Particle Physics at Colliders

Brian Shuve

Contents

1 Why Particle Colliders?

The goal of particle physics is to discover and characterize the fundamental building blocks that make up all matter. One of the most important scientific breakthroughs of the 19th Century was the discovery that everything in the universe was made up of only about 100 different elements. The rich diversity of materials that we see in the world come from assembling those building blocks in different ways; understanding how and why these pieces fit together has completely transformed our understanding of materials and led to, for example, the development of computers and the field of nanoengineering.

In the late 19th and early 20th centuries, studies showed that atoms themselves were made up of smaller constituent pieces such as protons and electrons, and their interactions could be described by a very short list of forces: electromagnetism, the weak and strong forces, and gravity. It is therefore natural to ask if protons and electrons are themselves made up of more fundamental particles. The answer (for protons) is yes, and particle physicists seek to understand the properties of these particles and discover whether there are any particles or forces in nature that we have not yet observed.

In order to understand how we study the properties of elementary particles, it is instructive to first consider an older and potentially more familiar example: the discovery of the atomic nucleus. In 1900, the existence of electrons was known and atoms were understood to be neutral, but it was not clear how the balancing positive charge was located in the atom. One theory, known as the "plum pudding" model, was that the positive charge of the atom was distributed uniformly inside of the atom (see Fig. 1). For those who have taken electromagnetism, Gauss' Law tells us that the electric field of a spherical ball of radius R at a distance d from the centre is determined by the enclosed charge, and so the electric field of the uniform positive charge in the plum pudding model is

$$\vec{E}(d) = \frac{k\,Q\,d}{R^3}\,\hat{r}, \tag{1}$$

where Q is the positive charge of the atom and k is a constant that characterizes the strength of the electromagnetic force. The vector $\hat{r}$ points radially out from the centre of the charge. According to this equation, the repulsive electric field actually gets smaller as you go further inside the ball of charge. Thus, if a high-momentum positively charged particle is shot at the atom, it would experience a progressively smaller repulsion as it traversed towards the centre of the atom, and consequently its trajectory would be only slightly deflected by electrostatic repulsion (as illustrated in the right pane of Fig. 1).

This should be contrasted with the hypothesis that the positive charge of the atom is concentrated in tiny clumps. In this case, we can treat the positive charge as a point charge at the centre of the atom, and the resulting electric field is given by Coulomb's Law,

$$\vec{E}(d) = \frac{k\,Q}{d^2}\,\hat{r}. \tag{2}$$

In this case, the closer a particle gets to the positive charge of the atom, the *larger* the repulsive force. In particular, a very large force can be applied that leads to a signifiant change to the trajectories of particles fired directly at the centre of the atom.

To test which idea is correct, Rutherford and his students, Geiger and Marsden, took a source of α particles, which are helium nuclei produced in radioactive decay with charge $+2e$ (where $-e$ is the electron charge). They directed the α particles towards a thin foil of gold (see Fig. 2). If the plum pudding model were correct, then each α particle would experience a small deflection. If instead the point-particle model of the nucleus were correct, some small fraction of the α particles passing close to the centre of the atom would experience an enormous repulsion and experience a large deflection.

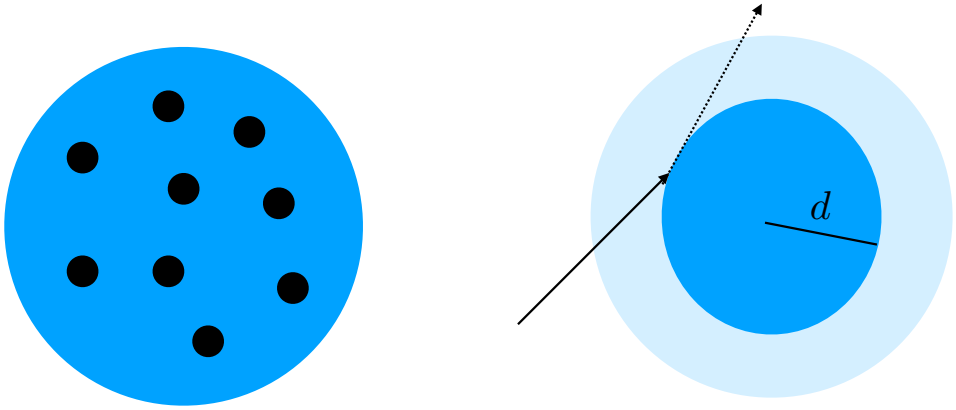

Figure 1: (Left) The "plum pudding" model of the atom, where the electrons (black blobs) are embedded in a positively charged matrix. It derives its name from the resemblance to a dessert with raisins embedded in it, where the raisins are the electrons. (Right) According to Gauss' Law, the electric field of a spherical object is proportional to the enclosed charge in accordance with Eq. (1). Therefore, the electric field is smaller for a particle passing through the haze of positive charge, and the plum pudding model only tends to predict small deflections of charged particles when fired at atoms.

A fluorescent screen was set up which gave off flashes of light when struck by α particles, and they used this light to measure where the α particles hit the screen and reconstruct their trajectories. The experimental setup is shown in Fig. 2.

What Rutherford *et al.* found was very different from the prediction of the plum pudding model. While most α particles passed straight through the gold foil, some were deflected at large angles (with some even larger than 90°!). This confirmed the point-particle model in which the positive charge of an atom is concentrated in the atomic nucleus: the large deflections of the α particles comes from α particles that get very close to the nucleus and experience a huge repulsive force. We now know that the positive charge is contained inside of protons, and they are confined within a nucleus; the radius of the gold nucleus is about 7×10^{-15} m (7 fm, or femtometres), to be compared with the atomic radius of 0.2 nm, and so the nucleus is approximately located at a point compared to the atom's size.

This example shows that particle scattering can provide information about the structure of matter. In fact, numerical predictions can also be made based on our theoretical understanding of known forces and compared with experiment. For example, Rutherford used electromagnetic theory to compute the number of deflected α particles per unit area at an angle ϕ relative to the incident beam:

$$\frac{N(\phi)}{\text{area}} = \frac{X\,n\,t}{16r^2\,\sin^4(\phi/2)}\left(\frac{2Q_\alpha Q_n}{mv^2}\right)^2, \tag{3}$$

where r is the distance between the source of α particles and the point of collision with the foil, X is the number of α particles hitting the foil, n is the number of atoms per unit volume in the foil, t is the foil thickness, Q_n is the charge of the nucleus, $Q_\alpha = +2$ is the alpha particle charge, and m and v are the mass and speed of the α particle. This formula can be compared to experiment; any deviation from the prediction suggests some combination of a) the nucleus is not pointlike; b) the electromagnetic force of a point charge is not an inverse square law (*i.e.,* Coulomb's Law is wrong); c) there is a new force beyond electromagnetism; d) there are some new particles that the α particle

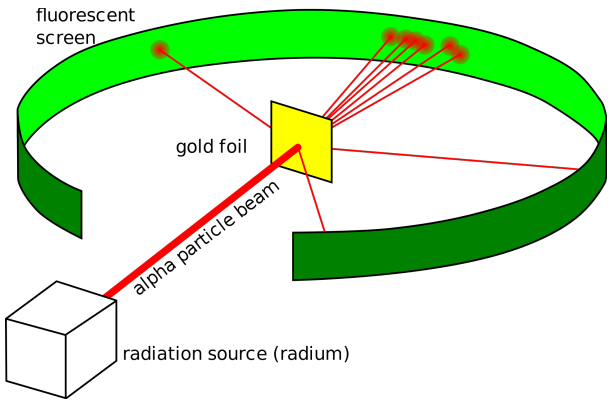

Figure 2: Schematic diagram of the Rutherford-Geiger-Marsden experiment, which discovered that the positive charge of atoms is concentrated in an atomic nucleus. Positively charged α particles are directed at gold atoms, and the positive charge distribution of the gold atoms is determined by observing whether the α particles undergo grazing collisions or wide-angle scattering (image source: Wikimedia Commons).

scatters against; etc. Thus, scattering can teach us a lot, both about the forces we know about as well as those we don't!

It is reasonable to wonder whether all of this is necessary. After all, most of what we know about the world comes just from using our sense of sight. Can't we just "look" inside of an atom (maybe using a powerful microscope to detect light scattered off the atom) instead of throwing things at it to see what comes out? The answer, unfortunately, is no: the kind of light that our eyes can see has a wavelength $\lambda \sim \mathcal{O}(100 - 1000 \text{ nm})$, whereas the atomic radius is much smaller than this (< 1 nm). The atom can affect the light wave (for example, by absorbing and re-emitting some of the light), but because the wavelength of light is much larger than the atom's size, the light is only sensitive to the atom as a whole and not its constituents. To be sensitive to the nuclear size, we would need light of a much smaller wavelength of 1 fm. Planck tells us that the characteristic energy of light is inversely proportional to the wavelength,

$$E = \frac{hc}{\lambda}, \tag{4}$$

where h is Planck's constant and c is the speed of light, and so we need light whose photons carry an energy that is over a thousand times more energetic than visible light, namely X-rays or γ rays. **In general, if we wish to look at smaller objects, we need to look at scattering of particles at ever larger energies.**

In the Rutherford-Geiger-Marsden experiment, α particles were used instead of light in the scattering experiment. Since α particles are helium nuclei, they are made up of protons and neutrons which are more commonly thought of as particles instead of waves. Is there a loophole that allows us to avoid going to high energies when probing small distances if we use particles? The answer again comes from quantum physics: there is no such thing as a "wave" or a "particle", but rather everything has characteristics of both. Thus, in spite of the fact that they are very "particle-like",

the α particles do indeed have a characteristic wavelength which was determined by de Broglie,

$$\lambda = \frac{h}{p}, \tag{5}$$

where p is the particle momentum. For the characteristic masses and velocities of α particles in the gold foil experiment, the wavelength is $\lambda \sim 10$ fm. Thus, the α particles were able to penetrate into the atoms and uncover their internal structure. However, α particles are insufficient to probe into the structure of the protons themselves because the wavelength is again too large. Quantum theory demands that we go to higher energies to learn about physics on smaller scales[1]. This is why the technical term for the field of particle physics and related subjects is **high-energy physics**. Our modern understanding of particle physics has shown us that particles such as the electron are pointlike down to distances of $\lesssim 0.01$ fm. Therefore, to "look inside" of the known particles we must accelerate and collider particles at truly immense energies.

There is one other reason to study ultra-high-energy collisions: special relativity tells us that the minimum energy that a particle of mass M can have is its rest energy, $E_{\text{rest}} = Mc^2$. In particle physics, we are interested to know whether there exist any heavy particles in nature that we have not yet discovered. For example, the most recently discovered particle is the Higgs boson, with mass $M_{\text{Higgs}} \approx 125 M_{\text{proton}}$. Thus, if we hope to produce a Higgs boson in a collision of protons, then the colliding protons must have energy that is very large relative to their rest energy. This puts us firmly in the domain of special relativity.

1.1 Guide to this Document

The goal of this document is to provide a brief introduction to many of the topics needed to understand the physics of colliders. Depending on what area of research you ultimately get involved in, you will likely need to explore some topics in more depth. However, the hope is that you will know enough on both the qualitative and quantitative level that you can begin to engage with the research literature and research topics.

I suggest that everyone read Sec. 2 and 3, at least in a cursory fashion (although those who have not taken multivariable calculus may have difficulty with some of the mathematical aspects in Sec. 3). These cover important concepts that show up in basically every aspect of particle physics. Pay particular attention to the use of relativistic invariants and natural units, as these form the fundamental building blocks of particle physics and will come up over and over (and over) again. While this material is in principle self-contained, it could be hard to follow if you have not already studied special relativity (or your relativity is a bit rusty).

Sec. 4 gives a background into the types of colliders we use and the importance of each one, as well as how we actually experimentally reconstruct particle collisions. For anyone doing simulations of particle collisions, the must-read section is Sec. 4.4, without which it will be impossible to understand on a practical level the output of computer simulations.

Sec. 5 is for those interested in or working on hadron (proton) colliders, as well as any particle collisions that produce quarks or gluons in the collision. It is meant to provide a qualitative overview of the relevant physics for understanding how these collisions work.

Finally, Sec. 6 focuses on drawing inferences from analyses of particle collisions. This section is still incomplete, but does introduce the Poisson distribution that is essential to the statistics of particle physics experiments.

[1] For more discussion of the correspondence between object sizes and energies, see Appendix A of my "Introduction to Particle Physics" notes.

2 Relativistic Kinematics of Particle Collisions

This section will review the necessary relativistic kinematics for collider physics, and introduce some kinematic variables that are commonly used in collider studies. My focus will be on motivating the ways in which we apply relativity to collider processes and getting as quickly as possible to useful results. See Griffiths, *Introduction to Elementary Particles*, Chapter 3 for a more detailed introduction to the topic; you may also wish to consult Helliwell's *Special Relativity* textbook, which is what we use to teach this topic at Harvey Mudd.

2.1 Four-Vectors and Lorentz Invariants

The theory of special relativity rests on two basic priniciples:

1. The laws of physics are the same for all inertial observers (*i.e.*, non-accelerating frames of reference);

2. The speed of light, c, is the same in all reference frames.[2]

This is a revolutionary theory that is far from our day-to-day experience. For instance, special relativity implies time dilation, length contraction, relativity of simultaneity, that velocities cannot be linearly added together, etc. This is both fascinating and the subject of much confusion (hence the multiplicity of "paradoxes" surrounding special relativity).

Rather than get hung up on details about what changes between different observers in special relativity, it is much more important to talk about what does not change between observers. Special relativity is a causal theory, which means that while two events A and B may appear to occur at different times for different observers, every observer should be able to tell if the action of event A caused event B to occur. For example, if event A is me pushing a button and event B is the Sun exploding, everyone should agree on whether my pushing the button caused the Sun to explode or not. Beyond this fanciful example, different observers should agree on things like: whether a collision occurs; the number of particles produced in the collision; the forces involved in the collision; the nature of the particles produced, etc.

We therefore wish to work as much as possible in terms of quantities that are the same for every observer in describing the underlying physics of particle collisions. We refer to such quantities as **Lorentz-invariant quantities** (or, in short, **Lorentz invariants)**, since they do not change under the Lorentz transformations of special relativity that change quantities from one observer's frame of reference to another. One major convenience of using Lorentz invariants is that it can be simplest to do theoretical calculations in one frame of reference and experiments in another. As long as we compare Lorentz-invariant quantities it doesn't matter which frame we use!

To start, we introduce the Lorentz transformations of time and space coordinates. If an event (like an exploding star, see Fig. 3) occurs at time t and position $\vec{x}$ for an observer in a reference frame $\mathcal{S}$, then for an observer in a different frame $\mathcal{S}'$ moving at velocity $\vec{v} = V\,\hat{x}$ relative to $\mathcal{S}$, the position and time coordinates in the $\mathcal{S}'$ frame are instead

$$t' = \gamma(t - Vx/c^2), \tag{6}$$
$$x' = \gamma(x - Vt), \tag{7}$$
$$y' = y, \tag{8}$$
$$z' = z, \tag{9}$$

[2] Actually, if we accept Maxwell's theory of electromagnetism, then only the first postulate is needed; the second follows as a consequence from the requirement that Maxwell's Equations be the same in every inertial reference frame.

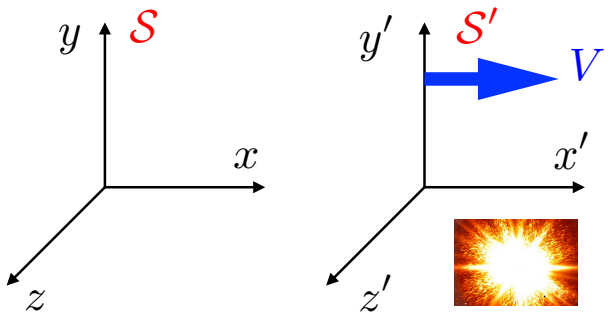

Figure 3: Observers in different frames $\mathcal{S}$ and $\mathcal{S}'$ moving at relative velocity V will disagree on the spatial position and time at which an event occurs, like an exploding star. The spacetime coordinates of the event are related by the Lorentz transformations, Eqs. (6)-(9).

where γ is defined as

$$\gamma = \frac{1}{\sqrt{1 - V^2/c^2}}. \tag{10}$$

We refer to a change from one inertial reference frame to another as a **boost**. In the above example, the boost was performed along the x-axis. A boost along an arbitrary axis can be found by performing a spatial rotation on the coordinates after the above boost.

Normally, in Newtonian mechanics we use three-vectors $\vec{x} = (x, y, z)$ to represent position. However, we see that the transformations between reference frames actually change the time variable as well, and so it is sensible to extend our usual "three-vectors" into **four-vectors** that include the time coordinate. The new position 4-vector, x^μ, carries a label μ which runs from 0 to 3 corresponding to the time and spatial components of the vector:

$$x^\mu \to (x^0, x^1, x^2, x^3) \equiv (ct, x, y, z). \tag{11}$$

Note that we have multiplied time by c to make each of the quantities in the position 4-vector have the same units. The Lorentz transformations then act as a matrix on the 4-vectors,

$$\begin{pmatrix} x'^0 \\ x'^1 \\ x'^2 \\ x'^3 \end{pmatrix} = \begin{pmatrix} \gamma & -\gamma\, V/c & 0 & 0 \\ -\gamma\, V/c & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x^0 \\ x^1 \\ x^2 \\ x^3 \end{pmatrix}. \tag{12}$$

This form of a transformation may be somewhat familiar to you. Consider, for example, the rotation of a 3-vector by an angle θ in the xy-plane (see Fig. 4). The transformation is

$$\begin{pmatrix} x' \\ y' \\ z' \end{pmatrix} = \begin{pmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix}. \tag{13}$$

Rotations of a coordinate system preserve both the lengths of individual vectors as well as the angles between vectors, and so this means that scalar products such as

$$\vec{a} \cdot \vec{b} \equiv a_x b_x + a_y b_y + a_z b_z = |\vec{a}||\vec{b}| \cos\Delta\phi \tag{14}$$

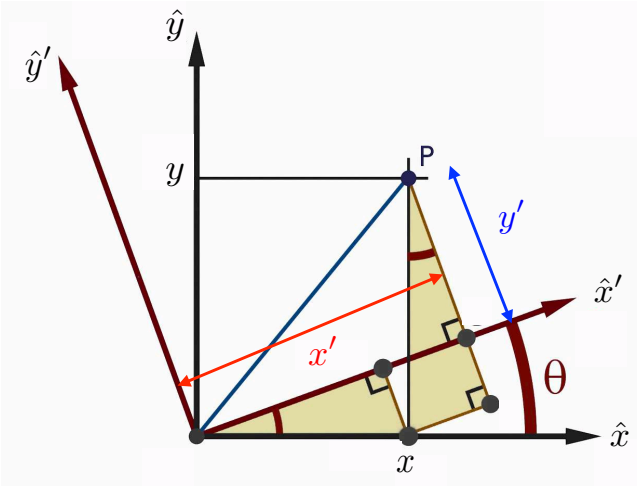

Figure 4: Rotation of the coordinate axes in the xy-plane. The location of the point P in the primed coordinate system relative to the unprimed coordinate system is given in Eq. (13).

are invariant under rotations (here, $\Delta\phi$ is the angle between the vectors). As long as we express physical quantities in terms of these dot products, they will hold in any coordinate systems related by rotations.

Since Lorentz transformations look very similar to rotations, we expect that there should be a scalar or dot product that we can use to combine the vectors into something that is invariant under coordinate transformations. In fact, the main difference between Lorentz transformations and rotations is that there is a relative minus sign on the $(2,1)$ entry of the rotation matrix that does not exist for the Lorentz transformation matrix. This suggests that the expression for the dot product between 4-vectors will have some minus signs relative to the dot product between 3-vectors. It turns out that the product of 4-vectors

$$a \cdot b = a^0 b^0 - a^1 b^1 - a^2 b^2 - a^3 b^3 \tag{15}$$

is a Lorentz-invariant quantity that is the same for any inertial frame, including both observers $\mathcal{S}$ and $\mathcal{S}'$. [3]. We can also denote the squared magnitude of a 4-vector as

$$a^2 \equiv a \cdot a = (a^0)^2 - (a^1)^2 - (a^2)^2 - (a^3)^2 = (a^0)^2 - \vec{a} \cdot \vec{a}; \tag{16}$$

unlike the magnitude of a 3-vector, the magnitude of a 4-vector can be zero or of either sign. The 4-vector dot product is also automatically invariant under spatial rotations, which do not affect a^0 and also do not change the 3-vector dot product $\vec{a} \cdot \vec{a}$.

Exercise 1: Show explicitly that Eq. (15) is invariant under the Lorentz transformation in Eq. (12), which represents boosts along the x direction. Argue that the $a \cdot b$ scalar product is also invariant under boosts in arbitrary directions.

[3] Some books use a different convention: $(a \cdot b)_{\text{alt}} = -(a^0)^2 + (a^1)^2 + (a^2)^2 + (a^3)^2 = -(a \cdot b)_{\text{particle physics}}$. Both are Lorentz invariants, but differ by a minus sign. Practitioners of general relativity and those who study gravity tend to prefer the alternate convention, whereas particle physicists use the convention where the minus sign is in front of the spatial components. We will see why shortly.

2.2 The Energy-Momentum 4-Vector

So far, we have shown that in special relativity, the position 3-vector $\vec{x}$ should actually be embedded into a spacetime 4-vector $(ct, \vec{x})$ because Lorentz transformations scramble the time and position components. In classical non-relativistic physics, we know of many other 3-vectors, such as momentum, velocity, currents, etc. When we consider relativistic physics, these objects also must be embedded into larger classes of vectors.

By far the most important case for us is momentum, so we focus on the treatment of momentum in special relativity. The reason is that momentum is conserved in collisions, and so studies of momentum allowed us to more easily solve for the kinematics of particle scattering. To embed $\vec{p}$ into a 4-vector, we need to associate it with an additional quantity that gets chosen as the 0^{th} component of the 4-vector. This quantity should be *invariant* under rotations, and therefore not have a spatial direction. Looking at Eq. (12), we see that a boost along the x-direction tends to reduce the magnitude of both the 0 and 1 component of the vector. Thus, we expect the 0^{th} component to be something that is reduced when the momentum goes down.

A simple guess is that the relevant quantity for the 0^{th} component of the 4-vector is energy, since the kinetic energy should indeed go down when the momentum decreases. Energy, however, has different units from $\vec{p}$, so we must divide energy by the physical constant c to keep every component of the 4-vector having the same units. This allows us to define the **energy-momentum 4-vector**,

$$(p^0, p^1, p^2, p^3) = (E/c, p_x, p_y, p_z). \tag{17}$$

Note that, at this point, we simply have made an educated guess that the 0^{th} component of the momentum 4-vector is energy; in other words, we have tacked on a quantity E to the momentum 3-vector to form a 4-vector, but we do not yet know whether E actually has *anything* to do with our conventional notion of energy. We must therefore show that E indeed reproduces what we expect for energy from non-relativistic mechanics.

As a first step, let's compute the Lorentz-invariant quantity $p^2 = p \cdot p$,

$$p^2 = \frac{E^2}{c^2} - p_x^2 - p_y^2 - p_z^2. \tag{18}$$

If p^μ represents the momentum 4-vector of a massive particle, then it is simplest to study the particle in the frame where it is at rest, $\vec{p} = 0$. Because p^2 is a Lorentz invariant, we can compute it in the frame where the particle is at rest (called the particle's **rest frame**) but it has the same value in any frame! In the rest frame, we have

$$p^2 = \frac{E_{\text{rest}}^2}{c^2} - |\vec{p}_{\text{rest}}|^2 = \frac{E_{\text{rest}}^2}{c^2}, \tag{19}$$

and so it must also be true in every frame. Thus, the Lorentz invariant quantity p^2 is related to the energy in the particle's rest frame. In relativity, the energy of a particle at rest is related to its mass, M, via $E_{\text{rest}} = Mc^2$, and so we have

$$p^2 = \frac{E_{\text{rest}}^2}{c^2} = M^2c^2. \tag{20}$$

In some different frame with $|\vec{p}| \neq 0$, we still have that $p^2 = M^2c^2$, but the Lorentz invariant computed in this frame is

$$p^2 = M^2c^2 = \frac{E^2}{c^2} - |\vec{p}|^2c^2. \tag{21}$$

Rearranging we have:

$$E^2 = |\vec{p}|^2 c^2 + M^2 c^4. \tag{22}$$

This is one of the most important equations in relativistic kinematics. It tells us that, in any frame, the particle's energy and magnitude of momentum are related to one another by Eq. (22): if we know one of these quantities, we can solve for the other provided we know the mass. Conversely, measuring the energy and momentum of a particle is sufficient to determine the mass.

Note that this formula also works if $M = 0$, in which case $p^2 = 0$ and $E = |\vec{p}|c$. This shows that massless particles travelling at the speed of light, such as photons, can carry momentum and energy. If $M = 0$, there is no frame at which the particle is at rest: it always moves at speed c; we will show this shortly.

For the $M \neq 0$ case, starting from the rest frame with $p^\mu = (Mc, \vec{0})$ and applying Eq. (12), we get in a different frame where the particle is travelling at speed V

$$E = \gamma Mc^2, \tag{23}$$
$$|\vec{p}| = \gamma MV. \tag{24}$$

Notice that the 3-momentum is no longer simply MV as Newtonian mechanics would tell us, but has an additional factor of $\gamma = 1/\sqrt{1 - V^2/c^2}$ in front. This means that even though a particle's speed is capped at c, its momentum can grow without bound. If we take the ratio of our expressions for $|\vec{p}|$ and E, we get

$$\frac{|\vec{p}|c}{E} = \frac{V}{c}. \tag{25}$$

Thus, if $|\vec{p}|c = E$, as is the case for a massless particle, then $V = c$. In other words, massless particles all travel at the speed of light, c!

To check whether these results are sensible, we can take the non-relativistic limit $V/c \ll 1$ and see whether the result matches our expectations from Newtonian mechanics. In this limit, $\gamma \to 1$, and the momentum acquires the usual Newtonian form at leading order in V. We perform a Taylor expansion in V/c on the expression for energy to get

$$E(V \ll c) \approx Mc^2 \left(1 + \frac{V^2}{2c^2}\right) = Mc^2 + \frac{1}{2}MV^2. \tag{26}$$

Thus, the energy also returns to the usual kinetic energy form, plus the contribution from the rest mass energy. The fact that we reproduce the usual results from non-relativistic mechanics means that we can indeed identify the 0^{th} component of the momentum 4-vector p^μ with our conventional notion of energy.

2.3 Systems With Multiple Particles

When we look at particle scattering, we can imagine that we take some initial particles A, B, ... with momenta p_A^μ, p_B^μ, ... and have them interact in some region. Just like the gold foil experiment, the interactions happen so quickly and in such a small space that we cannot directly observe the particles while the interaction is occurring; thus, we treat this region as an opaque box (see Fig. 5). On the other side of the box, another set of particles A', B', ... emerge with momenta $p_{A'}^\mu$, $p_{B'}^\mu$, ... We try to categorize what happens inside the box (*i.e.,* what interactions are happening) by analyzing the incoming and outgoing momenta. Our lives are sometimes made more difficult by the fact that we

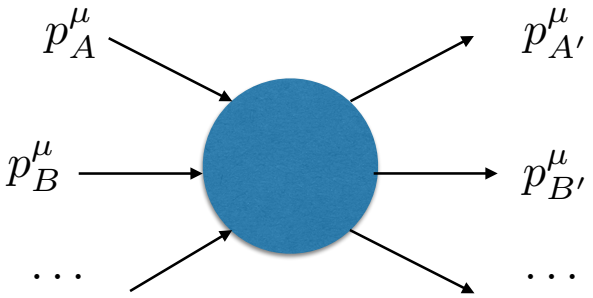

Figure 5: Schematic representation of the study of particle scattering. The process inside the blob is not directly seen and must be inferred by the changes in the particle types and momenta.

do not know all momenta perfectly (in the case of the gold foil experiment, for example, we know that the gold nuclei are initially at rest, but after the collision we only measure the α particles and not the nuclei).

One property that we can use is the fact that all known interactions (and all realistic conjectured interactions) obey conservation of energy and conservation of momentum[4]. These can be combined into **conservation of energy-momentum**, namely that for each of the 4-momentum components ranging from $\mu = 0$ to 3, we have

$$p_A^\mu + p_B^\mu + \ldots = p_{A'}^\mu + p_{B'}^\mu + \ldots. \tag{27}$$

When we detect the incoming/outgoing particles, they tend to be very far apart and are no longer interacting; thus, their only energy is in kinetic energy and mass energy. Kinetic energy is conserved if the incoming particles are the same as the outgoing particles, as the mass contribution to the energy is the same in this case; these are called **elastic collisions**. If particle species change (and consequently the masses), then kinetic energy is not conserved and we call these **inelastic collisions**.

The sum of 4-vectors is still a 4-vector. We can therefore consider an observer in a frame where the net 3-momentum of all objects taken together is zero. This frame of reference is called the **centre-of-mass (CM) frame**. In the CM frame, the total initial 4-momentum has the form

$$p_{\text{initial}}^\mu = (E_{\text{initial}}, \vec{0}). \tag{28}$$

By conservation of momentum, the final 4-momentum also has this form, $p_{\text{final}}^\mu = (E_{\text{final}}, 0)$. This frame is convenient because the sum of the 3-momenta vanishes separately in the initial and final states. The CM frame is usually the simplest reference frame for tackling calculations: when in doubt, try the CM frame!

Using conservation of energy-momentum, we can already see some techniques that are useful in particle physics. Suppose I have conjectured the existence of a new particle X with mass M_X which decays into particles A and B. Conservation of energy-momentum tells us that

$$p_X^\mu = p_A^\mu + p_B^\mu. \tag{29}$$

[4]For those familiar with theoretical mechanics, conservation of energy-momentum follows from the fact that interactions do not depend explicitly on position and time. Momentum is not conserved when there exists an external potential, $V(\vec{x})$; however, because we are dealing with fundamental particles and forces, the only interactions that exist are pairwise interactions that conserve momentum.

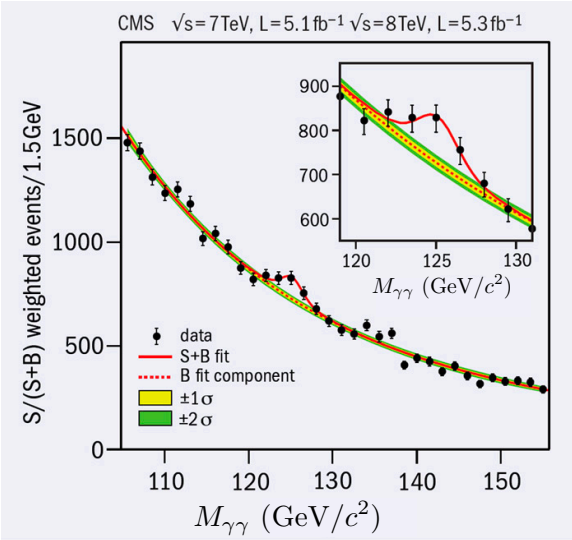

Figure 6: Search for the Higgs boson decaying into two photons by the CMS Collaboration at the Large Hadron Collider. Two photons are selected from each collision, and their 4-momenta are combined to form a quantity called the invariant mass, $M_{\gamma\gamma} = \sqrt{(p_{\gamma_1} + p_{\gamma_2})^2}/c$. If the photons originate from a Higgs boson decay, then conservation of 4-momentum tells us that $(p_{\gamma_1} + p_{\gamma_2})^2 = M_h^2 c^2$; indeed, we observe a bump at the Higgs mass of $M_h \approx 125$ GeV/c^2 corresponding to events from Higgs decays. The other diphoton events originate from photons emitted in other processes. The plot shown here was one of the main pieces of evidence that led to the announcement of the discovery of the Higgs boson (image source: CERN Courier).

If this is true, then the same is true for the square of the 4-momenta:

$$(p_A + p_B)^2 = p_X^2 = M_X^2 c^2, \tag{30}$$

where we have used $p_X^2 = M_X^2 c^2$ as derived for a single particle in the last section.

However, we have just computed a Lorentz-invariant quantity which is true in any reference frame. Thus, we can measure the A and B momenta in any reference frame, calculate this Lorentz-invariant quantity, and if particles A and B originate from the decay of a particle X then we should always find that $(p_A + p_B)^2 = M_X^2 c^2$. This method can also work for decays to 3 or more particles, and this is the most common method for identifying parent particles from the momenta of the child particles. As this example shows, it is often useful to square the expressions for conservation of energy-momentum to form Lorentz-invariant quantities.

Let us take a look at some examples. The Higgs boson was the last particle in the Standard Model to be discovered. It was conjectured to exist in the 1960s, but took over 50 years to find. The best way of discovering the Higgs boson had been hypothesized to be through its decays into photons, $h \to \gamma\gamma$. Particle collider experiments could therefore look for collisions in which a pair of photons emerged from the collision point; after measuring their 4-vectors, $p_{\gamma_1}^\mu$ and $p_{\gamma_2}^\mu$, it is possible to determine the 4-momentum of their sum, $p_{\text{sum}}^\mu = p_{\gamma_1}^\mu + p_{\gamma_2}^\mu$. If the two photons originated from the decay of a Higgs boson, then $p_{\text{sum}}^2/c^2 = (p_{\gamma_1} + p_{\gamma_2})^2/c^2 = M_h^2$. Therefore, by identifying a bunch of photon pairs that all reconstruct the same mass, it is possible to discovery the Higgs boson as well as determine its mass.

We may define a quantity called the **invariant mass**, which is just another word for the Lorentz invariant of a sum of 4-momenta, p_{sum}^μ. For example, the invariant mass of a pair of photons is defined

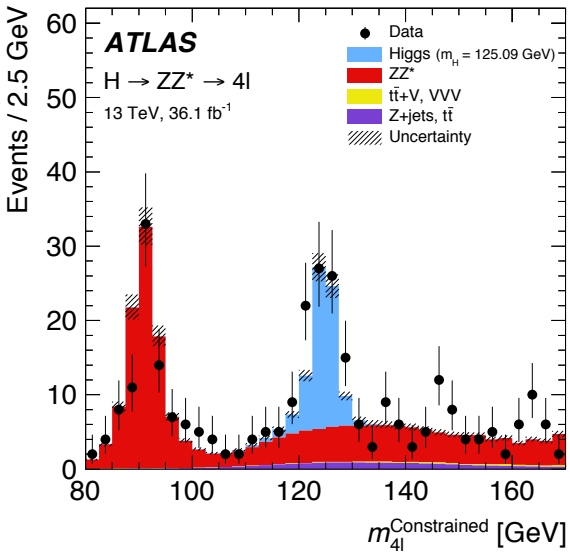

Figure 7: Search for the Higgs boson decaying into four leptons by the ATLAS Collaboration at the Large Hadron Collider. There are two bumps visible in the invariant mass of the 4 leptons. The one at 90 GeV/c^2 corresponds to $Z \to \ell^+\ell^-\ell^+\ell^-$ decays, while the one at 125 GeV/c^2 corresponds to $h \to \ell^+\ell^-\ell^+\ell^-$ decays. Source: `arXiv:1712.02304`.

as $M_{\gamma\gamma} = \sqrt{p^2_{\text{sum}}}/c$; this quantity $M_{\gamma\gamma}$ will be equal to the Higgs boson mass if the photons originate from Higgs decays. A plot of $M_{\gamma\gamma}$ as measured in proton-proton collisions at the LHC is shown in Fig. 6; while most of these photons are emitted in random electromagnetic process and consequently give random values of $M_{\gamma\gamma}$, we observe a large bump at $M_{\gamma\gamma} = 125$ GeV/c^2 which corresponds to photons originating from Higgs decays.

We can also discover traces of a heavy particle decaying into three or more particles. For example, one of the cleanest ways of discovering the Higgs boson is via its decays into 4 charged leptons. These can be $h \to e^+e^-e^+e^-$, $h \to e^+e^-\mu^+\mu^-$, or $h \to \mu^+\mu^-\mu^+\mu^-$ (here, μ^- is a muon and μ^+ is an antimuon). In this case, momentum conservation tells us that we should sum the momenta of each of the *four* leptons ℓ and then compute the Lorentz invariant:

$$(p^\mu_{\ell_1^+} + p^\mu_{\ell_1^-} + p^\mu_{\ell_2^+} + p^\mu_{\ell_2^-})^2 \equiv M^2_{4\ell}/c^2. \tag{31}$$

If the four leptons originate from the decay of a Higgs boson, we should see a bump at $M_{4\ell} = M_h$. In Fig. 7, such a bump is again evident for the Higgs boson at a mass of 125 GeV/c^2, consistent with the Higgs boson mass observed in diphoton decays in Fig. 6. This fact strongly suggests that we are seeing different decays of the same Higgs particle.

To close this sub-section, let us examine a result from the LHCb Collaboration at the Large Hadron Collider. With new techniques, they were able to record an unprecedented dataset of collisions that give rise to a muon, μ^-, and an antimuon, μ^+. They were then able to measure the muon and antimuon momenta and construct the invariant mass of the muon-antimuon pair, $M_{\mu\mu}$. A large number of decays of different particles $X \to \mu^+\mu^-$ with different masses can be identified by the bumps in the invariant mass plot: these particles include, from left to right, the η meson, the ρ and ω mesons, the ϕ meson, the J/ψ and $\psi(2S)$ meson, and the Υ mesons. The final bump on the far right is the Z boson.

As a useful piece of jargon, we often refer to the **spectrum** of the invariant mass, which simply

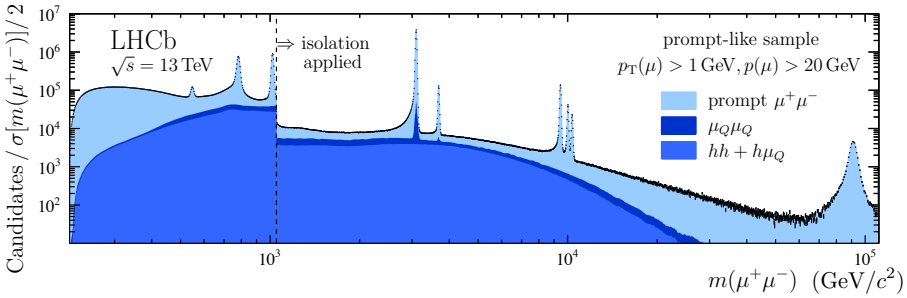

Figure 8: Plot in light blue of the invariant mass of a μ^+ and μ^- pair in proton collisions made by the LHCb Collaboration at the Large Hadron Collider. Each of the bumps corresponds to a particle decaying into a muon and an antimuon, including the Z boson (far right), Υ mesons (clustered around 10^4 MeV/c^2, and the J/ψ meson (huge peak around 3000 MeV/c^2). Source: `arXiv:1710.02867`

means a plot of how many different collision events are observed at each invariant mass. Note that the bumps are not infinitely narrow; the reason is that the detector can only measure the momentum of each muon to a certain precision, and the inaccuracy in reconstructing the muon momenta smears out the bump. There is another reason for bumps to have some finite width; we will return to this topic in Section 3.4.

Exercise 3: Show that a massless particle X can never decay into two particles A and B, even if $E_X/c^2 > M_A + M_B$ in a particular frame.

Exercise 4: Two massless A particles have the same magnitude of momentum $|\vec{p}|$. One travels in the $+\hat{z}$ direction, and the other travels in the direction $\cos\theta\hat{z} + \sin\theta\hat{x}$. What is the minimum angle, $\theta_{\min}$, at which they must collide in order to combine into a new particle B with mass M_B? Assume $|\vec{p}| \gg M_B c$. In terms of Lorentz invariants, explain why B cannot be produced for $\theta < \theta_{\min}$.

Exercise 5: (Griffiths 3.14) Particle A with energy E hits particle B at rest, producing $A + B \to C_1 + \ldots + C_n$. Calculate the minimum E for this reaction in terms of the various particle masses.

2.4 Natural Units

If you've gotten this far and haven't found any factors of c missing, then I should count myself very lucky. The reality is that many of these factors of c are annoying to carry around: for instance, we have to remember to divide E by c in the momentum 4-vector, and to multiply t by c in the position 4-vector. It would be simpler if we didn't have to deal with them at all.

For this to work, we need to define a set of units where $c = 1$. It is perfectly reasonable to define such a set of units, since c is the only fundamental constant that relates things like position and time, or energy, momentum, and mass. If I say "The proton has a mass of 1.6×10^{-10} J", there is no ambiguity as to my meaning, which more precisely is "The proton has a mass whose rest energy is equal to 1.6×10^{-10} J". There is no ambiguity because mass and rest energy map onto one another in a simple, one-to-one manner.

We call this set of units **natural units**. In natural units, we fix

$$c = 1, \tag{32}$$
$$\hbar = 1, \tag{33}$$

where $\hbar$ is the reduced Planck's constant, $\hbar = h/2\pi$. Recall that Planck's constant relates energy and time,

$$E = h\nu, \tag{34}$$

where ν is the frequency of light (in Hz = 1/s), and so using combinations of $\hbar$ and c it is possible to relate pretty much all physical quantities to a single unit.

By convention, we choose the physical unit to be a unit of energy, and in particle physics we use the energy unit of electron-Volts (1 eV $= 1.6 \times 10^{-19}$ J). Using Eq. (34) and setting $\hbar = 1$, we see that time has units of inverse energy. Similarly, since distance is related to time via $x^0 = ct$, we see that distance also has a unit of inverse energy. Physical dimensions of length, time, etc. can be found by remembering that $\hbar = c = 1$ in natural units, so you can multiply or divide by $\hbar$ and c wherever you like without changing the answer. Then, substitute in the physical values $\hbar = 6.58 \times 10^{-25}$ GeV $\cdot$ s and $c = 3 \times 10^8$ m to convert from natural units back into more familiar units.

Exercise 6: (a) A particle physicist tells you that the typical lifetime of a particle in its rest frame is $\tau = 10^{14}$ GeV^{-1}. *What is this time in seconds? (b) In the lab frame, this same particle has a speed* v. *What is the typical decay length in the lab frame?*

Exercise 7: (a) In natural units, we say "The wavelength of the photon is 1 MeV^{-1}*". What do we mean by this (in other words, translate the energy to a property of the wavelength). (b) In natural units, we say "The momentum of particle A is* 10 GeV*". What do we mean by this (in other words, translate the energy to a property of the momentum)?*

3 Rates in Particle Collisions

3.1 Introduction

The purpose of a particle collider is to smash particles together, observe the rate at which various processes happen, and compare them with theoretical predictions. For example, the Large Hadron Collider (LHC) is currently the world's most powerful collider experiment, at it collides pairs of protons at a CM energy of 13.6 TeV. The LHC experiments measure the rate at which protons collide and produce specific particles, such as quarks, electrons, or neutrinos. Since each of these particles experiences different forces (the quarks interact via the strong force, electrons via the electromagnetic and weak interactions, while the neutrinos interact only via the weak force), observing these rates allow us to test the predictions of the Standard Model (SM) forces. The SM not only predicts the overall rates of production of these particle types, but also allows us to calculate the typical momenta associated with each collision (this is often referred to as the **momentum spectrum** of particles produced in the collision). Through observations of the ingoing and outgoing momenta at the LHC, it is possible to compare experimental data with precise theoretical predictions: any deviation from the predicted momentum spectrum could herald the existence of a new particle or force of nature!

In order to make these comparisons, theorists and experimentalists need to agree on quantities to compute and measure. These folks often have different priorities, though: for an experimentalist, it is simplest to set up their apparatus in the most convenient fashion and count the number of observed events. The results are precise, but also specific to that particular experiment, making it difficult to compare the results of different experiments. Meanwhile, theorists would rather live in a dream world and calculate the probability of producing various particles under ideal conditions. The goal is to come up with a common benchmark that is convenient for theorists (having properties such as Lorentz invariance and being generally applicable to many different experiment types) and experimentalists (easy to convert their data into something that theorists and other experimentalists care about).

The most common way of expressing the rates of scattering processes is via the **scattering cross section**, σ. The term is related to "cross-sectional area", and the idea comes from the intuitive notion that something with a large cross sectional area is easy to hit if I throw something at it. A large cross section therefore corresponds to a large rate of scattering, while a small cross section corresponds to scattering being a rare process. The scattering cross section is a quantity that depends on the nature and momenta of the particles are taking part in the scattering, but is otherwise independent of the properties of the beam, such as how many particles are in the beam or how tightly they are compressed. This makes the cross section a meaningful quantity even if a given experiment colliders particles under many different conditions.

The cross section has units of area, which has dimensions of E^{-2} in natural units. We can translate a cross section into a number of scattering events if we are given the number of particles impinging on the collision target per unit area.

In the following sections, we make more precise what we mean by "cross section". We first use a simple example of elastic scattering in classical physics to motivate a particular definition of the cross section and see that it agrees with the intuitive notion introduced here. We then consider cross sections in a more general context as they are applied to particle physics processes. Finally, we discuss decay rates of unstable particles, which are also important processes that occur in particle collisions.

Some of this information, especially in Sec. 3.2, could be skipped on a first reading. What is most important is understanding how the number of scattering events when colliding particles can be related to a cross section. Similarly, some parts of Sec. 3.2 are challenging if you have not seen

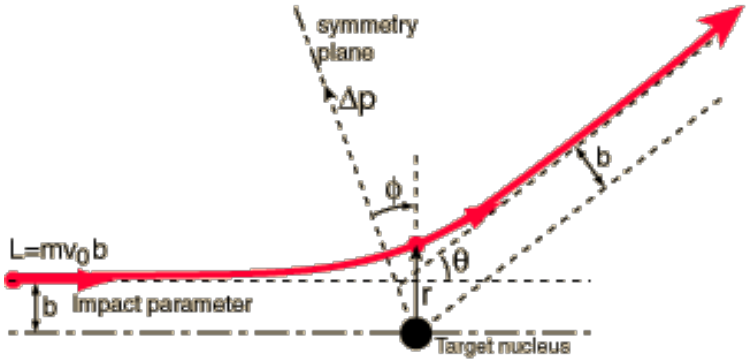

Figure 9: Set-up of a scattering problem. An incident particle of mass m is scattered off a target located at the origin. The target is assumed to be sufficiently heavy that it is stationary, and the resulting interaction is described by a potential energy $U(r)$ due to the force between the incoming particle and the target. The impact parameter, b, is the displacement of the incident particle relative to the origin, and the angle θ is the angle by which it is deflected in the scattering process. (Source: Hyperphysics)

multivariable calculus before, so this is something you can return to later.

3.2 Cross Sections in Classical Elastic Scattering

Single-Particle Scattering

The set-up of a general scattering problem in classical physics is shown in Fig. 9. A particle of energy E travel in the z-direction and impinges on a **scattering target**. The particle is usually not directed exactly at the centre of the target, but is instead offset by a perpendicular distance b known as the **impact parameter**. When a force exists between the incident particle and the scattering target (whether attractive or repulsive), the incident particle is deflected from its original trajectory with energy E'. In the simplest case, the scattering target is assumed to very heavy and is un-moved by the collision (even if not, it is possible to reduce a general 2-body scattering problem to an effective 1-body motion in a fixed potential, so this is an ok assumption). The scattering target can be represented in this situation by a time-independent potential $V(\vec{r})$ that acts on the incident particle as it moves through space. In the case of the Rutherford experiment where alpha particles are scattered off of gold nuclei, the scattering target is well represented by the Coulomb potential $V(\vec{r}) \sim Qe^2/r$.

To begin with, let us see what conservation laws have to say about the scattering process. Certainly, energy conservation should apply. Well before and after the scattering, the incident particle and scattering target are well separated, and so the potential associated with the force should be negligible ($V_{\text{initial}} = V_{\text{final}} = 0$)[5]. Since the potential does not change during the collision, conservation of energy requires the energy of the ingoing and outgoing particle to be the same ($E = E'$). In particular, if the particle does not change in mass, then the initial and final kinetic energies must be equal, and this is the scenario of **elastic scattering**. Apart from being a common occurrence

[5]If this were not true, then the incident particle would actually be bound to the scattering centre, and the scattering description would not be appropriate. For example, comets that are bound to the Sun should not be treated as "scattering" off the Sun, but rather as objects in orbit.

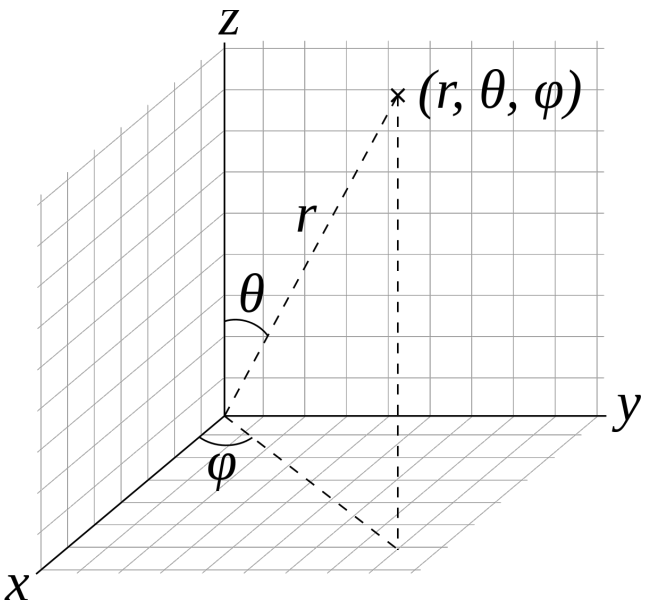

Figure 10: Spherical polar coordinates. The Cartesian (x, y, z) coordinate position of an object is described in terms of the distance from the origin, r, the polar angle, θ, and the azimuthal angle, ϕ. (Source: Wikimedia commons)

in physics, elastic scattering is relatively simple to analyze, because the collision only affects the directions (and not the energies) of the scattered particles.

Solving the scattering problem therefore reduces to the question of determining the scattered particle's direction given an initial energy, E, and impact parameter, b. The scattered particle's final direction is most easily described using a coordinate system with origin chosen to be located at the target (also known as the scattering centre). Then, the direction is described by two angles: a polar angle, θ, which describes the angle relative to the z-axis, and an azimuthal angle, ϕ, which describes the angle in the $x - y$ plane[6]. You will hopefully recognize these as the angles of a spherical polar coordinate system (see Fig. 10):

$$x = r \sin\theta \cos\phi, \tag{35}$$

$$y = r \sin\theta \sin\phi, \tag{36}$$

$$z = r \cos\theta. \tag{37}$$

Solving a single-particle scattering problem is therefore equivalent to finding the final θ and ϕ as a function of b and E. For systems with cylindrical symmetry (such as scattering of a particle incident in the $\hat{z}$ direction off a spherically symmetric potential), the scattering is uniform in ϕ and so the problem reduces to finding θ.

Let us consider a simple example: hard-sphere scattering. We assume that the scattered particle is incident on an "infinitely hard" sphere of radius R, meaning that the potential is

$$V(\vec{r}) = \infty, \quad r < R, \tag{38}$$

$$= 0, \quad r \geq R. \tag{39}$$

In words, the scattered particle is perfectly repelled when it gets closer than a distance of R to the sphere's centre. If the sphere is too heavy to be moved, conservation of energy requires that the

[6]Note that this is the opposite of the math convention, where ϕ is the polar angle and θ is the azimuthal angle.

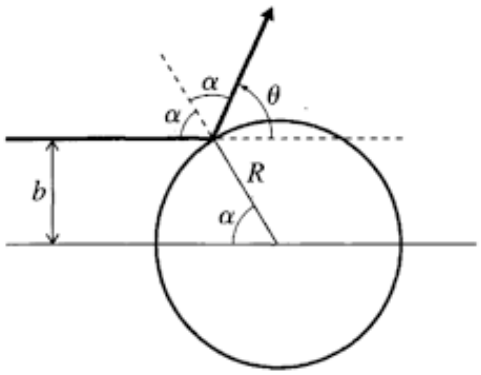

Figure 11: Diagram of scattering off a hard sphere of radius R. The angle of incidence is equal to the angle of reflection, and it is evident that $2\alpha + \theta = \pi$. (Source: J. Taylor, *Classical Mechanics*)

outgoing particle have the same magnitude momentum as the incoming particle. However, we expect that the potential should not be able to affect the momentum *parallel* to the surface of the sphere, because forces only act along potential gradients, *i.e.,* $\vec{F} = \vec{\nabla} V(r)$. Therefore, collision with the sphere will reverse the component of the momentum perpendicular to the sphere's surface, and the collision will leave intact the components of the momentum parallel to the surface (see Fig. 11).

According to the figure, we have for $b < R$,

$$\sin\alpha = \frac{b}{R}, \tag{40}$$

$$\pi = \theta + 2\alpha. \tag{41}$$

Thus, we can solve for θ as a function of b,

$$\theta(b) = \pi - 2\sin^{-1}\left(\frac{b}{R}\right) \quad (b < R). \tag{42}$$

For $b > R$, no collision takes place and $\theta(b) = 0$.

Note that this is a one-to-one mapping between θ and b, so we can also invert the relation to obtain

$$b(\theta) = R\sin\left[\frac{1}{2}(\pi - \theta)\right] = R\cos\frac{\theta}{2}. \tag{43}$$

The interpretation of this result is that, to scatter at an angle θ, the particle must be incident at a distance b. This may seem like a strange way of thinking about scattering, but it has a distinct advantage: the domain of the polar angle is always set by the dimensionless angles $0 < \theta < \pi$, while the corresponding value of b is set by the physical radius of the scattering target, R. For large R, the range of b that results in scattering is larger. We can thus use the above relation as a proxy of how "easy" it is to hit the target, with bigger R corresponding to more frequent scattering.

Scattering of Many Particles

We are almost never interested in the scattering of a single particle. One reason is that the information obtained from watching the trajectory of a single particle is very limited since we only see the relationship between b and θ for a single value of b. To determine the force law, we need to study

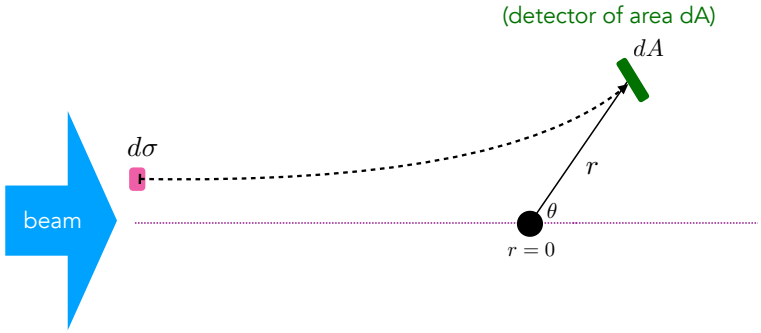

Figure 12: Set-up for scattering of a beam of particles. The incident flux, J_{incident} is the number of particles per unit area per unit time in the beam. The number of particles passing through the area $d\sigma$ in time dt is therefore $N_{\text{incident}} = J_{\text{incident}}\, d\sigma\, dt$. The particles are scattered by the target and end up hitting a detector of area $dA = r^2 \sin\theta d\theta d\phi$ located a distance r from the origin.

the trajectories of *many* particles with different values of b and θ. If we wish to observe very rare scattering processes, we need to scatter *immense* numbers of particles. At the LHC, the bunches of colliding protons each contain $\sim 10^{11}$ protons, while other experiments will collide up to 10^{15} electrons or protons on a target at once. Clearly, we need to extend our formalism to account for scattering of an entire beam of particles, and we wish to do so in a manner that allows us to easily characterize the scattering of particles for different beam configurations.

The problem is that the number of scattered particles depends on the details of the beam, and the beam is going to be *different* for every experiment! This is very inconvenient, because it means that theorists need to do a different calculation for every experiment to predict & interpret the outcome of every experiment. Therefore, we seek a beam-independent way of characterizing the physics. To do so, we need to define some properties of the beam and find a way of dividing out the intensity of the beam in our calculation.

The beam of particles is characterized by a **flux**, or number of particles impinging on a given area per unit time. We use J_{incident} to denote the incident particle flux, and it will be a function of the impact parameter that the beam passes through. The number of particles passing through an area $d\sigma$ in time dt is then

$$N_{\text{incident}}(d\sigma) = J_{\text{incident}}\, d\sigma\, dt. \tag{44}$$

For an elastic collision, these *same* particles are then scattered through a range of angles $d\theta$ and $d\phi$. We choose to position a detector to capture exactly the particles that passed through $d\sigma$ so that $N_{\text{scattered}} = N_{\text{incident}}$. This detector is positioned at a radius r from the scattering target. If we draw a sphere of radius r centred at the origin, the detector area that the particles pass through is $dA = r^2 d\Omega = r^2 \sin\theta\, d\theta\, d\phi$, where $d\Omega$ is the *solid angle* which is interpreted as the area subtended by the angles $d\theta$ and $d\phi$ on the unit sphere. The scattered flux is

$$J_{\text{scattered}}(d\Omega) = \frac{N_{\text{scattered}}}{dA\, dt} \tag{45}$$

$$= \frac{N_{\text{incident}}}{r^2 d\Omega\, dt} \tag{46}$$

$$= \frac{J_{\text{incident}}\, d\sigma}{r^2 d\Omega}, \tag{47}$$

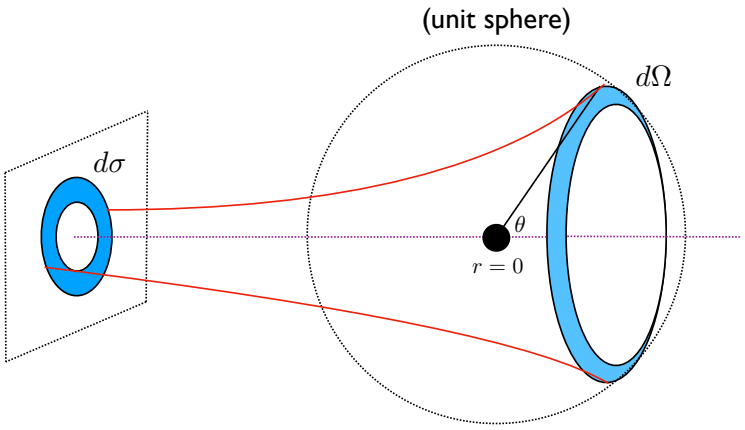

Figure 13: Illustration of the differential cross section. A collection of particles passing through the area $d\sigma$ and impact parameter b are scattered in the potential. They end up occupying an area on the unit sphere of magnitude $d\Omega = 2\pi \sin\theta d\theta$ where we have used the symmetry of the problem to integrate over all values of ϕ. While the number of particles passing through $d\sigma$ and $d\Omega$ depends on the beam, the ratio $d\sigma/d\Omega$ does not!

where we have used the fact that the detector has been designed so that it captures all of the scattered particles. The $1/r^2$ dependence merely comes from the fact that, if we look at surfaces very far from the scattering centre, the scattered beam is diluted because the same number of particles is passing through spheres of ever larger surface area. The overall process is illustrated in Fig. 13; in this figure, we have used the cylindrical symmetry of the system to integrate over $\int_0^{2\pi} d\phi = 2\pi$ in the azimuthal direction; this illustrates the motion of all the particles in the beam that have the same impact parameter b and pass through the area $d\sigma$.

We see that the ratio:

$$\frac{d\sigma}{d\Omega} = \frac{J_{\text{scattered}}\, r^2}{J_{\text{incident}}} \tag{48}$$

has the property that the overall intensity of the beam (as determined by the incident and scattered fluxes) cancels out. Everything on the right-hand side can be measured experimentally, but the ratio is independent of either the beam or the detector geometry. Therefore, theorists can calculate the quantity $d\sigma/d\Omega$ and then compare to the experimentally measured value.

Given an incident flux of particles and a detector positioned at a radius r, we have reduced the scattering problem to finding the quantity $d\sigma/d\Omega$. This is known as the **differential cross section**, and it tells us over how large of a beam area the incident particles must fall to be scattered into the solid angle $d\Omega$. The larger the differential cross section, the more particles get scattered in the given direction for a fixed beam intensity. It *does not depend* on the beam flux or its distribution in space, and depends only on the interaction that causes the incident particles to scatter into a particular angle. It is, admittedly, a funny and non-intuitive way of characterizing scattering, but it allows us to characterize scattering in a way that is independent of the beam. The information it carries is analogous to $b(\theta)$ for single-particle scattering.

In fact, we can now show that $d\sigma/d\Omega$ has the same information as $b(\theta)$. If we take σ to be the area of the beam from the origin out to a distance b, then the relation between them is $\sigma = \pi b^2$. The differential area of the ring $d\sigma$ corresponding to a change in impact parameter db is therefore

$$d\sigma = 2\pi b\, db \tag{49}$$

where we have once again integrated over ϕ. Similarly, $d\Omega = 2\pi \sin\theta \, d\theta$ after integrating over ϕ. We can take the ratio of the two differentials to get

$$\frac{d\sigma}{d\Omega} = \left| \frac{2\pi b \, db}{2\pi \sin\theta \, d\theta} \right| \tag{50}$$

$$= \frac{b}{\sin\theta} \left| \frac{db}{d\theta} \right|. \tag{51}$$

Note that we have had to introduce the absolute value signs because the derivative $db/d\theta$ is sometimes negative, whereas the ratio of areas is always positive.

Let's apply this to the specific case of hard-sphere scattering. Using Eq. (43), we have

$$\frac{d\sigma}{d\Omega} = \frac{b}{\sin\theta} \frac{R}{2} \left| \sin\frac{\theta}{2} \right| \tag{52}$$

$$= \frac{R^2}{4}. \tag{53}$$

This is a generalization of Eq. (43) by finding the *areas* of the incident beam that contribute to scattering at a given angle θ. Since $d\sigma/d\Omega$ is independent of θ, we find that each outgoing scattered angle receives contributions from equal areas of the beam.

From the differential cross section, we can derive a related quantity which is known as the **total cross section**,

$$\sigma = \int d\Omega \, \frac{d\sigma}{d\Omega}. \tag{54}$$

By integrating over all angles, it gives a characterization of what area of the original beam is scattered at all. In the case of the hard sphere, this is

$$\sigma = \frac{R^2}{4} \int_0^\pi d\theta \, \sin\theta \int_0^{2\pi} d\phi = \pi R^2. \tag{55}$$

The total cross section is equal to the cross-sectional area of the sphere when projected onto the $x-y$ plane! Therefore, we have a physical understanding of the total cross section in this case: if you fire a beam within an area σ of the target, it results in a collision with the hard sphere and the incident particles are scattered; however, incident particles outside of this area do not scatter.

For other scattering processes, the interpretation is less direct but the idea is the same: the total cross section characterizes the total area of the beam that gets scattered by the target. The larger the cross section, the more of the beam that gets scattered and so more scattering occurs (provided the beam is spread out over distances larger than R). By contrast, the differential cross section tells you how often scattering occurs in a given direction. For any given potential energy $U(r)$, it is possible to compute both of these quantities, and measuring them in the lab provides a test of our theoretical understanding of the underlying force.

3.3 Cross Sections in Particle Physics

Total Cross Section

The discussion in the last section was true for elastic scattering in classical mechanics. However, given any beam with flux J_{incident} delivered over a short time dt, the cross section is defined as

$$\sigma(AB \to CD\ldots) = \frac{N(AB \to CD\ldots)}{J_{\text{incident}} \, dt}. \tag{56}$$

The flux is often referred to as the instantaneous luminosity. Usually, however, we measure a number of collisions over many weeks, months, or years. In this case, we must integrate over time to get N:

$$N(AB \to CD\ldots) \;=\; \sigma(AB \to CD\ldots) \int J_{\text{incident}}\, dt. \tag{57}$$

We define the **integrated luminosity**, $\mathcal{L}$, as

$$\mathcal{L} = \int dt\, J_{\text{incident}}(t). \tag{58}$$

This gives rise to a very simple relationship between number of events, cross section, and integrated luminosity:

$$N(AB \to CD\ldots) \;=\; \sigma(AB \to CD\ldots)\, \mathcal{L}. \tag{59}$$

As a result, it is the integrated luminosity, $\mathcal{L}$, that is usually reported by experiments.

From the last section, we know that σ has units of area. Since the number of collision events is dimensionless, this means that $\mathcal{L}$ has units of inverse area. The standard unit of area in particle physics is the **barn**, and is denoted by the letter b. A barn is defined as

$$1 \text{ b} = 10^{-24} \text{ cm}^2 = 100 \text{ fm}^2. \tag{60}$$

It seems like a very arbitrary unit to use. However, its origins are in the scattering of nuclei. The characteristic size of a nucleus is ~ 10 fm, and so the typical cross sectional area is $\sim$ b. We see that the scattering cross section of nuclei has a *geometric* interpretation: if two nuclei happen to overlap in space, they typically scatter. The reason for this is that the strong force has a typical range of ~ 10 fm, which explains why nuclei tend to be this size. If nuclei cross at distances closer to this, the residual strong force gives rise to inter-nuclear interactions[7].

From Eq. (59) and the fact that the number of events has no units, we can deduce that the units of luminosity are **inverse barns**, b^{-1}. In particle physics, we tend to deal with interactions that are much weaker than the strong force, and so typical units of measure are nanobarns, picobarns, and femtobarns. Currently, the LHC has accumulated about 150 inverse femtobarns (fb^{-1}) of integrated luminosity in collisions, with plans to collect up to 3000 fb^{-1} by the end of its 20-year running.

The cross section is a quantity that can be computed from the theory of a given interaction to characterize how often initial-state particles $A+B$ will scatter into final-state particles $C+D+\ldots$ in a beam-independent way. The cross section is a Lorentz-invariant quantity, which we expect to be true from the fact that the number of collisions should not depend on the observer. Experiments collide particles A and B with a known flux and count the number of times the reaction $A+B \to C+D+\ldots$ occurs. This allows for a comparison between experiment and theory. Each process has its own cross section. For example, at a centre-of-mass collision energy of 14 TeV, the theoretically predicted cross section for the production of a top quark and antiquark pair in proton collisions is $\sigma(pp \to t\bar{t}) = 953.6$ pb, while the corresponding cross section for the production of a Z boson is $\sigma(pp \to Z) = 3.5$ nb.

[7]The name "barn" is also rather strange. It has its origins in the saying "They couldn't hit the broad side of a barn door", *i.e.*, that person has very bad aim. A barn is therefore interpreted as a huge cross sectional area that is hard to miss. Since the strong interactions are, well, strong, nuclear physicists decided to use this as the name for the unit of cross sections.

Differential Cross Section

The total cross section is an "inclusive" measure of the rate of a process, meaning that we are just adding up the total number of times a process happens without regard to details like what the energies of the final particles are or what angles at which they emerge from the collision. However, we saw in the last section that it is also possible to define a differential cross section in some observable, such as the scattering angle. Measurements of differential cross sections allow for more precise testing of the predictions of scattering mediated by a particular force, although they can be more challenging to calculate than total cross sections.

Because we deal with relativistic systems in particle physics, we have to be careful to define the reference frame when using differential cross sections. For example, we may wish to consider the differential cross section in the scattering angle for some process, and the angular dependence will depend on whether we use the lab frame, centre-of-mass frame, or some other frame. For example, the differential cross section at for $e^+e^- \to \mu^+\mu^-$ scattering in the centre-of-mass frame is (at high energies)

$$\frac{d\sigma}{d\Omega} = \frac{\alpha^2}{4(E_{e^-} + E_{e^+})^2}(1 + \cos^2\theta), \tag{61}$$

and $\alpha \approx 1/137$ is the fine-structure constant that characterizes the strength of the electromagnetic force. We see that scattering is more frequent when $\cos\theta = \pm 1$, *i.e.,* $\theta = 0, \pi$, while the scattering rate is lower when the polar angle is $\sim \pi/2$. The total cross section is found by integrating the above expression over all angles. To find the differential cross section in a different frame, we have to boost to a new coordinate system and express the rate in terms of the new polar angle θ'.

Exercise: Does the denominator of Eq. (61) change in a frame other than the CM frame? Why or why not? Calculate the total cross section for $\sigma(e^+e^- \to \mu^+\mu^-)$.

Often, differential cross sections are computed relative to variables other than the scattering angle. Both theorists and experimentalists like to compute cross sections that are differential in *Lorentz-invariant quantities* to make the comparison between theory and experiment easier. For instance, consider the process $pp \to e^+e^-\mu^+\mu^-$. There are various Lorentz-invariant quantities that can be computed, such as $p^2_{e^+e^-} \equiv (p_{e^+} + p_{e^-})^2$, $p^2_{\mu^+\mu^-}$, $p^2_{e^+\mu^-}$, etc. One can also calculate differential cross sections in multiple variables; an example of a **doubly differential cross section** is

$$\frac{d^2\sigma}{d(p^2_{e^+e^-})\, d(p^2_{\mu^+\mu^-})} \tag{62}$$

shows us how the cross section is distributed among the different values of final state momentum products. Because both the numerator and denominator are Lorentz invariants, the differential cross section is also the same in every frame.

Cross Sections and Quantum Mechanics

Eq. (59) is so important that it is worth repeating:

$$N(AB \to CD\ldots) \;\; = \;\; \sigma(AB \to CD\ldots)\,\mathcal{L}.$$

In classical mechanics, which is deterministic, the cross section has a very clear meaning: it quantifies how particles in a particular initial state get mapped into the final state after the collision. Any

particle impinging on an area $d\sigma$ will necessarily end up in the final state solid angle $d\Omega$. In quantum mechanics, however, determinism no longer applies, and it is no longer known with certainty whether a particular initial state will end up in a particular final state. Instead, there is a *probability* that a given initial state will end up in any number of final states.

In a quantum world, we can no longer interpret $N(AB \to CD\ldots)$ as *the* number of times $A+B$ will give rise to the final state $C+D+\ldots$. Rather, we must interpret it as the **expected**, or average, number of such collisions for a given cross section and integrated luminosity. The actual number of times a collision will occur this way is randomly drawn from some probability distribution[8] and the result will be different for each time you run the experiment. Statistical analysis is therefore necessary to compare an observed number of events with a predicted cross section.

3.4 Lifetimes, Decay Rates, and Widths

So far, we have talked about the rates at which various scattering processes occur when colliding particles together. However, there are other kinds of physical processes that can occur in particle physics. One that may be familiar to you is the process of radioactive decay, whereby an "unstable" heavy nucleus spontaneously emits an alpha, beta, or gamma ray and ends up in a more stable form. While cross sections describe $2 \to N$ scattering processes (where N is the number of final-state particles), decays are $1 \to N$ processes and require a different treatment.

From earlier studies of radioactivity, you probably have heard of the half-life of an unstable particle, which is the time it takes for half of the unstable particle to decay away. For a half-life of $t_{1/2}$, the fraction that remains after a time t is

$$f(t) = 2^{-t/t_{1/2}}. \tag{63}$$

Doing calculations with exponents of base other than e are a huge pain, so in particle physics we tend to use as the time constant the **proper lifetime**, τ, which is the time it takes for the unstable particle amount to decay to $1/e$ of its original abundance,

$$f(t) = e^{-t/\tau}. \tag{64}$$

When applied to a single particle in quantum mechanics, this gives the probability that a particle will survive to time t. The proper lifetime is a fundamental property of a particle; all Higgs bosons have the same lifetime as one another, and all neutrons have the same lifetime as one another, etc. The above solution $f(t)$ obeys the following differential equation,

$$\frac{df}{dt} = -\frac{1}{\tau} f(t), \tag{65}$$

which simply states that the decay rate depends only on the proper lifetime and the fraction of particles left at that point: the particle, as an elementary particle, does not age and so its relative probability of decay per unit time is the same at every instant of time.

In relativistic physics, time depends on the frame of reference used. In particular, an observer that sees a particle travelling near the speed of light will see that particle's clock slow down (this is the usual time dilation). The proper lifetime gives the characteristic decay time of the particle *in its rest frame*. Therefore, if the particle has a boost $\gamma = E/M$ in the lab frame, then the typical decay time is slowed by this amount,

$$f_{\text{lab}}(t) = e^{-t/(\gamma\tau)}. \tag{66}$$

[8]The probability distribution that correctly models the expected number of observed events is the Poisson distribution, which is derived in Sec. 6.3.

To avoid confusion, we *always* refer to the particle's own rest frame when defining its lifetime since this is a Lorentz-invariant prescription. Only massive particles can undergo $1 \to N$ decays consistent with special relativity, and so there is a unique rest frame in which to define the particle's lifetime.

In natural units, lifetimes have dimensions of 1/energy. The reciprocal quantity to the lifetime is known as the **width**, Γ,

$$\Gamma \equiv \frac{1}{\tau}, \tag{67}$$

and has units of energy. The width is also known as the **decay rate** because it gives the expected number of decays per unit time. Γ is called the width because of a particular quantum mechanical relationship between energy and time. This relationships states namely that particles which exist for only a finite time τ do not have a well-defined energy (and, in the rest frame of the particle, this translates into a spread of possible mass values), and the spread in the particle's energy, ΔE, is (in natural units)

$$\Delta E \sim \frac{1}{\tau}. \tag{68}$$

For example, consider a particle of mass M with lifetime τ. The mass simply represents the particle's energy in its rest frame. However, this energy can only be defined to within $\Delta E \sim 1/\tau = \Gamma$. Therefore, instead of always observing a particle with mass M, the particle actually has a *distribution* of masses within $\sim M \pm \Gamma$.

As an example, consider the Z boson mass as measured at the LEP experiment at CERN (see Fig. 14). The Z mass is measured in (for example) the reaction $e^+e^- \to Z$, followed by the decay to qarks $Z \to q\bar{q}$. The mass of the Z is reconstructed as a bump in the distribution of the centre-of-mass energy $\sqrt{(p_q + p_{\bar{q}})^2}$, since in the Z rest frame this quantity is just its mass energy M_Z. The spread of events in mass is not due to the imprecision of the detector, but rather due to the intrinsic spread of the Z boson mass-energy in its rest frame due to the finite lifetime. This "energy-time uncertainty" is well-known in quantum mechanics and can be rigorously derived in its application to particle lifetimes, but this derivation is beyond the scope of these notes.

The lifetime (or, equivalently, the width Γ) is important because many particles produced in high-energy collisions are unstable and decay on times ranging from 10^{-23} seconds to 1 second. The cross section therefore tells us the rate of production of various particles in the initial collision, while the decay rates tell us how those unstable particles decay throughout the detector. By combining these pieces of information, we can theoretically predict the various stages of particle production and decay in collider experiments.

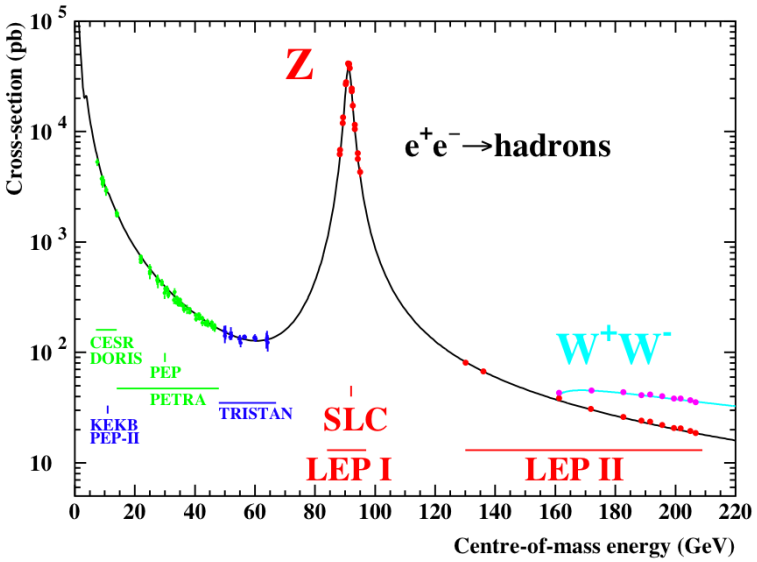

Figure 14: Rate of $e^+e^- \to q\bar{q}$ as a function of the CM-frame total initial beam energy. There is a large bump in the rate associated with $e^+e^- \to Z$, $Z \to q\bar{q}$ for CM energies around 91 GeV. The width of this bump is associated with the fact that the Z boson has a finite lifetime, which leads to an intrinsic uncertainty in the Z rest-frame energy. (Source: `arXiv:hep-ex/0509008`)

4 An Overview of Particle Colliders

4.1 Types of Accelerator

Before colliding particles at high energies, we must first *accelerate* them to high energies. The simplest way to accelerate a particle with electric charge Q, such as an electron or proton, is to apply an electric field $\vec{\mathcal{E}}$, which results in a force $\vec{F} = Q\vec{\mathcal{E}}$. The work done by the electric field over a distance $d\vec{r}$ is $\vec{F} \cdot d\vec{r}$. The change in energy of the particle is therefore

$$\Delta E \;=\; Q \int d\vec{r} \cdot \vec{\mathcal{E}}(\vec{r}). \tag{69}$$

To probe the inside of a proton, centre-of-mass energies $\gtrsim$ GeV are required, while producing the Higgs boson with $M_h = 125$ GeV requires even larger energies. Since the charge Q of a particle is handed to us by nature and is typically an integer multiple of the electron charge, e, we can only get large increases in the particle's kinetic energy by either applying a very strong electric field, or by applying the electric field over a very long distance.

The fact that we use eV-based units of energy is helpful in understanding the challenge of constructing particle accelerators. Accelerating a proton to 125 GeV nominally requires the application of 125 billion volts to a proton! The largest known electric fields are found inside of molecules and are of the order of 100 GV/cm. If this field could be applied to macroscopic systems, then only a very small accelerator would be needed to accelerate protons or electrons to 125 GeV. However, such strong electric fields tend to tear apart common materials. For instance, air experiences dielectric breakdown for fields exceeding 3 MV/m. Achieving very high electric fields over macroscopic distances remains a significant challenge, and so high-energy accelerators need to be *very long* to achieve the necessary beam energies.

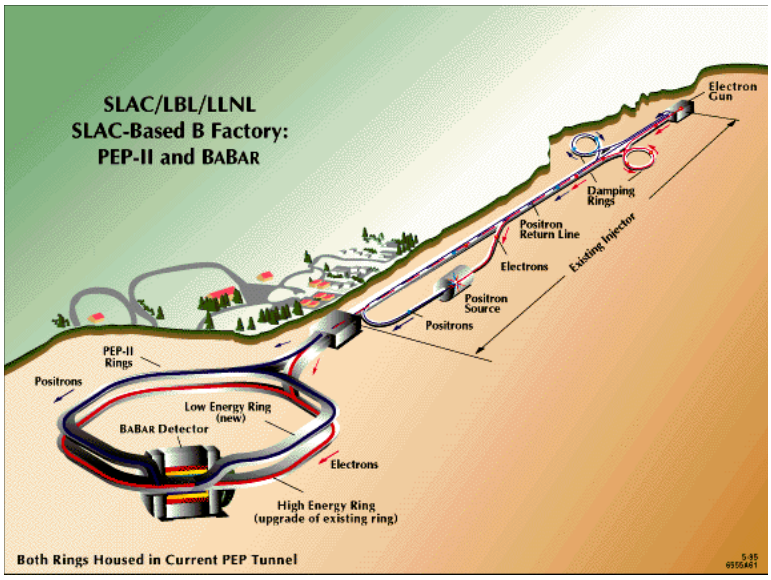

Figure 15: Configuration of the linear accelerator at SLAC near Stanford, CA. Electrons and positrons were accelerated over the two-mile length of the accelerator. They were then guided into a ring and collided into one another at the indicated detection point. The figure illustrates the configuration of the BaBar experiment, which collides electrons and positrons at a centre-of-mass energy of 10.6 GeV. Source: SLAC.

The simplest configuration is a **linear accelerator**. Acceleration is typically achieved via oscillating electromagnetic waves: the charged particles are only exposed to the electromagnetic wave when the electric field is pointing in the direction of acceleration, and shielded from the wave when it points in the opposite direction. This allows for the acceleration of charged particles over great distances. The longest and one of the most powerful linear accelerators in the world today is the Stanford Linear Accelerator located at SLAC National Accelerator Laboratory in Menlo Park, CA. At its peak, it was able to accelerate electrons and positrons to 50 GeV over a 3 km length (see Fig. 15). SLAC was built in the mid-1960s and was the site of many major discoveries in particle physics, including the discovery of the quark structure of protons and neutrons. There is currently a proposal to build a much larger International Linear Collider in Japan, which would be between 30-50 km in length and could potentially accelerate electrons and positrons to energies exceeding 500 GeV.

The other main accelerator type in particle physics is the **circular accelerator**. The idea is that, while linear colliders are restricted by engineering considerations to be $\lesssim 100$ km, beams that travel in a circle can be gradually accelerated over many, many trips around the accelerator ring. This gives a much larger integrated distance in Eq. (69). As a result, circular colliders can, in principle, achieve much higher energies than linear colliders. In practice, however, circular colliders face many challenges. For instance, the fact that the charged particles travel in a circle mean that they are constantly being accelerated, and accelerated charged particles emit "synchrotron" electromagnetic radiation that slows them down. Thus, each pass around the circle *both* accelerates the particle due to applied electric fields and slows it down due to synchrotron radiation. Additionally, very powerful magnetic fields are necessary to bend the charged particles in a circle while they are being accelerated. The longest and most powerful accelerator in the world, the Large Hadron Collider (LHC) at CERN, Geneva, Switzerland, accelerates protons to 6.8 TeV over a circumference of 27 km (see Fig. 16); it is also used to accelerate lead ions for specialized "heavy ion" runs. Prior to the LHC, the 27-km tunnel

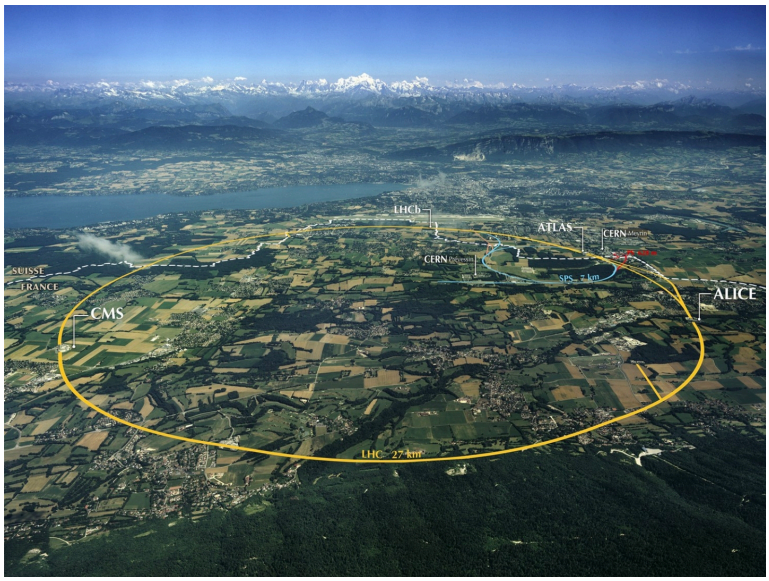

Figure 16: Configuration of the circular accelerator at the Large Hadron Collider (LHC), which spans the border between Switzerland and France. Protons are gradually accelerated over each pass around the 27-km accelerator. Beams of protons are accelerated in each direction: the two beams are then collided at one of four main interaction points. Situated at each interaction point is a detector that studies the debris coming out of the collision. Source: CERN.

was used for the Large Electron-Positron Collider (LEP), which accelerated electrons and positrons to energies of up to 104 GeV. Sites for a future circular accelerator are under consideration at CERN and in China.

What are the pros and cons of each type of acclerator? The main drawback of circular accelerators is the synchrotron radiation emitted by charged particles as they traverse the ring. For a circular collider, it is very challenging to accelerate electrons and positrons to energies exceeding ~ 250 GeV. The reason is that the radiated power per turn scales as

$$P \sim \frac{E^4}{R^2 M^4}. \tag{70}$$

Since the synchrotron radiation increases as the fourth power of energy, and making the ring larger only reduces the synchrotron energy quadratically, it quickly becomes prohibitive to build a sufficiently large ring for a circular collider of relatively low-mass electrons and positrons. Thus, a linear accelerator is the only option for going to energies $\gtrsim 1$ TeV for an electron-positron accelerator. Both linear and circular accelerators could be used for colliders at sub-TeV energies.

For protons, however, the story changes significantly. Because synchrotron radiation scales like M^{-4}, and the proton is some 2000 times heavier than the electron, the synchrotron radiation is much less important for a proton accelerator. This is why proton accelerators such as the LHC can reach vastly higher energies than electron accelerators. For protons, the main challenge is to provide sufficiently large magnetic fields to keep the protons moving in a circle; if such magnetic fields are not provided, the protons would simply punch through the sides of the beampipe. At the LHC, superconducting magnets with fields of up to 8 Tesla are required to fully accelerate the protons. Higher-intensity magnets are prohibitively expensive for the time being; a future proton accelerator with energies up to 100 TeV would only be possible with a longer tunnel than the LHC ($\sim$ 100

km), which would require a smaller amount of bending by each magnet, and sufficient research and development to make ~ 20 T magnets practical for future application at such large scales.

In recent years, there has been increased interest in building a **muon collider**. Muons have about a tenth of the proton mass and so share the reduction in synchrotron energy relative to electrons. Muons also have an advantage over protons: since protons are composite particles made up out of quarks, each quark in the proton only carries a small fraction of the total proton energy. On the other hand, muons are elementary particles and so their energy is not sub-divided among constituents. The result is that a muon collider could achieve similar sensitivity to energies and particle masses than proton colliders but with much shorter accelerators. Many challenges remain, including the fact that muons are short-lived particles that decay away in flight, but many particle physicists are now studying the possibility that a muon collider could be a good candidate for a future high-energy collider.

It is worth noting that, in particle physics, we don't always need to operate accelerators at the highest energies to make a new discovery. For example, there may exist new particles that are relatively light ($\lesssim 10$ GeV) but are produced only very, very rarely in a collision. In this case, it is preferable to have a **high-intensity** accelerator that does not reach the highest energies but can deliver orders of magnitude more collisions per second than the LHC. Examples of this include the "B-Factories", Belle II at KEK in Japan and the former BaBar Experiment at SLAC, which study in detail the decays of b quarks but could potentially discover other types of hidden particles. There are many other accelerators around the world that seek to understand lower-energy phenomena of the Standard Model and its extensions, either by precisely measuring rare processes among Standard Model particles or by looking for the direct production of new types of particles.

4.2 Types of Collider

Once high-energy particles, such as electrons, positrons, or protons, are accelerated to very high energies, they are smashed together. The outgoing particles from the collision are analyzed to reconstruct the physical processes occurring at the interaction point. The nature of the particles that are accelerated and collided depends on the physical processes that we wish to study. In this section, we review some of the most important types of collider; however, there are others that we don't discuss here.

Electron-positron collider: Electrons and positrons are, to our knowledge, fundamental particles without constituents. When they encounter one another, they annihilate and the resulting energy can be used to reconstitute new types of heavy particles, as well as other particles from the Standard Model. Typically, both the e^+ and the e^- vanish in the collision, and so the only particles observed in the detector are those produced in the collision, *i.e.*, those we care most about measuring. This is advantageous because the collision and its final-state products can be much better measured and understood in an electron-positron collider without any residual particles from the initial state. Another advantage of electron-positron colliders is that electrons and positrons only interact via the weak and electromagnetic forces, not the strong force, and so their collisions are (relatively) rare; this means we can record the outcome of most collisions in an electron-positron collider. This allows for a very "clean" environment for studying weak-interaction processes and other interactions that happen only very occasionally. Since particles not involved in the strong interactions are collectively known as leptons, an e^+e^- collider is an example of a **lepton collider**.

The primary disadvantage of electron-positron colliders is that it is challenging to accelerate electrons and positrons to very high energies in circular accelerators due to synchrotron radiation. The highest CM energies ever achieved in e^+e^- collisions was 209 GeV at the LEP collider at CERN

Figure 17: Event displays of particle collisions that allow us to visualize the outgoing products of the collision. (Left) Event display at the ALEPH experiment at LEP, which was an electron-positron collider (Source: ALEPH Collaboration website). (Right) Event display at the CMS experiment at the LHC. It is evident that there are many more particles produced per collision (and there can be over 100 separate proton-proton collisions occurring in a single picture!). (Source: CMS Collaboration website)

in the 1990s. This energy was insufficient to produce the mass energies of Higgs bosons or top quark-antiquark pairs at LEP, and going to higher energies to study such particles will require an extraordinary investment in a longer tunnel. e^+e^- colliders are able to reconstruct the final-state collision products to a high degree of precision, however, and the LEP collider (along with a similar e^+e^- collider at the SLAC accelerator known as SLC) provide the most precise data ever measured on various aspects of the SM forces and particles.

Proton-proton collider: As mentioned in Section 4.1, protons are easier to accelerate to higher energies than electrons due to the reduced synchrotron radiation. Therefore, proton colliders have the advantage of being able to access by far the highest energies ever attained in the lab. For example, the CM-frame total energy for proton-proton collisions at the LHC is over fifty times higher than the e^+e^- collisions at LEP, even though they use the same tunnel.

Any particle that is made up of quarks are known as **hadrons**, and pp colliders are a form of **hadron collider**. The primary drawback of hadron colliders is that there are *a lot* of strong-force interactions between protons at a pp collider. Unlike at e^+e^- colliders, which offer very clean environments, proton colliders send sprays of hadrons in all directions during a collision due to the various strong interactions that occur among the constituents of the colliding protons. Since we typically only care about one part of the collision, namely the final state produced in the interactions of a single quark or gluon pair between protons, it is very challenging to work backwards and reconstruct what happened in the part of the collision we are concerned with. For a comparison of lepton and hadron colliders in this regard, see Fig. 17.

An additional challenge of hadron colliders is that, unlike e^+e^- colliders where both initial state particles annihilate into the final state particles, remnants of the protons are left behind after the "primary" collision. This is because a proton collision typically occurs with a quark in one proton knocking out a quark from the other proton. Since a proton is made up of *several* quarks, there is a proton fragment left behind that emits large amounts of gluon radiation into the detector via the strong interactions. All of this suggests that hadron colliders are relatively "messy" environments, where it is difficult to precisely reconstruct the outcome of any given collision.

Hadron and electron colliders therefore fulfill complementary roles: hadron colliders are excellent

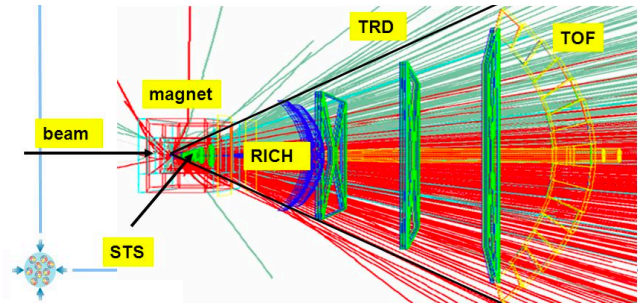

Figure 18: Schematic display of the proposed Compressed Baryonic Matter fixed-target experiment at Darmstadt. A high-energy particle impinges on a target at rest. Conservation of momentum dictates that the particles produced in the collision should travel in the same direction as the incident beam, and a detector is placed behind the target to identify any particles produced in the collision. (Source: CBM Collaboration)

at producing extremely heavy particles due to their world-leading energies. They are also most adept at producing new particles that interact via the strong force. Electron-positron colliders are limited in energy but allow precise studies of particle properties and interaction rates due to their clean environments. This is why hadron colliders are typically termed "discovery machines" (with the top quark discovered by the hadron collider at Fermilab, and the Higgs boson discovered by the LHC), while electron-positron colliders are "precision machines" that allow careful tests of the Standard Model.

Fixed-Target Colliders: The last type of particle collision we consider is one in which a beam of high-energy electrons, protons, or other particles are directed into a target at rest. For example, the earliest collision experiments at SLAC occurred with a high-energy electron beam directed into a target of deuterium gas at rest. The electrons then scatter off of the protons and (less frequently) electrons inside of the deuterium atoms. For an example of a fixed-target experiment, see Fig. 18. Neutrino beams are also sometimes directed into targets to study the weak interactions leading to neutrino-proton or neutrino-neutron scattering.

Because fixed-target collisions have one of the colliding particles at rest, the energy in the CM frame is much lower than in collisions between two high-energy particles. For instance, at the LHC two 6.8 TeV protons in opposite directions are collided into one another, giving a 13.6 TeV CM energy. By contrast, if a 6.8 TeV proton is directed into a target at rest, the CM energy is $E_{\rm CM} \approx \sqrt{2E_p^{\rm incident} M_p} \sim 100$ GeV. The lower CM energy at a fixed-target experiment is compensated for by a much larger increase in rate of collisions, or luminosity. The reason is that it is pretty easy to fire a high-energy beam into a target at rest, whereas precisely colliding two separate beams of high-energy particles is very difficult to do. In other words, if I fire an electron into an Avogadro's number of atoms, it is going to hit *something*, whereas two protons fired at one another are likely to miss.

Exercise: Consider a fixed-target experiment where particle A is collided at high energies into a target at rest made of particle B. Assuming that A is ultra-relativistic, $E_A \gg M_A$, show that the collision energy in the CM frame is $E_{\rm CM} = \sqrt{2E_A M_B}$.

The upshot is that fixed-target experiments are good for studying phenomena at lower energies but that require *a lot* of collisions. This includes looking for extremely rare processes such as new

particles that interact very weakly with electrons or protons. Fixed-target experiments are also logistically easier to implement, and so are typically cheaper than other kinds of colliders. In the end, a combination of lepton, hadron, and fixed-target colliders are used to study a variety of forces and particle properties at a wide range of energies.

4.3 Particle Detectors

Let's recall that we are ultimately interested in measuring:

1. Total cross sections of various processes, $\sigma(A + B \to C + D + \ldots)$. Or, in other words, the *fraction* of total collisions that give rise to the process $A + B \to C + D + \ldots$);

2. The differential cross sections with respect to various kinematic variables, such as scattering angles and combinations of final-state momenta. Or, in other words, what the distribution of final-state momenta (magnitude and directions) looks like.

Since each of these can be predicted by the theory of the Standard Model, measuring both can be used to test our current understanding of the known forces and potentially discover new interactions.

Presumably, we understand the kinematics of the initial-state particles because the collisions occur under controlled conditions. In order to reconstruct the cross sections, we therefore need to measure properties of the final-state particles coming out of the collision. In particular, we need to know:

1. The identity of the final-state particles. For instance, if I am measuring $\sigma(pp \to e^+e^-)$, I need to identify collisions with final-state electrons and separate them from other collisions that have final-state muons, photons, etc.

2. The momentum of each final-state particle. This allows us to reconstruct the angles and energies of each particle, and consequently to find out how often the collision occurs in a given kinematic configuration (*i.e.,* the differential cross sections).

Detectors at collider experiments are specifically engineered to accomplish these tasks in order to measure the cross sections to the desired precision.

First, we need to know what particles we are detecting. Most SM particles have lifetimes that are sufficiently short that they decay prior to hitting the sensitive parts of the detector. Therefore, there is a relatively short list of particles that we most commonly see coming out of the collision: electrons, muons, photons, protons, neutrons, charged pions (these are electrically charged hadrons that are a bound state of a quark and an anti-quark), and neutrinos. These particles span a wide range of properties:

- The masses of the particles range from $\sim$ GeV for protons and neutrons to $M = 0$ for photons (and, to an excellent approximation, neutrinos).

- The particles can be either electrically charged (protons, neutrons, pions, muons, electrons) or neutral (photons, neutrons, neutrinos). Charged particles can be positive or negative.

- They also have various charges under the other forces: the hadrons experience strong interactions, while electrons, muons, photons, and neutrinos experience only the electromagnetic and/or weak forces.

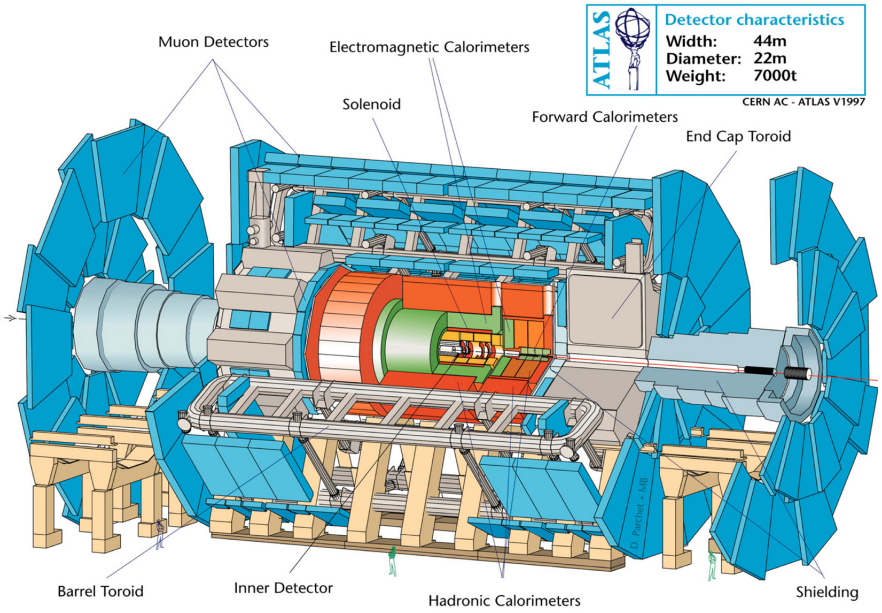

Figure 19: Schematic diagram of the ATLAS detector at the LHC. To allow you to get a sense of the scale, you can notice the standing human-size figures on the ground next to the detector. (Source: ATLAS Collaboration)

Since no two particles have identical properties, a detector can be made up of various components, each of which is able to measure *one* of the particle properties. By combining information from all detector components, it becomes possible to tell apart the different particle types.. This process is known as **particle identification** (often abbreviated as "PID"), and it is one of the most important parts of a collider experiment.

Along with simply identifying the type of particle, the detector is constructed to be able to measure each particle's momentum as well as it can. The detector therefore needs to be large enough that it can get a good read of the direction and energy possessed by each particle. For example, charged particles have their momenta measured by studying how a given particle deflects in a magnetic field. In order to measure this deflection to the needed accuracy, detectors are typically enormous; this accentuates any tiny magnetic deflection that a very high-momentum particle might experience. For example, the detectors at the LHC are up to 25 metres in diameter!

Detectors should be hermetic, meaning that they capture as many of the outgoing particles from the collision is possible. The reason is that, if we miss particles through gaps in the detector, we miss out on vital information about the process taking place: for example, it would be impossible to discover the Higgs boson decaying to $\gamma\gamma$ if in a typical collision one of the photons wasn't detected! It also hampers our ability to study the parts of the differential cross section that tends to shoot particles through the gaps; for instance, if we have a gap at an angle θ_0, we clearly cannot measure how often particles are emitted at the angle θ_0. Multi-purpose detectors such as those at the ATLAS and CMS experiments at the LHC are able to reconstruct particles moving in any direction out from the collision point, except for particles that are within $\lesssim 2^\circ$ of the beam. Such detectors are often referred to as having "4π coverage", because the total area of the unit sphere is 4π and this suggests that the detector is picking up particles emerging from the collision at (almost) any angle.

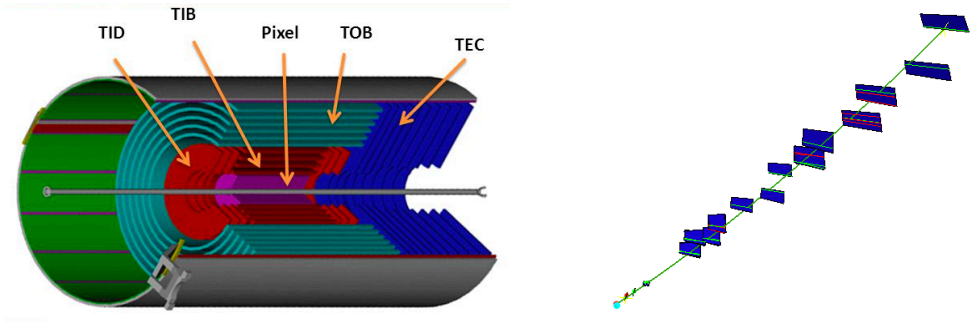

Figure 20: (Left) Diagram of the tracking system at the CMS experiment at the LHC. It is made up of some 10 million strips, each of which can "fire" when a charged particle passes through. (Right) The blue rectangles illustrate strips that have "fired" due to a passing charged particle. It is then possible to fit the trajectory of the particle and reconstruct its path based on these hits. (Source: CMS Collaboration)

In Fig. 19, we show a schematic diagram that illustrates the different sub-systems of the ATLAS detector at the LHC. Below, we will go into more detail about what each component does. It is difficult to really comprehend the scale of these experiments. For example, there are currently about 3000 scientists who are members of the ATLAS experiment. It takes a huge team to design, build, and operate such a machine, not to mention analyze the data it records. The LHC experiments are among the largest scientific collaborations in history.

Detector Components

Tracking system: The first order of business is to identify the precise point of the collision, conventionally know as the **primary vertex**. The reason is that doing so can allow us to identify particles that do not originate from the collision point and are not by-products of the initial interaction.

Tracking systems consist of strips of material such as silicon that are ionized when charged particles pass through and strip off electrons from atoms in the material; these electrons are collected and measured by applying an electric field to the sample, and this allows us to see if a high-energy charged particle has passed through the strip or not and when. If these strips are very finely segmented, then it is possible to tell where the charged particle passed through the tracker by seeing which parts of the strip "fire". Sophisticated algorithms then combine the hits in many strips to reconstruct the path of a single charged particle. The tracking system of the CMS experiment, along with a reconstructed track, are shown in Fig. 20. Measurements of the charge collected in each strip can give an estimate of how much ionization energy is deposited by the passing charged particle into the strip.

Typically, a very large magnetic field (2–4 Tesla) is applied to the inner part of the detector containing the tracker. Charged particle trajectories bend in the magnetic field, which means that they followed a *curved* path. The radius of curvature of the path is $R = |\vec{p}|/B$, where $|\vec{p}|$ is the momentum transverse to the magnetic field, B. Positively and negatively charged particles bend in opposite directions, and so the tracks are used to identify the sign of the electric charge as well. Therefore, the tracking system not only determines the paths of charged particles, but also can measure their momenta by looking at the track curvature. At the LHC, the tracking system has typical radius of ~ 1 m.

Calorimetry: Neutral particles leave no tracks. Likewise, very high-momentum particles do not curve much in the field, and so the tracker cannot provide a precise measurement of the energies of

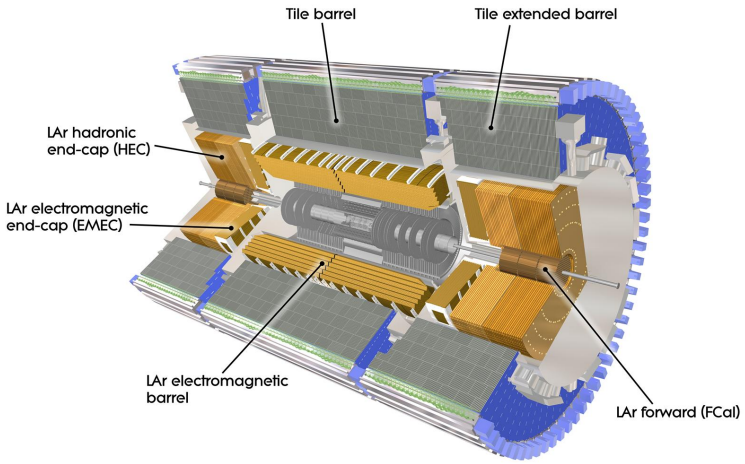

Figure 21: Diagram of the calorimeter system at the ATLAS experiment at the LHC. The inner, electromagnetic calorimeter uses liquid argon, while the outer hadronic calorimeter is made of tiles of plastic scintillator. The calorimeter needs to be quite deep to stop the energetic particles coming out of the collision. The tracking system is shown in grey in the middle. (Source: ATLAS Collaboration)

such particles. Instead, energies of these particles are measured by having them slam into a part of the detector called the **calorimeter**. The calorimeter consists of very dense materials and stops most particles that pass through. The kinetic energy of the particle "heats up" the material in the calorimeter as it is stopped, and by precisely measuring this energy deposition in the calorimeter, it is possible to infer the energy in the original particle. In addition to the dense materials that provide maximum stoppage power, the calorimeter needs some way of measuring the amount of the energy deposition. In some calorimeters, a material called a scintillator gives off light in an amount proportional to the deposited energy. By measuring this light, it is possible to determine the energy of the particle stopped in the calorimeter. Some particles (such as muons) are *not* stopped, but still give up some of their energy by interacting with detector materials; this is also measured.

There are typically two types of calorimeter. The **electromagnetic calorimeter (ECAL)** is designed to respond to electrically charged particles, while the **hadronic calorimeter (HCAL)** is designed to slow particles via nuclear scattering processes (*i.e.*, strong interactions). The calorimeters at the ATLAS experiment are shown in Fig. 21. When particles are stopped in a calorimeter, the energy is not deposited in a single instant. Instead, the incident particle emits photons and knocks loose some electrons or nuclei, which in turn deposit their energy downstream in the calorimeter in an ever-widening cone. This initiates what is known as a "shower" in the calorimeter, and different particles and energies tend to result in different shower shapes. This is useful in getting the most accurate energy measurement, as well as contributing to particle ID. An example of a shower inside of the electromagnetic calorimeter at the ATLAS experiment is shown in Fig. 22. The calorimeter system at ATLAS extends out to a radius of about 6 m.

Muon system: Most particles (electrons, photons, and hadrons) are stopped in the calorimeters. Hadrons are stopped because of their strong interactions with nuclei, while electrons and photons are very highly relativistic and lose their energies in the ECAL through forms of radiation known as

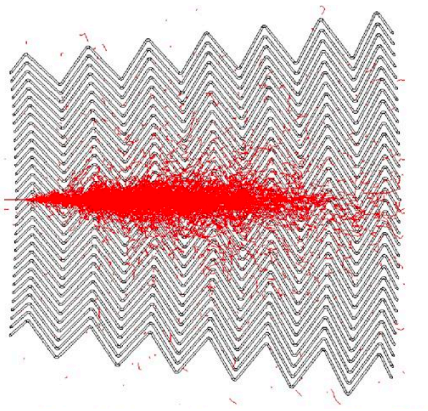

Figure 22: Simulation of a shower inside of the electromagnetic calorimeter of the ATLAS experiment at the LHC. An electron with 50 GeV of energy enters the calorimeter at left. This electron scatters off of atoms inside the calorimeter, leading to photon emission, production of e^+e^- pairs, and electrons produced from ionization. These in turn initiate their own new collisions, giving rise to a distinctive shower shape. (Source: ATLAS Collaboration)

bremsstrahlung and pair production. Muons, however, do not interact through the strong force and are not as relativistic as electrons, and therefore do not lose enough energy to be stopped. Outside of the calorimeters, a dedicated **muon system**, or muon spectrometer, is constructed to measure the trajectories of these muons.

The muon system is like a much larger version of the tracking system. Since the distance of the muon system is $\sim$ 10 m from the interaction point, the curvature of tracks induced by a magnetic field can be much more pronounced than in the inner tracking system, allowing for a more precise measurement of muon momentum when linked up with the inner track and energy depositions in the calorimeter.

Particle ID

By combining the measurements in the various detector components, it is possible to distinguish different types of particles. For instance, a photon is characterized by the absence of a track followed by a deposition of energy in the ECAL and none in the HCAL, while an electron is characterized by the *presence* of a track followed by a deposition in the ECAL and none in the HCAL. Charged and neutral hadrons like protons and neutrons are identified by whether or not there exists a track and the relative sizes of the energy depositions in the ECAL and HCAL. The particle ID methods for the CMS experiment are illustrated in Fig. 24.

An important particle to consider is the neutrino. It interacts only via the weak interactions, and it typically leaves the detector without scattering off of any of the detector components. Thus, a neutrino is *completely invisible* from the point of view of the detector. The only way we can infer its presence is through the fact that we do not measure it, and the disappearing particle appears to us as a violation of conservation of momentum. Since momentum conservation is a very well-established property of the Standard Model, the neutrino momentum is calculated as the "missing momentum" that should be there if momentum is conserved.

There are a few other particles worth mentioning that have distinctive features. A b-quark decays before hitting the detector, but its lifetime is long enough that it does not decay immediately after being produced. The lifetime of a b-quark is about 10^{-12} s, which if the b-quark were travelling

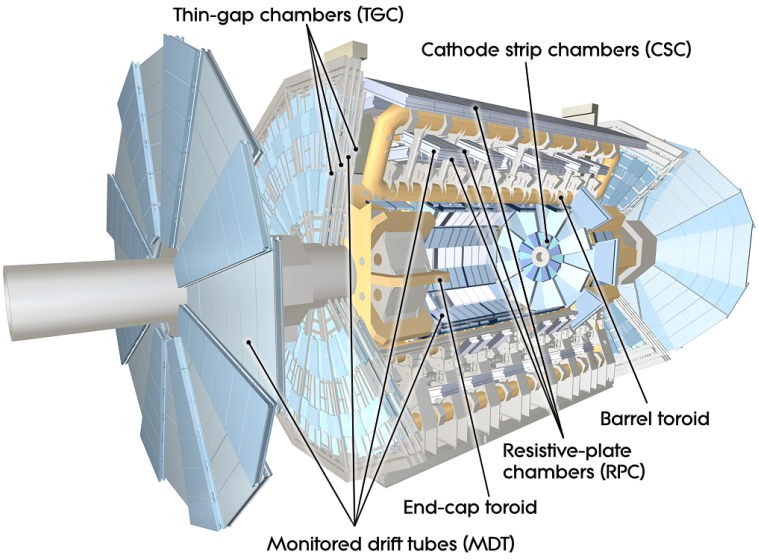

Figure 23: Diagram of the muon system, or muon spectrometer, of the ATLAS experiment at the LHC. The brown-yellow pipes are large toroidal magnets that generate a magnetic field in which the muons bend. Layers of tracking material allow an accurate reconstruction of the muon trajectory and (with the help magnetic field) their momentum. (Source: ATLAS Collaboration)

near the speed of light could correspond to a distance of $\sim 100 - 1000$ μm once taking into account relativistic time dilation. Thus, the presence of a b-quark can be inferred by looking at its decay products and noticing that they are **displaced** from the other tracks originating from the primary collision point by $\sim$ mm (see Fig. 25). Similarly, other long-lived particles that decay at distances $\gtrsim$ mm can also be identified: this is particularly useful for discovering new particles beyond the SM with long proper lifetimes ($c\tau \gtrsim 1$ mm).

With a perfect detector, it would be possible to state with certainty whether a passing particle is an electron, photon, or whatever else. However, we deal with real-life detectors with real-life noise, and so it is possible for one particle to fake another type. For example, an electron can scatter off some material in the detector and convert most of its energy into a photon: the detector might then register the particle as a "photon" even though it emerged from the collision point as an "electron" (see Fig. 26). Experimentalists characterize the possibility of mis-measurement with two related quantities:

- **Efficiency:** this is the probability that the detector will correctly identify the object. For example, the photon efficiency is the change that the detector registers a photon given that the collision produced a photon.

- **Mis-identification rate:** this is the probability that the detector will incorrectly identify object A as object B. For example, the electron-to-photon mis-identification rate is the percentage of real electrons that are reconstructed by the detector as photons instead.

Usually these rates are a function of the kinematics of the particle (energy, transverse momentum, and so on). Ideally, we want a 100% efficiency for, say, photons. However, the only way we could accomplish that is if we call *everything* a photon, in which case we would also have a 100% misidentification rate. Conversely, the only way we can get a 0% misidentification rate is if we call *nothing*

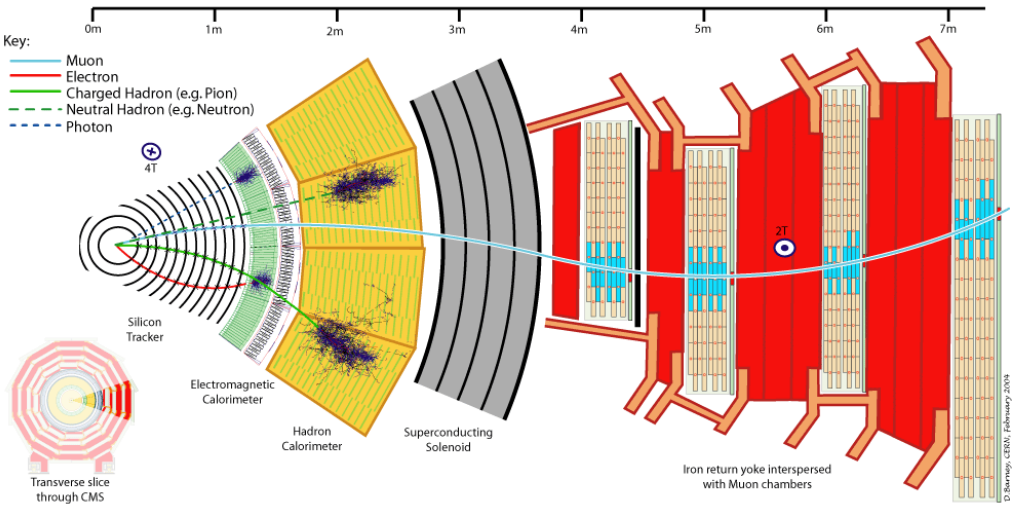

Figure 24: Schematic diagram illustrating the particle identification procedure in the CMS experiment. The entire detector is used to distinguish the traces left by muons, electrons, photons, charged and neutral hadrons. (Source: CMS Collaboration)

a photon, in which case we also have a 0% efficiency. Experimental analyses have to be carefully optimized to keep as many of the desired particles as they can while rejecting fakes; the optimal selection will vary depending on what the analysis is and the relevant underlying SM processes. For an example of real experimental misidentification rates and efficiencies for electrons, see Fig. 26, which shows the efficiency of a true photon being reconstructed by the detector as a photon, as well as the fake rate (the rate that a true electron is reconstructed mistakenly as a photon). We see that, at high p_{T}, the electron efficiency is approximately 90%, with 10% of photons misidentified as electrons.

The Trigger

So far, we have assumed that the detectors can reconstruct every event. However, collisions at colliders are very frequent, and reading out all components of the detector takes a significant amount of time, data bandwidth, and storage. For example, a typical recording of the whole detector in a collision at the LHC is approximately 1 MB in size. Since collisions happen at the LHC at a rate of 30 MHz (*i.e.,* 30 million times per second), this corresponds to approximately 30 TB of data per second! There is currently no way to read out and store all of this information, nor would we particularly want to since most times particles undergo "boring" grazing collisions that are not of interest to us.

The LHC experiments are designed to store and reconstruct events at a rate of about 1 kHZ (*i.e.,* 1000 collisions per second). This means that more than 9999 out of every 10000 events is thrown away without being recorded. The system that decides which collisions to discard and which to keep is called the **trigger**. This is one of the most important components of the detector: if the trigger

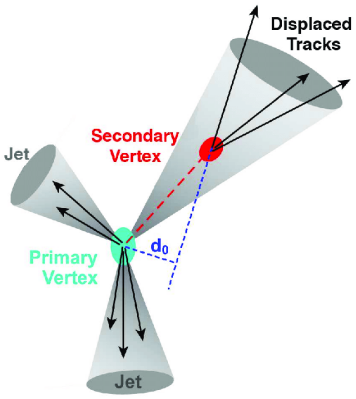

Figure 25: Schematic diagram of charged-particle tracks produced in a proton-proton collision. Most of the tracks originate from the collision point (the "primary vertex"), while the tracks coming from the long-lived b-quark decay originate from a separate, "secondary vertex". (Source: D0 Collaboration)

malfunctions, then potentially all of the most interesting data is thrown away!

The trigger is constructed out of multiple levels, each of which operates on successively larger amounts of information from the detector. The lowest level of the trigger is Level 1 (L1), and it takes a quick peek at each and every collision. It cannot read out detailed information from the whole detector, but instead looks for "interesting" things such as a large deposition of energy in one part of the calorimeter or the hint of an electron, photon, or muon based on hits in the calorimeter or muon system. Currently, no track information is available to the L1 trigger at ATLAS or CMS, but this may change in the future.

The L1 trigger operates at the 100 kHZ rate, and so about 1 in 300 events pass the L1 trigger. These events are moved on to the higher levels of the trigger; since there are fewer events per second reaching the higher-level triggers, it is possible to do more sophisticated reconstructions such as some type of track identification and more precise readouts of the detector information. This allows the final stage of the trigger, known as the High-Level Trigger (HLT) or Event Filter, to make the best informed decision about whether an event is worth keeping or not[9]. Because protons undergo copious strong interactions in scattering, the HLT tends to keep only very spectacular strong-interaction events, as well as events that show clear signs of a weak interaction or something else that is interesting (such as a Higgs boson). Under current conditions, an event passes the HLT if it contains a single electron or muon with a momentum perpendicular to the beam larger than 30 GeV; two charged leptons with momenta larger than 15-20 GeV each; a single photon with momentum perpendicular to the beam larger than 145 GeV; two photons with momenta larger than about 40 GeV. Collisions with only strong interactions only pass the trigger if the total deposited energy exceeds approximately 750 GeV!

[9]A notable exception at the LHC is the LHCb Experiment, which is specifically designed to study Standard Model particles such as b quarks. With a recent upgrade to the experiment, it is now possible to do a full real-time reconstruction of every single collision. They still must use a trigger to decide which collisions to keep, but fairly sophisticated algorithms can be run on each collision before making that decision. This is possible in part because the LHCb Experiment experiences a much lower rate of collisions, ~ 1 MHz., but also because of very sophisticated electronics and powerful high-throughput computing.

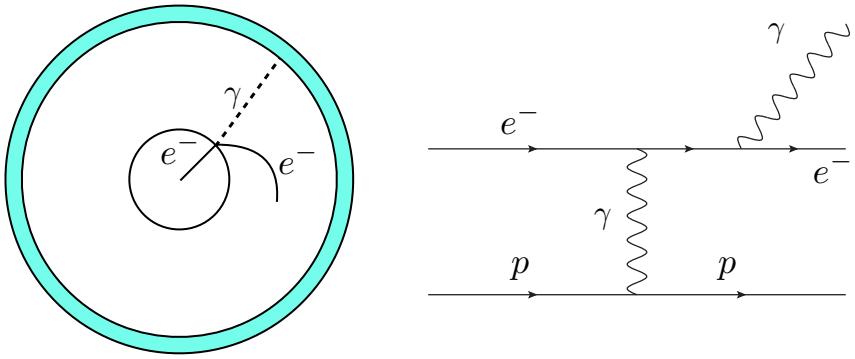

Figure 26: As an electron traverses the detector, it can scatter off of protons in the detector material, converting most of its energy into a photon. This leads to the misidentification of the true electron as a "photon" by the detector.

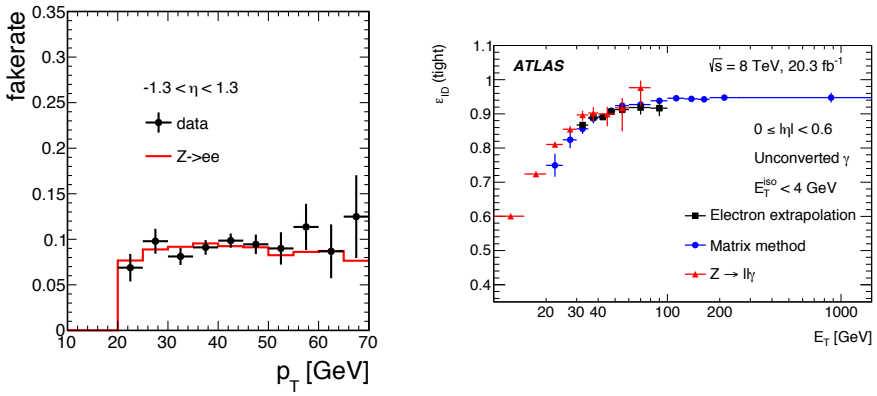

Figure 27: Electron-to-photon misidentification rate (fake rate, Source: J. Haase Diplomarbeit, ATLAS Collaboration) and photon efficiency in the ATLAS experiment, as a function of the electron momentum transverse to the beam axis (Source: `arXiv:1606.01813`)

It is good to know that the experiments are not merely keeping or rejecting events at random, but you may find aspects of the trigger program troubling. For instance, what if there is some signal of a new particle that we haven't thought to program into our trigger menu? This new particle could be produced at an enormous rate at the LHC, but if all of these collisions are thrown away, we would never find it! These questions have kept many theorists and experimentalists up at night, and we should be at least partly reassured by the fact that great care and thought have gone into trigger designs to avoid this nightmare scenario. Currently, it is hard to think of signatures that are missed by the current trigger strategy but that could actually be recorded onto disk without overwhelming the electronics. The triggers are also set up to be re-programmed (with the higher-level triggers easier to change since they are software based), so if a clever physicists such as yourself thinks of a new trigger idea to capture something interesting without keeping too much junk, the trigger can be updated to fill in those gaps.

There are other strategies to deal with rare signatures of new particles that have no way of passing the trigger. There are new techniques that go by the name of **trigger-level analysis**, in which small bits of events that do not pass the trigger are nevertheless recorded and examined. The problem with

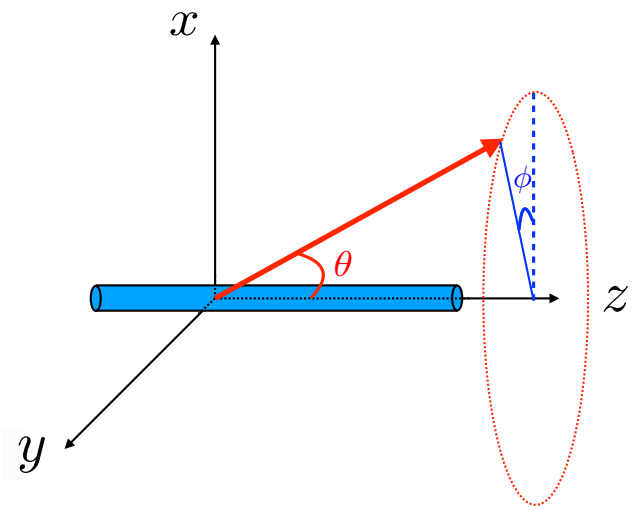

Figure 28: Spherical coordinate system used to describe directions of particle trajectories. The light blue cylinder represents the beam pipe; one of the beams moves in the z-direction and the other in the $-z$-direction. The particle trajectory is illustrated in red. The polar angle, θ, is the angle of the particle relative to the $+\hat{z}$ beam. The angle ϕ represents the angle of the momentum vector relative to the x-axis when projected into the xy-plane; if we hold θ fixed but change the orientation of the trajectory relative to the x-axis, we get the indicated dashed red circle lying in a plane perpendicular to $\hat{z}$. The azimuthal angle ϕ is the angle of the particle trajectory relative to the x-axis in this plane (*i.e.*, along the dashed circle).

this approach is that, as the full detector information is not stored, it is not possible to go back and assess whether the calibrations of your detector were working when the data was recorded. However, it is one powerful tool in the arsenal that can be used to try and ensure that no interesting signatures fall through the trigger gaps.

4.4 The Detector Coordinate System

So far, we have discussed how different components of the detector can measure the energy and momentum of particles produced in the collision. Since momentum is a vector, it has a *direction* associated with it, and we need to know how to specify this direction. In a particle collider, we already have a direction picked out for us, namely the direction of the beams. We define the direction of one of the beams to be the $\hat{z}$ direction (where the hat denotes a unit vector). This is illustrated in Fig. 28. We can specify the direction of an object by two angles relative to the beam: the angle θ relative to the beam, which is known as the **polar angle**; and the angle that the particle trajectory makes relative to the x-axis when projected into the xy-plane, which is known as the **azimuthal angle**. Note that the domain of θ is between 0 and π, while the domain of ϕ is between 0 and 2π.

The above angles form a spherical polar coordinate system. Suppose the magnitude of the momentum is $|\vec{p}|$. The projection of the momentum vector in the z-direction is $|\vec{p}|\cos\theta$. The radius of the red dotted circle in Fig. 28 (indicated by the dark blue line) is $|\vec{p}|\sin\theta$; this is the component of $\vec{p}$ lying in the plane perpendicular to the z axis. To obtain the x and y components of the momentum

vector, we use trigonometric relations involving the angle ϕ and get:

$$p_x = |\vec{p}| \sin\theta \cos\phi, \tag{71}$$
$$p_y = |\vec{p}| \sin\theta \sin\phi, \tag{72}$$
$$p_z = |\vec{p}| \cos\theta. \tag{73}$$

WARNING: The physics convention for spherical polar coordinates is different from the math convention; in particular, mathematicians swap the roles of θ and ϕ. We always use the convention where θ is the polar angle and ϕ is the azimuthal angle.

The momentum 4-vector in natural units is $p^\mu = (E, p_x, p_y, p_z)$. A measurement of $|\vec{p}|$, E, θ, and ϕ of a particle is sufficient to reconstruct the full 4-vector p^μ using Eqs. (71)-(73). Because $p^2 = M^2$, a measurement of each of the 4-vector components allows for the determination of a particle's mass. Alternately, if we know what kind of particle it is (say, an electron), then we already know what the mass is and a measurement of either E or $|\vec{p}|$ *and* θ and ϕ allow for the calculation of every quantity in the 4-vector using the relation $E^2 - |\vec{p}|^2 = M^2$.

Because the colliding beams exhibit cylindrical symmetry about the z-axis, there is nothing special about any particular value of ϕ; we expect all particle scattering rates to be the same for all values of ϕ. By contrast, particle scattering rates depend critically on θ. For example, grazing collisions are much more common that head-on interactions that scatter into large θ angles, and so the scattering rate becomes enormous at values of θ close to 0 or π. Most of the interactions we are interested in feature energetic collisions where much of the energy goes into producing new, heavy particles; this tends to give rise to final states with $\theta \sim \pi/2$. In any event, particle detectors can only get so close to the beam pipe without it being subjected to intense radiation from the beam.

In lepton colliders, the spherical coordinate system in Fig. 28 is used. At hadron colliders, there is an additional problem: while collisions in the lab are often done in the CM frame of the protons in which the total momentum of the two initial protons is zero, the actual collision takes place between individual quarks from each of the protons. These quarks carry less than the full proton momentum, so the lab is *not* in the zero-momentum frame of the colliding quarks. This means that, on an event-by-event basis, an additional Lorentz transformation is needed to get into the centre-of-mass frame.

We will not get into any details here, but the punchline is that it turns out to be simpler from the point of view of calculations to replace the angle θ with a new coordinate. This coordinate is called the **pseudorapidity**,

$$\eta = -\ln\left(\tan\frac{\theta}{2}\right). \tag{74}$$

As $\theta \to 0$ ($\theta \to \pi$), then $\eta \to \infty$ ($\eta \to -\infty$). Conversely, when $\theta = \pi/2$, $\eta = 0$. So collisions that significantly deflect the momentum of the incident particles occur at small values of η, while particles in the direction of the beam are at large values of θ. Because the relationship between η and θ is one-to-one, we can invert the relation to obtain

$$\theta = 2\tan^{-1}(e^{-\eta}). \tag{75}$$

A plot showing polar angles θ overlaid with pseudorapidities, η, is shown in Fig. 29. In most analyses at the multi-purpose LHC experiments (ATLAS and CMS), particles are required to have $|\eta| < 2.5$. In this range of angles, which constitutes about 160° out of the total available 180° in polar angles, track information for charged particles is available; for larger rapidities, only calorimeter information is available. Another LHC experiment, called LHCb, is specifically designed to look for particles that are somewhat aligned with the beam; this detector measure particles in the range $2 < \eta < 5$.

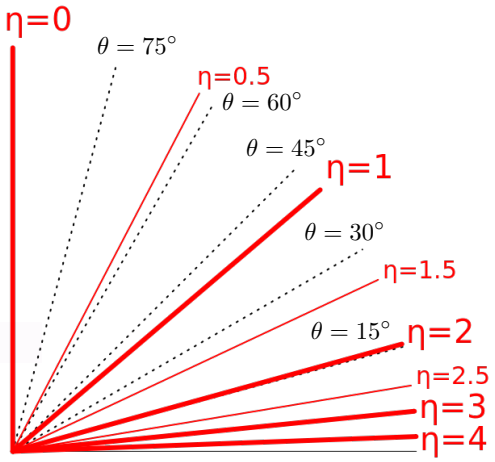

Figure 29: Comparison of the polar angle, θ, with the pseudorapidity, η. (Source: Wikimedia commons, with modifications)

(Optional) There is a related quantity known as the **rapidity**,

$$y \equiv \frac{1}{2}\ln\left(\frac{E+p_z}{E-p_z}\right). \tag{76}$$

It is possible to know that the difference in rapidity between two particles, $y_2 - y_1$, is invariant under Lorentz transformations along the z-direction. In the limit where a particle is massless, the rapidity becomes exactly equal to the pseudorapidity, η. The nice thing about η is that it is a purely geometric quantity, whereas y depends on the energy and momentum of the particle (and, consequently, its mass). Since most particles we will be dealing with are highly relativistic, we will use η exclusively in what follows, but you may see references to y in the research literature.

Transverse Momentum

As mentioned earlier, the most interesting collisions tend to have particles with large momentum perpendicular to the beam. Therefore, we are often interested in the **transverse momentum**, p_{T}, of a particle, which is defined as

$$p_{\mathrm{T}} \equiv \sqrt{p_x^2 + p_y^2} = |\vec{p}|\sin\theta. \tag{77}$$

Often, measurements of particle properties at a collider focus on transverse momentum. To specify the 4-momentum, we need 4 pieces of information. The information typically used at a hadron collider are the independent variables $(E, p_{\mathrm{T}}, \eta, \phi)$. Any other piece of information relevant for the momentum 4-vector can be derived from these quantities. For example,

$$p_x = p_{\mathrm{T}}\cos\phi, \tag{78}$$
$$p_y = p_{\mathrm{T}}\sin\phi, \tag{79}$$
$$p_z = p_{\mathrm{T}}\sinh\eta. \tag{80}$$

Identification of Neutrinos. Earlier, we mentioned that neutrinos escape unnoticed after a collision because they only interact via the weak force. Their existence can only be ascertained by an

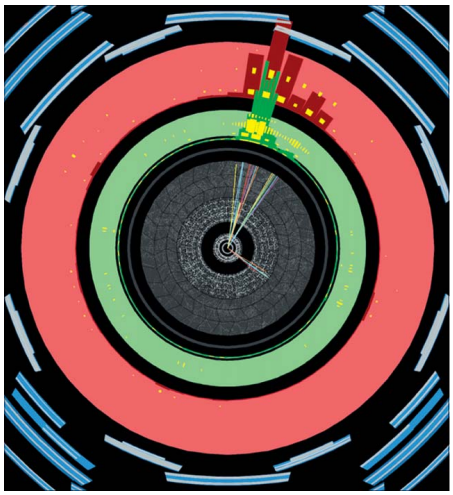

Figure 30: Event display of a collision with large missing momentum. The direction of the beam is into the page, and the picture shows the trajectories of particles in the transverse direction. It is evident that there is a large amount of momentum going towards the upper-right of the plot, while there is nothing going in the opposite direction. The missing momentum points towards the bottom-left. (Source: CERN Courier, ATLAS Collaboration)

apparent non-conservation of momentum by taking the initial net momentum and subtracting the observed momentum of every visible particle:

$$\vec{p}_\nu = \vec{p}_{\text{initial}} - \sum_{i \in \text{visible}} \vec{p}_i. \tag{81}$$

In hadron colliders, however, we have argued that we do not always know the initial momentum in the z-direction (because the quarks involved in the collision carry some *a priori* unknown fraction of the full proton momentum). Therefore, it is impossible at a hadron collider to determine $p_{\nu\,z}$. However, we know that there is zero net momentum initially in the direction transverse to the beams, and so we can identify any momentum carried off by neutrinos in the x and y directions:

$$\vec{p}_{\nu\,\mathrm{T}} = - \sum_{i \in \text{visible}} \vec{p}_{i\,\mathrm{T}}, \tag{82}$$

where the transverse vector is as defined in Eq. (77). The quantity in Eq. (77) is defined as **missing transverse momentum**, and the symbol is $\not{p}_{\mathrm{T}}$ (the slash designates "missing"). The more common name for missing transverse momentum is **missing transverse energy (MET)**, denoted by $\not{E}_{\mathrm{T}}$. It's a bit of a misnomer, because energy is a scalar quantity without a direction and so there is no sense in which there can "transverse energy", but the terminology is widely used and should always be interpreted as missing transverse momentum. $\not{E}_{\mathrm{T}}$ or $\not{p}_{\mathrm{T}}$ is useful for identifying neutrinos that are produced in a collider, as well as any other new invisible particles that may be produced (such as dark matter). An example of a collision event with large missing transverse momentum is shown in Fig. 30.

5 The Anatomy of a Hadron Collision

There are several advantages to colliding hadrons (and, in particular, protons) instead of electrons. Because protons and their constituents experience the strong force, proton colliders are ideal factories for producing new strongly interacting particles. As we've already seen, protons are also heavy enough that their acceleration is not limited by synchrotron radiation, and so much higher CM-frame collision energies are possible.

As we speak, the world's largest proton collider, the Large Hadron Collider (LHC) at CERN, is accumulating data at record-breaking energies. The hope is that evidence for new particles, forces, or other phenomena (like extra spatial dimensions!) are hiding in the data, and we just need to sift through the data in the right way or wait to collect enough data to find it. The analyses of LHC data proceed in the ways we've already described: theoretical calculations are done to predict SM cross sections, as well as the cross sections of hypothetical new particles, and the data are compared to see which hypothesis is the best fit. How this is done will be the subject of Sec. 6.

Before we get there, however, we have a more pressing concern. At high collision energies like at the LHC, the strong interactions are much "weaker" than at low energies due to the fact that the dimensionless strong interaction coupling gets smaller with higher energy collisions. This means that, at LHC energies, all of our calculations are done in terms of quarks and gluons in the initial and/or final states. This is all well and good, except:

1. Our initial scattering states are *not* single quarks or gluons. Instead, we collide protons, which are complicated bound states made up of many quarks and gluons. We somehow need to map our calculations with initial-state quarks and gluons onto the proton initial states.

2. Our detectors don't *observe* quarks or gluons. The confining property of the strong interactions means we can never see isolated quarks and gluons at macroscopic scales. Instead, the outgoing quarks and gluons from the collision get bound up into **hadrons** (such as protons, neutrons, pions, etc.), which are the collections of quarks and gluons that as a whole don't carry any colour charge. We somehow need to map our calculations with final-state quarks and gluons onto the observed protons, neutrons, and pions that fly out of the collision point and hit our detector.

Both of these considerations require us to cross the line from hadrons to quarks/gluons and vice-versa: because the strong interaction coupling becomes very strong at the transition between these phases of matter, we can no longer use Feynman diagrams to picture the particle interactions. Indeed, it is very hard to do any calculations at all of what happens in what we call the **non-perturbative** regime of the strong force. In light of these facts, it seems hopeless for us to actually make any concrete predictions for what we should see in our LHC detectors as the outcome of proton collisions.

All is not lost, however. If the collisions are happening at very high energies ($E_{\text{CM}} \sim 1$ TeV) and strongly interacting particles confine at much lower energies ($E_{\text{confine}} \sim 1$ GeV), then we might expect that some features of the high-energy quark/gluon calculation persist in our detector. For example, if I have a collision that sends a quark hurtling out of the collision point at very high energy in a particular direction, then even though it eventually gets bound up into hadrons, we should see a bunch of hadrons collectively shooting in that direction due to conservation of momentum (processes of $\Delta p \sim 1$ GeV can't appreciably change the momentum of a particle going $p \sim 1$ TeV). Similarly, there is some sense that we can think of the proton as a collection of quarks and gluons that allows us to map the initial-state proton into the quarks and gluons that participate in scattering. How we go about translating between the quark/gluon and hadron picture in hadron collisions will be the focus of this section.

Everything I describe in this section may appear to be more qualitative than in other sections. Part of the reason for this is that an advanced mathematical understanding of quantum field theory is necessary to access many of the technical details. Another reason is much more mundane: a lot of our picture of the non-perturbative nature of the strong force *is* based on intuition and hand-wavy explanations. This is not to say that there isn't a lot of quantitative data to back our current picture of the strong force, but rather that many of the developments have been based either on good theoretical guesses that give a reasonable fit to experimental data, or, phenomenological fits to the data with little solid theoretical backing for the ideas. Indeed, we can't even prove that the strong interactions confine quarks and gluons inside of hadrons: if you can do this, you will earn a million dollars from the Clay Mathematics Institute.

I don't work on many of these more technical issues, and so I will largely treat the procedures and ideas described here as a means to an end. However, if this topic piques your fancy, you should definitely consider pursuing research into perturbative and non-perturbative aspects of the strong force: as more time elapses and we have yet to find any new particles (apart from the Higgs boson) at the LHC, many theorists are working to unravel the mysteries of the strong force using the vast troves of LHC data.

Exercise: The strong interactions get weaker at higher energies. If a proton is accelerated to sufficiently high energy, it should therefore spontaneously fragment into a bunch of free quarks and gluons. True or false? Explain your reasoning.

5.1 The Parton Model of Hadrons

In this section, we examine the structure of hadrons (and, specifically, protons), describing them in terms of their constituent particles in a manner that can be used in calculations of high-energy collisions. First, however, I want to return to some of the ideas of Sec. 1 and look into how we know the structure of the proton in the first place. Let's start with an easier problem: what is the structure of the atom? This question is already answered, and we saw in Sec. 1 how the angular distribution of scattered alpha particles in turn depends on how the positive charge is arranged in the atom. If the positive charge were from some diffuse haze, most energetic particles would pass right through because they are sensitive to only a small volume of charge at each point in space, and this has a negligible effect on the trajectory. Instead, we see the alpha particles strongly deflected in some collisions, indicating that they are scattering off of some "hard" core particles. In the case of Rutherford's alpha particle experiment, that core is the nucleus.

If we perform a similar experiment to determine the proton's structure, we get similar results. Because quantum mechanics says that probing shorter distance scales requires going to higher energies, our incident particles need to be very energetic, $E \gtrsim$ GeV. Nevertheless, if we fire, say, electrons at a proton, we find that at high energies the rate and angular distribution of the scattered electrons resembles scattering off of small point-like particles residing inside of the proton. These proton constituents are collectively referred to as **partons**. We now know that the partons are actually quarks and gluons, but at the time of the early electron-proton scattering experiments, it was not yet known whether quarks were actual physical particles, and so the more generic name of parton was applied (and is still widely used).

Up until now, we have spoken of the proton quark (or parton) content in simple terms. For example, we know that the proton has charge +1 and is the lightest bound-state of three quarks because it is a baryon. This then suggests that the proton is composed only of the lightest quarks (u and d), and since the quarks have electric charges $Q_u = +2/3$ and $Q_d = -1/3$, the proton consists of a uud combination of quarks. Similarly, the neutron is neutral and is a udd combination of quarks.

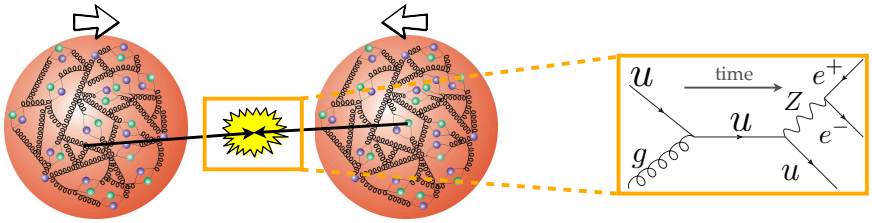

Figure 31: Schematic diagram illustrating two protons colliding. An up quark from one proton collides with a gluon from the other, producing a Z boson and an up quark.

As always, life is more complicated than this. The reason is that we know that quarks can emit gluons, which in turn can split into quark-antiquark pairs of any type of quark, of which there are 6: u, d, s, c, b, t. Because the strong interaction strength is *very large* for interactions at low momentum and energy (such as occur inside of a proton), then there is no penalty paid for these frequent splittings. Therefore, we have to think of the proton as not just a simple bound state of three quarks, but as a swarm of partons: quarks of all flavours as well as gluons. We can represent a proton as one of the blobs in Fig. 31, and we see it is made up of a churning mass of quarks, antiquarks, and gluons. When we collide two such protons, what is actually happening is that a parton from each object collides, and these form the initial states for our Feynman diagrams involving quarks and gluons.

What is remarkable about this is that it also means that there are *antiquarks* inside of the proton! Thus, to initiate a process $u\bar{u} \to Z$, to is sufficient to collide two *protons*; you don't necessarily need a proton and an antiproton! Of course, because the antiquarks come from gluon splittings inside of the proton, there are "fewer" of them (in a way that we will shortly quantify), but they are there nonetheless.

An important result that follows from this parton model is that the proton is made up not just of a few quarks, but also of gluons and, most importantly, *antiquarks*. Yes, there is antimatter inside of a proton! The reason is that the above process where gluons split to give quark-antiquark pairs necessarily leads to the production of antimatter. This means that, even if I collide two *protons* together, I might get a quark from one proton and an anti-quark from another proton to annihilate and give rise to something else entirely.

Exercise: If there are antiquarks inside of a proton, why don't they annihilate away the quarks inside of the proton? Does this cast doubt on proton stability?

The momentum 4-vector of the proton, which we call p_p^μ, is also split among all of the partons:

$$p_p^\mu = \sum_{i \in \text{partons}} p_i^\mu. \tag{83}$$

The three **valence partons**, which are the *uud* forming the core of the proton, carry most of the proton's momentum, but the components of the accompanying swarm typically have small momentum compared to the proton (remember that the strong interactions get stronger at low momentum/energy!).

Now, let's consider the scenario where a very energetic collision occurs involving a proton. For starters, we will consider the relatively simple case of an incident electron (with momentum p_e) and proton (with momentum p_p) colliding, and the resulting Lorentz-invariant momentum combination

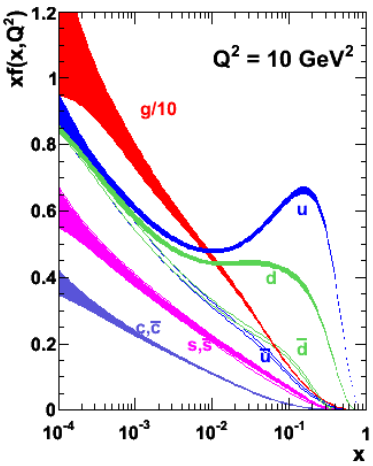

Figure 32: Parton distribution function, $f_i(x)$ (multiplied by x) for the proton. It is the probability of finding a parton of type i and momentum $|\vec{p}_i| = x|\vec{p}_p|$. Note that for high momentum fractions $x > 0.1$, we find mostly up and down quarks (with a 2:1 ratio in the probability), while at lower momentum fraction, we see gluons and antiquarks as well. (Source: MSTW Collaboration, `arXiv:0901.0002`)

$(p_e + p_p)^2 \gg M_p^2 \sim \text{GeV}^2$. In this case, the characteristic energy of the collision is much larger than the inverse size of the proton, and the electron is probing deep inside of the proton structure. According to our above discussion, what the electron should see is *not* a proton but instead a collection of partons. These partons are all interacting with one another characteristic time scale GeV^{-1} in natural units (because this is the typical energy of the strong interactions inside of the bound state), whereas the interaction of the electrons with the partons is much faster, $1/\sqrt{(p_e + p_p)^2} \ll \text{GeV}^{-1}$. In a simplistic application, the electron "hits" the partons in the space between partonic collisions with one another, and so we can essentially treat the partons as free particles in the initial state. This property is known as **factorization**: the time scale of the "hard" collision between the electron and the parton is much shorter than the time it takes for the other partons to react to the collision, and so we can to a good approximation ignore them. This means, for example, that if the electron ejects a quark from inside of the proton, this high-energy collision happens much faster than the confining force which tries to tug that quark back inside the proton.

To model the collision, we first calculate the cross section for an electron to scatter off of each parton. For example, suppose we are interested in the rate of an electron scattering off a proton and producing a new muon-antimuon pair in the collision. We can calculate the scattering cross section for each quark species, $e^- q \to e^- q \mu^+ \mu^-$, which is a function of the incident p_e and p_q. This is called the **partonic cross section**. Then, the total electron-proton cross section for this process is given by

$$\sigma(e^- p \to e^- + \text{hadrons} + \mu^+\mu^-) \quad = \sum_{q \in u,d,s,c,b,t,\bar{u},\bar{d},\bar{s},\bar{c},\bar{b},\bar{t}} \int_0^1 dx\, \sigma(e^- q \to e^- q \mu^+ \mu^-) \times f_q(x), \quad (84)$$

where $f_q(x)$ is the fraction of the proton that is composed of the parton q carrying momentum

$$|\vec{p}_q| \equiv x|\vec{p}_p|. \quad (85)$$

Let's unpack this formula a bit: it tells us that for each parton type q, we have to take the partonic cross section $\sigma(e^- q \to e^- q \mu^+ \mu^-)$ and multiply it by the chance of finding a parton q inside the

proton, then sum over all partons. We also have to account for the fact that $\sigma(e^- q \to e^- q \mu^+ \mu^-)$ depends on the parton momentum, while the total proton cross section should be independent of this. So, we actually have to take the cross section for a given p_q multiplied by the chance of finding a parton *with that momentum* inside of the proton, and then integrate over all possible momenta (ranging from $\vec{p}_q = 0$ to $\vec{p}_q = \vec{p}_p$). The function $f(x)$ is known as a **parton distribution function (PDF)**, because it tells us how the proton can be divided up into types of partons and what fraction of the total proton momentum the partons carry. The parton distribution function for the proton is shown in Fig. 32.

The PDFs must satisfy various constraints. For example, the total momenta of all the partons such should sum to the proton momentum due to momentum conservation. Therefore,

$$\sum_{A \in \text{partons}} \int_0^1 dx\, x\, f_A(x) \;=\; 1. \tag{86}$$

We also know that the proton consists of a net number of 2 u quarks and 1 d quark, plus some additional $u\bar{u}$ and $d\bar{d}$ pairs that come from gluon interactions. Thus, if there are Y pairs of $u\bar{u}$ from splitting, then the total number of u quarks is $2 + Y$ and the total number of $\bar{u}$ quarks is Y; the difference between u and $\bar{u}$ quarks is therefore 2. This gives the following constraints on the PDFs:

$$\int_0^1 dx\, [f_u(x) - f_{\bar{u}}(x)] \;=\; 2, \tag{87}$$

$$\int_0^1 dx\, [f_d(x) - f_{\bar{d}}(x)] \;=\; 1. \tag{88}$$

Finally, all other flavours of quarks come from gluons splitting to $q\bar{q}$ pairs, and so the PDF of the quark must equal the PDF of the anti-quark. For example, the charm quark contribution to the proton satisfies:

$$\int_0^1 dx\, [f_c(x) - f_{\bar{c}}(x)] \;=\; 0. \tag{89}$$

In general, the heavier the quark, the less it contributes to the PDF because there is a kinematic penalty to producing it in a gluon splitting inside of the quark.

Exercise: Suppose we want to study scattering of neutrinos off of protons via the exchange of a W^- boson, $\sigma(\nu_\mu p \to \mu^- + \text{hadrons})$. Which partons in the proton can ν_μ scatter off of? Draw the relevant Feynman diagram, and write the total hadronic cross section $\sigma(\nu_\mu p \to \mu^- + \text{hadrons})$ schematically as a sum/integral of the partonic cross sections over the PDFs.

We now generalize our result to the ultra-high-energy collision of *two* protons. It is exactly the same as before, except we now have one parton drawn from each proton that then interact. The result is that we need to draw a parton from a PDF for each proton. For example, consider the process $pp \to \bar{t}t$, where t is the top quark. Some of the parton-level Feynman diagrams are shown in Fig. 33. We then have

$$\sigma(pp \to \bar{t}t) \;=\; \int_0^1 dx_1 \int_0^1 dx_2\, f_g(x_1)\, f_g(x_2)\, \sigma(gg \to t\bar{t}) \tag{90}$$

$$+ \sum_{q=u,d,s,c,b,t} \int_0^1 dx_1 \int_0^1 dx_2\, f_q(x_1)\, f_{\bar{q}}(x_2)\, \sigma(q\bar{q} \to t\bar{t}), \tag{91}$$

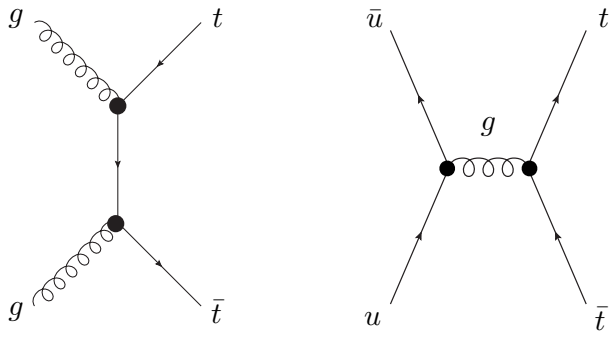

Figure 33: Example Feynman diagrams contributing to $pp \to t\bar{t}$ scattering from (left) gluon initial partons; (right) up quark-antiquark initial partons.

where x_1 is the fraction of the first proton momentum carried by the first parton, and x_2 is the fraction of the second proton momentum carried by the second parton.

In summary:

- The parton model says that, in high-energy scattering, we can think of an incident proton as a collection of free "partons", which includes all of the quarks and gluons;
- The fraction of the proton that is comprised of each parton is given by the parton distribution function (PDF), $f_i(x)$, where x is the fraction of the proton momentum carried by the parton i;
- Scattering cross sections involving initial-state protons are found by first computing the scattering cross sections involving each possible initial-state parton combination, and then integrating these rates over the momentum fraction x and summing over each type of parton.

5.2 The Parton Shower and Hadronization

In the context of the parton model, we saw how we can describe scattering involving initial-state protons in terms of the constituent partons, namely the quarks and gluons. In other words, we showed how we could take partonic scattering rates like $\sigma(e^- u \to \nu_e d)$ to the total hadronic scattering rate $\sigma(e^- p \to \nu_e d)$. However, we have not yet seen how to treat the *final-state* partons. In particular, we can think of the collision dramatically ejecting one (or more) of the partons from the proton, which then goes in a completely different direction. Describing its fate will require an investigation of how particles carrying colour charge (like quarks and gluons) get bound back up into hadrons like protons and neutrons.

The evolution of a "hard parton" (namely a parton that has a momentum significantly different from either incident proton) occurs in two distinct phases:

1. **The Parton Shower:** The initial collision occurs too quickly for the strong interactions to keep the parton contained inside the proton; however, as the final-state partons fly apart, enough time elapses that the strong interactions begin to kick in. The partons then begin to radiate a large number of gluons, which in turn split into $q\bar{q}$ pairs, which in turn radiate more gluons, etc. This is called a "shower" because of the cascade nature of the strongly interacting radiation. The end result of the parton shower is a large multiplicity of quarks and gluons: these are largely aligned with the directions of the original "hard" partons, but there are also low-energy quarks and gluons emitted in all directions.

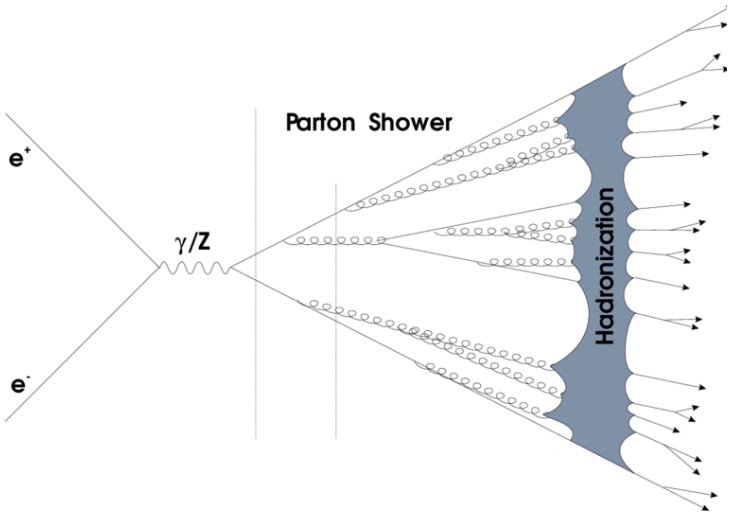

Figure 34: Depiction of the parton shower and hadronization, the processes by which final-state quarks and gluons get bound into final-state hadrons. (Source: C. Padilla Aranda, Ph.D. dissertation, ALEPH Collaboration)

2. **Hadronization:** Because the strong interactions become arbitrarily strong over long distances (typically larger than GeV^{-1}), the quarks and gluons that result from the parton shower begin to bind up into **hadrons** that do not carry any colour charge, such as protons, neutrons, and pions. The way in which this occurs is very mysterious and outside of the realm of applicability of most of our calculations. Physicists have come up with various physical arguments for how this occurs, and then derive a numerical model for how often certain configurations of quarks and gluons get converted into particular hadrons. However, these models have a lot of free parameters that are simply adjusted to match experimental data and give probabilitistic descriptions of what we expect from the data: we currently have no good way of performing hadronization that is correct on a collision-by-collision basis.

We show a graphical depiction of this process for the case of $e^+e^- \to Z \to q\bar{q}$ in Fig. 34. After showering and hadronization, the result is a collection of hadrons flying out of the collision point. These hadrons have fixed properties, such as mass and electric charge, and can be registered in our detector. It is therefore essential to perform the shower and hadronization steps to map our calculation of $\sigma(e^-p \to \nu_e d)$ onto the physical by-products of the collision and check that our calculations are correct.

We tackle each of these subjects in turn, and in the next section discuss how our physical understanding of showering and hadronization motivate better methods of reconstructing final-state partons at colliders.

The Parton Shower: The parton shower exists in the realm of the "perturbative" strong interactions, meaning that we can actually do controlled calculations of how the partons radiate as they move away from the collision point. The calculations are done using Feynman diagrams: recall that Feynman diagrams work with only a finite number of strong-interaction points (*i.e.,* gluon vertices) because each time you add a gluon, you pay a penalty of the dimensionless **strong coupling**, typically denoted by $\alpha_s < 1$. Therefore, diagrams with more gluon vertices have additional factors of α_s^n to some power n and are less important to the calculation.

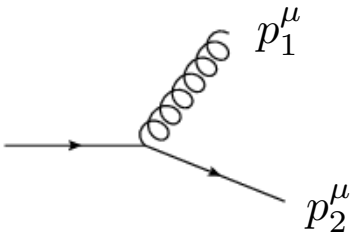

Figure 35: Feynman diagram illustrating an up quark radiating a gluon of momentum p_1, giving rise to a final up quark of momentum p_2.

While this is true, we also have to be mindful of the fact that there are additional properties of the strong interactions. For example, the strong couplings is secretly a function of the energy scale of the interaction μ, and the interaction strength gets *stronger* for lower energies, $\alpha_s(\mu_1) > \alpha_s(\mu_2)$ for $\mu_1 < \mu_2$. This means that not all diagrams with gluons are made equal, and particularly those involving *low-energy* gluons give a larger contribution. Instead of adding up all Feynman diagrams with a particular number of gluon insertions, it turns out that some of the diagrams count *more* than others, and so theorists have come up with ways of cleverly adding just those diagrams to give a better estimate of the rate.

It is perhaps not surprising that emitting a low-energy gluon is easier than emitting a high-energy one. Consider an example that is familiar to us: the hydrogen atom. Using Feynman diagrams arguments, one might suppose that an electron and a proton would only emit one (or a few) photons as they pass. However, we know that in real life, electrons and protons in an atom *continually* exchange photons to keep the electron and proton orbiting around one another. Indeed, we know that a free electron zooming through space comes with its own electric field, which is essentially a cloud of low-energy photons. So any object that carries electric charge is expected to come with some multiplicity of low-energy photons. Similarly, a particle with *colour charge* under the strong interaction is expected to set up its own colour charge field made up of a swarm of low-energy gluons.

What are the properties of this gluon emission? Well, we know that it preferentially happens at "small energy", but we now need to be careful: energy is not a Lorentz-invariant quantity. Additionally, since gluons are massless, and up and down quarks have small masses, the Lorentz invariant quantity $p^2 = M^2c^2$ is always small compared to the confinement energy for individual partons. What matters is the *relative* energies of two states, which is encoded in the Lorentz-invariant scalar product $p_1 \cdot p_2$. Let's consider the emission of a gluon from an up quark emerging from a collision, illustrated in Fig. 35. The Lorentz-invariant product is

$$\begin{aligned} p_1 \cdot p_2 &= E_1 E_2 - \vec{p_1} \cdot p_2 && (92)\\ &= E_1 E_2 - |\vec{p}_1||\vec{p}_2| \cos\theta && (93)\\ &= E_1 E_2 \left(1 - \cos\theta\right), && (94) \end{aligned}$$

where in the last line we have used the relativistic expression $E = |\vec{p}|$, which is appropriate for the (near-)massless outgoing partons. We see that there are two limits when this Lorentz-invariant quantity goes to zero, and hence we expect the strong interactions to be *very strong*:

1. The splitting is **soft** such that one of the parton energies goes to zero, $E_{1,2} \to 0$. Thus, we expect to see a lot of low-energy gluons (and, hence, hadrons) everywhere.

2. The splitting is **collinear**, meaning that $\theta \to 0$ and $\cos\theta \to 1$. If the two partons point in the same direction, there is a large possibility of gluon emission even if the gluon carries a lot of

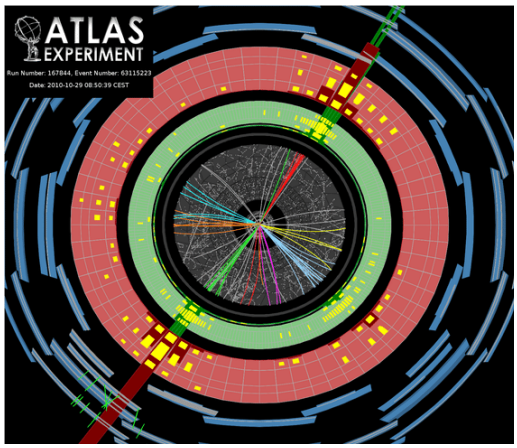

Figure 36: Event display from the ATLAS experiment of a collision with two high-p_{T} jets. This is a transverse slice of the detector (perpendicular to the beam axis). You can see two clear, collimated bursts of tracks and depositions in the calorimeters corresponding to the jets. Each of these is initiated by a parton, either a quark or a gluon. (Source: CERN Courier)

energy. We thus expect to see a lot of hadrons that are pointing in the same direction as the original quark.

And, indeed, this is what we see. Fig. 36 shows an event display from a collision where the different colours correspond to activated parts of the detector. It is possible to see two highly collimated sprays of particle tracks and depositions in the calorimeters. These are called **jets**. In experimental terms, a final-state quark or gluon always shows up as a jet. We also see numerous other, lower-momentum tracks that do not fall into the jet. These are partly due to the soft radiation from the parton shower. We will study jets in more detail in Sec. 5.3.

Hadronization: Hadronization is the process by which partons are bound back up into hadrons. The partons themselves carry colour charge, while the hadrons do not. Since the strong force is "confining", meaning that it is not possible to find a particle carrying colour charge separated at long distance from any other object with colour charge, everything we register in the detector is a hadron, while all of the outcomes of our earlier calculations are in terms of partons: quarks and gluons.

There are several physical pictures of how hadronization occurs. None of them can be derived from first principles and none fits the data perfectly well, but they are adequate for building simulations of particle collisions (provided that we try to minimize the sensitivity of our observables to the precise details of hadronization). The one I describe here is called the **string model**.

Before we dive into the physics of hadronization, however, I want to review some basic features of confining interactions. A confining interaction means that two particles cannot be indefinitely separated from one another. Another way of saying this is that the potential, $V(\vec{r}_1 - \vec{r}_2)$, between two particles must satisfy

$$V(\vec{r}_1 - \vec{r}_2) \to \infty \text{ when } |\vec{r}_1 - \vec{r}_2| \to \infty. \tag{95}$$

The infinite potential suggests that it is not physically possible to completely separate two such objects.

One simple potential that satisfies the property of confinement is the potential for a simple harmonic oscillator,

$$V(\vec{r}) \;=\; \frac{1}{2}k|\vec{r}|^2, \tag{96}$$

where k is the spring constant for the oscillator. This models the motion of two point masses connected by a spring. Regardless of the initial kinetic energy, the restoring force of the spring causes a decceleration of the relative motion, and eventually the particles begin moving back towards one another. We can therefore imagine our quarks and gluons being attached to one another by very stiff springs and being unable to be permanently separated.

The above argument holds if the spring is infinitely strong and cannot break. However, we can imagine another scenario where the spring is quite brittle. Once the potential energy in the spring exceeds a certain critical value, V_{break}, then the spring snaps in half. We now are left with two springs that are *below* the breaking point, and so the ends of the springs can continue to fly apart until the daughter springs in turn reach their breaking points.

We can now apply the spring analogy to our theory of partons. The potential from the strong force is like the spring: as particles with colour charge get sufficiently far apart, the simplest picture is that the colour force pulls the partons back together. Recall that the strong force is communicated by the exchange of gluons. The gluons therefore act (in a very simplistic picture) like our spring. However, we also know that gluons can split into quark-antiquark pairs, $g \to q\bar{q}$. Because the masses of the light quarks such as up and down are $M_q \sim$ MeV, the characteristic energy scale of the strong force is $E_{\text{strong}} \sim$ GeV, and the characteristic energy scale of partons coming out of an LHC collision are $E_{\text{collision}} \sim$ TeV, we find that the most likely thing to happen is that the gluons create quark-antiquark pairs to "break the spring" when two partons get too far apart. Eventually, enough $q\bar{q}$ pairs are created that all of the partons get bound up into colour-neutral hadrons. This is how the outgoing partons from the collision get converted into the hadrons that we see in the detector.

There are two ways that partons with colour can get bound up into hadrons that do not carry any net colour charge:

1. Any colour-anticolour combination is neutral under the strong force. For example, a quark that carries "red" colour charge will cancel out the colour charge of a "red" antiquark. Thus, a quark-antiquark pair of the same colour give rise to an object with no net charge under the strong force. We call such objects **mesons**, and they include particles like the pions, kaons, and rhos. Because they are made up of a combination of matter and antimatter, the quark-antiquark pair inside of the meson sometimes annihilate one another through the electromagnetic or weak forces, and consequently all mesons can decay into non-strongly interacting final states like electrons, photons, and neutrinos.

2. A combination of all three colours is also neutral under the strong force. For example, a collection of three quarks such that each quark carries one of red, blue, and green colour will give rise to a colour-neutral hadron. Similarly, a collection of a red, blue, and green antiquark is also colour-neutral. This type of hadron is called a **baryon**, and protons and neutrons are examples of baryons. Because they consist of either all quarks *or* all antiquarks, the lightest baryon (which is the proton) can be absolutely stable. Indeed, we have not observed any evidence for proton decay, and the lifetime of the proton must exceed something like 10^{24} years (recall that our universe is only itself about 10^{10} years old!). Apart from the proton and neutron, other examples of baryons include the hyperon and Δ baryons.

Advanced Exercise: The Δ^{++} baryon is understood as a bound state of 3 up quarks and has charge +2 (these are the valence quarks, and we ignore other partons in the Δ^{++}). The spins of the quarks are also aligned with one another and the quarks have zero net orbital angular momentum. Explain how the existence of this state is allowed given the Pauli Exclusion Principle, which states that identical fermions (such as electrons and quarks) are forbidden from occupying the same quantum state.

We now have a glimpse at how it is possible for computer programs to simulate high-energy collisions involving final-state partons. After the parton shower, the "string" model involving gluons mediating a spring-like force. This spring can snap into $q\bar{q}$ pairs, and the outcome of this succession of string snappings is a huge number of lower-energy partons that can be grouped into hadrons that carry no net colour charge. As you can imagine, this model is very rough and has a number of properties that we cannot calculate, such as the exact tensile force that the spring can sustain before snapping. In practice, our crude models are then tuned to the data: the various free parameters describing how the snapping and evolution of the springs occurs can be adjusted so that the outcome of the simulation reproduces the observed experimental results. Given the lack of theoretical control we have over these calculations, though, we don't want to take the outcome of the hadronization simulations *too* seriously, especially if we apply our simulation in new scenarios that we have never studied experimentally before. This is particularly challenging if we introduce, say, new particles beyond the Standard Model that can decay into partons or hadronize themselves: without having actually *seen* one of these particles, we can't say for certain what it will actually do, and so we should take the results of hadronization with a grain of salt.

5.3 Jets: Revealing Partons at High-Energy Colliders

The basic picture we have so far is:

1. The primary interaction is a collision of two partons, such as quarks and gluons, which reside inside the proton.

2. Any outgoing partons first undergo a parton shower, which results in a large number of other partons that are either high energy and preferentially emitted in the direction of the original parton, or are low energy and are emitted in any direction.

3. These partons can be thought of as connected by springs, which snap into additional $q\bar{q}$ pairs. The mess of partons are then grouped up into objects that carry no net colour charge known as hadrons.

4. The resulting hadrons are seen in the detector.

Stages 1 and 2 are reasonably well understood. Stage 3 is substantially less so. Stage 4 represents our actual experimental observations. Working from theoretical prediction to experimental observation (or vice-versa), we have to pass through the murkiness of Stage 3. It is therefore advantageous to try to pick ways of processing the data that are as *insensitive* as possible to the effects of hadronization.

It is simplest to start with what is *most* sensitive to the hadronization model, and then steer clear of these effects. For example, counting up the number of hadrons in the final state is probably a bad idea, because it is very sensitive to the details of how the springs snap and so on. Similarly, the precise position of all of the relative hadrons in the detector is not good either because that can also depend on how the spring evolution occurs. The detector has some spatial uncertainty of particle reconstruction and so we cannot any figure out where each hadron is with infinitely good precision anyway.

What we actually want to do is group together bunches of hadrons that are likely to have originated from the same initiating quark or gluon at the beginning of the parton shower. For example, because of the abundance of *collinear* radiation (along the direction of partons emitted from the primary collision), we expect to see a huge spray of hadrons along the direction of the original quark or gluon. If we combine these hadrons in a way that is insensitive to the number or precise position of each hadron but instead only cares about their collective motion, we will do a pretty good job of reconstructing the properties of the partons coming out of the original collision.

The sprays of hadrons going in the direction of the original parton are known as **jets**, and they were predicted in the 1970s before colliders were produced at sufficiently high energies to make jets. We already saw an example of jets in Fig. 36. As with many things in pattern recognition, it is often easier to pick out jets by eye than to come up with a robust mathematical algorithm of exactly what should be outside the jet for a general collision event without hitting all kinds of pathological cases that cause trouble. This is particularly true at a hadron collider, where the presence of the initial state partons in the event and their own radiation can mess things up. It was not until just before the LHC started running that satisfactory solutions were developed for grouping hadrons into jets in a way that was amenable to theoretical calculations and experimental observations, and where the algorithms were fast enough that they could be applied to large datasets such as at the LHC.

The process of merging hadrons together into jets is called **jet clustering**. In order to go over the currently used jet clustering algorithms, we must review two important kinematic quantities: the particle transverse momentum, p_{T}, which is the momentum transverse to the beam direction (see Sec. 4.4), and the angular distance between two particles,

$$\Delta R \equiv \sqrt{(\Delta\phi)^2 + (\Delta\eta)^2}, \tag{97}$$

where $\Delta\phi$ is the difference in polar angles between the particles, and $\Delta\eta$ is the difference of pseudorapidities from Eq. (74). When $\Delta R \gg 1$, the particles are moving off in different directions, while if $\Delta R \ll 1$, the particles are moving essentially in the same direction. The reason why we use the pseudorapidities in Eq. (97) is because of their nice properties under Lorentz transformations along the beam direction.

I will be describing a jet clustering algorithm know as the **anti-k_{T}** algorithm, developed in 2008. It is the most widely used algorithm for determining jets in the LHC experiments. It takes as an input the list of all final-state hadrons measured by the detector, as well as a **jet radius**, D. D is the characteristic size of the jet; the goal is to make it large enough that it contains most of the collinear radiation from the parton shower, but small enough that it does not pick up too much unrelated soft radiation from other partons. Typically, $D \approx 0.4 - 0.5$, although in some applications it can be much larger. The jets are then determined as follows:

1. Take a list of four-momenta, one for each hadron p_i^μ.

2. If there exists only one particle in the list, call it a jet and terminate the algorithm. If there are no particles in the list, terminate the algorithm.

3. For every pairing (i, j) of different particles, compute the distance measure

$$d_{ij} \equiv \min(1/p_{\mathrm{T},i}^2, 1/p_{\mathrm{T},j}^2)\,(\Delta R)_{ij}. \tag{98}$$

   Out of all (i, j) pairs, find the single combination with the smallest distance measure, d_{ij}^{min}.

4. For every particle i, compute the quantity

$$d_i \;=\; D/p_{\mathrm{T},i}^2. \tag{99}$$

   Find the particle with the smallest value of d_i, which we'll call d_i^{min}.

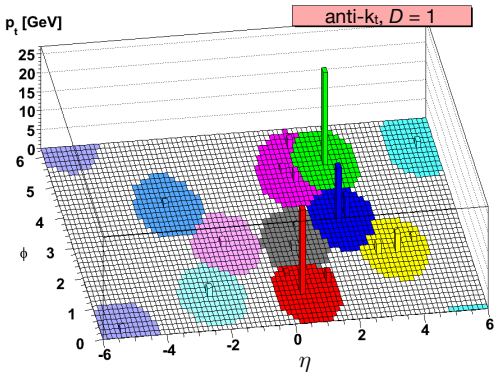

Figure 37: Plot showing depositions of transverse momentum in the detector as a function of ϕ and η. The height of the bars show directions where there is a large transverse momentum of objects hitting the detector in that direction. The shaded regions are jets found using the anti-k_{T} algorithm with jet radius $D = 1$. The result of the algorithm are nearly circular jets of radius ~ 1 centred on the most energetic object hitting that part of the detector. (Source: `arXiv:0802.1189`)

5. If $d_i^{\min} < d_{ij}^{\min}$, then call particle i a jet, remove it from the global list and return to Step #2. Otherwise, proceed to the next step.

6. Take the pairing with the *smallest* value of $d_{ij} = d_{ij}^{\min}$. Delete the entries i and j from the global list, and in their place add an entry to the list containing the sum of the two particles, $p_{\mathrm{new}}^{\mu} = p_i^{\mu} + p_j^{\mu}$. Return to Step #2.

The end result is a list of jets with a four-momentum associated with each jet. The algorithm is called an iterative clustering algorithm because pairs of hadrons are repeatedly combined into a single object at each step until we are left with a short list of jets at the end. The outcome of the anti-k_{T} algorithm applied to hadrons in a particular example collision is shown in Fig. 37.

Admittedly, this seems like a fairly strange algorithm, and so we now go through a step-by-step discussion. Step #1 involves making a global list of the particles measured by the detector. The goal is to eventually combine these and end up with a set of jets. In Step #2, we make the observation that if there is only one particle, then it is trivially a jet and we end our search there. Similarly, if there are no particles then we cannot identify any jets and terminate our search.

In Step #3, we define a "distance measure" that is used to quantify how close together different particles are. In this case, particles are close together if one of them has a very large transverse momentum (*i.e.,* $p_{\mathrm{T},i}$ is very large) or if the angular separation of the particles, $(\Delta R)_{ij}$, is very small; note that a large ratio of momenta and/or small angle separation of partons is exactly the type of splitting that is characteristic of the strong force. We wish to find the pair that is closest together; this is the pair that is most likely to be combined into a single object in the next iteration. In Step #4, we instead determine the characteristic distance that we would expect to find if a particle is a distance D away from all of its neighbours. Remember that if $(\Delta R)_{ij} > D$, we typically *don't* want the particles to end up in a jet together, whereas we do want them to end up in a jet if $(\Delta R)_{ij} < D$.

In Step #5, we take the closest pair of particles and compare the distance between them with the maximum characteristic distance we would expect for a jet of radius D. If the pair-wise distance is *larger* than the maximum characteristic distance for one of the particles, then this tells us that the pairing should not be combined because they are too far apart to be considered a single object. In this case, the particle with the smallest d_i (or, equivalently, with the largest $p_{\mathrm{T},i}$), is called a "jet"

and removed from the list. Otherwise, the pair of particles are combined into a single particle by summing their four-momenta, and the algorithm is iterated on the new list (which now has one fewer particles in it because we have combined two into one).

We see that our algorithm gives us a concrete set of rules that allow us to combine objects into jets, and that only allows a combination of two particles into a single jet if their angular separation is smaller than the jet radius parameter, D. In particular, our rules tell us that particles with large transverse momenta get clustered *first*, because they tend to have the smallest values of d_{ij}. This is a property of the anti-k_{T} algorithm: it tends to take high-momentum particles and quickly combine them with surrounding low-momentum particles. Because summing a low-momentum particle p^{μ}_{low} with a high-momentum particle p^{μ}_{high} gives a resulting momentum $p^{\mu}_{\mathrm{low}} + p^{\mu}_{\mathrm{high}} \approx p^{\mu}_{\mathrm{high}}$, the jet resulting from all of these combinations tends to still point in the direction of the original high-momentum particle(s). The result is that the jets look kind of like if you drew a circle of radius D around the highest-momentum original hadrons in the event (prior to any clustering). See Fig. ???

What do we do once we have grouped the hadrons into jets? Well, in a very simple and idealistic picture, each jet represents one energetic quark or gluon that originates from the primary collision. We can then do our theoretical calculations involving quarks and gluons and then try to look for the corresponding jets in real collisions. This is essentially what is done in every LHC collision study. We have therefore, more or less, come full circle: we started with an ideal picture of quarks and gluons and ended up with experimental objects (jets) that mimic the original quarks and gluons. It is important, however, to have gone through this whole journey for two reasons: one is that the jets do not *exactly* match the original partons involved in the collision, and so it is important to keep in mind what we are missing when we make this necessarily over-simplistic mapping. The second is that we need to know on a practical level how we can reconstruct the original partons from the spray of final-state hadrons seen by the detector.

6 Analysis of Collider Results

So far, we have learned how theories can make theoretical predictions of **cross sections** for the outcomes of particle scattering processes, and we have seen how collider experiments actually work to reconstruct what happens in individual scattering events. We now need to check and see whether or not the experimental observations match the theory!

As an example, let us meet Alice (a theorist) and Bob (an experimentalist). Alice has a theory for dark matter which predicts the existence of a new particle, χ, that is invisible to the detector and gives rise to missing transverse momentum. She estimates that if a recent LHC dataset were analyzed, there should be 100 events with $\not{p}_{\rm T} > 1$ TeV if her theory of dark matter is correct: 50 of these come from SM processes with Z bosons decaying to neutrinos, and the other 50 come from the production of the dark matter particle. If her theory is incorrect, then an LHC experiment should only see the 50 events arising from SM processes.

Bob is interested in this, so he designs a search where he looks only at events with $\not{p}_{\rm T} > 1$ TeV in his 2016 dataset. After carefully scrutinizing his data and ensuring he isn't making any major mistakes, he find that there are 87 events meeting the desired criteria. This certainly looks like more than Alice's prediction for SM processes alone. The question on everyone's mind is: has Bob discovered dark matter? While it's fun to make predictions and to do experiments, we actually want to be able to make a quantitative statement about whether the experimental evidence matches the theory or not!

First, however, we want to see what the qualitatively influences our interpretation of the outcome of the experiment. To do this, we need to understand the **uncertainty** in the measurement and its comparison with the theoretical predictions. Then, once we have quantified the uncertainty, we will need to do a **statistical analysis** to quantitatively determine with which hypothesis the experimental data agrees most given the uncertainties.

6.1 Uncertainties

There are three broad classes of uncertainty relevant for comparing theory with experiment:

1. **Theory Uncertainty:** In order to make a comparison between a theoretical prediction and an experimental result, we first have to do the theoretical calculation (correctly). The problem is that we almost never know how to get the *exact* theory answer; apart from the few examples of quantum mechanical systems that you may have seen in class, there are almost no physical scenarios that can be exactly solved. Instead, various approximation schemes and numerical techniques are used to extract the result. If we employ an approximation in doing a theoretical calculation, we cannot know for certain the value of the parts that we are *not* including in the calculation, although there are ways of estimating this contribution. Consider, for example, approximating the the function $f(x) = \sin x$ with $x \ll 1$ by its Taylor series,

$$f(x) \approx x - \frac{x^3}{3!} + \dots \tag{100}$$

   If we truncate the series at $f(x) \approx x$, one way of estimating the size of the neglected terms is by computing one more power (*i.e.,* the x^3 term) and seeing how big this is. If the series is well-approximated by the leading term, then we expect that the x^3 term should be small, the x^5 term should be even smaller, and a reasonable guess for the "theory uncertainty" on this value is approximately $x^3/3!$. There is no unique way to estimate this uncertainty (by definition, we can't know for certain the parts that we are missing!), and so any stated theory uncertainty should be taken with a grain of salt.

In our above example, we can consider the scenario where the theory uncertainty on Alice's calculation for both the dark matter production and SM contribution is 10%. If this is the case, we should more properly have stated her prediction as 100 ± 10 events if dark matter is made in the detector, and 50 ± 5 if it is Standard Model only. If the experimental result quoted is accurate and precise (*i.e.,* has small uncertainty), then we would say that the experimental evidence favours the dark matter hypothesis. If, instead, the theory uncertainty is 100% (which means 50 ± 50 for the SM hypothesis, 100 ± 100 for the DM hypothesis), then we could conclude that there is no way to distinguish between the two hypotheses since the observed 87 events is within the realm of possibility of each.

2. **Systematic Uncertainty:** The idea of systematic uncertainty can be summed up with the question "To what degree is my experiment measuring what I think it's measuring?" Suppose I am trying to determine the mass of an object by putting it on a very sensitive scale. If I have not properly tared the scale (*i.e.,* there is an offset in the result), then it doesn't matter how sensitive my scale is because I will always measure the mass wrong in a **systematic** way.

   A good experimentalist will work very hard to identify all of the biases in the experiment and correct for them. For the case of the biased scale, an experimentalist could use weights of *known mass* (called a control mass), measure them on the scale, and use this to identify a bias. Once this had been determined, they could simply subtract the offset value when measuring weights of unknown mass in order to improve the accuracy of the result. However, there is a limit to how effective this procedure is: for instance, we may not know the control mass to arbitrary precision, or there may be an error in extracting its mass in some other manner. The residual uncertainty in correcting for the bias is known as systematic uncertainty.

   Returning to the more abstract world of collider physics, one source of systematic uncertainty is that Bob's measurement of missing momentum might be offset. This is a reasonable concern, since missing momentum is by definition something that you are *not* observing, and so any mis-calibration in the detector can introduce a distortion in the measured missing momentum. The missing momentum can be calibrated by studying the response of the detector in other processes and extrapolating them to this measurement. Suppose that Bob does a lot of work and is able to determine that his detector measures the missing momentum correctly up to ± 50 GeV. If he *over-estimates* the missing momentum, then some events that otherwise wouldn't pass his selections would get bumped up artificially and mistakenly included in the analysis. By contrast, if he *under-estimates* the missing momentum, then there are events that would theoretically be predicted to pass selections but that are rejected.

   After considering the systematic uncertainties, Bob decides that he actually observed 87^{+20}_{-30} events. Note that, for systematic uncertainties (and theory uncertainties as well), the uncertainty can be asymmetric, meaning that the uncertainty interval is different above and below. This is, perhaps, because Bob is more sure that he is over-measuring rather than under-measuring. The agreement with the dark matter vs. SM hypotheses additionally depends on the theory uncertainty: if Alice is very sure that her calculation is correct, the measurement agrees more with the dark matter hypothesis, whereas if her calculation has an uncertainty of $\sim 20\%$, then the measurement is in agreement with both.

3. **Statistical Uncertainty:** Finally, we must account for random variations in the apparatus and measurement process over the course of the experiment. You are hopefully already familiar with this concept, knowing that if you are asked to measure the length of an object 20 times, you will likely get slightly different answers each time. "Random" fluctuations in temperature, air pressure, hand steadiness, etc. all account for these differences, and the idea is that doing

many trials and then averaging will give a better result than simply doing one measurement. It is important to note that in the context of a typical lab experiment, these effects are not *truly* random, but if they fluctuate sufficiently quickly, their impact on the measurement process is approximately random. Otherwise, they would be more properly classified as systematic effects.

There is one truly random aspect of nature, however, and that is the fact that we live in a quantum mechanical world. Unlike in classical mechanics, where a scattering calculation yields a definitive trajectory and outcome for each particle, quantum mechanics only gives a *probability* for each outcome. Running a collider experiment is therefore equivalent to tossing dice or flipping coins to determine the outcome of each collision, and each time you run it will yield a slightly different distribution of outcomes. In our example, the theoretical prediction of 100 events for dark matter or 50 events for the SM is only the **expected value** or mean number of events. Since each collider run will give a different number of observed events, we need to test how compatible the observed 87 event is with the statistical distribution accompanying each theoretical prediction.

For reasons that we will soon see, the statistical uncertainty typically scales like $\sqrt{N}$, where N is the number of events. Thus, the characteristic size of the statistical fluctuations for 87 events is $\sim \sqrt{87} \approx 9$ events. Thus, if the main uncertainty were statistical, the data would look to be much more in support of the DM hypothesis than the SM.

As theorists, we may be wondering why we care about anything other than the theory uncertainty. Surely this is something for the experimentalists to solve! However, if we are proposing a new way for experimentalists to look for a particle, or are investigating the precision with which an experiment can measure a cross section or other property of an interaction, we must be able to realistically assess how well we can test the theoretical predictions with the experimental results. An experimentalist may spend a year or more of their life on performing an experiment or analysis, and if we have neglected the effects of uncertainties to the extent where no conclusion can be drawn from the outcome of the experiment, then that is an enormous waste of time, effort, and money (not to mention that they will probably never listen to our advice again!). In contrast, if we can make a well-reasoned case that a particular search strategy will lead to the ability to draw *stronger* conclusions than existing methods, this will convince experimentalists to take up the case and perhaps make a discovery on our behalf!

6.2 Review of Statistics

In this section, we will take the somewhat rough qualitative idea of "uncertainty" and try to put it on a more quantitative footing. The notion of uncertainty suggests that a given experimental observation is actually one drawn from an entire **distribution** of possible outcomes, and we wish to describe the probability that a certain outcome originates from a given distribution. Ultimately, we would also like to go in reverse: to use a series of observations to figure out what underlying distribution they came from. The natural mathematical framework for doing all of this is statistics, and so we review a few of the most important concepts here.

We define a **random variable**, X, as some quantity that takes on different outcomes according to chance. For example, X could be the result of a coin toss or the roll of a pair of dice. X can be thought of as a set of possible outcomes, and the members of the set are $x \in X$. Let's assume for now that the members x of the set X take on discrete values such as integers. We define the **probability density/mass function**[10], $f(x)$, as the probability of obtaining each outcome x. For

[10]Strictly speaking, density is used to describe continuous random variables and mass is used to describe discrete

example, if X describes the outcome of a fair coin toss, then $f(\text{heads}) = f(\text{tails}) = 1/2$, while if X describes the outcome of the roll of a pair of fair dice, then $f(8) = 5/36$. Because the probability of *any* outcome should be one every time a random variable is drawn, the probability density function (PDF) must satisfy the **normalization condition**,

$$\sum_{x \in X} f(x) = 1. \tag{101}$$

The generalization to real numbers is fairly straightforward and involves interpreting $f(x)dx$ as the probability of getting an outcome between x and $x + dx$, and replacing all sums with integrals in the equations to follow.

The **mean** of the distribution is simply the sum over all possible outcomes weighted by their probability,

$$\langle X \rangle \;\; = \;\; \sum_{x \in X} x\, f(x). \tag{102}$$

We often use angle brackets around a quantity to denote its mean. The mean characterizes the average outcome of the distribution, but gives no sense of its spread. For example, if I find that I am teaching a class where the mean grade is 75%, I have no idea whether all of the students are around 75% or if half are at 100% and the other half at 50% (and the outcomes are obviously very different for each of these!). Note also that the mean might not correspond to an actual possible outcome. Let's say I take a poll of my class to find out how many cups of coffee each person has consumed that day. Each student will report an integer, but the mean might be a rational number like 1.5.

We now introduce the concept of **variance**, $\mathcal{V}$, which is related to the average deviation of each random outcome from the mean. Because this deviation can be positive or negative, and we do not wish for the positive and negative terms to cancel in the sum, we square the average deviation so that the variance is positive semi-definite:

$$\begin{aligned}
\mathcal{V} \;\; &\equiv \;\; \langle (X - \langle X \rangle)^2 \rangle & (103)\\
&= \;\; \sum_{x \in X} (x - \langle X \rangle)^2 \; f(x) & (104)\\
&= \;\; \langle X^2 \rangle - \langle X \rangle^2 & (105)\\
&= \;\; \sum_{x \in X} x^2\, f(x) - \left(\sum_{x \in X} x\, f(x) \right)^2, & (106)
\end{aligned}$$

where we have given several equivalent expressions for the variance. It is easiest to calculate this recursively (*i.e.,* first calculate the mean, then the variance). Because $\mathcal{V}$ has dimensions of X^2 and we may wish to compare the spread of outcomes with the mean, we define the **standard deviation**, σ, as the square root of the variance,

$$\sigma = \sqrt{\mathcal{V}} = \sqrt{\langle X^2 \rangle - \langle X \rangle^2}. \tag{107}$$

Often, we wish to find the probability that X falls in some range $x_1 \leq X \leq x_2$. This is found by summing

$$P(x_1 \leq X \leq x_2) = \sum_{x_1 \leq x \leq x_2} f(x). \tag{108}$$

variables. Given the other uses of "mass" in particle physics, however, you may see it referred to as density in both cases.

One particular value, known as the **cumulative distribution function**, is a function $\mathcal{C}(x)$ defined by

$$\mathcal{C}(x) \equiv P(X \leq x) = \sum_{y \leq x} f(y). \tag{109}$$

It is the probability of finding an outcome of the random variable that is *less than or equal to* x. Because probabilities are positive semi-definite and the PDF is normalized to one, then $\mathcal{C}(x)$ is a function that rises monotonically from zero to one.

A simple example of a probability distribution is the binomial distribution. This is a distribution which describes the outcome of N independent trials that have only *two* possible outcomes: "success" (1) and "failure" (0). This is the case if we take, for example, N consecutive coin tosses, where "success" is defined as finding heads. If the probability of tossing heads is p, then the binomial probability distribution for x successes is

$$f(x) \quad = \quad \frac{N!}{x!(N-x)!}\, p^x\, (1-p)^{N-x}. \tag{110}$$

Note that this equation only makes sense for $0 \leq x \leq N$, on other words the number of successes is bounded by 0 and the total number of trials.

Exercise: Show that the mean of the binomial distribution is Np, and the variance is $Np(1-p)$.

Another important probability distribution is the normal distribution. Unlike the previous examples, the normal distribution is the probability distribution of a *continuous* random variable X that can take on any real number value. The normal distribution has two parameters, μ and σ, and the PDF is

$$f(x) \quad = \quad \frac{1}{\sqrt{2\pi\sigma^2}}\, e^{-(x-\mu)^2/2\sigma^2}. \tag{111}$$

Because the variable is continuous, the PDF cannot be interpreted as the probability of finding the random variable "at" x because the measure for finding the variable exactly at x is zero. Instead, we consider an infinitesimal region dx around x and the probability for finding x in this range is $P(x \leq X \leq x + dx) = f(x)dx$. For a finite region, $x_1 \leq X \leq x_2$, the probability is found by integrating over this region

$$P(x_1 \leq X \leq x_2) = \int_{x_1}^{x_2} dx\, f(x). \tag{112}$$

The integral of the normal distribution is called the error function and has no analytic closed form expression; however, the values of these integrals are tabulated. The PDF is, however, normalized,

$$\int_{-\infty}^{\infty} dx\, \frac{1}{\sqrt{2\pi\sigma^2}}\, e^{-(x-\mu)^2/2\sigma^2} = 1. \tag{113}$$

The mean and variance are found as before, replacing the sums with integrals. The mean is calculated to be μ and the variance to be σ^2. The standard deviation of the normal distribution is simply σ.

The normal distribution has several nice properties. It is symmetric about the mean, and the PDF has a simple exponential form that scales simply with the number of standard deviations x is away from the mean. There is also the property that summing together a large number of

independent random variables tends to follow a normal distribution, regardless of what distribution those individual random variables follow. This is known as the central limit theorem, and it explains why normal distributions are ubiquitous in describing physical processes.

The probability of events that are normally distributed can be related to the number of standard deviations away from the mean:

$$\begin{aligned} P(X > \mu) &= 0.5, && (114)\\ P(X > \mu + \sigma) &= 0.159, && (115)\\ P(X > \mu + 2\sigma) &= 0.0228, && (116)\\ P(X > \mu + 3\sigma) &= 1.35 \times 10^{-3}, && (117)\\ P(X > \mu + 5\sigma) &= 2.87 \times 10^{-7}. && (118) \end{aligned}$$

What these numbers mean is that, if we *assume* that the normal distribution correctly models our system, then seeing a result that is 5σ larger than the mean in an experiment would only be expected to occur about one in 3 million experiments. This is very rare indeed! If we see such a large deviation, then it is strong evidence that our assumed normal distribution is *not* correct.

In particle physics, if we have distribution of an observable that is calculated from the SM and then we observe something that is so much larger than the expected value that it is only expected to occur every one in 3 million experiments, then we say that this is a "5-sigma deviation" from the SM expectation. This is typically the threshold for saying we have discovered a new particle or phenomenon, although whether people believe such a statistical excess depends on their confidence in how well the experiment was carried out and how well they estimated the errors.

6.3 The Poisson Distribution

In many particle physics experiments, we take very large data samples but wish to study comparatively rare processes. For example, the most important physics goal of the LHC was the discovery of the Higgs boson, and this was realized in July 2012. As already mentioned earlier, LHC collisions occur at the 30 MHz frequency, whereas the trigger can only record events at about 1 kHz and the rate of Higgs boson production is about 0.5 Hz. We wish to perform statistical analyses on these rare event rates to compare to theoretical predictions.

In a particle collider, collisions occur at well-separated points in space and time, so the outcome of one collision should not affect the result of another. This suggests that the number of Higgs boson events recorded to tape could be modelled by a binomial distribution where "success" counts as making a Higgs boson and "failure" counts as not having made a Higgs boson. Consider a time interval of one second. The number of trials would be $N = 3\times10^7$, while the probability of successfully finding a Higgs boson is about $p = 1.67 \times 10^{-8}$. The probability of finding x Higgs bosons during that one second of data is then

$$P(x) = \frac{(3 \times 10^7)!}{x!(3 \times 10^7 - x)!} \left(1.67 \times 10^{-8}\right)^x \left(1 - 1.67 \times 10^{-8}\right)^{3\times10^7 - x}. \quad (119)$$

I encourage you to go into your favourite numerical mathematics program, enter this for $x = 1$ or 2 and see what happens. It quickly becomes clear that pursuing this method of calculation will lead to a dead-end from practical limitations. For those who have studied statistics, you may wonder whether it is helpful to use the normal approximation to the binomial distribution, which is valid for $N \gg 1$. Our scenario clearly satisfies this condition, although we quickly run into another problem: the normal distribution is for a *continuous* random variable that takes on all real values, whereas there is no way to measure 0.7 or $\sqrt{2}$ Higgs bosons.

The solution comes in the form of the **Poisson distribution**, which is a distribution describing the outcomes of experiments that take on only integer values. The Poisson distribution is applicable in the limit where there are a very large number of trials (in our case, collisions), but the probability of success *per trial* is fixed and very small (in our case, the rate of Higgs boson production per collision). The outcomes of trials must, once again, be independent such that earlier results do not change the probabilities of subsequent measurements.

For Poisson-distributed processes, we don't care about the huge number of trials taken, but rather the relatively small number of these trials that yield an interesting result. We can therefore exchange the number of trials, N, for the expected number of *successful* trials, which we define to be a constant called λ:

$$\lambda = Np. \tag{120}$$

In the Poisson limit, $N \to \infty$ and $p \to 0$, while the product λ is held fixed[11]. We then have from the binomial distribution that

$$\begin{aligned} P(x) &= \frac{N!}{x!(N-x)!}\left(\frac{\lambda}{N}\right)^x\left(1-\frac{\lambda}{N}\right)^{N-x}, && (121)\\ &= \frac{N(N-1)\cdots(N-x+1)}{x!}\left(\frac{\lambda}{N}\right)^x\left(1-\frac{\lambda}{N}\right)^{N-x}, && (122)\\ &= \frac{N^x+\mathcal{O}(N^{x-1})}{x!}\frac{\lambda^x}{N^x}\left(1-\frac{\lambda}{N}\right)^{N}\left(1-\frac{\lambda}{N}\right)^{-x} && (123)\\ &= \frac{\lambda^x}{x!}\left(1-\frac{\lambda}{N}\right)^{N}\left(1-\frac{\lambda}{N}\right)^{-x}. && (124) \end{aligned}$$

Recall that

$$e^B \equiv \lim_{A\to\infty}\left(1+\frac{B}{A}\right)^A. \tag{125}$$

Taking $A \to \infty$ therefore gives

$$P(x;\lambda) = \frac{\lambda^x\, e^{-\lambda}}{x!}, \tag{126}$$

which is the PDF for the Poisson distribution. It is evident that $P(x;\lambda)$ is only valid for non-negative integer values of x.

Exercise: Check that the Poisson distribution, Eq. (126), is correctly normalized.

The mean and variance of the Poisson distribution can be calculated directly from Eq. (126). We can also use the fact that this is a limiting case of the binomial distribution, which gives $\langle X\rangle = Np = \lambda$ and $\mathcal{V} = Npq = \lambda(1-p) \approx \lambda$. The mean and variance are therefore the same for a Poisson distribution!

The standard deviation of a measurement of a Poisson process is $\sigma = \sqrt{\lambda}$. If we measure n Higgs boson events, a reasonable estimate for the parameter λ in the Poisson distribution is to take

[11] If this makes you feel uncomfortable, define some new quantity $M \equiv 1/N$ where $M \to 0$ as $N \to \infty$. This allows us to write $\lambda = p/M$, and the fact that λ is a non-zero constant is simply a standard result from calculus: it is possible to take a ratio of two very small quantities that is itself finite. A derivative is a classic example.

$\lambda = n$, in which case the uncertainty on the measurement would be approximately $\sigma = \sqrt{n}$. The relative uncertainty on the measurement is estimated to be $\sigma/\mu = 1/\sqrt{n}$, and we therefore see that collecting more events will allow us to reduce the relative statistical uncertainty on our measurement. For a process whose most important uncertainty is on the number of events observed due to random fluctuations, it is best to collect as much data as possible such that the expected number of events, λ, is very large. We have now also justified our conjecture in the above section on statistical uncertainty that the uncertainty scales as the square root of the number of events.

An important property of the Poisson distribution is that if I have two Poisson-distributed random variables, X and Y, with mean values λ_X and λ_Y respectively, then the sum $Z = X + Y$ is also Poisson-distributed with mean $\lambda_Z = \lambda_X + \lambda_Y$. We can think of it intuitively this way: suppose we divide our particle collisions into two runs, Run X and Run Y. The rate of Higgs boson production in each is simply determined by the number of collisions in each run, namely λ_X and λ_Y, and by the above arguments we expect them to each follow a Poisson distribution. However, if we don't care about whether the Higgs bosons were produced in any particular run and only want to know the total number, we could ignore the subdivision into runs X and Y and simply take the aggregate luminosity in $X + Y$. I would once again expect this to follow a Poisson distribution, but one with a larger expected number of Higgs bosons given by $\lambda_X + \lambda_Y$.

Exercise: For a random variable $Z = X + Y$, where X and Y are two independent Poisson random variables, use

$$P(z) = \sum_{x=0}^{z} P(x; \lambda_X) P(z - x; \lambda_Y), \tag{127}$$

as well as the binomial expansion,

$$(a + b)^n = \sum_{x=0}^{n} \frac{n!}{x!(n-x)!} a^x b^{n-x}, \tag{128}$$

to show that $P(z) = P(z; \lambda_X + \lambda_Y)$.

One of the powerful aspects of this result is that we can split data into smaller chunks or combine it into larger chunks and each piece obeys the Poisson distribution with the parameter λ adjusted according to the new expected number of events. If we have an experiment with $\lambda \gg 1$, then we can take $P(x; \lambda)$ and treat it as coming from the sum of λ *sub-experiments* which each have a mean of 1. We then use the central limit theorem, which states that the sum of a large number of independent random trials follows a *normal distribution* with mean $\mu = \lambda$ and standard deviation $\sigma = \sqrt{\lambda}$[12]. The corresponding normal distribution is extremely simple to use because of the tabulated probabilities of finding outcomes a certain number of standard deviations away from the mean as in Eq. (114). It is important to note that the normal approximation only makes sense for $\lambda > 30$ or so, since for smaller numbers of events the discreteness of the Poisson distribution is crucial to getting the correct probability, and the continuous normal distribution cannot be used as any kind of substitute.

[12] If you aren't familiar with the central limit theorem, one can also show that the Poisson distribution tends to a normal distribution for large λ using Sterling's approximation, $x! = \sqrt{2\pi x}\, x^x e^{-x}$ for $x \gg 1$.