Using the substructure of jets

Davison E. Soper
University of Oregon

with help from Zoltan Nagy and Michael Spannowsky

CTEQ, Madison, July 2011

- This talk is about analysis methods that use jet substructure.
- I first briefly define three analysis methods: mass drop plus filtering, trimming, & pruning.
- Then I examine the structure in nature that these methods try to exploit and that future methods can use.

Signal and background

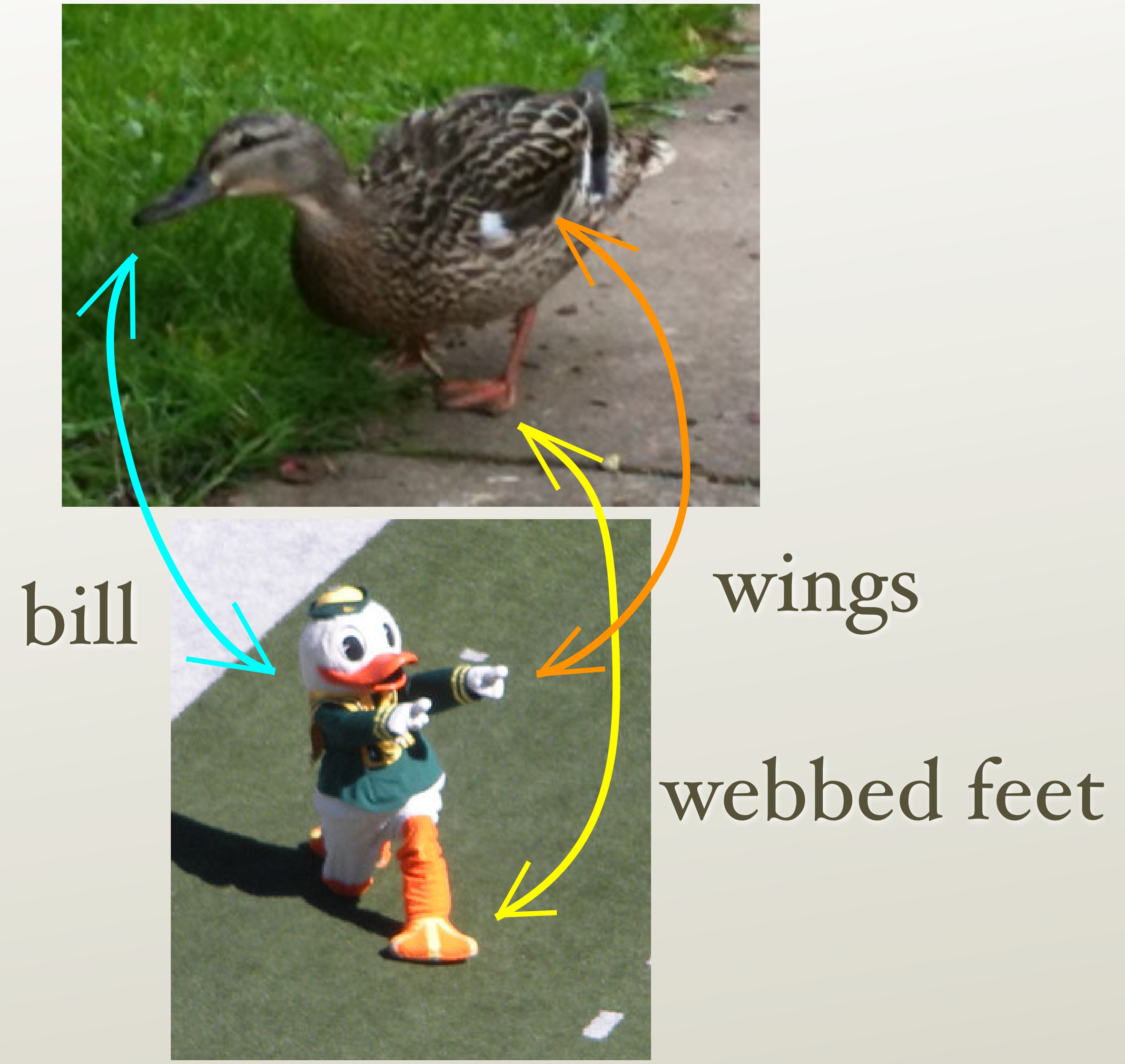
- We seek new physics signals in a standard model background.

- In Eugene OR, there are many background creatures.

- There are signal creatures that look a somewhat like background creatures, but there are very few of them.

- A signal creature.

- There are many features in common between signal and background creatures.
- We need to find the differences. (Eg. one of these can't swim.)



- Unfortunately, background creatures come with a range of mutations that make them sometimes look like a signal creatures.
- Thus we look need a statistical method to tell if there are signal creatures.

Signals and backgrounds at the LHC

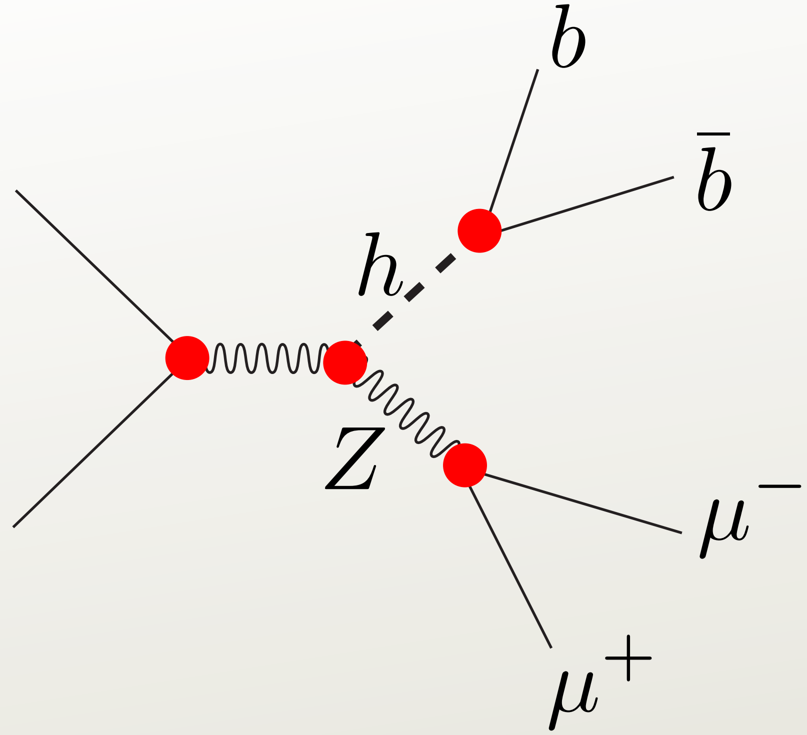
Signal & background have jets

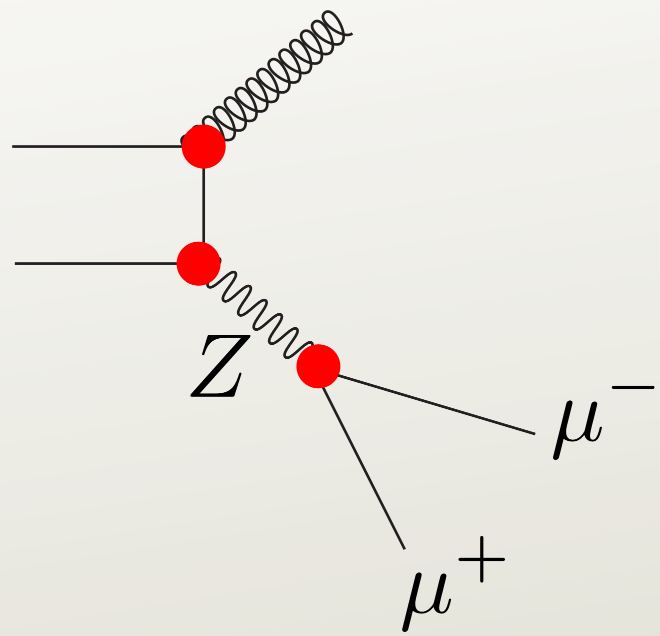
- Consider $p + p \to h + Z$.

$$h \searrow b + \bar{b}, \quad Z \searrow \ell^+ + \ell^-$$

Butterworth, Davison, Rubin, and Salam (2008)

- We demand that the Higgs have large p_T.
- That makes it easier to find.
- We may ask for b tags if we know where to look.
- Backgrounds: $p + p \to Z + jets$ and $p + p \to Z + Z$.
- The b and $\bar{b}$ are seen as jets.
- Background events have jets.





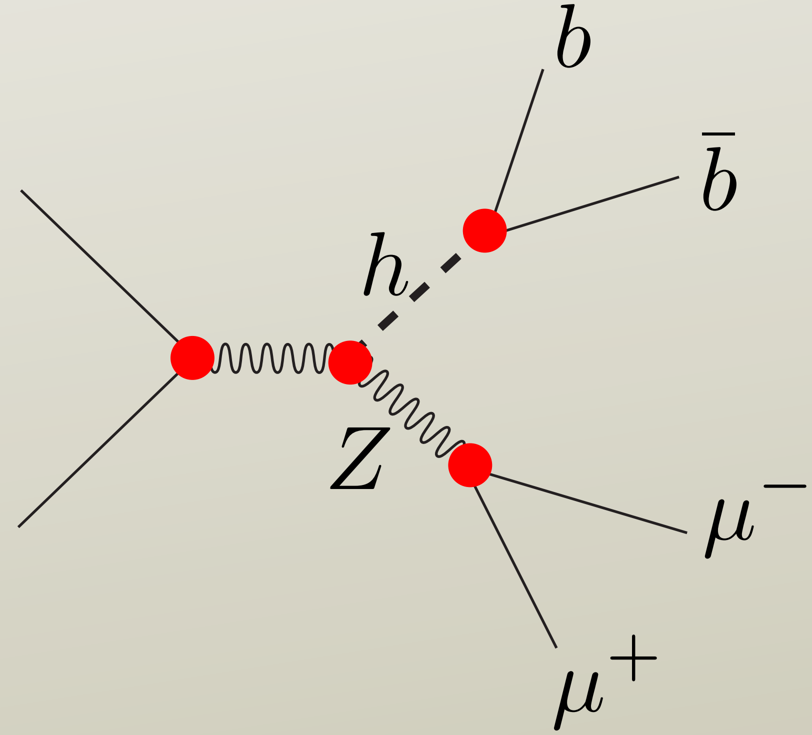
We want to find this. In a background of this.

- The gluon could split to two gluons, making the background look a lot like the signal.
- The gluon could split into $b\bar{b}$, making the background look even more like the signal.
- But the Higgs boson has a big mass, while a gluon has mass zero, so we can look for a mass bump.

Start with event selection

- Electron or muon pair near the Z mass.
- Large P_T of the lepton pair and of recoiling jet (>200 GeV).
- Large P_T implies that the possible Higgs decay products are easier to isolate: they are part of a (rather fat) jet.

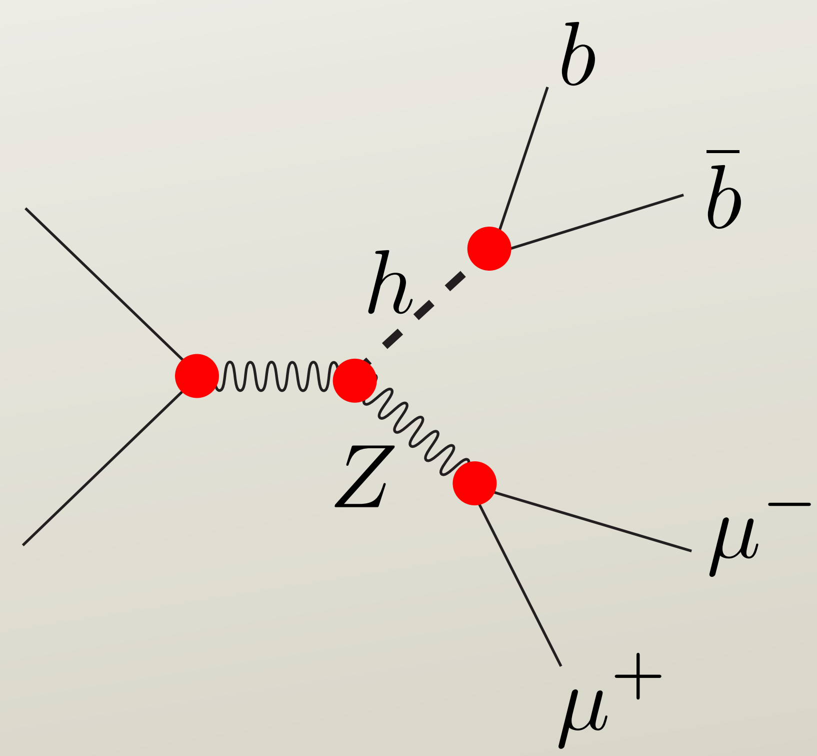
Butterworth, Davison, Rubin, and Salam (2008)

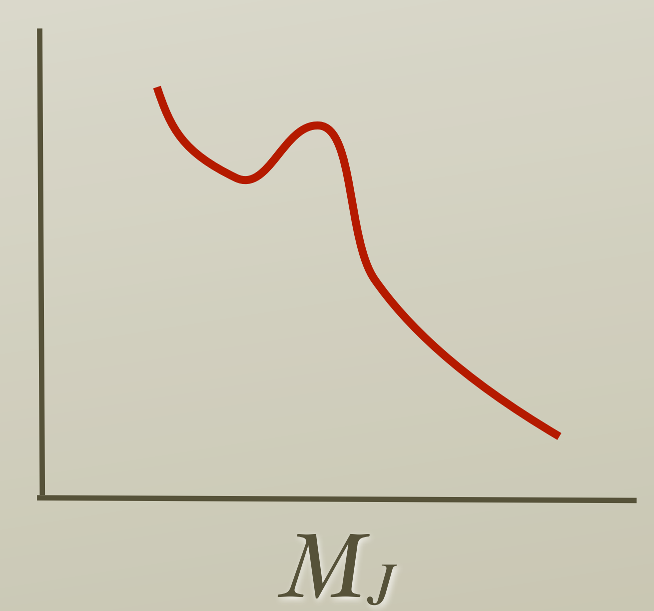


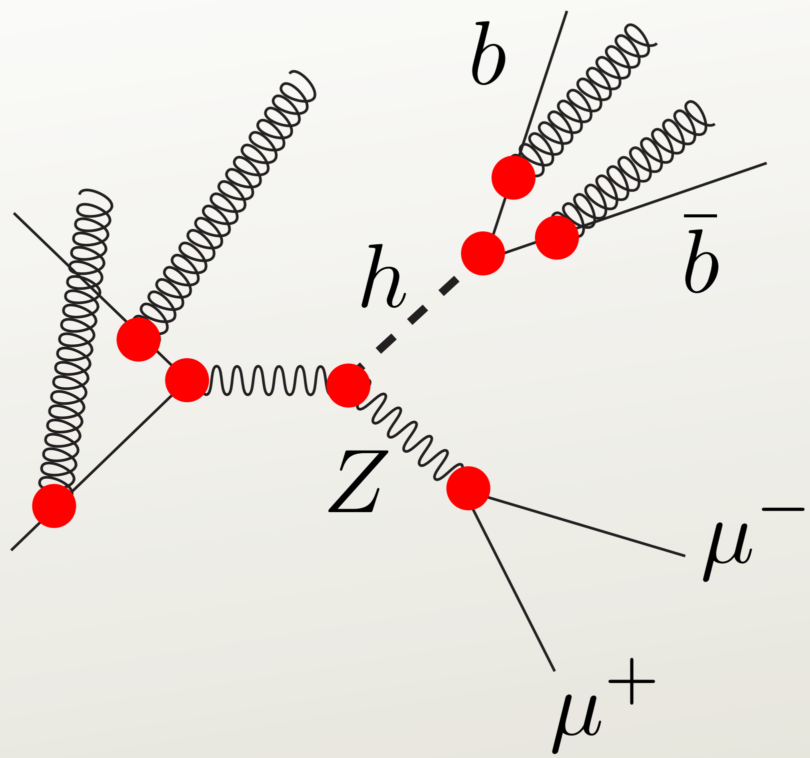
Define the fat jet

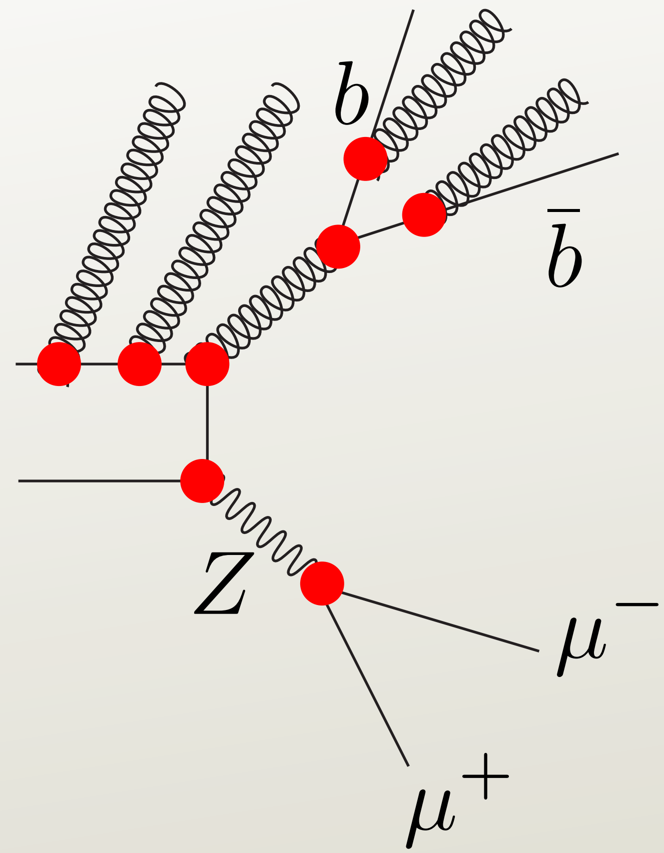
- Look for a high P_T jet using the Cambridge-Aachen (angle) algorithm with R=1.2.

- We might hope that the distribution of the mass of the fat jet shows a bump at the Higgs mass.

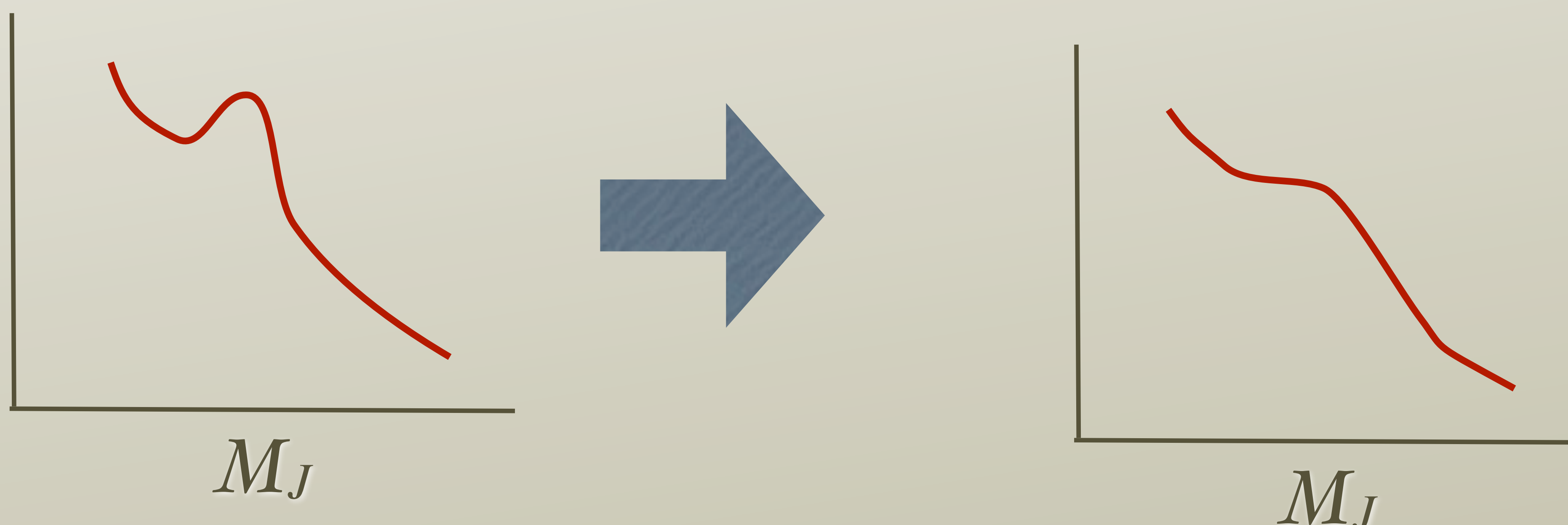






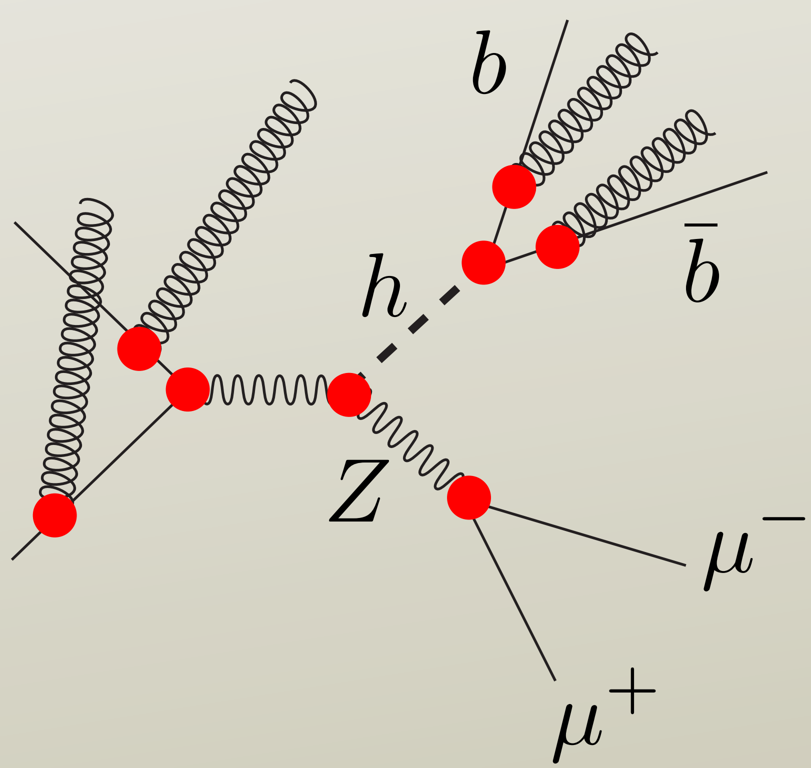


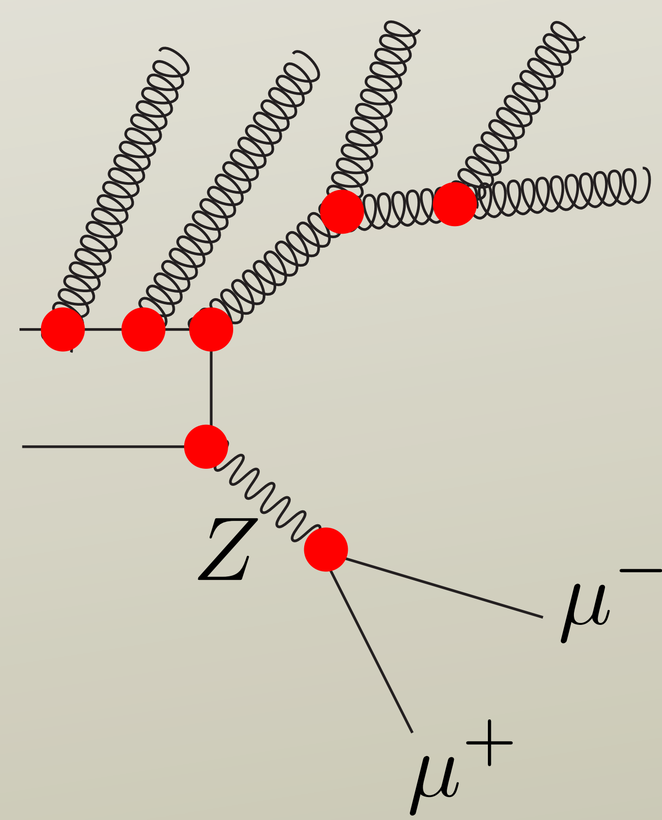
- Since QCD is operating, the mass bump gets smeared out.



Subjet analysis

- We would like to take apart the fat jet in order to get rid of the contaminating initial state radiation.



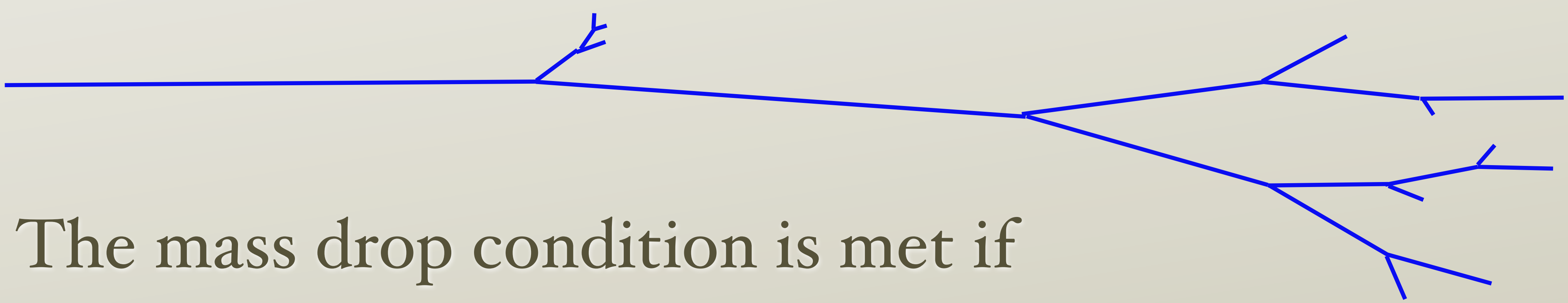


Saturday, July 9, 2011

Butterworth, Davison, Rubin, and Salam (2008)

Jet mass drop and filtering

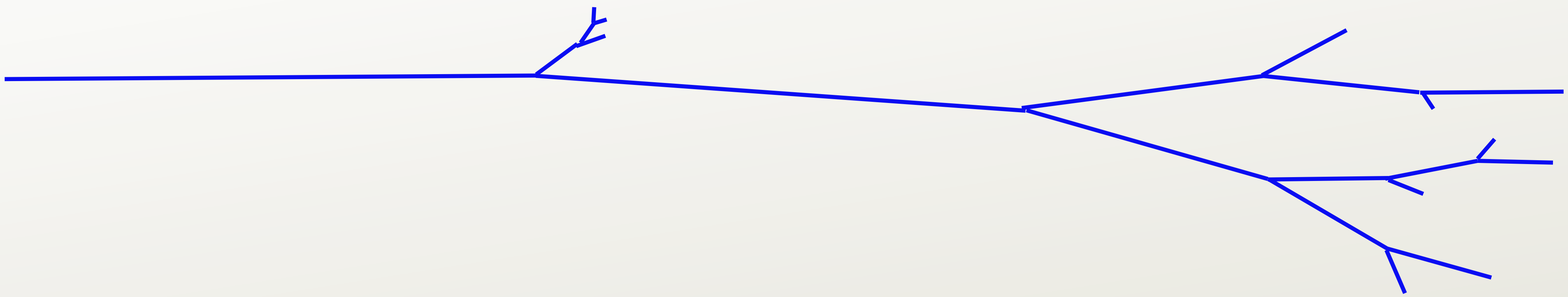
- Step I: mass drop.
  - Examine the C-A splitting tree, starting at the trunk.



  - The mass drop condition is met if

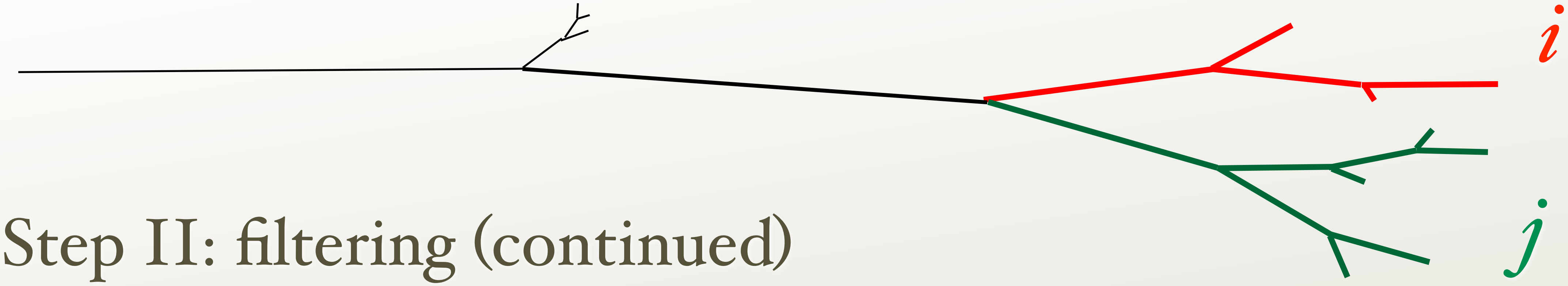
$$\max(M_i, M_j) < 0.67\, M_{\{i,j\}}$$

$$\min(p_{T,i}^2, p_{T,j}^2)[(y_i - y_j)^2 + (\phi_i - \phi_j)^2] > 0.09\, M_{\{i,j\}}^2$$



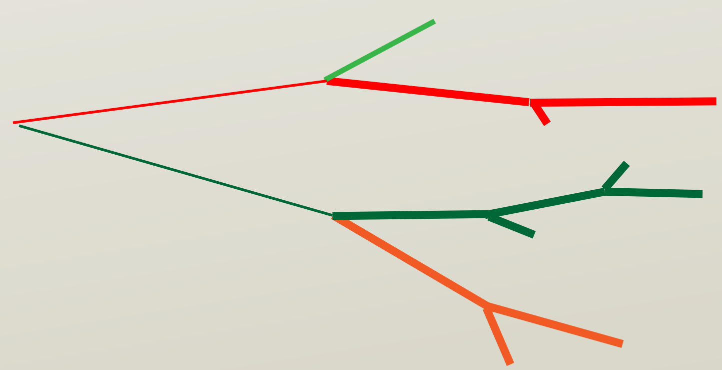
- If mass drop condition isn't met, drop smaller p_T daughter and keep looking.
- If it is never met, remove the event from your sample.

- Step II: filtering the prospective *b*-jets, *i* and *j*.
  - Are both prospective *b*-jets tagged as containing *b*-quarks?
  - (In simulating this, we assuming a *b*-tagging efficiency of 60% and a misstag probability of 2%)
  - If i and j are not *b*-tagged, reject the event.



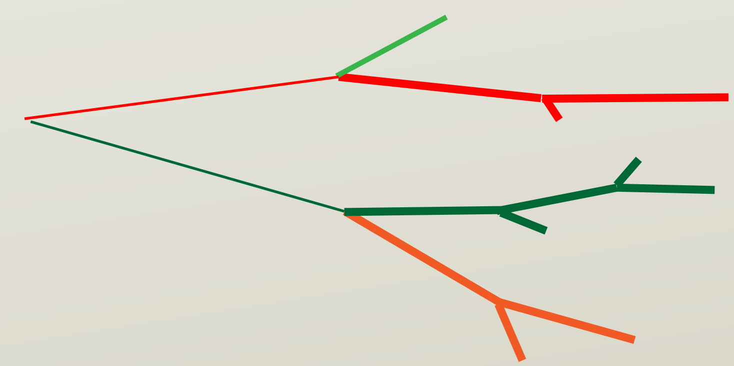
- Step II: filtering (continued)
  - Apply the C-A algorithm with to protojets *i* and *j* with

$$R = \min\left(\frac{1}{2}[(y_i - y_j)^2 + (\phi_i - \phi_j)^2]^{1/2}, 0.3\right)$$

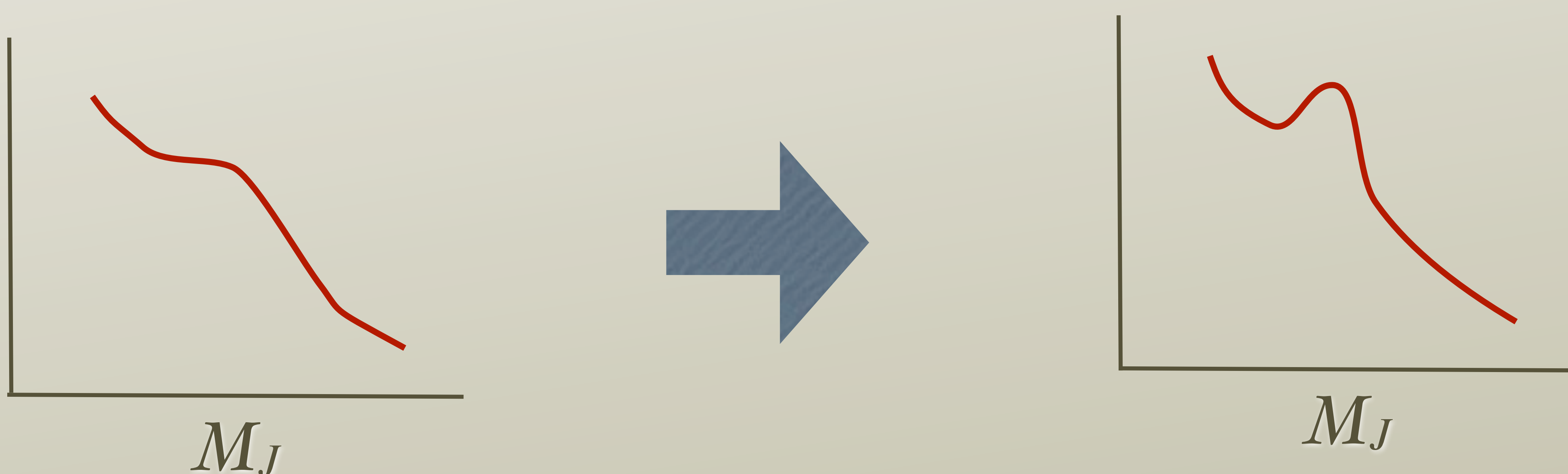


  - Are the two highest p_T subjets thus found tagged as containing b-quarks?
  - If not, throw out the event.

- Step II: filtering (continued some more)
  - Throw out all but the three highest p_T subjets thus found.
  - What remains is the filtered jet.



- Measure the mass of this filtered jet.



Krohn, Thaler, and Wang (2010)

Jet trimming

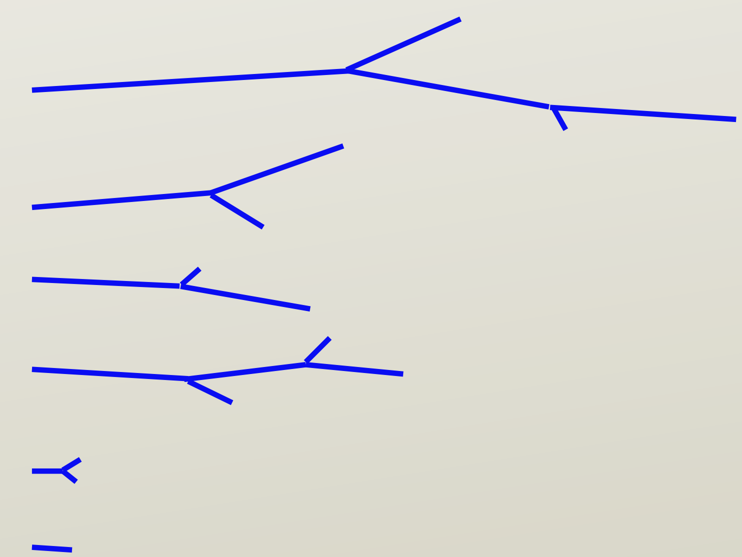
- Define the fat jet with the anti-k_T algorithm with R = 1.2.

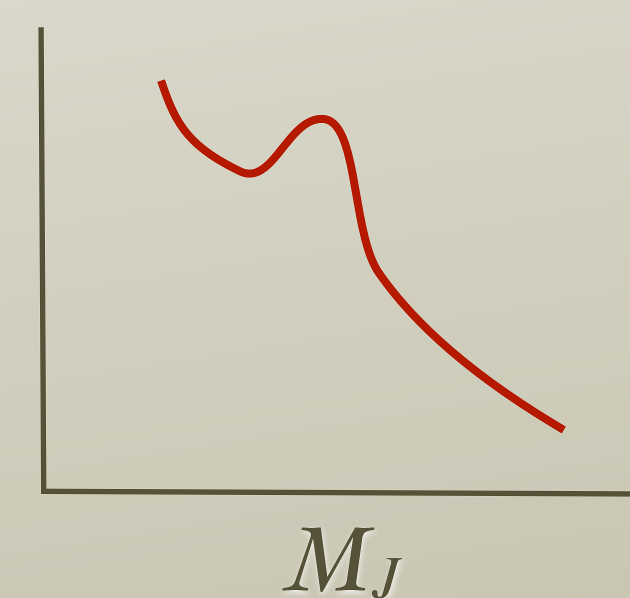
- Group starting protojets into successively fatter protojets with the k_T algorithm with R = 0.2.

- This gives several final jets, some with large p_T, some with small p_T.

- Now throw away the jets that are too soft. (Say, $p_T < 0.03\ P_T(\text{fat jet})$.)

- Measure the mass of this trimmed jet.





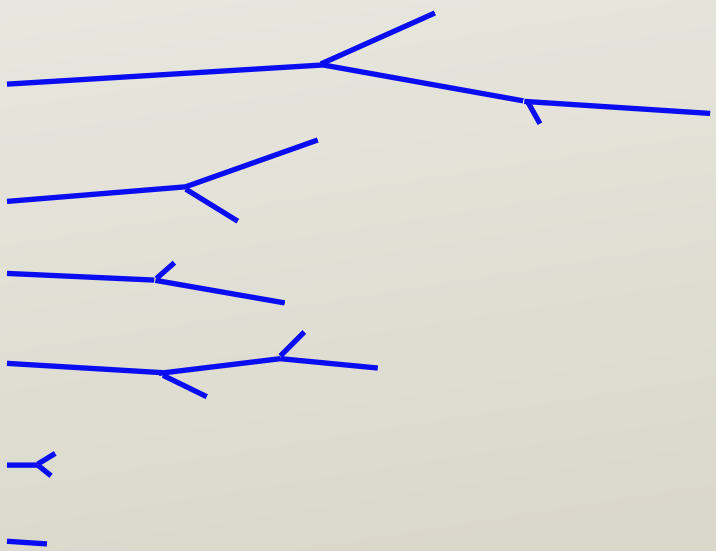
S.D. Ellis, Vermilion, and Walsh (2009)

Jet pruning

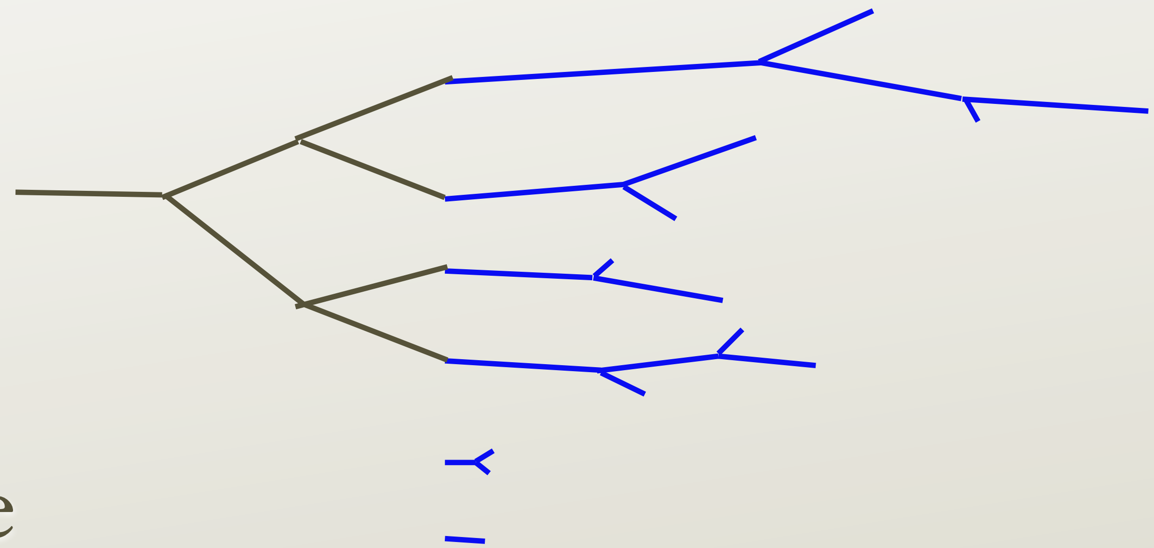
- Define the fat jet with the C-A algorithm with R = 1.2.
- Step I
  - Group starting protojets into successively fatter protojets with the C-A algorithm with

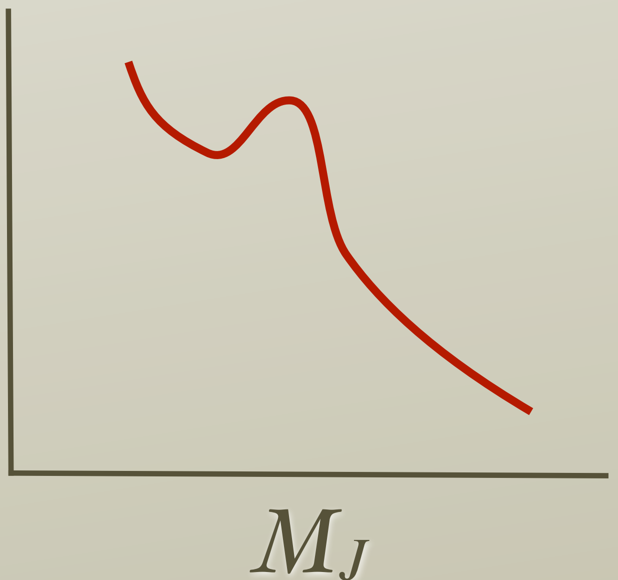
$$R = M(\text{fat jet})/P_T(\text{fat jet})$$

  - That is, merge protojets until no pair has an angle smaller than this.



- Step II
  - Continue merging up to larger angles. Examine
    $$z = \frac{\min(p_{T,i}, p_{T,j})}{|\vec{p}_{T,i} + \vec{p}_{T,j}|}$$
  - If $z < 0.1$, cancel the merge and omit the small p_T protojet.
  - Continue until all protojets are merged or eliminated.
- Measure the mass of this pruned jet.



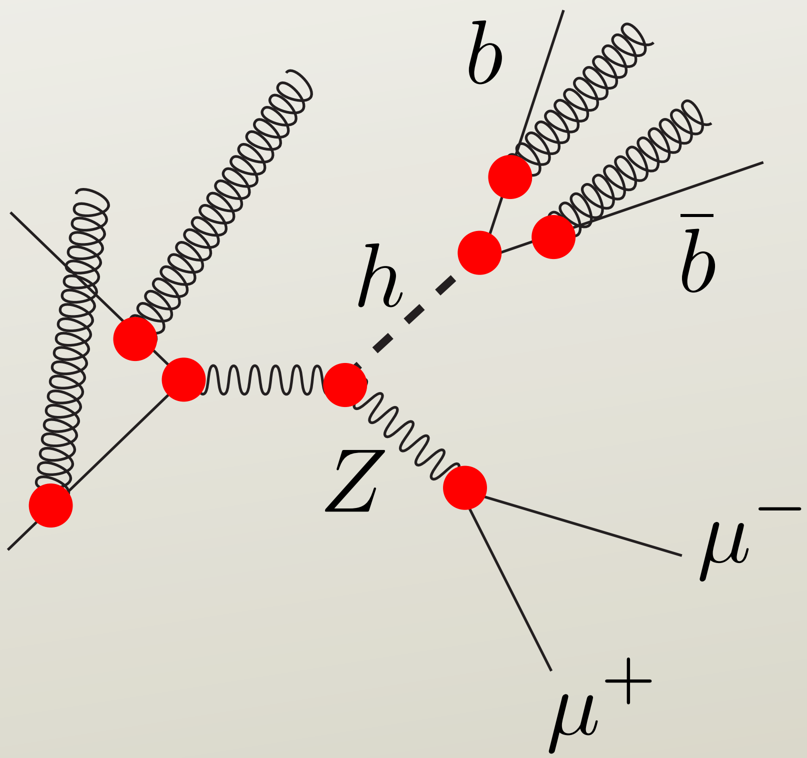


What gauge field theory has to say

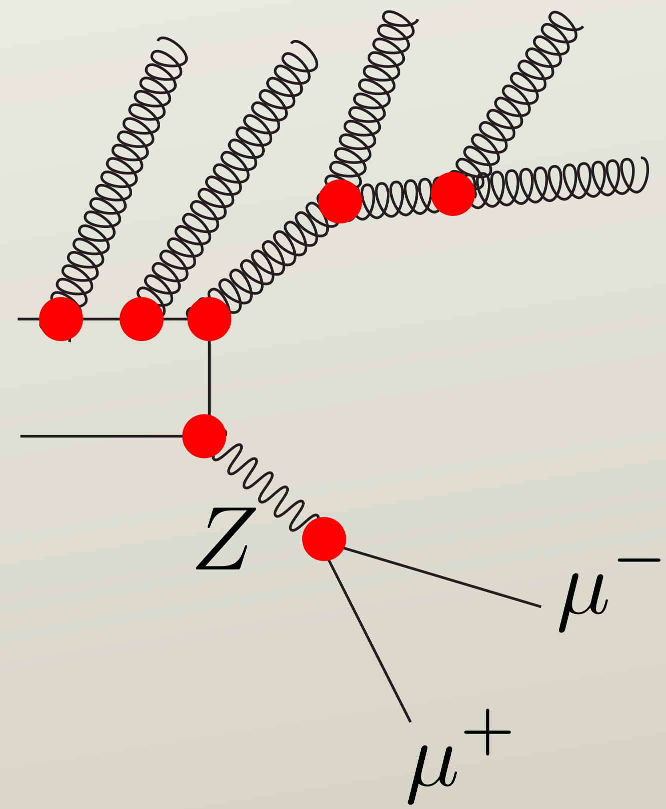
- One can use field theory to understand the characteristics of background events and the characteristics of signal events (with heavy particle decays).
- Schemes for selecting signal events depend on using the similarities and differences in signal and background events.
- I will try to outline the characteristics that can be used.

Initial state radiation

- One immediate similarity is that both signal and background events have initial state radiation.



A Higgs signal event



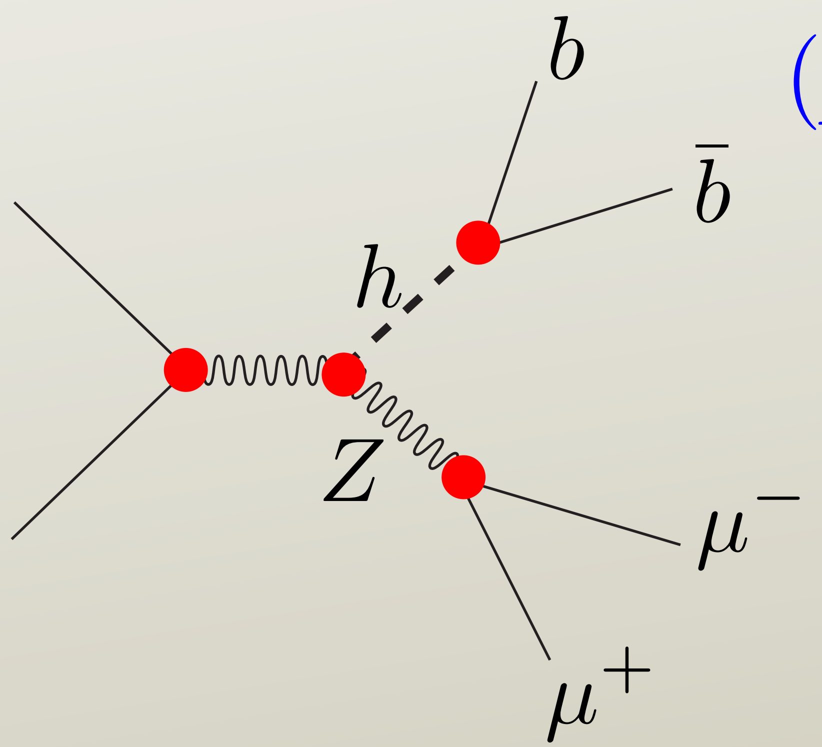
A similar background event

- This makes the signal/background separation difficult.

- More on this later.

Mass bumps

- One key feature of signal events is that the decay products of a heavy particle have a fixed mass.

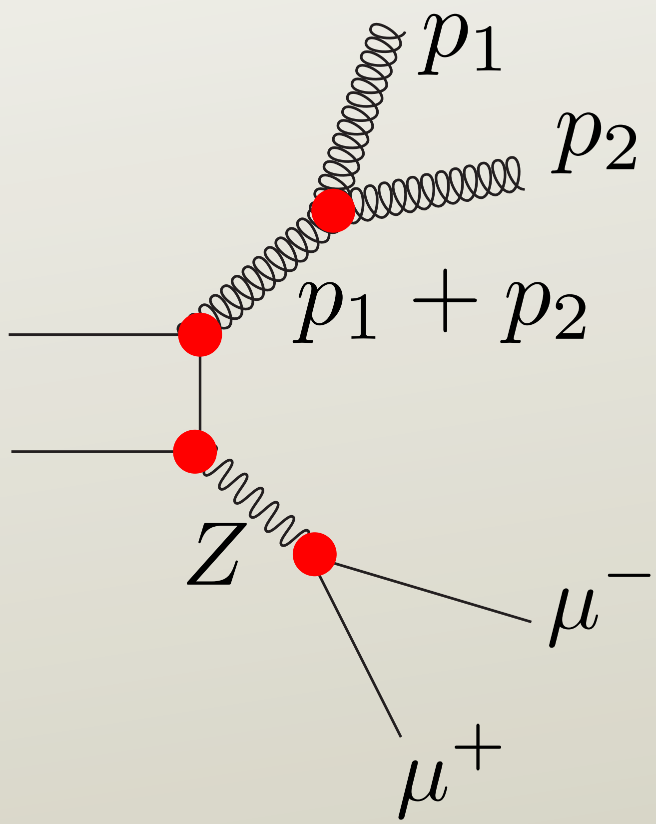


$$(p_b + p_{\bar{b}})^2 = M_h^2$$

- The theoretical uncertainty of this is $M_h \Gamma_h$.
- The practical uncertainty is much larger.

Mass in QCD splittings

- QCD splittings have a very different distribution as a function mass.



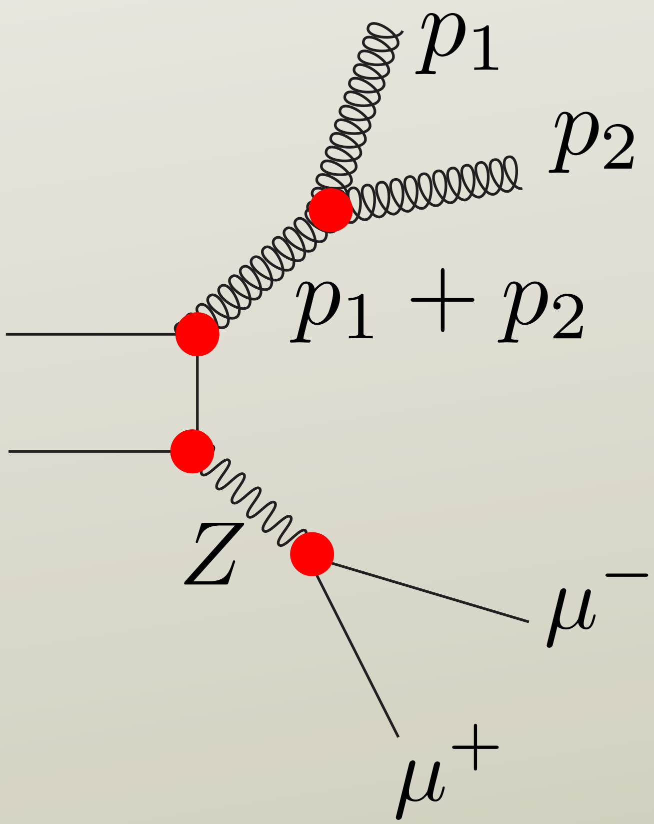
- I will use a scaled virtuality as a “hardness variable”:

$$v = \frac{(p_1 + p_2)^2}{(p_1 + p_2) \cdot Q_0}$$

- Q_0 is the momentum of the outgoing partons in the hard process.

QCD tree graph results

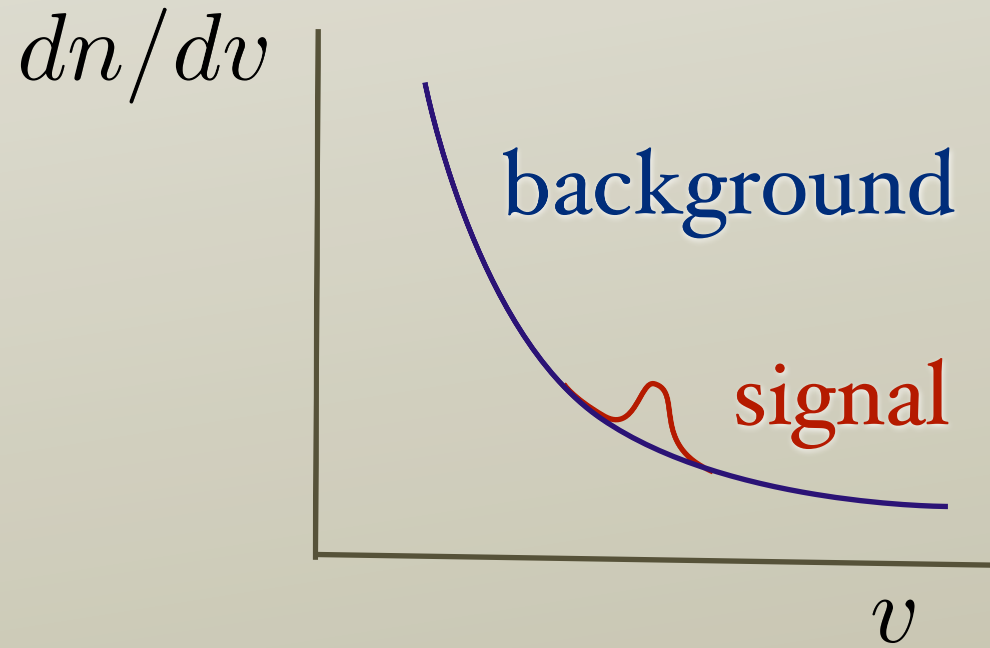
$$v = \frac{(p_1 + p_2)^2}{(p_1 + p_2) \cdot Q_0}$$



- Differential probability is

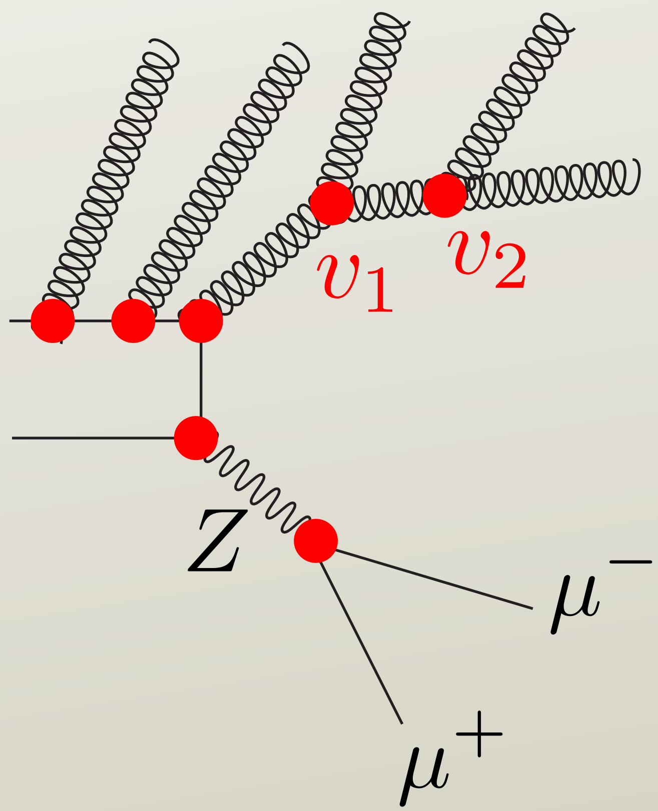
$$dn \propto \frac{dv}{v}\ \alpha_s \log(v/\tilde{v})$$

- The log (and $\tilde{v}$) comes from $\int dz$.



Nested splittings

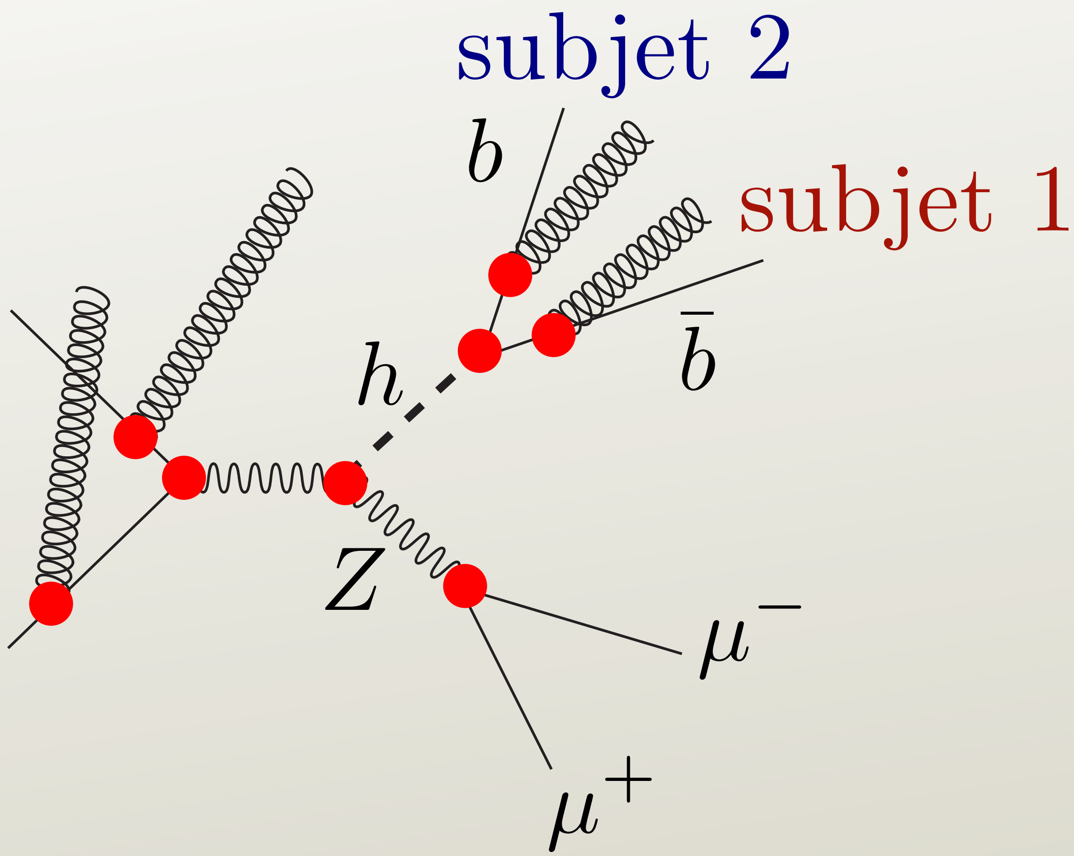
- The same structure can be (approximately) iterated.

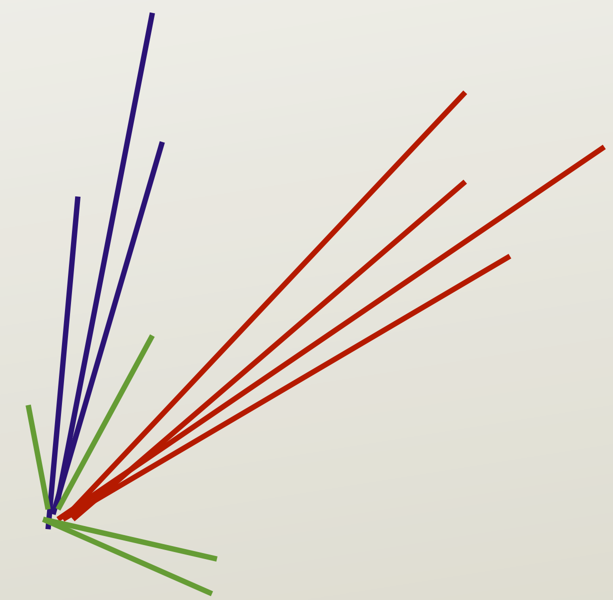


- $v_2 < v_1$.
- Set $v_2 = 0$ in splitting 1.

$$dn = \frac{dv_1}{v_1}\ \alpha_s \log(v_1/\tilde{v}_1) \times \frac{dv_2}{v_2}\ \alpha_s \log(v_1/\tilde{v}_2)\theta(v_2 < v_1)$$

An application of this



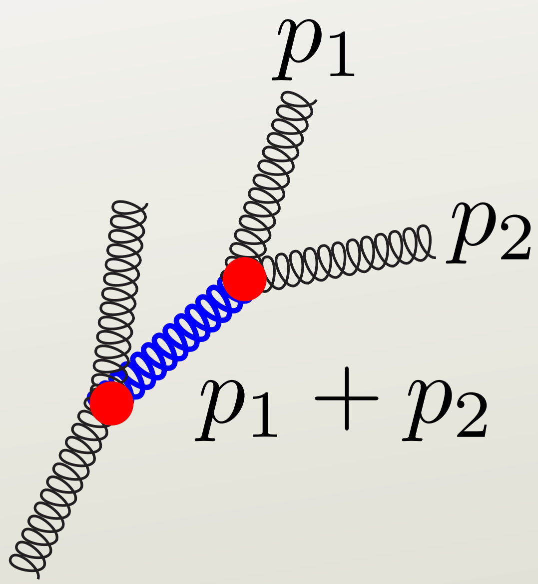


- A Higgs boson will likely decay to two low mass subjets.
- Try to find the low mass subjets.
- A "mass drop" condition can be part of this.

  (Butterworth, Davison, Rubin & Salam.)
- Measure the mass of subjet pair.

Sudakov factor

- A better approximation for dn includes a Sudakov factor.



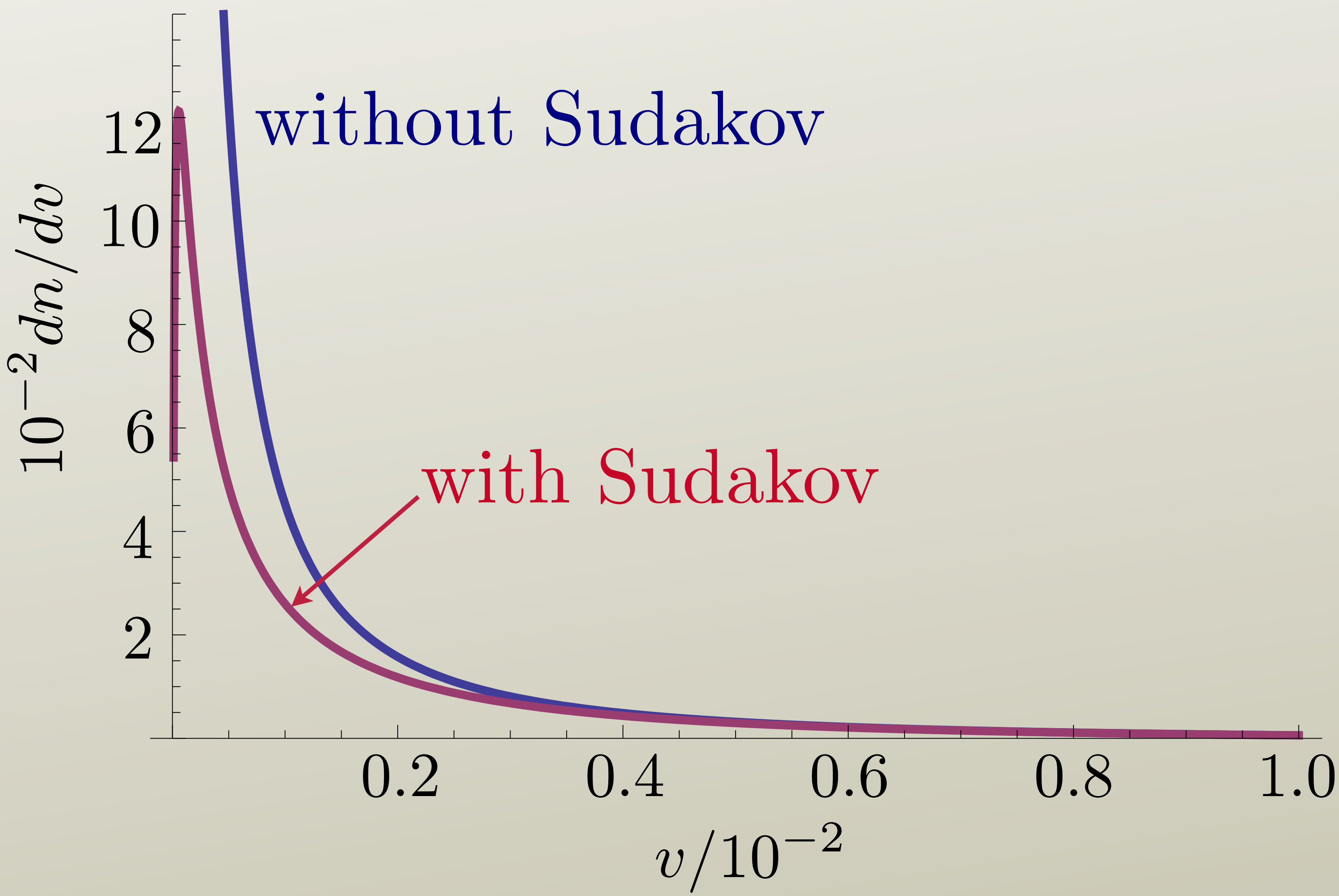
$$dn = \frac{dv}{v}\,\frac{\alpha_s}{2\pi}\,2C_{\rm A}\log(v/\tilde{v})$$
$$\exp\left(-\int_v^{v_1}\frac{d\bar{v}}{\bar{v}}\,\frac{\alpha_s}{2\pi}\,2C_{\rm A}\log(\bar{v}/\tilde{v})\right)$$

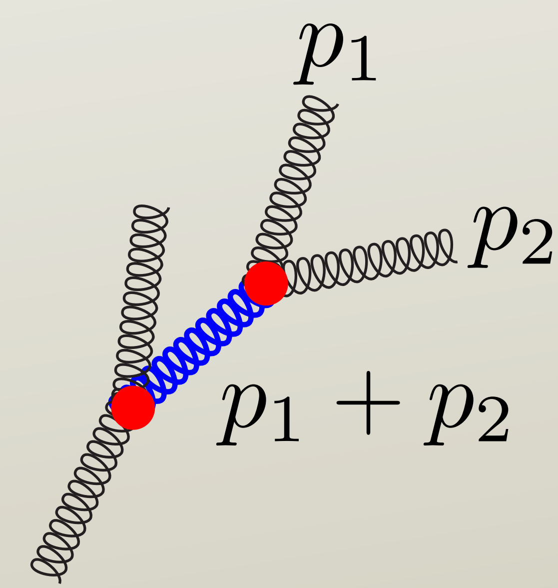
- We think of a renormalization group approach with a running resolution scale.

- The Sudakov factor accounts for virtual graphs and unresolved real parton emissions.

Example of effect of Sudakov

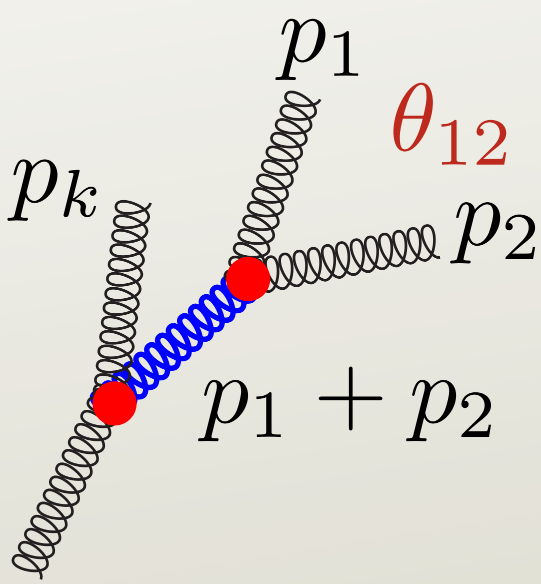
$$dn = \frac{dv}{v}\ \frac{\alpha_s}{2\pi}\, 2C_{\rm A}\log(v/\tilde{v})\exp\left(-\int_v^{v_1}\frac{d\bar{v}}{\bar{v}}\ \frac{\alpha_s}{2\pi}\, 2C_{\rm A}\log(\bar{v}/\tilde{v})\right)$$





$Q_0^0 = 1000$ GeV
$|\vec{p}_1 + \vec{p}_2| = 100$ GeV
$v_1 = 10^{-2}$
$\theta_{\max} = 0.3$

Distribution in angles



- For a given v, what is the distribution in the angle θ_{12} between $\vec{p}_1$ and $\vec{p}_2$?

- Kinematically,

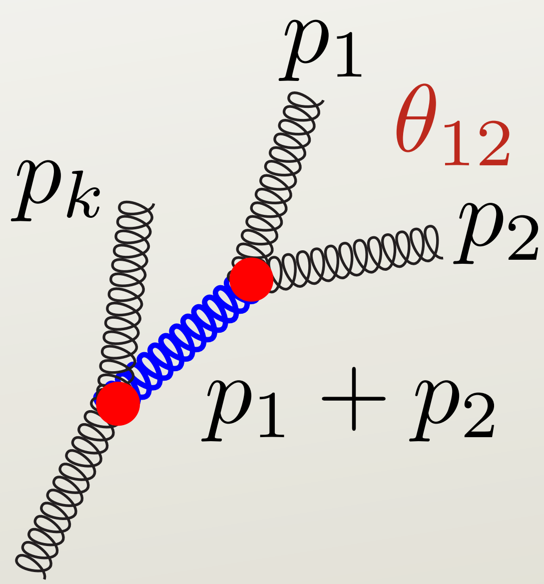
$$\theta_{12}^2 > v \, \frac{4Q_0^2}{(p_1 + p_2) \cdot Q_0}$$

- For larger, but not too large, θ_{12},

$$v = \frac{(p_1 + p_2)^2}{(p_1 + p_2) \cdot Q_0}$$

$$dn \propto \frac{d\theta_{12}^2}{\theta_{12}^2}$$

Large angles



- Large θ_{12} corresponds to $|\vec{p}_1| \ll |\vec{p}_2|$.
- (Or gluon 2 could be the soft one.)
- For emission of soft gluon 1, interference with emission from parton k is important.
- Parton k is the color connected partner.

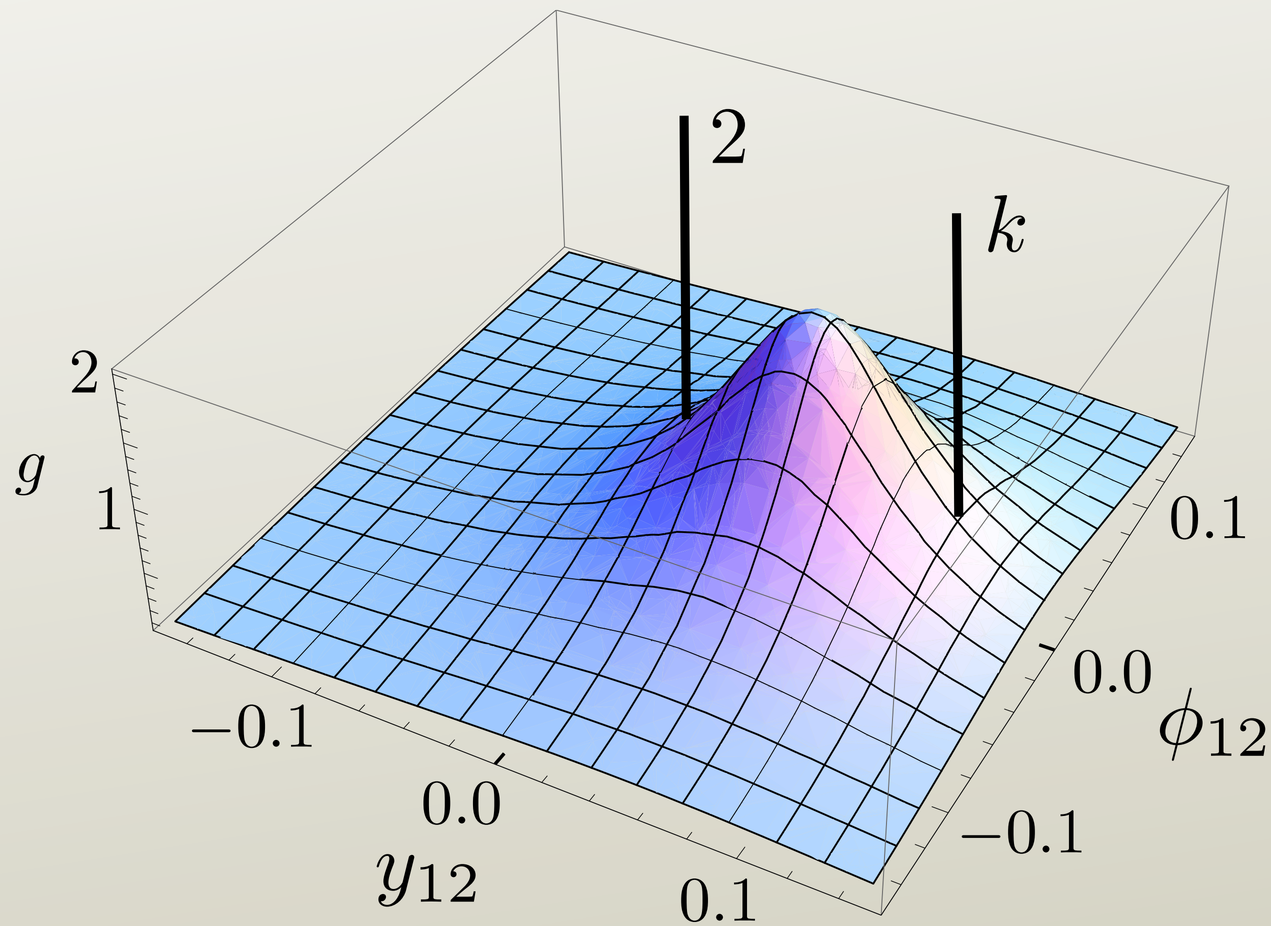
- With interference,

$$dn \propto \frac{d\theta_{12}^2}{\theta_{12}^2} g(\theta_{12}^2) \qquad g(\theta_{12}^2) = \frac{\theta_{2k}^2}{\theta_{12}^2 + \theta_{1k}^2}$$

Effect of the angle factor

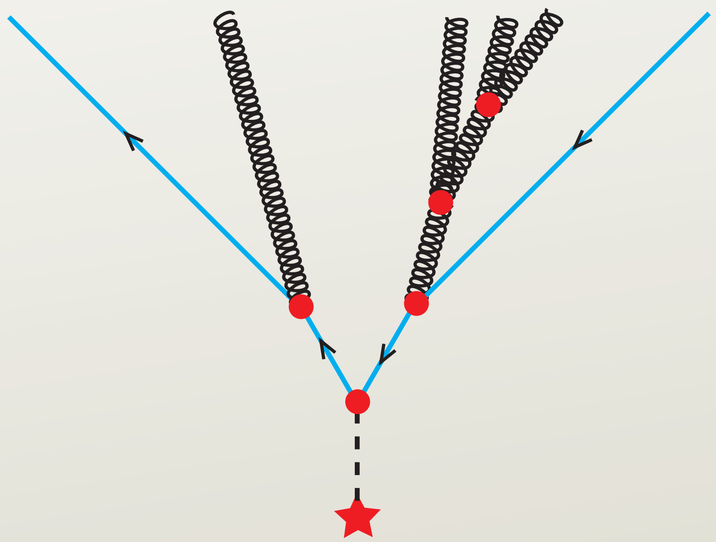
$$dn \propto \frac{d\theta_{12}^2}{\theta_{12}^2} g(\theta_{12}^2)$$

$$g(\theta_{12}^2) = \frac{\theta_{2k}^2}{\theta_{12}^2 + \theta_{1k}^2}$$



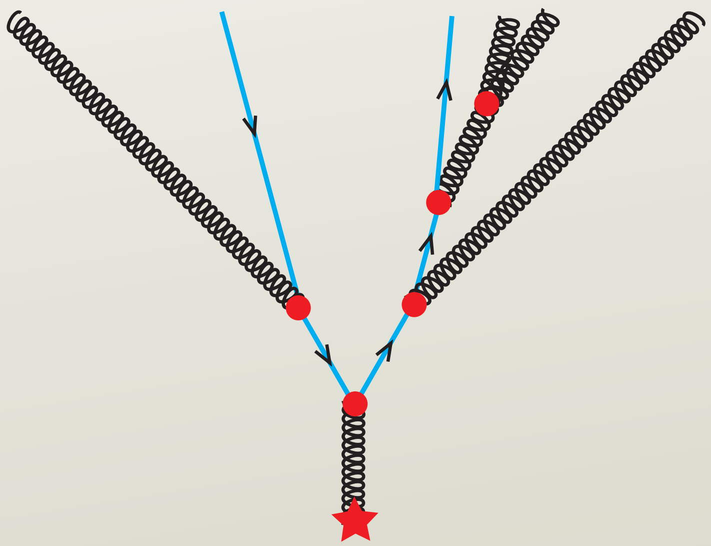
- Emission with $\theta_{12} \gg \theta_{2k}$ is suppressed.
- Emission between partons 2 and k is enhanced.

An application of this



After $H \to b + \bar{b}$

Soft gluons between b and $\bar{b}$



After $g \to b + \bar{b}$

Soft gluons away from b and $\bar{b}$

- *Cf.* "pull" (Gallicchio & Schwartz).

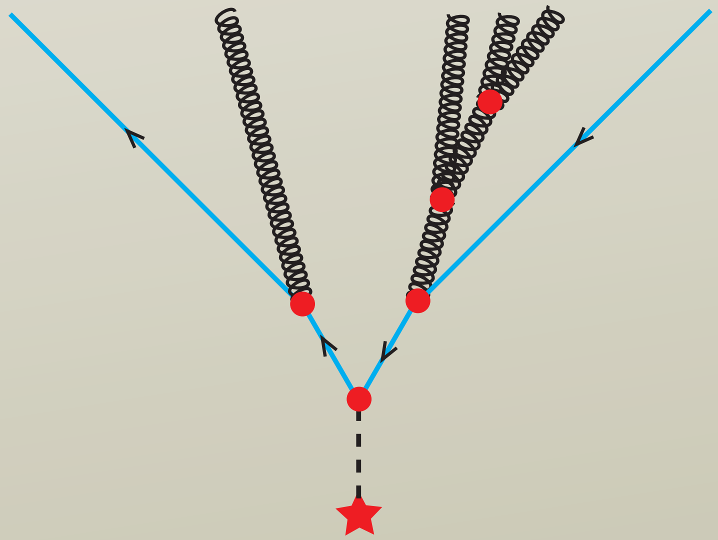
Gluons want to be soft

- Large angles $\implies$ small z:

$$z(1-z) = \frac{v}{\theta^2} \frac{Q_0^2}{(p_1+p_2)\cdot Q_0}$$

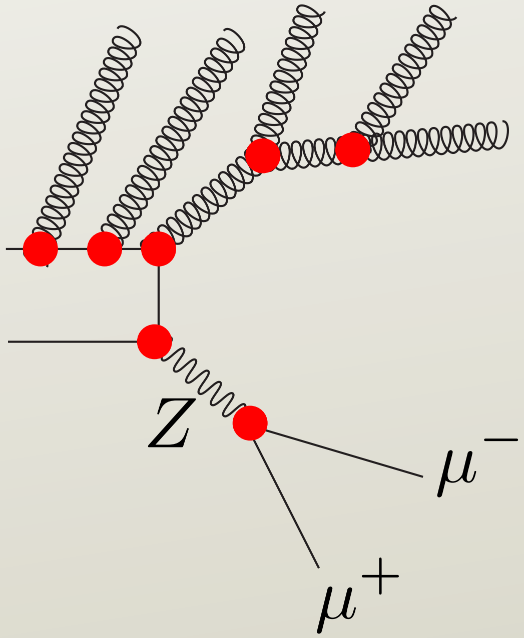
- So we expect lots of soft gluon radiation.

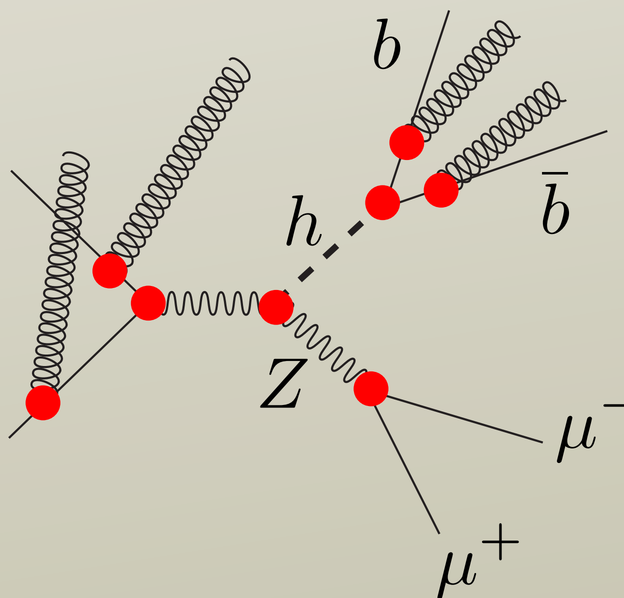
- However, soft gluon radiation is limited by "angular ordering" from $g(\theta)$.



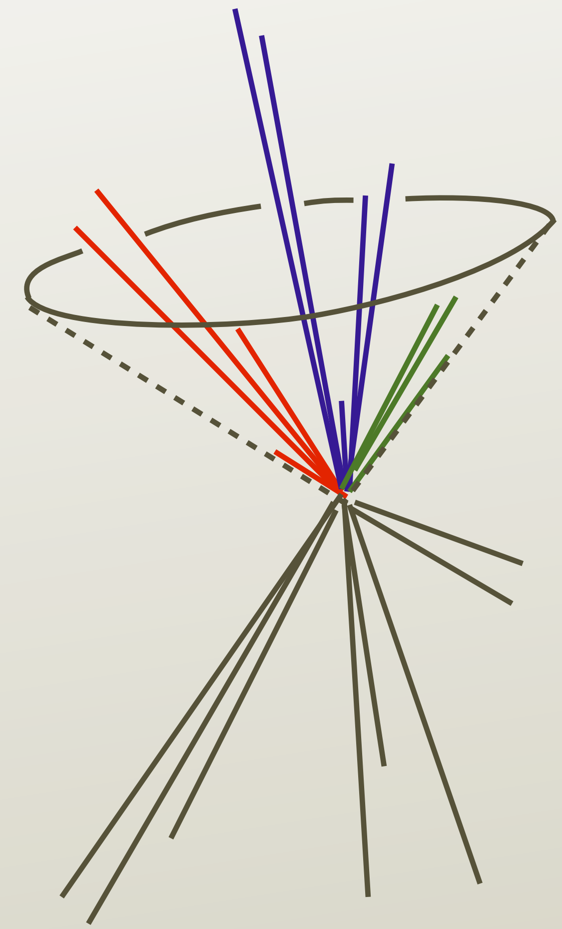
More soft gluons come from initial state radiation

- There is initial state radiation in both signal and background events.
- It can come at central rapidities.
- It is largely rather low transverse momentum.

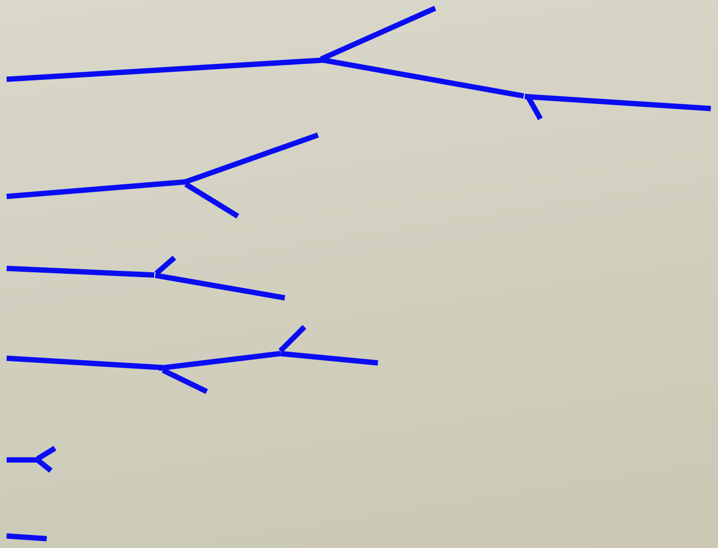




An application



- Look at all of the hadrons or calorimeter clusters in an angular region that might contain a boosted heavy particle.
- The k_T jet algorithm can put the starting clusters together into jets in something like a shower history.
- Use a fairly small angular size parameter.
- The smallest P_T jets are likely from initial state radiation.
- Throw them out. (Krohn, Thaler & Wang).



Another application

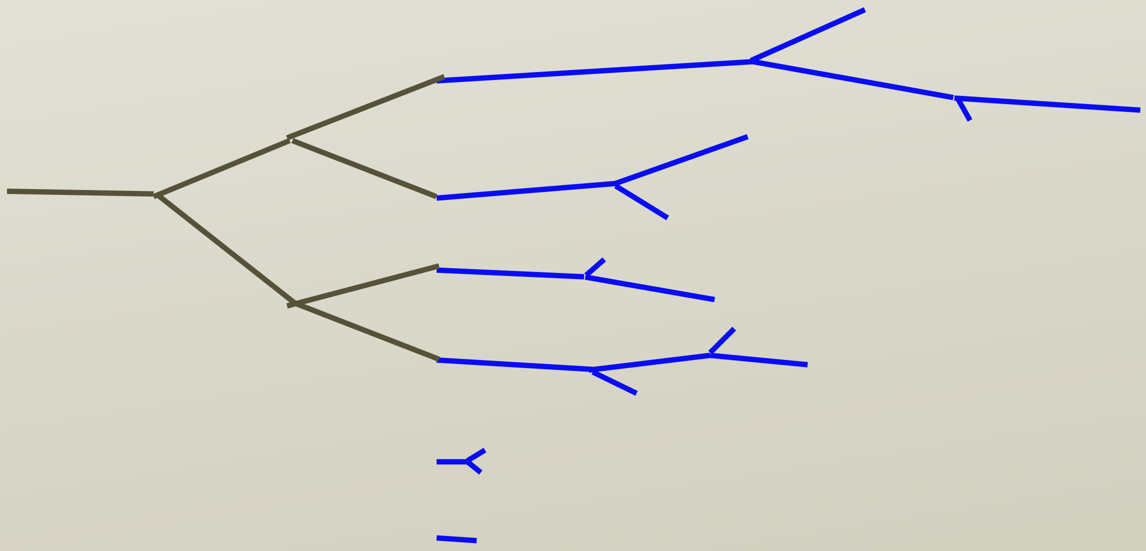
- It's hard to tell where soft gluons belong, so try to get rid of them.

- For instance, "pruning" joins protojets starting at small angular separations. But if

$$z = \frac{\min(p_{T,i}, p_{T,j})}{|\vec{p}_{T,i} + \vec{p}_{T,j}|}$$

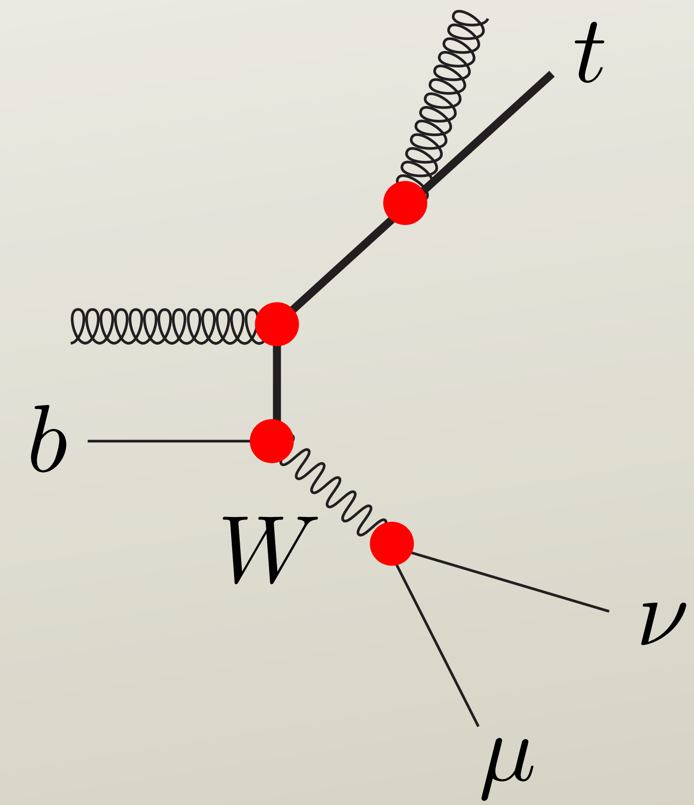
  is too small, one throws the softer protojet away.

  (Ellis, Vermilion & Walsh)



Heavy particles with color

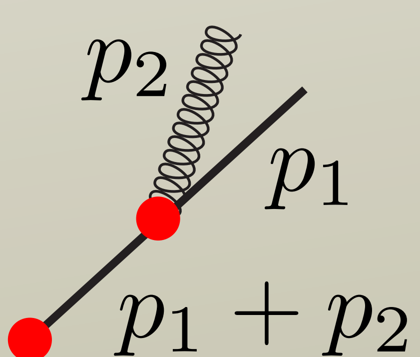
- A heavy particle with color will radiate gluons.
- There is singularity in $|\mathcal{M}|^2$.



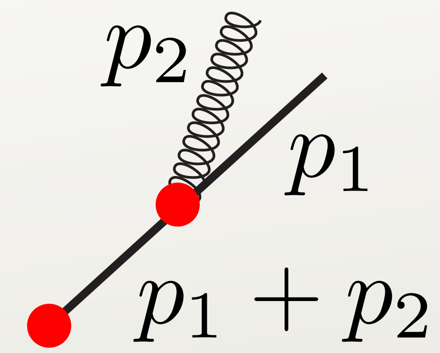
$$dn \propto \frac{dv}{v}\,\alpha_s \log(\theta_{\min}/\theta_{\max})$$

where the hardness v is now

$$v = \frac{(p_1+p_2)^2 - M^2}{(p_1+p_2)\cdot Q_0}$$



- The collinear singularity is cut off:

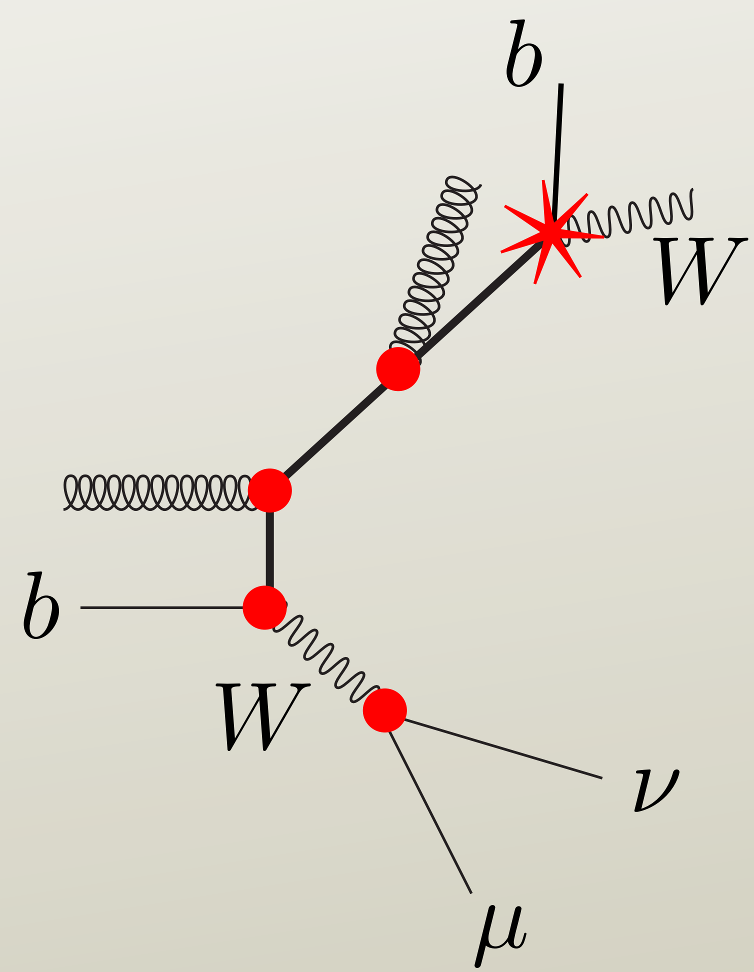


$$\theta^2_{\min} \sim \min\left(\frac{M^2 Q_0^2}{((p_1+p_2)\cdot Q_0)^2}, v\,\frac{4Q_0^2}{(p_1+p_2)\cdot Q_0} \right)$$

- There are no emissions that take longer than the particle lifetime:

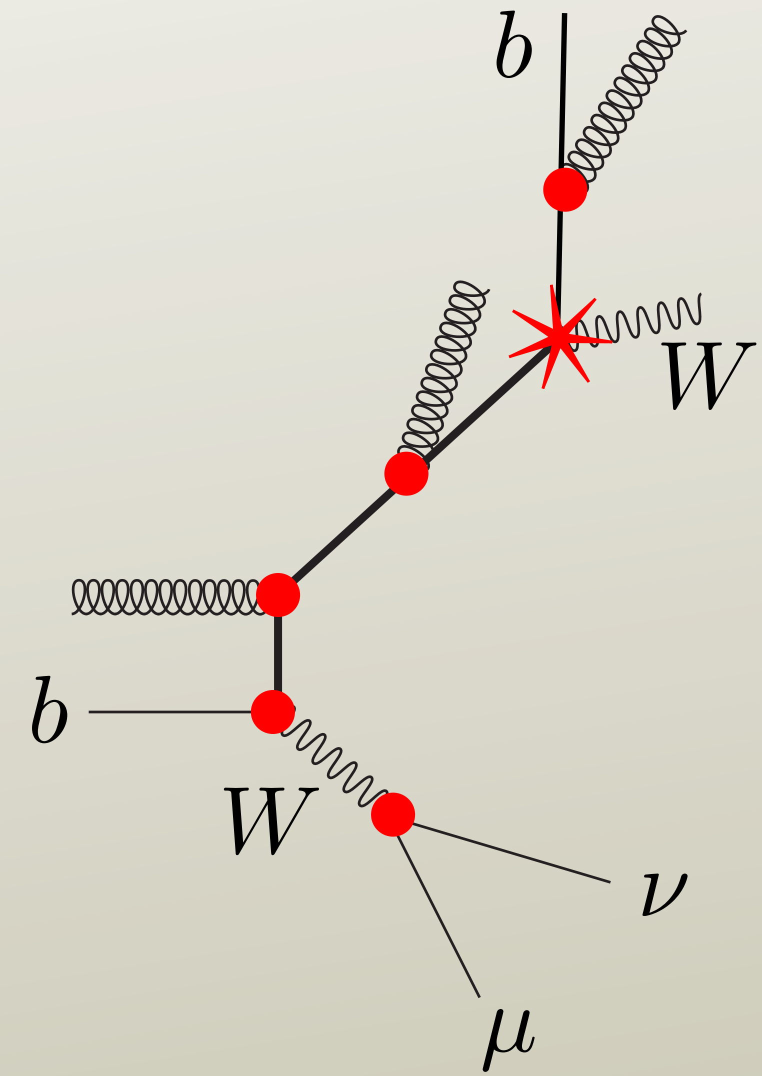
$$v > \frac{M\Gamma}{(p_1+p_2)\cdot Q_0}$$

Heavy particles decay



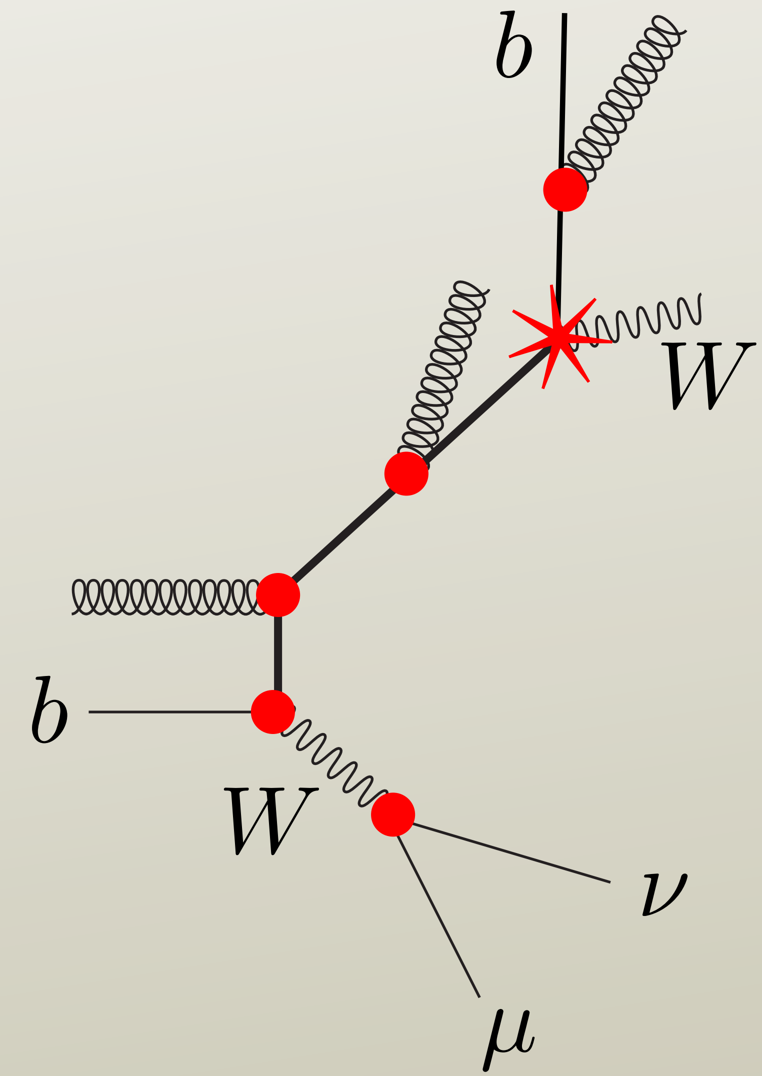
Daughters radiate

- In $t \to b + W$, a lot of energy is released.
- Color charge is accelerated.
- Gluons are radiated.
- The first radiation is from a color dipole of
  - The final state b-quark;
  - The "initial state" t-quark;

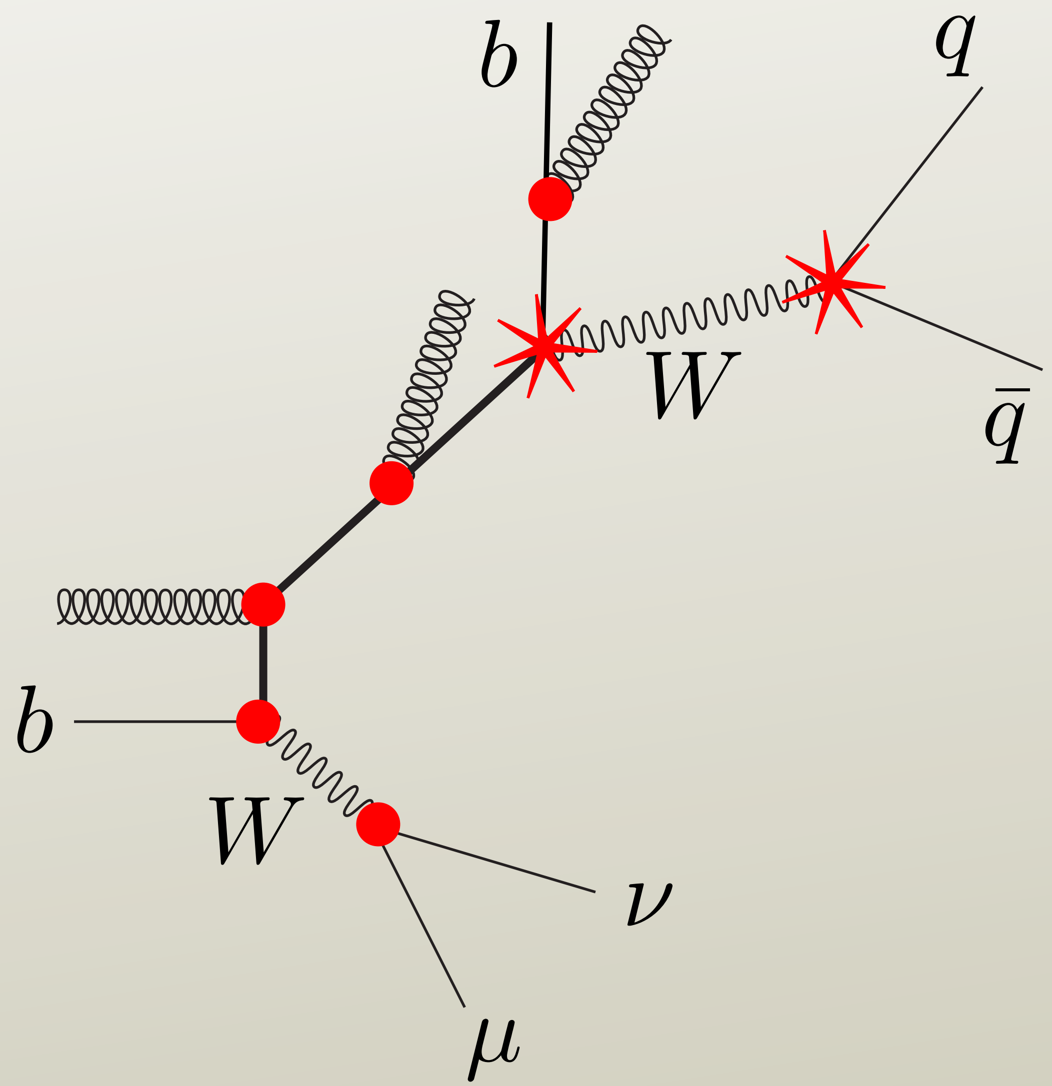


Where does the radiation go?

- The first radiation is from a color dipole of
  - The final state b-quark;
  - The "initial state" t-quark;
- Direction of g is likely collinear with the b.
- Direction of g is likely within angle $M/|\vec{P}_t|$ of the top.
- If the top is highly boosted, this is a narrow cone.

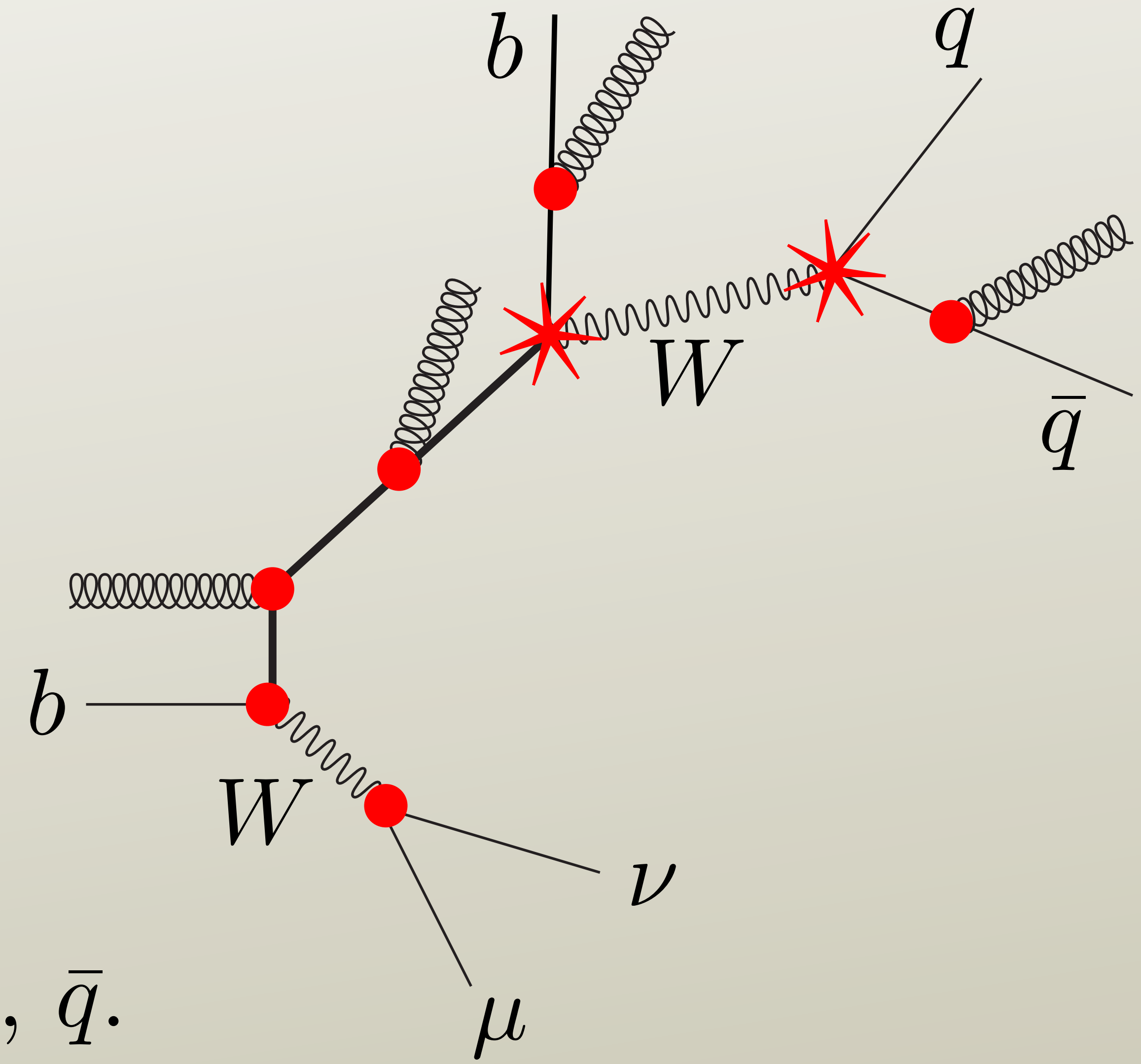


Heavy particles decay



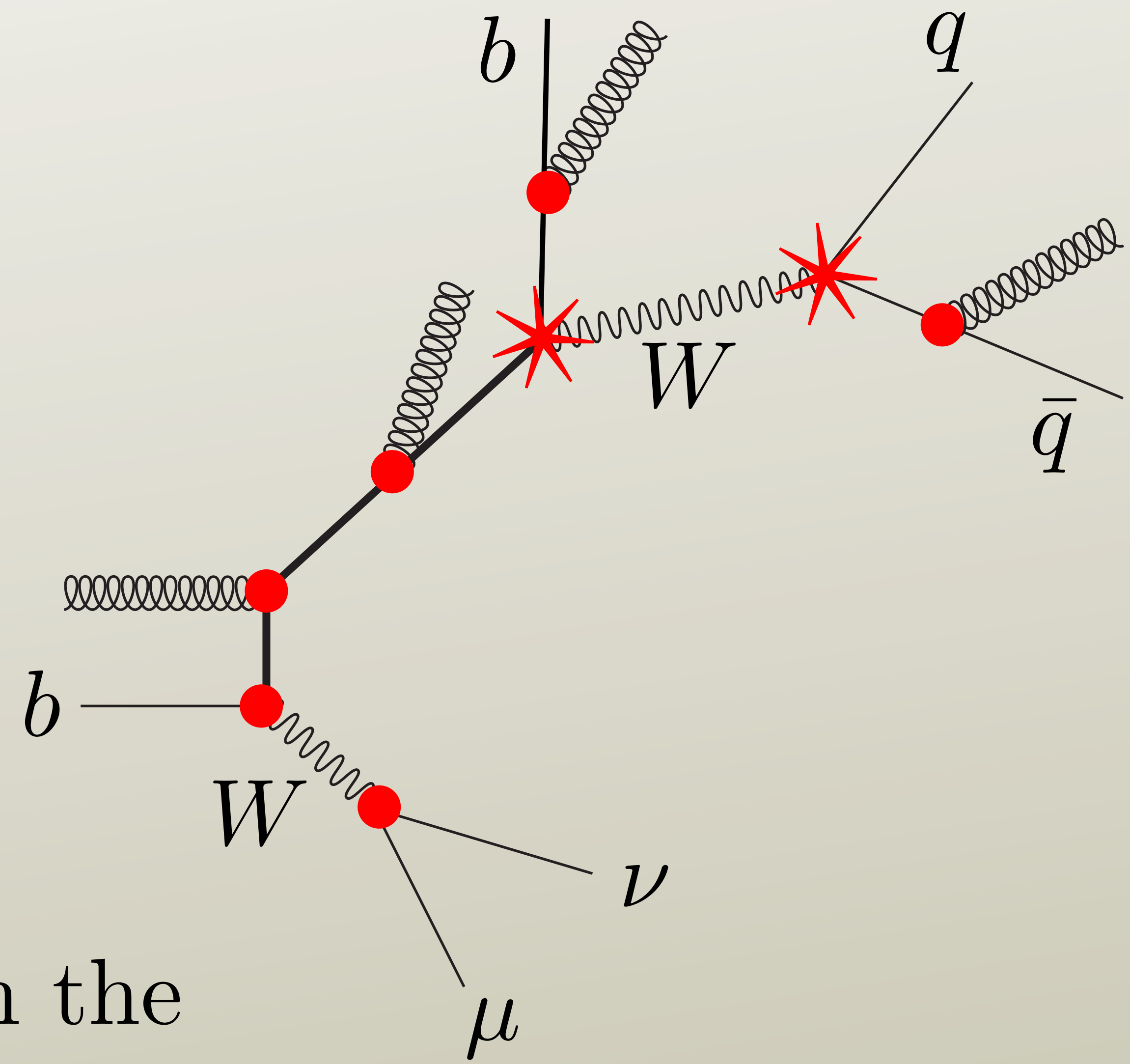
Daughters radiate

- In $W \to q + \bar{q}$, a lot of energy is released.
- Color charge is accelerated.
- Gluons are radiated.
- The first radiation is from a color dipole of
  - The final state quark, q;
  - The final state anti-quark, $\bar{q}$.



Where does the radiation go?

- The first radiation is from a color dipole of
  - The final state quark, q;
  - The final state anti-quark, $\bar{q}$.
- Direction of g is likely collinear with the q or $\bar{q}$.
- In any case, it is likely between the q and $\bar{q}$.

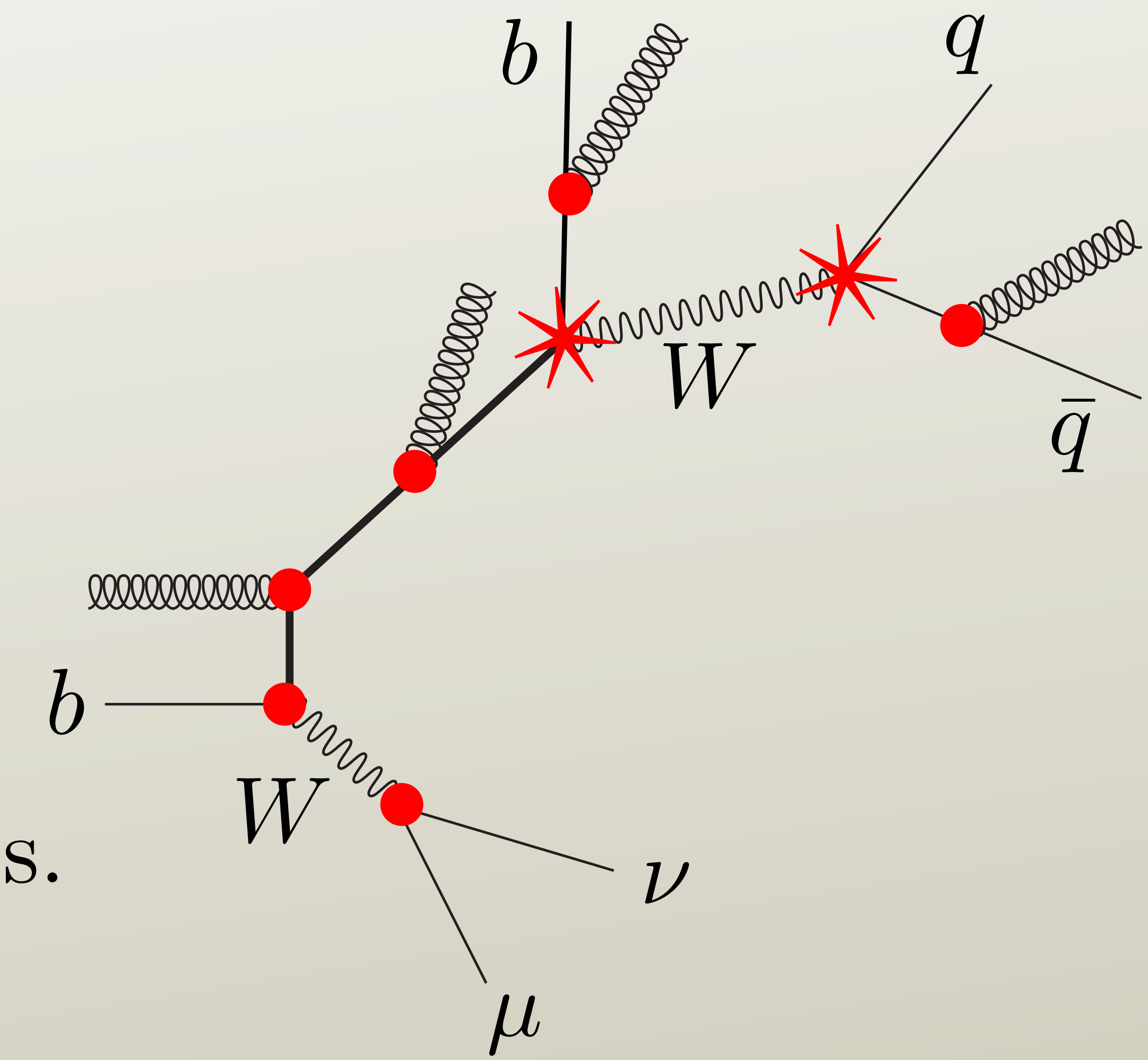


Very soft radiation

- For very soft emissions,

$$v < \frac{M_t \Gamma_t}{p_t \cdot Q_0} \sim \frac{M_W \Gamma_W}{p_W \cdot Q_0}$$

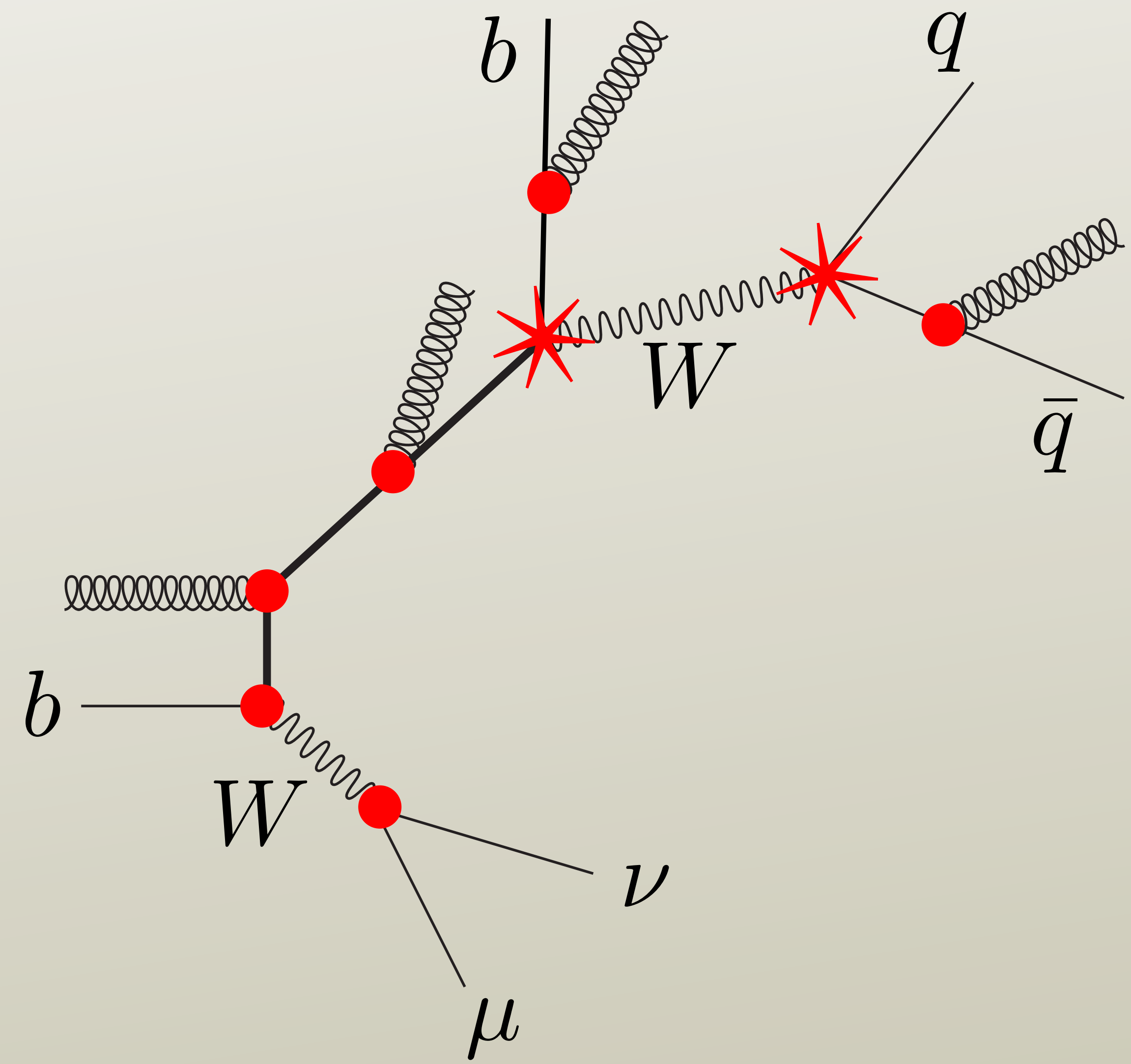
there is quantum interference between emissions from any two (color connected) particles.



- Probably the structure of these emissions is too complicated to be useful for separating signal from background.

Comments

- Since we know something abut where the collinear or soft radiation goes in a heavy particle decay, sequential heavy particle decays should have good signatures.
- We do have the problem of contamination from initial state radiation.
- Having highly boosted heavy objects helps with this.



Summary

- With a model for sequential heavy particle decays and knowledge of the structure of radiation in gauge theories, one can build filters for separating background events from signal events.