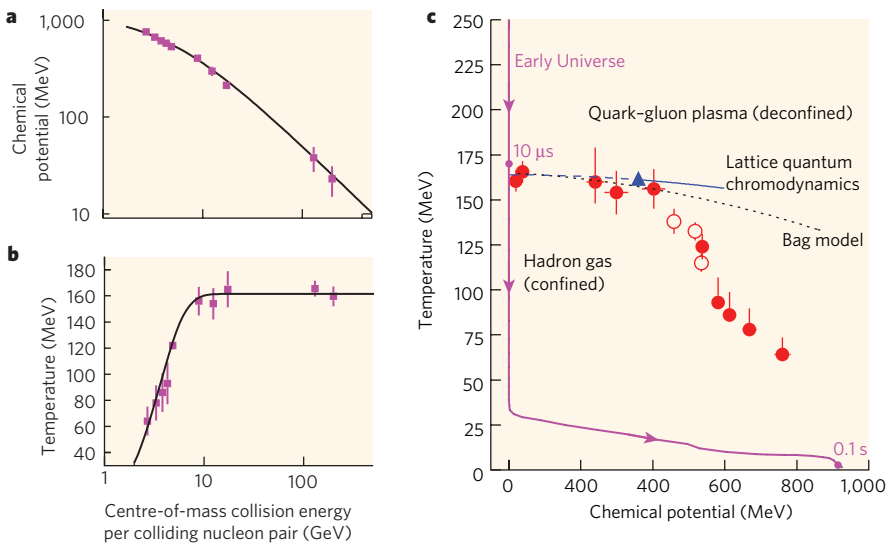


Figure 2 | Equilibrium parameters of the fireball. The energy dependence of chemical potential (**a**) and temperature (**b**), determined from a statistical analysis of hadron yields[18]. The temperature plateau at high collision energies suggests the presence of a phase boundary. Pink squares are results of individual experiments. **c**, The phase diagram of strongly interacting matter: the data points are obtained as in **a** and **b**. The evolution of the early Universe is shown, as are theoretical expectations such as from lattice quantum chromodynamics (blue line) and the bag model (dotted line) for the phase boundary between confined and deconfined matter. Red filled circles are from analysis of midrapidity data. Open circles are from analysis of 4π data. Blue triangle is possible position of a critical endpoint. **Figure reproduced, with permission, from ref. 18.**

the Alternating Gradient Synchrotron, SPS and RHIC. These yields can be described by assuming that all hadrons are formed only when the fireball reaches a specific equilibrium temperature, volume and baryon chemical potential (a measure of the energy change brought about by the addition of one more baryon to the system). Under these conditions, the hadron yields can be characterized in relatively simple terms by the thermodynamic grand-canonical ensemble or, in the special case of small particle numbers, by the canonical ensemble. Such conditions are dubbed the 'chemical freeze-out' scenario, in analogy to the production of bound particles as the early Universe cooled. Detailed analyses of the freeze-out can be found in refs 18 and 19, and a comprehensive review in ref. 20.

Importantly, the energies attained in the SPS and RHIC are also high enough to produce particles containing several strange quarks, including the Ω and $\bar{\Omega}$ baryons. Yields of these baryons agree very well with chemical-equilibrium calculations, and are much higher than in proton–proton collisions. The interpretation is that in heavy-ion collisions, the chemical freeze-out is caused by the quark–gluon plasma and its transition to normal matter, whereas this plasma is absent in collisions between protons.

With increasing centre-of-mass collision energy, the chemical potential decreases smoothly, so new baryons and antibaryons can be created with increasing ease (Fig. 2a). By contrast, although the temperature increases strongly at first, it plateaus rather abruptly near $\sqrt{s_{NN}} = 10$ GeV, at a value slightly higher than 160 MeV (Fig. 2b). This plateau supports Hagedorn's limiting-temperature hypothesis[5], and strongly suggests that a boundary — the phase boundary — is reached at a critical collision energy. Beyond that energy, all additional energy goes into heating the quark–gluon plasma which, in turn, cools again and freezes out at the phase boundary (critical temperature).

If the temperature of the collision fireball is plotted against its chemical potential, with one entry for each energy investigated, a phase diagram can be constructed for the strongly interacting matter contained within it (Fig. 2c). What emerges can be compared to various predictions of the position of the phase boundary taken[8–10] from lattice quantum chromodynamics and[21] from a simple 'bag model' of quarks' confinement into hadrons. For chemical potentials of less than about 400 MeV — corresponding to the critical energy discussed above — the temperatures and chemical potentials determined from the measured hadron yields coincide, within about 10 MeV uncertainty, with the phase boundary as determined from lattice quantum chromodynamics calculations. When the phase boundary is reached, all further points follow it — hadrons cannot be formed in the quark–gluon plasma, only as the plasma rehadronizes.

But could this just be coincidence? What mechanism enforces equilibrium at the phase boundary? Collision rates and the timescales of fireball expansion in the hadronic phase[22] imply that, at the energies used in the SPS and RHIC, equilibrium cannot be established in the hadronic medium. Rather, it is the phase transition between deconfined and confined matter that ensures chemical equilibrium through multi-particle collisions during hadronization. Alternatively, the plateau can be interpreted to arise[23,24] from the filling of phase space during hadronization. In either case, all current interpretations of the observed phenomena relate the chemical variables directly to the phase boundary. This implies that a fundamental parameter of quantum chromodynamics — namely the critical temperature for the deconfinement phase transition — has been determined experimentally to be close to 160 MeV, for small values of chemical potential.

This interpretation will be tested directly by experiments at the LHC. If the plateau phenomenon holds, as is to be expected from the above considerations, then the particle yields measured at LHC energy should, except for an overall volume parameter, agree closely with those measured at the much smaller RHIC energy. That would lend strong support to a phase boundary as the limiting agent.

The observed equilibrium is a strong indication that a matter-like medium is produced in high-energy collisions between nuclei. In collisions among particles such as leptons or nucleons, such equilibrium is not observed, at least not at the energies at which particles containing strange quarks can be produced[20], and hence no medium is formed. We finally note that, as is evident from Fig. 2c, in heavy-ion collisions the chemical freeze-out temperature is not universal but instead varies strongly at large values of the chemical potential. This implies that the properties of the medium change with energy, indicating a transition to a baryon-rich medium at low energies.

The phase transition at low baryon density is probably of the cross-over type[25]. General considerations, as well as results from lattice quantum chromodynamics, suggest the possibility of a first-order phase transition at higher baryon densities with a corresponding critical end-point as sketched in Fig. 2c. Experiments to search for the critical point are planned at the SPS, RHIC and the future Facility for Antiproton and Ion Research (FAIR) at the heavy-ion research centre GSI in Darmstadt, Germany.

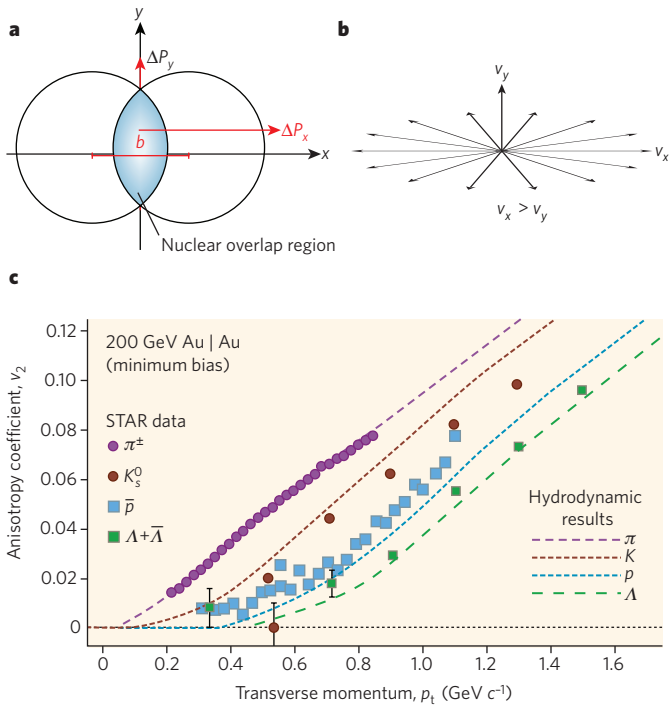


Figure 3 | Geometry of matter during a nuclear collision. a, The nuclear overlap region for semi-central collisions. Early in the collision, the pressure gradient is large in the plane of the collision, *x*. After some time, the large pressure gradient leads to a larger expansion velocity (v_x) in this direction (**b**). The expansion velocity profile in the *x*–*y* plane leads to highly asymmetrical particle emission, with azimuthal anisotropies in momenta perpendicular to the beam (p_t) of various particles. **c**, The distribution of these momenta can be quantified by a Fourier decomposition and parametrized by the second Fourier coefficient $v_2 = \langle\cos 2\varphi\rangle$ — also called the 'elliptical flow'[56–58] — in which the angle φ is measured relative to the direction of impact: higher transverse momenta are recorded for particles emerging in the reaction plane, whereas much lower momenta are observed perpendicular to the reaction plane. As a consequence, the v_2 coefficients are large and show a characteristic p_t dependence. Data for π mesons, *K* mesons, antiprotons ($\bar{p}$) and Λ baryons (with masses mc^2 of about 140, 495, 940 and 1,115 MeV, respectively) agree very well in their mass- and p_t-dependence with predictions[59–61] made with relativistic hydrodynamics and an equation of state determined by weakly interacting quarks and gluons. Although the data are not very sensitive to the particular equation of state used, equations of state based exclusively on hadrons do not lead to a satisfactory description of the data. The data shown are from the STAR experiment at RHIC[62]. **Part c reproduced, with permission, from ref. 62.**

Hydrodynamic expansion and cooling

If matter is formed in the moments after a nuclear collision, hydrodynamic flow effects should be seen owing to the strong pressure gradients present in it. At ultra-relativistic energies, two colliding nuclei are highly

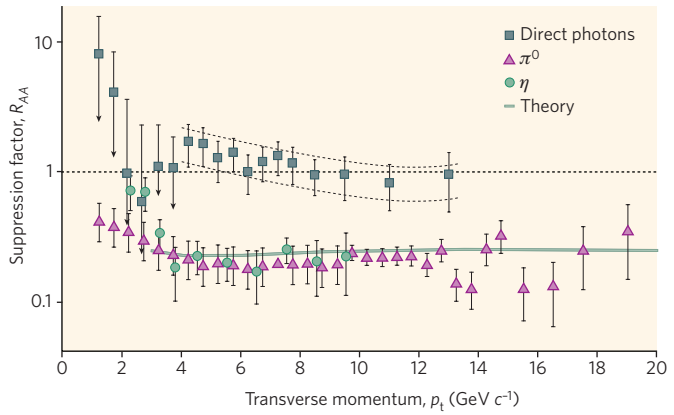


Figure 4 | Preliminary PHENIX results for the suppression factor R_{AA} out to high p_t for π^0 and η mesons. R_{AA} is the ratio of the number of events at different values of p_t for gold–gold collisions normalized to the number of events in proton–proton collisions, scaled by the number of collisions. The suppression of the gold–gold spectrum at high p_t is further evidence for the presence of a hot, dense medium on which jet partons scatter and lose momentum. Bars indicate statistical error. The dotted line at $R_{AA} = 1$ is the expected result when the photon spectrum is unmodified. A theoretical spectrum[63] is also shown, calculated under the assumptions that the precursor parton of the jet loses energy by radiation, and that the medium can be modelled as dense gluon gas. It agrees well with the experimental data. The results for photons produced directly in gold–gold and proton–proton collisions are also shown. These show no suppression; this is consistent with the idea of a gluon gas, as photons do not participate in strong interactions. Figure courtesy of the PHENIX collaboration.

Lorentz-contracted — at the RHIC energy by a factor of 100, at the LHC energy by a factor of 2,700. Consequently, the collision is very quick, lasting around 10^{-25} s. The geometry of the matter immediately after the collision is sketched in Fig. 3a; with increasing impact, b, the overlap zone becomes more and more aspherical in the plane perpendicular to the axis of the colliding beams. It attains an almond-like shape with a typical size in the perpendicular plane determined by the dimensions of the nuclei involved (the diameter of a lead nucleus is about 14 fm), whereas the extension in beam direction cannot be greater than the speed of light multiplied by the collision time — less than 1 fm.

This highly asymmetrical zone evolves by collisions between its constituents (quarks and gluons) until, after a time of about 1 fm c^{-1}, equilibrium is reached and a highly compressed, but still very asymmetrical, fireball is formed. The details of this initial phase are not well understood, but might involve highly coherent configurations of colour fields, generated by a 'colour glass condensate'[26], large fluctuations of which are thought to lead to extremely rapid equilibrium. Irrespective of the details of this highly complex evolution, some ground rules are clear if equilibrium is reached in a short enough time that the shape of the fireball remains essentially unchanged from the initial geometric overlap zone. In this case, the fireball's further evolution should be governed by the laws of relativistic hydrodynamics for a system with very strong pressure gradients, as well as by the equation of state that connects the variables such as volume, temperature and chemical potential that characterize the medium. When the original spatial correlation is transformed into a correlation in momentum (or velocity) space, this implies a very asymmetrical particle emission in the plane perpendicular to the axis of the colliding beams (Fig. 3b). The earlier the equilibrium, and with it the beginning of the hydrodynamic evolution, the larger the anisotropies will be.

What observations are to be expected if the fireball really does expand hydrodynamically, with a unique collective velocity for each fluid cell in the system? The transverse momenta (p_t) of the emitted particles are connected to the fluid velocity via $p_t = m\beta_f\gamma_f$ (with $\gamma_f = 1/\sqrt{1-\beta_f}$), in which m is the mass of the particle, β_f is the relativistic fluid velocity and γ_f is its Lorentz factor, and a characteristic mass-dependent flow pattern arises. The resulting mass ordering in the anisotropy coefficients (v_2) agrees closely with the experimental observations (Fig. 3c). The large anisotropy coefficients confirm the idea that the fireball reaches equilibrium rapidly.

This dramatic and unexpected success is the second strong pillar supporting the idea that ultra-relativistic nuclear collisions produce a collectively expanding medium in thermal equilibrium. We note here that the hydrodynamic calculations with which the experimental data are in such good agreement assume that fluid flow is non-viscous. When added to the hydrodynamic equations, even small viscosity destroys the agreement between data and calculations. Researchers have thus concluded that the matter made in the RHIC fireball is probably close to an ideal fluid[27–29].

If the fireball really does behave hydrodynamically at top RHIC energy, then the elliptical flow data from the LHC should be similar to those from RHIC. The anticipated much greater number of particles produced in each collision could be used to measure flow precisely for many types of particle. That could in turn allow the equation of state of the matter and its transport coefficients (such as viscosity) to be pinned down. Alternatively, as still argued by some authors[30], the flow pattern observed at RHIC is due to a cancellation between unusual initial conditions of the fireball owing to a colour glass condensate on the one hand and an imperfect thermal equilibrium on the other. If that is so, or if viscous effects do play a part, then the LHC data on elliptical flow could reach much larger values. In any case, the very large energy step when going from RHIC to the LHC should lead to important, and urgently needed, new information on the physics of the quark–gluon plasma.

An opaque matter

In collisions of heavy nuclei, hard scattering events — those with high momentum transfer — between the constituent 'partons' (quarks and gluons) liberated are expected to occur just as in collisions between protons. The number of such events, however, will scale with the number of individual proton–proton collisions for a given collision geometry: for head-on collisions of two equal nuclei of mass number A, the number of

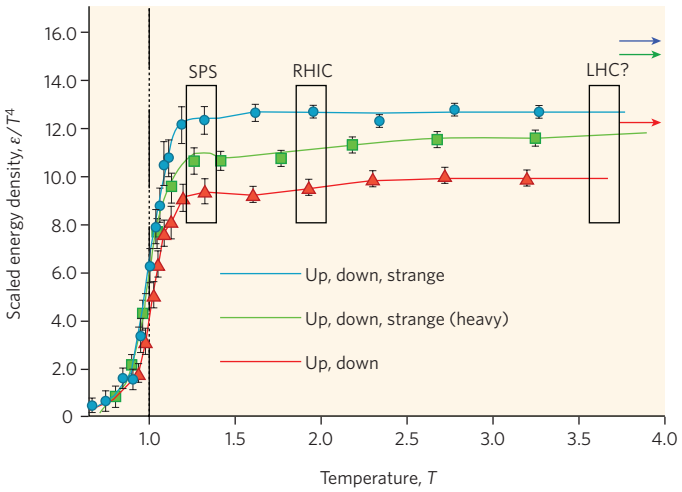


Figure 5 | Scaled energy density ε/T^4 as a function of the temperature calculated in lattice quantum chromodynamics[6]. For an ideal gas, the energy density is proportional to the fourth power of the temperature with the proportionality constant containing the number of degrees of freedom. The strong increase near the critical temperature (T_c, vertical line) indicates that the system is not only heated but that something dramatic happens: it undergoes a phase transition from hadronic matter to quark–gluon plasma with a corresponding large increase in the number of degrees of freedom. Above T_c, the quark–gluon plasma is only heated such that ε/T^4 is constant. The three lines are calculations for two light quark flavours (only up and down; red), three equally light flavours (up, down and strange; blue) and the most realistic case of two light flavours (up and down) and one more massive (strange) flavour (green). Coloured arrows show the expected values of scaled energy density at the Stefan-Boltzman limit. The regions labelled by accelerator facilities indicate maximum initial temperatures reached there. Figure reproduced, with permission, from ref. 65.

events will scale as $A^{4/3}$. Individual collisions between protons are thought to occur independently of each other, and their number can be computed from the distributions of the nuclear densities, the nuclear overlap for a given impact and the inelastic proton–proton cross-section.

Collisions of nuclei differ from collisions between protons in that the hard scattered partons may traverse the quark–gluon plasma before or during their hadronization into a jet. Jets are characteristic of collisions between protons in which two constituent partons scatter and recede from each other with a significant fraction of the initial beam momentum. In the plane transverse to the beams, the momenta are large and opposite in direction. The two scattered partons hadronize mainly into mesons that are emitted in a cone — the jet — around the direction of parton momentum. It was realized very early[31] that the quark–gluon plasma could modify jets resulting from collisions between nuclei. Calculations showed that a parton traversing a hot and dense medium consisting of other partons — that is, a quark–gluon plasma — should lose substantially more energy than one traversing cold nuclear matter[32–34]. This prediction appears to be borne out by data from all four experiments at RHIC.

A jet is much more difficult to see in a heavy-ion collision than after a collision between protons. The reason is the sheer number of particles produced: a single central (head-on) gold–gold collision generates about 5,000 charged particles, and unless the jet has very high (transverse) momentum, it will not stand out in the crowd. But the presence of jets will affect the overall transverse momentum distribution. At low transverse momenta, the spectrum in a heavy-ion collision is complex, as it is a superposition of hydrodynamic expansion effects and random thermal motion. Nevertheless, for particles of a particular species with transverse momenta that are significantly larger than their mass, the resulting spectrum is nearly exponential. The contribution of jets with high transverse momentum leads to a distinct power-law behaviour typically visible for values of transverse momentum of a few GeV or more.

To judge a possible modification of the shape of the spectrum in a high-energy nuclear collision, the transverse-momentum distribution of π mesons produced in central gold–gold collisions at RHIC can be compared with that measured in proton–proton collisions. To quantify this comparison, the ratio of the gold–gold-collision spectrum to the proton–proton-collision spectrum is scaled to the total number of inelastic collisions in the nuclear case, providing the suppression factor R_{AA}. For larger transverse momenta, this factor settles at about 0.2 (Fig. 4); that is, the production of high-momentum π mesons is suppressed by a factor of five in gold–gold collisions.

What is the origin of this suppression? The transverse-momentum spectrum for collisions between protons agrees well[35] with theoretical calculations that use next-to-leading-order quantum chromodynamic perturbation theory. When the spectra of deuteron–gold collisions of varying centrality are compared with the proton–proton spectrum, R_{AA} is 1 or larger (for more central collisions, values larger than 1 are even expected — a phenomenon known as the Cronin effect, caused by the scattering of partons before the hard collision). For peripheral gold–gold collisions, the values of R_{AA} also correspond well to the expectation from collisions between protons. The clear implication is that something special and new happens in central gold–gold collisions: the precursor parton of the jet produced must lose a lot of energy, causing the transverse-momentum spectrum of the mesons in the jet to fall off steeply.

Several researchers have shown that only calculations including large energy loss in the medium can account for these data. The clear implication is that the medium present in the collision fireball is hot and dense, and when partons pass through it, they lose energy. Both radiation of gluons and elastic scattering seem to be important here. In deuteron–gold collisions, by contrast, the jet sees at most cold nuclear matter (or a vacuum), and does not seem to be perturbed.

Calculating the energy loss of a fast parton in a quantum chromodynamic liquid, as suggested by the data discussed in the previous section, is beyond the current theoretical state-of-the-art. To gain insight into the underlying physics of energy loss, it is helpful to resort to another aspect of the medium: that it contains many gluons. Indeed, the RHIC data on parton energy loss are well explained by modelling the medium formed by the collision as an ultra-dense gluon gas with a density of the number of gluons (N_g) per rapidity interval of $dN_g/dy = 1{,}100$. Here, the rapidity y is a logarithmic measure of the gluon's longitudinal velocity, v. With the simple assumption that $v = z/t$ (z is the longitudinal space coordinate), Bjorken[36] showed how to map rapidity densities to spatial densities. The spatial gluon density in turn is linked directly to entropy density. Using relations from statistical mechanics for a relativistic gas of bosons (and fermions if quarks are included), the temperature and energy density can be obtained from these gluon densities. The high gluon densities needed to reproduce the observed gold–gold R_{AA} correspond to an initial temperature of about twice the critical temperature for the formation of a quark–gluon plasma. The initial energy densities of 14–20 GeV fm^{-3} are

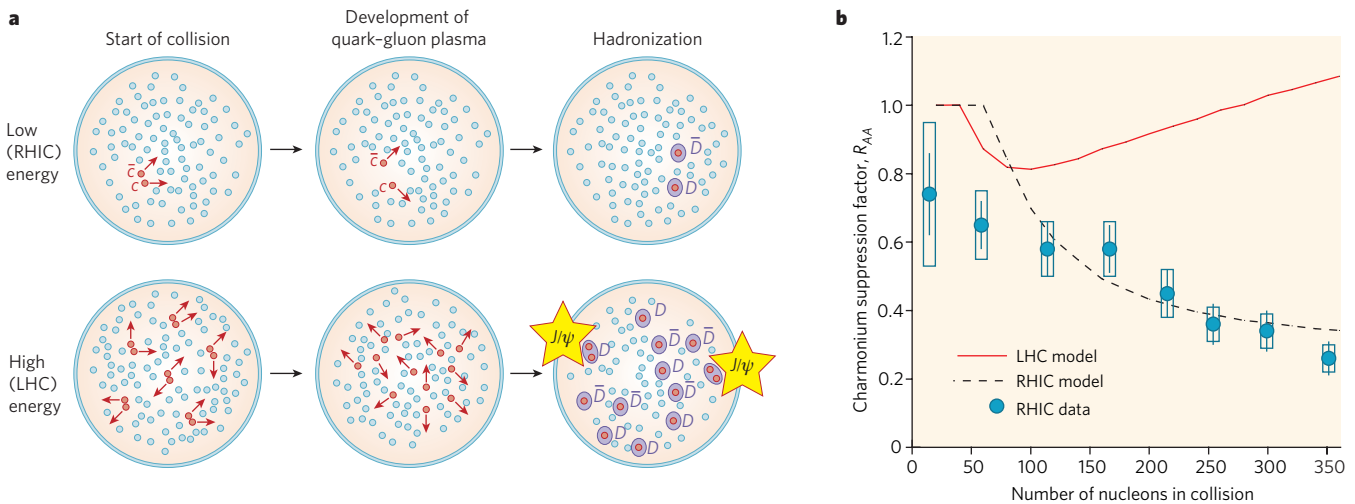


Figure 6 | Charmonium suppression. a, At low energies, the quark–gluon plasma screens interaction between the only pair of charm quark and antiquark produced (red dots) and any other two quarks (up, down, strange) will find themselves paired with the charm quark/antiquark in *D* mesons at hadronization (purple circles). At high energies, by contrast, many charm–anticharm pairs are produced in every collision and at hadronization, charm and anticharm quarks from different original pairs may combine to form a charmonium J/Ψ particle. Grey dots indicate light partons produced in the collision. **b**, Theory and experiment compared quantitatively. Model predictions[55] for the charmonium suppression factor agree well with recent RHIC data from the PHENIX collaboration[66]. Owing to the increased level of statistical recombination expected, enhancement rather than suppression is predicted for LHC conditions. What the experiments deliver will be a further crucial test of theories of the quark–gluon plasma. Part b reproduced, with permission, from ref. 55.

Box 1 | ALICE: the LHC's dedicated plasma hunter

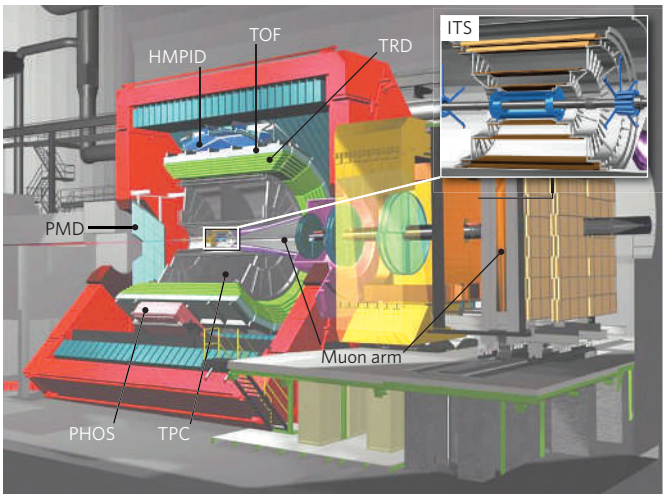
LHC's ALICE detector is dedicated to investigating nucleus–nucleus collisions at an energy of 2.8 TeV per nucleon in each of the colliding nuclei. The main part of the apparatus is housed in the world's largest solenoidal magnet, which generates a field strength of 0.5 T within a volume of 1,600 m^3.

Various detectors arranged in cylindrical shells around the interaction point (see figure) are designed to determine the identity and precise trajectory of the more than 10,000 charged particles propelled by a lead–lead nuclear collision into the active volume of the apparatus. The innermost detector is the inner tracking system (ITS), which consists of six layers of silicon detectors surrounding the 1-mm-thick beam pipe that encloses the ultra-high vacuum of the accelerator. These detectors are capable of high-precision tracking (resolution around 20 μm) so as to determine the decay vertex of short-lived particles carrying strange, charm, or bottom quantum numbers that typically decay within a few millimetres to centimetres of the primary interaction point.

The ITS is contained within, and mounted on, the cylindrical barrel of the time projection chamber (TPC). This is ALICE's major tracking device: it is the largest of its kind worldwide, with some 560,000 readout channels, and provides essentially continuous, three-dimensional tracking of charged particles between radii of 80 cm and 250 cm from the central interaction point.

Outside the TPC are two very large (with areas of around 150 m^2) particle identification detectors: the transition radiation detector (TRD), with more than 1 million channels and an on-board computer farm of a quarter-million central processing units for the triggering and identifying of electrons; and surrounding this, the time-of-flight (TOF) detector which can record the transit time between the interaction point and the detector surface at a resolution of better than 100 picoseconds.

The central barrel of ALICE is completed by dedicated detectors to measure photons (PHOS) and their distribution in the forward direction (PMD) and to identify high-momentum hadrons (HMPID), and by



further detectors to determine the position and time of the primary interaction point. Separated from the main detector in the forward direction of one of the accelerator beams, and behind a conical absorber that projects into the central barrel, is a muon detector with its own large dipole magnet. Because muons do not undergo strong reactions, and because those at relevant energies do not emit *bremsstrahlung*, they penetrate the absorber practically unscathed — unlike hadrons. Their momenta are then measured in tracking stations before, inside, and after the dipole magnet.

The huge scale and cutting-edge engineering of the ALICE detector should allow it to make a decisive contribution to understanding the properties of the medium that will be created in the LHC's high-energy collisions between nuclei. Image courtesy of the ALICE collaboration.

also well in line with the initial conditions required by the hydrodynamic models introduced earlier. Both the temperature and the energy density are well above the critical conditions calculated with lattice quantum chromodynamics (Fig. 5).

Another important cross-check for the gluon-gas interpretation is to compare the transverse-momentum spectrum of photons produced in hard initial parton scattering in gold–gold and in proton–proton collisions. For most values of transverse momentum, the corresponding R_{AA} factors are consistent with unity (Fig. 4). This is perfectly in line with an unmodified distribution as produced in initial hard scattering: the photon does not participate in the strong force, and so would traverse a gluon-dominated fireball without further interaction.

Similar analyses have been done for several different hadronic species. Generally, within the errors, all mesons behave just like π mesons (the data points for η mesons are shown in Fig. 4). In an intermediate range of 2–6 GeV, the suppression of baryons is significantly weaker, but for high transverse momentum it joins that for π mesons.

Another characteristic of hadron jets is the back-to-back correlation of two of them in the plane perpendicular to the colliding beams. Even without reconstructing the jet (the difficulty of doing this in a heavy-ion collision has been mentioned), such an analysis can be done by selecting just the particles that have the highest momentum in an event. Both the PHENIX and STAR experiments have undertaken such a 'leading particle' analysis, picking just one high-transverse-momentum trigger particle (say, in the range 4–6 GeV) and then checking for the distribution in azimuthal angle of all other 'associated' particles in a given range of transverse momenta. Proton–proton experiments produce two peaks, one at 0° and one at 180°. This is expected, as the associated particle could be a fragment of the same jet (the 0° peak), or a leading particle of the second jet diametrically opposite (the 180° peak).

In gold–gold collisions, a dramatic change is seen[37,38] when the transverse momentum of the associated particle is varied. At very high momenta, the two expected peaks are present. For lower particle momenta (2–4 GeV), however, the peak at 180° broadens to the point that it is barely visible. For even lower momenta, it develops into a dip with two pronounced peaks, one on either side about 1 radian apart. A very similar observation has been made at the top SPS energy[39,40]. The interpretation of this feature of the data is still hotly debated, but an interesting suggestion has been made[41,42]: the dip with the two satellite peaks could indicate the presence of a shock wave in the form of a Mach cone caused by a supersonic parton traversing the quark–gluon plasma. If borne out, this could lead to the determination of the speed of sound in the plasma.

The modelling of the parton energy loss in the quark–gluon plasma is still somewhat schematic, and leaves open a range of theoretical possibilities and ways of implementation. To sharpen the interpretation, it would be good to have individual measurements of the properties of jets stemming from all different quark flavours, as well as from gluons. For heavy quarks (charm and bottom), an important step has been made by the PHENIX and STAR collaborations at RHIC. Both experiments have measured transverse-momentum spectra of electrons stemming from decays of D and B mesons (each of which contain a charm or bottom quark) into electrons plus anything else. The R_{AA} values have been determined for these spectra[43,44], too, and have been surprising: they are very close to the values determined for mesons that involve only up, down and strange quarks. Theoretically, energy loss by radiation should be much lower for the charm and bottom quarks than for the up, down and strange quarks and gluons owing to their much larger masses. Since then, it has been realized that energy loss by scattering is probably of comparable importance to energy loss by radiation[45,46], which improves the quantitative situation somewhat, but doesn't resolve the puzzle.

With the start of experiments at the LHC, matters will change dramatically: because of the much higher beam energy, jet production will be enhanced by many orders of magnitude compared with the situation at RHIC energy. Estimates[47] based on solving quantum chromodynamics by perturbation theory imply, for example, an enhancement of

more than four orders of magnitude at $p_t = 100$ GeV. Thus a whole new range of transverse-momentum values between 20 and 250 GeV will become accessible. Such measurements can then be used to discriminate between the various theoretical scenarios competing to describe parton energy loss, either by determination of R_{AA} or, uniquely for the LHC, by direct reconstruction of the jet in a collision between nuclei. This should allow jet probes to be developed into quantitative tools to determine the parton density of the matter formed.

Charmonia as harbingers of deconfinement?

The particles collectively known as charmonia — bound states of heavy charm quarks and antiquarks — have a special role in research into the quark–gluon plasma. In 1986, Satz and Matsui[48] realized that the high density of gluons in a quark–gluon plasma should destroy charmonium systems, in a process analogous to Debye screening of the electromagnetic field in a plasma through the presence of electric charges. The suppression of charmonia (compared with their production in the absence of a quark–gluon plasma) was thus proposed as a 'smoking gun' signature for plasma formation in nuclear collisions at high energy. Measurements at the SPS[49] did indeed provide evidence for such suppression in central collisions between heavy nuclei. No suppression was found in grazing collisions or in collisions between light nuclei, in which a plasma is not expected to form. But absorption of charmonium in the nuclear medium, as well as its break-up by hadrons produced in the collision, is also a mechanism that could lead to charmonium suppression even in the absence of plasma formation[50], and the interpretation of the SPS data remains inconclusive.

This situation took an interesting turn in 2000, when researchers realized that the large number of charm-quark pairs produced in nuclear collisions at collider energies could lead to new ways to produce charmonium, either through statistical production at the phase boundary[51,52], or through coalescence of charm quarks in the plasma[53]. At low energy, the mean number of charm-quark pairs produced in a collision is much fewer than 1, implying that charmonium is formed, if at all, always from charm quarks of this one pair. Because the number of charm quarks in a collision at LHC energy is expected to reach about 200, charm quarks from different pairs can combine to form charmonium (Fig. 6a). This works effectively only if a charm quark can travel a substantial distance in the plasma to 'meet' its prospective partner. Under these conditions, charmonium production scales quadratically with the number of charm-quark pairs, so enhancement, rather than strong suppression, is predicted for LHC energy[54,55] (Fig. 6b). If observed, this would be a spectacular fingerprint of a high-energy quark–gluon plasma, in which charm quarks are effectively deconfined. Again, as in most other cases, the data from the LHC will be decisive in settling the issue.

Looking forward

The data from the SPS and particularly the RHIC accelerator have taught us that central nuclear collisions at high energy produce a medium made up of partonic matter in equilibrium and possessing collective properties. The medium flows much like an ideal liquid and is dense enough to dissipate most of the energy of a 20 GeV parton. Future experiments at RHIC will further elucidate some of the aspects discussed above, in particular in the heavy-quark sector. With their 30 times higher energy, lead–lead collisions at LHC in the specially designed ALICE detector (Box 1) will produce this new state of matter at unprecedented energy densities and temperatures and over very large volumes compared with the size of the largest stable nuclei. With the planned experiments, the LHC heavy-ion community looks forward with anticipation to elucidating the properties of such partonic fireballs. The prize is unravelling the mystery of the matter that formed a fraction of a nanosecond after the Big Bang, and disappeared just 10 microseconds later. ■

1. Gross, D. J. & Wilczek, F. Ultraviolet behavior of non-abelian gauge theories. *Phys. Rev. Lett.* **30,** 1343–1346 (1973).
2. Politzer, H. J. Reliable perturbative results for strong interactions? *Phys. Rev. Lett.* **30,** 1346–1349 (1973).
3. Cabibbo, N. & Parisi, G. Exponential hadronic spectrum and quark liberation. *Phys. Lett. B* **59,** 67–69 (1975).
4. Collins, J. C. & Perry, M. J. Superdense matter: neutrons or asymptotically free quarks? *Phys. Rev. Lett.* **34,** 1353–1356 (1975).
5. Hagedorn, R. Statistical thermodynamics of strong interactions at high energies. *Nuovo Cimento Suppl.* **3,** 147–186 (1965).
6. Karsch, F., Laermann, E. & Peikert, A. Quark mass and flavour dependence of the QCD phase transition. *Nucl. Phys. B* **605,** 579–599 (2001).
7. Aoki, A., Fodor, Z., Katz, S. D. & Szabo, K. K. The QCD transition temperature: results with physical masses in the continuum limit. *Phys. Lett B.* **643,** 46–54 (2006).
8. Fodor, Z. & Katz, S. Critical point of QCD at finite T and μ, lattice results for physical quark masses. *J. High Energy Phys.* **4,** 050 (2004).
9. Allton, C. R. *et al.* The QCD thermal phase transition in the presence of a small chemical potential. Preprint at <http://arxiv.org/abs/hep-lat/0204010> (2002).
10. Ejiri, S. *et al.* Study of QCD thermodynamics at finite density by Taylor expansion. Preprint at <http://arxiv.org/abs/hep-lat/0312006> (2003).
11. CERN. New state of matter created at CERN. <http://press.web.cern.ch/press/PressReleases/Releases2000/PR01.00EQuarkGluonMatter.html> (2000).
12. Heinz, U. & Jacob, M. Evidence for a new state of matter: an assessment of the results from the CERN lead beam programme. Preprint at <http://arxiv.org/abs/nucl-th/0002042> (2000).
13. Arsene, I. *et al.* Quark–gluon plasma and color glass condensate at RHIC? The perspective from the BRAHMS experiment. *Nucl. Phys. A* **757,** 1–27 (2005).
14. Back, B. B. *et al.* The PHOBOS perspective on discoveries at RHIC. *Nucl. Phys. A* **757,** 28–101 (2005).
15. Adams, J. *et al.* Experimental and theoretical challenges in the search for the quark–gluon plasma: the STAR Collaboration's critical assessment of the evidence from RHIC collisions. *Nucl. Phys. A* **757,** 102–183 (2005).
16. Adcox, K. *et al.* Formation of dense partonic matter in relativistic nucleus–nucleus collisions at RHIC: experimental evaluation by the PHENIX Collaboration. *Nucl. Phys. A* **757,** 184–283 (2005).
17. Csörgö, T., Dávid, G., Lévai. P. & Papp, G. (eds) Quark matter 2005 — proceedings of the 18th international conference on ultra-relativistic nucleus–nucleus collisions. *Nucl. Phys. A* **774,** 1–968 (2006).
18. Andronic, A., Braun-Munzinger, P. & Stachel, J. Hadron production in central nucleus–nucleus collisions at chemical freeze-out. *Nucl. Phys. A* **772,** 167–199 (2006).
19. Becattini, B., Gadzicki, M., Keranen, A., Manninen, J. & Stock, R. Chemical equilibrium study in nucleus–nucleus collisions at relativistic energies. *Phys. Rev. C* **69,** 024905 (2004).
20. Braun-Munzinger, P., Redlich, K. & Stachel, J. in *Quark–Gluon Plasma 3* (eds Hwa, R. C. & Wang, X. N.), 491–599 (World Scientific, Singapore, 2004).
21. Braun-Munzinger, P. & Stachel, J. Probing the phase boundary between hadronic matter and the quark-gluon plasma in relativistic heavy-ion collisions. *Nucl. Phys. A* **606,** 320–328 (1996).
22. Braun-Munzinger, P., Stachel, J. & Wetterich, C. Chemical freeze-out and the QCD phase transition temperature. *Phys. Lett. B.* **596,** 61–69 (2004).
23. Stock, R. The parton to hadron phase transition observed in Pb+Pb collisions at 158 GeV per nucleon. *Phys. Lett. B* **456,** 277–282 (1999).
24. Heinz, U. Hadronic observables: theoretical highlights. *Nucl. Phys. A* **638,** c357–364 (1998).
25. Aoki, Y., Fodor, Z., Katz, S. D. & Szabo, K. K. The order of the quantum chromodynamics transition predicted by the standard model of particle physics. *Nature* **443,** 675–678 (2006).
26. Gyulassy, M. & McLerran, L. New forms of QCD matter discovered at RHIC. *Nucl. Phys. A* **750,** 30–63 (2005).
27. Baier, R., Romatschke, P. Causal viscous hydrodynamics for central heavy-ion collisions. Preprint at <http: //arxiv.org/nucl-th/0610108> (2006).
28. Romatschke, P. Causal viscous hydrodynamics for central heavy-ion collisions II: meson spectra and HBT radii. Preprint at <http://arxiv.org/nucl-th/0701032> (2007).
29. Heinz, U., Song, H. & Chaudhuri, A. K. Dissipative hydrodynamics for viscous relativistic fluids. *Phys. Rev. C* **73,** 034904 (2006).
30. Bhalerao, R. S., Blaizot, J. P., Borghini, N. & Ollitrault, J. Y. Elliptic flow and incomplete equilibration at RHIC. *Phys. Lett. B* **627,** 49–54 (2005).
31. Bjorken, J. D. Fermilab-PUB-82-59-THY (1982) and erratum (unpublished).
32. Wang, X. N. & Gyulassy, M. Gluon shadowing and jet quenching in *A+A* collisions at $\sqrt{s} = 200A$ GeV. *Phys. Rev. Lett.* **68,** 1480–1483 (1992).
33. Baier, H., Dokshitzer, Y. L., Mueller, A. H., Peigne, S. & Schiff, D. Radiative energy loss of high energy quarks and gluons in a finite-volume quark-gluon plasma. *Nucl. Phys. B* **483,** 291–320 (1997).
34. Baier, H., Dokshitzer, Y. L., Mueller, A. H., Peigne, S. & Schiff, D. Radiative energy loss and $p_\perp$-broadening of high energy partons in nuclei. *Nucl. Phys. B* **484,** 265–282 (1997).
35. d'Enterria, D. Quantum chromo (many-body) dynamics probed in the hard sector at RHIC. Preprint at <http://arxiv.org/abs/nucl-ex/0406012> (2004).
36. Bjorken, J. D. Highly relativistic nucleus–nucleus collisions: the central rapidity region. *Phys. Rev. D* **27,** 140–151 (1983).
37. Adler, S. S. *et al.* (PHENIX collaboration). A detailed study of high-p_T neutral pion suppression and azimuthal anisotropy in Au+Au Collisions at $\sqrt{s_{NN}} = 200$ GeV. Preprint at <http://arxiv.org/abs/nucl-ex/0611007> (2006).
38. Mischke, A. (for the STAR collaboration). High-p_T hadron production and triggered particle correlations. Preprint at <http://arxiv.org/abs/nucl-ex/0605031> (2006).
39. Adamova, D. *et al.* (CERES collaboration). Semihard scattering unraveled from collective dynamics by two-pion azimuthal correlations in 158*A* GeV/*c* Pb+Au collisions. *Phys. Rev. Lett.* **92,** 032301 (2004).
40. Ploskon, M. (for the CERES collaboration). Two particle azimuthal correlations at high transverse momentum in Pb–Au at 158 A GeV/c. Preprint at <http://arxiv.org/abs/nucl-ex/0701023nucl-ex/0701023> (2007).
41. Stöcker, H. Collective flow signals the quark–gluon plasma. *Nucl. Phys. A* **750,** 121–147 (2005).

42. Casalderrey-Solana, J., Shuryak, E. & Teaney, D. Conical flow induced by quenched QCD jets. Preprint at <http://arxiv.org/abs/hep-ph/0411315> (2004).
43. Adare, A. *et al.* (the PHENIX collaboration). Energy loss and flow of heavy quarks in Au+Au collisions at $\sqrt{s_{NN}}$ = 200 GeV. Preprint at <http://arxiv.org/abs/nucl-ex/0611018> (2006).
44. Zhang, H. *et al.* Heavy flavour production at STAR. *J. Phys. G* **32,** S29–S34 (2006).
45. Zapp, K., Ingelman, G., Rathmans, J. & Stachel, J. Jet quenching from soft QCD scattering in the quark–gluon plasma. *Phys. Lett. B.* **637,** 179–184 (2006).
46. Adil, A., Gyulassy, M., Horowitz, W. A. & Wicks, S. Collisional energy loss of non asymptotic jets in a QGP. Preprint at <http://arxiv.org/abs/nucl-th/0606010> (2006).
47. Vitev, I. Contribution to CERN yellow report on hard probes in heavy ion collisions at the LHC. Preprint at <http://arxiv.org/abs/hep-ph/0310274> (2003).
48. Satz, H. & Matsui, T. J/ψ suppression by quark-gluon plasma formation. *Phys. Lett. B* **178,** 416–422 (1986).
49. Abreu. M. C. *et al.* Transverse momentum distributions of $J/\psi, \psi'$. Drell–Yan and continuum dimuons produced in Pb–Pb interactions at the SPS. *Phys. Lett. B* **499,** 85–96 (2001).
50. Capella, A., Kaidalov, A. B. & Sousa, D. Why is the J/ψ suppression enhanced at large transverse energy? *Phys. Rev. C* **65,** 054908 (2002).
51. Braun-Munzinger, P. & Stachel, J. (Non)thermal aspects of charmonium production and a new look at J/ψ suppression *Phys. Lett. B* **490,** 196–202 (2000).
52. Braun-Munzinger, P. & Stachel, J. On charm production near the phase boundary. *Nucl. Phys. A* **690,** 119–126 (2001).
53. Thews, R. L., Schroedter, M. & Rafelski, J. Enhanced J/ψ production in deconfined quark matter. *Phys. Rev. C* **63,** 054905 (2001).
54. Andronic, A., Braun-Munzinger, P., Redlich, K. & Stachel, J. *Nucl. Phys. A.* Preprint at <http://arxiv.org/abs/nucl-th/0611023> (2006).
55. Andronic, A., Braun-Munzinger, P., Redlich, K. & Stachel, J. *Phys. Lett. B.* Preprint at <http://arxiv.org/abs/nucl-th/0701079> (2007).
56. Barrette, J. *et al.* Observation of anisotropic event shapes and transverse flow in Au–Au collisions at AGS energy. *Phys. Rev. Lett.* **73,** 2532–2536 (1994).
57. Voloshin, S. & Zhang, Y. C. Flow study in relativistic nuclear collisions by Fourier expansion of azimuthal particle distributions. *Z. Phys. C* **70,** 665–671 (2007).
58. Poskanzer, A. M & Voloshin, S. A. Methods for analyzing anisotropic flow in relativistic nuclear collisions. *Phys. Rev. C* **58,** 1671–1678 (1998).
59. Huovinen, P. *et al.* Radial and elliptic flow at RHIC: further predictions. *Phys. Lett. B* **503,** 58–64 (2001).
60. Teaney, D., Lauret, J. & Shuryak, E. A hydrodynamic description of heavy ion collisions at the SPS and RHIC. Preprint at <http://arxiv.org/abs/nucl-th/0110037> (2001).
61. Kolb, P. & Heinz, U. in *Quark–Gluon Plasma 3* (eds Hwa, R. C. & Wang, X. N.) 634–714 (World Scientific, Singapore, 2004).
62. Adams, J. *et al.* (the STAR collaboration). Azimuthal anisotropy in Au+Au collisions at $\sqrt{s_{NN}}$ = 200 GeV. *Phys. Rev.* **C72,** 014904 (2005).
63. Vitev, I. Jet quenching in relativistic heavy ion collisions *J. Phys. G.* Preprint at <http://arxiv.org/abs/hep-ph/0503221> (2005).
64. Akiba, Y. Probing the properties of dense partonic matter at RHIC. *Nucl. Phys. A* **774,** 403–408 (2006).
65. Kolb, P. & Heinz, U. in *Quark–Gluon Plasma 3* (eds Hwa, R. C. & Wang, X. N.) 1–59 (World Scientific, Singapore, 2004).
66. Adare, A. *et al.* (PHENIX collaboration). J/ψ production vs centrality, transverse momentum, and rapidity in Au+Au collisions at $\sqrt{s_{NN}}$ = 200 GeV. Preprint at <http://arxiv.org/abs/nucl-ex/0611020nucl-ex/0611020> (2006).

Author Information Reprints and permissions information is available at npg.nature.com/reprintsandpermissions. The authors declare no competing interests. Correspondence should be addressed to J.S. (stachel@physi.uni-heidelberg.de).

NATURE|Vol 448|19 July 2007|doi:10.1038/nature06081

The God particle *et al.*

Leon Lederman

The territory of the Large Hadron Collider might be populated not just by the Higgs particle but also by all manner of other exotic apparitions.

The birth of particle physics — that is, high-energy physics — can be dated to about 1950, offspring of the marriage of nuclear physics and the study of cosmic rays. It exploited techniques and technology from both disciplines, and its objective was to identify the primordial particles of nature — those from which all matter is made — and codify the laws of physics that oversee their properties and social behaviours.

Progress in high-energy physics has always been mortgaged to the requirements of the ever more powerful particle accelerators required to reveal the inner life of the particle 'zoo'. Around 1950, the world energy record was held by a synchrocyclotron accelerator that accelerated protons to 400 megaelectronvolts (MeV) around a circular path. That was enough to shatter atomic nuclei and produce copious quantities of pions and muons. These particles, first discovered in investigations of cosmic rays, were indicators of a vast complexity to come.

The record-breaking 'atom-smasher' of 1950 was constructed by the physics department of Columbia University on the Nevis estate, about 30 miles north of the university, bordering the Hudson River. At its dedication, by the then University president, Dwight Eisenhower, a series of relays brought the Nevis synchrocyclotron to life, as attested by a Geiger counter emitting an amplified series of clicks. My job, as a new graduate student intent on using the accelerator for my PhD research, was to stand by with a radioactive source in case the machine failed. It did, of course, and, as I had misplaced the source, the dawn of this new era in particle physics was delayed, in the ears of the assembled company, by five very embarrassing minutes.

The Nevis machine soon left this small setback behind. This machine, and others that followed all over the world, ensured that by 1995 the cutting-edge energy domain had climbed by a factor of more than 2,000 over those early days. The probe of choice was protons at an energy of 900 gigaelectronvolts (GeV) in head-on collision with antiprotons of the same energy. These collisions, which occurred at a rate of almost 10^6 per second, took place in the 6.3-km-circumference ring of the Tevatron accelerator, at Fermilab, in Batavia, Illinois. The principle of conservation of momentum tells us that a head-on collision is much more violent than is aiming one beam at a stationary target, as the Nevis machine had done. It is the difference between a large speeding truck colliding with a ping-pong ball and two equally huge trucks involved in a full-on collision. In the first case, nothing much happens to the truck, and the ping-pong ball recoils rapidly, none the worse for wear. In the second case, bumpers, mirrors, radios and steering wheels fly off in all directions. Picking through the debris left by such an impact gives us a good grasp of how the truck's interior was put together.

In 1979, as the new director of Fermilab, I made the decision that protons should smash into antiprotons in the new accelerator. The Tevatron's success was crowned in 1995 with its discovery of the last and heaviest of the expected fundamental matter particles: the top quark. The ultimate product of the increasingly savage collisions at Fermilab and elsewhere in the years between 1950 and 1995 was the seemingly complete, self-contained and self-consistent table of nature's fundamental particles — the 'standard model' (see page 270).

With this table now seemingly replete, why is there still hunger for further discovery? Why is the Tevatron now running 24 hours a day, 7 days a week at huge, historic collision rates for fear of what its soon-to-be rival, the Large Hadron Collider (LHC), might find? There is a palpable sense of expectation in the control rooms of the Tevatron; in the construction activities of the physicists from all over the world participating in the LHC project; in the coming together of the massive detectors that will provide the eyes of the LHC; and, most especially, in the quiet rooms populated by theoretical physicists.

A new frontier

By 2009, it is reasonable to expect that the LHC will have claimed the crown of king of the accelerators. When the LHC is completed, the frontier of particle physics will be at a total collision energy of 14 teraelectronvolts (TeV), far beyond the energies reached by the Tevatron. At this frontier, the past decade of research in high-energy physics and in experimental astrophysics tells us that the known, explored world of the standard model and the summed achievements of the past half-century will be behind us. Defects in our theoretical construct — founded on the twin pillars of quantum field theory and Albert Einstein's general theory of relativity — that have become ever more apparent, but that could until now be ignored, will have to be confronted. We are reaching what the medieval map-maker would have denoted *terra incognita*.

By far the most popular expected denizen of these unknown lands is the Higgs particle, which was postulated to tidy up the glitch in the standard model known as electroweak symmetry breaking. Ever since Einstein published his special theory of relativity in 1905, theorists have had great respect for symmetry. At the heart of special relativity is the idea of Lorentz symmetry: that all laws of physics should be the same for all observers moving with constant relative velocities. The equation $E = mc^2$ is a direct consequence of this symmetry. Today, you cannot visit a high-energy physics laboratory without stumbling over symmetry on your way in. As in art and architecture, symmetry in this sense is an aesthetic concept: we believe that nature is best described in equations that are as simple, beautiful, compact and universal as possible. According to this way of thinking, the W and Z particles, which carry the weak nuclear force, and the photon, carrier of the electromagnetic force, should combine to show electroweak symmetry, and all should have zero mass.

> "We are reaching what the medieval map-maker would have denoted *terra incognita*."

Unless the aesthetes are fundamentally wrong, therefore, the fact that electroweak symmetry isn't perfect — because the W and Z particles are heavy — means that something is acting to break the symmetry. This something will give mass not only to the W and Z particles but also to all other particles except those few (such as photons) that can escape its clutches. The effect can be compared to running swiftly on hard ground versus knee-high through oil. In oil, your motion is slower, as if your mass had increased. The Higgs particle is that oil, and a Higgs 'field' is spread across the entire Universe.

In 1993, I co-authored a book on the history and status of high-energy physics. Then, as now, this mysterious Higgs field haunted us. As well

as explaining why particles had mass, the Higgs field allowed theorists to calculate reactions that, lacking such a speculative field, yielded nonsensical values. The beauty of the Higgs idea stimulated us to name the book *The God Particle*. "Besides," as my editor explained with an eye on the sales figures, "no one has ever heard of Higgs."

Ever more precise data emerging from the Tevatron indicate that the Higgs particle itself is not very heavy. It should be relatively easy to produce at the huge energy of the LHC. If so, questions abound. Is the Higgs alone, or is there a whole family of Higgs-type particles? Does the Higgs really give mass to everything: not just to the *W* and *Z* particles but to quarks and charged leptons as well? That really would be the key to a unified theory embracing the gamut of particle physics. But even just knowing the role of the Higgs in breaking electroweak symmetry will allow us to gain understanding of that symmetry and add consistency to quantum field theory. For the sake of its own consistency, the standard model needs something like a Higgs.

The Higgs is not, of course, the be-all and end-all of the LHC. There is also the question of supersymmetry (SUSY). SUSY is a theory that maintains that every particle in the fermion enclosure of the particle zoo (the quarks, the leptons and the composite particles with an odd number of half-integer spins) has a heavier twin in the boson enclosure (where photons, gluons, and the *W* and *Z* particles currently reside). Thus, the electron (a fermion) has a supersymmetric boson partner, known as a 'selectron', and so on. Theorists love SUSY for her elegance. The LHC will allow us to establish whether SUSY exists or not: even if 'squarks' and 'gluinos' are as heavy as 2.5 TeV, the LHC will find them.

And then there is the question of the extra space dimensions predicted by string theory — that herculean attempt to unify quantum theory and gravitation. For these new dimensions to exist, yet for us to be unaware of them, they must be 'curled up' incredibly small. Theoretically, some might be just big enough to be detected at the LHC through the escape of (gravitational) energy into them.

A speculative laundry list

To me, these three factors — the Higgs particles, supersymmetric particles and new dimensions — are the discoveries most likely to emerge from the first five or so years of LHC operations. But there is a long, more speculative laundry list of objects that might be illuminated by the powerful beams of the LHC. Most of these are speculative in the extreme.

Dark matter origins

Dark matter is one cosmological discovery that has shaken up particle physics, giving rise to many a joint conference with an 'inner space/outer space' theme. The rotational speed of galaxies requires more gravitational 'stuff' than is accounted for by the shining stars. Measurements during the past decade have yielded the information that about 25% of the Universe's mass must be this dark stuff. Neutrinos, which were the initial prime suspects because huge quantities of them were known to have been left over from the Big Bang, are not massive enough. Over time, other exotic candidates — dead suns, black holes and large planets (known as Jupiters) — have been ruled out.

Theorists have supplied us with a plethora of possible solutions, mostly out of their bag of supersymmetric particles. What is known about dark matter is that, first, there is lots of it; second, it does not shine; and, third, it has gravitational force. It is certainly possible that particles will emerge from the collisions of the LHC that will both gladden the hearts of SUSY theorists and account for dark matter.

Dark energy origins

This is, potentially, the elephant in the control room of the LHC. We haven't a clue as to what it is, but we know what it does: it maintains a continuous outward push on the matter of the Universe, sustaining and increasing the expansion rate, and thereby counteracting the gravitational attraction that should be slowing the expansion. It might not be dark, and it might not be energy. But it accounts for more than 70% of the mass of the Universe, so its identification is an important objective. Illumination by the LHC would be a seminal discovery.

Compositeness

Increasingly precise experiments, in the spirit of Ernest Rutherford's scattering experiments on the substructure of gold atoms, have attempted to detect some sort of substructure to the quintessential electron, using progressively more powerful microscopes, each capable of 'seeing' objects smaller than its predecessor: 10^{-18}, 10^{-19}, 10^{-20} cm. This is the maximum scale for an electron's radius (and therefore any internal structure it might have). By necessity, we are now comfortable with the hypothesis that all standard-model particles have zero radius and so no substructure. But this doesn't preclude a future machine detecting a finite size: the higher its energy, the smaller the domain searched.

It is also possible to imagine a sort of substructure that would escape detection by scattering experiments. If one were ever to detect a quark so structured that it could have a higher energy quantum state, and so might absorb energy from the scattered protons, that would be a just-fancy-that moment!

Technicolour

'Technicolour' refers not to the glowing shades in which the fantasy land appears in *The Wizard of Oz* but rather to the quantum field theory postulated as an alternative to the Higgs hypothesis for explaining the masses of the *W* and *Z* particles. Theories that have not yet been confirmed experimentally are judged by their mathematical 'elegance' and their economy in predicting new particles. Supersymmetry predicts a doubling in the total number of particles and must therefore be considered uneconomical. From the point of view of the ambitious experimental physicist desperate to make discoveries, however, SUSY is a godsend. Technicolour predicts a new strong force and a large number of new particles (although fewer than SUSY), whose 'signatures' could stand out above backgrounds of the complex collisions that the LHC will produce.

Strong scattering

One of the wonders of the Higgs hypothesis for high-energy physics is that it cures a particular 'pathology' present in certain predictions: for example, there are infinities in the cross-sections (a measure of the probability of a process) for the scattering of two *W* bosons. The presence of the Higgs would cure this 'disease'. Thus, experiments such as *W*–*W* scattering (I don't know how to do these, but I am sure there are experiments that would include infinite contributions in a Higgs-less theory) should be carried out at the highest energies. If Higgs cannot be discovered, such experiments will be crucial to establish what the missing ingredient that we need to make our theories sensible looks like.

New gauge bosons

Our old force carriers are photons, the *W* and *Z* particles, and gluons. It is strongly assumed that gravitons are 'almost' discovered — although not by suspicious conservatives. Perhaps, a decade into the LHC's operation, our skills in precise analysis of collisions at 14 TeV will have been honed such that we can discover a new force, and so a new boson, predicted by a theorist now in a good high school. At present, our theories don't need such bosons, but that doesn't mean they don't exist.

Right-handed neutrinos

The neutrinos we know and love have less than one millionth the mass of the electron and are left-handed — that is, their spin direction is opposite to their momentum. To be accepted gracefully into current theory, a right-handed neutrino — the spin and momentum of which would be parallel — must be very massive. In addition, we must lose the distinction between the massive neutrino and its antimatter twin. A discovery of such particles at the LHC would be a fantastic step forwards in our quest for a theory of everything. It would, for example, have a bearing on the 'origin of matter' dilemma — that is, why is there a small excess of matter over antimatter, and, by extension, how did we come to be here?

Mini black holes

Black-hole physics deals with the astronomical phenomenon of a massive sun using up its nuclear fuel and eventually collapsing, if it is heavy enough, into a black hole. Of more interest to particle physicists are smaller black holes, left over from the Big Bang, which may well exist in and around our Galaxy. At first pass, even such mini black holes must be much more massive than any imaginable accelerator could reach. But the existence of extra dimensions of finite size, as proposed by string theory, would lower the energy required to produce these hypothetical particles. The idea of the LHC as a mini-black-hole factory is not as worrying as it sounds; they will quickly evaporate through the radiation of energy (Hawking radiation).

What did we leave out?

Theoretical physicists are an imaginative group, and each of these exotic suggestions has its proponents and its naysayers. But the history of the sort of step that the LHC will be making teaches us that, more often than not, a discovery will be made that was not anticipated by theorists. That discovery will change our theories beyond imagination. Fifty years spent investigating the standard model have taught me that, by year ten of the LHC's physics, many an expected and unexpected discovery could well have been celebrated with champagne drunk from styrofoam cups. ■

Leon Lederman is at the Department of Biological, Chemical, and Physical Sciences, Illinois Institute of Technology, 3300 South Federal Street, Chicago, Illinois 60616-3793, USA.

Author Information Reprints and permissions information is available at npg.nature.com/reprintsandpermissions. The author declares no competing interests. Correspondence should be addressed to the author (lederman@fnal.gov).

The making of the standard model

Gerard 't Hooft

A seemingly temporary solution to almost a century of questions has become one of physics' greatest successes.

The standard model of particle physics is more than a model. It is a detailed theory that encompasses nearly all that is known about the subatomic particles and forces in a concise set of principles and equations. The extensive research that culminated in this model includes numerous small and large triumphs. Extremely delicate experiments, as well as tedious theoretical calculations — demanding the utmost of human ingenuity — have been essential to achieve this success.

Prehistory

The beginning of the twentieth century was marked by the advent of two new theories in physics[1]. First, Albert Einstein had the remarkable insight that the laws of mechanics can be adjusted to reflect the principle of relativity of motion, despite the fact that light is transmitted at a finite speed. His theoretical construction was called the special theory of relativity, and for the first time, it was evident that purely theoretical, logical arguments can revolutionize our view of nature. The second theory originated from attempts to subject the laws found by James Maxwell for the continuum of electric and magnetic fields to the laws of statistical mechanics. It was Max Planck who first understood how to solve this: the only way to understand how heat can generate radiation is to assume that energy must be quantized. This theory became known as quantum mechanics.

At first, it was thought that quantum mechanics would apply only to atoms and the radiation emitted by their electrons. But, gradually, it became clear that the laws of quantum mechanics had to be completely universal to make sense. This idea of universality was in common with Einstein's theories of relativity. In particular, quantum mechanics had to apply not only to electrons but also to the particles that reside in atomic nuclei.

It was clear, right from the beginning, that the two new theoretical constructions would need to be combined into one. The vast amounts of energy found to inhabit atomic nuclei implied that 'relativistic quantum mechanics' had to apply to atomic nuclei in particular. Thus, a new problem became evident and soon garnered worldwide attention: how is quantum mechanics reconciled with special relativity? This question kept physicists busy for most of the rest of the century, and it was not completely answered until the standard model saw the light of day.

The early days

By 1969, the reconciliation of quantum mechanics with special relativity was still a central issue[2], but much more had been discovered through experimental observation[3]. Matter particles (see page 270) had been divided into leptons and hadrons. The known leptons were the electron, the muon and their two neutrinos (these last assumed to be massless); hadrons, such as protons and pions, obeyed the conservation laws of quantum numbers known as 'strangeness' and 'isospin'. Hadrons are divided into mesons, which can be described loosely as an association of a quark and an antiquark, and baryons, which can be simply depicted as being made up of either three quarks or three antiquarks. The symmetry of strong interactions between subatomic particles was known to be approximated by the 'eightfold way' (Fig. 1). And it seemed that all hadrons had infinite series of excited states, in which angular momentum was bounded by the square of the mass measured in units of ~1 gigaelectronvolt (Fig. 2). This feature of all hadrons was telling us something important about strong interactions, but the first attempts to understand it consisted of rather abstract formalisms.

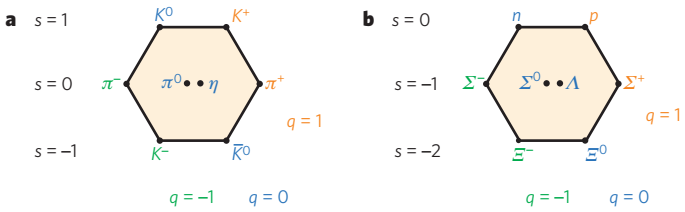


Figure 1 | The eightfold way. Spin-zero mesons (**a**) and spin-half baryons (**b**) can be grouped according to their electric charge, *q*, and strangeness, *s*, to form octets (which are now understood to represent the flavour symmetries between the quark constituents of both mesons and baryons).

It was also known that there are weak forces and electromagnetic forces, to which subatomic particles owe some of their properties. However, only the electromagnetic force was understood in sufficient detail for extremely precise calculations to be checked against accurate experimental observations. Theorists had tried to devise methods to subject not only the electromagnetic force but also other forces to the laws of quantum mechanics and special relativity. Despite their efforts over nearly 50 years, attempts to improve this 'quantum field theory' to include weak interactions failed bitterly. And describing the strong interactions between mesons and baryons drove them to despair.

The theorists at that time therefore concluded that quantum field theory should be dismissed as a possible way of addressing the dynamics of particle interactions. We now know that this was a misjudgement. Their mistrust of quantum fields was, however, understandable: in all known quantum field systems, there were divergences in the high-energy domain, making these systems unsuitable for describing strong interactions. Yet it was clear that strong interactions, such as those that hold a nucleus together, do exist. The error made by the theorists was that this 'bad' high-energy behaviour was thought to be an unavoidable, universal feature of all quantum field theories[4].

Because of this widespread objection to quantum field theories, few theorists ventured to investigate field theoretical methods. They should have realized that their objections could be swept away when the forces are weak. Indeed, the weak force was the first subatomic force to be formulated using the new 'gauge theories'[2]. Such theories had been proposed in 1954 by Chen Ning Yang and Robert Mills (Fig. 3), who were inspired by the fact that the two basic forces of nature that were well understood, gravity and electromagnetism, were both based on the principle of local gauge invariance: that is, that symmetry transformations can be performed in one region of space-time without affecting what happens in another. This beautiful idea got off to a slow start, even after Peter Higgs, François Englert and Robert Brout realized in 1964 how the structure of the vacuum can be modified by the field of a scalar (spin-zero) particle, which came to be called the Higgs particle. With the inclusion of the Higgs particle, the Yang–Mills field equations could

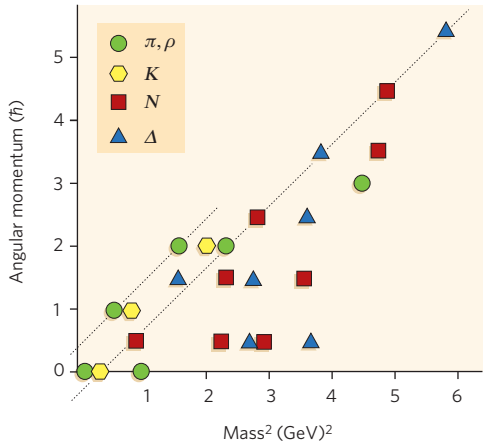


Figure 2 | A hint at the nature of the strong force. All strongly interacting particles and their excited states seem to have an angular momentum (in units of the reduced Planck constant $\hbar$) that is less or about equal to the square of their mass (measured in gigaelectronvolts, GeV). The limits for various particle species form lines that seem to be straight and parallel. *N*, nucleon (which includes neutrons and protons).

now be used to describe the weak force accurately; the force would be carried by the quanta of the Yang–Mills field, which had gained mass by this 'Brout–Englert–Higgs mechanism'. Reasonably realistic models in which exactly this happens were proposed by Abdus Salam, Sheldon Glashow and Steven Weinberg in the 1960s.

The 1970s

In 1971, Martinus Veltman and I demonstrated that it is exactly these theories (in which the mass of the Yang–Mills gauge quanta is attributed to the field of a Higgs particle) that are 'renormalizable', and it seems that this was all that was needed for a full rehabilitation of quantum field theory to begin[4]. Renormalization is the mathematical description of the requirement for distinguishing, at a fundamental level, the algebraic mass terms and coupling terms in the equations from the actual physical masses and charges of the particles. The choice of values for these algebraic parameters depends crucially on the smallest distance scales taken into account in the theory. So, if it were insisted that all particles are truly point-like — that is, the smallest distance scale should be zero — then these algebraic parameters would need to be infinite. The infinite interactions were needed to cancel the infinitely strong self-interactions of particles that the equations inevitably lead to. But the mathematical procedure of cancelling infinite forces against one another needed to be performed with considerable care. Many theorists did not understand what was going on and aired their strong suspicions that 'all this' had to be 'rubbish'.

We were learning not only how to construct realistic and logically coherent models but also how to study the behaviour of these theories over short distances by using the concept of the 'renormalization group'. Introduced by Ernst Stückelberg and André Petermann in 1953, this mathematical procedure allows one to go from one distance scale to another. It is used both in condensed-matter theory and in elementary particle physics, for which it was pioneered by Curtis Callan and Kurt Symanzik. A function that can be computed for every theory, named β-function by Callan and Symanzik, determines what might happen: if β is positive, the strengths of the couplings are increased at shorter distances; if β is negative, they are weakened. The error that I mentioned earlier was that all quantum field theories were thought to have positive β-functions. Indeed, it was claimed that this could be proved. Owing to various miscommunications, earlier calculations that yielded negative β-functions (including calculations by me) were systematically ignored, until in 1973, David Politzer, David Gross and Frank Wilczek published their findings that, for Yang–Mills theories, β is generally negative. Therefore, the strength of interactions would be reduced at short distances, making them controllable — a property that was named asymptotic freedom. Until this point, Yang–Mills theories had been understood to describe only electromagnetic and weak interactions. But the discovery of asymptotic freedom immediately turned Yang–Mills theory into a prime candidate for describing strong interactions as well.

In fact, experimental observations had been pointing in the same direction. A Yang–Mills structure not only fitted beautifully with the algebraic symmetries that had been established for the strong force (such as the eightfold way) but also could be deduced from observations made at the Stanford Linear Accelerator Center (SLAC), in California, where strong interactions seemed to show scaling behaviour, as though their strength diminished at short distances (known as Bjorken scaling)[4]. Indeed, theorists had concluded that no quantum field theory would be suitable for the strong force — until the asymptotic freedom of Yang–Mills fields was uncovered.

Basically, Yang–Mills fields are a generalization of the electromagnetic field, for which Maxwell had determined the equations a century earlier. Particles carry a generalized type of electric charge, allowing them not only to be accelerated by the Yang–Mills fields but also to be transmuted into other kinds of particle under the influence of these fields. Thus, electrons can transform into neutrinos, protons into neutrons and so on, as a result of the weak force. The strong force is understood as a new kind of field acting on quarks, which are the building blocks of protons and neutrons inside the atomic nuclei. In addition to ordinary electric charges, quarks also carry a threefold charge, which is reminiscent of colour vision (and hence they are usually called red, green and blue). For this reason, the Yang–Mills theory for the strong force is called quantum chromodynamics, the Greek word *chromos* meaning colour.

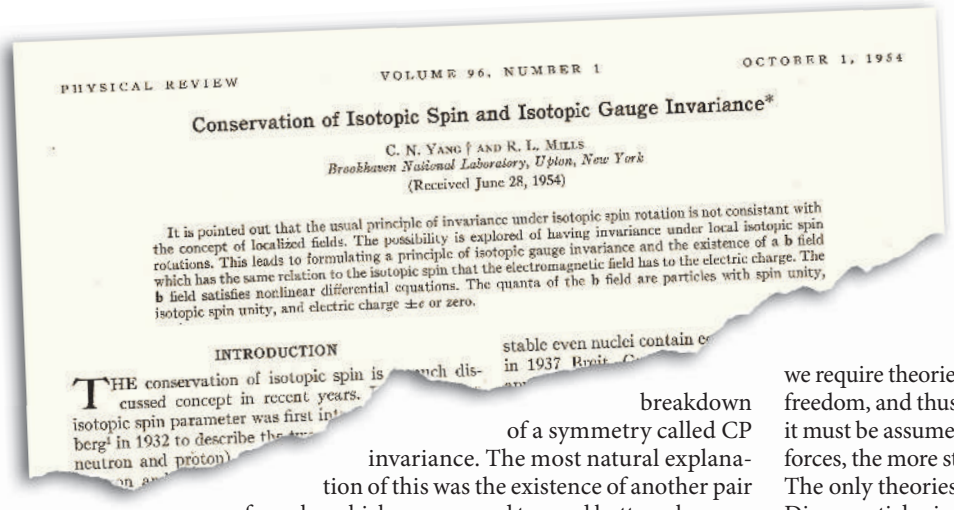
Getting the details right

For the first time, all known particles and all known forces between them could be cast in a single model. This model described three closely related Yang–Mills systems for the three major forces (strong, weak and electromagnetic), one Higgs field and several matter fields. These matter fields were Dirac fields, describing the four known leptons and the three known quarks (up, down and strange), all of which have half a unit of spin. According to this theory, the Dirac particles cannot interact directly with one another but only by exchanging energy quanta of the Yang–Mills field. The interactions between Yang–Mills fields and matter fields are identical for all particle types; only the Higgs field couples differently to the different matter fields. And only in this way is differentiation brought about between the various kinds of particle according to this new insight. By breaking the symmetry of the vacuum, the Higgs field could also give masses to the Yang–Mills quanta. But even the Higgs field is allowed to have only a limited number of interaction coefficients, so this model had only a small number of adjustable parameters: the masses of the quarks and leptons and a handful of 'mixing parameters'. The gravitational force, being excessively weak when acting between individual particles, could be included only to the extent that it acts classically.

The early versions of this model had other deficiencies. One of these was the remarkable absence of interactions due to the exchange of the neutral component, the Z boson, of the weak Yang–Mills quanta (the charged components being the W^+ and W^- bosons). These 'neutral current interactions' were detected for electrons and neutrinos in pivotal experiments at CERN in 1973 (Fig. 4). But they should also have caused strangeness-changing interactions among hadrons, and the existence of these was excluded by experimental observations. A possible remedy to this problem had already been proposed by Glashow, John Iliopoulos and Luciano Maiani in 1969, but this required a drastic revision of the model: the addition of a fourth quark, which was named charm.

The discovery of a series of new particles in 1974, beginning with the J/ψ particle at SLAC and at Brookhaven National Laboratory (the Alternating Gradient Synchrotron, in Upton, New York), marked a revolution of sorts. These new particles contained the elusive charm quark. Furthermore, their properties dramatically confirmed quantum chromodynamics and asymptotic freedom.

More details were then added. A rare type of transition observed in a special type of K meson called a K_L meson seemed to imply



PHYSICAL REVIEW VOLUME 96, NUMBER 1 OCTOBER 1, 1954

Conservation of Isotopic Spin and Isotopic Gauge Invariance*

C. N. YANG† AND R. L. MILLS
Brookhaven National Laboratory, Upton, New York
(Received June 28, 1954)

It is pointed out that the usual principle of invariance under isotopic spin rotation is not consistant with the concept of localized fields. The possibility is explored of having invariance under local isotopic spin rotations. This leads to formulating a principle of isotopic gauge invariance and the existence of a b field which has the same relation to the isotopic spin that the electromagnetic field has to the electric charge. The b field satisfies nonlinear differential equations. The quanta of the b field are particles with spin unity, isotopic spin unity, and electric charge ±e or zero.

INTRODUCTION

THE conservation of isotopic spin is ...cussed concept in recent years. ... isotopic spin parameter was first int... berg[1] in 1932 to describe the ... neutron and proton) ... stable even nuclei contain e... in 1937 Breit ...

Figure 3 | Yang–Mills gauge theory. The field equations introduced by Chen Ning Yang and Robert Mills in 1954 became the basis for the three forces of the standard model — electromagnetic, weak and strong. Image reproduced, with permission, from ref. 7.

breakdown of a symmetry called CP invariance. The most natural explanation of this was the existence of another pair of quarks, which were named top and bottom, because only the delicate interplay of at least six quarks with the Higgs field could give rise to the observed CP breakdown. Mathematical consistency of the scheme also required the existence of more leptons. The tau lepton, and its associated neutrino, were discovered and confirmed around 1978. The bottom quark (in 1977) and, finally, the top quark (in 1995) were also proved to exist.

Thus, a picture emerged of three generations, each containing two species of lepton and two species of quark. All of these particles interact with three types of Yang–Mills field and one Higgs field and, of course, with gravity. This picture was subsequently referred to as the standard model. In the 1970s, it was generally thought that this standard model would merely be a stepping stone. Amazingly, however, no improvements seemed necessary to explain the subsequent series of experimental observations. The standard model became a 'standard theory' — an accurate and realistic description of all of the particles and forces that could be detected.

One further detail did need to be added. The standard model was originally designed to accommodate only strictly massless neutrinos, but there was one anomaly — the neutrino flux from the Sun[5]. Pivotal observations announced in 1998, made using the Kamiokande detector, in Japan, showed that neutrinos can mix and therefore must have mass. Adding neutrino mass terms to the standard model was, however, only a minor repair and not totally unexpected, although it did add more parameters to the model. The earlier version had 20 fundamentally freely adjustable constants (parameters) in it; now, this number would need to be increased to at least 26.

Super theories

By the 1980s, it was understood that quantum field theories are perfect frameworks for the detailed modelling of all known particles. Indeed, if we require theories with only a limited number of elementary degrees of freedom, and thus a finite number of freely adjustable parameters, then it must be assumed that all forces are renormalizable. But, for all strong forces, the more stringent condition of asymptotic freedom is required. The only theories with these desired properties are theories in which Dirac particles interact exclusively with Yang–Mills fields and (where needed) with Higgs fields. This is now regarded as the answer to that problem of more than half a century ago — how to reconcile quantum mechanics with special relativity.

The mere fact, however, that these three Yang–Mills field systems are based on exactly the same general gauge principle, acting on the same sets of Dirac particles, has inspired many researchers not to stop here but to search for more common denominators. Can we find a completely unified field theory? Such theories have been sought before, notably by Einstein and by Werner Heisenberg in their later years, but their efforts were bound to fail because the Yang–Mills theories were then unknown. Now, it seems that we have the key to doing a much better job.

Indeed, we do have clues towards constructing a unified field theory. Despite its stunning successes, there are weaknesses in the standard model. Mathematically, the model is nearly, but not quite, perfect. Also, from a physics point of view, there are problematic features. One is the occurrence of gigantic differences in scale: some particles are extremely heavy, whereas others are extremely light. At distance scales that are short compared with the Compton wavelength of the heaviest particles — the cut-off scale below which field theories become important for these particles — there seems to be a crucial 'fine-tuning' among the effective couplings. And, most importantly, the quantum effects of the gravitational force are not included. These issues are the focus of new generations of theoretical proposals[6]. Might there be a new symmetry — a 'supersymmetry' — between Dirac particles and the force-carrying particles? Might particles turn out to be string-like rather than point-like? Or will a new generation of particle accelerators reveal that quarks and leptons are composites?

In the strongest possible terms, as theorists, we now urge our friends in experimental science to do whatever they can to obtain further information on the properties of nature's building blocks at the tiniest possible scales. In our business, this means reaching for the highest attainable energies: the Large Hadron Collider will make such a step. We can hardly wait. ■

Gerard 't Hooft is at the Institute for Theoretical Physics, Utrecht University and the Spinoza Institute, Post Office Box 80.195, 3508 TD Utrecht, The Netherlands.

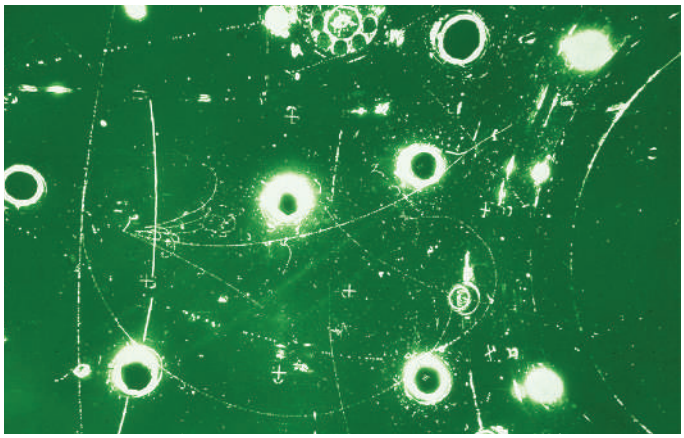


Figure 4 | The neutral current. An image from the heavy-liquid bubble chamber Gargamelle, at CERN, in 1973. The curling tracks reveal the interaction of a neutrino with a nucleon through the neutral current of *Z* exchange. Image reproduced with permission from CERN.

1. Pais, A. *Niels Bohr's Times, in Physics, Philosophy, and Polity* (Clarendon, Oxford, 1991).
2. Crease, R. P. & Mann, C. C. *The Second Creation: Makers of the Revolution in Twentieth-Century Physics* (Macmillan, New York, 1986).
3. Källén, G. *Elementary Particle Physics* (Addison-Wesley, Reading, Massachusetts, 1964).
4. Hoddeson, L., Brown, L. M., Riordan, M. & Dresden, M. (eds) *The Rise of the Standard Model: a History of Particle Physics from 1964 to 1979* (Cambridge Univ. Press, Cambridge, 1997).
5. Bahcall, J. N. *Neutrino Astrophysics* (Cambridge Univ. Press, Cambridge, 1989).
6. Ross, G. G. *Grand Unified Theories* (Perseus, Reading, Massachusetts, 2003).
7. Yang, C. N. & Mills, R. L. Conservation of isotopic spin and isotopic gauge variance. *Phys. Rev.* **96,** 191–195 (1954).

Author Information Reprints and permissions information is available at npg.nature.com/reprintsandpermissions. The author declares no competing financial interests. Correspondence should be addressed to the author (g.thooft@phys.uu.nl).

NATURE|Vol 448|19 July 2007|doi:10.1038/nature06075

High-energy colliders and the rise of the standard model

Terry Wyatt[1]

Over the past quarter of a century, experiments at high-energy particle colliders have established the standard model as the precise theory of particle interactions up to the 100 GeV scale. A series of important experimental discoveries and measurements have filled in most of the missing pieces and tested the predictions of the standard model with great precision.

The standard model of particle physics describes the Universe as being composed of a rather small number of different types of elementary particle (see page 270) that interact in a small number of well-defined different ways.

Interactions among the elementary particles are represented by Feynman diagrams such as those in Fig. 1a. These show the annihilation of an electron–positron (e^+e^-) pair to produce a fermion–antifermion pair (such as a quark–antiquark or lepton–antilepton pair), and such interactions are examples of the 'electroweak' interaction, which is propagated by the photon, $W^{\pm}$ and Z bosons. All of the fermions participate in the electroweak interaction; certain 'self-interactions' among the photon, W and Z bosons may also take place.

Quarks, but not leptons, also participate in the strong interaction, which is propagated by gluons and described by the theory of quantum chromodynamics (QCD). Collectively, quarks and gluons are referred to as 'partons'. Quarks may carry one of three 'colours', which in the strong interaction are the analogue of charge; antiquarks carry the equivalent anticolour. A particular feature of the strong interaction is that coloured quarks cannot exist as free particles for more than about 10^{-24} s. The particles we observe in our detectors are hadrons — collections of quarks and/or antiquarks that have no net colour. There are two basic types of hadron: mesons contain a quark and an antiquark (of opposite colour); baryons contain three quarks (one of each colour). When a high-energy quark or gluon is produced, it is observed as a collimated 'jet' of hadrons.

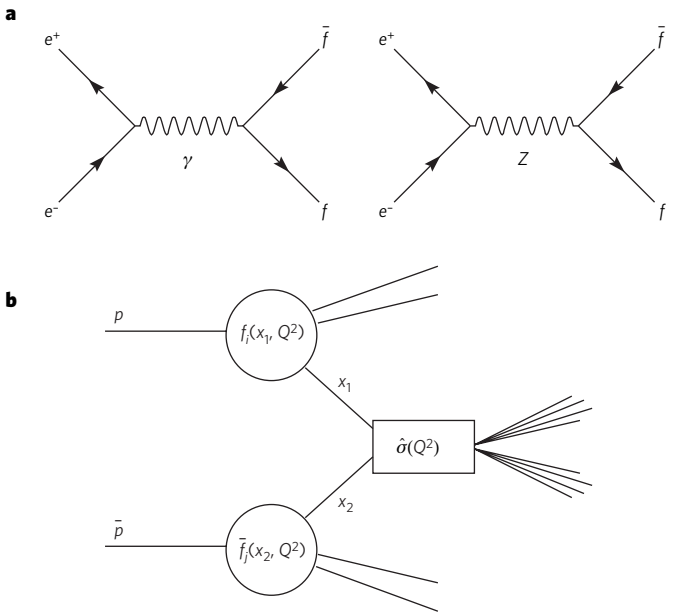


Figure 1 | Particle interactions. a, The lowest-order Feynman diagrams for the process $e^+e^- \rightarrow f\bar{f}$, where f is any elementary fermion (quark or lepton). **b**, Schematic view of a high-energy $p\bar{p}$ collision. Part b reproduced, with permission, from ref. 16.

The Higgs mechanism is introduced into the standard model to allow elementary particles to have non-zero masses, through their interaction with the Higgs field, while maintaining the gauge invariance of the model. A consequence of including the Higgs mechanism is that a massive, spin-zero Higgs boson is also predicted to exist.

If the mathematical structure of the standard model is taken as a given (although, of course, it represents a considerable amount of empirical input!), then all particle couplings are predicted in terms of a relatively small number of 'free' parameters that must be determined by experiments. For example, the strong interaction is determined by the value of a single coupling constant, denoted α_s. In the electroweak sector, the physically observed photon and Z boson arise from a linear superposition of two hypothetical particles: the W^0, the electrically neutral partner of the $W^{\pm}$, and another neutral boson, B^0 (of the so-called 'hypercharge' interaction). A rotation angle is defined between the W^0/B^0 and Z/photon, known as the electroweak mixing angle, θ_W, which describes the relative strengths of the electromagnetic and weak interaction. The interactions of the photon, Z and W are then determined by three free parameters. Logically, these can be thought of as the coupling constants of the weak and hypercharge interactions and the electroweak mixing angle. The masses of the W and Z bosons can also be predicted in terms of these parameters (with the photon and gluon required to be massless by gauge invariance). The masses of the 12 fermions and the Higgs boson are not predicted and thus represent additional free parameters that must be determined by experiment.

A particular feature of the electroweak interactions is that the couplings of the fermions to the W and Z depend on their handedness or helicity. The $W^{\pm}$ couples only to left-handed (negative-helicity) fermions and right-handed (positive-helicity) antifermions. The Z couples to both left- and right-handed fermions, but with a different coupling constant in each case.

In simple terms, the basic aims of particle physics are to find direct experimental evidence for each of the elementary particles and to make as precise as possible measurements of their various properties

[1]Particle Physics Group, School of Physics and Astronomy, University of Manchester, Manchester, UK.

(masses, coupling strengths and so on). Because the number of different experimental measurements that can be made is much larger than the number of free parameters in the standard model, we are dealing with an 'over-constrained' system. That is, our experimental measurements not only determine the values of the free parameters of the standard model, they also provide stringent tests of the consistency of the model's predictions.

Electrons versus protons

For more than a quarter of a century, the high-energy frontier of particle physics has been dominated by experiments performed at particle–antiparticle colliders. In these accelerators, beams of electrons and positrons, or protons (p) and antiprotons ($\bar{p}$), travel with equal and opposite momenta and collide head-on in the centre of the particle detectors.

Experiments at electron colliders have several advantages over those at proton colliders, which stem from the fact that electrons are elementary particles. When an e^+e^- pair annihilates, the initial state is well defined and, if the pair collide at equal and opposite momentum, the centre-of-mass energy of the system (E_{cm}) is equal to the sum of the beam energies. E_{cm} is the energy available to produce the final-state particles.

Electrons participate only in the electroweak interaction. This means that the total e^+e^- annihilation cross-section is small, so event rates in experiments are low, but essentially every annihilation event is 'interesting', and the observed events are relatively simple to analyse. Initial-state *bremsstrahlung* (radiation from the beam particles) can reduce the available centre-of-mass energy, but because this is a purely electromagnetic process it can be calculated with great precision, and it introduces no significant systematic uncertainties into the analysis of annihilation events.

The disadvantage of using electrons as beam particles is their small rest mass. When high-energy electrons are accelerated, they lose energy (producing synchrotron radiation), and that energy loss must be compensated by the machine's accelerating cavities. The energy radiated by a charged particle in performing a circular orbit of radius, R, is proportional to γ^4/R, where γ is the ratio of the particle's total energy to its rest mass, m_0c^2. Even though the world's largest particle accelerator, the Large Electron–Positron Collider (LEP), at CERN, had a circumference of 27 km, its maximum beam energy of around 104 GeV was limited by the fact that each particle radiated about 2 GeV per turn. By contrast, the large rest mass of the proton means that synchrotron energy loss is not a significant limiting factor for proton–antiproton colliders. For example, the world's highest energy collider at present is the Tevatron proton–antiproton collider, at Fermilab (Batavia, Illinois), which, with a circumference of only 6 km, achieves a beam energy of 1,000 GeV (or 1 TeV); the Large Hadron Collider (LHC), using two proton beams in the 27-km LEP tunnel, will achieve beam energies of 7 TeV.

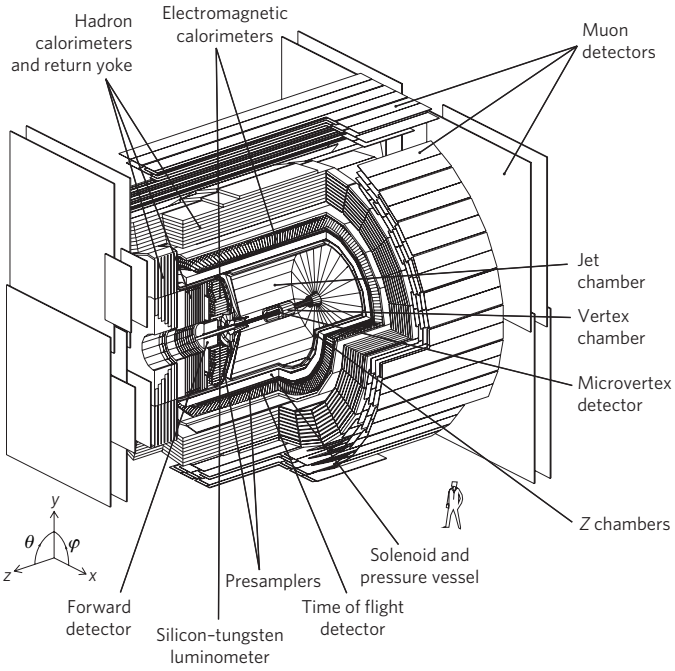


Figure 2 | The OPAL experiment at LEP. The typical, hermetic design of this detector comprises central track detectors inside a solenoid, calorimeters and — the outermost layers — muon detectors.

Although the beam energies of proton colliders may be much higher, for experiments at these colliders there are a number of challenges that stem from the fact that protons and antiprotons are strongly interacting, composite particles. A high-energy proton–antiproton collision is shown schematically in Fig. 1b. The highest energy collisions take place between a valence quark from the proton and an antiquark from the antiproton. These colliding partons carry fractions x_1 and x_2 of the momentum of the incoming proton and antiproton, respectively. The energy, Q, in the parton–parton centre-of-mass frame is given by $Q^2 = x_1x_2E_{cm}^2$. The probability of a proton containing a parton of type i at the appropriate values of x_1 and Q^2 is given by a 'parton distribution function' (PDF), $f_i(x_1, Q^2)$. The cross-section for the parton–parton collision to produce a given

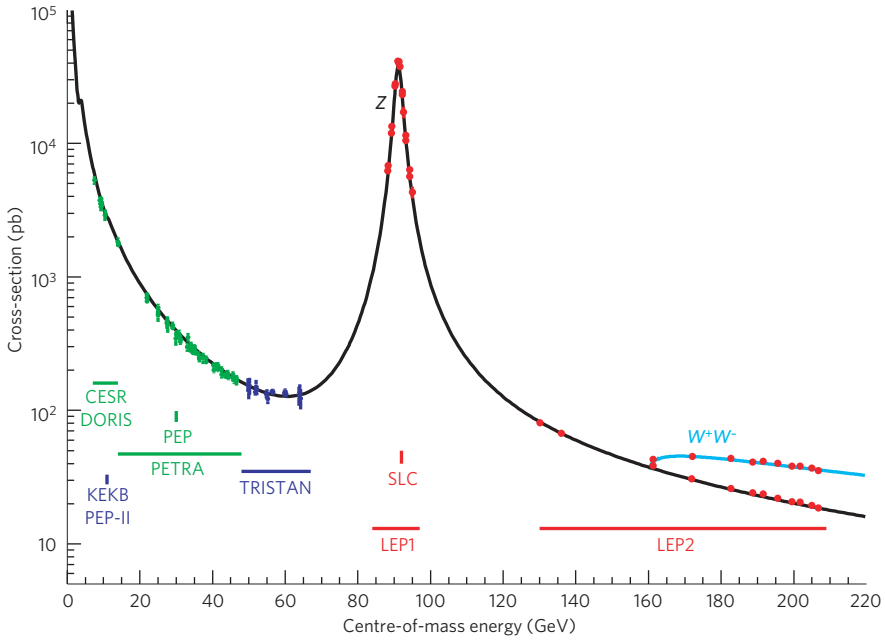


Figure 3 | The cross-section for e^+e^- annihilation to hadrons as a function of E_{cm}. The solid line is the prediction of the standard model, and the points are the experimental measurements. Also indicated are the energy ranges of various e^+e^- accelerators. (A cross-section of 1 pb = 10^{-40} m^2.) Figure reproduced, with permission, from ref. 12.

final state is denoted by $\hat{\sigma}(Q^2)$. To determine the cross-section, σ, for the proton–antiproton collision to produce this final state, we have to sum over all possible combinations of incoming partons and integrate over the momentum fractions x_1 and x_2:

$$\sigma = \sum_{i,j=q,\bar{q},g} \int dx_1 dx_2 f_i(x_1, Q^2) \cdot \bar{f}_j(x_2, Q^2) \cdot \hat{\sigma}(Q^2)$$

Therefore, the proton and antiproton beams, at a fixed beam energy, can be thought of as broadband beams of partons.

The total cross-section for proton–antiproton collisions at high energy is huge, and the event rate is consequently large — at the Tevatron, for example, about 10 collisions take place each time the bunches of protons and antiprotons meet and cross each other in the circular machine. Such bunch crossings take place 1.7 million times each second. But most of these collisions are rather uninteresting, because they result from a low momentum transfer between the proton and antiproton. Interesting processes, such as those containing W or Z bosons, are produced at a much lower rate and can be difficult to observe above the huge background.

Furthermore, the PDFs cannot be calculated from first principles in QCD. They can, however, be fixed by experimental measurements. A great deal of information on PDFs has come from the H1 and ZEUS experiments at the HERA collider, at DESY (Hamburg). At HERA, 27.5-GeV beams of electrons or positrons collide with a 920-GeV beam of protons, to produce 320 GeV in the centre-of-mass frame. The electrons and positrons provide a clean (electroweak-interaction) probe of the proton structure, and hence the PDFs, at these energies; the measured PDFs can then be extrapolated, using, for example, the so-called DGLAP evolution equations of QCD, to the much higher energies that are relevant at the Tevatron and the LHC.

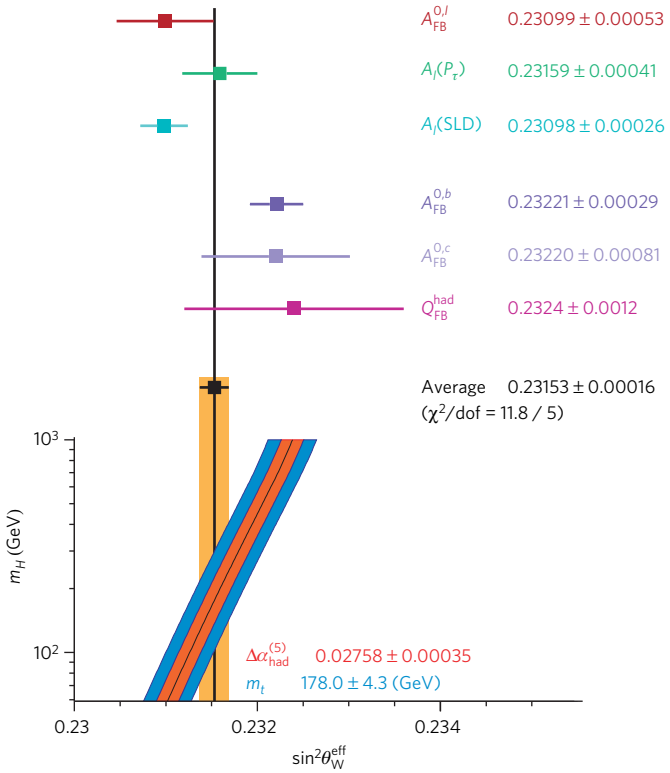


Figure 4 | Comparison of the effective electroweak mixing angle, $\sin^2\theta_W^{eff}$, derived from six classes of asymmetry measurements. These six classes are: A_{FB}, from leptonic final states at LEP; A_l (P_τ), derived from τ polarization measurements at LEP; A_l (SLD), derived from A_{LR} and from leptonic-final-state A_{LRFB} measurements at the SLC; and A_{FB}, from $b\bar{b}$ and $c\bar{c}$ final states at LEP. Q_{FB}^{had} is an average forward–backward charge asymmetry in hadronic events at LEP, without any attempt to distinguish individual quark flavours. Also shown is the standard-model prediction for $\sin^2\theta_W^{eff}$ as a function of m_H. The additional uncertainty of the standard-model prediction is parametric and dominated by the uncertainties in $\Delta\alpha_{had}^{(5)}(m_Z^2)$ (the correction for the effects of light-quark loops) and m_t, shown as bands. The total width of the band is the linear sum of these effects. Figure reproduced, with permission, from ref. 12. dof, degrees of freedom.

A further complication is that the initial-state partons have a high probability of radiating gluons before they collide. To some extent, this can be compensated by tuning Monte Carlo simulations of the collisions to those events that include leptonically decaying W and Z bosons (in which there is no complication from the possibility of final-state gluon *bremsstrahlung*). Nevertheless, the uncertainties associated with the lack of precise predictions for initial-state gluon *bremsstrahlung* represent a significant source of sytematic uncertainty in many analyses.

Proton and electron colliders are thus complementary: proton colliders offer the energy reach to make discoveries; electron colliders provide a cleaner experimental environment in which it is easier to make precise measurements.

Experiments at high-energy particle colliders typically share many common features, which are motivated by the requirements of the various measurements to be made. The basic aims are to detect with high efficiency each particle produced in the high-energy collision, to measure as accurately as possible its energy and momentum and to determine its particle type. No single detector type can achieve all of the above for all types of particle. Therefore, an experiment comprises a number of different detector systems, each of which has a specialized function. For example, at the centre of most experiments are detectors that measure the tracks produced by charged particles. Calorimeters are used for energy measurement, and muon systems are used for specific identification of those particles. An important feature of such detectors is their hermetic nature, which allows any apparent imbalance in the net transverse momentum of the visible particles to be ascribed to the production of weakly interacting particles, such as neutrinos. Fig. 2 shows, as an example, a cut-away view of the OPAL experiment at LEP, which is typical of detector design.

Discoveries and mounting evidence

By the late 1970s, the majority of the elementary fermions had been discovered. In particular, the discovery in the mid-1970s of the bottom quark and the tau lepton firmly established the existence of a third generation of fermions. However, there was only indirect evidence for the existence of two members of that generation: the top quark and the tau neutrino.

By contrast, among the elementary bosons only the photon had been observed directly as a physical particle. Although there was strong indirect evidence for the existence of the gluon, the first direct evidence came from the observations in 1979 by the JADE, Mark-J, TASSO and PLUTO experiments at the 30–35-GeV e^+e^- collider PETRA, at DESY. These experiments found events containing three hadronic jets, which correspond to the quark and antiquark produced in the e^+e^- collision, plus a gluon radiated from one of the quarks. The W and Z bosons were observed directly for the first time in 1983, by the UA1 and UA2 experiments[1–4] at the 560–640-GeV Super Proton Synchroton (SPS) proton–antiproton collider at CERN — a collider project that was conceived for the specific purpose of finding these particles and was rewarded with the 1984 Nobel Prize in Physics. The masses of the W and Z measured by the UA1 and UA2 experiments were found to be consistent with expectations, which was a beautiful confirmation of the standard model in electroweak interactions.

The scene was then set in 1989 for the 90-GeV e^+e^- colliders, LEP1 at CERN and SLC at the Stanford Linear Accelerator Center (SLAC; California). The ALEPH, DELPHI, L3 and OPAL experiments at LEP1 and the SLD experiment at the SLC performed measurements of Z production and decay that still today form the cornerstone of the precise tests of the electroweak standard model. Measurements of the Z mass elevated it to one of the most precisely known quantities within the standard model. Measurements of the total decay width of the Z (to all possible particle types) and the partial decay widths into each visible final state (that is, all final states except for $Z \rightarrow \nu\bar{\nu}$) allowed the number

of light neutrino species to be fixed at three. This observation effectively confirmed the existence of the tau neutrino as a distinct physical particle within the three-generation standard model and ruled out the existence of a fourth generation of fermions, unless the neutrino from that generation has a mass greater than half that of the Z. (Direct observation of the tau neutrino was finally reported[5] in 2001, in a different style of experiment using a proton beam directed at a fixed, tungsten target to produce neutrinos.)

Experimenters at all of the high-energy colliders since the days of PETRA had searched unsuccessfully for direct evidence for the existence of the top quark. These searches continued at 'Run I' of the 1.8 TeV Tevatron, which began in 1988. By 1994, the lower limit on the top quark mass from direct searches had reached[6,7] about 130 GeV. By that time, there was also considerable indirect evidence for the existence of the top quark. For example, measurements of the electroweak couplings of the bottom quark were consistent with the hypothesis that it formed one half of a pair of third-generation quarks within the standard model. Furthermore, fits to the precise electroweak data from LEP1 and SLC gave self-consistent results within the standard model only if the effects of a top quark with a mass of between about 155 and 195 GeV were included.

The top quark was directly observed for the first time in 1995, by the CDF and DØ collaborations[8,9] at the Tevatron. The first measurements gave its mass as 180 ± 15 GeV, consistent with the indirect determinations described above. This consistency represented a powerful confirmation of the electroweak standard model as an accurate picture of elementary particle physics.

In the second phase of the LEP programme, running between 1996 and 2000 with a vastly upgraded system of radio-frequency (RF) accelerating cavities, a maximum E_{cm} of nearly 209 GeV was reached. This allowed the production of W^+W^- pairs, enabling the ALEPH, DELPHI, L3 and OPAL experiments at LEP2 to measure the mass, m_W, and many other properties of the W boson.

In 2002, the second phase of running, Run II, began at the Tevatron. The accelerator complex was upgraded to deliver a slightly higher E_{cm} of 1.96 TeV and, more importantly, a greatly increased luminosity; the CDF and DØ detectors were also upgraded.

Now the most urgent question in particle physics (maybe in physics as a whole) is: where is the Higgs? Just as with the top quark, this question is being attacked on two fronts. Adding information from the direct measurements of the mass of the W boson and the top quark (from LEP2 and the Tevatron) to the precise electroweak measurements (from LEP1 and SLC) improves the precision with which the standard model can be tested. The overall fit gives self-consistent results only if the effects of a moderately light Higgs boson are included. Currently, a value for the Higgs mass of about 80 GeV is preferred, with an upper limit[10], at a 95% confidence level, of 144 GeV; further improvements to the mass measurements from the Tevatron may narrow the confidence interval. Direct searches for the Higgs boson were performed at LEP. The best available direct lower limit[11] on the Higgs mass is currently 114 GeV (95% confidence level) from the search for $e^+e^- \rightarrow ZH$ at LEP2. This already excludes a large part of the confidence interval allowed by the standard-model fit. Direct searches for Higgs production are currently the subject of intense effort at the Tevatron, and sensitivity to masses beyond the LEP2 limit is expected in the near future.

Precise tests of the standard model

A central part of the particle physics programme over the past quarter of a century has been to test the consistency of the standard model through precise measurement of many of its parameters. Precise theoretical calculations, implemented through computer codes of high technical precision, and a careful assessment of residual theoretical uncertainties are also essential elements in efforts to confront the standard model using precise data.

Let us return to the simple process shown in Fig. 1a, the annihilation of an e^+e^- pair to produce a fermion–antifermion pair through an electroweak interaction mediated by the photon or Z boson. Particles that

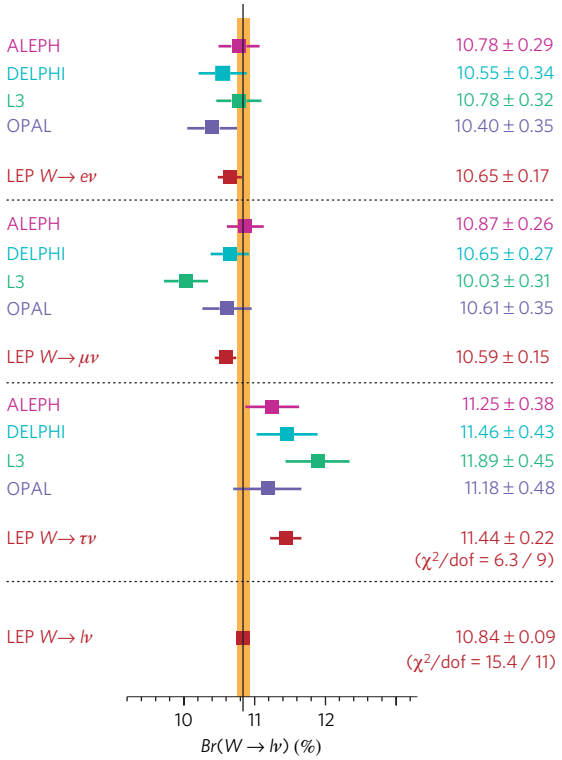


Figure 5 | Leptonic branching ratios. The measurements of the four LEP experiments of branching ratios for W decays to $e\nu$, $\mu\nu$ and $\tau\nu$ final states separately and for all lepton types combined. Figure reproduced, with permission, from ref. 14.

appear as internal lines in a Feynman diagram, such as the photon or Z in Fig. 1a, are 'virtual' particles — that is, they are not constrained to their 'physical' mass. However, the more virtual the particle becomes — the further away it is from its physical mass — the smaller the resultant amplitude for the process. Fig. 3 shows the cross-section for e^+e^- annihilation as a function of centre-of-mass energy, E_{cm}, based on data from several colliders including LEP and SLC. At low values of E_{cm}, the cross-section is dominated by the photon-exchange diagram (an exchanged Z would be highly virtual and the corresponding amplitude highly suppressed). With increasing E_{cm}, the cross-section falls as the exchanged photon becomes more and more virtual. At around 60 GeV, the amplitudes for photon and Z exchange are of comparable magnitude. As E_{cm} approaches the mass of the Z (91 GeV), the cross-section is dominated by the Z exchange diagram and reaches a peak, called the 'Z pole'.

The very large number of Z decays (around 20 million) collected by the experiments at LEP1 has allowed precise measurements of the couplings of the fermions to be made. The SLC delivered a much smaller number of Z decays (around 600,000) to the SLD experiment. However, the SLC delivered a longitudinally polarized e^- beam, which collided with an unpolarized e^+ beam, whereas at LEP both beams were unpolarized. The dependence on handedness of the fermion couplings has enabled SLD to make measurements, using polarized beams, that were in some respects competitive with and complementary to the measurements made at LEP1. (The results quoted in this section are all taken from ref. 12 unless explicitly stated otherwise.)

A number of important electroweak quantities have been determined from measurements around the Z pole at LEP1. The mass of the Z, m_Z, is related to the position of the peak in the cross-section, and total decay width of the Z, Γ_Z, is related to the width of the peak. The accuracy with which $m_Z = (91.1875 \pm 0.0021)$ GeV has been measured is limited by the accuracy with which the mean energy of the colliding beams is known over the entire data-taking period. Achieving such precision was a considerable challenge and resulted from a successful collaboration between physicists from both the LEP experiments and the accelerator. The energy

of a single circulating beam was determined to a high accuracy during dedicated calibrations, using the technique of resonant depolarization. However, such calibrations could be performed only every few days and gave the beam energy only at that specific point in time. The challenge was to propagate this precise knowledge of the beam energy over several days of accelerator running.

The circumference of the beam orbit is fixed by the frequency with which the RF accelerating cavities are excited. This frequency is very stable. The energy of the beams is then determined by the integral around the accelerator ring of the vertical component of the magnetic field experienced by the beams. This vertical magnetic field is produced mainly by the main 'bending' dipole magnets, but there is also a contribution from the large number of quadrupole magnets in the machine if the beam is not perfectly centred as it passes through them. If the position of the beam with respect to the quadrupoles changes over a period of hours or days this can affect the beam energy by a significant amount. Lunar tides, high rainfall in the nearby Jura mountains and changes in the water level of Lake Geneva all caused sufficient physical distortion of the accelerator (changing its radius by a few parts in 10^{-9}) to produce a measureable effect on the beam energy.

Erratic electric currents flowing in the accelerator beam pipe also affected the dipole fields over periods of many hours during which beams were circulating in the accelerator. Measurements of the spatial distribution of these currents around the ring established that they were produced by leakage currents from trains running on the Geneva-to-Bellegarde line. Understanding these various effects meant that a model could be developed to predict the beam energy as a function of time during data collection. Ultimately, residual uncertainties in the beam-energy calibration introduced systematic uncertainties of 0.0017 GeV in m_Z and 0.0012 GeV in Γ_Z, correlated among the four experiments.

The total decay width, $\Gamma_Z = (2.4952 \pm 0.0023)$ GeV, is given by the sum of the partial decay widths for each possible type of final-state fermion–antifermion pair. By measuring Γ_Z and the partial decay widths for each visible final state (quarks and charged leptons), the partial decay width to invisible final states (which in the standard model are neutrino–antineutrino pairs) can be determined. This number may be interpreted as a measurement of the number of types of light neutrino produced in Z decay, $N_\nu = 2.9840 \pm 0.0082$. This result requires the measurement of absolute cross-sections. These require a precise determination of the 'luminosity' of the accelerator, which is achieved by measuring the rate of low-angle electron–positron scattering. That the necessary precision of order 10^{-4} was achieved in these measurements represents a great success for theorists and experimentalists engaged in this joint project.

The rate of Z decays to quark–antiquark final states is enhanced by a factor related to α_s, the strong coupling constant, $(1 + \alpha_s^2/\pi + \ldots)$. Thus, a precise measurement of α_s can be made: $\alpha_s = 0.118 \pm 0.03$. This is in agreement with other precise determinations[13], such as those from event shapes (which are sensitive to the amount of final-state gluon radiation), and represents an important consistency test of QCD.

Asymmetries

Another class of electroweak measurement made at LEP1 and the SLC is of various asymmetries that are sensitive to the difference between the left- and right-handed couplings. One of the most sensitive of these electroweak measurements, and also one of the easiest to understand, is the so-called left–right asymmetry, A_{LR}. This is measured with polarized e^- beams at the SLC and is defined as:

$$A_{LR} = \frac{\sigma_L - \sigma_R}{\sigma_L + \sigma_R}$$

where σ_L (σ_R) is the cross-section for any given final state with a 100% left-hand (right-hand) polarized incoming electron beam. In practice, 100% polarization is not achievable, but it can be easily shown that if the magnitude of the (luminosity-weighted) average e^- beam polarization is $\langle P_e \rangle$ then the measured asymmetry, A_{LR}^{meas}, is given by $A_{LR}^{meas} = \langle P_e \rangle A_{LR}$. At the SLC, $\langle P_e \rangle = 70$–80% was regularly achieved.

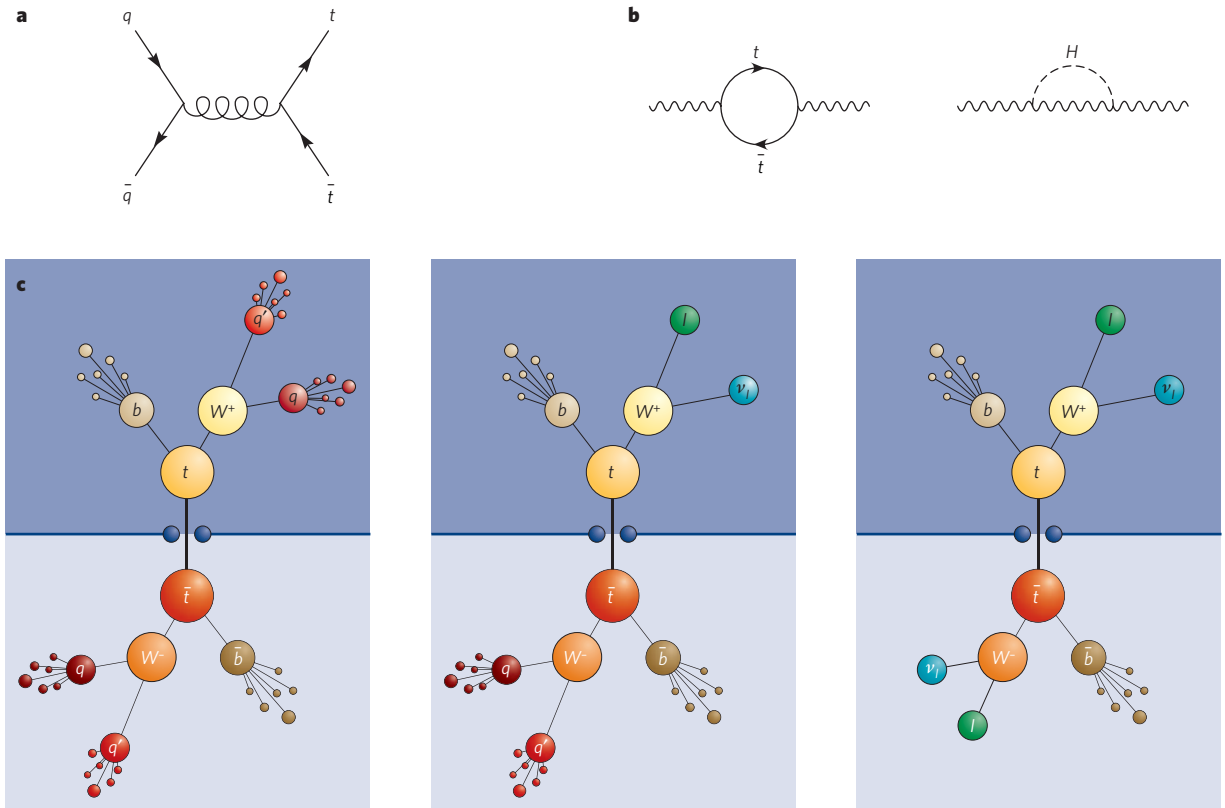


Figure 6 | Top-quark production, and virtual loops. a, The Feynman diagram for $q\bar{q}$ annihilation to produce a $t\bar{t}$ pair. **b**, Virtual loops involving t quarks and Higgs bosons. The left-hand diagram may modify a process involving the propagation of a photon or Z; the right-hand, the propagation of a W or Z. **c**, The possible event signatures for $t\bar{t}$ production. From left to right, panels show 'all-jets', 'lepton + jets' and 'di-lepton'.

The advantage of defining A_{LR} as above is that many factors — such as the dependence on the final-state couplings, acceptance of the detector, and so on — cancel in the ratio (as long as the experimental acceptance and $<P_e>$ are independent of the sign of the beam polarization). For such measurements at the Z pole, corrections (which are usually small) must be made to account for the photon-exchange diagram (Fig. 1a) and for the interference between the photon- and Z-exchange diagrams. In addition, a correction has to be applied for the fact that *bremsstrahlung* from the incoming e^+e^- results in an average annihilation centre-of-mass energy that is lower than the nominal E_{cm} of the colliding beams. Results are corrected to correspond to $E_{cm} = m_Z$, and these 'pole' cross-sections and asymmetries are therefore to be interpreted as corresponding to pure Z exchange at exactly $E_{cm} = m_Z$; they are sometimes denoted by adding the superscript '0' to the corresponding variable name, for example, A_{LR}^0.

As we have seen above, the fact that left- and right-handed e^- have different couplings from the Z produces an asymmetry between the annihilation cross-section for left- and right-hand-polarized incoming e^- beams. In addition, the difference between the left- and right-handed fermion couplings produces asymmetries in the angular distributions of the outgoing fermions. Consider an incoming e^- beam that is 100% left-hand polarized: angular-momentum conservation requires that this can annihilate only with the right-handed component of the incoming e^+ beam to produce Zs that are 100% polarized in the direction opposite to the incoming e^- beam. Angular-momentum conservation in the decay of the Z has the consequence that the preferred direction for the outgoing fermions to emerge is along the direction of the incoming e^- beam (the 'forward' direction) for left-handed fermions and in the opposite direction (the 'backward' direction) for right-handed fermions.

Using polarized electrons, as at the SLC, it is possible to define the 'left–right forward–backward' asymmetry,

$$A_{LRFB} \equiv \frac{(\sigma_F - \sigma_B)_L - (\sigma_F - \sigma_B)_R}{(\sigma_F + \sigma_B)_L + (\sigma_F + \sigma_B)_R}$$

As before, the measured asymmetry, A_{LRFB}^{meas}, is given by $A_{LRFB}^{meas} = <P_e> A_{LRFB}$.

At LEP, the e^- and e^+ beams were unpolarized. That is, there were equal numbers of left- and right-handed incoming beam particles. Nevertheless, the fact that left- and right-handed e^- have different couplings to the Z produces an asymmetry between the numbers of left- and right-hand incoming e^- that annihilate. Thus, the produced Zs are partially polarized along the direction of the incoming beams, and the difference between the left- and right-handed fermion couplings produces a forward–backward asymmetry, A_{FB}, in the angular distributions of the outgoing fermions, which is given by:

$$A_{FB} \equiv \frac{(\sigma_F - \sigma_B)}{(\sigma_F + \sigma_B)}$$

The forward–backward asymmetry with unpolarized beams, A_{FB}, mixes the couplings of the initial- and final-state particles. This makes A_{FB} intrinsically a less sensitive measure of the electroweak mixing angle, θ_W, (in the form $\sin^2\theta_W$) than measurements possible with polarized beams. However, the much larger samples of Zs available at the LEP experiments compensate for this lack of intrinsic sensitivity.

To measure A_{LRFB} and A_{FB}, it is necessary to isolate a sample of Z decays to a particular fermion type and to distinguish the fermion from the antifermion. In the case of Z decays to charged leptons this is fairly straightforward: events containing a high-momentum e^-e^+, $\mu^-\mu^+$ or $\tau^-\tau^+$ pair may be readily distinguished from one another and from other backgrounds; the electric charge distinguishes the lepton from the antilepton. In the case of Z decays to quarks, precise measurements of A_{FB} are only really possible in the $c\bar{c}$ and $b\bar{b}$ final states.

In most cases it is not possible to determine the handedness of the final-state particles (hence observables are usually summed over this quantity). The one exception is for final-state tau leptons, where the momenta of the observed tau decay products are correlated with the handedness of the produced tau.

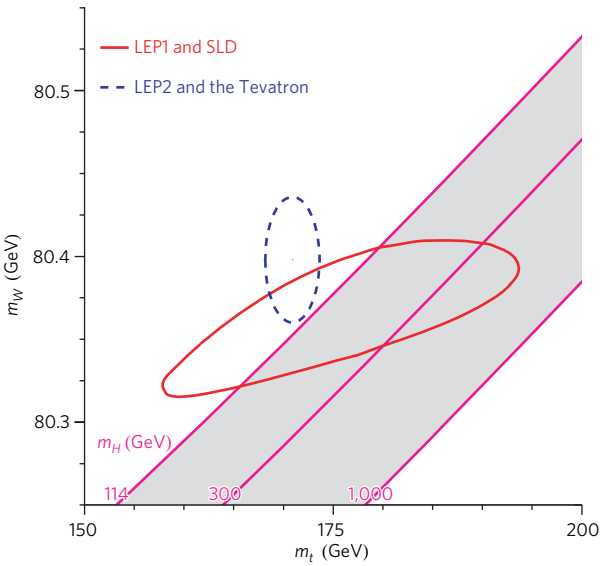


Figure 7 | Contours at 68% confidence level showing the direct (LEP2 and the Tevatron) and indirect (LEP1 and SLC) measurements of m_W and m_t. The shaded band shows the predictions of the standard model for various values of m_H. Figure reproduced, with permission, from ref. 10.

All of the asymmetry measurements discussed are sensitive to the difference between the left- and right-handed fermion couplings and thus to the value of $\sin^2\theta_W$. The degree to which the different classes of asymmetry measurements yield consistent values of $\sin^2\theta_W$ — as illustrated in Fig. 4 —represents an important consistency check of the standard model.

Consistency of the standard model

In W^+W^- events at LEP2, the value of m_W is obtained by directly reconstructing the invariant mass of the pair of particles produced in the W decay. In principle, the two final states with high branching ratios — $q\bar{q}l\bar{\nu}$ and $q\bar{q}q\bar{q}$ — give similar statistical sensitivity. However, in the $q\bar{q}q\bar{q}$ channel, uncertainties associated with strong interactions and Bose–Einstein correlations between the products of the two hadronically decaying Ws render the measurement of m_W in this channel less precise. The combination of results[14] from the four LEP experiments yields $m_W = (80.376 \pm 0.033)$ GeV. Other properties of the W (such as the branching ratios shown in Fig. 5) were measured[14] at LEP2.

At the Tevatron, only the leptonic decays $W \to e\nu$ and $W \to \mu\nu$ can be used to measure m_W. CDF has produced the first preliminary measurement of m_W using the Run II data accumulated so far, and it has an uncertainty to match that of a single LEP experiment. Including data from Run I, the Tevatron average[10] is $m_W = (80.429 \pm 0.039)$ GeV. Combining the LEP and Tevatron values gives the 'world average'[10] as $m_W = (80.398 \pm 0.025)$ GeV.

The most important process for producing top quarks in $p\bar{p}$ collisions is shown in Fig. 6a. The dominant decay of the top quark is $t \to Wb$ and possible signatures of $t\bar{t}$ production are shown schematically in Fig. 6c. If one W decays leptonically and one W decays hadronically, a final state is produced containing a high-transverse-momentum lepton, missing transverse momentum (due to the undetected neutrino) and four high-transverse-momentum jets. This occurs in about 46% of $t\bar{t}$ pairs produced, and this so-called 'lepton + jets' channel yields the most precise measurement of m_t. The combination[15] of CDF and DØ measurements gives $m_t = (170.9 \pm 1.8)$ GeV. This precision of around 1% makes m_t by far the most precisely known quark mass. The ultimate precision expected for the Tevatron measurements is around 20 MeV for m_W and around 1 GeV on m_t; to equal such precision at the LHC will take much time and concerted effort.

It is interesting to understand how experiments can produce evidence for the existence of a particle, and even constrain its mass and couplings,

even though they have insufficient energy to produce the particle directly. The indirect effects of the top quark and the Higgs boson may be observed at LEP/SLC because of the existence of processes such as those shown in Fig. 6b. The possibility of such 'radiative corrections' modifies the simple 'lowest-order' picture of e^+e^- annihilation in Fig. 1a, and experimentally observable effects become sensitive to the masses and couplings of virtual particles in such loops. For example, it is usual to consider $\sin^2\theta_W^{\text{eff}}$, an 'effective' parameter that absorbs the effect of the radiative corrections but allows the basic form of the coupling equations involving $\sin^2\theta_W$ to stay the same. The correction to $\sin^2\theta_W$ can be calculated in the standard model; it depends on the square of the top quark mass, m_t, but only logarithmically on the Higgs mass, m_H. An illustration of these effects is given in the lower half of Fig. 4, in which the experimentally measured value of $\sin^2\theta_W^{\text{eff}}$ is compared with the prediction of the standard model as a function of m_H.

The contours in Fig. 7 show the world-average direct measurements of m_W and m_t compared with the indirect values of those quantities extracted from the standard-model fit to the LEP and SLC data. The shaded band shows the predictions of the standard model for various values of m_H. The fact that the direct and indirect values of m_W and m_t agree is a triumph of the standard model.

An even more stringent test of the consistency of the standard-model fit to all available high-energy electroweak data is shown in Fig. 8: each measured quantity is compared with its value obtained from the fit. The largest single deviation is seen for $A_{FB}^{0,b}$ (the forward–backward asymmetry for Z decays to bottom quarks) measured at LEP1, but, particularly given the number of measurements considered, a discrepancy of 2.8 standard deviations in one of them does not meet the threshold required for claiming a significant departure from the standard model.

The increased samples of $t\bar{t}$ events available at Run II have allowed measurements of the cross-section for $t\bar{t}$ production and of t quark properties, such as spin, electric charge and decay branching ratios, that are

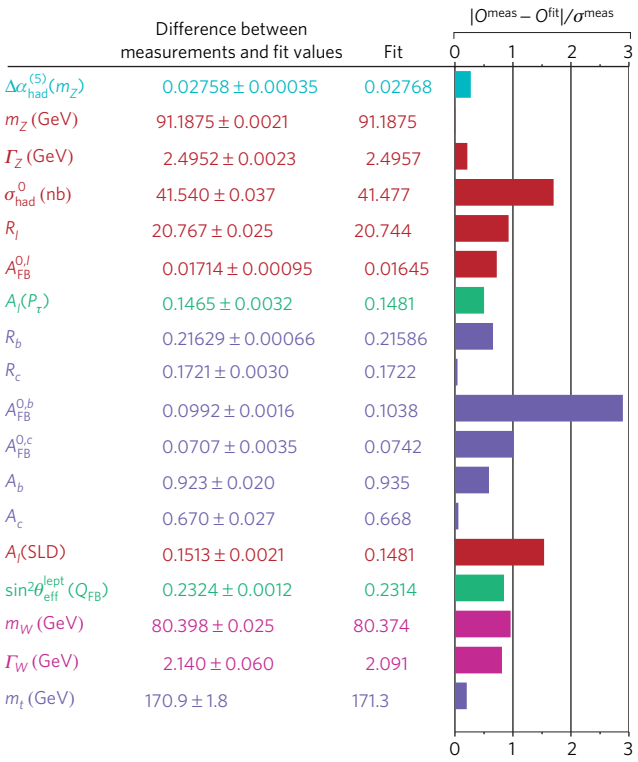


Figure 8 | A test of the consistency of the standard-model fit to all available high-energy electroweak precise data. Each measured observable (O^{meas}) quantity is compared with the value obtained from the fit (O^{fit}). Also shown graphically is the difference between measurement and fit values in number of standard deviations. Colours indicate groups of similar variables. Figure reproduced, with permission, from ref. 10. For full definitions of each quantity, see ref. 10.

consistent with those expected in the standard model[16]. The Tevatron experiments are also detecting processes with ever smaller cross-sections— which bodes well for developing the sensitivity of the searches for the Higgs boson at this collider. The CDF experiment has detected the associated production of WZ pairs[17] and has found the first evidence at a hadron collider for the production of ZZ pairs[18]; the DØ experiment has found the first evidence for electroweak production of single top quarks[19], enabling the first direct determination of the $t \to Wb$ coupling.

The years ahead

The next few years will be an exciting time in experimental particle physics, with first collisions at the 14 TeV proton–proton collider, the LHC, scheduled for 2008. Until then, as the world's current highest-energy collider, the Tevatron has a monopoly on direct searches for new physics at a high-mass scale and can perform the most stringent tests of the point-like nature of the fundamental particles.

The Tevatron will run at least until late 2009; its mantle will not pass to the LHC overnight. Except for a few special cases that could produce the most spectacular, unmistakable signatures, it will take time to understand and calibrate the LHC accelerator and detectors.

It is hard to imagine that new physics beyond the standard model will not be found at the LHC. What form that new physics will take is harder to imagine. We know from the past 30 years' work that all theories predicting any observable effects beyond the predictions of the standard model were quickly disposed of by experiment. This means that no matter what is to come, the standard model will remain at least an extremely accurate 'approximation' to the physics of elementary particles at scales up to a few hundred GeV. ■

1. Arnison, G. *et al.* (UA1 Collaboration). Experimental observation of isolated large transverse energy electrons with associated missing energy at $\sqrt{s} = 540$ GeV. *Phys. Lett. B* **122,** 103–116 (1983).
2. Banner, M. *et al.* (UA2 Collaboration). Observation of single isolated electrons of high transverse momentum in events with missing transverse energy at the CERN $\bar{p}p$ collider. *Phys. Lett. B* **122,** 476–485 (1983).
3. Arnison, G. *et al.* (UA1 Collaboration). Experimental observation of lepton pairs of invariant mass around 95 GeV/c^2 at the CERN SPS collider. *Phys. Lett. B* **126,** 398–410 (1983).
4. Bagnaia, P. *et al.* (UA2 Collaboration). Evidence for $Z^0 \to e^+e^-$ at the CERN p collider. *Phys. Lett. B* **129,** 130–140 (1983).
5. Kodama, K. *et al.* (DONUT Collaboration). Observation of tau neutrino interactions. *Phys. Lett. B* **504,** 218–224 (2001).
6. Abachi, S. *et al.* (DØ Collaboration). Top quark search with the DØ 1992–1993 data sample. *Phys. Rev. D* **52,** 4877–4919 (1995).
7. Abachi, S. *et al.* (DØ Collaboration). Search for the top quark in $p\bar{p}$ collisions at $\sqrt{s} = 1.8$ TeV *Phys. Rev. Lett.* **72,** 2138–2142 (1994).
8. Abe, F. *et al.* (CDF Collaboration). Observation of top quark production in $p\bar{p}$ collisions with the collider detector at Fermilab. *Phys. Rev. Lett.* **74,** 2626–2631 (1995).
9. Abachi, S. *et al.* (DØ Collaboration). Observation of the top quark. *Phys. Rev. Lett.* **74,** 2632–2637 (1995).
10. LEP Electroweak Working Group. *LEP Electroweak Working Group*. <http://lepewwg.web.cern.ch/LEPEWWG/> (2007).
11. ALEPH Collaboration, DELPHI Collaborations, L3 Collaboration, OPAL Collaboration and The LEP Working Group for Higgs Boson Searches. Search for the standard model Higgs boson at LEP. *Phys. Lett. B* **565,** 61–75 (2003).
12. ALEPH, DELPHI, L3, OPAL, SLD collaborations, LEP Electroweak Working Group, the SLD Electroweak and Heavy Flavour Groups. Precision electroweak measurements on the Z resonance. *Phys. Rep.* **427,** 257–454 (2006).
13. Jones, R. W. L. Final α_s combinations from the LEP QCD working group. *Nucl. Phys. B Proc.* (suppl.) **152,** 15–22 (2006).
14. The LEP Collaborations: ALEPH Collaboration, DELPHI Collaboration, L3 Collaboration, OPAL Collaboration, the LEP Electroweak Working Group. A combination of preliminary electroweak measurements and constraints on the standard model. Preprint at <http://arxiv.org/abs/hep-ex/0612034> (2006).
15. Tevatron Electroweak Working Group (for the CDF and DØ Collaborations). A combination of CDF and DØ results on the mass of the top quark. Preprint at <http://arxiv.org/abs/hep-ex/0703034v1> (2007).
16. Quadt, A. Top quark physics at hadron colliders. *Eur. Phys. J. C* **48,** 835–1000 (2006)
17. Abulencia, A. *et al.* (CDF Collaboration). Observation of WZ production. Preprint at <http://arxiv.org/abs/hep-ex/0702027>.
18. CDF Collaboration. Evidence for ZZ production in $p\bar{p}$ at $\sqrt{s} = 1.96$ TeV. CDF Note 8775 (preliminary). <http://fcdfwww.fnal.gov/physics/ewk/2007/ZZ/ZZ_comb_public_note.ps>.
19. Abazov, V. M. *et al.* (DØ Collaboration). Evidence for production of single top quarks and first direct measurement of |Vtb|. *Phys. Rev. Lett.* **98,** 181802 (2007).

Author information Reprints and permissions information is available at npg.nature.com/reprintsandpermissions. The author declares no competing financial interests. Correspondence should be addressed to T.W. (twyatt@fnal.gov; Terry.Wyatt@manchester.ac.uk).

How the LHC came to be

Chris Llewellyn Smith

Approval of a project the size of the Large Hadron Collider is an exercise in politics and high finance.

The idea of following CERN's Large Electron–Positron Collider (LEP) with a Large Hadron Collider (LHC), housed in the same tunnel, dates back at least to 1977, only two years after LEP itself was conceived. The importance of not compromising the energy of an eventual LHC was one of the arguments for insisting on a relatively long tunnel in the discussions that led to the approval of LEP in 1981.

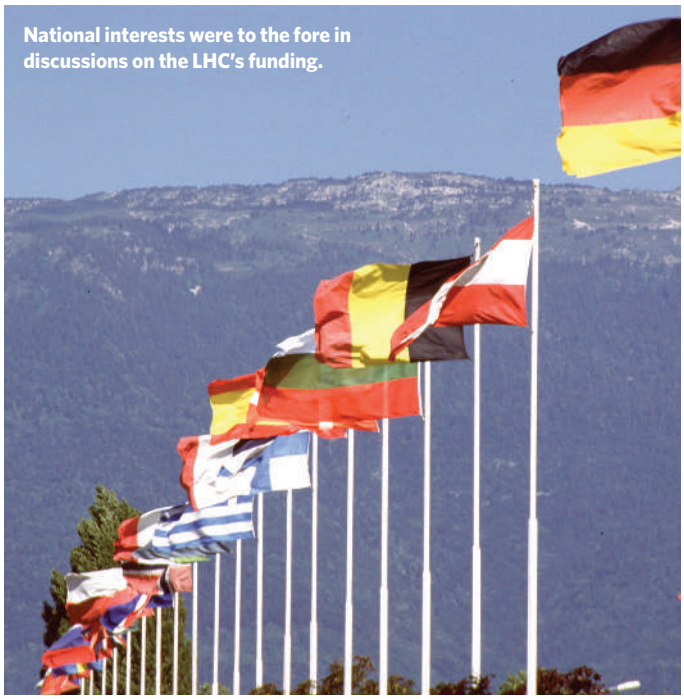
Early discussions of the LHC were dominated by sometimes acrimonious competition and comparisons with the proposed 40 teraelectronvolt (TeV) Superconducting Super Collider (SSC) in the United States. Serious work on the SSC was kick-started by American reactions to the discovery of the carriers of the weak force, the $W^{\pm}$ and Z bosons, at CERN in 1983. CERN's discovery was greeted by a *New York Times* editorial entitled "Europe 3, US Not Even Z-Zero", and a call from the President's science adviser for the United States to "regain leadership" in high-energy physics.

Viewed from Europe, this was provocative. 40 TeV is more than twice the energy that could possibly be reached by a hadron collider installed in the LEP tunnel, and many Europeans suspected that this was why 40 TeV was chosen. Furthermore, CERN's leadership favoured the next really large accelerator being an inter-regional facility. There was agreement on the vital importance of being able to explore new phenomena of up to 1 TeV — the energy below which the Higgs boson, or whatever else generates the mass of all particles, should be discovered. But how much better this could be done at the SSC than at the lower energy LHC was hotly debated. The energy domain that can be explored by a hadron collider is less than that of the accelerated particles, which is shared between their constituents. However, the much higher intensity (or luminosity, in the terminology of particle physics) planned for the LHC could in principle compensate for it having lower energy than the SSC — with denser bunches of accelerated particles there is a greater chance of collisions between constituents with large fractions of their parents' energy — although another decade of intensive research and development was needed to establish that experiments are possible at such high luminosity.

"Not everyone was convinced that the collider was justified, or would ever be funded"

The extreme European view was that the SSC was irresponsible as it would cost five times as much as the LHC without providing much more physics. A global plan, on the other hand, would provide complementary facilities — a large linear electron–positron collider in the United States and the LHC at CERN, which would use the LEP tunnel and other existing infrastructure — and would therefore be much cheaper than the SSC. A typical American response was to refute the claim that the LHC could do much the same physics for one-fifth of the cost, and to meet with scepticism any concern for American taxpayers. Meanwhile, senior Japanese physicists who argued that the SSC should be international were told that it was a national facility. They did not forget this when the United States later asked Japan to contribute US$2 billion.

I thought the extreme European position was unrealistic. The United States wanted a new project that could reach 1 TeV as soon as possible and a large hadron collider was the only realistic option at the time. The technology was not then available to build a 1 TeV linear collider: the 0.1 TeV linear collider at Stanford, the world's first, had not been built, and even the International Linear Collider now being proposed will initially reach only 0.5 TeV. On the other hand, my experience as adviser to the Kendrew Committee, which was then considering whether the United Kingdom should remain in CERN, had made me acutely conscious of growing pressure on funding for particle physics. I doubted that the SSC — which seemed profligate to me, despite its enormous potential — would ever be funded.



National interests were to the fore in discussions on the LHC's funding.

CERN

In fact, the SSC was endorsed by President Ronald Reagan in January 1987, with a price tag of $4.4 billion. In May 1990, when the cost had risen to $7.9 billion, the House of Representatives voted to limit the federal contribution to $5 billion, with the rest to be found from the state of Texas ($1 billion) — the proposed site of the SSC — and from overseas (where none was found). The SSC was defeated in the House in June 1992, but later revived by the Senate; this happened again in June 1993, by which time the General Office of Accounting estimated the cost as $11 billion. It was cancelled in October 1993.

The right machine for the future

Meanwhile at CERN, research and development started on the very demanding LHC magnets in 1988. This was recommended by a Long Range Planning Group chaired by Carlo Rubbia, who shared the 1984 Nobel Prize in physics for the discovery of the W and Z bosons, and who became Director-General of CERN in 1989. Rubbia argued that the LHC would provide healthy competition for the SSC at a relatively modest cost and that it would be more versatile and bring important

additional new physics. As well as accelerating protons, it would be able to accelerate heavy ions to world-beating energies at little extra cost, and LHC protons could be collided with electrons in LEP at much higher energy than in the Hadron Electron Ring Accelerator (HERA) then being built in Hamburg (this option was abandoned in 1995 when it was decided that, after it was eventually closed, LEP should be removed to make it easier to install the LHC).

Rubbia's powerful advocacy and a series of workshops built up pan-European enthusiasm for the LHC, although not everyone was convinced that the collider was justified, or would ever be funded, in parallel with the SSC. I swallowed any doubts and supported the LHC as a member (1986–92) and chairman (1990–92) of CERN's scientific policy committee. I thought the versatility argument was good, and that the LHC should be supported — at the very least as an insurance policy in case the SSC ran into trouble.

In December 1991, the CERN council adopted a resolution that recognized the LHC as "the right machine for the advance of the subject and the future of CERN" and asked Rubbia to come forward with a complete proposal before the end of 1993. I was due to succeed Rubbia as CERN's Director-General at the beginning of 1994, and in early May 1993 he handed me the responsibility of preparing and presenting the LHC proposal. The outlook was not encouraging: a new, detailed costing was significantly bigger than previous estimates; the personnel requested by CERN group leaders to build the LHC would have required a 20% increase in staff; attitudes to high-energy physics were hardening in several CERN member states; and the CERN council had just agreed to a temporary reduction in Germany's contribution on the grounds that reunification was proving very costly.

Bid for approval

Over the summer and autumn of 1993, Lyn Evans, by then nominated as LHC director, proposed several modifications to the design that reduced the cost, and with the help of many others I identified reductions in the rest of the CERN programme that would free up money and manpower. Costing was difficult as this was before most of the research and development for the LHC had been completed — for instance, the first full-length dipole magnet was not tested until December 1993. It was also before approval of the experimental programme, which became more ambitious after a large influx of American researchers joined proposals for LHC experiments after the SSC was cancelled. Our 1993 costing therefore underestimated the eventual specification and cost of the underground areas that were to house the experiments.

The plan I presented to the CERN council in December 1993 foresaw LHC construction, with commissioning in 2002, on the basis of a humped budget, with full compensation for inflation of materials costs. The hump was to come from a mixture, still to be defined, of a general budget increase, additional voluntary contributions from some member states and contributions from non-member states. The plan was generally well received, although it was clear that Germany and the United Kingdom were very unlikely to agree a budget increase, and we were asked to come back with proposals to reduce costs further and indications of how much non-member states might be willing to contribute.

We developed proposals to delay LHC commissioning until 2003 or 2004, stage the construction of the detectors and, while maintaining priority for LEP and the very ambitious LEP upgrade (the first phase of which was not complete), reduce other parts of the CERN programme to a bare minimum over the coming years, with complete closure for one year. In June 1994 we requested approval of the LHC from the CERN Council. The date of commissioning was to be decided later, depending on what voluntary and non-member state contributions were obtained. Encouragingly, a US panel had by then recommended that "the government should declare its intention to join other nations in constructing the LHC" (although the suggested contribution was disappointingly small), and positive signals had been received from Japan, Russia and India. Seventeen member states voted to approve the LHC. The vote was left open, however, because the other two member states — Germany and the United Kingdom — would not accept the proposed budgetary conditions, and demanded substantial additional voluntary contributions from the host states (Switzerland and France), who, they considered, gained disproportionate benefits from CERN.

The missing-magnet machine

Over the next six months, difficult discussions ensued between CERN, the host states, and Germany and Britain (at one point, the Director General of the UK Research Councils, John Cadogan, told me I would be "staring into the abyss" if we could not reduce the cost of the LHC). Some movement on the part of France and Switzerland was beginning to ease the position when Germany and Britain announced that they could only approve the LHC under a planning assumption of 2% inflation to be compensated by 1% indexation — in other words, a 1% annual budget reduction in real terms — and continuation of the German rebate for

Even before the 27-kilometre tunnel for the Large Electron–Positron collider was built, thoughts were turning to its successor, the Large Hadron Collider.

CERN

some years. These conditions seemingly made it impossible to launch the LHC while upgrading and exploiting LEP, and maintaining CERN's small programme of excellent fixed-target experiments.

Our response was to propose keeping down the annual cost by building a 'missing-magnet machine', in which a third of the dipole magnets would be omitted in a first phase, thereby saving some 300 million Swiss francs (US$240 million). The machine would have operated at two-thirds of full energy for some years, before the remaining magnets were installed. Although the final cost would have been more, and results from phase two would have eclipsed the physics from phase one, the two-stage LHC would nevertheless have been a world-beating facility — and it was the only option available.

I was asked whether the two-stage LHC would be viable under the proposed budget conditions. My first reply amounted to a "yes, but…". It was made clear that qualifications would prevent approval. So I took a deep breath and replied yes, adding that the conditions would be acceptable if accompanied by an assurance that any contributions from non-member states would be used to speed up and improve the project, not to allow reductions in the member states' contributions. The CERN council explicitly gave this assurance, and on 16 December 1994 approved the LHC, on the basis of a two-stage construction plan, to be reviewed in 1997 in light of what contributions had been obtained from non-member states, and of the budget conditions required by Germany and Britain, including the continuation of the German rebate. Generous French and Swiss offers to make in-kind contributions and to increase their contributions by 2% a year were crucial factors.

Reassured by a letter from the then British Minister of Science, David Hunt, which described the conditions as "realistic, fair and sustainable", it seemed that all that remained was to try to identify further internal cutbacks, so as to turn our reply to the question of viability into a genuinely unqualified yes, and to seek contributions from non-member states. This we did. Negotiations with Japan — which made the first-ever substantial contribution of a non-member state to a CERN accelerator in June 1995 — and with Russia, India, Canada and the United States went well. By the middle of 1996, we were becoming confident that single-stage construction would be possible, and raised with the council the possibility of bringing forward the 1997 review to December 1996.

Then, in July 1996, out of the blue, the German government announced that, to help ease the financial burden of reunification, it intended to reduce all international science subscriptions. A particularly large cut (8.5% for two years and 9.3% thereafter) was proposed for CERN, despite the fact that Germany was already enjoying a 'reunification rebate'. The possibility of limiting the reduction to just the German contribution was scuppered when the UK government announced that, the minister's letter notwithstanding, it had always seen the 1997 review as another opportunity to look for reductions in the CERN budget. It called for "the largest possible reduction" claiming that this could be achieved "without damaging CERN's scientific mission or endangering the LHC". The particle-physics community in the United Kingdom was reluctant to challenge this assertion, as they were told that reductions in the CERN budget would be their source of funding for participation in LHC experiments.

Forwards through deficit

CERN's public response to the proposed budget cut was muted by fear of shaking the confidence of the non-member states, who had been assured that their contributions would not be used to allow reductions in the member states' contributions. The United States, which had repeatedly asked for reassurances on the viability and sustainability of CERN's planning and funding, was a particular worry. A US contribution to the LHC, significantly larger than suggested by the 1994 panel, had just been negotiated, but a formal agreement had not yet been drafted, let alone signed.

The subsequent discussions were rough. At one stage, Germany threatened to leave CERN if its demands were not met exactly, and even prepared a letter of withdrawal that leaked to the press. There were suggestions that the United Kingdom might also use the threat of withdrawal as a negotiating tactic. I repeatedly told the CERN council that cuts of the magnitude proposed would destroy the LHC's viability, although — as inflation had been zero since 1994, thanks to the strength of the Swiss franc — we had built up a reserve of around 2% in the budget, which could be sacrificed without making matters worse than they had been at the end of 1994. Drawing attention to this reserve was perhaps a tactical error, as it seemed to make a 9% reduction somewhat less out of reach. Given the United Kingdom and Germany's determination, however, I don't think it made any difference to the final outcome.

The crunch came in October when with Horst Wenninger, who played a central role in putting together the LHC proposal, I met the German minister with responsibility for CERN, Jürgen Rüttgers. We explained that the LHC could not survive the proposed budget cut. Rüttgers was uncompromising, but finally asked whether there was any

conceivable way to avoid this conclusion. We replied that it could be avoided if CERN were allowed to take out loans, an idea that Germany had previously vetoed: deficit financing would be extremely risky, but would be acceptable if accompanied by clear acknowledgement of the risks by the CERN council and approval of single-stage construction, which was not only desirable scientifically, but necessary to ensure the contributions required from the non-member states.

A sting in the tail

During the next meeting of delegations to CERN, Germany declared that "a greater degree of risk would inevitably have to accompany the LHC". Others, while accepting deficit financing, acknowledged the risks in similar or stronger terms, and single-stage construction of the LHC, with completion foreseen in 2005, was approved on 20 December 1997. After intensive lobbying of other member states, the accompanying budget reduction was marginally smaller than requested by Germany. The CERN council also imposed a one-year 2% 'crisis levy' on the salaries of CERN staff, even though they had hardly risen since 1993, while further efficiency savings and economies were sought.

We duly got on with this job, but there was a major distraction and sting in the tail. The new chairman of the US House of Representative's Science Committee, James Sensenbrenner, who was suspicious of international projects as a result of his experience with the International Space Station, declared that the proposed agreement between the US Department of Energy and CERN was unsatisfactory. I believed that without the United States, the hard-won European agreement to build the LHC might unravel, whereas their involvement would make it secure. Some very anxious months ensued. In the end it turned out that Sensenbrenner was satisfied with modifications that strengthened the United States' protection against unforeseen events without changing the magnitude of its contribution, and a US–CERN agreement was finally signed in December 1997.

When I left CERN at the end of 1998, the final phase of the LEP upgrade was still being completed, there was a vigorous ongoing LEP programme, and agreement had been obtained to operate LEP for an additional, final year in 2000. Nevertheless, it was clear that the time had come to prepare to reorganize CERN for the post-LEP era on a basis focused on the LHC project. On the face of it, things were going well, half the LHC contracts (by value) having been placed, at prices (in aggregate) just below the estimates.

It was obvious, however, that the budgetary position was extremely fragile. The deficit financing of the LHC was hyper-sensitive to small changes in the timing of contracts. Although the assumptions on timing, manpower and costs, made under great pressure in 1994 and 1996, had not been unreasonable, they had tended to be on the optimistic side, and the LHC proposal had contained no contingency funding because it would not have been accepted by some of the CERN member states. Furthermore, although the tenders for the underground civil engineering had come in below expectations, such contracts are almost always subject to revisions due to unexpected geological conditions, and the contracts for the most demanding and largest single item — the dipole magnets — had not been placed.

The LHC is an extremely challenging project, and the delays that were later produced by problems with the civil engineering and other factors should not have been a surprise. Given that it was approved in the research and development phase with no contingency, and given the 1996 discussion of risks, the 2001 revelation that the LHC would go significantly over budget should also not have been a surprise — although the lack of any warning made it a huge shock.

Lessons for the future

What lessons can be learned from the LHC saga? First, the SSC debacle strongly suggests that new projects should, if possible, be sited at existing laboratories, where they can use existing infrastructure and be spared the challenges of setting up a new laboratory and having to recruit all the key staff from scratch. Second, potential partners should be brought in at the start on equal terms or, if this is not possible, their contributions should bring added value, and they should be offered a 'voice' in the governance, as was done for the LHC. Third, approving large projects is particularly difficult for international organizations: at any time at least one partner may have economic difficulties or be out of sympathy with the organization. Fourth, stability is crucial for successful planning and execution of major projects. The events of 1996 upset orderly management, as well as shaking the confidence of CERN's staff and of potential partners in the non-member states.

Finally, it is not wise to approve projects without contingency on the basis of optimistic assumptions, although it may be worth it if — as in the case of the LHC — this is the only way to get them approved. This point was made in extreme terms, which certainly do not apply to the LHC, in a 2003 supplement to *The Times* of London on the world's great construction projects, which asserted that "If those involved didn't lie about the cost, they would never be built."

The LHC is now almost built, thanks to the dedication of the CERN staff, at a final materials cost only some 20% more than foreseen in 1993; not bad for a high-tech project approved in the research and development phase. It is a fantastic project, and I am confident that the LHC will perform superbly. ■

Chris Llewellyn Smith was Director-General of CERN from 1994 to 1998. He is currently Director of the UK Atomic Energy Authority Culham Division, Culham Science Centre, Abingdon OX14 3DB, UK.

Author Information Reprints and permissions information is available at npg.nature.com/reprintsandpermissions. The author declares no competing interests. Correspondence should be addressed to the author (chris.llewellyn-smith@ukaea.org.uk).

Building a behemoth

Oliver Brüning[1] & Paul Collier[1]

The Large Hadron Collider makes extensive use of existing CERN infrastructure but is in many respects an unprecedented undertaking. It is a proton–proton collider; therefore, it requires two separate accelerator rings with magnetic fields of opposite polarity to guide the two beams in opposite directions around its 27-km circumference. In addition, the extraordinary energies and collision rates that it has been designed to attain pose huge challenges for controlling the beam and protecting the accelerator.

The main objective of the Large Hadron Collider (LHC) is to explore the validity of the standard model of particle physics at unprecedented collision energies and rates. The design performance envisages roughly 30 million proton–proton collisions per second, spaced by intervals of 25 ns, with centre-of-mass collision energies of 14 TeV that are seven times larger than those of any previous accelerator. Reaching and maintaining this level of performance means that the LHC collider itself — although building on experiences gained at previous accelerators such as the Tevatron, at Fermilab (Batavia, Illinois), and HERA, at DESY (Hamburg) — requires a range of novel features that stretch existing technologies to the limit.

Colliders can, in principle, be designed for many different particle species (see page 270): electrons, positrons, protons, antiprotons and ions are all used in existing machines. The Tevatron, which at present defines the energy frontier for particle colliders, operates with proton and antiproton beams. By contrast, the Large Electron–Positron Collider (LEP), the last collider project at CERN, used leptons in the form of electron and positron beams. Each choice has its advantages and disadvantages. On the one hand, because leptons are elementary particles, the centre-of-mass collision energies in machines such as the LEP are precisely defined and therefore are well suited to high-precision experiments. On the other hand, the hadrons that are smashed together by the Tevatron and the LHC are composite particles, and the collisions actually occur between constituent quarks and gluons, each carrying only a proportion of the total proton energy. The centre-of-mass energy of these collisions can vary significantly, so they are not as well suited for high-precision experiments. The hadron colliders, however, offer tremendous potential for the discovery of as-yet unknown particles, because they admit the possibility of collisions over a wide range of much higher energies than is otherwise possible. Protons are relatively heavy and so lose less energy than leptons do while following a curved trajectory in a strong magnetic field. This fact, coupled with the use of superconducting magnet technology, allows the construction of a relatively compact and efficient circular machine, in which the particle beams can collide with each other at each turn. During the lifetime of the LHC, it is planned to operate with both proton and heavy-ion (lead) beams. In this review, we discuss the crucial features of the LHC that should ensure the stability and longevity of the machine while it hosts the uniquely violent collisions of these beams.

Collision energy and beam luminosity

The crucial parameters for a collider such as the LHC are the collision energy and the event rate. Taking into account the partitioning of the proton's energy between its constituent particles (that is, quarks and gluons), the choice of a proton beam energy of 7 TeV at the LHC means that average centre-of-mass collision energies will be greater than 1 TeV. To maximize the total number of events seen by the detectors, a high collision rate is also required, meaning in turn high intensities. The production rates that are achievable for antiprotons at present are too low for the design performance of the LHC; therefore, two counter-rotating proton beams are used. As a result, unlike the Tevatron, the LHC needs two separate vacuum chambers with magnetic fields of opposite polarity to deflect the counter-rotating beams in the same direction.

The number of collision events that can be delivered to the LHC experiments is given by the product of the event cross-section (which

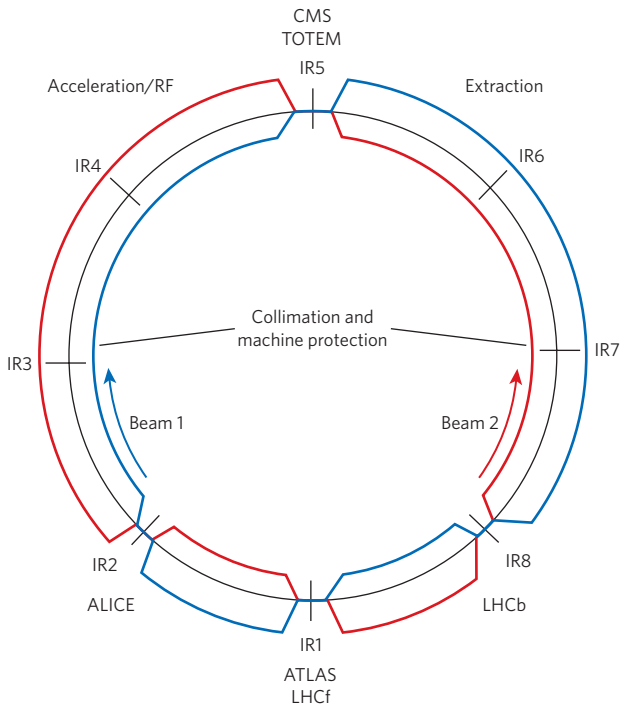


Figure 1 | Layout of the LHC collider. Two proton beams rotate in opposite directions around the ring, crossing at the designated interaction regions (IRs). Four of these (IR1, IR2, IR5 and IR8) contain the various experiments (ALICE, ATLAS, CMS, LHCb, LHCf and TOTEM). IR4 contains the radio-frequency (RF) acceleration equipment, and IR3 and IR7 contain equipment for collimation and for protecting the machine from stray beam particles. IR6 houses the beam abort system, where the LHC beam can be extracted from the machine and its energy absorbed safely.

[1]Accelerators and Beams Department, CERN, CH-1211 Geneva 23, Switzerland.

is a measure of the probability that a collision will produce a particular event of interest) and the machine luminosity, L. This is determined entirely by the proton beam parameters:

$$L = \frac{f_{rev} n_b N^2}{\sigma_x \sigma_y} F(\Phi, \sigma_{x,y}, \sigma_s)$$

Here, σ_x and σ_y are the transverse root mean squared (r.m.s.) beam sizes at the interaction points; f_{rev}, the revolution frequency; n_b, the number of particle packages ('bunches'); N, the number of particles within each bunch; and F, a geometric reduction factor that depends on the crossing angle of the two beams (Φ), the transverse r.m.s. beam size ($\sigma_{x,y}$) and the r.m.s. bunch length (σ_s). To provide more than one hadronic event per beam crossing, the design luminosity of the LHC has been set to $L = 10^{34}$ cm^{-2}s^{-1}. This translates as 2,808 bunches, each containing 1.15×10^{11} protons, a transverse r.m.s. beam size of 16 μm, an r.m.s. bunch length of 7.5 cm and a total crossing angle of 320 μrad at the interaction points. For the programme involving lead-ion collisions, L will be 10^{27} cm^{-2}s^{-1} at a centre-of-mass energy of 1,148 TeV. In this case, each ring of the LHC will contain 592 bunches, each with 7×10^7 lead ions. The transverse beam sizes will be similar to those of the proton beams.

The LEP tunnel

The LHC features six experiments (Fig. 1): two high-luminosity experiments (ATLAS[1] and CMS[2]); two supplementary experiments at low scattering angles (LHCf[3] and TOTEM[4]), which are near ATLAS and CMS, respectively; one B-meson experiment (LHCb[5]); and one dedicated ion physics experiment (ALICE[6,7]).

To make best use of the existing infrastructure at CERN, the LHC is being built in the 27-km-long LEP tunnel[8]. Approximately 22 km of the LEP tunnel consist of curved sections, or arcs, in which bending dipole magnets can be installed. The remaining 5 km consist of eight straight interaction regions that provide space for the experiments, injection and extraction elements for the proton beams, acceleration devices and dedicated 'cleaning' insertions that collimate the beam and protect the superconducting magnets from stray particles.

Dipoles and quadrupoles

Not all of the tunnel's curved sections can be used for the installation of dipole magnets. In addition to the bending fields of the dipole magnets, a circular accelerator also requires a focusing mechanism that keeps the particles centred on the design orbit. There are basically two types of circular accelerator: pulsed machines and storage rings. A storage ring is a circular accelerator where the beam may be kept for a significant time in steady conditions. In the case of the LHC, this will be several hours. Most modern storage rings use the concept of strong focusing[9,10], in which dedicated quadrupole magnets provide field components that are proportional to the deviation of the particles from the design orbit. The resulting Lorentz force prevents divergent trajectories: the particles, instead, oscillate around the design orbit as they circulate in the storage ring. The number of transverse oscillations per revolution is an important operational parameter and is referred to as the machine tune, Q. The stronger the focusing, the smaller the oscillation amplitudes (and thus the transverse r.m.s. beam size) and the larger Q is. In the longitudinal direction, the electric field supplied by a radio-frequency resonator focuses the particles into bunches and accelerates them. The LHC has two such systems, one for each beam, in one of the ring's straight sections (IR4) (Fig. 1).

The design of accelerator magnets becomes easier and less expensive for small magnet apertures, so the natural inclination is to increase the number of focusing elements in the machine to minimize the transverse beam size. But a careful balance must be struck between maximizing the

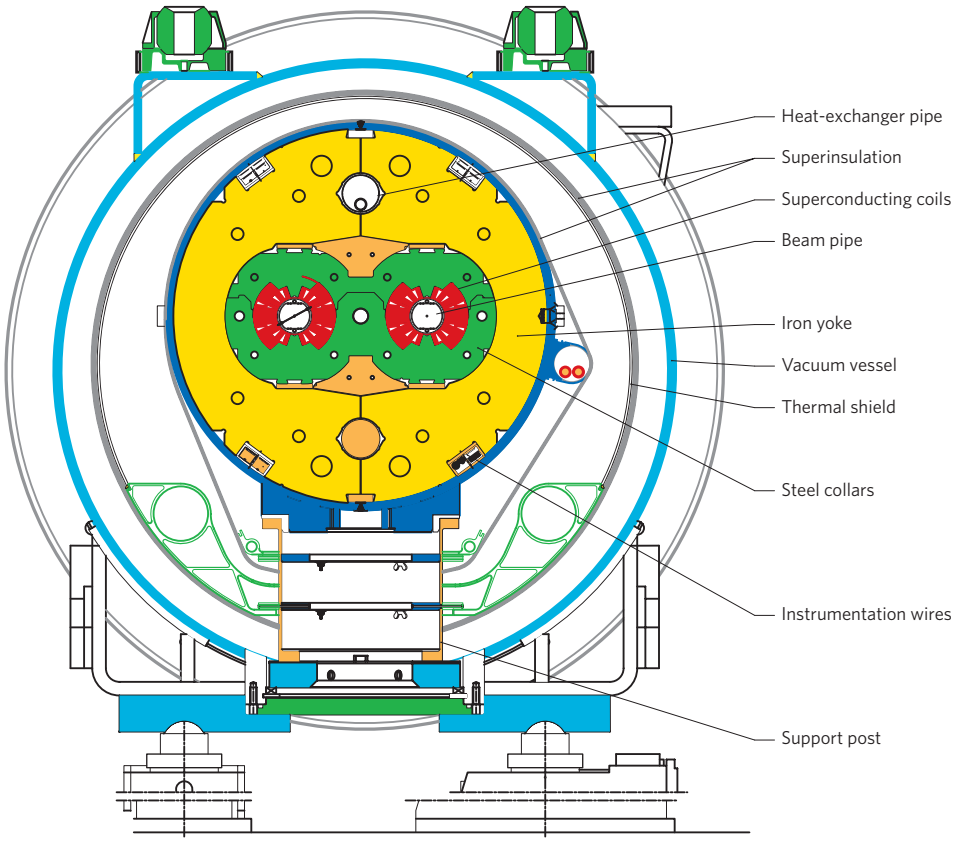


Figure 2 | Cross-section of the two-in-one design for the main LHC magnets. In the centre are the two beam pipes, separated by 194 mm. The superconducting coils (red) are held in place by collars (green) and surrounded by the magnet yoke (yellow). Together, these form the cold mass of the magnet, which is insulated in a vacuum vessel (outer blue circle) to minimize heat uptake from the surroundings.Image reproduced, with permission, from ref. 11.

space for dipole installation and providing sufficient space for transverse beam focusing. In the LHC, ~80% of the arc length is taken up by the dipole magnets, allowing the maximum transverse r.m.s. beam size in the arcs to be kept below 1.3 mm.

A two-in-one design

The combination of the length of the existing tunnel and the required beam energy sets the scale for the strength of the bending magnetic fields in the main magnets of the LHC. Keeping the 7-TeV proton beams on their closed orbits implies bending fields of 8.4 T, ~30,000 times stronger than Earth's magnetic field at its surface. Such fields are at the limit of the existing superconducting magnet technology. To confine two counter-rotating proton beams, two separate magnet apertures with opposite field orientations must be squeezed into the 3.76-m diameter of the existing LEP tunnel. The LHC therefore adopted a novel two-in-one magnet design, in which the two magnetic coils have a common infrastructure and cryostat[11] (Fig. 2).

This design provides a compact structure, with a cryostat diameter of 0.914 m, that fits two separate beam apertures into the relatively small existing machine tunnel. But it also couples the construction constraints of the two magnets, imposing new challenges and tighter tolerances on their production. This is the first time this has been done, so there is no existing experience to build on.

To minimize the number of magnet interconnections, and therefore the space lost for dipole field installations, the LHC uses 30-tonne, 15-m-long dipole magnets, which are more than twice as long as the dipole magnets in previous accelerators (~6 m for the Tevatron and HERA[12,13]). These large dimensions imposed tighter geometric constraints on the construction, transportation and installation of the magnets (Fig. 3). Each of the 8 arcs of the LHC consists of 46 repeating series of 1 quadrupole and 3 dipole magnets. Each magnet is manufactured using a niobium–titanium (NbTi)-based superconducting cable (Box 1).

Measures against magnetic quench

The operating temperature and field strength of 1.9 K and 8.4 T mean that the LHC has a very small thermal margin before the superconducting state is lost. Even small particle losses or other thermal instabilities inside the magnets can cause local heating of the material. After a section of the NbTi cable becomes a normal conductor, ohmic losses increase the operating temperature still further, an effect known as magnet quench. Testing for when a quench occurs has been an important part of the pre-installation tests of all of the LHC magnets, but efficient operation of the collider demands that the likelihood of this happening during operation is minimized.

The small tolerances for temperature fluctuations and energy deposition in the magnet coils at the LHC are combined with the extremely high energy densities inside the magnet system. The total stored electromagnetic energy — 8.5 GJ for the dipole circuits alone — is more than ten times greater than the previous record of 0.7 GJ, set by HERA[12]. The damage potential to the accelerator hardware from this stored energy is enormous: just 1 MJ is enough energy to melt 2 kg of copper.

In case of a magnet quench, this stored energy must be extracted and dissipated quickly in a controlled manner. By separating the main LHC magnet circuits into eight independent powering sectors, the stored electromagnetic energy per sector falls to that seen in existing superconducting storage rings. The drawback of this division into sectors is that it requires accurate synchronization of the different magnet sectors during operation. Existing storage rings avoid this synchronization problem by powering all main magnets in series in a central circuit. The LHC will enter new territory in this respect.

Damage to individual magnet units during a quench is avoided by a dedicated magnet protection system that monitors the voltage drop across each magnet unit. As soon as any part of the magnet cable loses its superconducting state, the voltage drop across the magnet will become non-zero. This jump will activate special heaters inside the magnet to bring the whole magnet into a normal conducting state, thus spreading the quench over the whole magnet length. A dedicated quench diode dissipates the stored electromagnetic energy before it can damage the magnet coils.

Figure 3 | Installing the LHC magnets. a, An LHC dipole ready for installation at the CERN site. **b**, Transport of LHC magnets in the tunnel, alongside installed elements, illustrating the tight space conditions for installation. Images reproduced with permission from CERN.

The stored energy in the proton beams themselves is another dangerous source of energy deposition in the superconducting magnet coils. At 7 TeV and an intensity of 3.23×10^{14} protons, the kinetic energy of each of the LHC beams is 362 MJ. Safe beam extraction in case of problems during machine operation, or at the end of a period of operation for data taking by the experimental detectors (physics fill), is assured by two installations: the beam abort system, and the machine protection system. The ring of the LHC has a dedicated beam abort system, formed of specially designed absorber blocks capable of absorbing the full beam intensities at 7 TeV without damage. The machine protection system constantly monitors all critical machine parameters and initiates a beam abort if the parameters exceed the acceptable operation tolerances or if the beam losses along the storage ring become too large.

Beam lifetimes

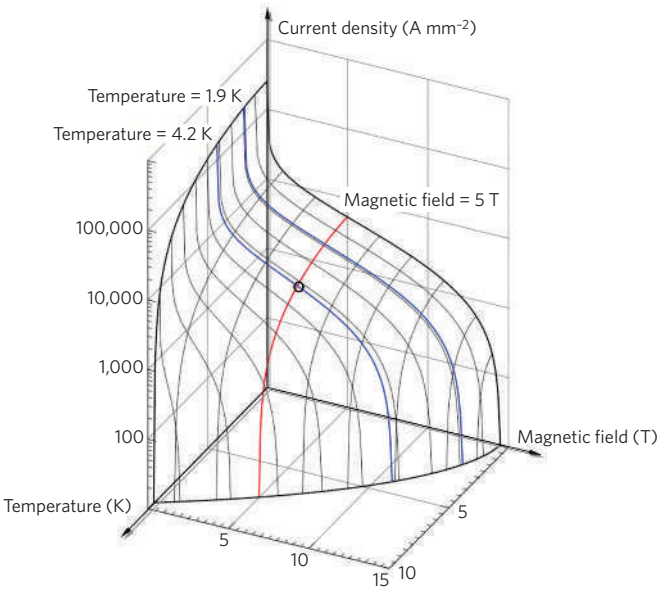
Beam intensity — and hence luminosity — decays during the operation of an accelerator in colliding mode. After these parameters become too small for efficient operation, the beams are discarded using the beam abort system, and a new fill of proton beams needs to be prepared, injected and accelerated. One of the main and unavoidable causes of reductions in beam intensity is the collisions inside the detectors themselves, because these cause the disintegration of beam particles. The rate of this disintegration is given by the product of the machine luminosity, the total cross-section for an inelastic interaction of two protons at 7 TeV, and the number of collision points. Assuming a total inelastic cross-section of 10^{-25} cm^2 at 7 TeV and two main interaction points with

Box 1 | The LHC superconductor

The superconducting magnets of the LHC are superlative devices: had the LHC been made of conventional magnets, it would have needed to be 120 km long to achieve the same energies, at the cost of a much greater electricity consumption.

Like all superconductors, NbTi (the material used for the cables of the LHC magnets) is a superconductor only if its operational parameters — temperature, current density and ambient magnetic field — are within certain bounds[15–17]. The critical surface below which this combination of parameters must lie is shown in the figure. At the preferred operating temperature for most existing superconducting accelerators such as HERA and the Tevatron[11,12], 4.2 K, the critical magnetic field is around 5 T.

The temperature of the LHC, 1.9 K, allows the generation of the required 8.4-T magnetic field using a current density of 1.5–2 kA mm^{-2} inside the superconducting cables. It also allows the use of superfluid helium, which has high thermal conductivity, as a coolant. A helium inventory of 120 tonnes or more will be needed to cool the total magnet mass of 37,000 tonnes, the largest such inventory in the world. Figure courtesy of L. Bottura (CERN).



a luminosity of 10^{34} $cm^{-2}s^{-1}$, the beam intensities will have dropped to half of their initial values after ~45 hours.

Particles are also lost through perturbations and resonances in the proton motion that deflect particles away from the design orbit. These are generated, for example, at the collision points, where a particle in one beam is exposed to the Coulomb field of the opposing beam, or by field imperfections in the main magnets. Thanks to the focusing mechanism of the quadrupole magnets, these deflections do not lead directly to particle losses but, initially, just to an oscillation around the design orbit. But consecutive perturbations can add up coherently if the particle oscillations are in resonance with the revolution frequency, in which case the oscillation amplitudes can grow until the particles are lost when they reach the boundary of the LHC vacuum system.

Two approaches minimize this amplitude growth. First, extreme care is taken during the magnet design, construction and installation in order to minimize any imperfections in the machine. Second, the LHC is equipped with dedicated circuits that allow the correction of the most dominant residual field errors. All magnets are measured before their installation to develop an accurate magnetic model of the entire machine's operation. In total, the LHC features 112 correction circuits per beam (not including simple steering magnets for an adjustment of the central orbit), and all of these must be adjusted during operation. To compound the difficulty, the field errors of a superconducting magnet are not constant but vary with time as a function of the magnet's powering history.

Other causes of a reduction in luminosity during operation include the scattering of protons on residual gas molecules inside the beam vacuum system, and the Coulomb scattering of the protons inside each bunch as they perform longitudinal and transverse oscillations while circulating inside the storage ring. The rate of collisions with residual gas molecules depends on the pressure and gas composition inside the machine vacuum system. An efficient operation with proton beams requires vacuum levels below 10^{15} molecules per cubic metre for all gas components (H_2, He, CO, CO_2 and so on), corresponding to a pressure of less than 10^{-7} Pa at 5 K. (Atmospheric pressure is ~10^5 Pa at sea level.) This, in turn, demands an elaborate system of different vacuum pumps. In its final phase, the pumping mainly relies on the cryo-pumping of the cold surfaces that exist at the boundary between the beam vacuum and the helium in the superconducting magnets, similar to the way that ice builds up on the surfaces of the freezing compartment of a household refrigerator.

Collimation

There are therefore several unavoidable mechanisms causing a continuous loss of particles during LHC operation through a relatively slow drift to larger oscillation amplitudes. Two dedicated collimation insertions with specially designed absorber blocks mop up these stray particles before they can reach the cold aperture of the superconducting magnets (and so possibly cause a magnet quench). This mopping up must be done with high efficiency so that only 1 in 10,000 particles that hit the primary collimators end up inside the cold aperture. Such a high cleaning efficiency requires extremely tight tolerances for the main machine parameters during operation, as well as the use of a complex two-stage collimation system with additional dedicated absorbers at crucial locations. The LHC is the first high-energy collider that requires a beam collimation during all stages of the operation to protect its machine elements — previous colliders only required a beam collimation during the physics run, mainly to reduce the background in the experiments. For additional safety, therefore, the collimator jaws at the LHC are made of fibre-reinforced graphite so as to be able to withstand the direct impact of a large proportion of the 7 TeV beam.

Working up to full beam strength

After the LHC proton beams have been prepared and injected into the accelerator using the existing accelerators at CERN[14] (Box 2), the acceleration can be initiated. This acceleration relies on a synchronous change of the machine settings with the increasing dipole field. In the case of the LHC, the final beam energy is more than 15 times greater than that of the injected beam (7,000 GeV compared with 450 GeV). With a high impedance in the main magnet circuits, the process of increasing the magnet current, and therefore the energy, is slow in the LHC, taking ~20 minutes. During this operational phase, the transverse beam dimensions shrink as the rigidity of the beam increases.

The injection and acceleration takes place with the beams separated in the experimental regions and with a lower focusing strength in these areas than in the final configuration for luminosity production. After reaching high energy, a synchronized change in the settings of the focusing elements is made at each interaction point to reduce the beam spot size at the interaction point. As a result, the beam size in the adjacent final focusing quadrupoles increases. In the final configuration, the aperture of these elements is smaller than in the rest of the machine and must be protected by further reducing the collimation gap.

The final step before the experiments can begin taking data is to remove the separation scheme and to bring the beams into collision in each experimental area. Careful optimization is required to align the beams correctly and to maximize the overlap of the 16-μm beam spots.

Once data taking has started, the luminosity will decrease as the intensity falls. In fact, because the luminosity is proportional to the square of the intensity, a reduction in the intensity by ~30% will halve the luminosity. The goal for efficient machine operation is a luminosity lifetime that is considerably longer than the average time for preparing a new fill of proton beams. Assuming an exponential decay of luminosity, and an

intensity lifetime of 45 hours, the luminosity lifetime will be ~15 hours for the LHC. The overall collider efficiency depends on the ratio of the run length and the average turnaround time. Assuming a 5-hour turnaround time, the optimum run length will be ~10 hours.

The commissioning process

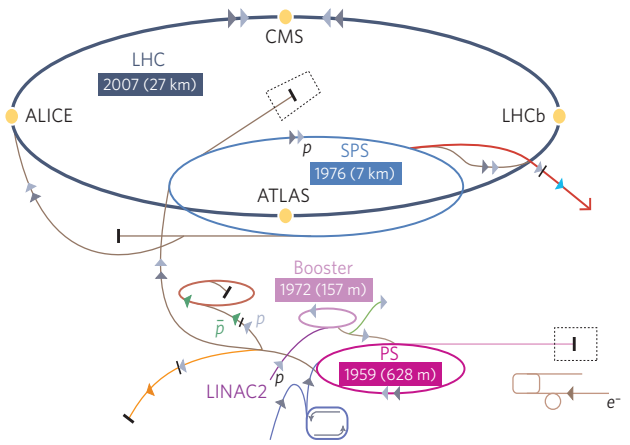
The LHC is a huge, complex facility, and careful and precise control of all machine elements is necessary. Commissioning the whole machine is a challenge in itself. Careful commissioning of each individual set of accelerator hardware will be followed by rigorous system tests and integrated operation of the whole accelerator before any beam is injected. In the early phases of beam operation, the complexity can be reduced by limiting the number of bunches, the intensity per bunch and even the final energy. At each stage in the commissioning process, the equipment and protection systems must be tested and run to allow operation with beam while minimizing the risk of damage to the accelerator itself.

For the first operation of the LHC at 7 TeV, there will be a single bunch in each ring. From there, a staged increase in the number of bunches is intended, with schemes for 43, 156 and 936 bunches per ring envisaged before arriving at the final number of 2,808 bunches per ring. The simplest scheme, with 43 bunches per ring and an intensity per bunch around half the nominal value, represents a stored energy that is already comparable to that of the Tevatron.

Box 2 | Preparing the LHC beam

The beam of the LHC starts off in a 50-MeV linear accelerator, LINAC2 (see figure). It is then passed to a multi-ring booster synchrotron for acceleration to 1.4 GeV, and then to the 628-m-circumference Proton Synchroton (PS) machine to reach 26 GeV. During acceleration in the PS, the bunch pattern and spacing needed for the LHC are generated by splitting the low-energy bunches. A final transfer is made to the 7-km Super Proton Synchroton (SPS) machine, where the beam is further accelerated to 450 GeV. At this point, it is ready for injection into the LHC. The cycle takes ~20 s and creates a ribbon, or train of bunches, with a total kinetic energy of more than 2 MJ. This is ~8% of the beam needed to fill an LHC ring completely, so the whole cycle must be repeated 12 times per ring.

The transfer of the bunch trains from the SPS to the LHC is one of the most dangerous phases of the operational cycle of the LHC. The injected beam already has sufficient energy to damage the LHC equipment, and the transfer involves the use of fast kicker magnets to abruptly change the trajectory of the beam to move it out of the SPS, down a 3-km transfer line, and into the LHC. Any mis-steering here could be disastrous, so a low-intensity 'pilot' beam is injected into the machine first. This is used to measure and correct the machine parameters before the full-intensity injection sequence is allowed to start. Each injection is positioned in the LHC circumference so as to generate the complete pattern for each beam. During the 8 minutes needed to fill the LHC completely, the stability of the whole complex is critical and must be carefully monitored. Figure modified with permission from CERN.



During the first full year of LHC operation, the number of bunches and the intensity per bunch will be increased slowly. It is hoped that a luminosity of 10^{33} cm^{-2}s^{-1}, or 10% of the nominal value, will be reached during this time. In subsequent runs, the performance will be slowly increased towards the nominal value as understanding of the machine and control of the machine parameters is refined.

The LHC is a machine in which all technologies are stretched towards their limit, and it has been built, in many cases, with very small operational margins in the equipment. It is probable that upgrades to certain accelerator components will be made during the lifetime of the machine. Some of these will be designed to re-introduce operational margins in crucial areas in which machine efficiency can be improved. Others will be designed to increase the nominal performance of the machine.

The outlook

The LHC is designed to push back the frontiers of our knowledge of fundamental particle physics. With the requirement of providing both high energies and high beam intensities, there are many challenges that had to be overcome to produce a viable design for the complete machine. Realizing the designs for each component of the accelerator has often, in turn, pushed back the technical boundaries for the design and performance of the individual accelerator systems. The sheer size and complexity of the complete machine makes the commissioning and operation of the LHC a challenge in itself. But its many technical innovations mean that the LHC should be capable of helping us to explore — and, we hope, answer — some of the most fundamental questions in particle physics today. ■

1. Armstrong, W. W. *et al.* (The ATLAS Collaboration). ATLAS: *Technical Proposal for a General-Purpose* pp *Experiment at the Large Hadron Collider at CERN. CERN* <http://cdswebdev.cern.ch/record/290968> (1994).
2. Della Negra, M., Petrilli, A., Hervé, A. & Foà, L. *CMS Physics: Technical Design Report. Vol. 1: Detector Performance and Software. CERN* <http://cdswebdev.cern.ch/record/922757> (2006).
3. Mukari, Y., Itow, Y. & Sako, T. *LHCf Experiment: Technical Design Report. CERN* <http://cdswebdev.cern.ch/record/926196> (2006).
4. Berardi, V. *et al. Total Cross-Section, Elastic Scattering and Diffractive Dissociation at the Large Hadron Collider at CERN: TOTEM Technical Design Report. CERN* <http://cdswebdev.cern.ch/record/704349> (2004).
5. Amato, S. *et al.* (The LHCb Collaboration). *LHCb Technical Proposal, a Large Hadron Collider Beauty Experiment for Precision Measurements of CP Violation and Rare Decays CERN/LHCC 98-4. CERN* <http://cdsweb.cern.ch/record/622031> (1998).
6. Carminati, F. *et al.* (The ALICE collaboration). ALICE: physics performance report, volume I. *J. Phys. G* **30,** 1517–1763, doi:10.1088/0954-3899/30/11/00 (2004).
7. Alessandro, B. *et al.* (The ALICE collaboration). ALICE: physics performance report, volume II. *J. Phys. G* **32,** 1295–2040, doi:10.1088/0954-3899/32/10/001 (2006).
8. CERN. *LEP Design Report. CERN* <http://cdsweb.cern.ch/record/102083> (1984).
9. Christofilos, N. C. Focusing system for ions and electrons. US patent 2,736,799 (1950); reprinted in Livingston, M. S. *The Development of High Energy Accelerators* (Dover, New York, 1966).
10. Courant, E. D. & Snyder, H. S. Theory of the alternating gradient synchrotron. *Ann. Phys. (Leipz.)* **3,** 1–48 (1958).
11. Brüning, O. S. *et al.* (eds) *LHC Design Report Vol. 1: The LHC Main Ring. CERN* <http://cdsweb.cern.ch/record/782076> (2004).
12. Cole, F. T. *et al. A Fermilab Superconducting Accelerator Design Report*. Beams Document 1888, Fermilab, Batavia Illinois <http://beamdocs.fnal.gov/AD-public/DocDB/ShowDocument?docid=1888> (1979).
13. *HERA — A Proposal for a Large Electron Proton Colliding Beam Facility at DESY.* DESY Hamburg, Germany. *SLAC* <http://www.slac.stanford.edu/spires/find/hep/www?key=897230> (1981).
14. Benedikt, M., Collier, P., Mertens, V., Poole, J. & Schindl, K. (eds) *LHC Design Report Vol. 3: The LHC Injector Chain. CERN* <http://cdsweb.cern.ch/record/823808> (2004).
15. Wilson, M. N. *Superconducting Magnets* (Clarendon, Oxford, 1983).
16. Bottura, L. A practical fit for the critical surface of NbTi. *IEEE Trans. Appl. Supercon.* **10,** 1054–1057 (2000).
17. Leroy, D. Review of the R&D and supply of the LHC superconducting cables. *IEEE Trans. Appl. Supercon.* **16,** 1152–1159 (2006).

Acknowledgements

The authors would like to acknowledge the many thousands of people at CERN and in collaborating institutes who have contributed to the design and construction of the LHC machine and experiments.

Author Information Reprints and permissions information is available at npg.nature.com/reprintsandpermissions. The authors declare no competing financial interests. Correspondence should be addressed to P.C. (paul.collier@cern.ch).

NATURE|Vol 448|19 July 2007|doi:10.1038/nature06078

Detector challenges at the LHC

Steinar Stapnes[1,2]

The best way to study the existence of the Higgs boson, supersymmetry and grand unified theories, and perhaps the physics of dark matter and dark energy, is at the TeV scale. This is the energy scale that will be explored at the Large Hadron Collider. This machine will generate the energy and rate of collisions that might provide evidence of new fundamental physics. It also brings with it the formidable challenge of building detectors that can record a large variety of detailed measurements in the inhospitable environment close to the collisions points of the machine.

Four main experiments have been designed and constructed for the Large Hadron Collider (LHC) machine: ATLAS, CMS, LHCb and ALICE. ATLAS and CMS are large general-purpose experiments. LHCb will study *b*-quark systems, produced predominantly in the forward direction, and ALICE is designed specifically for studies of heavy-ion collisions (see page 302).

This review focuses on the challenges — related to tracking, calorimetry, muon detection, triggering and data acquisition — faced by the designers and builders of the general-purpose detectors ATLAS and CMS, as well as some of the particular issues for the more specialized detector LHCb.

Experimental measurements at the LHC

Inside the 27-km ring of the LHC, bunches of 10^{11} protons will collide 40 million times per second to provide 14-TeV proton–proton collisions at the LHC design luminosity of 10^{34} cm^{-2}s^{-1} (see page 285). With an inelastic proton–proton cross-section of about 100 mb (~10^{-25} cm^2), this gives 25 events per bunch crossing, or a total rate of 10^9 inelastic events per second. This means that around 1,000 charged particles will emerge from the collision points every 25 ns, within a volume defined by $|\eta| < 2.5$, where pseudorapidity, η, is related to the polar angle relative to the beam axis, θ, by $\eta = -\ln[\tan(\theta/2)]$ (Fig. 1a).

This formidable luminosity and interaction rate are necessary, because the expected cross-sections are small for many of the LHC benchmark processes (such as Higgs production and decay, and some of the processes needed to search for and explore new physics scenarios such as supersymmetry and extra dimensions). They also raise a serious experimental difficulty; every candidate event for new physics will, on average, be accompanied by 25 inelastic events occurring simultaneously in the detector.

The very nature of proton–proton collisions creates a further difficulty. The cross-sections for producing jets of particles, through quark interactions governed by quantum chromodynamics (QCD), are large compared with the rare processes being sought — several orders of magnitude larger, even for jet production above 500 GeV. Therefore, one has to look for characteristic experimental signatures, such as final states involving one or more leptons, or photons, or with missing transverse energy or secondary vertices, to avoid being drowned by QCD background processes. Searching for such final states among already rare events imposes further demands on the luminosity needed and on the detectors' particle identification capabilities.

Specific requirements for the LHC detector systems[1–8] have been defined using a set of benchmark processes that covers most of the new phenomena that one might hope to observe at the TeV scale. The first such process is the production of the standard-model Higgs boson, which is particularly important because there is a wide range of decay modes possible, depending on the mass, m_H, of the Higgs boson. If m_H is low (less than 180 GeV, which is twice the mass of the *Z* boson), the natural width is only a few MeV, and the observed width will be defined by the instrumental resolution. The dominant decay mode into hadrons is difficult to isolate, because of the QCD background. Therefore, the two-photon decay channel will be important, as will other channels, including associated productions such as *ttH*, *WH*, *ZH* (see page 270), for which a lepton from the decay of the accompanying particle will be used for triggering and background rejection.

Above 130 GeV, Higgs decay into *ZZ* (one Z being virtual when m_H is below the *ZZ* threshold), with its four-lepton final state, will be the most interesting channel. Above 600 GeV or so, *WW* or *ZZ* decays into jets or into states involving neutrinos (leading to missing transverse energy because the neutrinos are undetected) are needed to extract a signal; for m_H close to 1 TeV, it becomes necessary to tag 'forward' jets, in the region $2 > |\eta| > 5$, from the *WW* or *ZZ* fusion production mechanism. The Higgs might not even be of the standard-model variety; detection of some of the Higgs particles of the 'minimal supersymmetric extension of the standard model' (MSSM) would require very good sensitivity to processes involving tau leptons and *b* quarks.

If supersymmetric particles such as squarks and gluinos are produced at the LHC, their decays would involve cascades that always contain a lightest stable supersymmetric particle, or LSP (if R-parity is conserved). Because the LSP interacts very weakly with the detector, the experiments would measure a significant missing transverse energy in the final state. The rest of the cascade results in a number of leptons and jets.

Several new models, motivated by theories of quantum gravity, propose the existence of extra dimensions[1–4]. In terms of experimental signatures, the emission of gravitons that escape into extra dimensions would result in missing transverse energy; furthermore, Regge-like excitations could manifest themselves as *Z*-like resonances with ~TeV separations in mass. Other experimental signatures could be anomalous, high-mass dijet production and mini-black-hole production with very spectacular decays involving democratic production of jets, leptons, photons, neutrinos, and *W* and *Z* bosons.

The LHC will also allow studies of QCD, electroweak and flavour physics. For example, *t* quarks will be produced at the LHC at a rate measurable in hertz. New, heavy gauge bosons (W' and Z') could be accessible at masses up to 5–6 TeV. To study their leptonic decays, high-resolution lepton measurements and charge identification are needed

[1]Department of Physics, University of Oslo, 0316 Blindern, Oslo, Norway. [2]Department of Physics, CERN, CH-1211 Geneva, Switzerland.

up to transverse momenta of a few TeV. Another new-physics signature could be jets produced with very high transverse momenta; if quarks are composite rather than fundamental particles, deviations from QCD expectations in the jet cross-sections could result.

Detector concepts

The necessary detection capabilities for all of these experimental signatures lead to a stringent set of design requirements. First of all, owing to the experimental conditions at the LHC, the detectors need fast, radiation-tolerant electronics and sensor elements. In addition, high granularity of the detectors is needed to be able to handle the particle fluxes and to reduce the influence of overlapping events. Good charged-particle momentum resolution and reconstruction efficiency in the tracking system, and in the inner tracker specifically, are essential. For efficient high-level triggering and offline tagging of taus and *b* quarks (which decay a short distance from the primary interaction vertex at which they are produced), pixel detectors close to the interaction region are needed to observe the distinctive secondary vertices.

Two key requirements are good electromagnetic calorimetry for electron and photon identification and measurements, and full-coverage hadronic calorimetry for accurate jet and missing-transverse-energy measurements. Likewise, good muon identification and momentum resolution over a wide range of momenta, and the ability to determine unambiguously the charge of muons with high transverse momentum, are essential. Finally, triggering the event readout on the presence of leptons, jets, photons or missing transverse energy — and at low transverse-momentum thresholds to ensure high efficiencies for most of the physics processes of interest at LHC — is an absolute requirement to reduce the data rate (a few hundred collisions out of the 40 million taking place every second are finally kept) to a level that can be handled offline.

The layout of the ATLAS detector[1,2] is shown in Fig. 1c. It has an inner, thin, superconducting solenoid surrounding the inner detector cavity, and large, superconducting, air-core toroids, consisting of independent coils arranged with an eight-fold symmetry, outside the calorimeters. The inner detector comprises a large silicon system (pixels and strips) and a gas-based transition-radiation 'straw' tracker. The calorimeters use liquid-argon technology for the electromagnetic measurements and also for hadronic measurements in the endcaps of the detector. An iron/scintillator system provides hadronic calorimetery in the central part of the detector. The muon system is based on gas detectors and has precise tracking chambers and trigger chambers for a robust and efficient muon trigger. The ATLAS detector has a radius of 13 m and is 46 m long, with a weight of 7,000 tonnes.

The design of CMS[3,4] is shown in Fig. 1b . The main distinguishing features of CMS are a high-field solenoid housing a full silicon-based inner tracking system (pixels and strips), a fully active, scintillating crystal electromagnetic calorimeter, and a compact scintillator/brass hadronic calorimeter. Outside the solenoid, there is a hadronic 'tail-catcher' in the central region, and an iron-core muon spectrometer sitting in the return field of the powerful solenoid, with tracking chambers and trigger

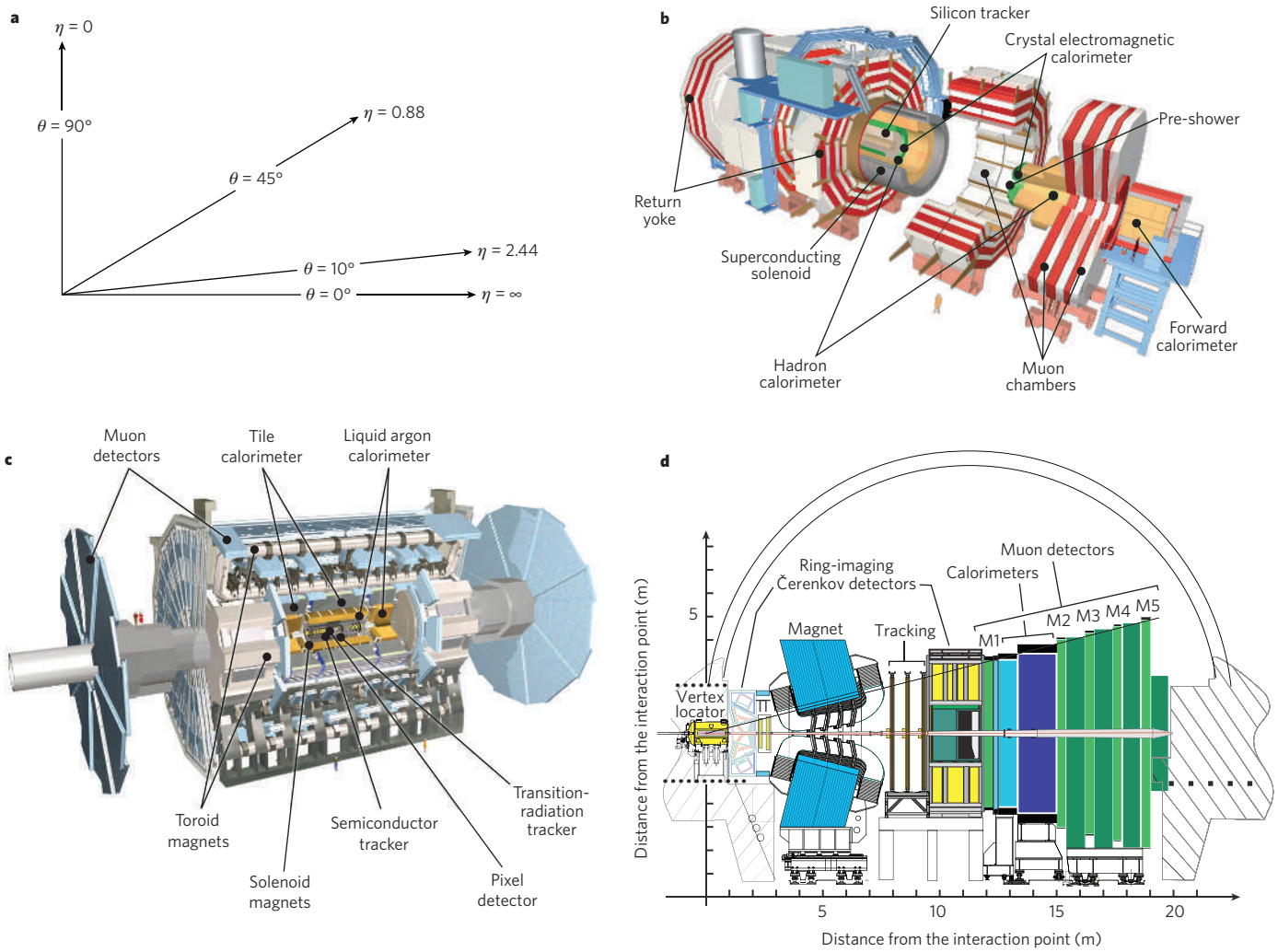


Figure 1 | Detector design. The complex experimental apparatus comprises different detector components, each optimized for a particular task. Regions of the detector volume are commonly described using the variable pseudorapidity, η, which is related to the polar angle, θ, as shown in **a**. The geometry and basic elements of the general-purpose LHC detectors, CMS (**b**) and ATLAS (**c**), are similar, but the layout of the more specialized detector LHCb (**d**) is optimized for detecting the production of *b* quarks in the forward direction. Images b–d reproduced with permission from CERN.

chambers. The CMS system is more compact than ATLAS, and has a radius of 7.5 m and length of 24 m, but weighs 12,000 tonnes.

LHCb[8] is designed to operate at a luminosity of 2×10^{32} cm^{-2}s^{-1}, and its instrumentation is concentrated in the forward direction, between 10 and 300 mrad (Fig. 1d), because this is the region in which the pairs of b and $\bar{b}$ quarks that it aims to study are predominantly produced. The LHCb detector has a silicon vertex detector around the interaction region; then a tracking system consisting of silicon microstrip detectors and a straw tracker, and it includes a dipole magnet. It also has two ring-imaging Čerenkov detectors, positioned in front of and after the tracking system, for charged-hadron identification; a calorimeter system and finally a muon system. LHCb, in particular, has to trigger efficiently on the secondary vertices that are the signature of b quarks, and so its vertex and tracking detectors are factored into an early stage of its trigger scheme.

To construct these large detectors requires substantial resources. The ATLAS and CMS communities each consist of more than 150 universities and institutions from about 35 countries with about 2,000 collaborators per experiment. (LHCb is a factor of three smaller.) Research and development for the LHC detectors began around 1990; the construction projects were approved in 1996 and started in earnest around 1998. The effort to build all of the detector components has involved physicists all over the world, with groups of geographically distributed institutes taking responsibility for the construction of various parts according to their specific expertise and capabilities, as well as involving a large network of industrial partners.

Inner detectors

The ATLAS and CMS inner detectors (Fig. 2) are contained in central solenoid fields of 2 T and 4 T, respectively. They provide efficient tracking of charged particles within the pseudorapidity range $|\eta| < 2.5$, allowing momentum measurement and the reconstruction of primary and secondary vertices. Both systems are largely based on silicon, with high-granularity pixel systems at the smallest radii, and silicon-strip detectors at larger ones. ATLAS has a 'straw' tracker at the largest radius.

Silicon detectors are p–n junction diodes that are operated at reverse bias[9]. This forms a sensitive region depleted of mobile charge and sets up an electric field that sweeps charge (electron–hole pairs) liberated by radiation towards the electrodes. Detectors typically use an asymmetric structure, for example, a highly doped p electrode and a lightly doped n region (p-i-n), so that the depletion region extends predominantly into the lightly doped volume. By adding highly doped n electrodes (n-i-n) at the back, the back side can also be read out. Integrated circuit technology allows the formation of high-density micrometre-scale electrodes on large (10–15 cm in diameter) wafers, providing excellent position resolution. Furthermore, the density of silicon and its small ionization energy (the energy needed to created an electron–hole pair) result in adequate signals with active layers only 200–300 μm thick, and the charge mobility is such that the signals are also fast (typically tens of nanoseconds).

The main challenges for the inner detector parts are the high particle rates, the radiation tolerance needed and the control of ageing effects. The ATLAS and CMS trackers had to be designed to withstand high radiation doses (500–1,000 kGy) for the innermost pixel layers, and up to 100 kGy for the systems farther away from the interaction point, after 10 years of operation). As a result, the development of the integrated front-end electronics for these systems has been a major problem to solve over several years and design iterations. These circuits must be fast, radiation tolerant and low power, and are integrated on low-mass modules where cooling and material limitations are severe. Several rounds of testbeam measurements and rigorous irradiation programmes have been necessary to prove that the circuits will function in their final assemblies, as well as after high irradiation.

A similarly stringent research and development programme was needed for the silicon sensors themselves[10,11], for which the major difficulty is bulk radiation damage. The relevant parameter is the accumulated dose in the volume of the inner detector, which varies between 10^{15} and 10^{14} cm^{-2} from the innermost layers to the outer ones (where n_{eq} means the number of equivalent particles, normalized using non-ionizing energy loss cross-sections to the damage expected to be caused by 1 MeV neutrons). These radiation doses have severe consequences for the silicon sensors (as they do for all other module components and for the thermal design of the system). They cause increased leakage current and changes in effective doping, and therefore changes in depletion and operation voltages, leading to type inversion for n-type sensors. After type inversion, the effective doping also shows an increase with time following irradiation (reverse annealing) that is temperature dependent. To maintain the operation voltage within reasonable limits, the sensors are therefore kept cold (–10 °C to 0 °C) throughout their lifetime, which has the added benefit of reducing the leakage current.

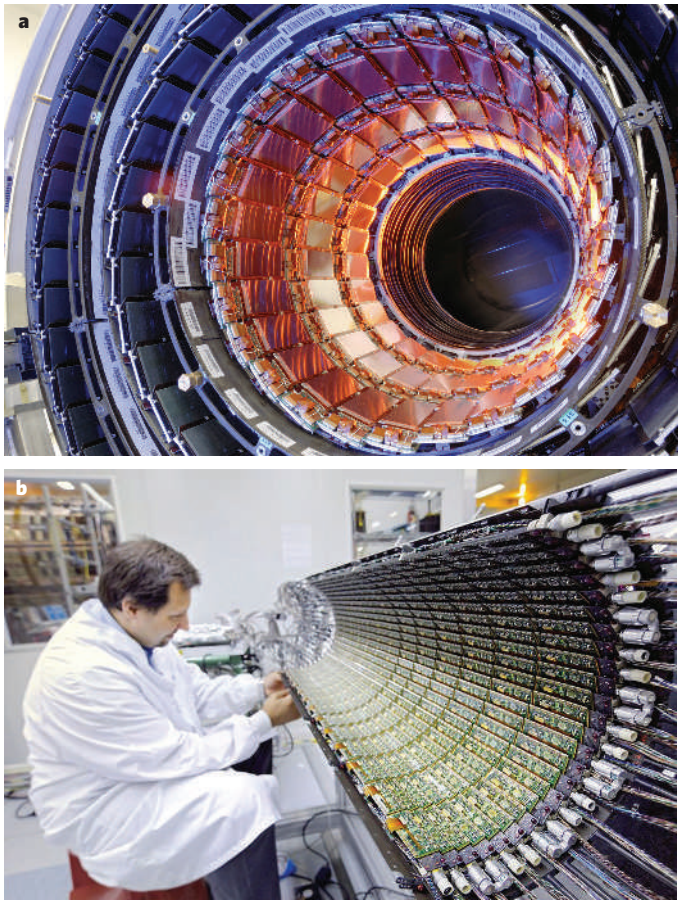


Figure 2 | Tracking systems. a, Silicon microstrip detectors in the CMS barrel region. **b**, The ATLAS pixel detector during the final assembly stage. Images reproduced with permission from CERN.

All of these effects have been carefully mapped out, and various design options have been evaluated in prototypes. Of particular interest are n-i-n silicon sensors, in which the charge-collection region grows with bias voltage from the n-implant side after type inversion following irradiation. Therefore, high efficiency can be obtained from an under-depleted detector. This allows a system to be specified with a lower maximum operating voltage. The ATLAS and CMS pixel systems and the LHCb vertex detector use such sensors. These require double-sided processing and are relatively complex and costly, so for the large-area silicon-strip systems, simpler, single-sided p-i-n designs have been adopted.

Considering the flux of charged particles at increasing radii around the LHC beams, three detector regions are defined in ATLAS and CMS. In the first of these, closest to the interaction point where the particle flux is highest, there are silicon pixel detectors (Fig. 2b), whose cell sizes of 50×400 μm^2 and 100×150 μm^2 in ATLAS and CMS, respectively, give an occupancy of about 10^{-4} per pixel per bunch crossing. To improve the measurement of secondary vertices (typically from b-quark decays) an innermost layer of pixels has been introduced as close to the beam as is practical, at a radius of about 4.5 cm. The lifetime of this layer will

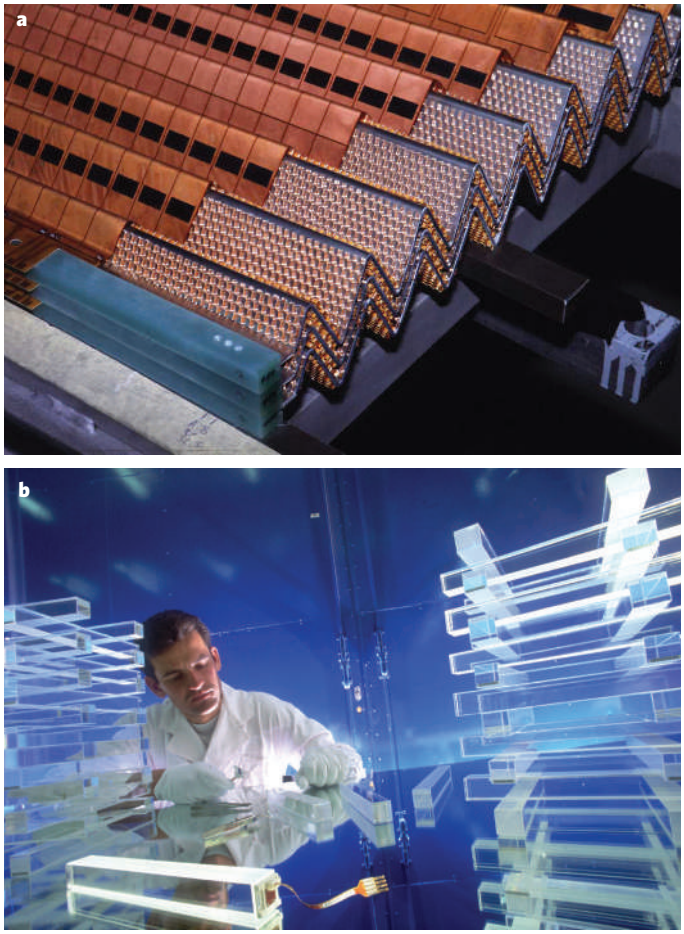


Figure 3 | Calorimetry: different approaches. a, The layers of the ATLAS electromagnetic calorimeter have an 'accordion' geometry. **b**, Tens of thousands of lead tungstate crystals have been prepared and tested, before being assembled into the electromagnetic calorimeter of the CMS detector. Images reproduced with permission from CERN.

be limited, owing to radiation damage, and it may need to be replaced after a few years. The pixel systems in ATLAS and CMS are very much larger than any comparable existing system. The ATLAS pixel system covers about 2 m^2 and has 80 million channels; the CMS pixel system is only slightly smaller.

Pixel detectors are expensive and have high power density, so at a certain radius and system size, silicon microstrip systems become the preferred technology. In the intermediate tracking region of ATLAS and CMS, at a radius of 20–55 cm, the particle flux becomes low enough to use silicon microstrip detectors. Barrel cylinders and endcap discs provide coverage out to about $|\eta| = 2.5$. Strip dimensions of 10–12 cm × 80–120 μm lead to an occupancy of 1–3% per bunch crossing. Both trackers use stereo angle in some of the strip layers (that is, strips placed at a small angle with respect to the z axis, 40 mrad and 100 mrad for ATLAS and CMS, respectively) to improve the resolution in z. In these microstrip systems, it has been essential to find a good balance between the pitch of the cells (determining resolution and occupancy), radiation effects, capacitive load (noise), material length and costs.

Finally, in the outermost region (beyond about 55 cm), the particle flux has dropped sufficiently to allow the use of larger-pitch silicon microstrips in the CMS tracker, with a maximum cell size of 25 cm × 180 μm while keeping the occupancy to about 1%. There are six layers of such microstrip modules in the barrel, accompanied by nine endcap discs providing coverage out to about $|\eta| = 2.5$, amounting to 15,400 modules and 9.6 million channels, and spanning a total detector area of more than 200 m^2.

For ATLAS, at radii greater than 56 cm, a large number of tracking points (typically 36 per track) is provided by the 'straw' tracker — 300,000 straw-tubes embedded in fibre or foil radiators and filled with a xenon-based gas mixture. This detector allows continuous track following with less material per point, and also has electron identification capabilities. X-ray photons are produced through transition radiation because highly relativistic particles such as electrons traverse the straw tracker's multiple interfaces.

Another silicon microstrip detector at LHC, the Vertex Locator[12] of LHCb, has some special features that present significant challenges. The 42 double-sided, half-moon-shaped detector modules are placed at a radial distance from the beam (8 mm) that is smaller than the aperture required by the LHC during injection and must therefore be retractable. For minimizing the material between the interaction region and the detectors, the silicon sensors are inside a thin aluminum box at a pressure of less than 10^{-4} mbar (10^{-2} Pa). The n-i-n sensors used have rφ geometry with pitch (40–140 μm) depending on the radius.

Calorimeters

The calorimeters absorb and measure the energies of electrons, photons and hadrons. In the design of the electromagnetic calorimeters for both ATLAS and CMS, the emphasis is on good resolution for photon and electron energy, position and direction measurements, and wide geometric coverage (up to $|\eta|$ close to 3.0). In the QCD-dominated environment of the LHC, the ability to reject neutral pions is crucial for photon and electron identification. It is also important to have efficient photon and lepton isolation measurements at high luminosities. For the hadronic calorimeters, the emphasis is on good jet-energy measurements, and full coverage (to $|\eta| = 5$) to be able to ascribe the observation of significant missing transverse energy to non-interacting particles (such as neutrinos, or light neutralinos from supersymmetric-particle cascade decays) rather than to losses in the forward regions. Last but not least, the quantities measured in the calorimeters play a crucial part in the trigger of the experiment as signatures of significant parts of the new physics sought at the LHC.

These considerations bring stringent requirements for high granularity and low noise in the calorimeters. The major technical difficulties for the calorimeters are related to the radiation doses (reaching 200 kGy for the electromagnetic part and 1,000 kGy for the hadronic part at the highest η), the sampling speed and the dynamic range needed to measure with low noise and good resolution over a wide energy range.

The ATLAS calorimetry consists of an electromagnetic calorimeter covering the pseudorapidity region $|\eta| < 3.2$, a hadronic barrel calorimeter covering $|\eta| < 1.7$, hadronic endcap calorimeters covering $1.5 < |\eta| < 3.2$, and forward calorimeters covering $3.1 < |\eta| < 4.9$, as shown in Fig. 1c. Over the pseudorapidity range $|\eta| < 1.8$, a presampler is installed in front of the electromagnetic calorimeter to correct for energy loss upstream.

The electromagnetic calorimeter system consists of layers of lead (creating an electromagnetic shower and absorbing particles' energy), interleaved with liquid argon (providing a sampling measurement of the energy-deposition) at a temperature of 89 K. The system's 'accordion' geometry[1,2] provides complete azimuthal symmetry, without cracks, and has been optimized for the high sampling rate environment of the LHC (Fig. 3a). The barrel section is sealed within a barrel cryostat, which also contains the central solenoid, surrounding the inner detector. The endcap modules are contained in two endcap cryostats that also contain the endcap hadronic and forward calorimeters.

The hadronic barrel calorimeter is a cylinder divided into three sections: the central barrel and two identical extended barrels. It is again based on a sampling technique, but uses plastic scintillator tiles embedded in an iron absorber. The vertical tile geometry makes it easier to transfer the light out of the scintillator to photomultipliers and achieves good longitudinal segmentation.

At larger pseudorapidities, closer to the beam pipe where higher radiation resistance is needed, liquid-argon technology is chosen for all calorimetry, for its intrinsic radiation tolerance. The hadronic endcap calorimeter is a copper/liquid-argon detector with parallel-plate

geometry, and the forward calorimeter is a dense liquid-argon calorimeter with rod-shaped electrodes in a tungsten matrix.

The approximately 200,000 signals from all of the liquid-argon calorimeters leave the cryostats through cold-to-warm feedthroughs located between the barrel and the extended barrel tile calorimeters, and at the back of each endcap. The barrel and extended barrel-tile calorimeters both support the liquid-argon cryostats and act as the flux return for the solenoid.

The CMS calorimeter system contrasts with that of ATLAS, because of its compactness[3,4]. In CMS, the solenoid is positioned outside the calorimeter, reducing the material in front of it, but also limiting the thickness of the calorimeter itself, and in particular the number of interaction lengths available to absorb hadronic showers.

The electromagnetic calorimeter, with coverage in pseudorapidity up to $|\eta| < 3.0$, comprises around 80,000 crystals of lead tungstate (Fig. 3b). These crystals have a high density and short radiation length, which make for a compact and high-resolution calorimeter. The main challenges are related to ageing/radiation effects and to temperature control (to the level of a tenth of a degree) to make full use the excellent intrinsic resolution of the system. The scintillation light is detected by silicon avalanche photodiodes in the barrel region and by vacuum phototriodes in the endcap region. A 'pre-shower' system, of silicon strip and lead layers, is installed in front of the endcaps to aid rejection of neutral pion signatures.

Surrounding the electromagnetic calorimeter is a brass/scintillator sampling hadron calorimeter, with coverage up to $|\eta| < 3.0$. Brass has a relatively short interaction length, is easy to machine and is non-magnetic. The scintillation light is converted by wavelength-shifting fibres embedded in the scintillator tiles and channelled to novel photodetectors known as hybrid photodiodes, which can operate in high axial magnetic fields.

Even with such compact electromagnetic and hadronic calorimeters in the barrel region, the total interaction length is limited to 7.2 λ (where λ is the interaction length or mean free path of a particle in the material) at $\eta = 0$ inside the solenoid coil. For this reason, a 'tail catcher' has been added around the coil to complement the hadronic calorimetry and to provide better protection against the escape (or 'punch-through') of hadronic energy into the muon system beyond.

The CMS forward calorimeter, constructed from steel and quartz fibres, is situated 11 m from the interaction point, thereby minimizing the amount of radiation and charge density in the detector during operation. The Čerenkov light emitted in the quartz fibres is detected by photomultipliers. The forward calorimeters ensure full geometric coverage up to $|\eta| = 5.0$, for the measurement of the transverse energy in the event and forward jet measurements.

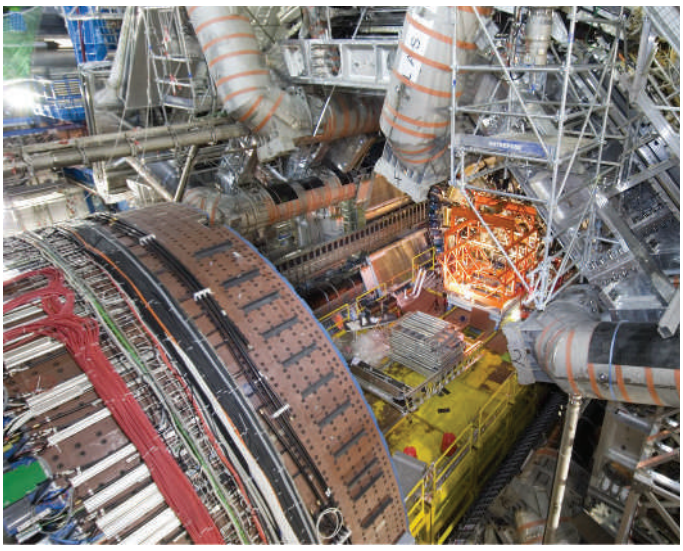


Figure 4 | Final integration. The components of the ATLAS detector are installed in the experiment's underground cavern. Here, part of the inner detector has just been moved inside the barrel calorimeter and toroid systems, while the endcap calorimeters (in the foreground) are kept in an open position to allow access. Image reproduced with permission from CERN.

Muon systems

The outermost detector layers of ATLAS and CMS, and the farthermost layers of LHCb, are dedicated to the measurement of the directions and momenta of high-energy muons, which escape from the calorimeters. Muons form a robust, clean and unambiguous signature of much of the physics of interest at the LHC.

Both ATLAS and CMS have had their overall detector designs optimized and adapted to trigger on and reconstruct muons at the highest luminosities of the LHC. (LHCb will operate, deliberately, at lower luminosity.) Muons must be measured with high efficiency and momentum resolution at low energies (such as in *B*-physics studies), at intermediate energies (for example, in the search for a Higgs decay into four muons), and at very high energies (to identify multi-TeV resonances such as a *Z*′). Wide pseudorapidity coverage is also important, and the ability to trigger on muons with energies of 5–10 GeV is crucial for several of the key physics goals. Finally, good timing resolution and the ability to identify in which proton-bunch crossing the muons were produced are absolute requirements, putting important constraints on the technological solutions chosen for the muon systems.

The muon systems in all three experiments are large-area gas-based detectors (several thousand square metres of multilayer chambers each in ATLAS and CMS). The chambers are divided into two sets, one intended for precise measurements of muon tracks and the other dedicated to triggering on muons. The sheer size of the systems means that there are significant technical challenges related to the stability and alignment of the chambers and to the careful mapping of the detectors' magnetic fields over large volumes. The radiation levels for the muon chambers are much less severe than for the inner detectors or calorimeters, but there are still concerns about ageing of the systems and also the neutron radiation environment of the experimental halls in which the detectors sit. The designs of the beam pipe and the shielding elements in the forward direction have been carefully optimized to reduce the neutron-induced background rates in the muon chambers.

Although the muon-chamber technologies chosen for ATLAS and CMS have many similarities, the magnet configuration in the two experiments is quite different. The ATLAS air-core toroid system, with a long barrel and two inserted endcap magnets, generates a large-volume magnetic field with strong bending power within a light and open structure. Multiple-scattering effects are thereby minimized, and excellent muon momentum resolution is achieved with three stations of high-precision muon-tracking chambers, covering up to $|\eta| = 2.7$. Over most of the range in pseudorapidity, the measurement of track coordinates (in the principal bending direction of the magnetic field) is performed by monitored drift tubes[1,2]. This technology provides robust and reliable operation, thanks to the mechanical isolation of each sense wire from its neighbours in the gas-filled drift volumes of the individual tubes. At large pseudorapidities and close to the interaction point, where the rate and background conditions are more difficult, cathode-strip chambers with higher granularity strip readout are used. The muon trigger system, with a fast time response and covering $|\eta| < 2.4$, comprises resistive-plate chambers in the barrel and thin gap chambers in the endcap regions. As well as triggering, these chambers provide a measurement of a second track coordinate orthogonal to the one measured by the high-precision chambers. In addition to the muon-chamber measurements, the inner detector measurements in the central solenoid of ATLAS contribute to the combined muon momentum resolution of the experiment.

The superconducting solenoid inside CMS is, at 13 m long with a 5.9-m inner diameter, the largest of its kind[3,4]. To achieve good momentum resolution without making overly stringent demands on muon-chamber resolution and alignment, the solenoid will operate at a high magnetic field of 4 T. In CMS, centrally produced muons are measured three times: in the inner tracker, after the coil, and in the return flux, into which four muon 'stations' are integrated, achieving robustness

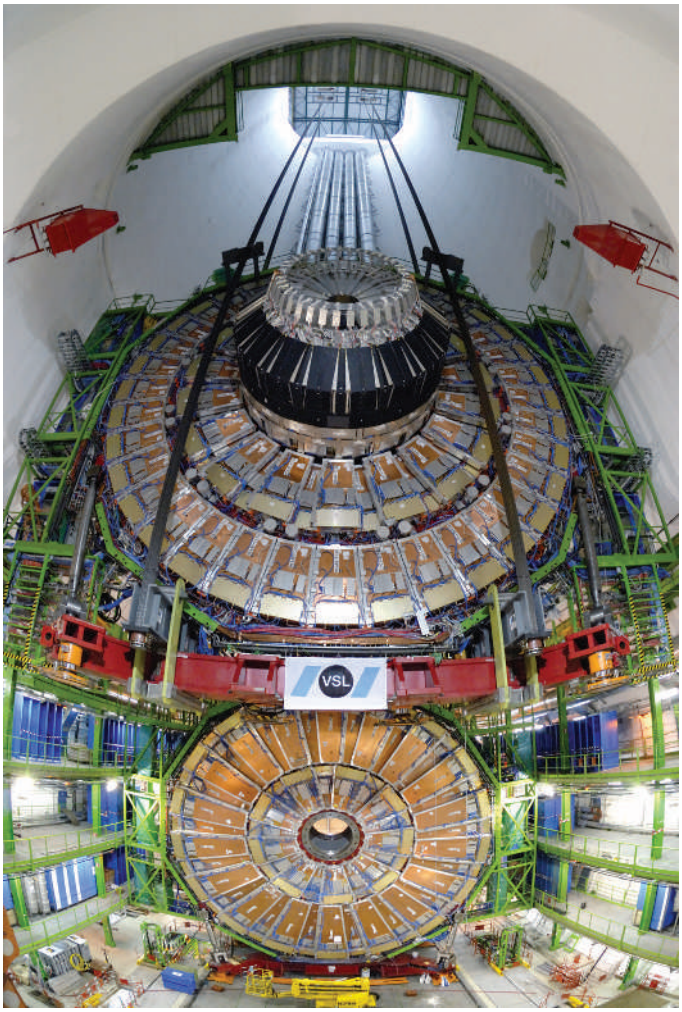


Figure 5 | Going underground. A large unit of the CMS detector, an endcap disc with muon chambers and part of the hadronic calorimeter, is lowered 100 m into its final position in the cavern. Image reproduced with permission from CERN.

and full geometric coverage. Three types of gaseous detector are used: drift tubes in the barrel region ($|\eta| < 1.2$), where the neutron-induced background is small, the muon rate is low and the residual magnetic field in the chambers is low; cathode-strip chambers in the two endcaps ($|\eta| < 2.4$), where the muon rate, the neutron-induced background and the magnetic field are all high; and resistive-plate chambers, in both the barrel and the endcap regions, to provide a fast response with good time resolution and to identify unambiguously the correct bunch crossing (albeit with coarser position resolution).

The intrinsic resolution of the high-precision chambers is in the range 60–150 μm (ref. 6), but the overall performance over the large areas involved (particularly at the highest momenta) depends on how well the muon chambers are aligned with respect to each other and with respect to the overall detector. The high accuracy of the ATLAS stand-alone muon measurement necessitates a precision of 30 μm on the alignment; in CMS, the different muon chambers need to be aligned with respect to each other and to the central tracking system to within 100–500 μm.

Both experiments have intricate hardware systems installed that are designed to measure the relative positions of chambers that contribute to the measurement of the same tracks, but also to monitor any displacements during the detector operation. For instance, in ATLAS, about 5,000 optical alignment sensors and 1,800 magnetic field sensors will track the movements of the chambers and will map and track the magnetic field to an accuracy of approximately 20 G (2 mT) throughout the detector volume. In the CMS, the solenoid magnetic field is more uniform but nevertheless the field is carefully monitored by around 80 sensors. Around 1,400 alignment sensors will provide independent monitoring of the tracking detector geometry with respect to an internal light-based reference system. The final alignment values will be obtained with the large statistics of muon tracks traversing the muon chambers.

The LHCb muon system consists of five stations. Multiwire proportional chambers are used throughout, except for the innermost region closest to the beamline of the first station. This station is placed in front of the calorimeters and represents a significant challenge in terms of material budget, space constraints, rate capability and radiation tolerance. The innermost region of this station, where the particle rates are highest, is equipped with triple-GEM (gas electron multiplier) detectors[13] with pad readout that are particularly suited for tracking in a high particle rate environment.

Triggering and readout

At design luminosity, the LHC will create 10^9 proton–proton events per second, but data storage and processing capabilities are such that data from only about 100–200 carefully selected events per second (each of these interesting events is accompanied by an average of 25 overlapping proton–proton events in the same bunch crossing) can be recorded offline for complete analysis. Hence, there is a need for a trigger system to select only the most important physics signatures and achieve a rejection factor of nearly 10^7.

The trigger systems for the LHC experiments have distinct levels. The first level, based on custom-built processors, uses a limited amount of the total detector information to make a decision in 2.5/3.2 μs (ATLAS/CMS) on whether to continue the processing of an event or not, reducing the data rate to around 100 kHz. Higher levels, using a network of several thousand commercial processors and fast switches and networks, access gradually more information and run algorithms that resemble offline data analysis to achieve the final reduction. The total amount of data recorded for each event will be roughly 1.5 megabytes, at a final rate of 150–200 Hz. This adds up to an annual data volume of the order of 10 petabytes for the LHC experiments.

The challenges to be faced in real-time data collection and reduction are many. The synchronization of the individual parts of the detector — and there are several thousand units to time in — must be accurate to better than a nanosecond, taking into account the flight times of particles to the individual sensor elements. At later stages, there is the second synchronization challenge of assembling all of the data for a particular bunch crossing from various parts of the detector into a complete event.

There is no chance of processing and selecting events within the 25 ns available between successive proton-bunch crossings. Furthermore, the sizes of the detectors and of the underground caverns in which they sit impose a minimum transit time between the detector electronics and trigger electronics. The first-level trigger calculations themselves need to be sufficiently sophisticated to identify clear physics signatures; the decision is based on the presence in the calorimeters or muon detectors of 'trigger primitive' objects, such as photons, electrons, muons and jets above pre-set transverse-energy or transverse-momentum thresholds. It also employs global sums of transverse energy and missing transverse energy. During the transit and processing time — less than 2.5/3.2 μs for ATLAS/CMS — the detector data must be time-stamped and held in buffers.

After an event is accepted by the first-level trigger system, the data from the pipelines are transferred from the detector electronics into readout buffers. The further processing involves signal processing, zero suppression and data compression while the events are examined by a farm of commodity processors consisting of several thousands of central processing units. The design and implementation of the processor farm, switching network, control software and trigger application software are major challenges. The event fragments must be directed from the readout buffers to a single processor and buffer node, using fast switches and networks, in order to perform more detailed calculations of the critical parameters of the event and to reduce the final rate further.

Even after such a large online reduction, huge amounts of data will be recorded. It was soon clear that the required level of computing resources

could be provided only by a significant number of computing centres working in unison with the CERN on-site computing facilities. Off-site facilities will be vital to the operation of the experiments to an unprecedented extent. Hence, over the past five years, GRID infrastructure for data processing and storage has been developed[14,15].

The GRID solution is geographically distributed and relies on three tiers[6]. The various tiers have clear responsibilities. The raw data output from the high-level trigger is processed and reconstructed in a Tier 0 computing facility at CERN, producing reconstructed data. Most detector and physics studies, with the exception of calibration and alignment procedures, will rely on this format. A copy of the raw data plus the reconstructed data is then sent to the Tier 1 centres around the world. These share the archiving of a second copy of the raw data, provide the reprocessing capacity and access to the various streams of reconstructed data (corresponding to the major trigger signatures) and allow data access and processing by the experiments' physics-analysis groups. The analysis data produced at Tier 0 and 1 are derived from the reconstructed data and are a reduced-event representation, intended to be sufficient for most physics analyses.

The Tier 2 facilities, each linked to a specific Tier 1 centre, are smaller but more numerous and are used for analysis, data-calibration activities and Monte Carlo simulations. Furthermore, the Tier 1 analysis data are copied to Tier 2 to improve access to them for physics analysis; by contrast, the Tier 1 centres provide safe storage of the large data sets produced at Tier 2 (for example, simulation data). A final level in the hierarchy is provided by individual group clusters and computers used for analysis.

This machinery for processing and analysis is being set up, and tests already indicate that the transfer speed between CERN and Tier 1 that is essential for initial running can be achieved, and that large-scale data production can be carried out in the GRID framework.

Ready to start

Installation of the LHC detectors in their underground caverns began in 2004. The components of ATLAS have been assembled and tested in the experiment's cavern, at 'Point 1' on the LHC ring (Fig. 4), and LHCb in the cavern at 'Point 8'. By contrast, the bulk of the CMS system was assembled and tested on the surface, before being lowered 100 metres into its cavern at 'Point 5' in 15 large lift operations (Fig. 5).

All of the LHC experiments have been operating their detector elements as much as possible on the surface and in 'test beams', and also, following installation, in the underground caverns. The flux of muons from cosmic rays provides a useful test of systems such as the calorimeters, inner detectors and muon systems to check alignment, calibration and the integration of data collection. In parallel, the GRID computing infrastructure and organisation are being planned, implemented and tested. For ATLAS, CMS and LHCb, major exercises of their data processing, software and computing infrastructure have been performed, and more are planned in the run-up to the introduction of beams into the LHC in 2008.

The first collisions at a centre-of-mass energy of 14 TeV are expected by mid-2008. As soon as a luminosity of 10^{32} $cm^{-2}s^{-1}$ is reached, within days the LHC can produce data sets — of W and Z bosons, t quarks, high-transverse-momentum jets, and even supersymmetric particles — that will surpass those of any previous or existing accelerator. At that point, the main physics goals of the LHC will be in full focus, and the aim will be to collect as much data in 2008 as possible.

The decade-long period of detector development and construction is coming to an end. Many of the fundamental challenges addressed by the experiment builders at LHC have been solved successfully, most notably in the development of fast, radiation-tolerant and sufficiently granular detector systems and electronics, integrated in turn in large systems that exceed any existing detector in specification (size, speed and channel count, in particular). Technology advances in computing, switches, networks and software have allowed the development of sophisticated trigger, data acquisition and GRID systems to handle the LHC data rates and volumes — it was far from obvious that this could be achieved when the building of the experiments was initially approved. The accelerator, detectors and off-line systems now need to be completed in their underground areas, commissioned fully and operated efficiently.

A further challenge overcome in the LHC project is the successful collaboration in each experimental team of as many as 2,000 scientists from all over the world, working together for a decade and using their resources and skills efficiently. Thanks to their efforts, and those of the teams building the LHC itself, a new era of research in experimental particle physics is finally within reach. The community can now look forward to the new challenges posed in interpreting the data from the LHC — challenges as great as those that have been faced in the building of the detectors. There is good reason to be optimistic, and the potential rewards, in terms of physics discoveries, make it well worth the effort. ■

1. ATLAS Collaboration. *ATLAS Detector and Physics Performance: Technical Design Report*. Report No. CERN/LHCC/99-15 (CERN, Geneva, 1999).
2. ATLAS Collaboration. *ATLAS: Technical Proposal for a General-Purpose p p Experiment at the Large Hadron Collider at CERN*. Report No. CERN/LHCC/94-43 (CERN, Geneva, 1994).
3. CMS Collaboration. *CMS Physics Technical Design Report* Vol. I. Report No. CERN/LHCC/2006-01 (CERN, Geneva, 2006).
4. CMS Collaboration. *CMS Technical Proposal*. Report No. CERN/LHCC/94-38 (CERN, Geneva, 1994).
5. Ellis, N. & Virdee, T. S. Experimental challenges in high luminosity collider physics. *Annu. Rev. Nucl. Part. Sci.* **44,** 609–653 (1994).
6. Froidevaux, D. & Sphicas, P. General purpose detectors for the Large Hadron Collider. *Annu. Rev. Nucl. Part. Sci.* **56,** 375–440 (2006).
7. Gianotti, F. *European School of High-Energy Physics, CERN Yellow Report*. Report No. CERN-2000-007 219–244 (CERN, Geneva, 1999).
8. LHCb Collaboration. *LHCb Technical Proposal*. Report No. CERN/LHCC/98-104 (CERN, Geneva, 1998).
9. Yao, W.-M. *et al*. Review of particle physics. *J. Phys. G* **33,** 1 (2006).
10. Doležal, Z. *et al*. The silicon microstrip sensors of the ATLAS semiconductor tracker. *Nucl. Instrum. Methods* **578,** 98–118 (2007).
11. CMS Collaboration. *Addendum to the CMS Tracker TDR*. Report No. CERN/LHCC 2000-016 (CERN, Geneva, 2000).
12. LHCb Collaboration. *LHCb TDR 5* Report No. CERN/LHCC 2001-011 (CERN, Geneva, 2001).
13. LHCb Collaboration. *Second Addendum to the Muon System Technical Design Report*. Report No. CERN/LHCC/2005-0012 (CERN, Geneva, 2005).
14. Foster, I. & Kesselman, C. *The Grid: Blueprint for a New Computing Infrastructure* (Morgan & Kaufmann, San Francisco, 1998).
15. LCG project. *LHC Computing Grid Technical Design Report*. Report No. CERN/LHCC/2005-024 (CERN, Geneva, 2005).

Acknowledgements This review article is largely based on the very complete documentation already existing for the ATLAS, CMS and LHCb experiments, and on refs 1–8 in particular. Furthermore, with around 5,000 scientists involved in the construction of the experiments described, I can only cover a small part of the challenges, excitement and difficulties involved in making them a reality. The best I can do is therefore to acknowledge all of the members of these collaborations, and in particular those who have helped me with corrections and comments, and apologize for all the dedicated sub-projects and work I had to leave out of this review.

Author information Reprints and permissions information is available at npg.nature.com/reprintsandpermissions. The author declares no competing financial interests. Correspondence should be addressed to the author (steinar.stapnes@cern.ch).