Title: Jet substructure in BSM physics

Date: Feb 21, 2013 09:10 AM

URL: http://pirsa.org/13020143

Abstract:

SEARCHES FOR NEW PHYSICS DURING THE SHUTDOWN

Jay Wacker
SLAC
February 21, 2013

Applications of Jet Substructure to New Physics Searches
Perimeter Institute

OUTLINE

- Motivation
- Substructure Variables
- Multijet Searches
- Data Driven Backgrounds
- Outlook

MOTIVATION

First round of the LHC is complete

New physics didn't jump out

2 years to energy upgrade

Parked data being analyzed

HIDDEN SIGNALS

Searches that use lots of MET or leptons
are spectacularly clean

Having explored those channels,
now it's prudent to examine ones
that don't have those handles

Need handles to distinguish signals from QCD/Top

CLASSIFICATION OF SIGNALS

Simple Signals

- Only a few final state particles
- Must make use of all kinematic information for discovery
- High backgrounds

Complicated Signals

- Lots of final state particles
- Use of kinematic information too subtle for discovery
- Irreducible backgrounds low, but challenge to separate S/B

SIMPLE SIGNALS

2-Jet Resonance

Pair Produced 2-Jet Resonances

Resonant Produced Pair of 2-Jet Resonances

Pair Produced 3-jet Resonances

Resonant Produced Pair of 3-Jet Resonances

COMPLICATED SIGNALS

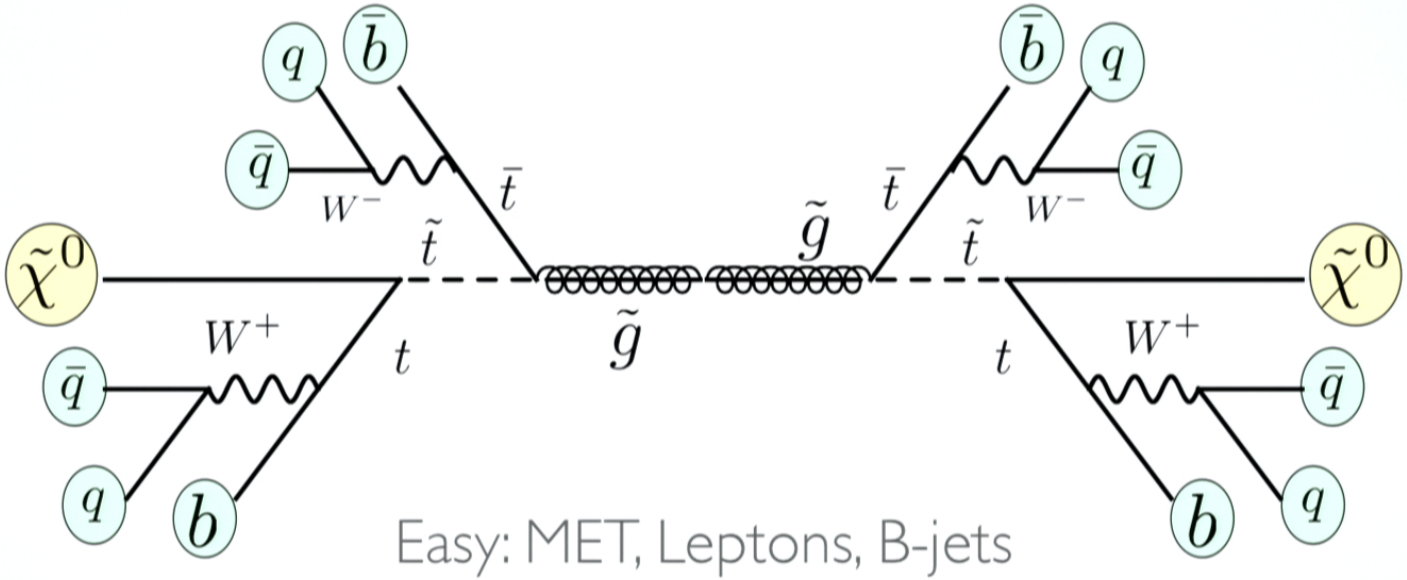
Frequently occur when lightest particles are hard to find

Low Production Rate

Poor Signal to Background

COMPLICATED SIGNALS

Susy Cascade Decays

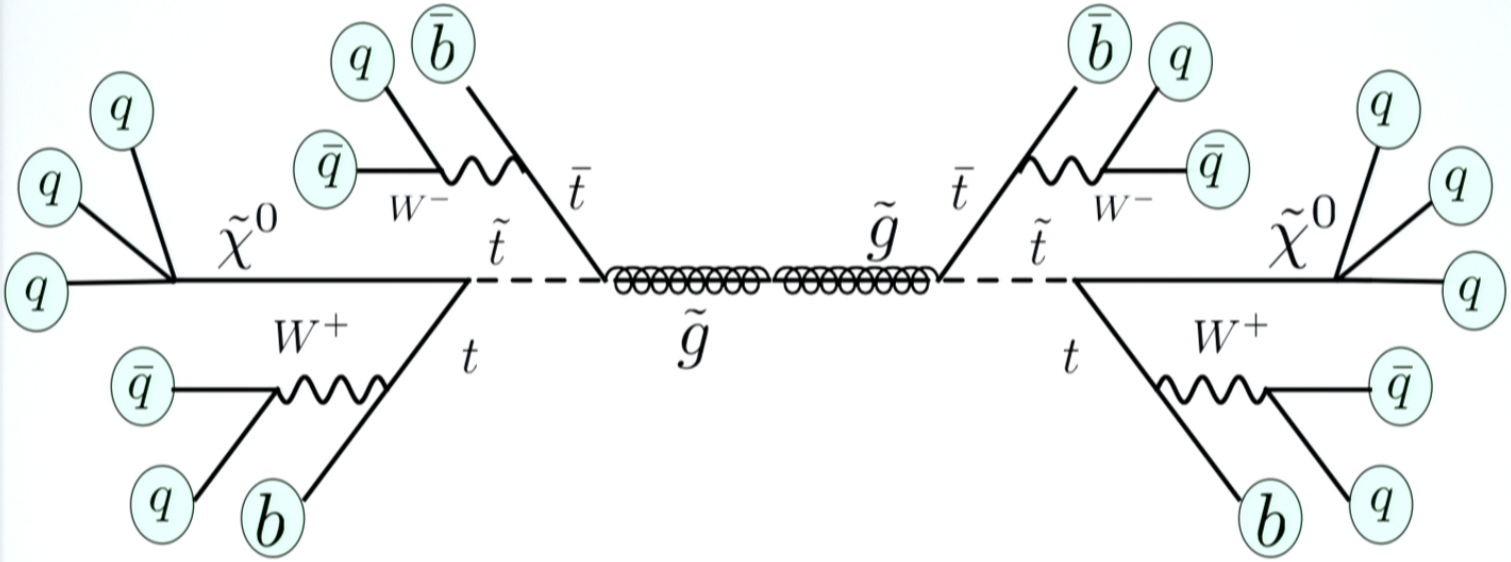


Easy: MET, Leptons, B-jets

Reconstruction not part of discovery process

COMPLICATED SIGNALS

Susy Cascade Decays With RPV



Easy: Leptons, B-jets

COMPLICATED SIGNALS

New Theories of Strong Dynamics

	SU(N)	SM	Spin
q	N	3	1/2
q^c	N^*	3^*	1/2

	SM	Spin
π	8	0
ρ	8	1

$\pi \to \quad gg, \quad t\bar{t}$ $\qquad\qquad$ $\rho \to \pi\pi$

	π→ gg	π→ tt*
$pp \to \pi\pi$	4g	4t
$pp \to \rho$	4g	4t
$pp \to \rho\rho$	8g	8t

COMPLICATED SIGNALS

Add Flavor

	SU(N)	SM	Flavor	Spin
q	N	3	2	1/2
q^c	N^*	3^*	2	1/2

	SM	Flavor	Spin
π	8	1	0
ρ	8	1	1
π_f	8	3	0
ρ_f	8	3	1
π_0	1	3	0
ρ_0	1	3	1

Colored Resonances Heavier

Long cascade decays & signals depend on details

COMPLICATED SIGNALS

Add Flavor

	SU(N)	SM	Flavor	Spin
q	N	3	2	1/2
q^c	N^*	3^*	2	1/2

	SM	Flavor	Spin
π	8	1	0
ρ	8	1	1
π_f	8	3	0
ρ_f	8	3	1
π_0	1	3	0
ρ_0	1	3	1

Colored Resonances Heavier

Long cascade decays & signals depend on details

THE CHALLENGE OF JETS

Thin (R=0.4) jets don't have much structure

Described by p_T η φ

Signals appear at high multiplicity

Phase space is 3^{NJ}

Parton Shower produces jets too

THE ROLE OF FAT JETS

Fat (R=1.0) jets have lots of structure

Described by p_T η φ + Jet Shapes

High multiplicity isn't possible

Signals appear as anomalies in jet shapes

Parton Shower doesn't change multiplicity

THE ROLE OF FAT JETS

Fat (R=1.0) jets have lots of structure

Described by p_T η φ + Jet Shapes

High multiplicity isn't possible

Signals appear as anomalies in jet shapes

Parton Shower doesn't change multiplicity

SUBSTRUCTURE VARIABLES

Lots of jet variables now proposed

Mass, Girth, etc.

N-Subjettiness

Color Flow Variables

N-Subjets

N-SUBJETS

Having a measure of the number of subjets in a jet is quite useful

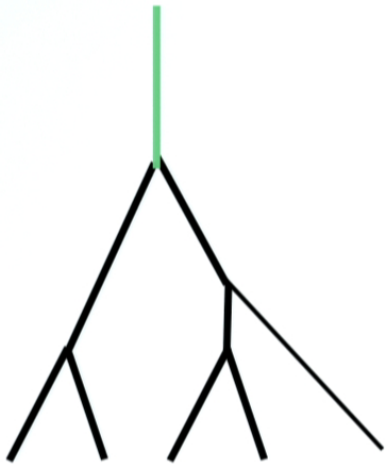
"What about N-Subjettiness?"

Not answering this question.
Not easy to turn $\tau_1, \ldots, \tau_n$ into a discrete number

N	τ1	τ21	τ31	τ41
1	<13%			
2	13% - 25%			
2	>25%	<40%		
3	25% - 45%	>40%	<35%	
3	>45%	>40%	<35%	<20%
4	>45%	>40%	<35%	>20%
4	25% - 40%	>40%	>35%	<40%
4	40% - 50%	>40%	35% - 45%	<40%
5	>25%	>40%	>35%	>40%
5	25% - 50%	>40%	>45%	<40%
5	>50%	>40%	35% - 45%	>40%

Subjet Counting

by reversing the CA tree



If $m_j < m_{cut}$, j is a subjet

If not, uncluster last step

If $R_{j1j2} < R_{cut}$, j is a subjet

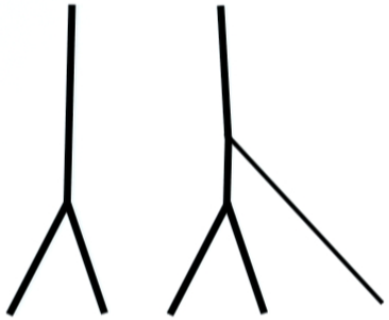
If $p_{T\,j2} < y_{cut}\ p_{T\,j}$, j_2 throw out

If $p_{T\,j2} < p_{T\,cut}$, j_2 throw out

Iterate

Subjet Counting

by reversing the CA tree



If $m_j < m_{cut}$, j is a subjet

If not, uncluster last step

If $R_{j1j2} < R_{cut}$, j is a subjet

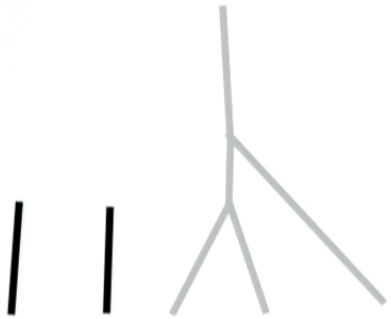
If $p_{T\,j2} < y_{cut}\ p_{T\,j}$, j_2 throw out

If $p_{T\,j2} < p_{T\,cut}$, j_2 throw out

Iterate

Subjet Counting

by reversing the CA tree



If $m_j < m_{cut}$, j is a subjet

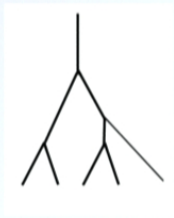
If not, uncluster last step

If $R_{j1j2} < R_{cut}$, j is a subjet

If $p_{T\,j2} < y_{cut}\ p_{T\,j}$, j_2 throw out

If $p_{T\,j2} < p_{T\,cut}$, j_2 throw out

Iterate

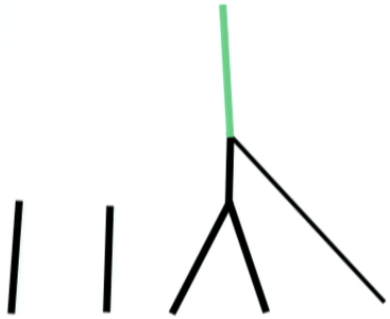


Subjet Counting

by reversing the CA tree

If $m_j < m_{cut}$, j is a subjet

If not, uncluster last step



If $R_{j1j2} < R_{cut}$, j is a subjet

If $p_{T\,j2} < y_{cut}\ p_{T\,j}$, j_2 throw out

If $p_{T\,j2} < p_{T\,cut}$, j_2 throw out

Iterate

Subjet Counting

by reversing the CA tree

If $m_j < m_{cut}$, j is a subjet

If not, uncluster last step

If $R_{j1j2} < R_{cut}$, j is a subjet

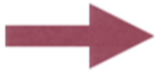
If $p_{T\,j2} < y_{cut}\ p_{T\,j}$, j_2 throw out

If $p_{T\,j2} < p_{T\,cut}$, j_2 throw out

Iterate

Subjet Counting

by reversing the CA tree



If $m_j < m_{cut}$, j is a subjet

If not, uncluster last step

If $R_{j1j2} < R_{cut}$, j is a subjet

If $p_{T\,j2} < y_{cut}\ p_{T\,j}$, j_2 throw out

If $p_{T\,j2} < p_{T\,cut}$, j_2 throw out

Iterate

Subjet Counting

by reversing the CA tree

If $m_j < m_{cut}$, j is a subjet

If not, uncluster last step

If $R_{j1j2} < R_{cut}$, j is a subjet

If $p_{T\,j2} < y_{cut}\ p_{T\,j}$, j_2 throw out

If $p_{T\,j2} < p_{T\,cut}$, j_2 throw out

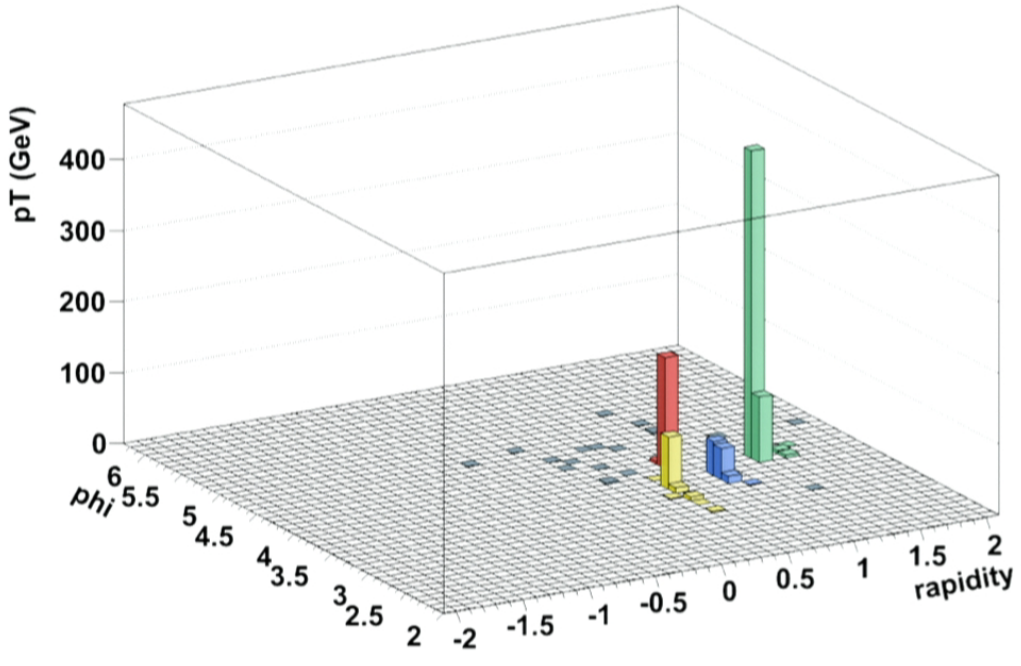
Iterate

3 Subjets

Values Used In Counting Algorithm

(Not strongly sensitive)

m_{cut} = 30 GeV R_{cut} = 0.15 y_{cut} = 0.10 $P_{T\ cut}$ = 30 GeV



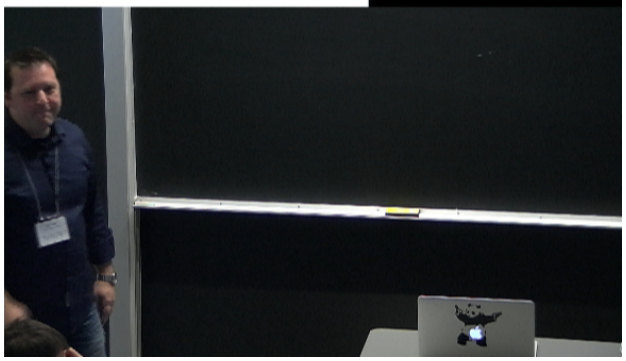
N-SUBJETS

Having a measure of the number of subjets in a jet is quite useful

"What about N-Subjettiness?"

Not answering this question.
Not easy to turn $\tau_1, \ldots, \tau_n$ into a discrete number

N	τ1	τ21	τ31	τ41
1	<13%			
2	13% - 25%			
2	>25%	<40%		
3	25% - 45%	>40%	<35%	
3	>45%	>40%	<35%	<20%
4	>45%	>40%	<35%	>20%
4	25% - 40%	>40%	>35%	<40%
4	40% - 50%	>40%	35% - 45%	<40%
5	>25%	>40%	>35%	>40%
5	25% - 50%	>40%	>45%	<40%
5	>50%	>40%	35% - 45%	>40%



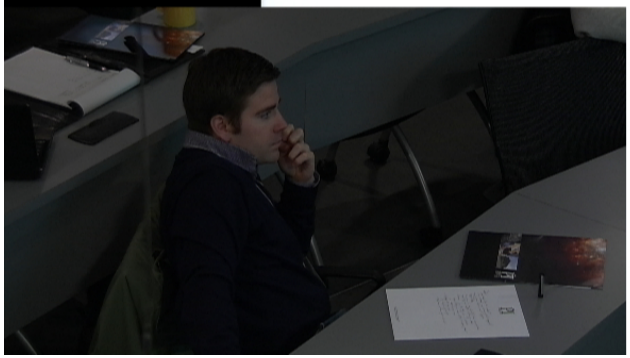
N-SUBJETS

Having a measure of the number of subjets in a jet is quite useful

"What about N-Subjettiness?"

Not answering this question.
Not easy to turn $\tau_1, \ldots, \tau_n$ into a discrete number

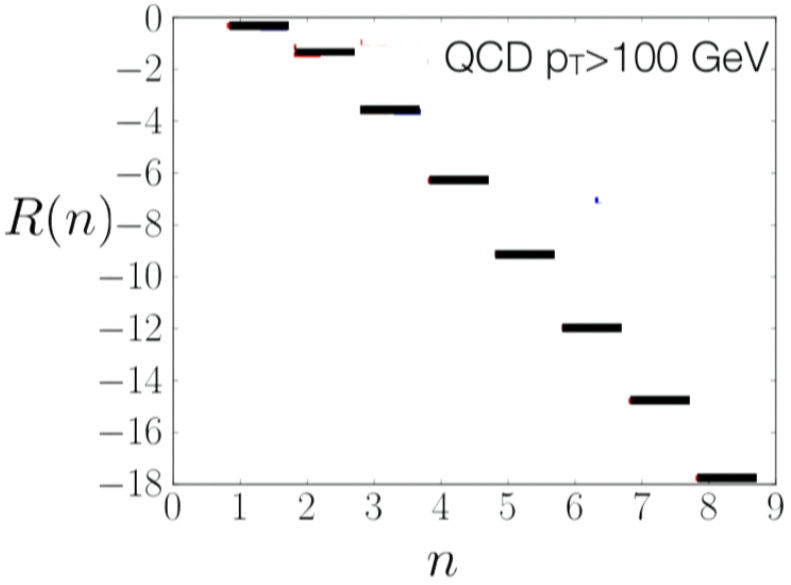
N	τ1	τ21	τ31	τ41
1	<13%			
2	13% - 25%			
2	>25%	<40%		
3	25% - 45%	>40%	<35%	
3	>45%	>40%	<35%	<20%
4	>45%	>40%	<35%	>20%
4	25% - 40%	>40%	>35%	<40%
4	40% - 50%	>40%	35% - 45%	<40%
5	>25%	>40%	>35%	>40%
5	25% - 50%	>40%	>45%	<40%
5	>50%	>40%	35% - 45%	>40%

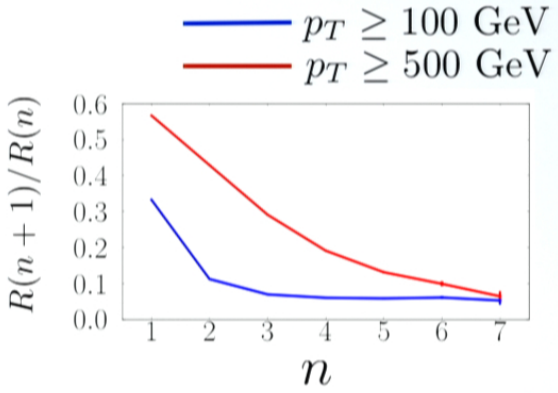


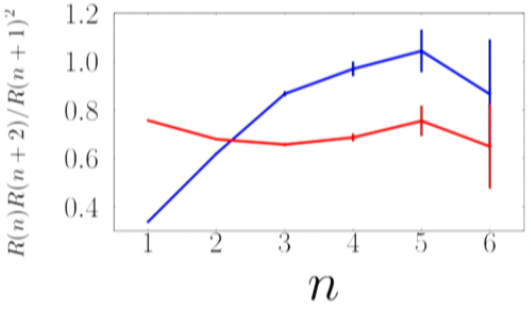
$N_{subjets}$ is not Number of Jets

Does not exhibit “stair case scaling”

Many fewer subjets than jets



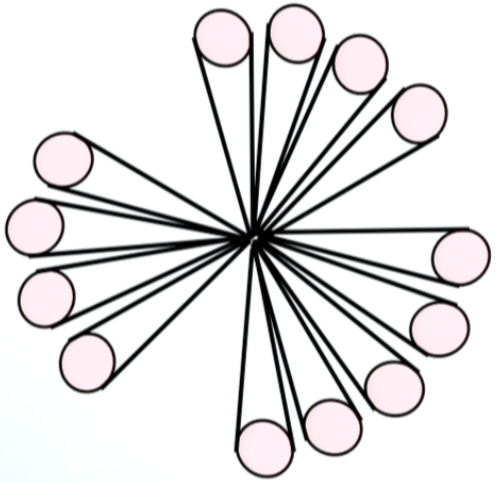




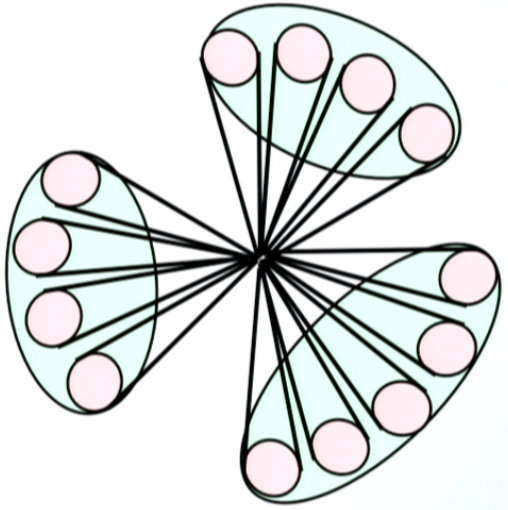
MULTIJET SEARCHES

High multiplicity final states will
appear as multiple fat jets,
but with unusual fat jets

13 Jet Event



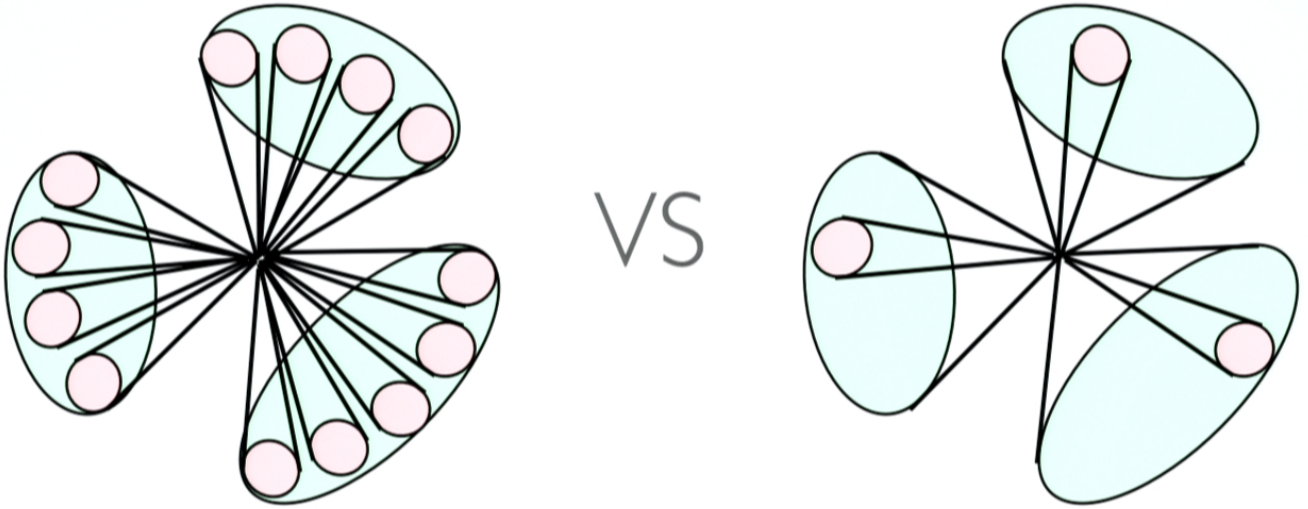
3 Jet Event



DISCOVERY PROCESS

Signal

Background



COMBINING SUBSTRUCTURE

After requiring N-fat jets (3, 4, 5)
moderate to high pT

Can try any set of variables,
will use Mass & N-subjets

Fat jets are isolated from each other

Have N pairs of (mass, $n_{subjets}$)

USING MASS

Our studies have mostly had some MET,
meaning that top becomes an important BG

Top has an intrinsic mass scale of ~350GeV

Sum of Jet Masses $$M_J = \sum_{n=1}^{N_J} m_{j_n}$$

If dealing with MET-less signal, other combinations may be better

Perhaps using summing m/pT

USING MASS

Our studies have mostly had some MET,
meaning that top becomes an important BG

Top has an intrinsic mass scale of ~350GeV

Sum of Jet Masses $$M_J = \sum_{n=1}^{N_J} m_{j_n}$$

If dealing with MET-less signal, other combinations may be better

Perhaps using summing m/pT

M_J as a replacement for H_T

$$H_T = \sum E_{T\,i} = \sum (p_{T\,i}^2 + m_{j_i}^2)^{\frac{1}{2}}$$

Signal	Background
$m_j/p_T \sim 1$	$m_j/p_T \lesssim 1$
$H_T \sim M_J$	$H_T \gtrsim M_J$

Signal has higher M_J for fixed H_T

Keep less background at same signal efficiency

Never does parametrically worse

USING SUBJETS

Simplest thing to do with subjets is

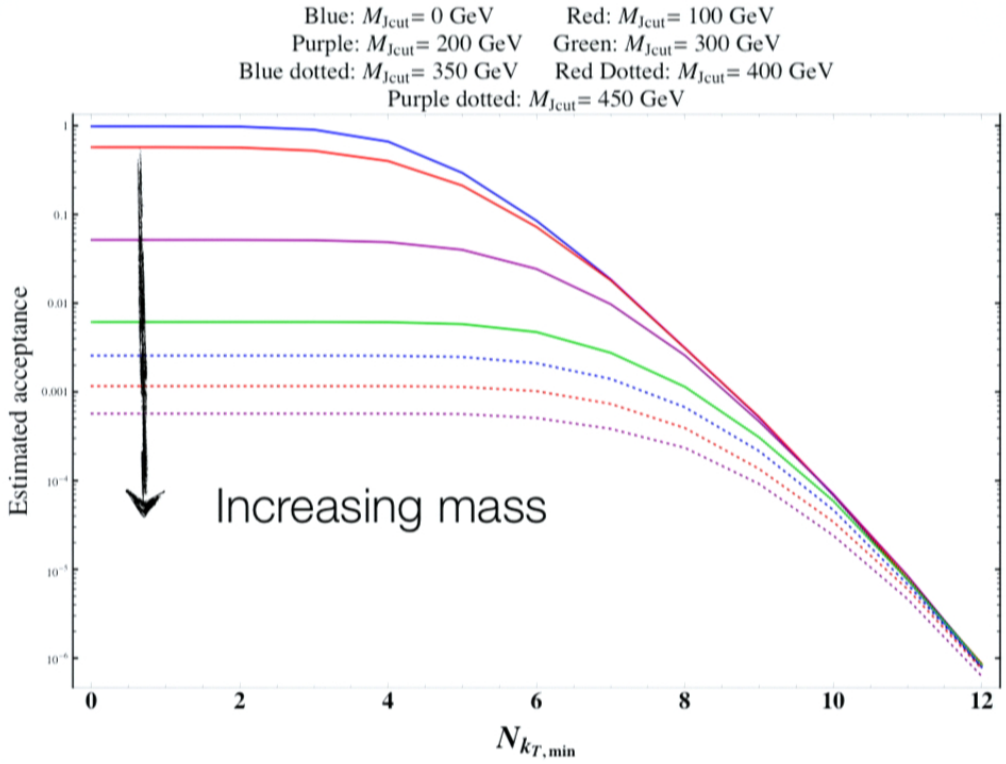
$$N_J = \sum_i n_i$$

Not the same as the number of jets in an event

Steeper fall off in multiplicity due to grooming

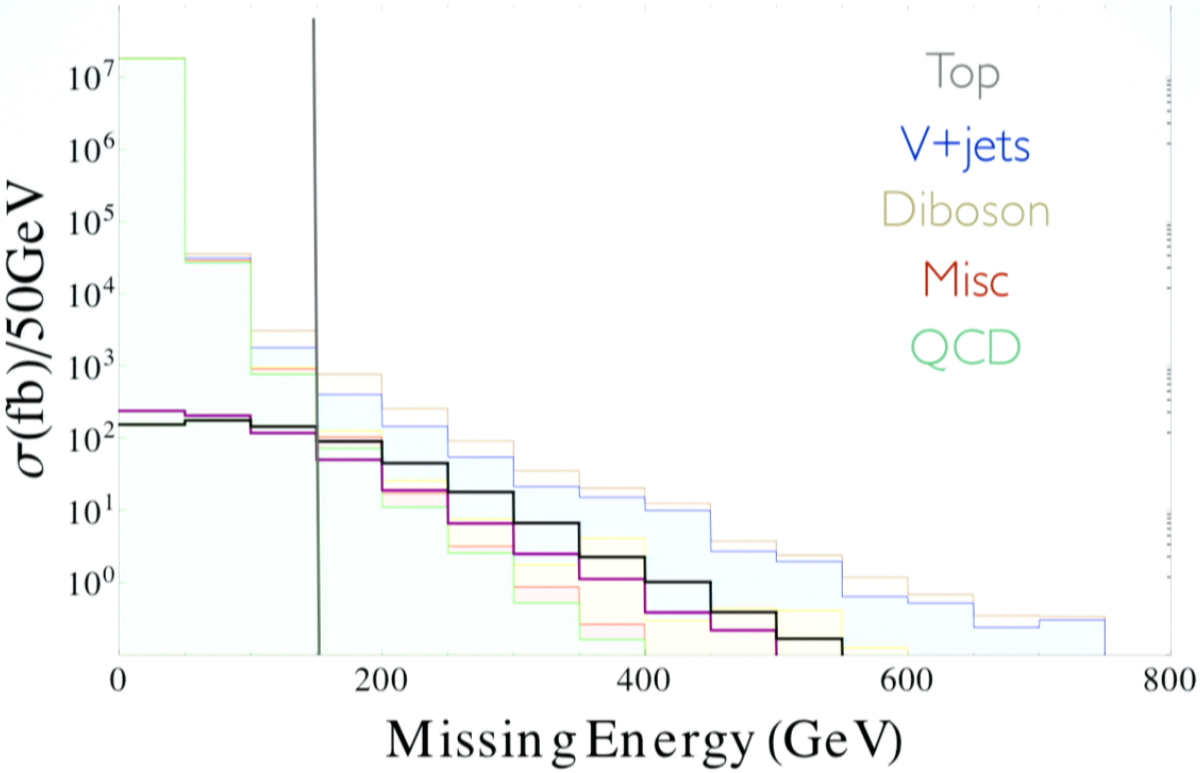
Mass vs N-Subjets

Mass and N-subjets are correlated,
but most mass generated with 2 subjets



4 Fat Jets, $p_T > 100$ GeV

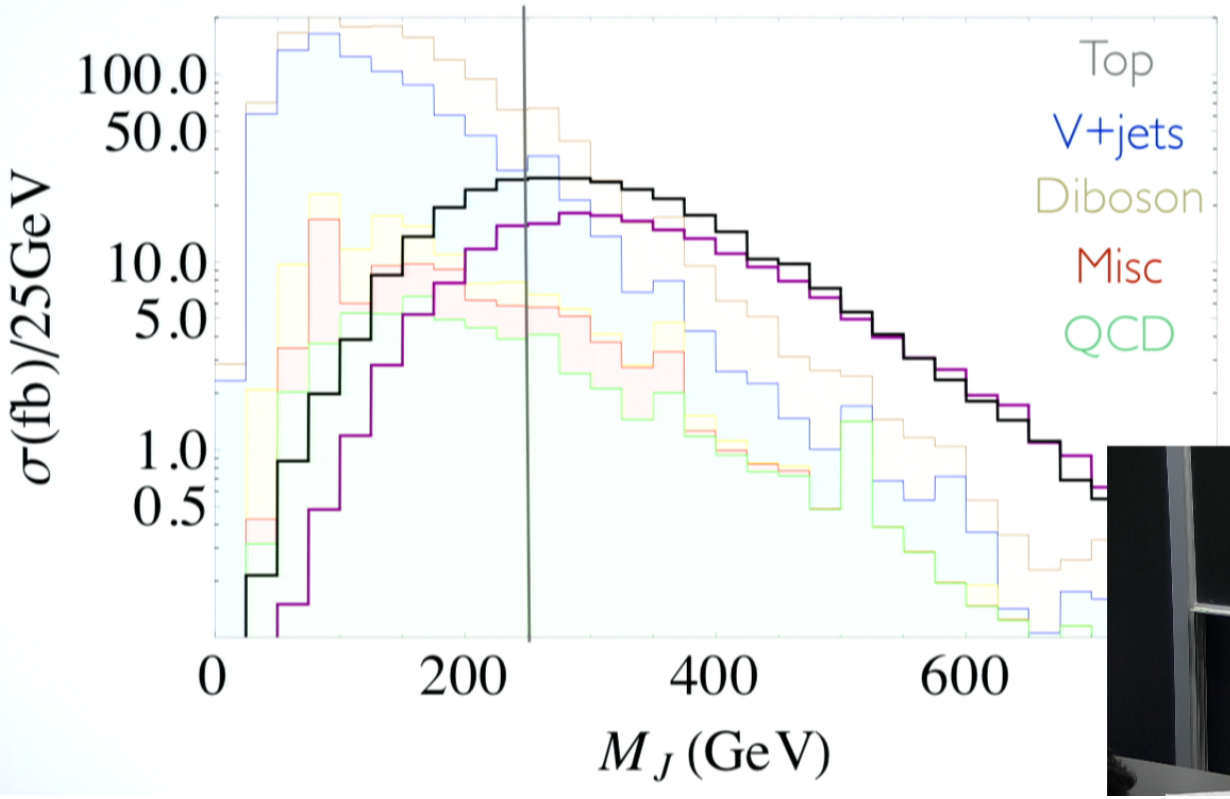
$\not{E}_T$ Distribution

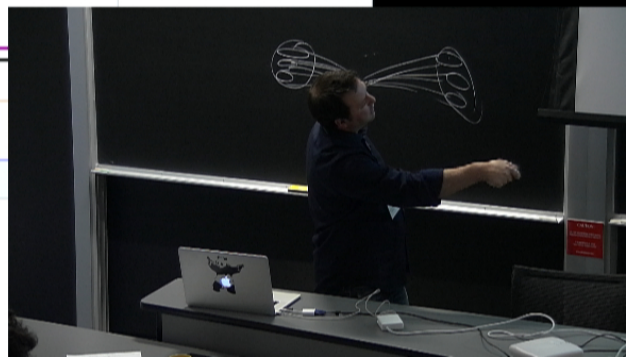


4 Fat Jets, $p_T > 100$ GeV

After $\not{E}_T > 150$ GeV

M_J Distribution

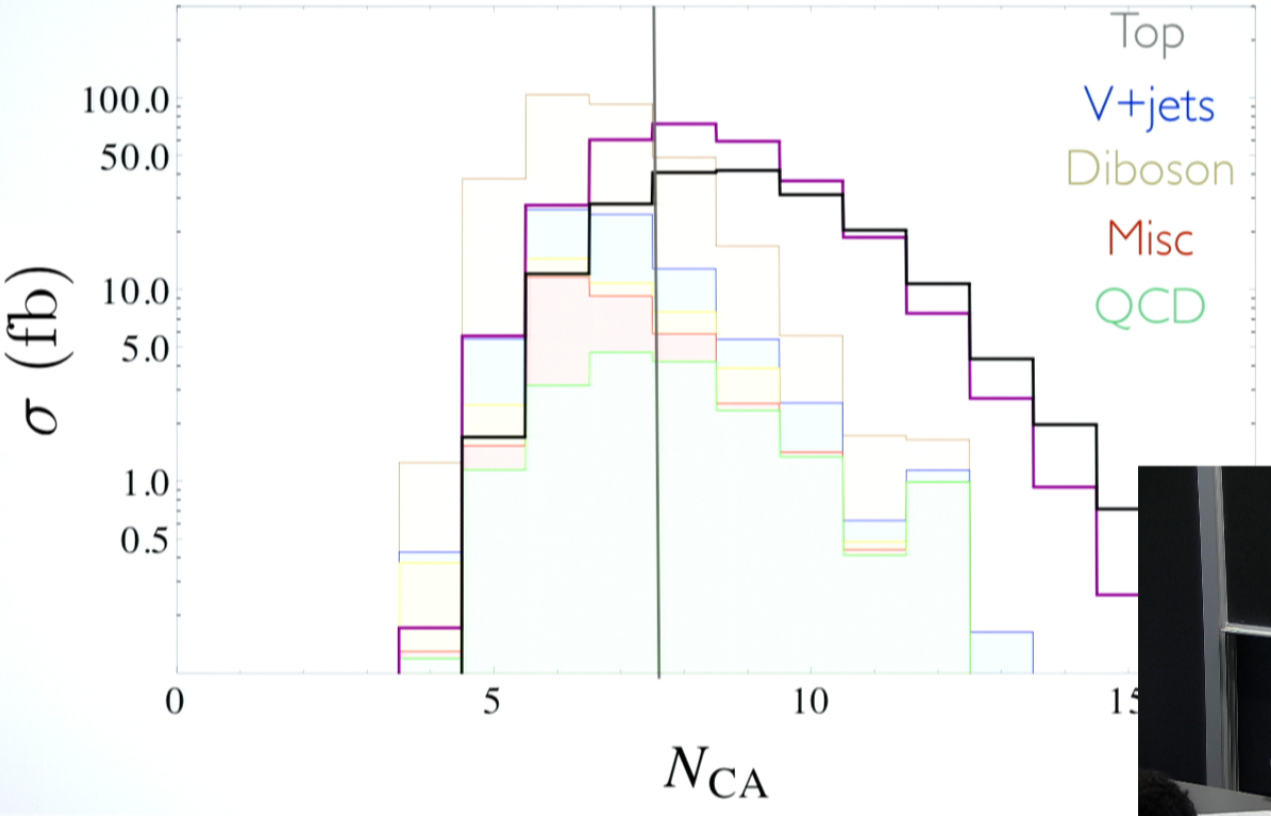




4 Fat Jets, $p_T > 100$ GeV

After $\not{E}_T > 150$ GeV & $M_J > 280$ GeV

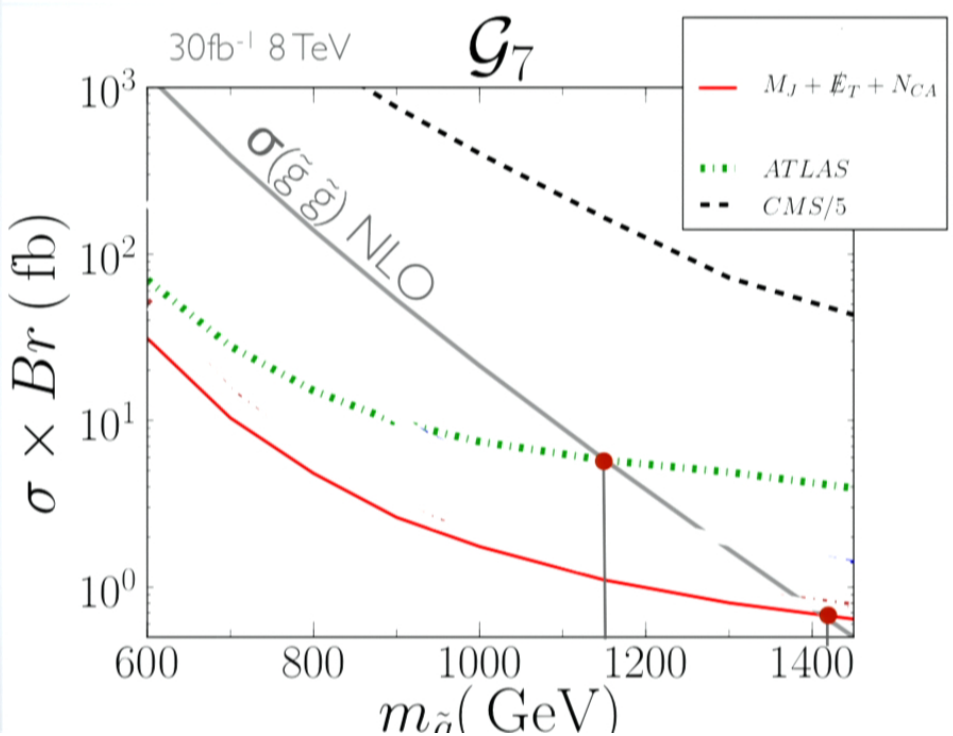
N_J Distribution

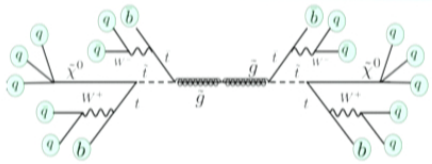


Improvements of N_J vs M_J only Search

$$\not{E}_T > 125 \text{ GeV} \qquad M_J \geq 425 \text{ GeV} \qquad N_J > 14$$

$\sigma_{SM} \approx 0.07$fb





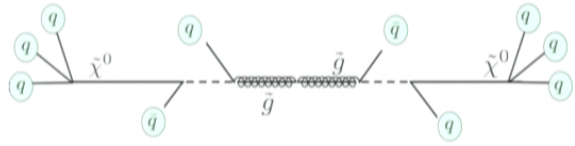
Factor of 8 improvement in cross section, factor of 64 less luminosity

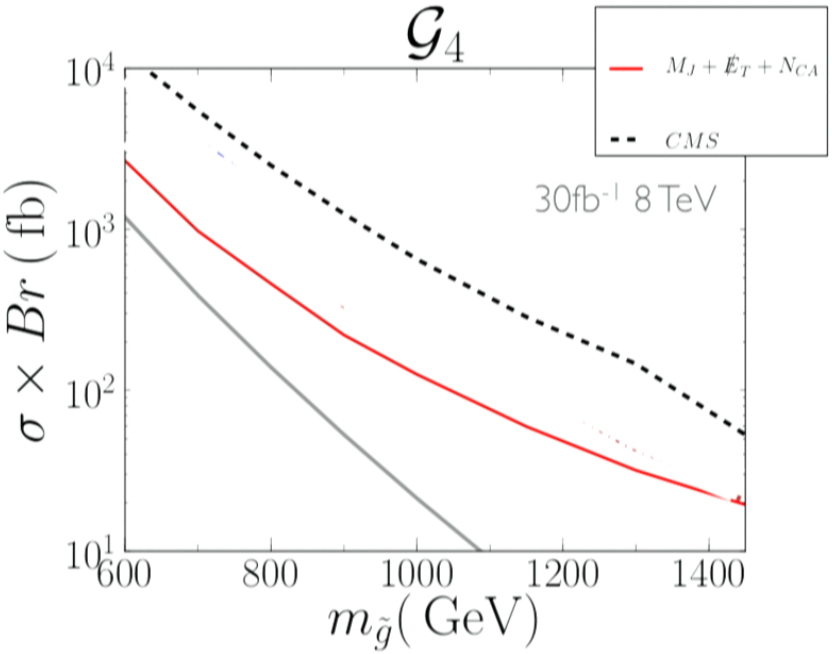
Improvements of N_J vs M_J only Search

$M_J + N_J$

$M_J \geq 1050$ GeV $N_J > 14$

$\sigma_{SM} \simeq 0.7$fb





Factor of 5 improvement in cross section, requiring 25x less luminosity

Can be combined with other requirements

b-tags leptons

Having these new techniques
requires that less weight is placed
upon these objects

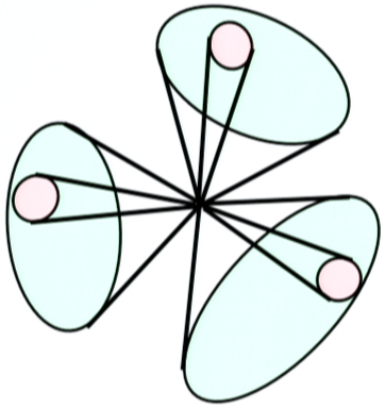
May be single lepton or b-tag can tame QCD

OUTLINE

- Motivation
- Substructure Variables
- Multijet Searches
- Data Driven Backgrounds
- Outlook

DATA DRIVEN BACKGROUNDS

QCD jets only have small correlations



$$x = m_j/p_T$$

$$P_3(x_1, x_2, x_3) \simeq P_1(x_1)P_1(x_2)P_1(x_3)$$

Measure in one sample and extrapolate

Also can use other control regions (MET/leptons/bjets)

Would like a calculation to understand correlations

Should measure in multiple settings (q vs g composition)

Natural "Data-Driven" approach to backgrounds

$$P_1(x; p_T)$$

Now use in the multijet sample

Predict event-by-event acceptances

(probability an event passes cut)

$$A(p_{T\,1}, p_{T\,2}, p_{T\,3}) = \int_{M_J > m_{\rm cut}} d^3x \quad P_1(x_1; p_{T\,1}) P_1(x_2; p_{T\,2}) P_1(x_3; p_{T\,3})$$

Can make an M_J prediction based upon the events *measured*

Don't need to be able to calculate M_J distribution from first principles

Natural "Data-Driven" approach to backgrounds

$$P_1(x; p_T)$$

Now use in the multijet sample

Predict event-by-event acceptances

(probability an event passes cut)

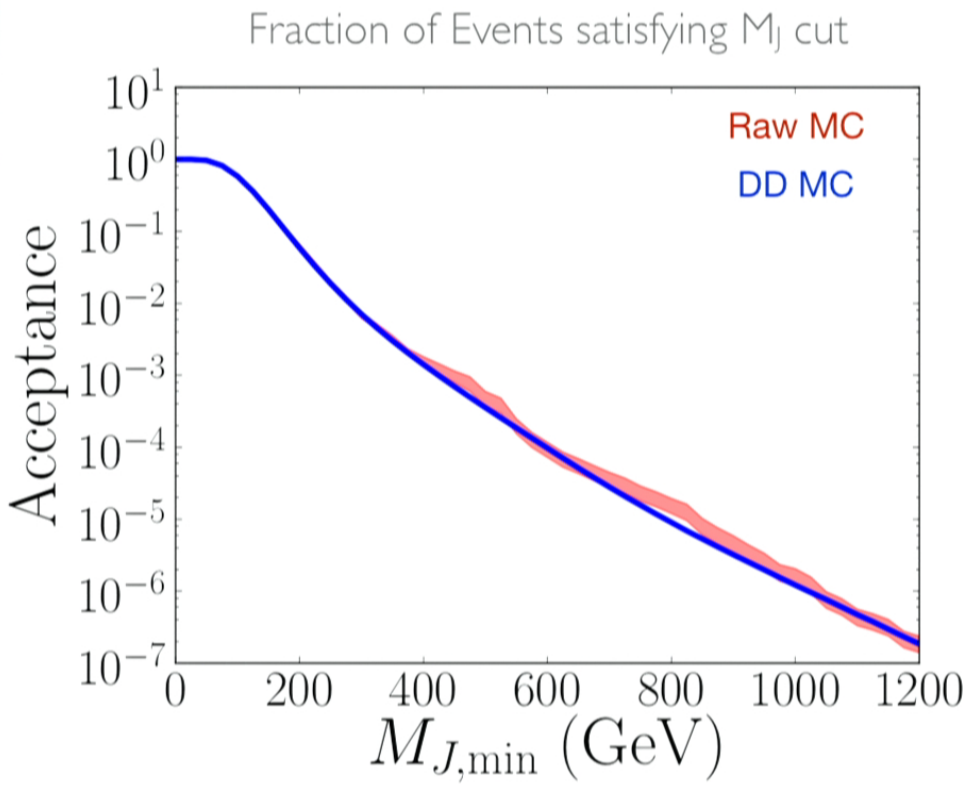
$$A(p_{T\,1}, p_{T\,2}, p_{T\,3}) = \int_{M_J > m_{\rm cut}} d^3x \quad P_1(x_1; p_{T\,1}) P_1(x_2; p_{T\,2}) P_1(x_3; p_{T\,3})$$

Can make an M_J prediction based upon the events *measured*

Don't need to be able to calculate M_J distribution from first principles

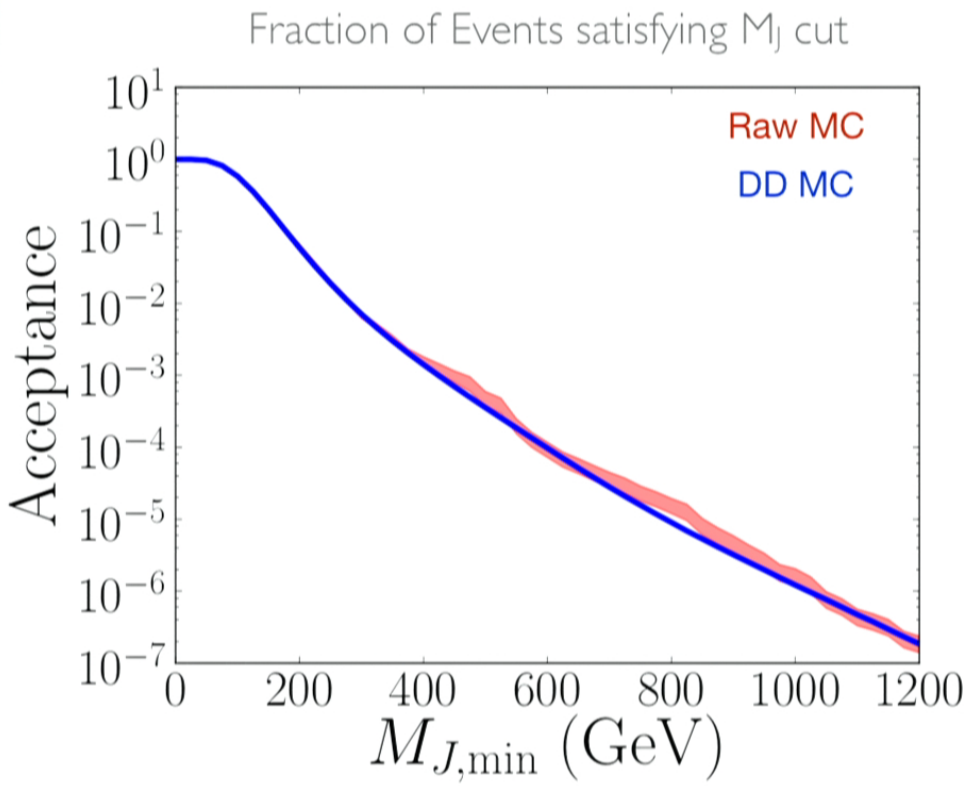
Works well in Monte Carlo

< 20% systematic differences



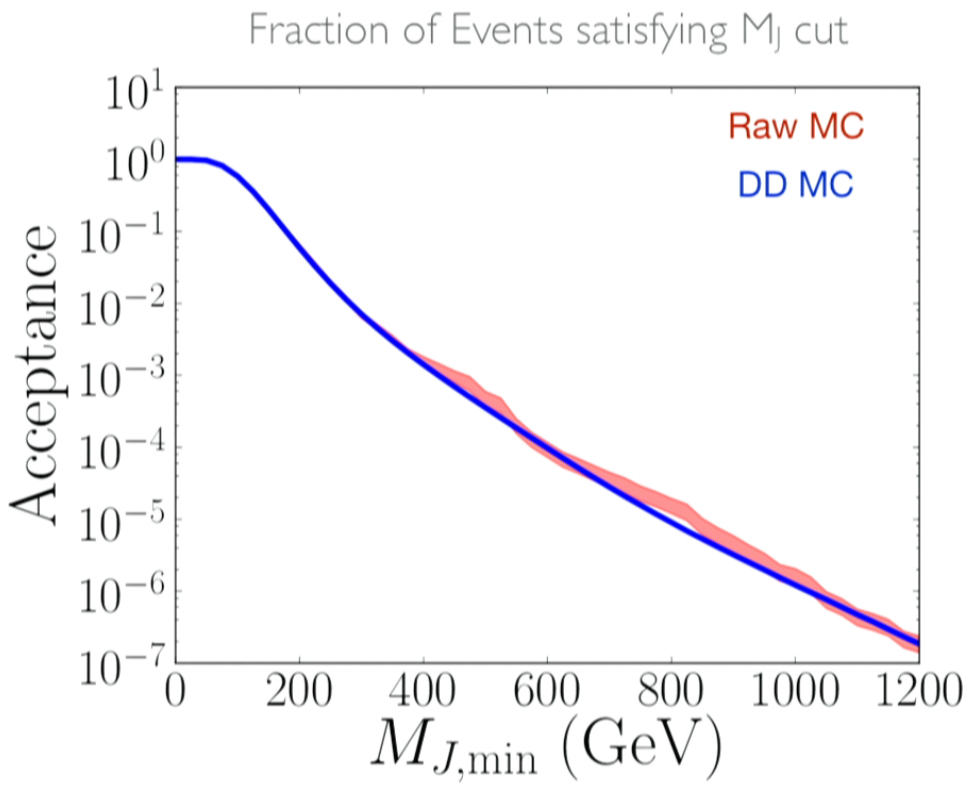
Works well in Monte Carlo

< 20% systematic differences



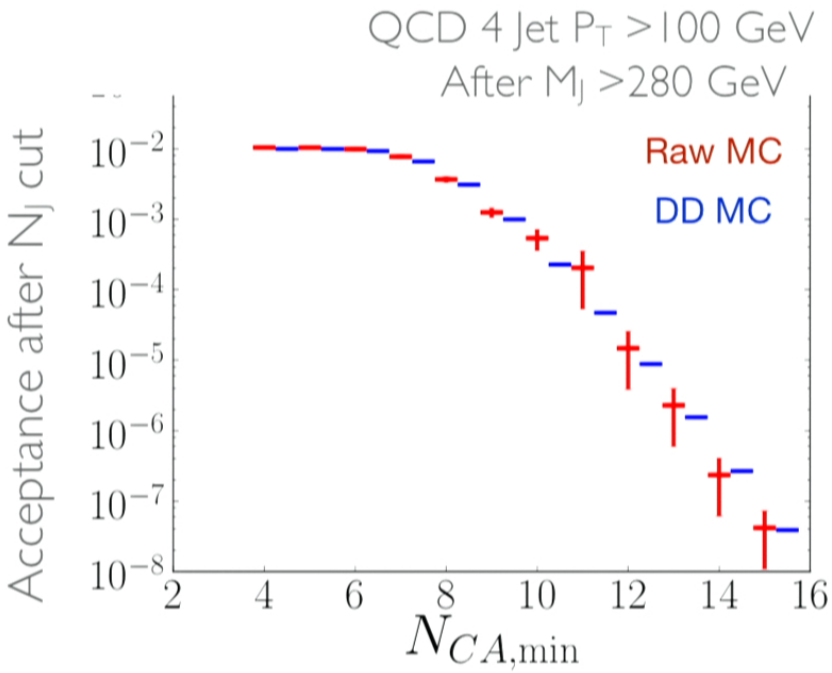
Works well in Monte Carlo

< 20% systematic differences



Can use Data Driven approach to M_J & N_J

$$\rho(m/p_T, n; p_T)$$



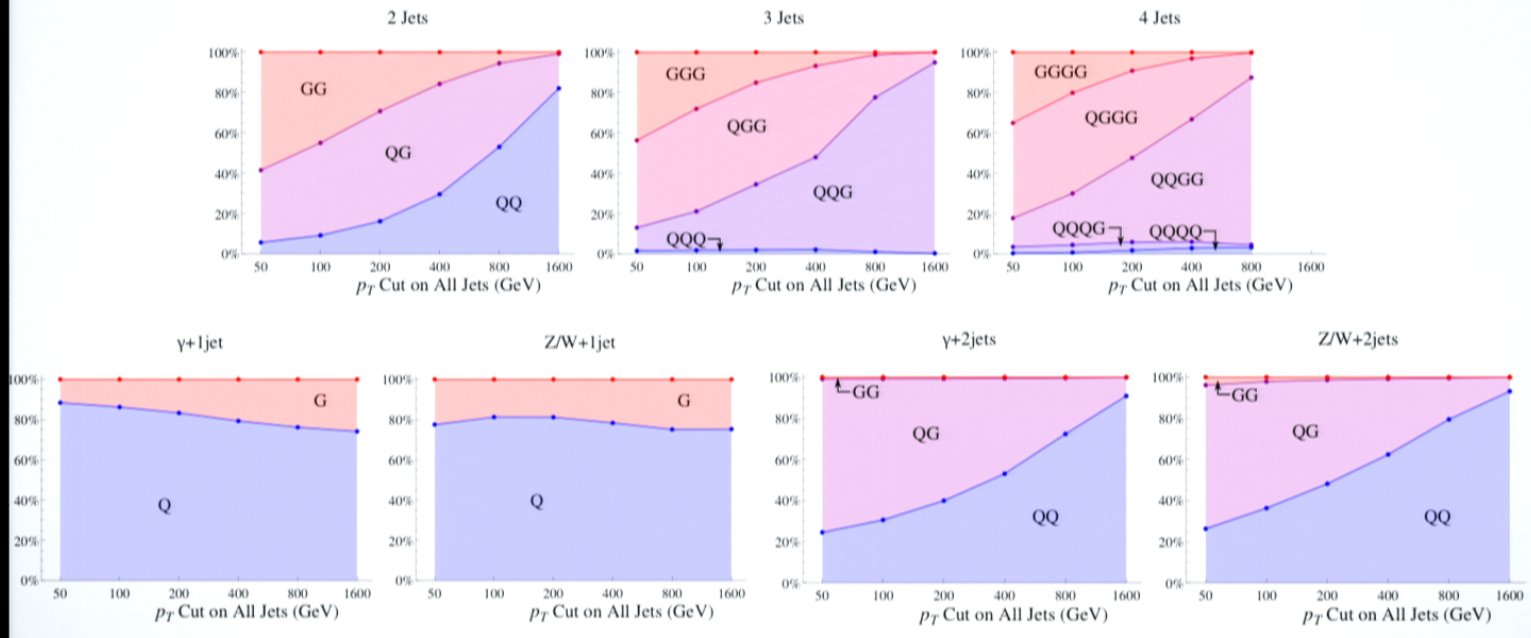
Can reduce QCD by 2^+ orders of magnitude

DESPERATELY SEEKING CORRELATIONS

Have seen no evidence yet of correlations

Actively looking for them

Looking at samples with different compositions



OUTLOOK

Approach of looking for multiple jets with anomalous properties has not been optimized

Appears that factors of 3-10 in expected sensitivity on $\sigma \times$ Br are possible

Use of data-driven approach should be explored

Q v. G Templates & Validation

Using N-subjets can be useful at classifying jets

e.g. don't compute τ_4 if n<5