"An UnOrthodox Introduction to QCD"
(cf. hep-th/0309149)

CTEQ Summer School Lecture on Jet Substructure

Andrew Larkoski

It is extremely challenging to introduce any subject in one hour, and the field of jet substructure is especially challenging because its purview now encompasses much/most of the physics program of the LHC (Higgs, BSM, SM measurements, fragmentation, heavy flavor, heavy ions, etc.). So, this lecture will be narrow in scope and ignore much/most of the applications of this exciting field. While you have probably been introduced to QCD from the gauge principle, analogies with electromagnetism, and finally its (high-energy) Lagrangian, I want to take a different approach here. Jet substructure, at its most fundamental, is a study of QCD in the near soft (low-energy) and collinear (small-angle) limits, and the Lagrangian of QCD isn't a natural starting point for studying this. Instead of a "top-down" approach, I want to emphasize a "bottom-up" approach, starting from some natural, simple assumptions about the behavior of ~~high~~ QCD at high energies. We'll see that this will be remarkably powerful.

To proceed, we need to make two reasonable assumptions. These are important enough that I will call them axioms, for the purposes of this lecture.

Axiom 1: At high energies, the coupling of QCD, α_s, is small. Therefore quarks and gluons are good quasi-particles.

This axiom essentially means that the perturbation theory of QCD (defined by Feynman diagrams, for example) is a good approximation. It is sensible to describe final states in terms of quarks and gluons as corrections to this picture are small because α_s is small.

Axiom 2: At high energies, QCD has no intrinsic scales. Quarks and gluons are massless, and so QCD is (approximately) a conformal, or scale-invariant quantum field theory.

We know that this axiom strictly isn't true. While quarks and gluons may be very low mass or massless, hadrons are massive. Also, the coupling α_s runs with energy, spoiling true scale-invariance. Nevertheless, because α_s is small (by axiom 1) these deviations from scale invariance can be thought of as corrections. This is how we will treat them in this lecture, and I will discuss how to fix-up this picture later.

With these axioms established, I would now like to do some simple calculations. Let's calculate the probability for a quark to emit a gluon:

$P(q \to qg) = \Big|$  $\Big|^2$

We will express this probability in terms of the phase space variables of the final state (the quark and the gluon). What are these phase space variables? For two particles, two-body phase space is two-dimensional. Each particle has a four-vector momentum which is required to be on-shell and massless. Additionally, the sum of those four-vector momenta is the total initial momentum. This leaves two degrees of freedom, or two phase space variables. We will choose these phase space variables to be the energy of the gluon, E_g, and the invariant mass of the final quark and gluon, m^2. Then,

$$P(E_g, m^2) = \Bigg| \quad \Bigg|^2$$

Note that $m^2 = 2 p_q \cdot p_g = 2 E_q E_g (1 - \cos\theta_{qg})$.

What can this probability be? Our assumption of scale-invariance helps us out. Scale invariance means that the probability is unchanged if the energy/mass scales are multiplied by a factor $\lambda > 0$:

$$P(\lambda E_g, \lambda^2 m^2)\, d(\lambda E_g)\, d(\lambda^2 m^2) = P(E_g, m^2)\, dE_g\, dm^2.$$

What could this function be? The simplest function that one can write down is:

$$P(E_g, m^2)\, dE_g\, dm^2 = \frac{\alpha_s C_F}{\pi} \frac{dE_g}{E_g} \frac{dm^2}{m^2}$$

Before continuing, I should say a couple things about this expression.

First, the overall factor of $\frac{\alpha_s C_F}{\pi}$ is the strength of which a gluon couples to a quark; C_F is the color factor that represents the amount of color that the quark carries (called the fundamental representation Casimir). We'll come back to this later. Note also that we could multiply this expression by any function of E_g^2/m^2 and still maintain scale-invariance. This will be important for detailed studies, but there is a well-defined approximation in which we can ignore such terms. This is called the "double-logarithmic approximation" or DLA.

With this DLA probability in hand, let's change variables to dimensionless quantities, as they are a bit nicer to work with. Let's express the probability in terms of the gluon's energy fraction, z, and the angle $\theta_{qg} \equiv \theta$ between the quark and gluon:

$$z = \frac{E_g}{E_q + E_g}, \quad 1 - \cos\theta = \frac{m^2}{2E_q E_g}$$

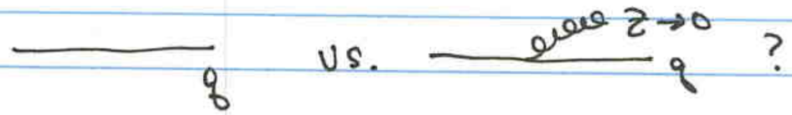
$$\Rightarrow P(z, \theta) = \frac{\alpha_s C_F}{\pi} \frac{dz}{z} \frac{d\cos\theta}{1-\cos\theta}$$

Let's even go one step further and work in the small angle limit, $\theta \ll 1$. Then,

$$P(z,\theta) \rightarrow \frac{\alpha_s C_F}{\pi} \frac{dz}{z} \frac{d\theta^2}{\theta^2}$$

This expression tells us a huge amount of physics. Note that the probability diverges when either $z \to 0$ or $\theta \to 0$, in the soft and/or collinear limits. It seems weird for a probability to diverge, but we just have to reinterpret it.

Consider, for example, the soft limit, $z \to 0$. If the energy of the gluon $E_g \to 0$, then what distinguishes that final state from just the quark, with no gluon?



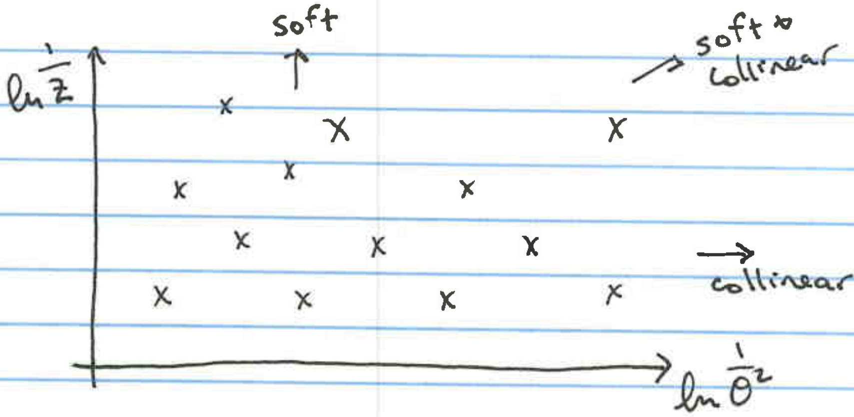
Is there a measurement we can do to distinguish these systems? The answer is no! They become degenerate in the $z \to 0$ limit. Indeed, Feynman diagram perturbation theory is degenerate perturbation theory, which is why the probability diverges in the $z \to 0$ limit. There is no measurement we can do to distinguish a system with no gluons, one 0-energy gluon, two 0-energy gluons, etc. Results and predictions in degenerate perturbation theory are only finite when we sum up <u>all</u> degenerate states. We will see how to do this in a second. As $z \to 0$, we should not interpret $P(z, \theta^2)$ as a probability, but rather as an expectation value of the number of soft/low energy gluons emitted from the quark.

Similar arguments follow for the collinear limit, $\theta^2 \to 0$, but I won't discuss that in detail.

Let's rewrite the probability in an enlightening way:

$$P(z,\theta^2)\, dz\, d\theta^2 = \frac{\alpha_s C_F}{\pi} \frac{dz}{z} \frac{d\theta^2}{\theta^2} = \frac{\alpha_s C_F}{\pi} d(\ln z)\, d(\ln \theta^2)$$

That is, emissions of soft/collinear gluons are uniformly distributed in the $(\ln z, \ln \theta^2)$ plane! There's a very nice way to visualize this, in what is called a "Lund diagram." This is:



Here, each x denotes another gluon emission off of the quark, and the emissions x are uniformly distributed in the plane. This is a semi-infinite plane and depending on how we approach ∞, we are sensitive to a different singular limit. Moving vertically in the plane is the soft limit, horizontally is the collinear limit, and diagonally is the soft and collinear limit.

At this point, I should emphasize that this uniform distribution of emissions is special to our approximations. Including a running coupling, higher-order effects, hadronization (which cuts off this picture at some point), etc, will change this picture. Nevertheless, there is a sense in which all of those things are corrections to this simple picture. Additionally, filling out this plane is the goal of Monte Carlo parton shower programs, like Pythia and Herwig. They each employ different methods for doing so, but their fundamental goal is the same. See Andrzej's talk for more!

We could stop here, but I want to do a non-trivial calculation since we have set up this framework. Let's calculate the distribution of the ratio of the invariant mass of the quark-gluons system to its total energy:

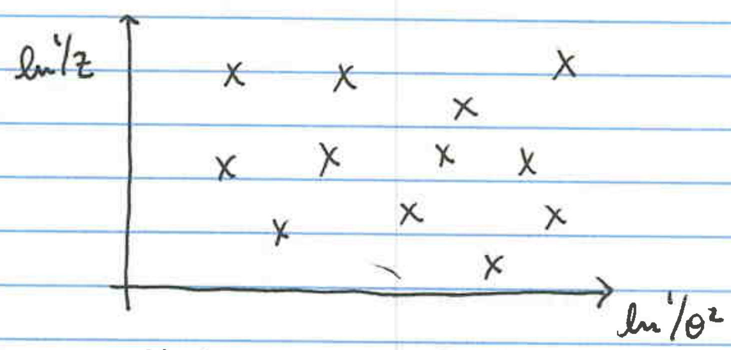
$$\tau = \frac{m^2}{E^2}.$$

In our phase space coordinates and with our assumptions, the observable τ is:

$$\tau = \sum_{i=\text{gluon}} z_i \theta_i^2 \,, \quad \text{where } z_i \text{ is the energy}$$

fraction of the i^{th} gluon and θ_i is the angle of the i^{th} gluon to the quark. (I use the symbol τ for this observable because it is identical to thrust in the soft and/or collinear limits.) The sum runs over all emitted gluons / x emissions in the $(\ln z, \ln \theta^2)$ plane. We will calculate the cumulative probability distribution, $P(x < \tau)$; that is, the probability the measured value of this observable is <u>less</u> than some value τ.

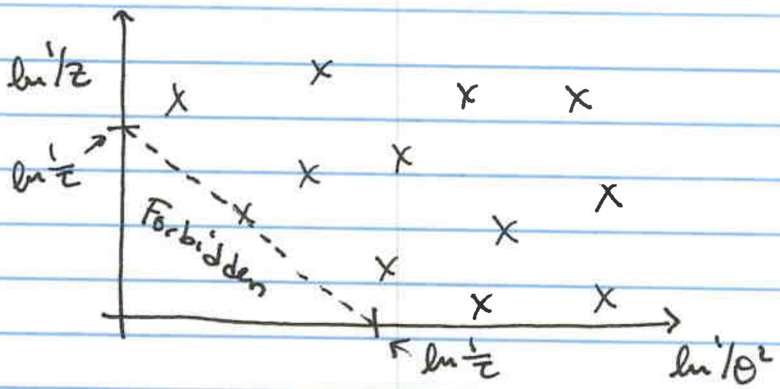
To do this, note that the emissions are uniformly distributed in $(\ln z, \ln \theta^2)$. This means that in "real" space (z, θ^2), emissions are <u>exponentially</u> far apart! This will help dramatically simplify our task. Because of this observation there is a <u>single</u> emission that dominates the value of τ, and all others provided tiny corrections. So, with emissions in the plane as:



there will be one that dominates the value of τ:

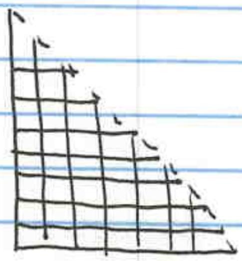
$\tau \approx z\theta^2$. Note that a fixed value of τ on this plane corresponds to a line:

$\ln \tau = \ln z + \ln \theta^2$. This line then corresponds to:



All emissions above the line are tiny corrections, there is one emission on the line, and no emissions below the line. If there were emissions below the line, then the measured value of τ would have increased. So, for calculating the cumulative probability, we must calculate the probability that there were no emissions below the line.

This probability is easy to calculate. We can imagine breaking up the forbidden triangle into many tiny regions:



The probability for emission into any one of those regions is proportional to the area of the region:

$$P(\text{emit in region } i) = \frac{\alpha_s C_F}{\pi} \cdot (\text{Area of region } i)$$

Therefore, the probability of no emission is 1 minus this:

$$P(\text{no emit in region } i) = 1 - \frac{\alpha_s C_F}{\pi} \cdot (\text{Area of region } i)$$

If we break up the forbidden triangle into N equal-area regions, then the area of any one region is

$$\text{Area of region } i = \frac{\frac{1}{2}\ln^2\tau}{N},$$ because the area of

the triangle is $\frac{1}{2}\ln^2\tau$. Then, to forbid any emission in all regions, we multiply these probabilities together:

$$P(\text{no emissions}) = \left(1 - \frac{\frac{\alpha_s}{\pi}\frac{C_F}{2}\ln^2\tau}{N}\right)^N$$

Taking the limit as $N \to \infty$, this transmogrifies into an exponential:

$$P(\text{no emissions}) = \exp\left[-\frac{\alpha_s}{\pi}\frac{C_F}{2}\ln^2\tau\right].$$

This is just equal to the cumulative probability:

$$P(x<\tau) = \exp\left[-\frac{\alpha_s}{\pi}\frac{C_F}{2}\ln^2\tau\right]$$

Note that this is exponentially suppressed as $\tau \to 0$. This object is called the Sudakov form factor.

To find the probability distribution, we just differentiate:

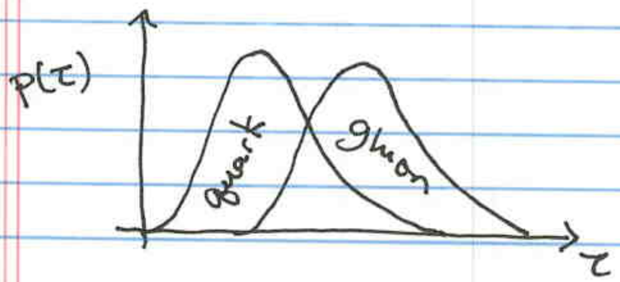
$$p(\tau) = \frac{d}{dt}\exp\left[-\frac{\alpha_s}{\pi}\frac{C_F}{2}\ln^2\tau\right] = -\frac{\alpha_s C_F}{\pi}\frac{\ln\tau}{\tau}\exp\left[-\frac{\alpha_s}{\pi}\frac{C_F}{2}\ln^2\tau\right].$$

We've tamed all the infinities! The Sudakov form factor is an explicit sum over all degenerate states with soft/collinear gluon emission. The probability distribution is finite, and in fact 0 for $\tau \to 0$.

Before concluding, I want to connect this to a fundamental problem in jet physics: discrimination of quark-initiated jets from gluon-initiated jets. We can perform the same exercise for gluon jets, and we find the cumulative distribution:

$$P_g(x<\tau) = \exp\left[-\frac{\alpha_s}{\pi}\frac{C_A}{2}\ln^2\tau\right].$$

The only change is replacing C_F by C_A, which is the color Casimir for the adjoint representation (the color carried by a gluon). Schematically, the distributions of τ for quark and gluon jets look like:



The ratio between the average values of these distributions is controlled by the ratio of C_A to C_F.

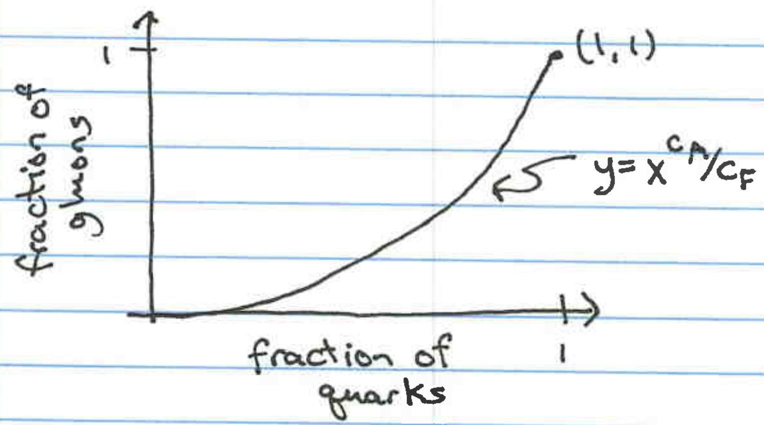
To separate quarks from gluons, we can make a cut on τ, and only keep those events to the left of the cut. The fraction kept is just given by the appropriate cumulative distribution:

$$P_q(x<\tau) = \exp\left[-\frac{\alpha_s}{\pi}\frac{C_F}{2}\ln^2\tau_{cut}\right]$$

$$P_g(x<\tau) = \exp\left[-\frac{\alpha_s}{\pi}\frac{C_A}{2}\ln^2\tau_{cut}\right] = \left(P_q(x<\tau)\right)^{C_A/C_F}$$

That is the fraction of gluons kept is found by raising the fraction of quarks kept to the C_A/C_F power!

We can nicely display this information in a receiver operating characteristic (ROC) plot:



This plot just displays the quark versus gluon efficiencies with this cut. In QCD, $C_F = 4/3$ and $C_A = 3$ and so the ROC curve for quark/gluon discrimination is:

$$y = x^{9/4}.$$

This can be improved somewhat by designing better observables or including higher-order effects, but is a benchmark for expectation.

We've gotten a lot of mileage out of our two simple axioms!