LECTURE ONE:

The LHC accelerator, and the ATLAS and CMS detectors

The Large Hadron Collider is the largest and most complex scientific project ever attempted. LHC discoveries will likely dominate the development of particle physics for the rest of your professional careers. These lectures are a concise introduction to collider physics at the LHC. I will begin by discussing basic facts about the LHC accelerator and the two general purpose detectors ATLAS and CMS. This will make clear many basic features of LHC experiments, which in turn determine how (or if) discoveries can be made.

Let's begin with a few impressive bullets:

- Last year alone CERN spent about 1 billion euros on LHC, which is about 4 times the total budget of Fermilab.
- The 26.7 km LHC tunnel will soon contain more than 1200 superconducting dipole bending magnets. Each magnet is 14 meters long and will produce 8.33 Tesla dipole fields. This is nearly twice the strength of the Fermilab Tevatron dipoles.
- The 31,000 tons of superconducting dipoles need to be operated at 1.9° K, despite the inclusion of cables carrying up to 15,000 amps of current. This requires the largest cryosystem ever built, with 90 tons of liquid helium and over a thousand tons of liquid nitrogen.

To understand LHC physics, you need to understand some of the technical challenges of operating both the LHC accelerator and the five LHC detectors. Time constraints will only allow me to discuss the two general purpose detectors ATLAS and CMS, but I encourage you to explore the websites of ALICE (a heavy ion experiment), LHCb (a hadronic B physics experiment) and TOTEM (a forward detector for diffractive, elastic and total cross section measurements).

The LHC accelerator

The LHC is actually two accelerators: twin synchrotrons that accelerate beams of protons to 7 TeV in opposite directions around the 27 km underground ring. Once accelerated, the beams are brought into collison in four interaction regions, producing proton-proton collisions with 14 TeV total center-of-mass (CM) energy. Thus the LHC is a proton *collider*, as opposed to the single beam accelerators popular in the 1950-70s which directed proton beams into a *fixed target*. The advantages of using particle colliders are enormous, as you will verify in your homework.

To get an idea of how far accelerator–based particle physics has progressed, consult the "Livingston Plot" shown in Figure 1. This shows the world's major accelerators according to the year they were completed versus their fixed-target equivalent energy in eV. At the lower left are the two classic breakthroughs of the 1930s: the Cockcroft-Walton machine, which was basically a 600,000 volt DC battery, and the cyclotron of Lawrence and Livingston. At the upper right are the Fermilab Tevatron collider, which currently collides protons and antiprotons at 2 TeV CM energy, and the 14 TeV LHC. Note that the effective energy improvement shown in this plot is more than 11 orders of magnitude; this is about the same as the increase in computing speed over this same period, from the mechanical IBM machines of the 1930s to the Teraflop computers of today.

All conventional accelerators are based on linear acceleration with high power high frequency electromagnetic pulses. These radio frequency (MHz to GHz) "buckets" contain and accelerate bunches of charged particles. A cyclotron is the simplest circular accelerator, in which RF acceleration is combined with dipole bending magnets. The approximately constant vertically oriented dipole fields cause the charged particles to move in an approximately circular path.

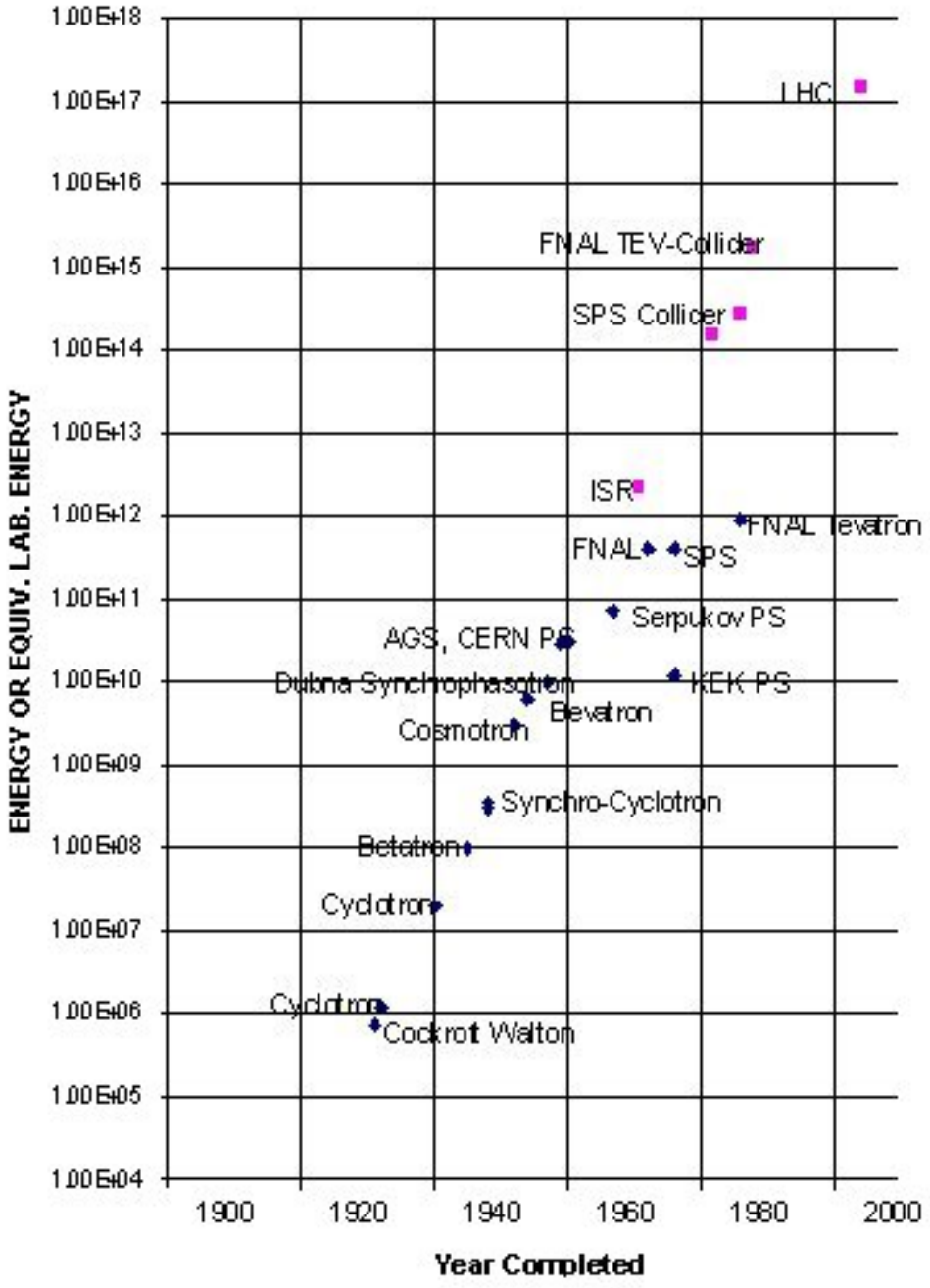

Figure 1: Livingston plot of particle accelerators.

This has the advantage that a small number of RF accelerating cavities can be used to accelerate the same bunches over and over again. A synchrotron is basically a cyclotron with the added feature that the magnet strength is ramped up as the circulating particles are accelerated; this makes it possible to keep the bunches on a path of constant radius, which allows the aperture of the high field magnets to be of managable size. In principle a synchrotron could ramp up from zero field, but in practice synchrotrons require that the bunches be pre-accelerated somewhere else, then injected into the synchrotron (itself a tricky feat) at some reasonably high energy.

For the LHC, a complex of older CERN accelerators will pre-accelerate proton bunches from rest (a bottle of hydrogen) to 450 GeV per proton. For each "store", 2808 bunches will then be injected into each of the twin LHC synchrotrons, which will accelerate them to 7 TeV per proton. The two beams will then be focused and brought into collision at the four interaction points. Although there will be 100 billion protons per bunch, most of them will pass by each other without scattering. Thus collisions of the same bunches can continue for many hours.

Ironically the power used to accelerate the proton bunches is not significant: only 256 kilowatts of RF power is needed to accelerate each beam, which is hardly more than the 140 kW of refrigeration power in the cryosystem. The accelerating protons lose energy by emitting synchrotron radiation (which has to be absorbed by the cryosystem), but this only amounts to 3.6 kW per beam; in your homework you will see that this would be a very different story if we tried to operate the LHC with electron beams.

In other respects the power and energy figures for the LHC are very impressive. For example, let us calculate the total energy of each LHC beam:

$$2808 \text{ bunches } \times 1.15 \times 10^{11} \text{ protons per bunch } \times 7 \text{ TeV} \quad (1)$$

$$= 0.185\, M_{\text{Planck}} = 362 \text{ Megajoules} . \quad (2)$$

Here I have expressed the result both in string units and in MKS units. Now consider what would happen if there were a beam accident, *e.g.*, the beam was accidently steered into a focusing magnet near the interaction point. During the brief time that it would take for all 2808 bunches to reach the scene of the accident, a spot about 1/10 of a millimeter wide on the unfortunate magnet would be subjected to 4 Terawatts of proton beam power. The result would be rather similar to detonating 86 kg of TNT (1 kg of TNT = 4.2 Megajoules). As we will see, the interaction points are surrounded by sensitive detector components which took about a decade to build and cost upwards of 500 billion euros in the case of ATLAS and CMS. Thus the acceptable number of beam accidents at the LHC is zero. By comparison, during the startup of Run II of the Fermilab Tevatron, one serious beam accident did occur, resulting in damage to a detector that, scaled up to ATLAS or CMS at the LHC, would indeed be catastrophic.

The reason why I have gone through this exercise (other than to scare you and the CERN management) is to explain why the LHC startup will be slow and gradual. Although the accelerator should be completed and ready to turn on by September of this year, the first "physics run" is not scheduled until July of 2008. It is not likely that the LHC will operate routinely to its full design specifications until 2010.

A related issue is the luminosity and integrated luminosity of the LHC collider. The luminosity of a collider is the number of collisions per second per interaction region, divided by the total cross section of protons to collide. Cross section has units of area, so luminosity can be expressed in units of $cm^{-2}\ sec^{-1}$. A simple formula for the luminosity is

$$\mathcal{L} = \frac{n_b \cdot N_L \cdot N_R \cdot f_{\rm rev}}{A_T^{\rm eff}} \ , \tag{3}$$

where $n_b = 2808 =$ number of proton bunches, $N_L = N_R = 1.15 \times 10^{11}$ is the number of protons per bunch in each direction, and $f_{\rm rev}$ is the frequency of revolution around the ring:

$$f_{\rm rev} = \frac{c}{27\ {\rm km}} \simeq 10^4\ {\rm Hz} \ . \tag{4}$$

Here I have also used the approximation that the protons are traveling at the speed of light, although in fact they are traveling 9.7 km/hour slower than the speed of light. The remaining parameter $A_T^{\rm eff}$ is just the effective transverse area of the proton beam ("effective" because the beam profile doesn't have a sharp edge); the design spec for this is given by:

$$A_T^{\rm eff} = 4\pi\sigma_b^2, \quad \sigma_b = 16\ {\rm microns} \ . \tag{5}$$

Plugging all this in we get the design luminosity of the LHC:

$$\mathcal{L} = \frac{2808 \cdot 10^{11} \cdot 10^{11} \cdot 10^4\ sec^{-1}}{4\pi \cdot 2.6 \cdot 10^{-6}\ cm^2} \simeq 10^{34}\ cm^{-2}\ sec^{-1} \ . \tag{6}$$

Now let's estimate the total cross section, which multiplied by $\mathcal{L}$ will then give us the event rate for 14 TeV proton-proton collisions. Long range inelastic strong interactions are modeled pretty well by meson exchange (as in the original Yukawa theory!). The Compton wavelength of the lightest meson, the pion, is about 1/140 MeV. Here I am starting to use particle physics units:

$$\frac{\hbar c}{1 \text{ GeV}} \simeq 2 \times 10^{-14}\ cm = 0.2 \text{ fermi} \simeq \text{size of proton} . \tag{7}$$

Since the strength of the strong interactions is about 1 in natural units, a guesstimate of the total inelastic cross section is:

$$\sigma \sim \pi \cdot (0.2 \text{ fermi} \cdot 938/140)^2 \tag{8}$$

$$\sim 6 \cdot 10^{-26}\ cm^2 = 60 \text{ millibarns} , \tag{9}$$

where I used the standard unit of cross section: 1 barn $= 10^{-24}\ cm^2$. Believe it or not this crude estimate agrees very well with data (including the implicit fact that the total cross section is roughly constant with energy at very high energies).

Rounding up, we now have an estimate of the number of collisions per second at each LHC interaction region:

$$\mathcal{L} \cdot \sigma \sim 10^{34}\ cm^{-2}\ sec^{-1} \cdot 10^{-25}\ cm^2 = 10^9 \text{ Hz} . \tag{10}$$

This result is disturbingly high. How can our experimental colleagues hope to keep detect, record and analyze all of these LHC collisions, when they are occuring a billion times per second? The short answer is that they can't; this critical issue will reappear in more detail later, when I discuss data streams and triggers.

Another important fact is that the proton-proton collision rate that we just calculated is not the same thing as the rate that proton bunches pass through each other in a given interaction region. This rate is given by:

$$\frac{2808 \cdot c}{27000} \simeq 3 \cdot 10^7 \text{ Hz} . \tag{11}$$

Actually the bunch crossing rate at LHC design luminosity will be a little higher than this, about 40 Megahertz, because the bunches are themselves bunched together to make some gaps for aborting and injecting the beams. Comparing this result with the previous one, we see that each bunch crossing will produce, simultaneously, 20–30 proton-proton scatterings. This raises another daunting problem for experimenters, since it means that each recorded "event" will actually consist of roughly 25 different 14 TeV collisions.

Fortunately a typical hard scattering of two protons has a large impact parameter and produces relatively low momentum particles in the final state. These are called "minimum bias" collisions. An event with one spectacular hard scattering is unlikely to contain another one. Less fortunate is the fact that minbias collisions involve poorly understood nonperturbative QCD (not much better than we did above), so to subtract the minbias pollution of interesting events, we will have to model minbias physics from actual LHC data. This was done successfully with Tevatron data by Rick Field (a theorist!). Although minbias collisions are "soft" (*i.e.* they produce relatively low momentum particles), their effects are quite significant. A typical LHC event will

contain a total of about 1 TeV of "soft" energy from minbias collisions; by comparison, the most energetic *hard* scattering ever recorded prior to LHC was a 1.36 TeV collision at the Tevatron.

Another problem raised by our simple calculation is that LHC events will occur every 25 nanoseconds. In 25 nanoseconds, light travels only 7.5 meters; as we will see, this is less than the physical size of the ATLAS and CMS detectors. Thus no electronic system could possibly read out data from an entire detector, much less do anything with it, before the next event is upon it. This means that at least some detector components must read out data integrated over more than one bunch crossing. This increased "pile-up" of minbias collisions in the data stream causes additional pollution of interesting events.

Of course, both of the above problems could be eliminated by running the LHC at lower luminosity. How much luminosity do we need? What really matters is the rate for "interesting" events (*e.g.* production of Higgs bosons, superpartners, etc.). To compute this, we need the concept of *integrated luminosity*: $\int \mathcal{L} dt$. The integrated luminosity over some period of time, times the cross section for a given type of interesting event, times the efficiency of your detector to record such events, estimates the number of those interesting events that will be "written to tape". Integrated luminosity has units of cm^{-2}, but a more convenient unit is the inverse femtobarn:

$$1\ fb^{-1} = \frac{1}{10^{-15} \cdot 1\ \text{barn}} = 10^{39}\ cm^{-2}\ . \tag{12}$$

Let's compute how much integrated luminosity the LHC will accumulate per year when running at design luminosity. One year $= \pi \times 10^7$ seconds, however accelerators do not produce collisions all of the time, and the luminosity of a store degrades with time. A good rule of thumb is to use the "Snowmass year" $= 10^7$ seconds. So $\mathcal{L} \cdot 10^7$ seconds $= 10^{41}\ cm^{-2} = 100 fb^{-1}$. By comparison, the total integrated luminosity produced at the Fermilab Tevatron after ten years running is slightly over 2 fb^{-1}.

Now let's look at the cross sections for some interesting processes, as computed for pp collisions at 14 TeV:

Process	*c*ross section in fb
a pair of 500 GeV jets	10,000
SUSY with 1 TeV superpartners	3,000
light Higgs boson	2,500
heavy Higgs boson	1,000
3 TeV Z' decaying to muons	2

Although we have not yet included detector efficiencies (which range from less than a percent to more than 50% depending on the process), these numbers are quite encouraging. Multiplied by 100 fb^{-1} per year, they indicate that thousands or tens of thousands of new physics events could be detected per year.

As already mentioned, it is not expected that the LHC will reach design luminosity until 2010, and even then the integrated luminosity projection is 40 fb^{-1} per year, not 100. For the first physics run in 2008, the estimated integrated luminosity is about 1 fb^{-1}. To understand the discovery potential with this first golden sample of 14 TeV events, we need to learn more about detectors and about Standard Model backgrounds.

What do collider detectors detect?

They do not detect W or Z bosons directly, since these decay on a time scale of about 3×10^{-25} seconds. They do not detect quarks or gluons directly, since these hadronize on a time scale of a few times $\hbar/\Lambda_{\rm QCD} \sim 2 \times 10^{-24}$ seconds. In fact a high energy quark or gluon can produce hundreds of hadrons, which form a *jet* of particles moving approximately in the same direction as their parent parton. Here is a simple summary of what collider detectors actually detect:

• **Electrons:** An electron moving through material loses energy by ionizing atoms; the additional electrons thus produced can be collected in such a way that the position of the high energy particle is determined. Thus an electron leaves a *track* in a device generically known as a "tracker". If a strong magnetic field is in place, pointing along the direction of the beampipe, an electron moving away from the beam will curve in the plane orthogonal to the beam:

$$\frac{d\vec{p}}{dt} = -e\frac{\vec{v}}{c} \times \vec{B} \, , \tag{13}$$

so *e.g.* for an electron whose initial transverse motion is along the x-axis:

$$\vec{x}(t) = -R\left(\hat{x}\sin\omega_B t + \hat{y}\left(1 - \cos\omega_B t\right)\right) \, , \tag{14}$$

where the bending radius R is related to the precession frequency ω_B by: $R = c/\omega_B = pc/eB$. Measuring this curved track thus tells us the momentum of the electron, and whether it is an electron or a positron. For electron energies above ~ 10 MeV, bremsstrahlung becomes the dominant energy loss mechanism, while photons thus produced convert back to e^+e^- pairs due to the presence of the material. This cascades into an electromagnetic shower. Such showers are initiated, contained, and detected in devices called *calorimeters*, which obviously have to be positioned at a larger radius from the interaction point than the tracker. Calorimeters measure the total energy of the initial high energy electron.

• **Photons:** For a collider detector, a high energy photon is like an electron without a track. If the calorimeter is finely segmented, we can make up for this by a position measurement from the calorimeter.

• **Hadrons:** High energy pions and kaons live long enough to initiate showers in calorimeters, as do protons and neutrons. These hadronic showers can penetrate farther (and are broader) than electromagnetic showers, thus collider detectors have larger hadronic calorimeters surrounding their electromagnetic calorimeters. Charged hadrons will leave tracks in the tracker. Electrons are discriminated from hadrons by the fact that their showers are typically at least 90% contained in the electromagnetic calorimeters (which obviously are designed such that this will be true!). Photons are discriminated from neutral hadrons in the same way, although high energy neutral pions present a tough case, since they look like two roughly collinear photons.

• **Muons:** Muons have the same interactions as electrons, but because they are 200 times heavier they do not initiate electromagnetic showers. With a lifetime in the lab frame of $2.2\gamma \times 10^{-6}$ seconds, they are essentially stable particles. Thus high energy muons leave tracks in the tracker, a GeV or so of ionization loss in the two calorimeters, and continue on their merry way. For very high energy muons, the tracker may give a poor measurement of the momentum, and at any rate we no longer have an energy measurement from the calorimeters. Thus collider detectors require huge muon systems, outside their calorimeters, designed to measure the trajectory and momentum of muons.

• **Taus:** Taus can decay into a muon or electron plus two neutrinos, in which case they are detected as muons or electrons. Taus also have hadronic decays, which produce either one or three charged tracks. These "hadronic taus" can be somewhat discriminated (on a statistical basis) from hadronic jets initiated by quarks and leptons.

• **Heavy flavor:** The mean distance traveled by a B meson before decaying is 0.5γ millimeters, while it is 0.3γ millimeters for a D meson. With fine tracking devices sufficiently close to the beampipe, it is possible to distinguish the displaced secondary vertices of the charged particles from these decays. In this sense detectors can "see" heavy flavor mesons; this ability is known as b-tagging and charm-tagging. An alternative technique for heavy flavor tagging is to look for muons inside the hadronic jet: semi-leptonic decays of Bs and Ds produce muons, while jets initiated by light quarks or gluons rarely contain muons.

• **Missing energy:** Although neutrinos interact too weakly to be seen in a collider detector, modern detectors can infer the presence of high energy neutrinos by measuring "missing transverse energy". What is actually measured is missing transverse momentum, applying momentum conservation to all of the observed products of the pp scattering. Since the momentum of the colliding pp pair is almost entirely longitudinal (along the beam) rather than transverse, the transverse momenta of the scattering products should add up to zero. An imbalance is attributed to an undetected particle (such as a neutrino). Missing energy analyses are performed with calorimetry, not tracking, hence the slight misnomer "missing transverse energy" or MET. MET is actually a vector in the plane transverse to the beam. If the MET vector points to a crack or gap in the calorimeters, the interpretation of the event is ambiguous. This is one reason why modern collider detectors are constructed to be as hermetic as possible. Note that it is not possible to do a missing energy analysis in the straightforward sense of comparing energy in with energy out. In a pp collision only one parton from each proton has a hard scattering (usually). The energy of this collision is not known - it is much less than the full pp CM energy. The remnants of the protons (the "underlying event") are poorly measured, since they mostly go off down the beam pipe.

The ATLAS and CMS detectors

From the discussion above one can roughly predict what the ATLAS and CMS detectors should look like. Moving radially outward from the interaction region, they should have a tracking system, an electromagnetic calorimeter, a hadronic calorimeter, and a muon system. There should also be one or more large magnets, to provide bending fields for the tracker and the muon system. One might imagine that the optimal shape for a hermetic detector is a sphere, but the desire for very large uniform magnetic fields dictates instead that the main detector is a cylinder or "barrel" centered on the beampipe and the interaction region, supplemented by "endcaps" which provide coverage in the "forward" regions.

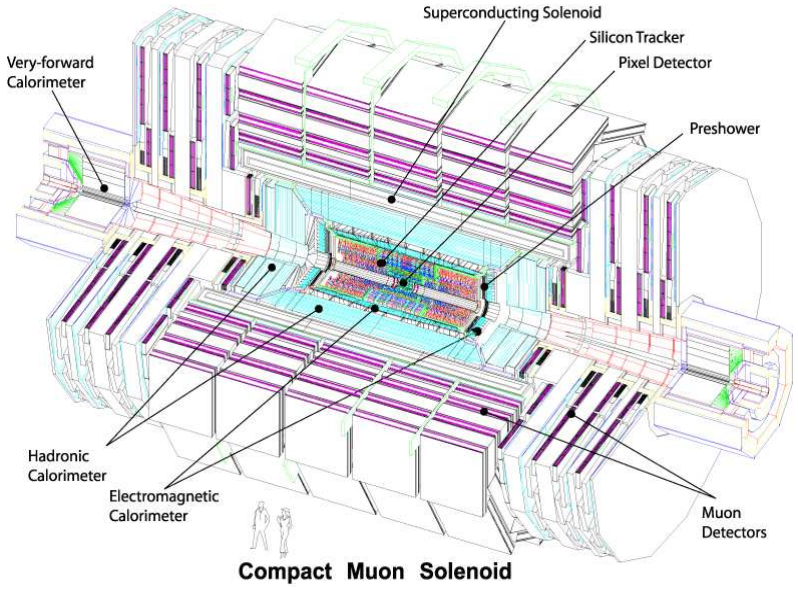

Figure 2: Schematic of the CMS detector.

The schematic of the CMS detector in Figure 2 shows all of these elements:

• **Tracker:** In CMS the tracking system consists entirely of silicon sensors. These have two great advantages. The first is that they provide

spatial resolution as good as a few microns, even for complex events with hundreds of tracks. The part of the tracker closest to the beampipe is called the pixel detector in Figure 2; it allows b-tagging by finding displaced vertices. The second advantage is that the sensors (including their readout chips) are radiation hard. They will be able to withstand the megarad radiation doses that they will be exposed to after several years of LHC operation (for comparison the lethal dose of this kind of radiation is about 100 rads).

• **Electromagnetic calorimeter:** The CMS "ECAL" consists of over 80,000 large crystals of lead-tungstate. This transparent metal (!) is an exquisite material for detecting electromagnetic showers, resolving difficult events such as a Higgs decaying to two photons.

• **Hadronic calorimeter:** The CMS "HCAL" has scintillator detection layers interleaved with dense absorption layers made of brass. The CMS HCAL contains five times more copper alloy than the Statue of Liberty.

• **Muon system:** The CMS muon system is the largest part of the detector. It consists of several different components, balancing the need to measure muons well with the need to measure them quickly (remember the 25 nanoseconds!).

• **Magnet:** CMS uses a single large long solenoid magnet to provide a uniform 4 Tesla bending field for both the tracking system and the muon system. This superconducting solenoid is, by far, the most energetic magnet ever built; in operation it carries 3 Gigajoules of stored energy.

ATLAS versus CMS

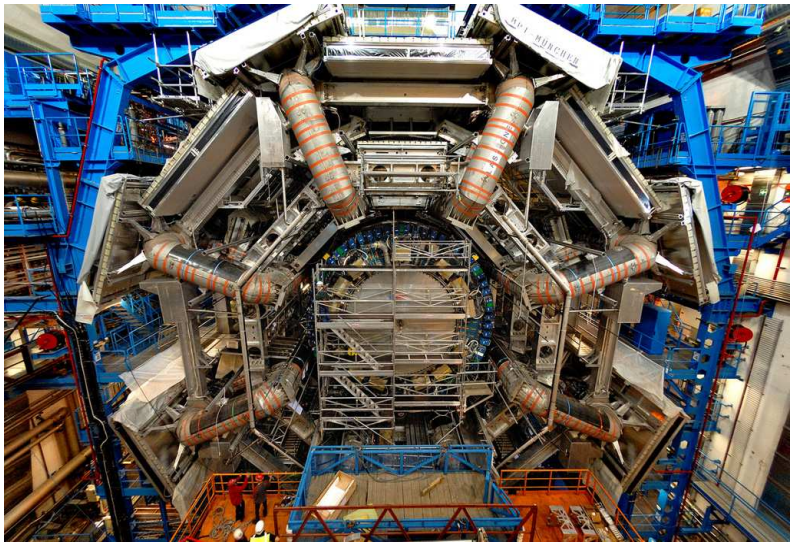

Figure 3: The ATLAS detector in October 2006.

Object	*W*eight in tons
Boeing 747 (fully loaded)	200
Endeavor space shuttle	368
ATLAS detector	7,000
Eiffel Tower	7,300
USS John McCain (destroyer warship)	8,300
CMS detector	12,500

The CMS detector has roughly the volume of a space shuttle but weighs 30 times more: 12,500 tons altogether. The ATLAS detector, shown in Figure 3, has more than 10 times the volume of CMS, but weighs only half as much. As can be seen in the figure, ATLAS has a barrel detector somewhat like that of CMS, but it is surrounded by eight huge superconducting air-core toroidal magnets. The structure is so spread out that if you put ATLAS into a plastic bag and dropped it into Lac Léman, it would float.

This raises an interesting question: since ATLAS and CMS are supposed to do the same physics, why such a radical difference between their designs? As we have seen, the muon systems are the largest parts of collider detectors, and thus their design has the largest effect on the overall detector configuration. For these general purpose LHC detectors, the muon systems were designed to meet one crucial benchmark: the ability to measure the momentum of a 1 TeV muon with no worse than 10% accuracy. As it happens, there are two equally good technical solutions to this requirement.

To see this, we first observe that there are two obvious strategies for measuring the momenta of high energy muons. The first is to use the same solenoid magnet bending field that the tracker relies on. The second is to use the fact that muons penetrate beyond the calorimeters, and build bending magnets into the muon system itself.

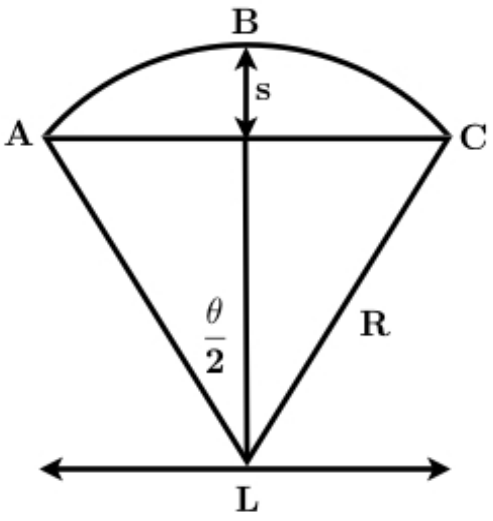

Figure 4: Measuring the sagitta of a deflected muon.

To compare these strategies, we need to know how the momentum resolution depends on the B field and on the distance L over which the bending takes place. This is easily done by consulting Figure 4, which

shows a simplified diagram of a magnetically curved trajectory which is measured at three points A, B and C over some distance L. The bending radius, as we calculated before, is given by $R = pc/eB$, which I will assume is much larger than L. Thus the deflection angle θ is approximated by:

$$\frac{\theta}{2} \simeq \sin\frac{\theta}{2} = \frac{L/2}{R} \ . \tag{15}$$

What is actually measured is the sagitta s shown in Figure 4:

$$s = R - R\cos\frac{\theta}{2} = 2R\sin^2\frac{\theta}{4} \simeq \frac{R\theta^2}{8} \simeq \frac{eBL^2}{8pc} \ . \tag{16}$$

Since the sagitta is a constant divided by the muon momentum p, the measurement errors are related by

$$\frac{|\Delta p|}{p} = \frac{|\Delta s|}{s} = \frac{8c|\Delta s|}{eBL^2} \ , \tag{17}$$

where Δs is a constant that just depends on how much money you spent on your high resolution muon tracking.

Now we can compare the strategies of CMS and ATLAS for measuring the momenta of high energy muons. CMS relies on the same solenoid magnet used by the tracker. Muons moving radially traverse a distance L of about 3 meters through the 4 Tesla bending field, giving $BL^2 = 36$ Tesla-meters^2. Outside of the solenoid, they traverse about another 3 meters, about half of that distance encountering the solenoid return flux in the iron magnet yoke, which is about a 2 Tesla field (with the opposite orientation); this gives another $BL^2 \simeq 5$ Tesla-meters^2. A combination of the tracker and muon detectors in and around the yoke measure these two bendings.

In ATLAS the tracker solenoid is 2 Tesla with a radius of only about 1 meter, giving $BL^2 \simeq 2$. Outside the calorimeters, the air-core toroidal magnets provide a bending field of about 1 Tesla over a radial distance

of about 6 meters, giving $BL^2 = 36$ Tesla-meters^2. Note that the tracker is not useless for the muon momentum measurement, since it can determine the location of the interaction point.

So indeed there are two equally good but radically different solutions to the same problem. Other differences between the ATLAS and CMS detectors follow from this choice, *e.g.* the ATLAS calorimeters can be larger since they don't have to be contained within a solenoid.

In this brief survey of ATLAS and CMS, I have omitted the essential systems which handle the enormous data streams (the raw data rates are ~ 100 Terabytes/second). I will describe these after we have learned more about what high energy collisons will look like at LHC.

References

[1] J. J. Livingood, *Principles of Cyclic Particle Accelerators*, D. Van Nostrand Company, Princeton, 1961.

[2] C. Grupen, *Particle Detectors*, Cambridge University Press, Cambridge, 1996.

[3] D. Green, *High P_T Physics at Hadron Colliders*, Cambridge University Press, Cambridge, 2005.

[4] The CMS Collaboration, *The Compact Muon Solenoid Letter of Intent*, CERN/LHCC 92-3, 1992.

[5] The ATLAS Collaboration, *The ATLAS Technical Proposal*, CERN/LHCC 94-43, 1994.

[6] The CMS Collaboration, *CMS Muon Project Technical Design Report*, CERN/LHCC 97-32, 1997.

[7] The ATLAS Collaboration, *ATLAS Muon Spectrometer Technical Design Report*, CERN/LHCC 97-22, 1997.

LECTURE TWO:

The basics of pp collisons

Electrons (as far as we know) are elementary particles, whereas protons are composite objects whose internal structure is complicated and not understood from first principles. If one were not worried about cost or luminosity, one would almost always prefer a nice clean e^+e^- collider over a messy pp collider. However as you showed in your homework it is simply not possible to build a circular e^+e^- collider with energy much greater than LEP, and we will have to wait more than decade before a next generation linear e^+e^- collider appears: the ILC. Meanwhile at LHC we must figure out how to understand messy collisions of strongly-interacting composites.

The answer to the question: "What's inside the proton?" depends on how hard your probe is, *i.e.*, the 4-momentum squared Q^2 transferred in the process. At very low Q^2 protons are probably best thought of as flux tubes - hadronic strings! - although this picture is only qualitative. For Q^2 greater than about $(2\ \mathrm{GeV})^2$, perturbative QCD becomes a good description of processes which probe the proton. Then we find that the proton contains 3 valence quarks (two u's and a d), a bunch of gluons, and "sea" quarks $u, \bar{u}, d, \bar{d}, s, \bar{s}, c, \bar{c}, b, \bar{b}$ from virtual pair production in the gluonic field. Collectively all these particles are called *partons*. Each parton carries some fraction of the longitudinal momentum and energy of a 7 TeV LHC proton, a little bit of transverse momentum p_T, and part of the angular momentum (getting the total angular momentum of the partons to add up to the spin 1/2 of the proton is still an unsolved problem). In any given pp collision, all of these quantities, as well as the identities of the partons, are unknown. So how can we predict anything about pp collisions?

deep inelastic scattering

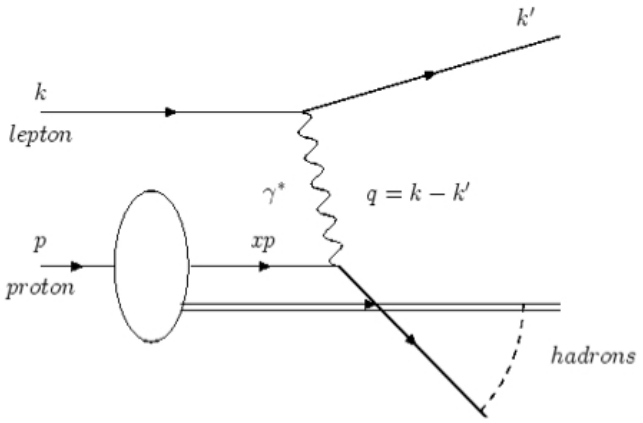

Figure 1: Schematic for deep inelastic scattering of a charged lepton with a proton.

Let's start with a simpler process, known as deep inelastic scattering (DIS), pictured in Figure 1. DIS probes the quarks (but not the gluons) inside the proton via electromagnetic interactions whose leading order vertex just involves a virtual photon coupling to a charged parton. If P_μ denotes the 4-momentum of the proton, then the 4-momentum of the initial quark can be written

$$p_\mu = x\, P_\mu \ , \quad 0 < x < 1 \ , \tag{1}$$

where we are ignoring (at leading order) the possibility that the initial quark has some p_T. The momentum fraction x is called "Bjorken x". For a hard scattering we can treat the electron as a massless particle and write the 4-momenta (k_0, k_z, k_x, k_y) of the intial and final state electrons as

$$e_i^- \ : \ (k, k, 0, 0) \ , \quad e_f^- \ : \ (k', k'\cos\theta, k'\sin\theta, 0) \ , \tag{2}$$

where without loss of generality I have taken the final state electron to scatter in the $x - z$ plane. Thus the virtual photon exchanged in the

lowest order diagram has 4-momemtum q_μ given by

$$q_\mu = (k - k', k - k'\cos\theta, -k'\sin\theta, 0) \ , \tag{3}$$

which means that the final state quark has 4-momentum $p_\mu + q_\mu$. Conventionally Q^2 of DIS is defined as $Q^2 = -q^2 > 0$.

I will ignore the rest of the proton and just try to compute this hard scattering subprocess. As always we want to use Lorentz invariant kinematic variables as much as possible. The Mandelstam invariants corresponding to the DIS subprocess are:

$$\begin{aligned} \hat{s} &= (p_\mu + k_\mu)^2 \simeq 2p \cdot k \simeq xs \ , \\ \hat{t} &= (k_\mu - k'_\mu)^2 = -Q^2 \simeq -2kk'(1 - \cos\theta) \ , \end{aligned} \tag{4}$$

where hatted variables belong to the partonic subprocess, while $s = (P_\mu + k_\mu)^2$ is the total squared 4-momentum of the whole process. In the approximations I have treated both the electron and the quark as massless. Note the invariant kinematic variables x, Q^2 carry the same information as $\hat{s}, \hat{t}$ respectively.

The total cross section for DIS can be written:

$$\sigma_{ep\to eX} = \int_0^1 dx \int d\hat{t} \sum_i f_i(x, Q^2) \, \frac{d^2\hat{\sigma}_{eq\to eq}(\hat{s}\hat{t})}{dx \, d\hat{t}} \ , \tag{5}$$

where the sum is over all valence and virtual quarks inside the proton. The differential cross section $d^2\hat{\sigma}/dxd\hat{t}$ for the subprocess hard scattering can be computed in perturbation theory.

The *parton distribution functions* $f_i(x, Q^2)$ give the relative probablity of finding a parton of type i with energy fraction x in a hard scattering characterized by Q^2. These functions represent nonperturbative information about the structure of the proton. Nobody knows how to compute them from QCD, not even lattice gauge theorists or string theorists.

Even worse, the x dependence of the "pdfs" depends on universal

features of protons, but the Q^2 dependence comes from details of the probing process. So at first glance it appears that we can say nothing about DIS in regimes where we haven't aleady measured it!

Fortunately, in a hard scattering the electron probes the proton on a time scale that is short compared to $\hbar/\Lambda_{\rm QCD} \sim 2 \times 10^{-24}$ seconds. This means that, to leading order in α_s, DIS only depends on a universal fixed distribution of independent quarks:

$$f_i(x, Q^2) \to f_i(x) \quad \text{as} \;\; Q^2 \to \infty \; . \tag{6}$$

This feature is known as Bjorken scaling, and the approximation that we are making here is the "naive parton model" of Feynman.

The tree-level matrix element for the hard scattering is

$$i\mathcal{M} = \bar{u}^{s_f}(k_f)(-ie\gamma^\mu)u^{s_i}(k_i)\left(\frac{-i\eta_{\mu\nu}}{q^2}\right)\bar{u}^{r_f}(p+q)(-ieQ_q\gamma^{nu})u^{r_i}(p)\,, \tag{7}$$

where Q_q is the electric charge of the quark, *i.e.*, either $2/3$ or $-1/3$.

Square this matrix element, sum/average or initial/final spins, and plug into the formula for the total cross section. I will change variables from $\hat{t}$ to $y = -\hat{t}/\hat{s}$, using the nice property that y, like x, runs from 0 to 1. Thus:

$$\begin{aligned} \sigma_{ep\to eX} &= \int_0^1 dx \int_0^1 dy \, \frac{d^2\sigma}{dxdy} \; , \\ \frac{d^2\sigma}{dxdy} &= \left[\sum_i x f_i(x) Q_{q_i}^2\right] \frac{2\pi\alpha^2 s}{Q^4}[1 + (1-y)^2] \; . \end{aligned} \tag{8}$$

Thus to zeroeth order in α_s we have factorized all our of ignorance of nonperturbative QCD into a single quantity, called a "structure function":

$$F_2(x) = \sum_i x f_i(x) Q_{q_i}^2 \; . \tag{9}$$

parton evolution and factorization scales

When we go to the first nontrivial order in α_s this simple picture seems to break down. At first order quarks can emit gluons. This leads to infrared divergences in QCD from the emission of soft gluons from hard partons, or from an emitted gluon becoming collinear with the parent parton. The divergences from gluon emission by final state quarks are guaranteed to cancel (by the Kinoshita-Lee-Nauenberg Theorem) as long as we only look at inclusive final states that sum over all possible hadrons (thus the notation $ep \to eX$). Any quantity which has such an insensitivity to long-distance QCD effects in the final state is called "infrared safe".

However in DIS there is an *uncancelled* infrared divergence in the *initial* state. This happens because the photon probe, which only sees quarks, distinguishes between a quark and a collinear quark+gluon pair with the same total momentum. So *e.g.* in dimensional regularization the structure function has a piece that diverges like α_s/ϵ. Alternatively, we can regulate this divergence by requiring initial state gluons to have a minimum p_T "kick" p_T^{min} where $p_T^{min} \ll \Lambda_{QCD}$. But this means that $\mathcal{O}(\alpha_s)$ results have large logarithms:

$$\log \left[Q^2/(p_T^{min})^2 \right] \ . \tag{10}$$

This is a genuine breakdown of perturbative QCD. Physically, it is coming from the fact that perturbative QCD treats initial state quarks as on-shell particles, with $p^2 \to 0$ in the massless approximation, but actual quarks inside the proton are never truely on-shell, due to confinement.

The sneaky way to recover from this failure is to introduce an arbitrary *factorization scale* μ_F, and write

$$\log \left[\frac{Q^2}{(p_T^{min})^2} \right] = \log \frac{Q^2}{\mu_F^2} + \log \frac{\mu_F^2}{(p_T^{min})^2} \ . \tag{11}$$

The second term – and thus the singularity – can be absorbed into a redefinition of the pdfs:

$$f_i(x) \to f_i(x, \mu_F^2) \; . \tag{12}$$

Obviously physics doesn't depend on the value of μ_F, so the μ_F dependence of the of the pdfs must cancel the μ_F dependence of the hard subprocess. Order by order in α_s we now define renormalized pdfs at the factorization scale which are free of initial state singularities.

Thus even though we can't compute the pdfs, once you measure the pdfs at one scale (which really means one value of Q^2), you can compute them at any other scale; this is called "Altarelli-Parisi" evolution (or "DGLAP"). As with the usual renormalization, the quantities we are computing (pdfs in this case, renormalized couplings in the usual case) had infinities which we subtracted off, the price being a renormalization scheme dependence which makes fixed-order results convention dependent. In the pdf case we know that the true underlying theory (nonperturbative QCD) gets rid of the infinities from the start, the same role played by string theory in the usual case of ultraviolet divergences.

the master equation for pp collisons

Now we know enough QCD to write down the master equation for high energy pp collisions:

$$\sigma(p_1, p_2) = \sum_{i,j} \int_0^1 dx_1 \int_0^1 dx_2 \;\; f_i(x_1, \mu_F^2) f_j(x_2, \mu_F^2) \tag{13}$$

$$\times \frac{d^2}{dx_1 dx_2} \left[\hat{\sigma}_{ij}^{\rm LO}(x_1 p_1, x_2 p_2, \alpha_s(\mu_R)) + \hat{\sigma}_{ij}^{\rm NLO-LO}(x_1 p_1, x_2 p_2, \mu_R, \mu_F) + \hat{\sigma}_{ij}^{\rm NNLO-NLO} + \ldots \right] ,$$

where p_1, p_2 are the 4-momenta of the incoming protons, x_1, x_2 are the parton momentum fractions, μ_R is the renormalization scale, μ_F is the factorization scale, and $\hat{\sigma}_{ij}$ is the cross section for the hard subprocess

involving partons of type i and j. In practice $\hat{\sigma}_{ij}$ is computed from matrix elements to some fixed order in α_s: leading order (LO), next-to-leading-order (NLO) or, if you are very strong, NNLO.

For LO matrix elements there is no dependence on the factorization scale, and but there a dependence on the renormalization scale from the running of α_s. As we saw for DIS, at LO we can also treat the pdfs in the naive parton approximation, where they have no dependence on factorization scale.

At NLO one gets a much better approximation to the physical result; NLO results are often 20% to 50% larger or smaller than LO results. NLO results also have less dependence on the renormalization scale, since at NLO the μ_R dependence of loop diagrams partially cancels the μ_R dependence in α_s. These features are illustrated in Figure 2 for the case of $W+$ jet production. On the other hand, at NLO the answer depends on the choice of factorization scale and the choice of factorization scheme.

If one ignores the NLO dependence on factorization, as if often done implicitly by setting $\mu_F = \mu_R$, then NLO total cross sections typically are nearly constant, as in Figure 2. If you give this constant to an experimenter, he/she will forget all the hard work you did and use this constant to define a "K factor": $K = \sigma^{NLO}/\sigma^{LO}$. This K factor is then used to rescale the results coming out of LO Monte Carlo event generators that simulate the process of interest for LHC.

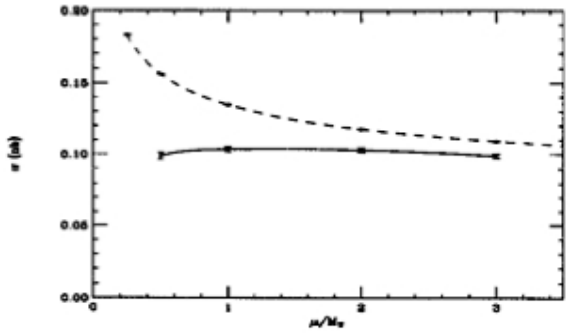

Figure 2: The renormalization scale dependence of the $W + 1$ jet cross section at the Tevatron as a function of the ratio of the renormalization scale and the W mass. The dashed line is LO, and the solid line is NLO. From W. Giele, E.W.N. Glover and D. Kosower, Fermilab-Conf-92/213-T.

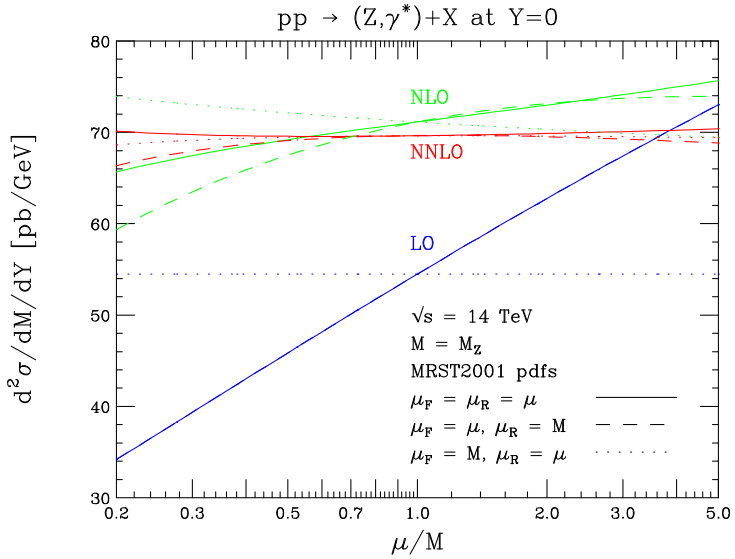

Figure 3: Renormalization and factorization scale dependence for inclusive Drell-Yan production at the LHC. The vertical axis is the doubly differential cross section where M is the lepton pair invariant mass and Y is the rapidity, which is fixed to zero in this plot. The dashed curves show the LO, NLO and NNLO dependence on the factorization scale. The solid curves show the combined renormalization and factorization scale dependence with $\mu_R = \mu_F$. From C. Anastasiou, L. Dixon, K. Melnikov and F. Petriello, hep-ph/0312266.

At NNLO (see Figure 3) one gets a result with a reduced dependence on μ_F (not to mention an even more reduced dependence on the renormalization scale). Given a NNLO result, one can pick the optimal value for μ_F such that the NLO and NNLO results agree. This is similar in spirit to the K factor method. This trick is important because NLO event generators are beginning to become available for LHC physics, whereas NNLO event generators are out of the realm of possibility for (at least) the LHC era.

Drell-Yan

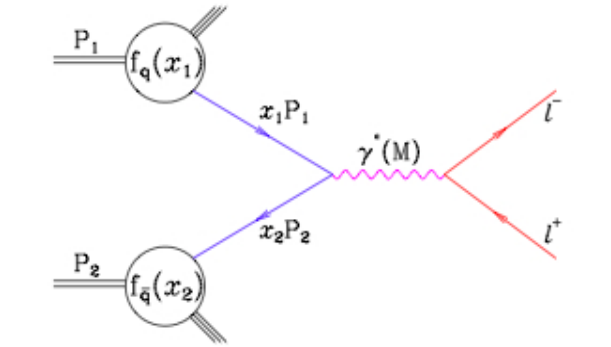

Figure 4: Schematic of Drell-Yan production.

Drell-Yan means the production of a pair of hard muons or electrons through quark-antiquark annihilation into a virtual photon or Z boson (see Figure 4). This process has a clean final state that is relatively easy to detect experimentally. Letting k_μ, k'_μ denote the 4-momenta of the leptons, the *invariant mass* squared of the dilepton pair is given by

$$M^2 = (k_\mu + k'_\mu)(k^\mu + k'^\mu) \ . \tag{14}$$

In the data, a plot of the lepton invariant mass versus number of events should show a strong peak around $M^2 = M_Z^2 = (91.19 \text{ GeV})^2$. Such a peak is indeed seen in the Tevatron data (see Figure 5).

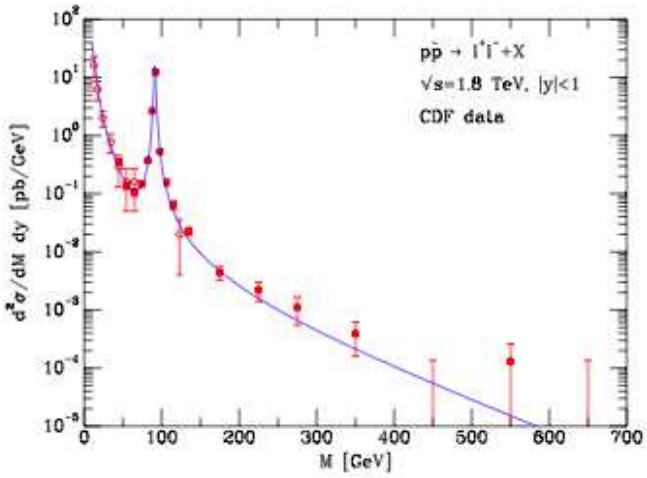

Figure 5: CDF data for Drell-Yan compared to NLO theory.

At the LHC we will have the capability to observe Drell-Yan events with dilepton invariant masses up to about 5 TeV. This is a promising discovery channel, where to first approximation one just looks for an extra peak at large invariant mass. This would likely indicate a new heavy $U(1)$ vector boson, generically called a Z'. Such Z's are one of the more generic qualitative predictions of string theory, since string theory (usually) starts with a large gauge group, while most of the gauge breaking mechanisms do not reduce the rank of the gauge group. Thus in string theory there is no reason not to expect that at least one extra $U(1)$ of the visible sector survives down to the TeV scale.

With this motivation, let's look at the Drell-Yan process in more detail. For the moment we will ignore the Z and only look at the contribution from photons. At leading order, which is α^2 in QED and zero powers of α_s, Drell-Yan is related to DIS by crossing symmetry; it is also the inverse of the process $e^+e^- \to q\bar{q}$; the latter is given by:

$$\sigma(e^+e^- \to q\bar{q}) = \frac{4\pi\alpha^2}{3s}\left(3\sum_q Q_q^2\right) , \tag{15}$$

where q labels the possible quarks in the final state with electric charge Q_q, and the factor of 3 is counting colors. This cross section is the leading order contribution to the famous infrared safe ratio R:

$$R = \frac{\sigma(e^+e^- \to \text{hadrons})}{\sigma(e^+e^- \to \mu^+\mu^-)} \ . \tag{16}$$

Of course Drell-Yan itself has infrared singularities in the initial state, just like DIS. So once we go beyond leading order in Drell-Yan we will have to introduce a factorization scale and renormalized pdfs.

For Drell-Yan the corresponding LO subprocess formula is

$$\sigma(q\bar{q} \to \ell^+\ell^-) = \frac{4\pi\alpha^2}{3\hat{s}} \left(\frac{1}{3} \sum_q Q_q^2 \right) \ , \tag{17}$$

where s has become $\hat{s}$. Note the color factor of 3 has become $1/3$, because now the quark and antiquark colors must match up to make a color singlet. Putting in the pdfs, the LO cross section is:

$$\sigma(pp \to \ell^+\ell^- + X) = \int_0^1 dx_1 dx_2 \left(\sum_q f_q(x_1) f_{\bar{q}}(x_2) \frac{1}{3} Q_q^2 \right) \frac{d^2\hat{\sigma}_{q\bar{q}\to\ell^+\ell^-}(\hat{s})}{dx_1 dx_2} \ , \tag{18}$$

where

$$\frac{d^2\hat{\sigma}_{q\bar{q}\to\ell^+\ell^-}(\hat{s})}{dx_1 dx_2} = \frac{4\pi\alpha^2}{3\hat{s}} \ , \quad \hat{s} = x_1 x_2 s \ . \tag{19}$$

This is a simple but useless result, since what we really want to predict is the invariant mass distribution of the lepton pair. To get this we need to do a little kinematics.

The 4-momenta of the initial state partons in the lab frame are:

$$\begin{aligned} p_\mu &= x_1 \frac{\sqrt{s}}{2}(1,1,0,0) \ , \\ p'_\mu &= x_2 \frac{\sqrt{s}}{2}(1,-1,0,0) \ , \end{aligned} \tag{20}$$

where I have used the leading order approximation that the parton

momenta are purely longitudinal. The lab frame is the center-of-mass frame for the pp collision (literally by construction!) but it is not the center-of-mass frame for the partonic subprocess. Let's do a longitudinal Lorentz boost to the subprocess CM frame. For the two partons:

$$\begin{pmatrix} E^{CM} \\ p_z^{CM} \end{pmatrix} = \begin{pmatrix} \cosh Y & -\sinh Y \\ -\sinh Y & \cosh Y \end{pmatrix} \begin{pmatrix} x_1 \frac{\sqrt{s}}{2} \\ x_1 \frac{\sqrt{s}}{2} \end{pmatrix} ,$$
$$\begin{pmatrix} E^{CM} \\ -p_z^{CM} \end{pmatrix} = \begin{pmatrix} \cosh Y & -\sinh Y \\ -\sinh Y & \cosh Y \end{pmatrix} \begin{pmatrix} x_2 \frac{\sqrt{s}}{2} \\ -x_2 \frac{\sqrt{s}}{2} \end{pmatrix} . \qquad (21)$$

But $E^{CM} = p_z^{CM} = \sqrt{\hat{s}}/2$, so the above tells us how to write the boost parameter Y as a function of x_1 and x_2:

$$Y = \frac{1}{2} \ln \left(\frac{x_1}{x_2} \right) , \quad \hat{s} = x_1 x_2 s . \qquad (22)$$

Taking the final state leptons to be massless, their 4-momenta in the subprocess CM frame can be written:

$$\begin{aligned} k_\mu &= \frac{\sqrt{\hat{s}}}{2} (1, \cos\theta, \sin\theta, 0) , \\ k'_\mu &= \frac{\sqrt{\hat{s}}}{2} (1, -\cos\theta, -\sin\theta, 0) . \end{aligned} \qquad (23)$$

Obviously at leading order the dilepton invariant mass M is just equal to $\sqrt{\hat{s}}$. Thus a complete set of kinematic variables is M^2, Y and $\cos\theta$.

The polar angle θ, which here is the scattering angle in the subprocess CM frame, can be replaced by the *pseudorapidity* η, defined as

$$\eta = \ln \cot \frac{\theta}{2} . \qquad (24)$$

Thus scattering into the forward/backward regions corresponds to large positive/negative η. The only problem with η is that it does not transform elegantly under longitudinal boosts. From this point of view

a better variable is the *rapidity*:

$$y = \frac{1}{2}\ln\frac{E+p_z}{E-p_z} \ . \tag{25}$$

If I take a final state lepton with rapidity y in the subprocess CM frame, then in the lab frame its rapidity is $y+Y$. Note that for massless particles rapidity = pseudorapidity.

What we actually measure in the experiment are the lab frame pseudorapidities of the two leptons, their azimuthal angles, and their transverse momenta. The transverse momentum p_T is obviously invariant under longitudinal boosts; in terms of our subprocess CM frame variables it is given by:

$$p_T = \frac{\sqrt{\hat{s}}}{2}\sin\theta \ . \tag{26}$$

The lab frame pseudorapidities are nearly the same as the two lab frame rapidities; half the sum and half the difference of these give Y and y respectively. Thus p_T, y and Y are a complete set of observables for DY kinematics, as are p_T, M^2, and Y. Using the massless approximation, the lepton 4-vectors in the subprocess CM frame are then

$$\begin{aligned} k_\mu &= p_T(\cosh y, \sinh y, 1, 0) = (\tfrac{M}{2}, \sqrt{\tfrac{M^2}{4} - p_T^2}, p_T, 0) \ , \\ k'_\mu &= p_T(\cosh y, -\sinh y, -1, 0) = (\tfrac{M}{2}, -\sqrt{\tfrac{M^2}{4} - p_T^2}, -p_T, 0) \ . \end{aligned} \tag{27}$$

Changing variables from x_1, x_2 to M^2, Y in our previous cross section formula, we can compute the differential cross section $d\sigma/dM^2$. The Jacobian is

$$\left|\frac{\partial(x_1,x_2)}{\partial(M^2,Y)}\right| = \frac{x_1 x_2}{M^2} \ , \tag{28}$$

so the DY cross section formula becomes:

$$\sigma(pp \to \ell^+\ell^- + X) = \int_0^s dM^2 \ \frac{d\sigma}{dM^2} \ , \tag{29}$$

$$\frac{d\sigma}{dM^2} = \frac{4\pi\alpha^2}{3M^4} \int_{-\frac{1}{2}\ln\frac{s}{M^2}}^{\frac{1}{2}\ln\frac{s}{M^2}} dY \left[\sum_q x_1 f_q(x_1)\, x_2 f_{\bar{q}}(x_2)\, \frac{1}{3} Q_q^2 \right] .$$

Note that the differential cross section drops off like $1/M^4$, as we might have guessed by dimensional analysis. Also note that the effective pdfs are $x\, f_q(x)$, not $f_q(x)$, thus you will usually see pdfs plotted in x versus $xf_q(x)$.

You can find the order α_s corrections to Drell-Yan worked out in Ellis, Stirling and Webber. At this order the pdfs depend on a factorization scale: $f(x) \to f(x, \mu_F)$. A reasonable procedure is to set $\mu_F = \mu_R = M$, and use the popular $\overline{MS}$ renormalization scheme to subtract off the ubiquitous $\ln 4\pi - \gamma_E$ finite parts from dimensional regularization.

At NLO the kinematics of DY change. At leading order there were two independent kinematic variables x_1, x_2, or equivalently $\hat{s}, Y$, or equivalently M^2, Y. However at order α_s the dilepton invariant mass squared M^2 is no longer equal to $\hat{s}$; this is because at this order we include initial state radiation of a gluon from a quark, and we include a new subprocess:

$$q + g \to q^* \to q + \gamma^*/Z \to \ell^+\ell^- + X \; . \tag{30}$$

Thus M^2 is now independent from $\hat{s}$; conventionally we write the new kinematic variable as the dimensionless ratio $\tau \equiv M^2/s$. The complete order $\alpha_s\alpha^2$ result for DY (still ignoring the Z) is then given by:

$$\frac{d\sigma_{pp\to\ell^+\ell^- X}}{dM^2} = \frac{4\pi\alpha^2}{9} \frac{1}{M^2\hat{s}} \int_0^1 dx_1 \int_0^1 dx_2 \tag{31}$$
$$\times \left[\left(\sum_q f_q(x_1) f_{\bar{q}}(x_2) Q_q^2 \right) \left[\hat{s}\, \delta(M^2 - x_1 x_2 s) + \frac{\alpha_s(M^2)}{2\pi} D_q\left(\frac{M^2}{\hat{s}}\right) \right] \right.$$
$$\left. + \sum_{q/\bar{q}} \left(f_{q/\bar{q}}(x_1) f_g(x_2) + f_g(x_1) f_{q/\bar{q}}(x_2) \right) Q_q^2 \, \frac{\alpha_s(M^2)}{2\pi} D_g\left(\frac{M^2}{\hat{s}}\right) \right] ,$$

where the functions $D_q(M^2/\hat{s})$ and $D_g(M^2/\hat{s}$ are given in section 9.2 of Ellis, Stirling and Webber. Notice that to this order we need to know the pdfs for the gluon, as well as for the quarks and antiquarks.

To include the effect of the Z in Drell-Yan production, we replace the leading order QED process by the combined leading order electroweak process, using the vector and axial vector charges of the quarks and leptons to the Z:

$$\begin{aligned} V_f &= T_f^3 - 2Q_f \sin^2\theta_W \ , \\ A_f &= T_f^3 \ , \end{aligned} \tag{32}$$

where $T_f = \pm 1/2$ is the third component of weak isospin. Including the $Z - \gamma$ interference, the net result is to make the following replacement in our previous expressions:

$$Q_q^2 \to Q_q^2 - 2Q_q V_\ell V_q \chi_1(\hat{s}) + (A_\ell^2 + V_\ell^2)(A_q^2 + V_q^2)\chi_2(\hat{s}) \ , \tag{33}$$

where

$$\chi_1(\hat{s}) = \frac{\kappa \hat{s}(\hat{s} - M_Z^2)}{(\hat{s} - M_Z^2)^2 + \Gamma_Z^2 M_Z^2}, \quad \chi_2(\hat{s}) = \frac{\kappa^2 \hat{s}^2}{(\hat{s} - M_Z^2)^2 + \Gamma_Z^2 M_Z^2} \ , \tag{34}$$

and

$$\kappa \equiv \frac{\sqrt{2} G_F M_Z^2}{16\pi\alpha} \ . \tag{35}$$

Notice that the Breit-Wigner resonance of the Z is parametrized by a pole mass M_Z and a total decay width Γ_Z.

Figure 5 shows the excellent agreement between the NLO Standard Model prediction for Drell-Yan (using pdfs measured in other processes like DIS) and actual Tevatron data. Drell-Yan has a simple final state with a built-in calibration from the Z peak. This makes DY a promising channel for early discoveries at the LHC.

Even for Drell-Yan, however, there are several important worries for the LHC experiments:

- The theory predictions rely on knowing the pdfs to good accuracy. Especially for the gluon pdfs, this can only be done by using LHC measurements themselves.
- The Drell-Yan final state also gets contributions from other Standard Model processes, especially (at LHC) from $t\bar{t}$ production.
- A Z' could be very heavy and have a very broad width. In that case it may first be detected not as a resonance peak *per se*, but rather as a dip or shoulder in the high invariant mass tail of the early LHC data sets.
- A Z' (or Randall-Sundrum graviton resonance) could be very heavy and have a narrow width. In that case you need to understand your detector and QCD better to disentangle the Z' width from detector resolution and effects of initial and final state radiation.

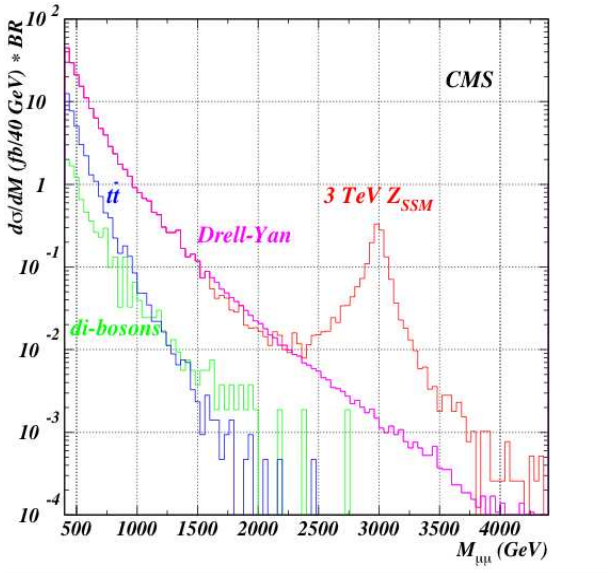

Figure 6: Simulated dimuon signal of a 3 TeV Z' at CMS, superimposed on the Standard Model background from Drell-Yan, $t\bar{t}$ and diboson production. From the CMS SUSYBSM webpage.

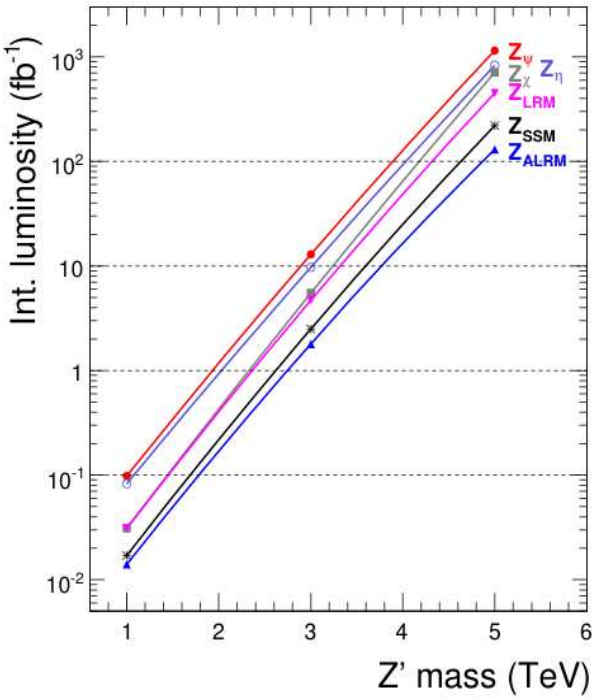

Figure 7: Estimated integrated luminosity required for a 5-sigma Z' discovery in the dimuon channel at CMS, for various Z' models. Taken from the CMS Physics TDR.

Figure 6 shows a CMS simulation of a 3 TeV Z' with the same couplings as a Z to quarks and leptons. Figure 7 shows how the CMS sensitivity to a Z' depends on the mass and the couplings. From this figure we see that even with the $\sim 1 fb^{-1}$ data sets from the 2008 run, a Z' discovery is possible for masses as high as 3 TeV.

A Z' is not the only resonance that could appear in the Drell-Yan final state at the LHC. One must also consider the possibility of resonances from heavy particles that have spin 0, 2, 3, etc. These could be Randall-Sundrum gravitons, string modes, or who knows what.

To distinguish these possibilities, a good start is to look at the angular distribution of the final state leptons. Let θ denote the polar angle of the negatively charged lepton with respect to the positive beam axis in the subprocess CM frame. Since the initial and final state particles are

highly relativistic, they are approximately helicity eigenstates. The lepton angular distributions are just determined geometrically by projecting the initial helicity state with total angular momentum J onto a helicity state with the z-axis rotated by θ. The probabilities for various spin J resonances and helicity combinations are given by the Wigner d-functions $d^J_{1,\pm 1}(\theta)$, which are tabulated in, *e.g.*, the Particle Data Book. One then averages over initial helicities and sums over final ones. For example, in the simplest case of parity conserving interactions, the normalized angular distributions for $q\bar{q} \to \ell^+\ell^-$ via a spin J resonance are given by:

Spin	*p*robability density
0	1
1	$\frac{3}{8}(1+\cos^2\theta)$
2	$\frac{5}{8}(1-3\cos^2\theta+4\cos^4\theta)$

Here we have ignored interference effects, higher order corrections, and the possibility of resonances from gg initial states. For parity-violating couplings like those of the Z, odd powers of $\cos\theta$ will also appear, with coefficients that depend on the couplings. Such effects are thus more model dependent but have the advantage of producing distinctive forward-backward asymmetries.

Using this kind of analysis, it should be be possible to discriminate between different kinds of resonances with real LHC data [3, 4].

References

[1] R. K. Ellis, W. J. Stirling and B. R. Webber, *QCD and Collider Physics*, Cambridge University Press, Cambridge, 1996.

[2] The CMS Collaboration, *CMS Physics Technical Design Report*, Volume II, CERN/LHCC 2006-021, 2006.

[3] B. C. Allanach, K. Odagiri, M. A. Parker and B. R. Webber, "Searching for narrow graviton resonances with the ATLAS detector at the JHEP **0009**, 019 (2000) [arXiv:hep-ph/0006114].

[4] R. Cousins, J. Mumford, J. Tucker and V. Valuev, "Spin discrimination of new heavy resonances at the LHC," JHEP **0511** (2005) 046.

LECTURE THREE:
How to discover top at the LHC

This lecture is devoted entirely to understanding $t\bar{t}$ production and detection at the LHC. Here is why:

- The top quark is the heaviest known particle. Its semileptonic decays produce large missing energy in association with high energy leptons and jets, just like (R-parity conserving) supersymmetry, not to mention Little Higgs with T-parity and Universal Extra Dimensions. If you don't know how to discover top at the LHC, then you don't know how to discover supersymmetry.
- Top production and decay is much simpler than supersymmetry, and has already been seen in large samples at the Tevatron.
- Top is a background (often the dominant background) to many possible new physics signals at LHC.
- The top samples at LHC may themselves provide discoveries, either from exotics which decay to top, or from new physics which mimics some properties of top (*e.g.* SUSY) and thus contaminates the top sample.

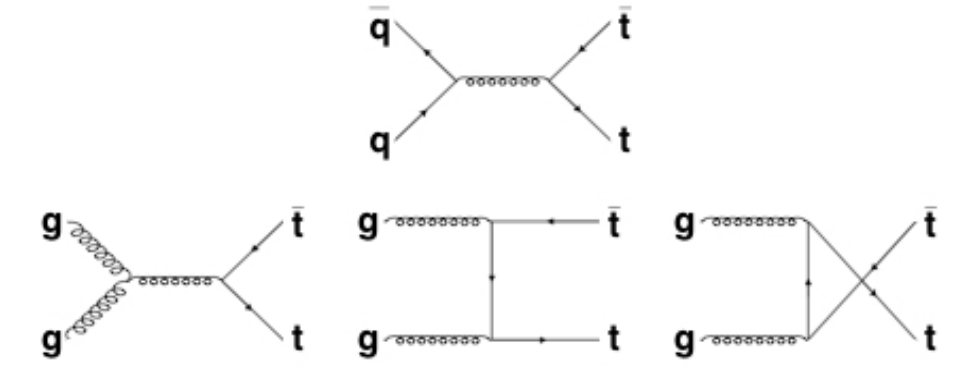

Figure 1: The leading order diagrams for hadronic production of $t\bar{t}$.

Figure 1 shows the leading order diagrams for the production of $t\bar{t}$ pairs at the LHC or the Tevatron. The initial hard scattering involves either two gluons or a quark and an antiquark. Since the top mass is 171.4 GeV, the threshold for $t\bar{t}$ production is around $\sqrt{\hat{s}} = \sqrt{x_1 x_2 s} = 350$ GeV. At the Tevatron, this can arise *e.g.* when a valence quark with $x \simeq 0.4$ from a proton hits a valence antiquark with $x \simeq 0.4$ from an antiproton. Consulting the parton distributions in Figure 2, you can make a rough estimate that $t\bar{t}$ production at the Tevatron is about 85% from $q\bar{q}$ and 15% from gg; this is indeed correct.

At the LHC the relevant parton momentum fractions are smaller (since $\sqrt{s} = 14$ TeV), and $q\bar{q}$ annihilation requires that we scatter a sea antiquark from one of the incident protons. Both of these facts favor gg as the dominant inital state for $t\bar{t}$ production, and indeed about 90% of $t\bar{t}$ pairs at LHC will come from gg.

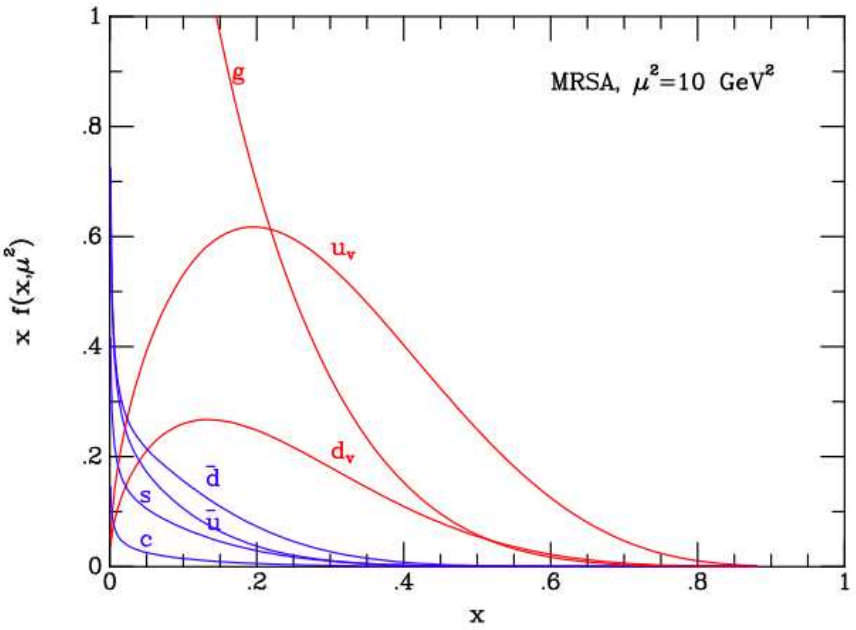

Figure 2: Parton distributions for the proton. Taken from R. K. Ellis, 2006 TASI lectures.

The table below shows estimates of the $t\bar{t}$ cross section for the LHC for

a nominal top mass of 175 GeV [2]. These were computed at NLO in QCD, and also included a resummation of soft gluon effects. For simplicity the renormalization and factorization scales are taken to be equal, and are varied between $m_t/2$ and $2m_t$ to give an idea of the remaining scale dependence. As you can see, this uncertainty is about 6% .

$\mu_{\mathbf{R}} = \mu_{\mathbf{F}}$	$t$otal cross section in pb
$m_t/2$	883
m_t	825
$2m_t$	782

In addition, these calculations used a particular choice for the pdfs (the MRST 1999 fits). These pdfs come from fits to data, but existing data do not constrain the gluon pdf very well in the range of x relevant to $t\bar{t}$ production at the LHC. By varying the pdf fits, it is estimated that this contributes a further uncertainty of $\pm 10\%$ to the total cross section. Other processes at the LHC, especially inclusive jet production, will provide better estimates of the gluon pdfs. However applying this data will introduce an additional uncertainty since it is not obvious how the scale dependence of jet production translates to $t\bar{t}$.

In any event, the $t\bar{t}$ cross section for LHC is very large: almost a nanobarn. This is more than 100 times the $t\bar{t}$ cross section at the Tevatron, and about 1000 times larger than the inclusive cross section for 1 TeV superpartners. At design luminosity the LHC will produce nearly 10 top events per second.

This is impressive, but we saw in Lecture 1 that the total collision rate the LHC will be about a Gigahertz. To see top, the experiments need a very fast way to trigger on and read out the one event per 4 million that contains top. To do this, we need to know a lot more about the experimental signature of the $t\bar{t}$ final state.

top decays

Top quarks decay nearly 100% of the time to a W and a b quark:

$$t \to W^+ b \,, \qquad \bar{t} \to W^- \bar{b} \,. \tag{1}$$

This weak decay is suppressed by G_F but enhanced by m_t^3:

$$\Gamma_t = \frac{G_F m_t^3}{8\pi\sqrt{2}} |V_{tb}|^2 \left(1 - \frac{M_W^2}{m_t^2}\right)^2 \left(1 + 2\frac{M_W^2}{m_t^2}\right)^2 \left[1 - \frac{2\alpha_s}{3\pi}\left(\frac{2\pi^2}{3} - \frac{5}{2}\right)\right] , \tag{2}$$

neglecting terms of order m_b^2/m_t^2, $\alpha_s M_W^2/m_t^2$, and higher orders in α_s. Plugging in the numbers gives a top width of about 1.76 GeV, which means that the top lifetime is short compared to $\hbar/\Lambda_{\rm QCD}$. Thus the top quark decays before it has a chance to hadronize. The b quark however does hadronize, producing a B meson and other hadrons, which eventually become a jet. As mentioned in Lecture 1, the B meson lifetime is sufficiently long that b-jets can be distinguished (with some efficiency) from jets initiated by lighter quarks.

Of course the W boson is highly unstable and decays rapidly to either two leptons or two quarks. The relative branching fractions are shown in the table. As I said in Lecture 1, at a hadron collider "lepton" is usually taken to mean an electron or a muon, not a tau or a neutrino. So top quarks decay 22% of the time to a "lepton" plus a b-jet plus missing energy, and about 2/3 of the time to a b-jet plus two non-b-jets.

W boson decay modes	
$d$ecay mode	$b$ranching fraction
$e + \nu_e$	11%
$\mu + \nu_\mu$	11%
$\tau + \nu_\tau$	11%
$q + \bar{q}$	68%

Thus, neglecting taus, $t\bar{t}$ events produce three distinct final states:

- *dileptons channel:* About 5% of the time, the final state is two b-jets, two high energy leptons with opposite sign charge, and large missing transverse energy.
- *leptons+jets channel:* About 30% of the time, the final state is two b-jets, two non-b-jets, one high energy e or μ and missing transverse energy.
- *all jets channel:* About 44% of the time, the final state is two b-jets and four non-b-jets.

triggers

At design luminosity ATLAS and CMS will see an event (each event having ~ 25 hard scatterings) 40 million times per second. These events will be read out by fast detector components with response times in the range 20-50 nanoseconds. However the rate at which full events will be written to tape for later off-line analysis will be on the order of 100 Hz! Sophisticated trigger and data acquisition systems are designed to filter the events, attempting to save all of the ones which are potentially interesting, along with a prescaled fraction of generic events. The first layer of filtering is called the Level 1 trigger, which is designed to make coarse but rapid assessments of the general properties of each event. The event rate coming out of Level 1 will be about 0.1 Megahertz. The slowness of the speed of light is overcome at Level 1 by "pipelining" the data. The data stream moves along an electronic pipeline at the end of which it gets dumped unless a positive decision is made to divert it. The time to traverse the pipeline is long compared to 25 nanoseconds.

After Level 1 there are higher level triggers; these have enough time to analyze more detailed and integrated features of the events. Although the triggers involve fast electronics they are also programmable,

allowing a flexible "trigger menu". A typical entry in a trigger menu would be "all events containing a candidate electron with $p_T > 25$ GeV." Since many different triggers are desired, any one should not exceed about 10 Hz in rate. If the rates turn out to be too high in the real experiment, the trigger thresholds (25 GeV in the example above) will have to be raised.

In considering the discovery prospects for any new phenomenon at the LHC, the first question is whether this phenomenon (usually a heavy particle) will be produced at the LHC with reasonable rates. The second question is whether we can trigger on this phenomenon. If you can't trigger on it, you can't discover it, unless the phenomenon is so copiously produced that it can be seen in generic pre-scaled samples (this is unlikely since the pre-scalings are by factors as large as a million).

How do we trigger on $t\bar{t}$ events? The characteristics of these events are: multiple jets at least two of which are b-jets, high p_T leptons, and missing transverse energy. The simplest trigger would be to require just a jet with large transverse energy, $E_T \equiv E \sin\theta_{lab}$, plus anything. While such a trigger will exist, the threshold will have to be set at something like 400 GeV, which means this sample will miss a substantial fraction of $t\bar{t}$ events. Since the all jets and leptons+jets channels contain at least four jets, another approach is to design a trigger asking for at least four jets. However is this case the estimated E_T thresholds to avoid pre-scaling are 110 GeV per jet, which is still uncomfortably high. For the all jets channel the only other improvement we can make in the trigger is to require either more jets or that at least one jet is b-tagged. The b-tagging is done either by looking in the tracker for a displaced vertex from a B meson decay (a "vertex tag"), or by taking advantage of the fact that 22% of b-jets have a muon inside them (a "soft muon tag") which is rarely true for light quark jets. The trouble with including b-tagging in your trigger is that, of course, only a fraction of

actual b-jets are identified; an additional problem is that you had better not rely on complicated triggers until you understand your detector with great confidence. CMS estimates that they can trigger on $t\bar{t}$ in the all jets channel with an efficiency of 17% , but in truth this remains to be worked out once we have real data.

For the lepton+jets channel we can do much better by triggering on either a single electron or muon candidate with p_T larger than 20-25 GeV. CMS estimates that with such a trigger they can collect $t\bar{t}$ events from the leptons+jets channel with 62% efficiency. The difference between this number and 100% comes mostly from the limited geometrical "acceptance" of the detector, *e.g.* only muons with pseudorapidity in the range $-2.4 \leq \eta \leq +2.4$ are detected. For the dilepton channel one has the possibility of using dilepton trigger samples, which ask for either two electrons or two muons. These triggers have the advantage that the p_T thresholds can be cut roughly in half. On the other hand CMS estimates that they can achieve 80% trigger efficiency for $t\bar{t}$ events in the dilepton channel, just using the single lepton triggers.

Last but not least, we could consider using a MET trigger, *i.e.* triggering on events with missing transverse energy above some threshold. This is a very important trigger for lots of new physics. However it is not efficient for $t\bar{t}$, since the threshold will have to be set at something like 200 GeV. This can be lowered to 80 GeV by also requiring a hard jet (> 180 GeV), but this will still miss most of the $t\bar{t}$ events, where the missing energy is just neutrinos from W decay. Missing energy as a trigger creates problems, since in addition to collecting exotic new physics (neutralinos, gravitons, gravitinos etc.) it also collects a huge number of "garbage" events, where *e.g.* your jets were poorly measured, cosmic rays or beam halo sprayed a shower into your detector, etc. Figure 3 shows the ratio of garbage to real missing energy as observed at the Tevatron in a mature experiment.

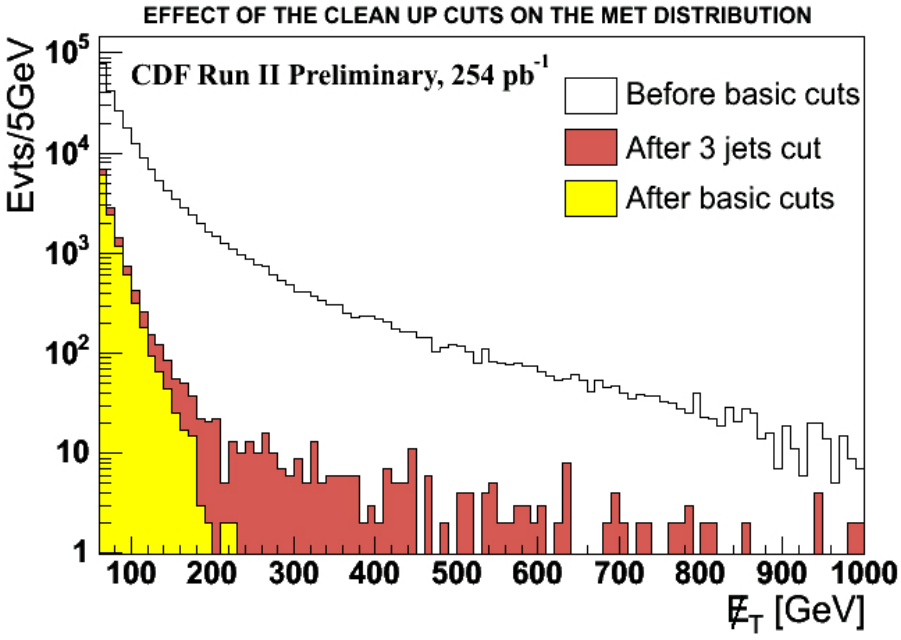

Figure 3: CDF missing energy events before and after clean up. Notice the log scale.

kinematics

In the lab frame the 4-momenta of the top and anti-top quarks can be written:

$$\begin{aligned} k_\mu &= (m_T \cosh y_t, m_T \sinh y_t, p_T \cos\phi, p_T \sin\phi) \ , \\ k'_\mu &= (m_T \cosh y_{\bar{t}}, m_T \sinh y_{\bar{t}}, -p_T \cos\phi, -p_T \sin\phi) \ , \end{aligned} \tag{3}$$

where $y_t, y_{\bar{t}}$ are the lab frame rapidities and the transverse mass m_T is defined by

$$m_T = \sqrt{p_T^2 + m_t^2} \ . \tag{4}$$

This parametrization has a very simple relationship to the 4-vectors in the subprocess CM frame:

$$k_\mu = (m_T \cosh y, m_T \sinh y, p_T \cos\phi, p_T \sin\phi) \ ,$$

$$k'_\mu \quad = \quad (m_T \cosh y, -m_T \sinh y, -p_T \cos\phi, -p_T \sin\phi) \ , \tag{5}$$

since

$$y = \frac{1}{2}(y_t - y_{\bar{t}}) \ , \quad Y = \frac{1}{2}(y_t + y_{\bar{t}}) \ . \tag{6}$$

Similarly I can write:

$$\sqrt{\hat{s}} = 2 m_T \cosh y \ , \tag{7}$$

and of course we can also obtain the parton momentum fractions x_1, x_2 using:

$$x_1 = \sqrt{\frac{\hat{s}}{s}}\, \mathrm{e}^{Y} \ , \quad x_2 = \sqrt{\frac{\hat{s}}{s}}\, \mathrm{e}^{-Y} \ . \tag{8}$$

Let's revisit the leading order Feynman diagrams for $t\bar{t}$ production shown in Figure 1. Each diagram has one off-shell particle (either a gluon or a top quark). The denominators of their propagators will be

$$\begin{aligned} (p_1 + p_2)^2 &= 2 p_1 \cdot p_2 = 4 m_T^2 \cosh^2 y \ , \\ (p_1 - k)^2 - m_t^2 &= -2 p_1 \cdot k = -2 m_T^2 \mathrm{e}^{-y} \cosh y \ , \\ (p_2 - k)^2 - m_t^2 &= -2 p_2 \cdot k = -2 m_T^2 \mathrm{e}^{y} \cosh y \ . \end{aligned} \tag{9}$$

Notice that all of these propagators are off-shell by an amount whose magnitude is at least as big as m_T, which is at least as big as m_t. This means that the initial state QCD process has a built-in cutoff which is much larger than Λ_{QCD}, ensuring the validity of perturbative QCD. This also indicates that we should take the renormalization scale equal to something like m_t or m_T.

Now we write down the leading order cross section for $t\bar{t}$ production:

$$\frac{d^3\sigma}{dy_t dy_{\bar{t}} dp_T^2} = \frac{1}{256\pi^2 m_T^4 \cosh^4 y} \sum_{i,j} f_i(x_1, \mu_F^2) f_j(x_2, \mu_F^2) \overline{\sum} |\mathcal{M}_{ij}|^2, \tag{10}$$

where the bar over the sum indicates averaging/summing over initial/final spins and colors. The matrix elements are of course

different for the $q\bar{q}$ and gg processes:

$$\overline{\sum}|\mathcal{M}_{q\bar{q}}|^2 = \frac{2g_s^4}{9}\left(\frac{1}{\cosh^2 y}\right)\left[\cosh 2y + \frac{m_t^2}{m_T^2}\right], \quad (11)$$

$$\overline{\sum}|\mathcal{M}_{gg}|^2 = \frac{g_s^4}{48}\left(\frac{8\cosh 2y - 1}{\cosh^2 y}\right)\left[\cosh 2y + 2\frac{m_t^2}{m_T^2} - 2\frac{m_t^4}{m_T^4}\right].$$

These formulae tell us the basic kinematic features of $t\bar{t}$ production. The $q\bar{q}$ contribution to the cross section damps like $1/\cosh^4 y$ for large y, and the gg contribution to the cross section damps like $1/\cosh^2 y$. Thus the $t\bar{t}$ pairs tend to be produced at right angles to the beam in the subprocess CM of frame, which means that their rapidity difference in the lab frame tends to be small.

For fixed rapidity, we also see that the cross section damps like $1/m_T^4$ for large p_T, with a further suppression from the dropoff of the pdfs as m_T and thus x_1, x_2 get larger. The most likely values for the p_T of the top quarks are therefore of order m_t.

looking at the data

In the real experiment, of course, we do not see top quarks. In a candidate $t\bar{t}$ event, we would like to be able to *reconstruct* the top quarks from the raw data. I will skip the first (and most diffciult) steps of this process, which convert the raw hits in the detector into the reconstructed "physics objects" that I introduced in the first lecture: electrons, muons, photons, taus, jets, heavy flavor jets, and missing transverse energy (MET). Part of this reconstruction assigns three numbers to each of these objects: a p_T or E_T, a pseudorapidity η, and an azimuthal angle ϕ. This amounts to assigning 4-vectors to each of these physics objects, since we know their masses. For jets we can cook up a "jet mass" by treating each cluster of activated calorimeter cells as a massless object, assigning it a 4-vector, then adding up these 4-vectors

to compute the invariant mass of the jet. The other special case is the MET physics object, which is only assigned two numbers: a missing E_T and a ϕ.

The result of this procedure is that for each event the off-line analysis produces a list of n-tuples, each entry giving the identity of the physics object, its 4-vector, and additional information if known such as the electric charge, the number of charged tracks in a jet, information about a b-tag, etc.

Event #37110285

	object	eta	phi	p_T in GeV	n-tracks	
1	muon	0.580	1.571	6.18	-1	
2	muon	2.091	5.222	22.87	-1	
3	muon	0.539	1.611	12.23	1	
4	jet	0.570	1.555	123.20	16	
5	jet	1.103	4.542	158.10	34	
6	jet	1.463	1.922	74.99	11	
7	jet	1.045	3.983	42.49	23	
8	b-jet	0.758	5.015	37.08	23	(vertex tag)
9	jet	2.002	5.228	36.39	22	
10	jet	0.972	0.376	27.39	4	
11	MET		1.265	66.30		

An example of such an event listing is shown above. Was this a $t\bar{t}$ event? It certainly could be. There is a hard muon, 66 GeV of MET, a vertex b-tagged jet, and another hard jet that seems to have two muons in it, consistent with being another b-jet. There are also 5 additional jets, more than the two we need for the lepton+jets channel. However having extra jets is not too worrisome. For one thing, the number of jets in an event depends on your definition of a jet, something that the Standard Model doesn't specify. In addition, gluon radiation in the initial or final state can easily produce extra jets. Furthermore, we have only

considered $t\bar{t}$ production at leading order. At NLO one can *e.g.* produce $t\bar{t}$ from a qg initial state, which automatically implies an extra hard jet.

With any single event, there is no way to be sure of its identity. For example the CDF experiment observed events in 1992 in the lepton+jets channel that looked just like $t\bar{t}$ with $m_t = 175$ GeV, but they did not declare discovery of top until 1995, with a larger sample of candidate events. Discoveries at colliders have to be done statistically, evaluating the hypothesis that an observed sample was produced by a combination of statistical fluctuation and systematic errors, rather than by some new physics. Traditionally "discovery" is defined as a data sample such that only a 5 sigma fluctuation or more of combined effects can explain it in the absence of new physics. This definition may seem highly conservative, but it is based on experience over the decades with many 3 sigma fluctuations that have eventually gone away upon further analysis or with the inclusion of more data. The basic problem is that analyses of complex experiments often contain unsuspected or underestimated systematic errors, so that 2-3 sigma fluctuations occur much more often than Gaussian statistics would predict.

Notice that this definition of discovery does not involve a statistical analysis of the putative new physics signal but rather of all the other things that could fake it. These are usually divided into "physics backgrounds" and "detector backgrounds". For $t\bar{t}$ in the lepton+jets channel, the largest physics background is from the direct production of a W boson in association with 4 jets. An example of a detector background would be a jet misidentified as a lepton, combined with fake MET from jet mismeasurement; these are unlikely occurences but they involve pure QCD jet production which has a cross section a million times larger than real $t\bar{t}$. Detector backgrounds are difficult to estimate, requiring both detailed detector simulation and real data samples.

Physics backgrounds are estimated by theorists. Consider for example the $W + 4$ jets process, which is a dominant background at LHC for

both $t\bar{t}$ and for supersymmetry. Obviously some team of theorists must have computed this at NLO a decade ago, and probably is close to having it at NNLO? Unfortunately not. NLO calculations of Standard Model processes involving 7 external legs are incredibly challenging. Even using your favorite string-inspired twistor tricks to evaluate amplitudes in perturbative QCD, we are still a long way from a calculation of $W + 4$ jets at NLO. Part of the problem, by the way, is sociological: there are 300 string theory students at this school, but only ~ 30 people in the world working on Standard Model calculations of basic importance for LHC discovery.

event generation

As we have just seen, discovery at the LHC will rely on a detailed understanding of both physics and detector backgrounds. In addition, we would like to be able to make clever "cuts" on the data samples that will enhance the signals and suppress the backgrounds. In order to design such an analysis we obviously need a detailed understanding of the signals. All of this has to be done first at the level of computing matrix elements, but then somehow translated into simulating how real events should look in the ATLAS and CMS detectors. Simply put, we need three kinds of tools:

- Matrix element generators.
- Showering Monte Carlo programs to convert parton-level initial and final states into fully hadronized events.
- Detector simulations.

For tree level matrix elements there are now available several very powerful public programs. These include MadGraph/MadEvent and CompHEP/CalcHEP, which allow you to start at the Lagrangian level, then automatically generate Feynman graphs and compute leading

order matrix elements. If tomorrow I invent a new model for Terascale physics, "hyperSUSY with J parity", within a couple of weeks I could probably incorporate it into one or both of these programs. Another tool called ALPGEN is limited to Standard Model processes, but computes leading order matrix elements for high multiplicity processes like $W + 4$ jets. There are also matrix element generators for a large and growing number of Standard Model processes at NLO. For example MCFM can do $W + 2$ jets at NLO, and $t\bar{t}$ at NLO. At present only a few experts can use NLO generators for full event generation.

The dominant general purpose showering Monte Carlo programs are Herwig and Pythia, with growing competition from a young upstart called Sherpa. These programs use built-in matrix elements or take input from matrix element generators like MadEvent. They then create full events with hadronized final states, producing a standard output that is easily interfaced with detector simulations. The showering models try to simulate in an ad hoc way what real QCD does, causing initial and final state partons to emit gluons, which in turn emit more gluons or split into $q\bar{q}$ pairs. The result of such showering is depicted in Figure 4. It is important to note that the showering algorithms require that hard gluon emissions come before softer ones, and of course eventually all of the particles in the shower are soft.

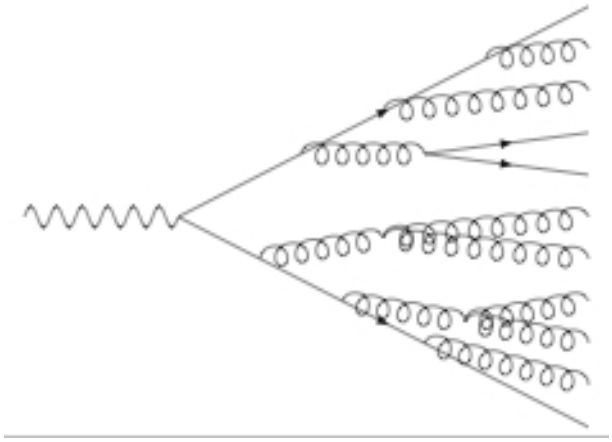

Figure 4: A schematic view of a parton shower.

Pythia incorporates the Lund model of string fragmentation to convert parton showers into hadrons. Note that this is the most used application of string theory in particle physics to date! In this model gluons are treated as a $(3, \bar{3})$ color octet equivalent to a $q\bar{q}$ pair. As $q\bar{q}$ pairs separate, a color flux tube is assumed to join them, which is modeled as a relativistic string with a string tension $\kappa \sim 1$ Gev/fermi. An ad hoc probabilistic recipe causes these strings to fragment at some point, producing new $q\bar{q}$ pairs, or pairs of diquarks to make it easier to make baryons. This process of making strings and fragmenting them continues until there are no concentrations of energy left able to compete with Λ_{QCD}. At that point a model assembles all the partons into hadrons.

ATLAS and CMS use full-blown detailed detector simulations that consume enormous quantities of CPU. As a theorist you will never be allowed anywhere near these programs. However two much simpler programs are available for theorists to play with. For ATLAS there is AcerDET, while for CMS or ATLAS there is PGS 4. These are parametrized detector simulations that only roughly approximate the real thing. However they are simple to use, and you can read the Fortan code to see what the program is doing, which is educational.

discovering top in simulated data

Now let's analyze some simulated data. To get the simulated data, I generated 380,000 $t\bar{t}$ events on my laptop using Pythia 6.4. Pythia only uses leading order matrix elements, so for a real study one would compare results with other generators. From the NLO cross section I know that 380,000 events corresponds to about 0.5 fb^{-1}, which is in the ballpark of what ATLAS and CMS may have in 2008. I fed the output of Pythia into PGS 4.0, which produces simulated event listings similar to the one that I showed above.

I want to figure out a strategy to analyze this data such that I expose its true identity as $t\bar{t}$, a strategy which in a real data sample would maximize the top signal at the expense of backgrounds. Again to do a serious study I would actually have to simulate physics backgrounds like $W + 4$ jets, and use a better detector simulation.

I am going to follow a 2004 ATLAS study which was mentioned in a paper by Fabiola Gianotti and Michelangelo Mangano [8]. The assumption is that in the first physics run b-tagging may be unreliable or unavailable, so we won't use it in our analysis. Muon measurements should be under control, as well as basic jet measurements, except that the calibration of jet energies may still be shaky for multijet events like $t\bar{t}$.

So a conservative approach to take our 380,000 events and only keep the ones which have at least one muon (with $p_T > 20$ GeV otherwise we won't trigger on it) and with exactly four jets (with $E_T > 40$ GeV). Thus we are only going to analyze events that would correspond to relatively clean $t\bar{t}$ in the lepton+jets channel. We also require that the muon and jets have $|\eta| < 2.4$, so that they are actually detected and (one hopes) well-measured.

Of our initial 380,000 events, only 10,829 pass all these requirements.

This is only 3% efficiency for our generic signal events, but in line with what is done in real analyses *e.g.* at the Tevatron.

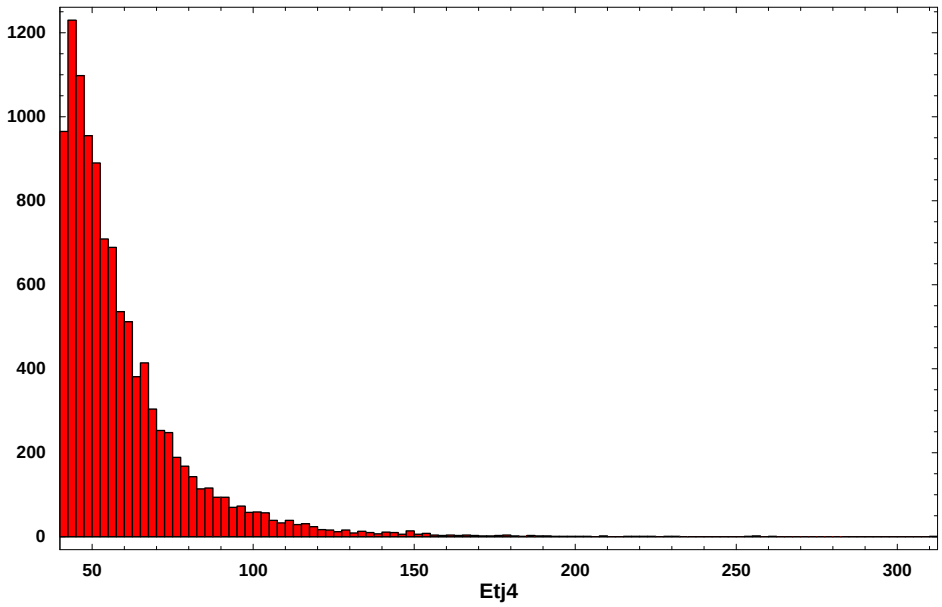

Figure 5: E_T distribution of the 4th leading jet

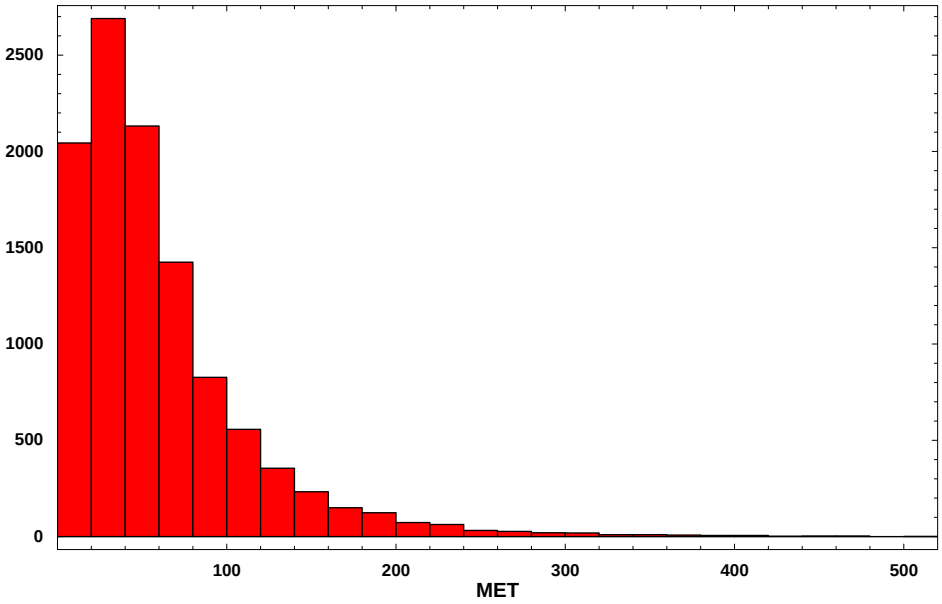

Figure 6: Distribution of missing transverse energy.

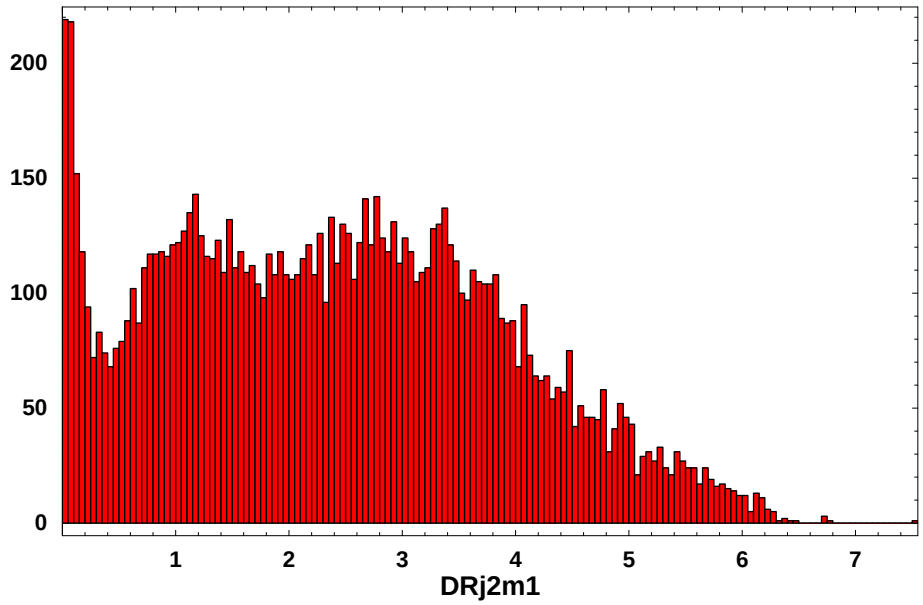

Figure 7: The separation ΔR between the leading muon and the 2nd leading jet.

For these 10,829 events I have made a bunch of histograms using Mathematica (a real collider physicist would use ROOT for this). Figure 5 shows the E_T distribution of the 4th leading jet (*i.e.* the least energetic jet) in the selected events. This verifies my previous claim that a four jet trigger that requires the fourth jet to have $E_T > 110$ GeV is not very efficient for $t\bar{t}$. Figure 6 shows the distribution of MET. This verifies my claim that a 200 GeV MET trigger is not very efficient for $t\bar{t}$. It is also interesting that the neutrino sometimes manages to produce MET larger than 400 GeV, even in this small sample.

Figure 7 shows a histogram of ΔR, which is a useful estimate of the opening angle between two objects in the subprocess CM frame:

$$\Delta R = \sqrt{(\eta_1 - \eta_2)^2 + (\phi_1 - \phi_2)^2} \ . \tag{12}$$

Thus for example two massless objects produced back-to-back in the azimuthal plane with zero rapidity have $\Delta R = \pi$. Figure 7 shows ΔR for between the most energetic muon and the 2nd leading jet. We notice two things. First is that the distribution is pretty flat. Second is the

spike near zero; these correspond to events in which the leading muon is inside this jet. Those events are probably not lepton+jets events; they are probably all jets events with the muon coming inside a b-jet. Although not really essential for this analysis, we can get rid of this pollution by making a muon "isolation cut", where we only keep events that have $\Delta R > 0.7$ between the leading muon and any of the jets.

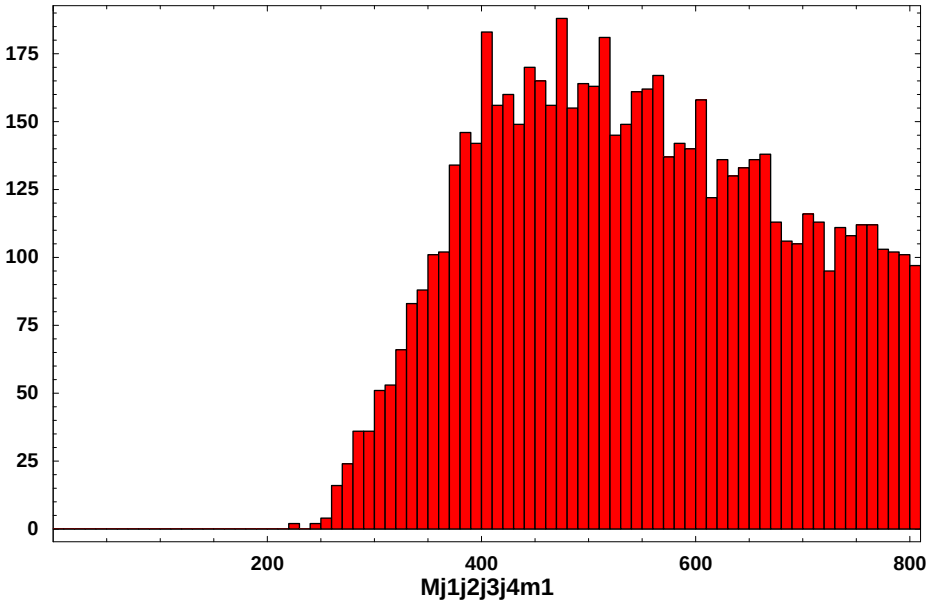

Figure 8: Total visible invariant mass.

Figure 8 shows the total visible invariant mass in our events. This is an estimate of the subprocess CM energy $\sqrt{\hat{s}}$, although of course it is missing the contribution from the neutrino. What do we expect for this distribution? Recall that $\sqrt{\hat{s}} = 2m_T \cosh y$, and we said that y tends to be close to zero. We also said that p_T likes to be of order m_t, so m_T likes to be around $\sqrt{2}m_t \simeq 500$ GeV. Thus we predict that this distribution should turn on around $2m_t \simeq 350$ GeV, and have a broad peak around 500 GeV. The figure shows both of these features.

Now what we will really want to do is reconstruct at least one of the two top quarks. The easiest one to reconstruct in this channel is the one

that decayed to $b + W \to 3$ jets, since we know the full 4-vectors for all the jets. However we do not know in each event which 3 of the 4 jets reconstructs a top.

Thinking a bit more about kinematics, the b-jets will tend to be more energetic than the two jets from the hadronic W decay (why?). So of the four possible combinations of 3 jets, j1j2j3 and j1j2j4 are *least* likely to reconstruct a top. So let's just look at the other two combinations: j1j3j4 and j2j3j4. We treat this as a physics object whose 4-vector is just the sum of the 4-vectors of the 3 jets.

As a last bit of cleverness, notice that if we have the wrong combination, then two of the jets are actually b-jets, which like their parent top quarks like to be back-to-back in the azimuthal plane. So these wrong combinations give a 4-vector in which the *net* p_T tends to be smaller. So let's only look at events in which one of our candidate top objects j1j3j4 or j2j3j4 has the *largest* net p_T of any three jet combination. Together with the muon isolation cut, this leaves us with 4,422 events.

After these selections, Figure 9 shows the distribution in invariant mass of our candidate reconstructed tops. There is a clear peak in the bin at 170-180 GeV. Even if I add physics backgrounds like $W+$ jets, this peak is an unambiguous signal of top. Indeed if the experiments find this peak at *e.g.* 190 GeV, they will recalibrate their jet measurements until the peak comes out closer to 175: top is a calibration tool!

Figure 10 is the analogous plot from [8], showing also the $W + 4$ jets background simulated with ALPGEN. Our results agree beautifully, and we can even detect a typo in [8]: the integrated luminosity for their plot is apparently 1.5 fb^{-1}, not 150 pb^{-1}.

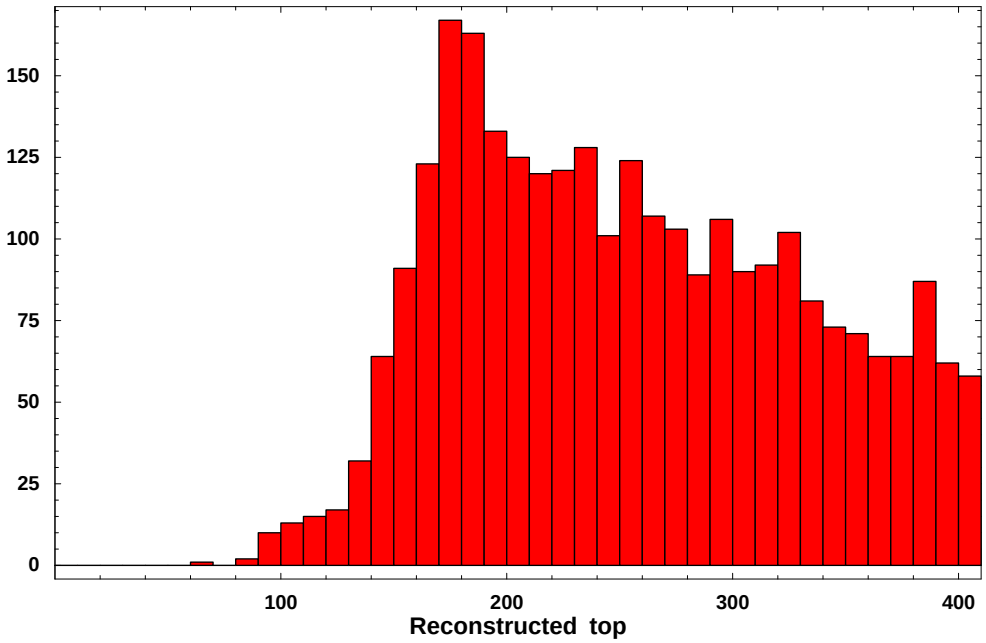

Figure 9: Three jet invariant mass after the "max p_T" selection.

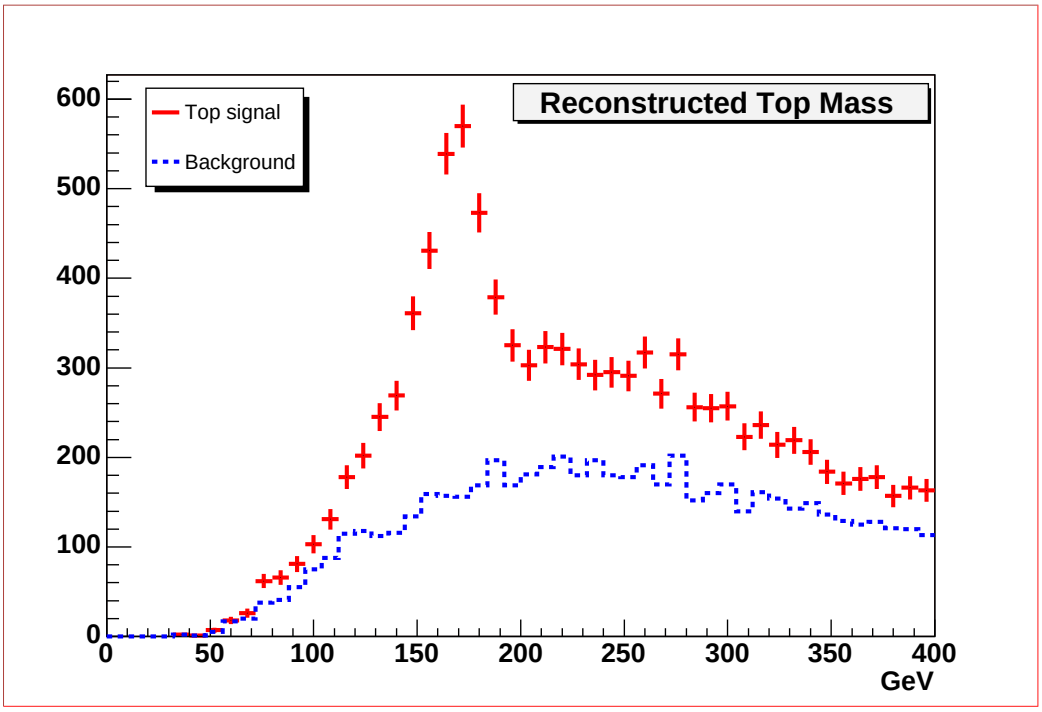

Figure 10: Three jet invariant mass after selections, from F. Gianotti and M. Mangano [8].

There is a lot more that we can learn from studying top. Of course, when we move on to look for supersymmetry at the LHC, $t\bar{t}$ becomes a background that we want to reduce, not enhance. SUSY events involve multiple jets, leptons and MET, and we will use the same kind of detailed kinematic analysis that we used here to expose a SUSY signal and reconstruct the identity and masses of as many superpartners as possible.

References

[1] R. K. Ellis, W. J. Stirling and B. R. Webber, *QCD and Collider Physics*, Cambridge University Press, Cambridge, 1996.

[2] M. Beneke *et al.*, "Top quark physics," arXiv:hep-ph/0003033.

[3] The ATLAS Collaboration, *ATLAS High-Level Trigger Data Acquisition and Controls Technical Design Report*, CERN/LHCC 2003-022, 2003.

[4] The CMS Collaboration, *CMS Physics Technical Design Report*, Volume II, CERN/LHCC 2006-021, 2006.

[5] M. A. Dobbs *et al.*, "Les Houches guidebook to Monte Carlo generators for hadron collider physics," arXiv:hep-ph/0403045.

[6] T. Sjostrand, "Monte Carlo generators," arXiv:hep-ph/0611247.

[7] M. L. Mangano and T. J. Stelzer, "Tools For The Simulation Of Hard Hadronic Collisions," Ann. Rev. Nucl. Part. Sci. **55**, 555 (2005).

[8] F. Gianotti and M. L. Mangano, "LHC physics: The first one-two year(s)," arXiv:hep-ph/0504221.