Lecture 4: A Case Study in Jet Substructure

The last decade has seen a transformation in how we think about jets and jet substructure. Many clever observables have been developed to improve experimental robustness/performance and enable theoretical control. There is a growing catalog of precise jet measurements/calculations, and with new machine learning techniques, no signs of a slow down.

I want to end these lectures with a classic but simple calculation that uses the ingredients you've learned. It is a jet discrimination task that has a history going back almost four decades.

Quark Jets vs. Gluon Jets.

At a cartoon level, this problem is simple.

q  $C_F = 4/3$

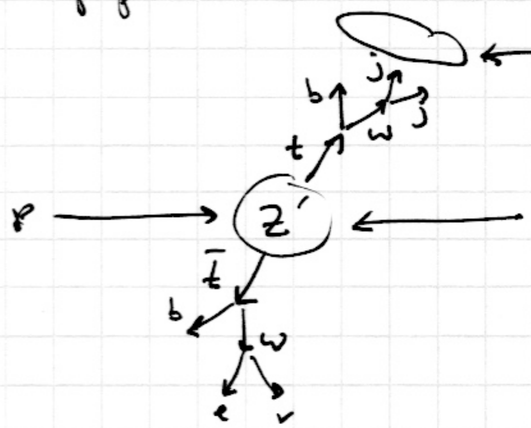
g  $C_A = 3$

Gluons have $9/4$ more radiation than quarks.

Putting aside questions about the intrinsic definition of "quark jet" and "gluon jet", I want to show you a calculation of how well you can discriminate these categories by measuring the jet mass. (Strictly speaking, a 2-point correlator.)

For additional context, jet classification became a hot topic once the new physics we were searching for became much heavier than $m_{W/Z/H/t}$.

E.g. $pp \to Z' \to t\bar{t}$

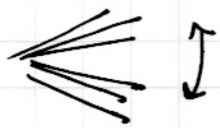


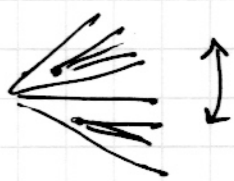
decay products of top contained in $\Delta R \sim \frac{2m_t}{p_T}$

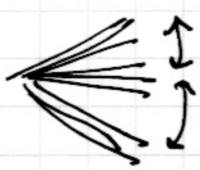
$m_{Z'} \uparrow \Rightarrow p_T^{top} \uparrow$

So eventually, hadronic top decay product's were reconstructed as single 3-prong "fat jet"

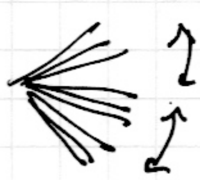
By now, boosted jet tagging is standard.

W/Z  2 prong

Higgs  double b-tagged fat jet

top  3 prong. b-tag

light Z'  2 prong again, different mass

RPV gluino/neutralino  3 prong again, different mass.

In this context, quark/gluon discrimination is particularly challenging, since no "topological" distinction. Nevertheless, it is highly relevant for BSM searches.

E.g. gluino cascade decays

$$pp \to \tilde{g}\,\tilde{g}$$

$\tilde{g} \to q\,\tilde{q}$, $\quad \tilde{q} \to q\,\tilde{\chi}^0$

$\tilde{g} \to q\,\tilde{q}$, $\quad \tilde{q} \to q'\,\tilde{\chi}^+$, $\quad \tilde{\chi}^+ \to W^+\,\tilde{\chi}^0$, $\quad W^+ \to q\bar{q}'$

Example of a quark-enriched signal, where background can be gluon dominated.

We will focus on soft & collinear limit of QCD where FSR splitting probability is

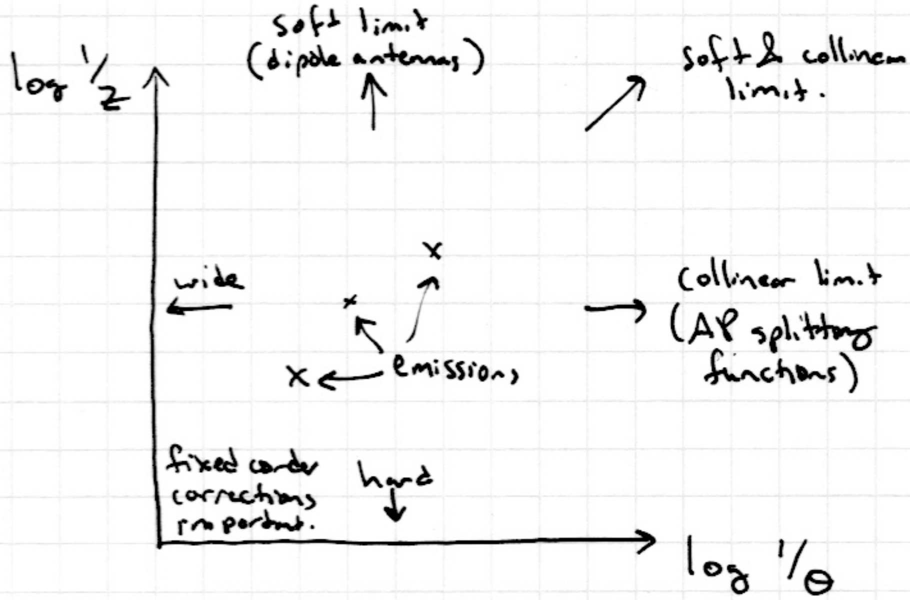
$$dP_{i\to ig} = \frac{2\alpha_s}{\pi}\, C_i\, \frac{dz}{z}\,\frac{d\theta}{\theta}$$

In this "strongly-ordered" (or "leading log" or "double log") limit, we can do simple but insightful calculations.

Very low accuracy, but gives a flavor for the types of questions you can ask / answer.

See SCET, direct resummation, ... for ways to systematically improve this picture.

Key insight: splitting kernel yields uniform emissions in $(\log 1/\theta, \log 1/z)$ plane (sometimes called Lund plane)



Recursive application of $dP_{i\to jk}$ is how parton shower algorithms work, essentially taking factorization to logical extreme.

This picture captures some information at <u>all</u> orders in α_s.

Often more realistic qualitatively than fixed-order methods.

Main Insight: Jet is not just a single parton

(Closer to Wilson-line-wrapped eikonal parton)

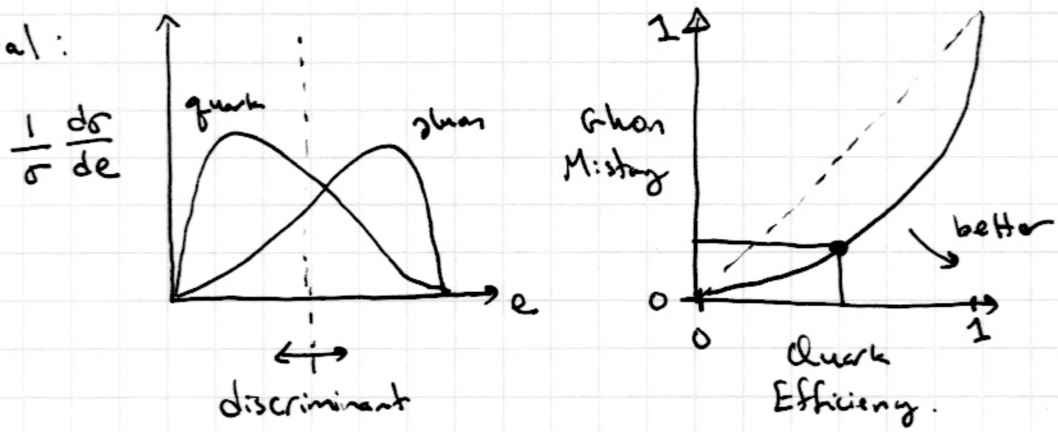
Very instructive to always keep in mind the soft gluon haze.

Concrete question: How well can you tag quark jets and reject gluon jets?

$C_F < C_A$, so gluon jets should be "fatter"

Need to define observables sensitive to this difference.

Goal:



Ideally we would:

- <u>Predict</u> from first-principles QCD
- <u>Validate</u> in Monte Carlo parton showers
- <u>Test</u> in LHC data

Choice of discriminant observable?

→ IRC safe (in order to perform perturbative calculations)

→ Relatively insensitive to hadronization and other nonperturbative effects

$$\delta e \sim \left(\frac{\Lambda_{QCD}}{p_T^{jet}}\right)^{\#}$$

Here is a simple one-parameter family of quark/gluon discriminants

Energy-Energy Correlation Function

$$e_2^{(\beta)} = \frac{\sum_{ij} p_{T_i} p_{T_j} R_{ij}^{\beta}}{\left(\sum_i p_{T_i}\right)^2 R^{\beta}}$$

e_2 ↑ 2-point correlator for 1-prong testing

R^β ↑ convenient to normalize by jet radius

Angular exponent β should satisfy $\beta > 0$ for IRC safety

(Also known as $C_1^{(\beta)}$.)

Quick check for IRC safety:

- Soft safe? Yes, e_2 unaffected by $p_{T_i} \to 0$
- Collinear safe? Yes, e_2 is additive so $p_T \to p_{T,1} + p_{T,2}$ has no effect (for $\beta > 0$)

What is quark/gluon discrimination power for e_2?

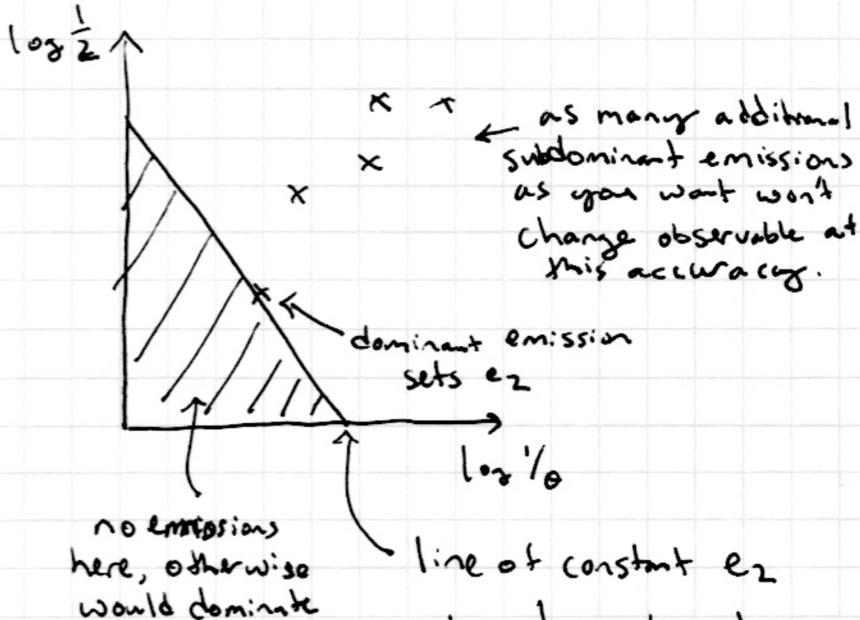
We will work in extreme soft & collinear & strongly-ordered limit of QCD. (This is borderline physical, but good for intuition.

$$e_2 \approx \frac{p_{T1}\, p_{T2}\, R_{12}^{\beta}}{\left(p_{T1} + p_{T2}\right)^2 R^{\beta}}$$

$$\approx z\, \theta^{\beta} \qquad \text{for } p_{T2} \ll p_{T1}$$

with $z = \dfrac{p_{T2}}{p_{T1} + p_{T2}}$ $\quad \theta = \dfrac{R_{12}}{R}$

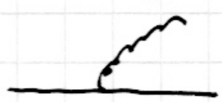
We can plot this observable on the $(\log 1/\theta, \log \frac{1}{z})$ plane.



$$\log \frac{1}{e_2} = \log \frac{1}{z} + \beta \log \frac{1}{\theta}$$

Recall: $dP_{i \to ig} = \frac{2\alpha_s}{\pi} C_i \frac{dz}{z} \frac{d\theta}{\theta}$, so we just have to compute probability to get specific value of e_2, right?

But wait, $dP_{i \to ig}$ is for real emissions. Where are virtual diagrams?

Real: 

at finite z, θ

Virtual: 

effectively at $z \to 0, \theta \to 0$
$(\log \frac{1}{\theta}, \log \frac{1}{z}) \to (\infty, \infty)$

Because our observable is IRC safe, we know divergences have to cancel, but how does this work?

Easiest way to think about this is in terms of probabilities

$$P_{emit} + P_{no\text{-}emit} = 1$$

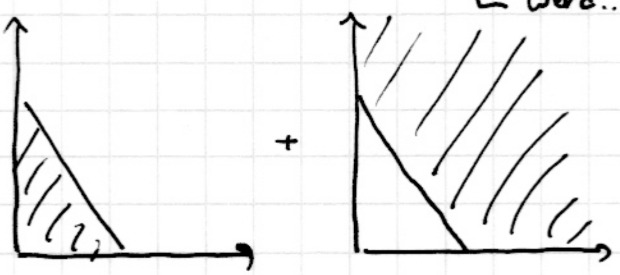
At Born order:

$$\mathcal{O}(\alpha_s^0) \qquad P^{(0)}_{noemit} = 1 \qquad \xrightarrow{\text{just}}$$

At next order:

$$\mathcal{O}(\alpha_s^1) \qquad P^{(1)}_{emit} + P^{(1)}_{no\text{-}emit} = 0$$

weird...



has to equal zero!

Now we just have an exercise in probability!

Chance to get any value of e_2 <u>less</u> than e_2^{max} is

$$\Sigma_q(e_2^{max}) = 1 + $$  $$+ \frac{1}{2}$$  $$^2 + \ldots$$

for quark · no emission at $\mathcal{O}(\alpha_s^0)$ · all emissions below e_2^{max} at $\mathcal{O}(\alpha_s^1)$ · symmetry factor · both emissions below e_2^{max} at $\mathcal{O}(\alpha_s^2)$

Then, using fact that  = − , we have:

real plus Virtual · just real

$$\Sigma_q(e_2^{max}) = \int_0^{e_2^{max}} p(e_2)\, de_2 = \exp\left[-\frac{2\alpha_s}{\pi} C_F \left(\text{area of}\right.\right.$$  $$\left.\left.\right)\right]$$

Probably most of you have never done a field theory calculation like that before!

Final result:
$$\Sigma_q\left(e_2^{max}\right) = \exp\left[-\frac{\alpha_s}{\pi}\frac{C_F}{\beta}\log^2\frac{1}{e_2^{max}}\right]$$

called Sudakov form factor

double logarithmic because of soft & collinear singularities.

For gluons, just swap $C_F \to C_A$.

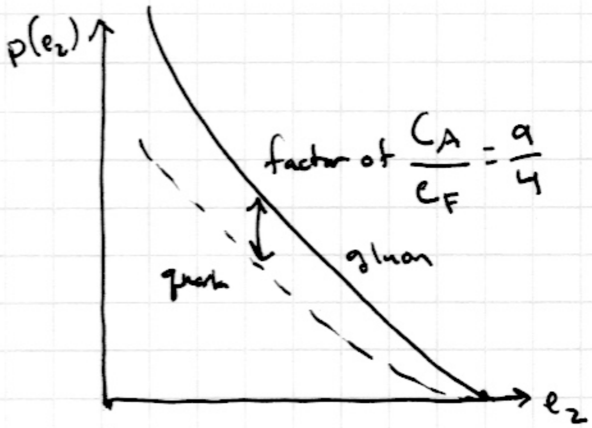
This is a (baby) example of a resummed calculation where you capture some information to all orders in α_s.
(can be systematically improved using approaches like SCET)

To go from cumulative distribution Σ to probability distribution, just take a derivative.

$$p(e_2) = \frac{1}{\sigma}\frac{d\sigma}{de_2} = \frac{\partial}{\partial e_2}\Sigma(e_2)$$

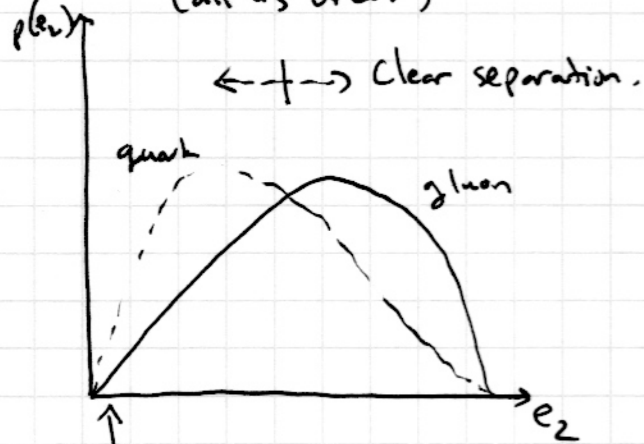
$$p_q(e_2) = \underbrace{\frac{2\alpha_s}{\pi}\frac{C_F}{\beta}\frac{1}{e_2}\log\frac{1}{e_2}}_{\text{at fixed order in } \alpha_s\text{, this goes singular as } e_2\to 0}\;\underbrace{\exp\left[-\frac{\alpha_s}{\pi}\frac{C_F}{\beta}\log^2\frac{1}{e_2}\right]}_{\text{at all orders in } \alpha_s\text{, you get Sudakov form factor that regulates singularity, giving much more physical result.}}$$

At fixed α_s order



↳ plus/delta function from virtual diagrams to ensure normalized distributions

In strongly-ordered limit (all α_s orders)



↳ Sudakov supression

Finally, we can compute discrimination power.

Place a cut at e_2^{cut}

$e_2 < e_2^{cut} \Rightarrow$ "quark"

$e_2 > e_2^{cut} \Rightarrow$ "gluon"

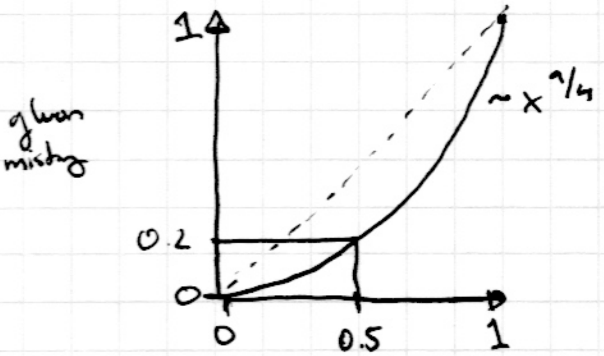
We want to know gluon mistag rate $\left(\Sigma_g(e_2^{cut})\right)$ as a function of quark efficiency $\left(\Sigma_q(e_2^{cut})\right)$.

At this order, we have simple "Casimir scaling"

$$\underset{\text{gluon}}{\Sigma_g} = \left(\underset{\text{quark}}{\Sigma_q}\right)^{C_A/C_F \to 9/4}$$

(This simple relation is violated at higher orders. Note that at this order, it is independent of β.)

Final mistag vs. efficiency curve:



Huge range of observables exhibit Casimir scaling, which is why quark/gluon discrimination is challenging.

But there are methods to achieve improved performance, which you can ask me about offline.

This calculation neglected many important higher-order effects.

- → Multiple Emissions
- → Color Coherence
- → Subleading terms in AP splitting function
- → Fixed-order corrections
- → Running α_s
- → Non-global logarithms
- → Hadronization Effects
- → Underlying event contamination
- → ...

Mostly, I hope you've gained some appreciation of why we need to define observables and why it is beneficial to work to all orders in α_s.

Concluding Thoughts

Here are three things I want you to remember about QCD and collider physics.

1. Have to think about observables. Just knowing scattering amplitudes is not enough to make predictions.

   Right now, these is no "theory of observables", so you have to assess things case by case.

2. Factorization is crucial for making predictions.

   You can't predict where every pion goes. Need to choose observables that respect, e.g., PDFs, otherwise you can't (yet) make first-principles predictions

3. You see (quasi-)stable particles in your detector, not Standard Model states

   Good observables on hadrons yield good proxies for QCD partons. Jet algorithms are one way to construct such proxies.

   If you want to use perturbative proxies, need to use observables that are infrared and collinear safe.

Finally, I started these lectures with the master formula:

$$\sigma_{obs} = \frac{1}{2E_{cm}^2} \sum_{n=2}^{\infty} \int d\Phi_n \left| \mathcal{M}_{AB \to 12\cdots n} \right|^2 f_{obs}(\Phi_n)$$

Are there things you can do with collider data that don't fall into this framework?

Yes! There are data analysis techniques (e.g. from machine learning) that aren't based on histograms.

In many ways, frontiers of collider physics will involve thinking about the whole "space of measurements" and striking right balance between experimental robustness and theoretical calculability.

I hope I've given you some sense of the richness of QCD and collider physics in these lectures!