Finding New Physics Using Jet Substructure: Two Examples

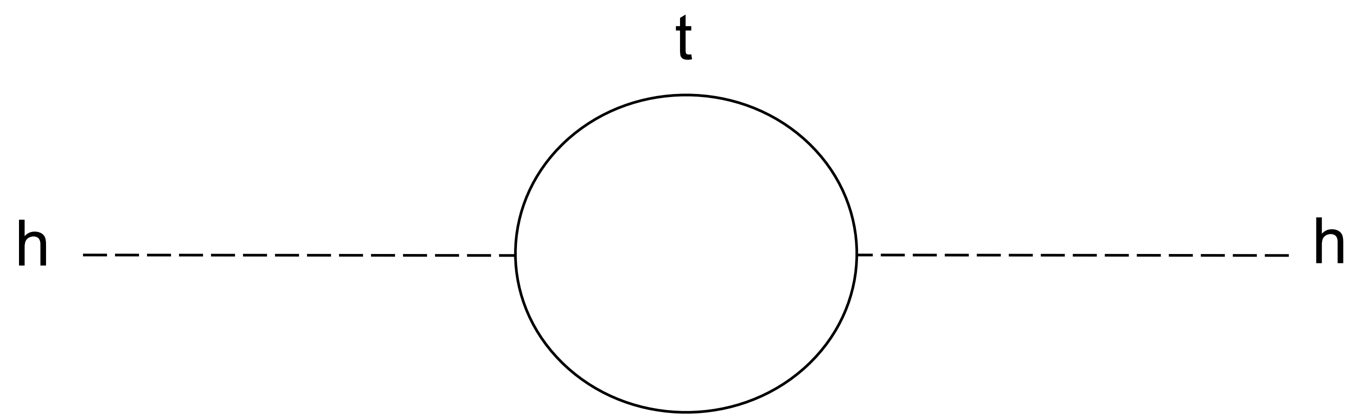
David Kaplan
ETH
6 Jan 2011

Nature is Fine-Tuned!

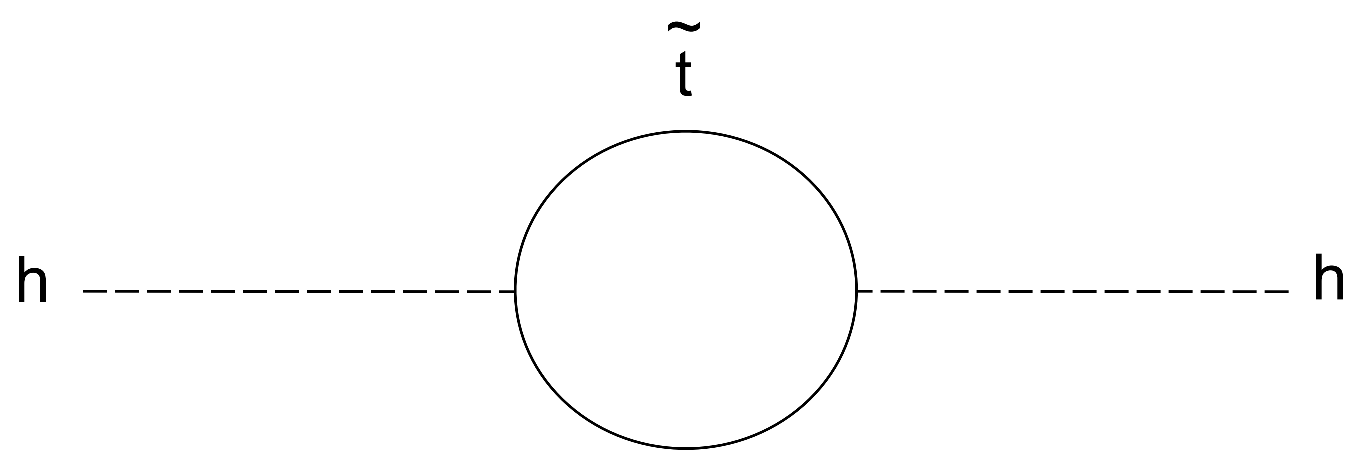
(or really ugly)

Non-discovery of the Higgs or any new particles beyond the standard model has us concerned.

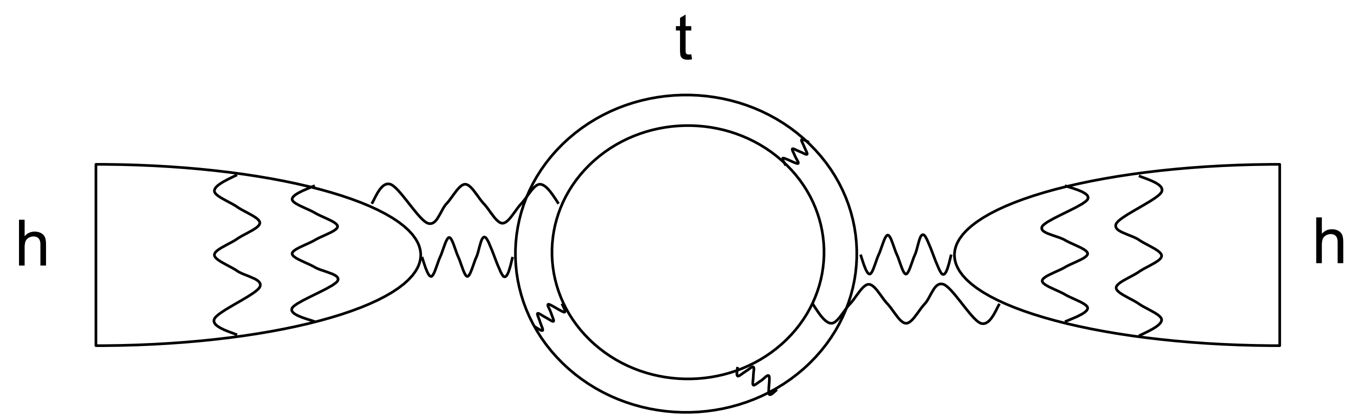
Higgs Mass Corrections



Higgs Mass Corrections

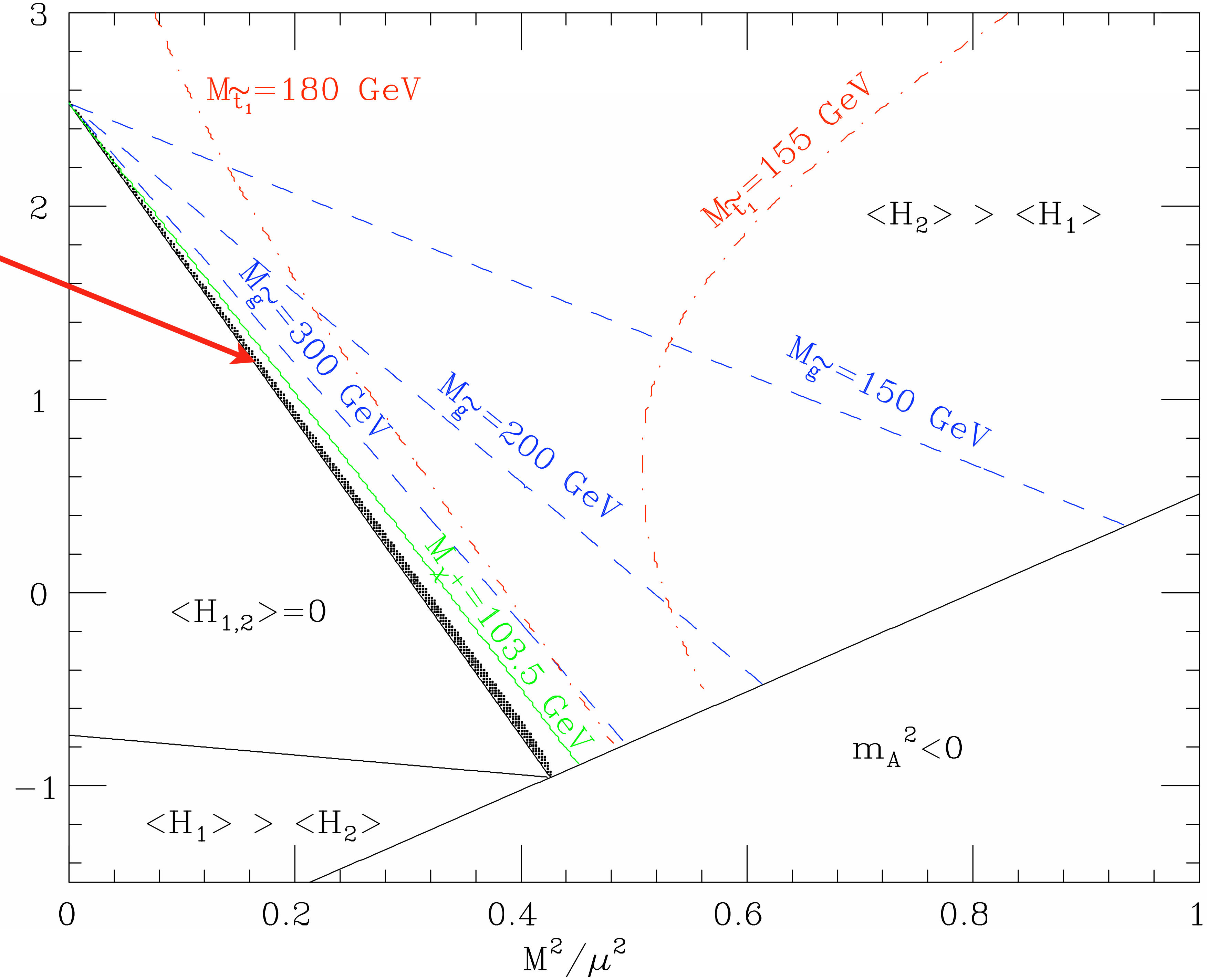


Higgs Mass Corrections



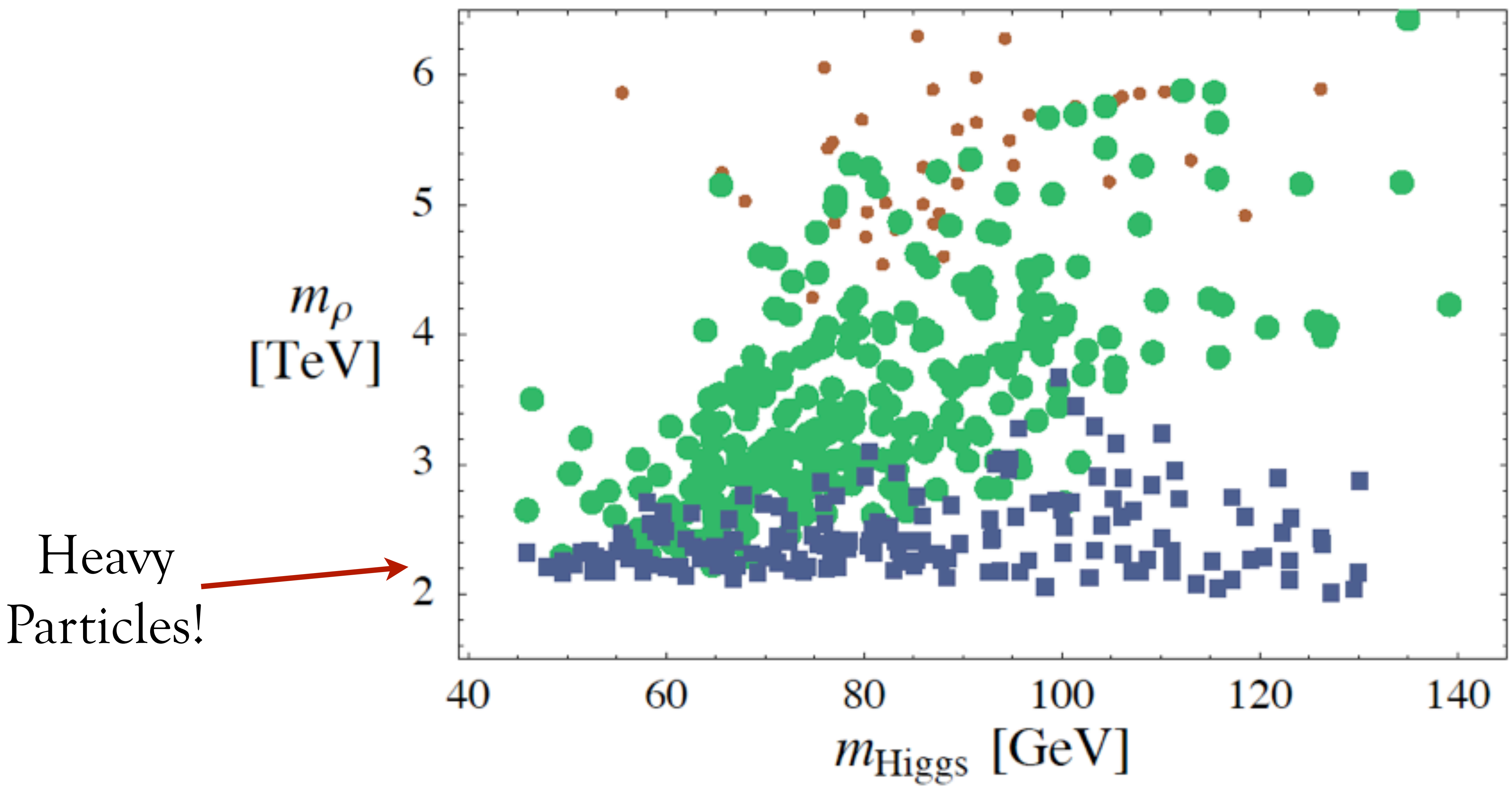
mSUGRA

Higgs mass >114 GeV



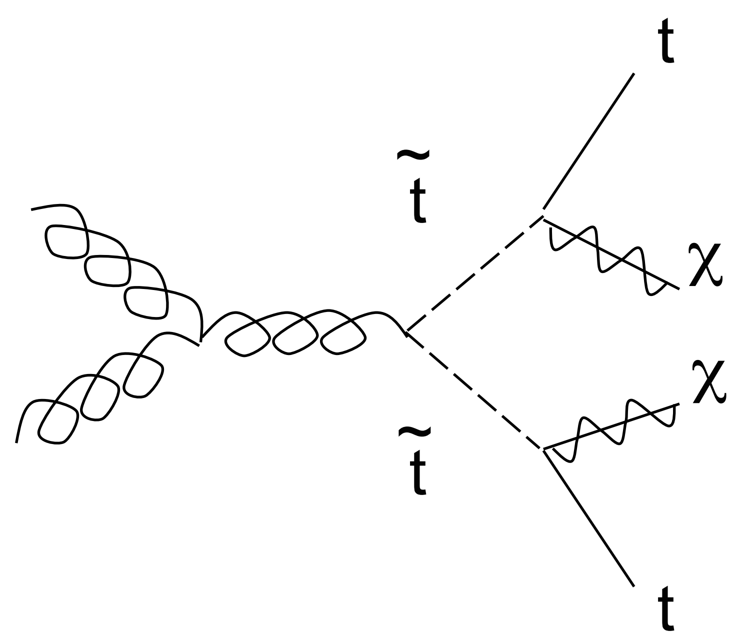
Giudice, Rattazzi '06

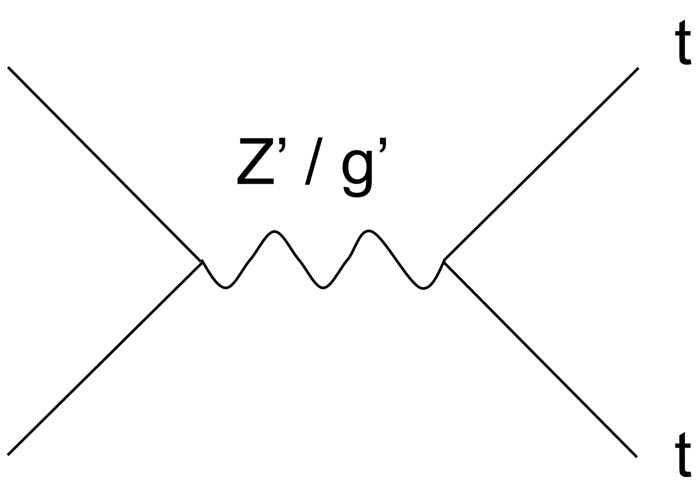
Composite Models



'Minimal Composite Higgs Model', Agashe, Contino, and Pomarol (2005)

Tops Produced



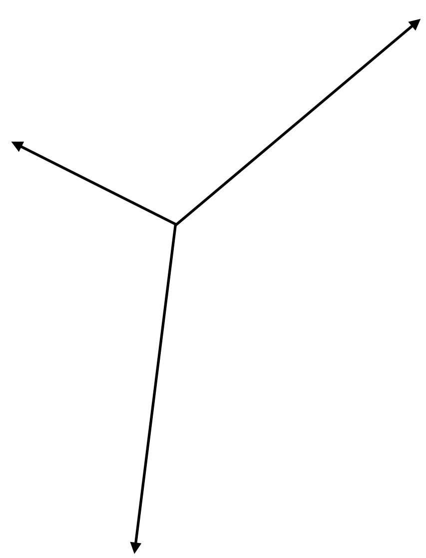


Why High Pt?

- SM works too well
  - Direct searches
  - Electroweak precision
- Suggests new particles above multi-hundred GeV
- LHC is a TeV-scale machine
  - We must watch out for *any* kinds of new heavy objects
  - Top coupling is a valuable measurement

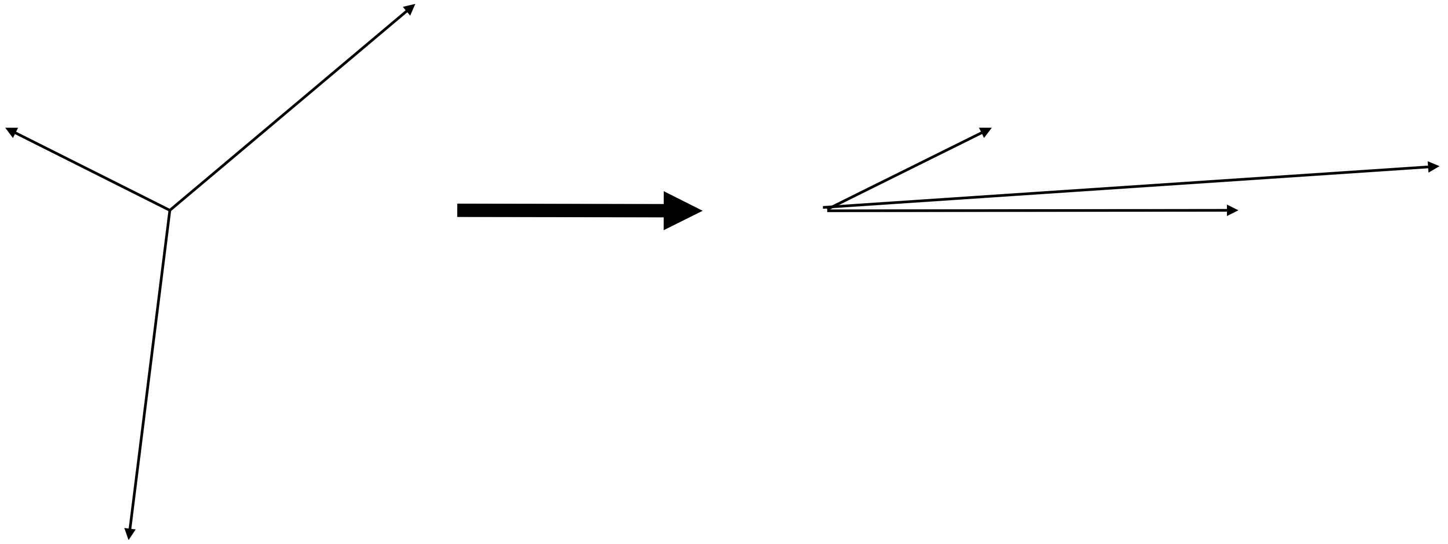
Tagging Tops

$$t \longrightarrow W\,b \begin{cases} l\,\bar{\nu}\,b \\ q\,\bar{q}'\,b \end{cases}$$

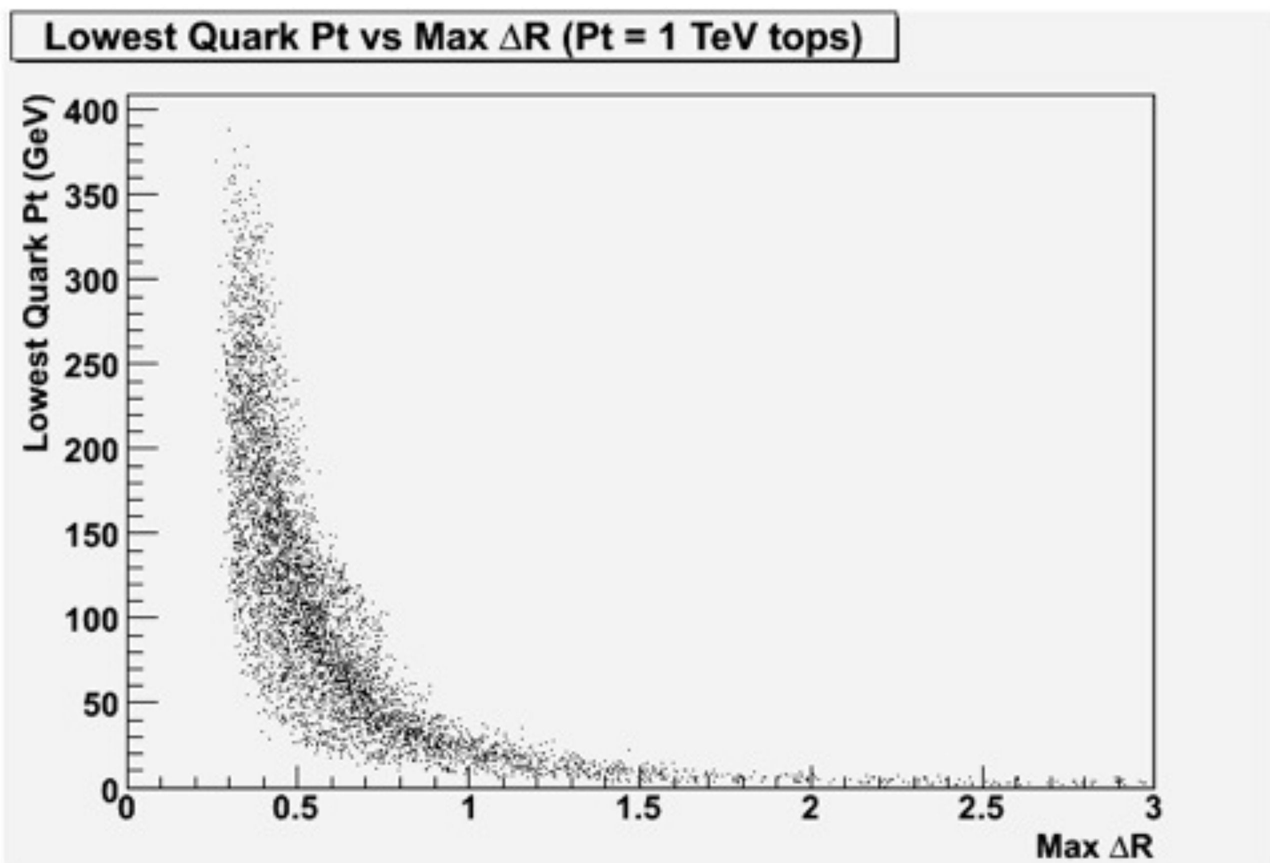


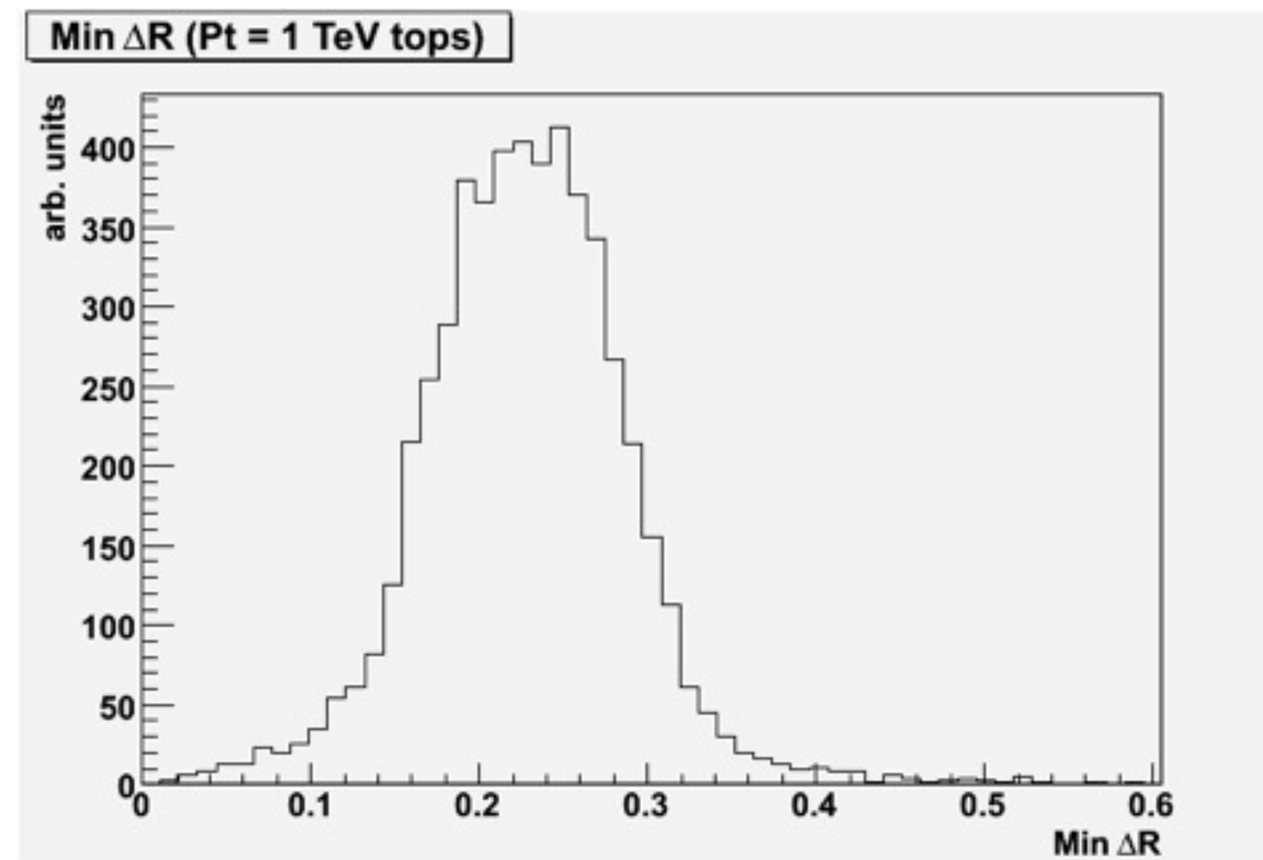
- Find 3 hard objects
- ID b-jet using displaced vertices
- Reconstruct top mass and W mass

...At High Pt



Boosted Top Kinematics at 1 TeV





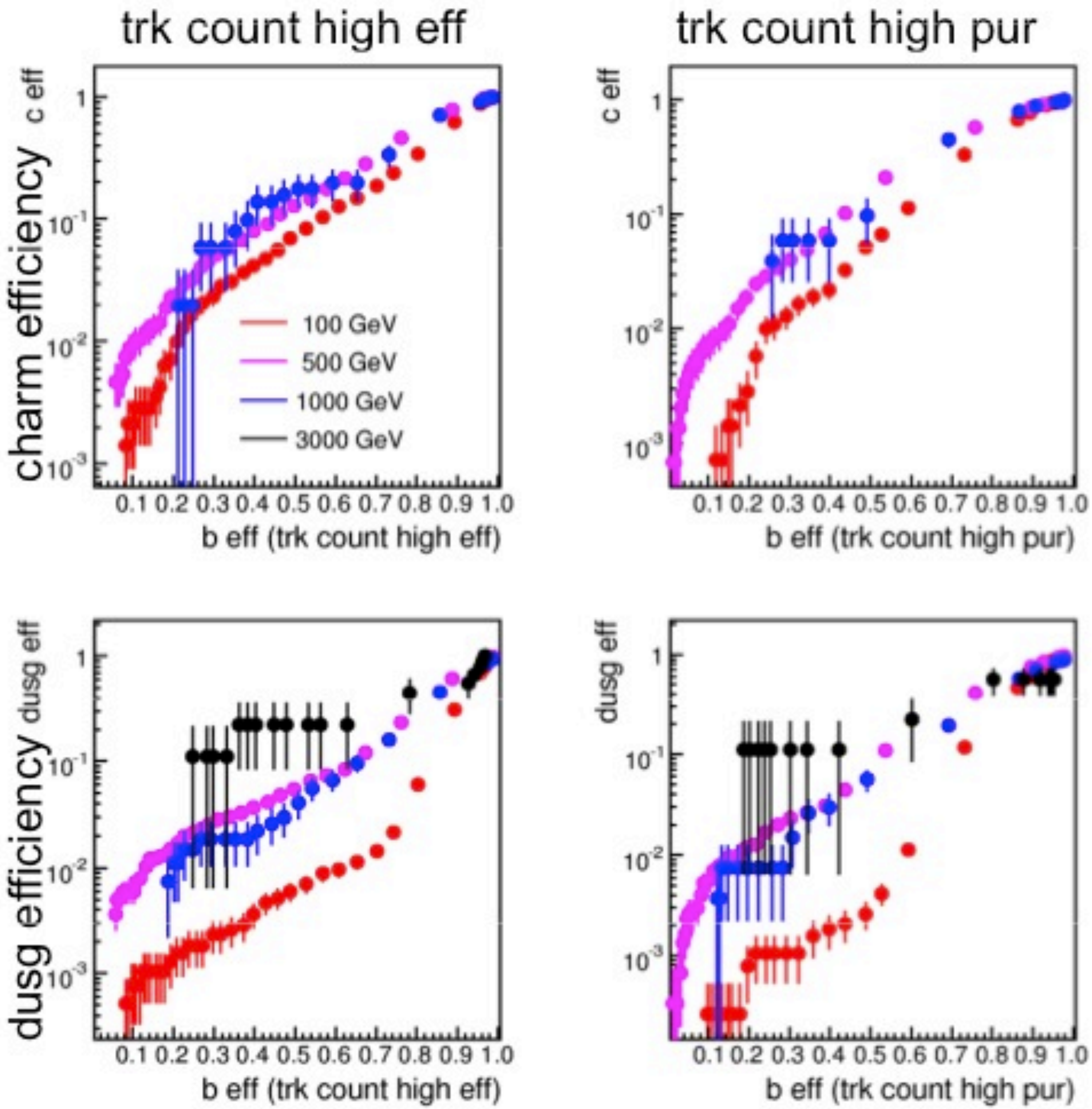
$\Delta R = \mathrm{sqrt}(\Delta\eta^2 + \Delta\phi^2)$

Standard Resonance Search

l+jets

- Cone jets of fixed size
  - Capture variable # of top decay products per jet
  - Lose kinematic info
  - Have to be very careful with the lepton (esp. electrons)
  - Subject to backgrounds you maybe shouldn't be worrying about
- Degraded b-tagging
  - At high Pt, tracks are crowded
  - Fake displaced vertices are a big issue, still under investigation
  - 1 TeV top: 20% b-tag / ~1% udsg mistag
    - Progressively worse at higher Pt
- Total signal efficiency ~ 1%

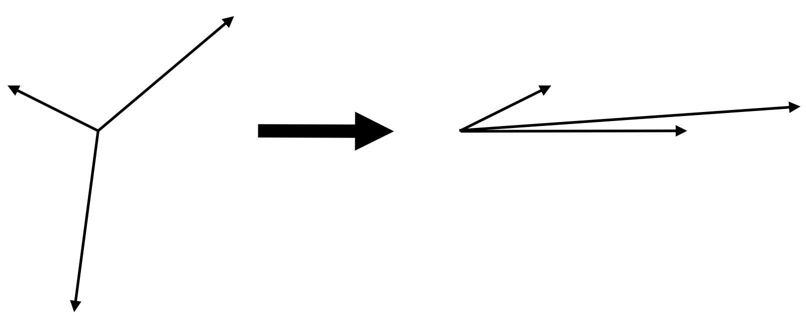
B-mistags for high-pT tops



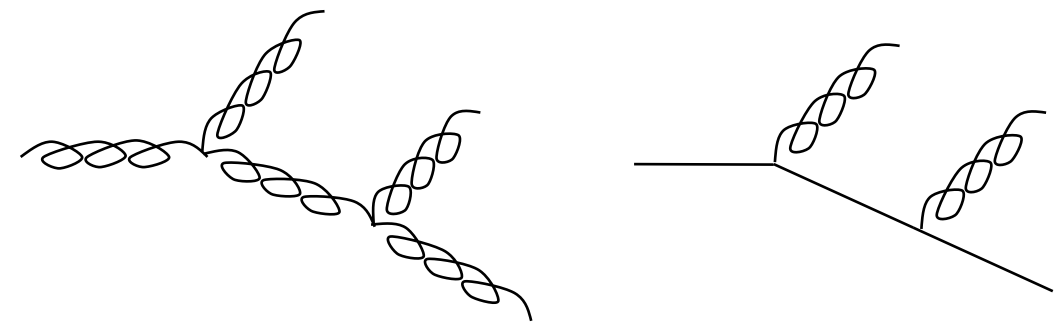
Mission Statement

- We would like some way to look inside these jets and use as much info as possible
- We would also like to free ourselves of reliance on b-tagging

Hadronic Tops vs Light Jets



- 3 hard partons
- Mass = m_t
- On-shell W
- ~Isotropic in top frame, comparable energies in lab

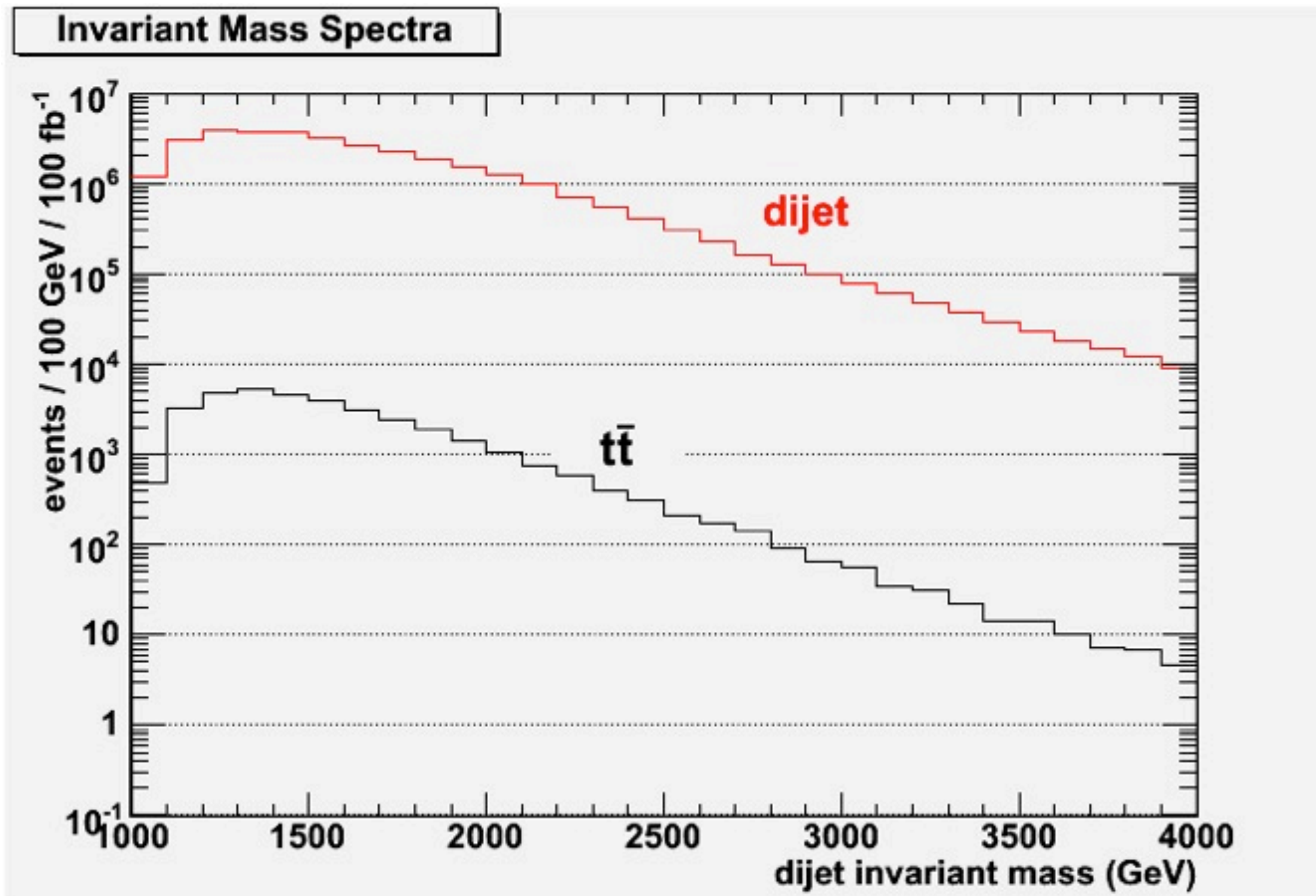


Variable # hard partons

Continuum of masses

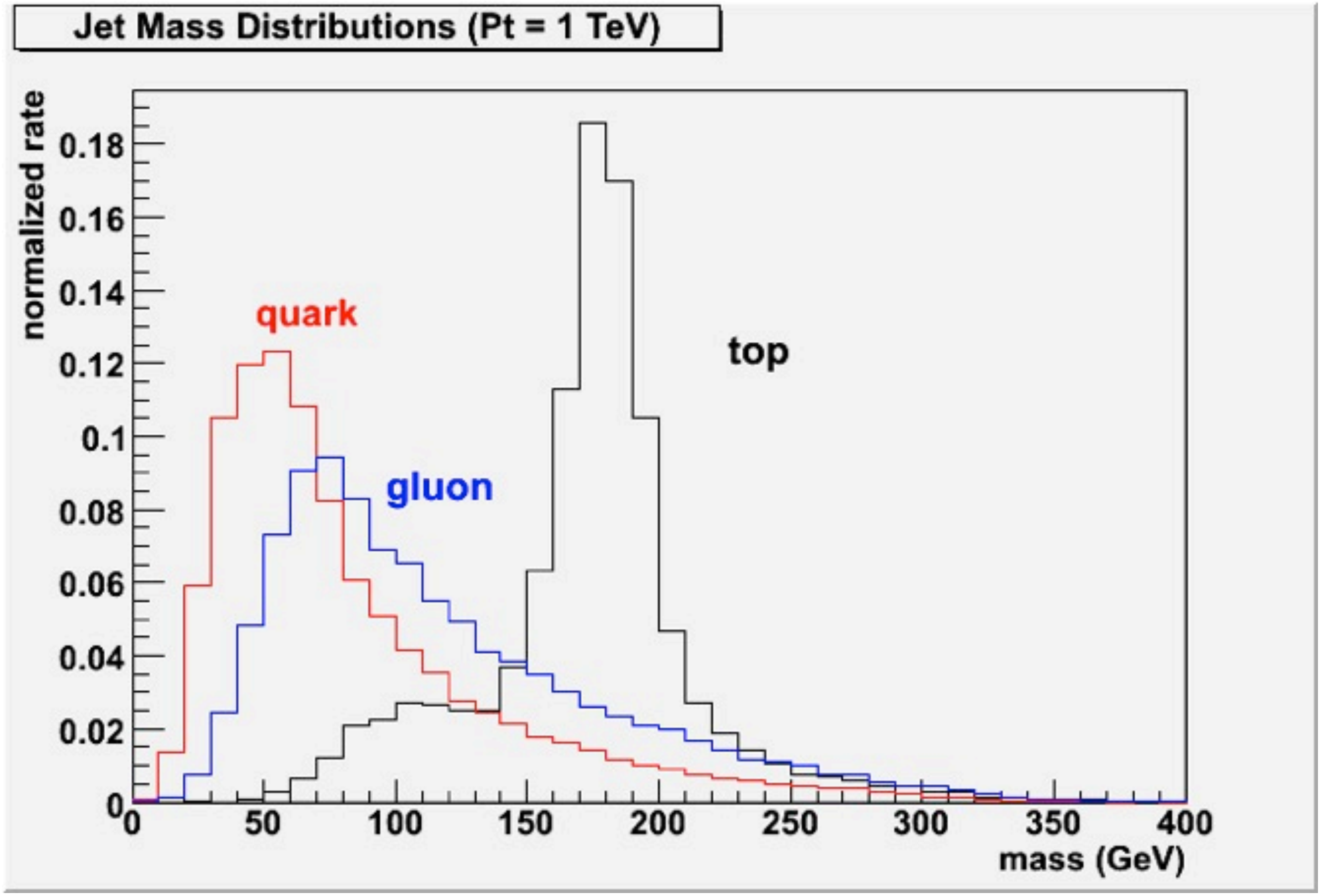
Soft/collinear singularities

Dijet Mass Spectrum

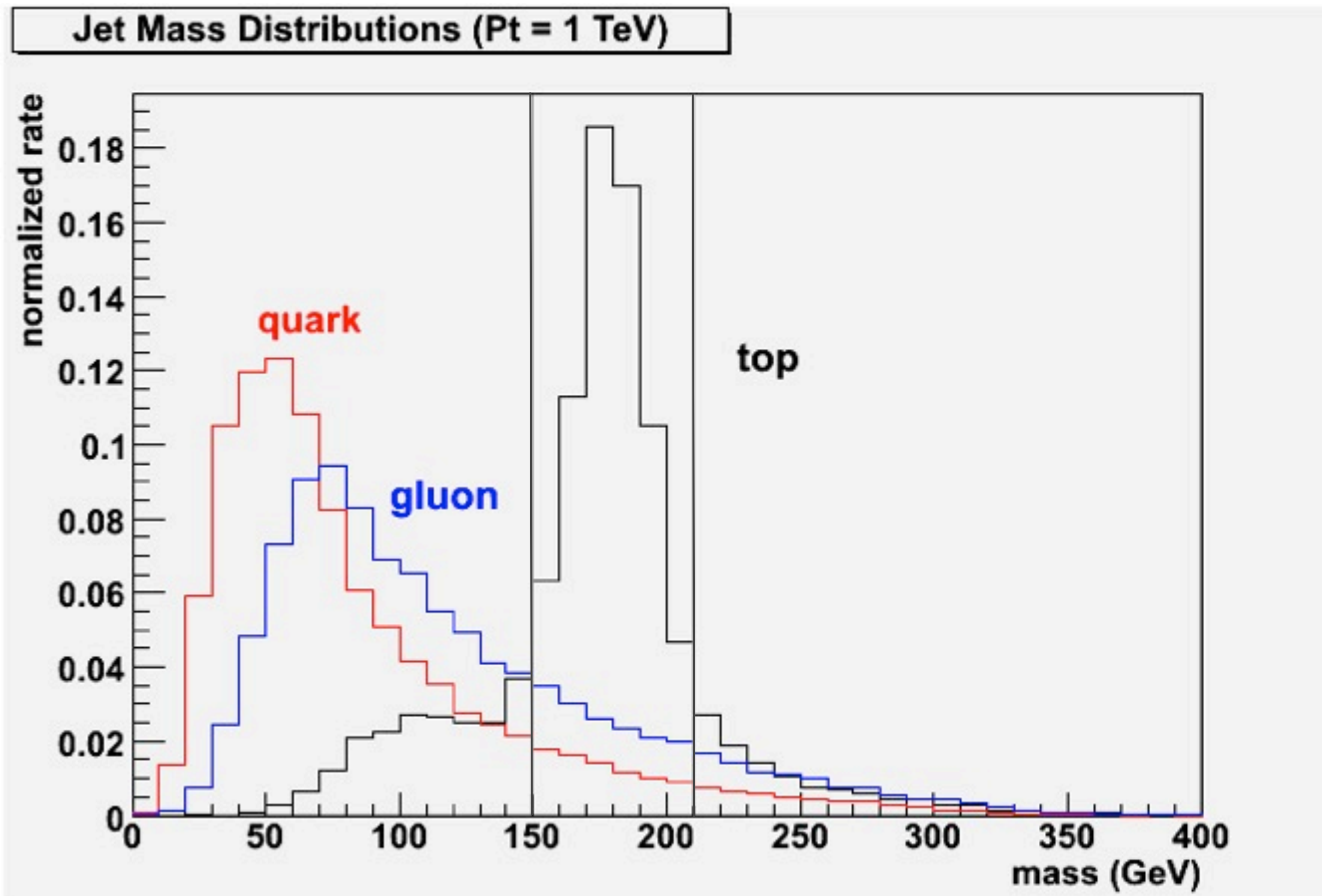


- All-hadronic tops
- PYTHIA 6.4 continuum QCD and top pair
- Pt > max(500 GeV, m/4)

First Pass: Mass Cut

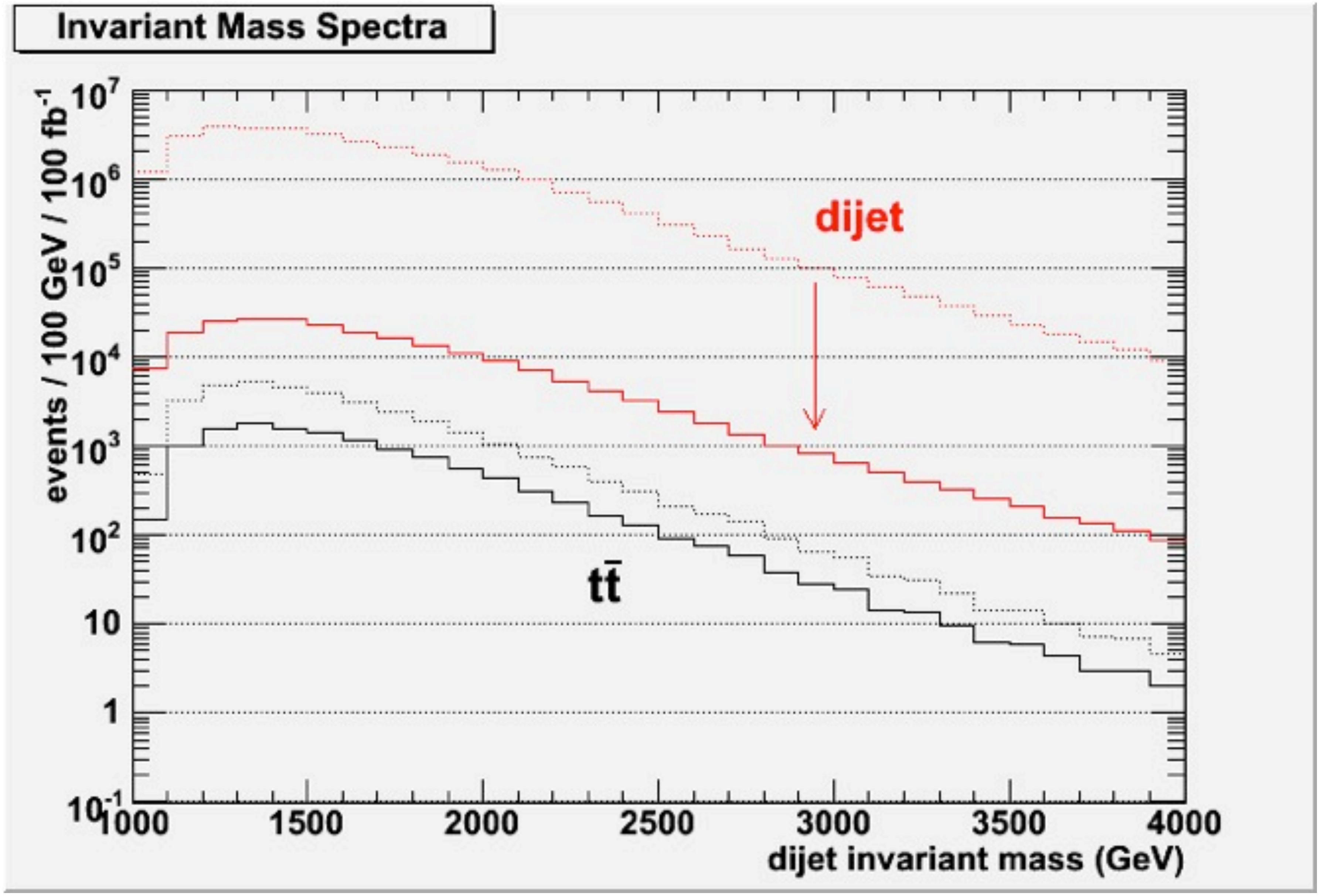


First Pass: Mass Cut

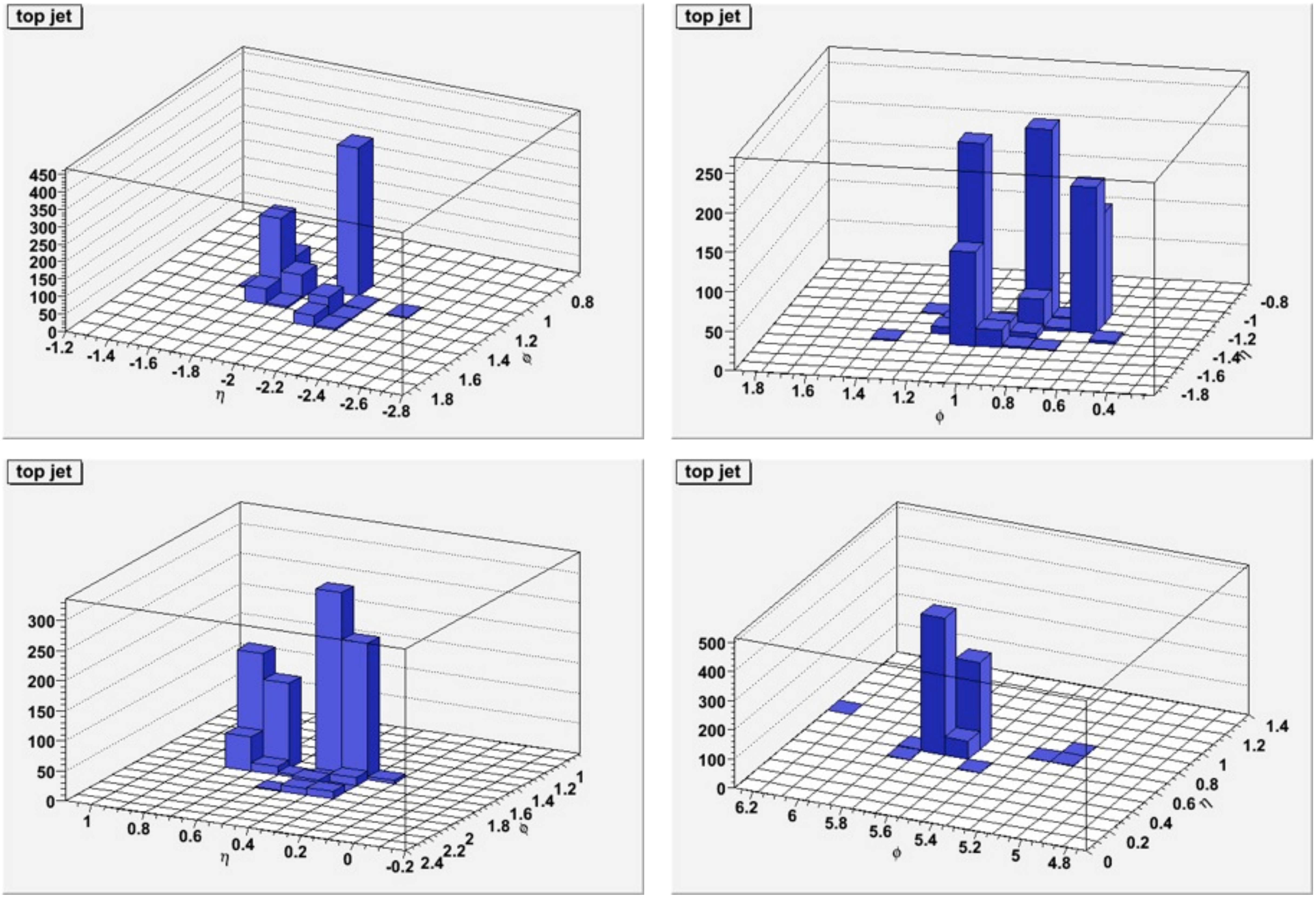


- 68% top
- 8% quark
- 16% gluon

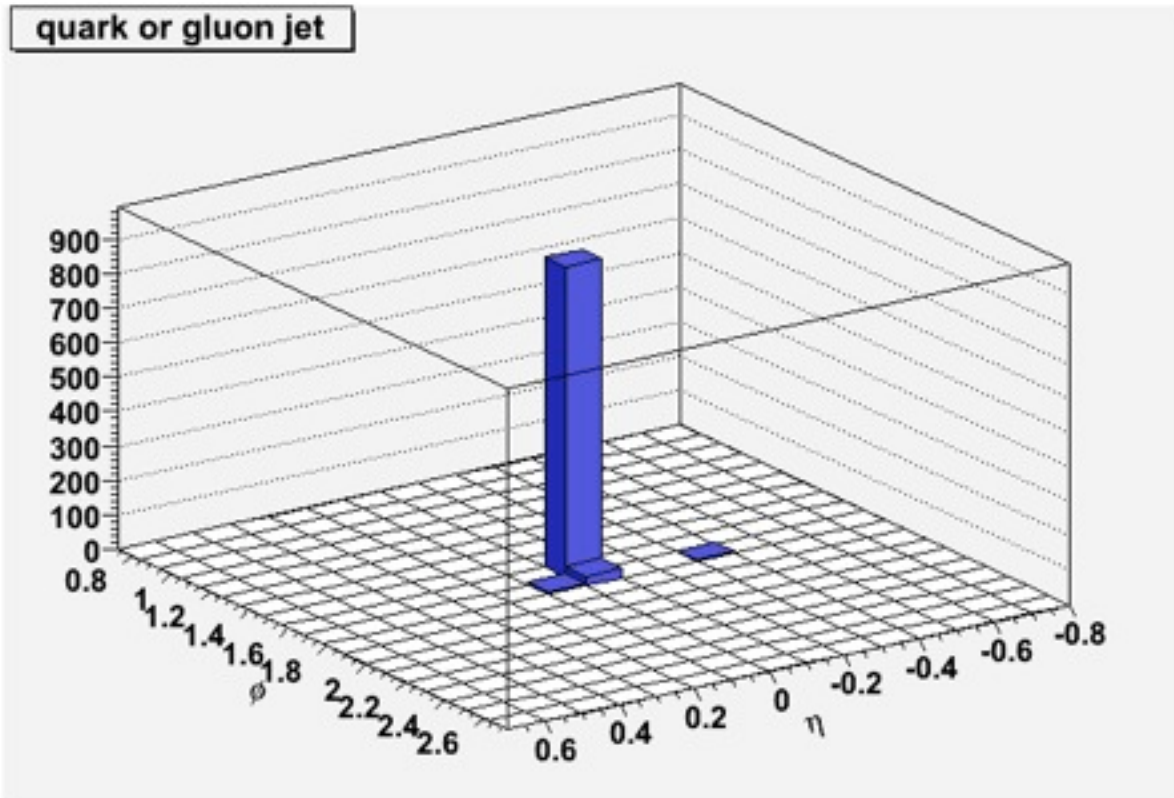
Dijet Mass Spectrum, Again

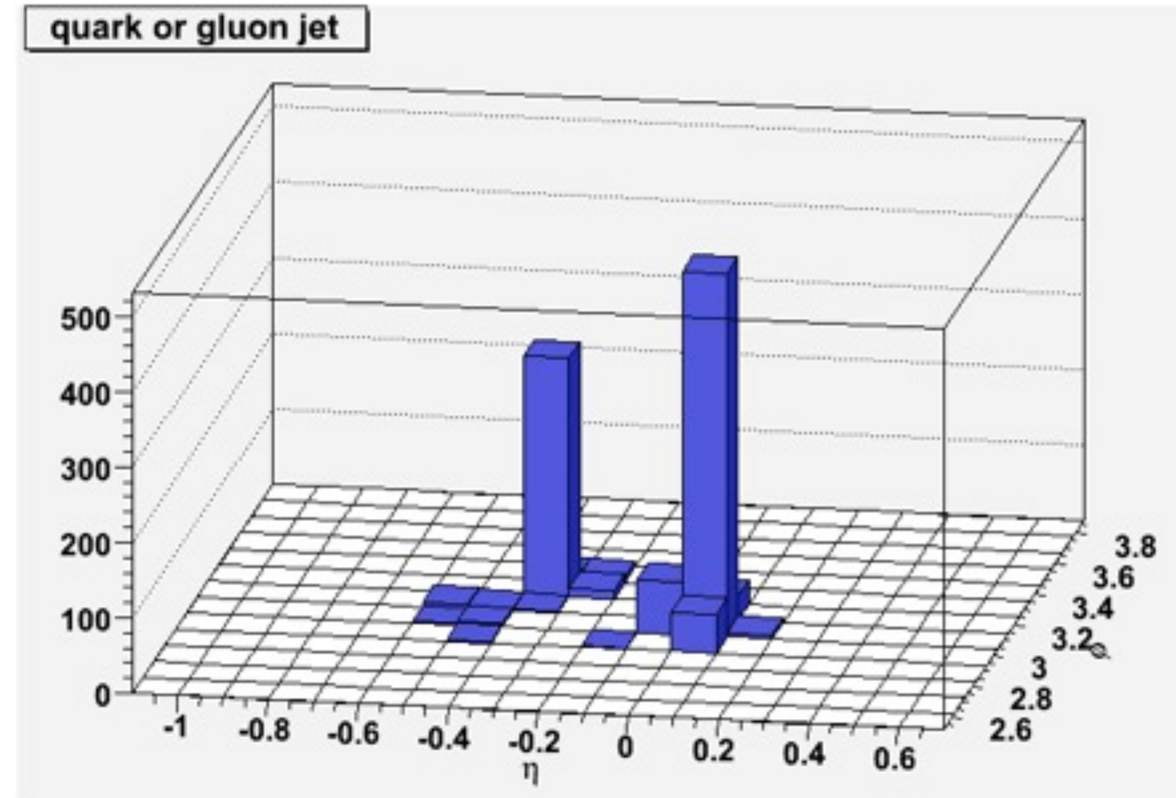


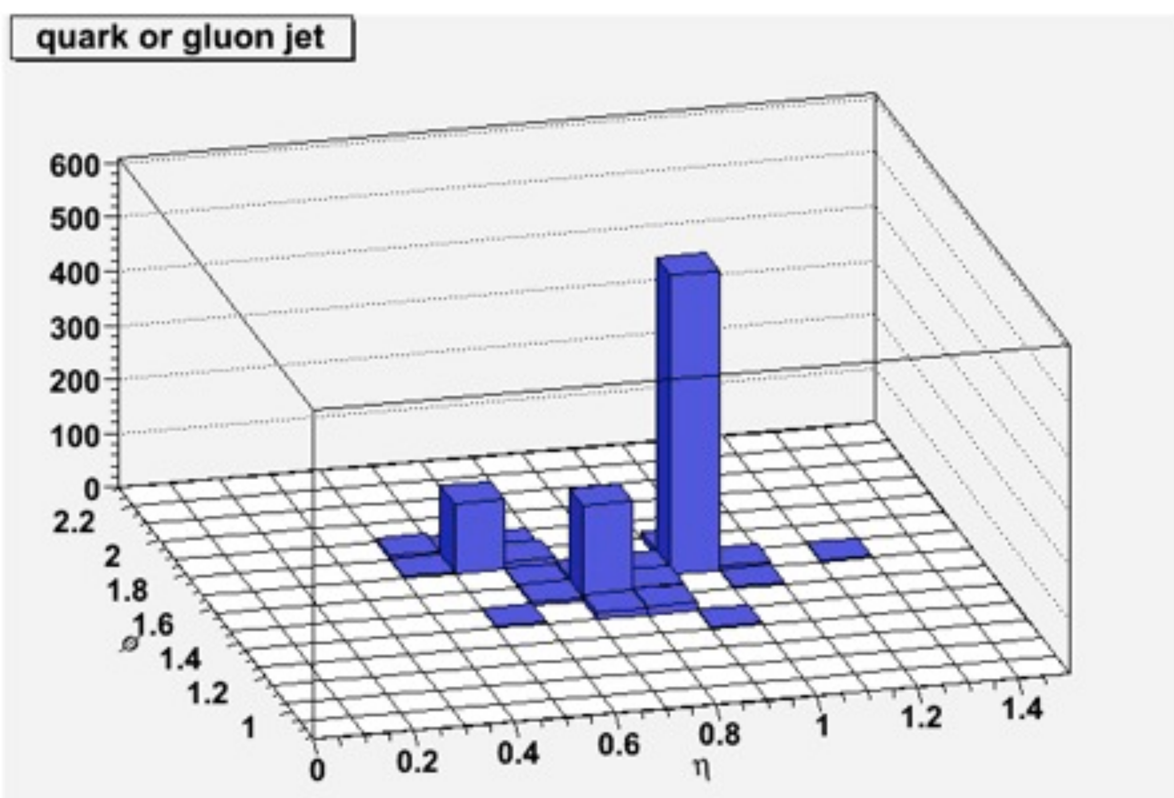
1 TeV Top-Jet Gallery

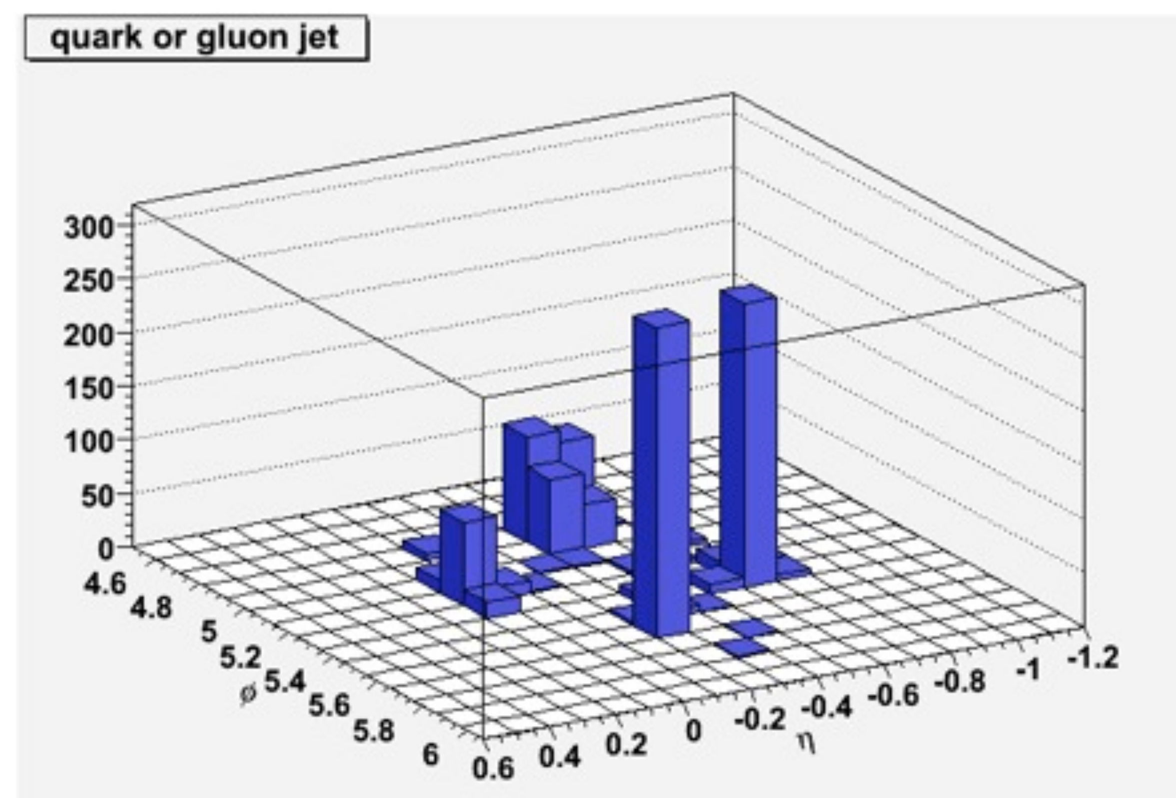


1 TeV Light-Jet Gallery









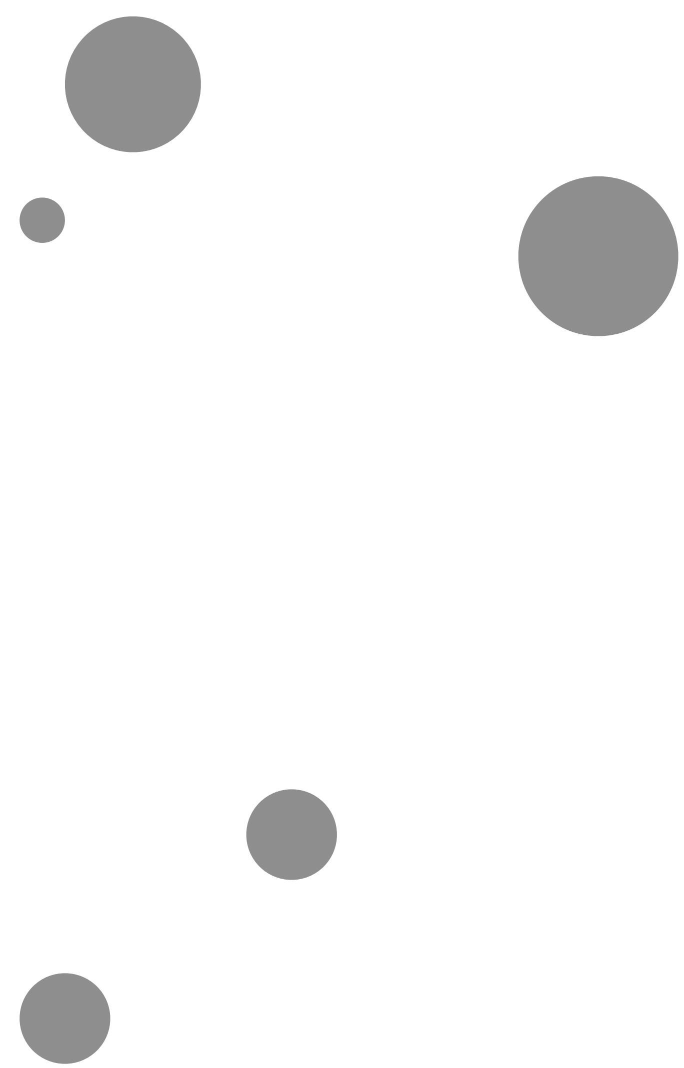
Our Take on the Problem

- Exploit the excellent calorimeter granularity of CMS and ATLAS to isolate the hard partons at $\Delta R \sim 0.1$
  - If they can be picked out by eye, they can be picked out by a computer program
- Employ both multiplicity and full kinematics as discriminators, as for low Pt
  - But give up on b-tagging
- Give up some conventional notions of what constitutes a "jet"

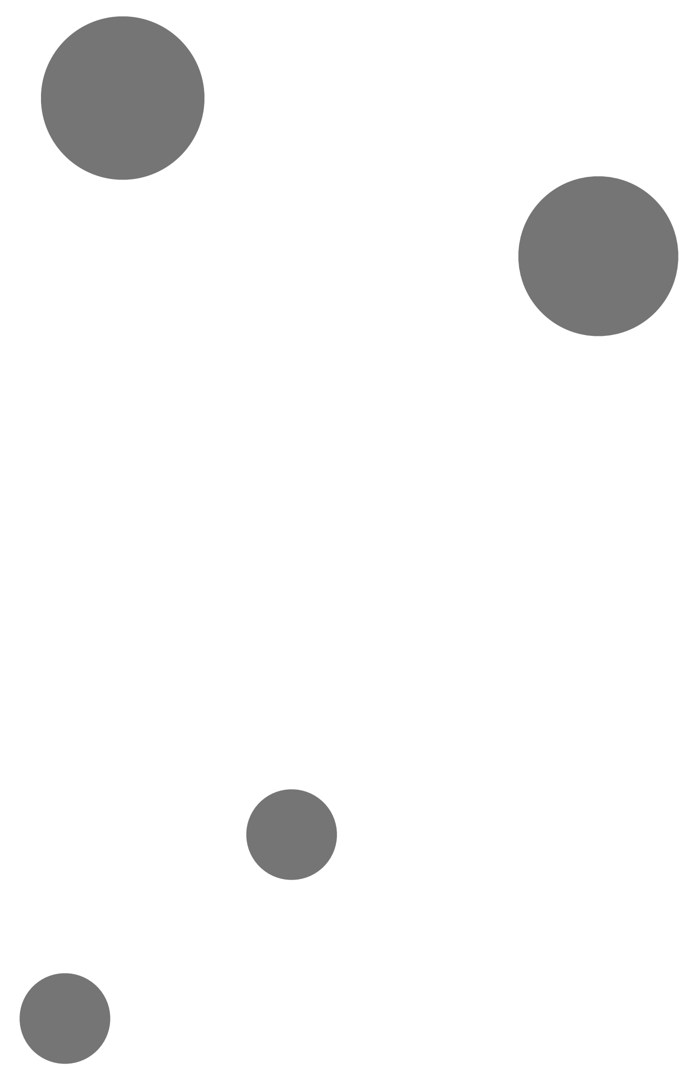
Cambridge/Aachen Algorithm

0. Calorimeter cells = massless 4-vectors
1. Calculate distance ΔR_{ij} between all pairs of 4-vectors
2. Stop if all $\Delta R_{ij} > R$, otherwise add together the closest pair and go back to Step 1

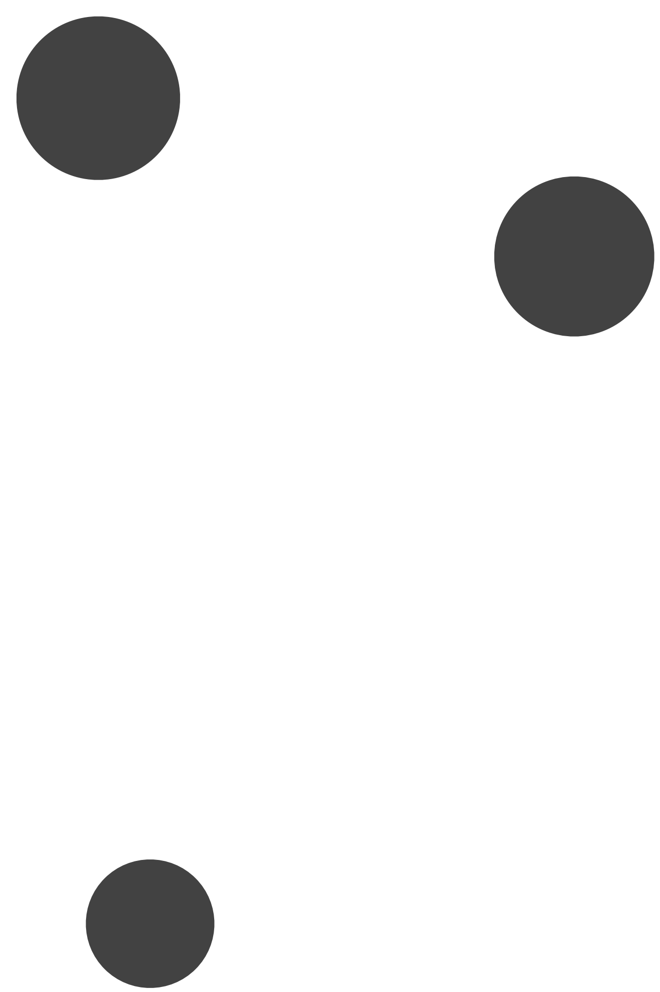
Cambridge/Aachen Algorithm



Cambridge/Aachen Algorithm



Cambridge/Aachen Algorithm



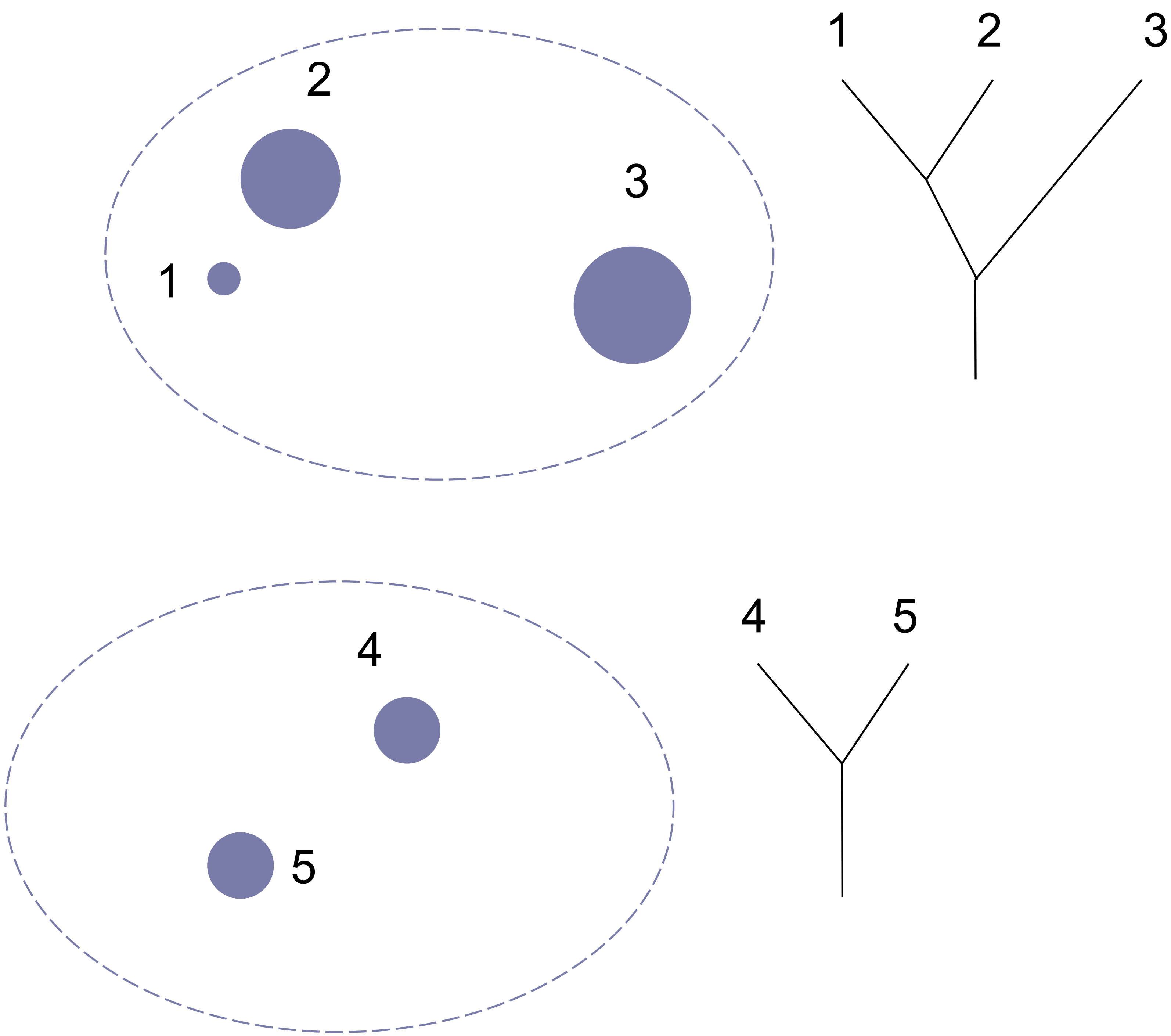
Cambridge/Aachen Algorithm

Cambridge/Aachen Algorithm

Cambridge/Aachen Algorithm

- Constructs a "parton shower history"
  - Sequence of 2-to-1 recombinations
  - OR…Sequence of 1-to-2 splits
- IR safe
- Identical in philosophy to k_t algorithm, but purely geometric measure
  - C/A: $D_{ij} = \Delta R_{ij}$
  - k_t : $D_{ij} = \min(p_{Ti},p_{Tj}) * \Delta R_{ij}$
- C/A is ideal for this kind of study, though k_t can also be used (just have to be more careful)

Clustering History Trees



Our Inspiration

- Jon Butterworth, et al boosted Higgs ID
- Cluster/decluster strategy
  - First cluster event with a large R using C/A
  - Identify a jet of interest, then *reverse* the clustering steps in search of substructure
- Extract "subjets"
  - Hard clusters of energy inside a jet
  - Size/distance is not fixed
- They proceed to recluster the jet…we will work directly with the subjets

J. Butterworth, A. Davison, M. Rubin, G. Salam, PRL 100:242001,2008
[arXiv:0802.2470 [hep-ph]]

fastjet

- Authors: M. Cacciari, G. Salam, G. Soyez
- Standalone C++ code
- 3 sequential recombination jet algorithms, various options
  - k_t, C/A, anti-k_t
- Stores the entire clustering tree
- Very user-friendly

http://www.lpthe.jussieu.fr/~salam/fastjet/

Our Algorithm, Part I

0. Cluster event with C/A and look at individual jets
   1. Decluster jet one step. Throw away softer object if its Pt $< \delta_p$ and continue declustering.
   2. Stop declustering if:
      1. Both objects Pt $> \delta_p$. These are **subjets**.
      2. Both objects Pt $< \delta_p$
      3. Objects are "too close": $|\Delta N_\eta| + |\Delta N_\phi| < \delta_N$
      4. Only one object is left

(items 2–4:) declustering fails, rebuild original jet

Our Algorithm, Part II

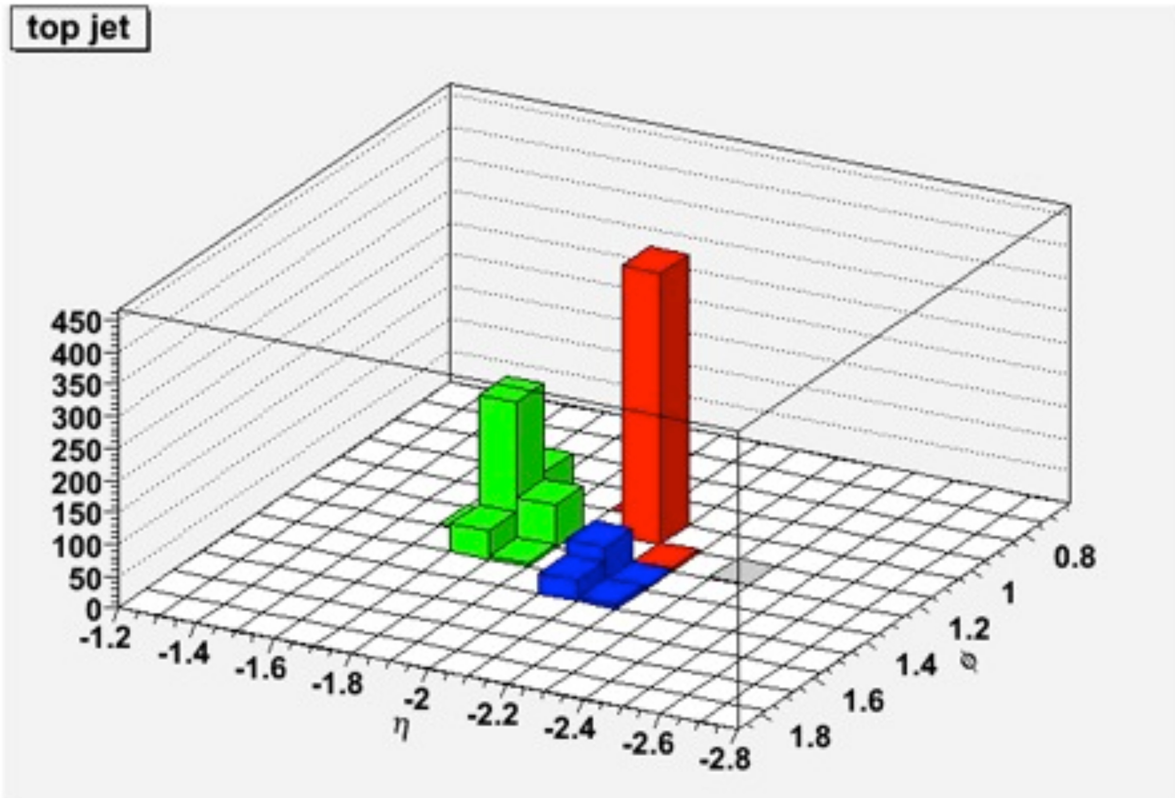
3. If the jet breaks into two subjets, repeat declustering on those subjets
4. Keep cases with 3 or 4 final subjets (4th is rare and tends to be soft)
5. Apply kinematic cuts

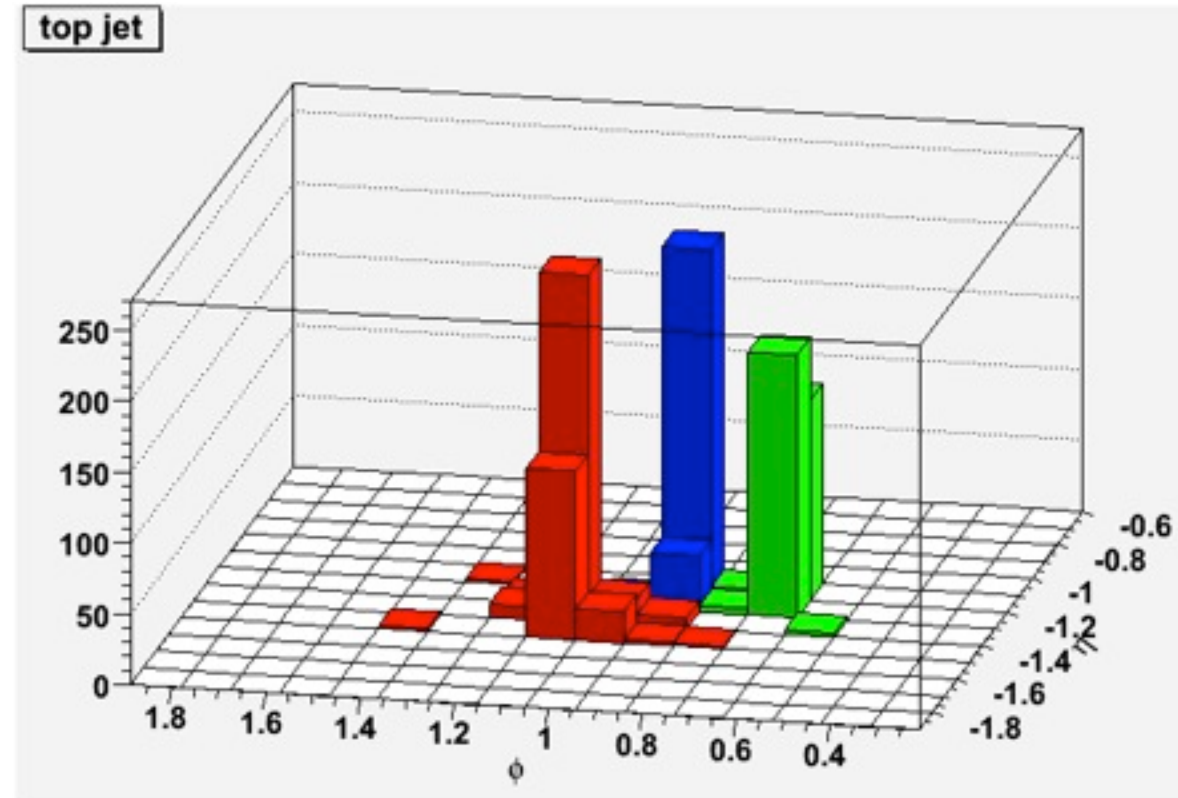
Details

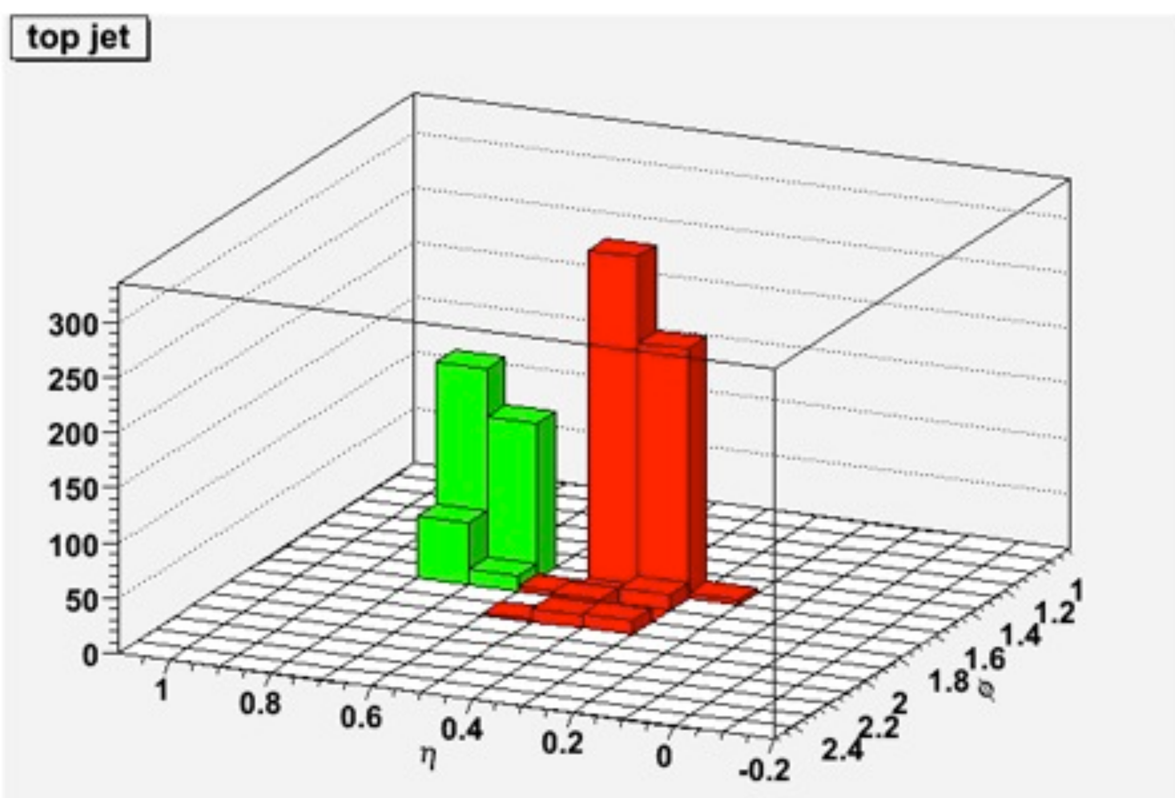
- C/A R parameter, δ_p, and δ_N picked according to event $H_T > \{1, 1.6, 2.6\}$ TeV
  - R = {0.8, 0.6, 0.4}
  - δ_p = Pt * {0.10, 0.05, 0.05}

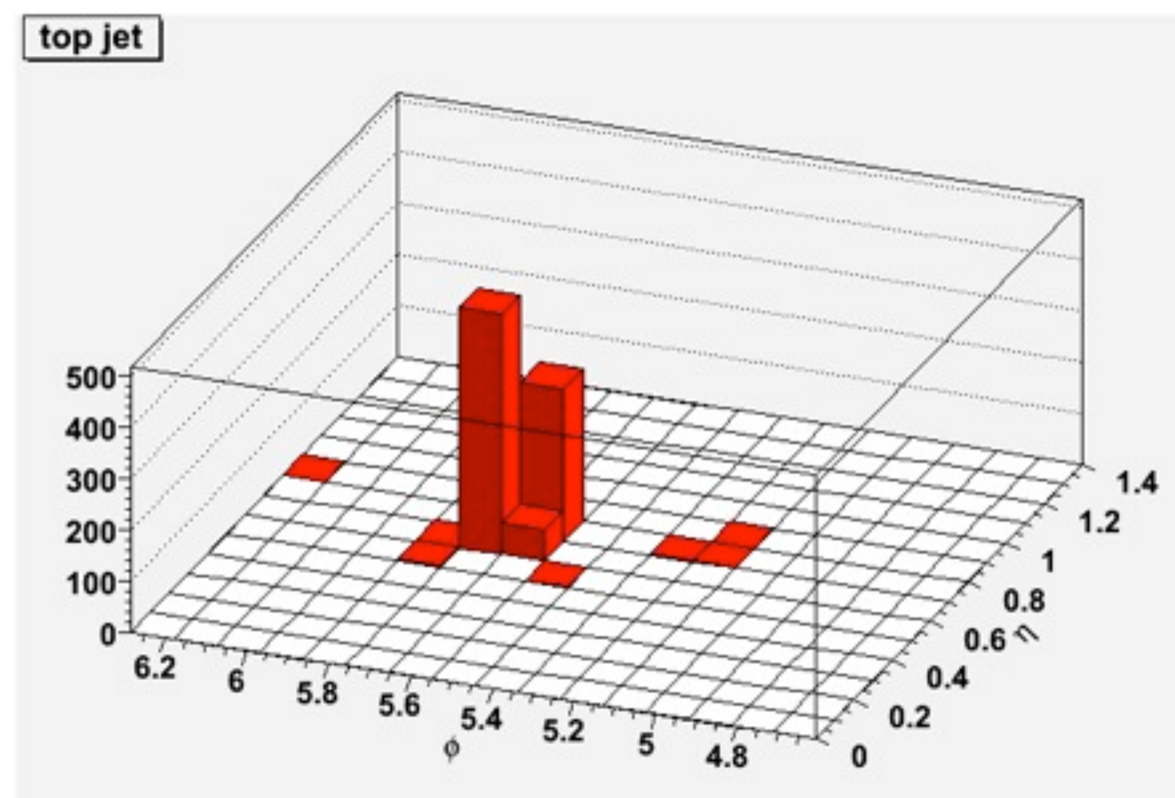
    Same δ_p for both declustering stages
  - δ_N = {1.9, 1.9, 1.9}
- Also jet reconstruction criteria
  - Pt > max(500 GeV, 0.7*H_T/2)
  - $|\eta| < 2.5$

1 TeV Top-Jet Gallery

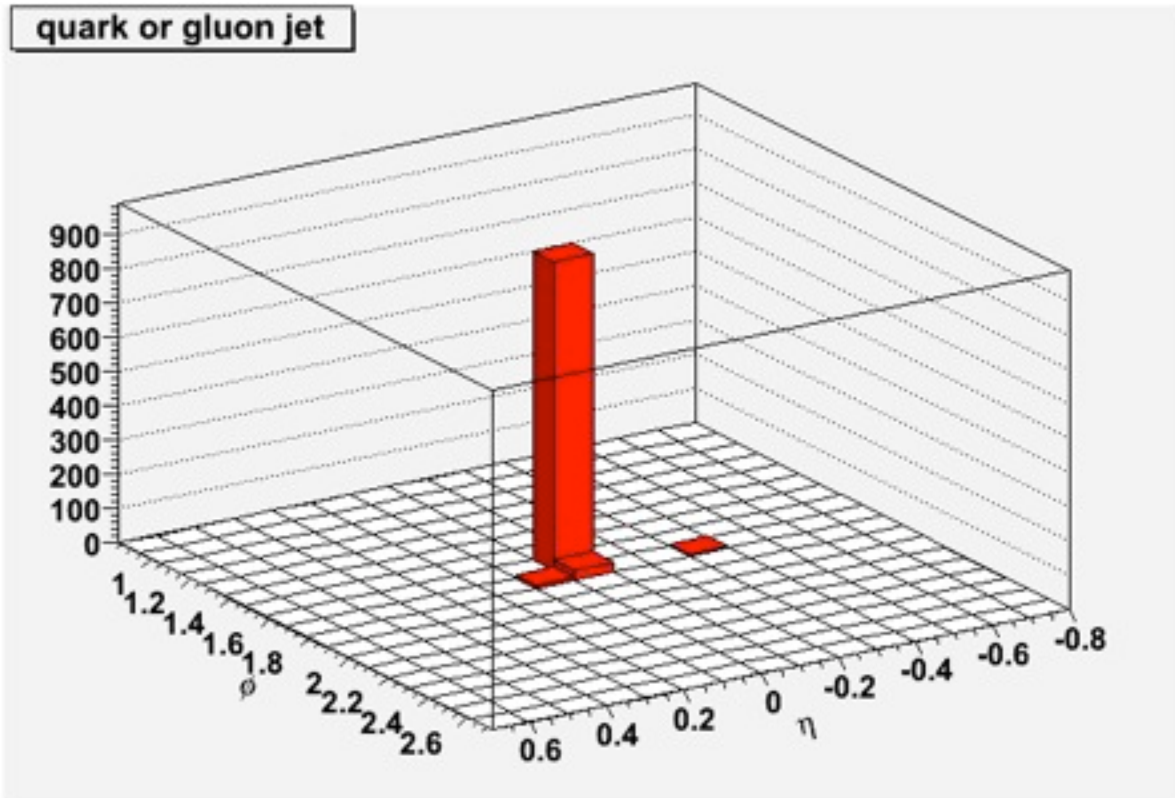


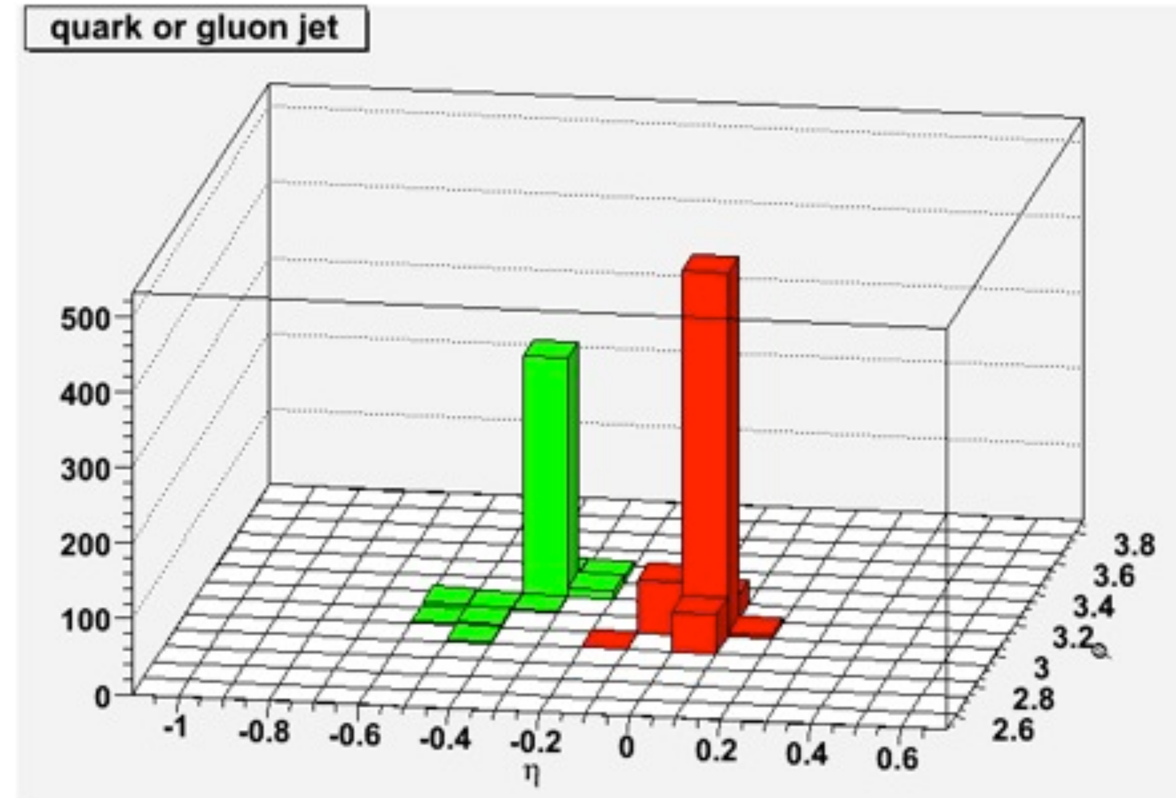


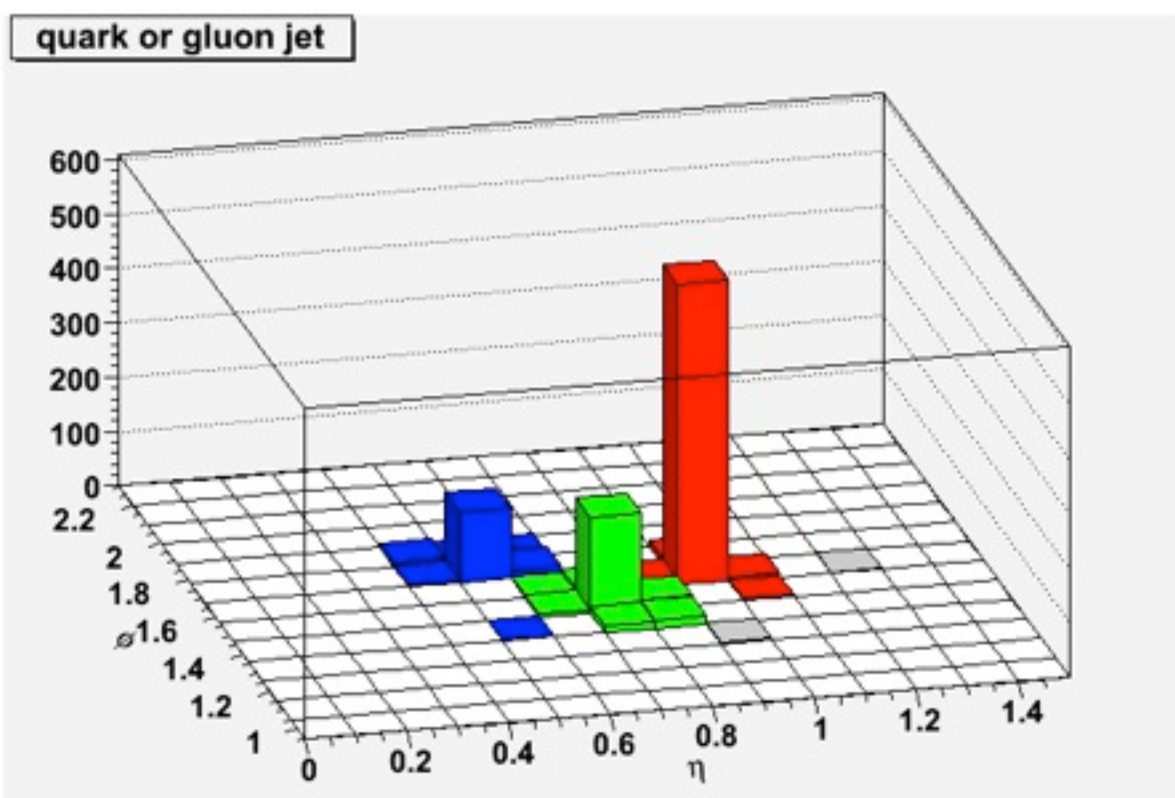


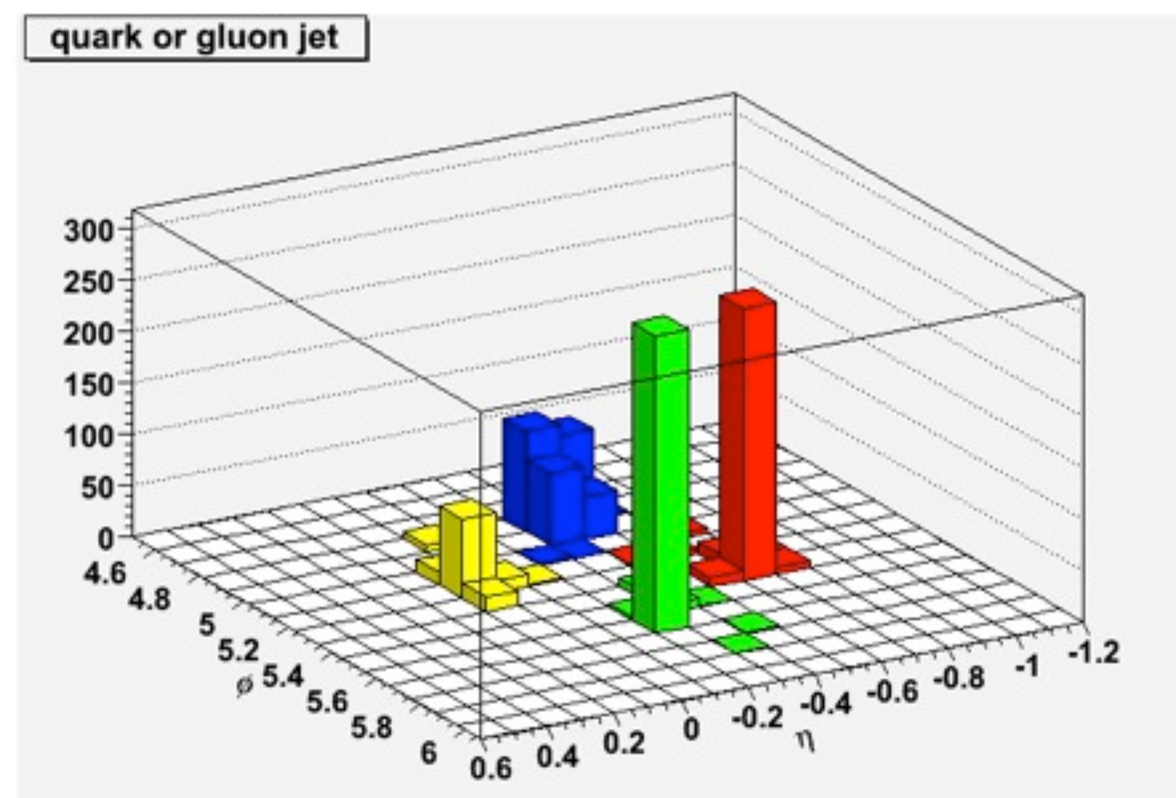


1 TeV Light-Jet Gallery

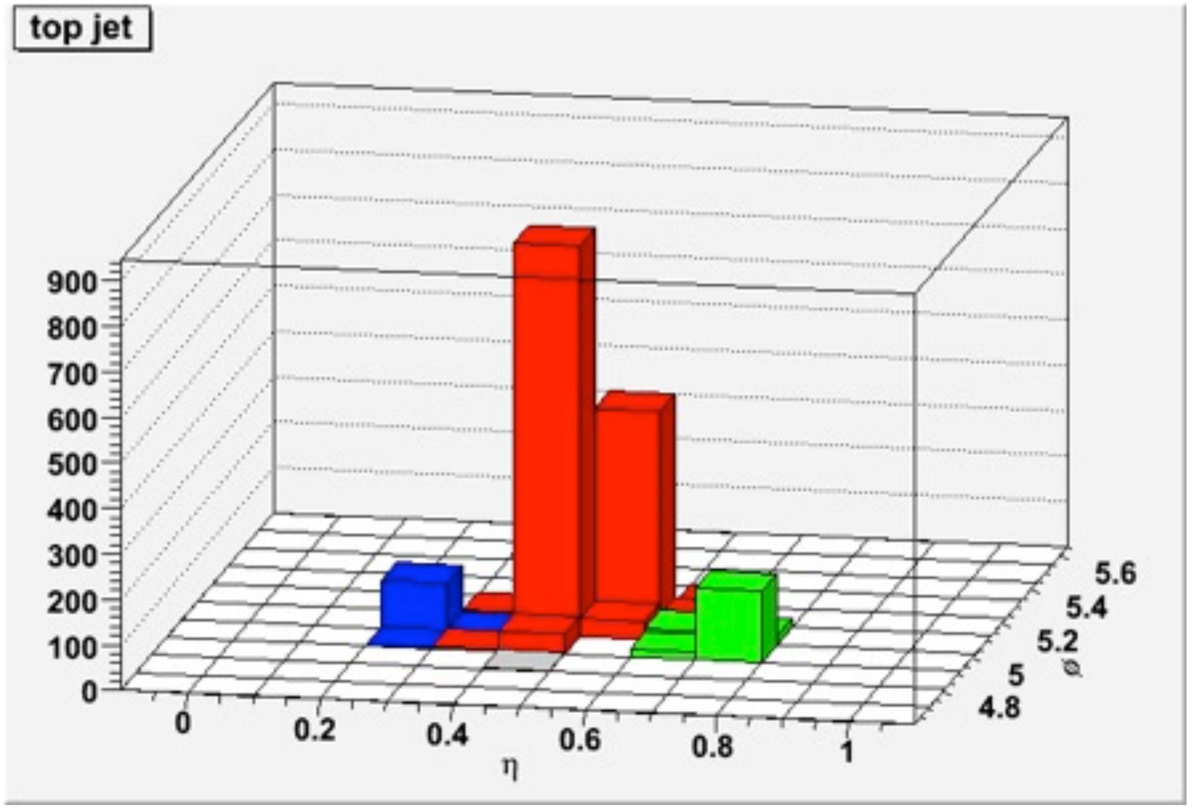


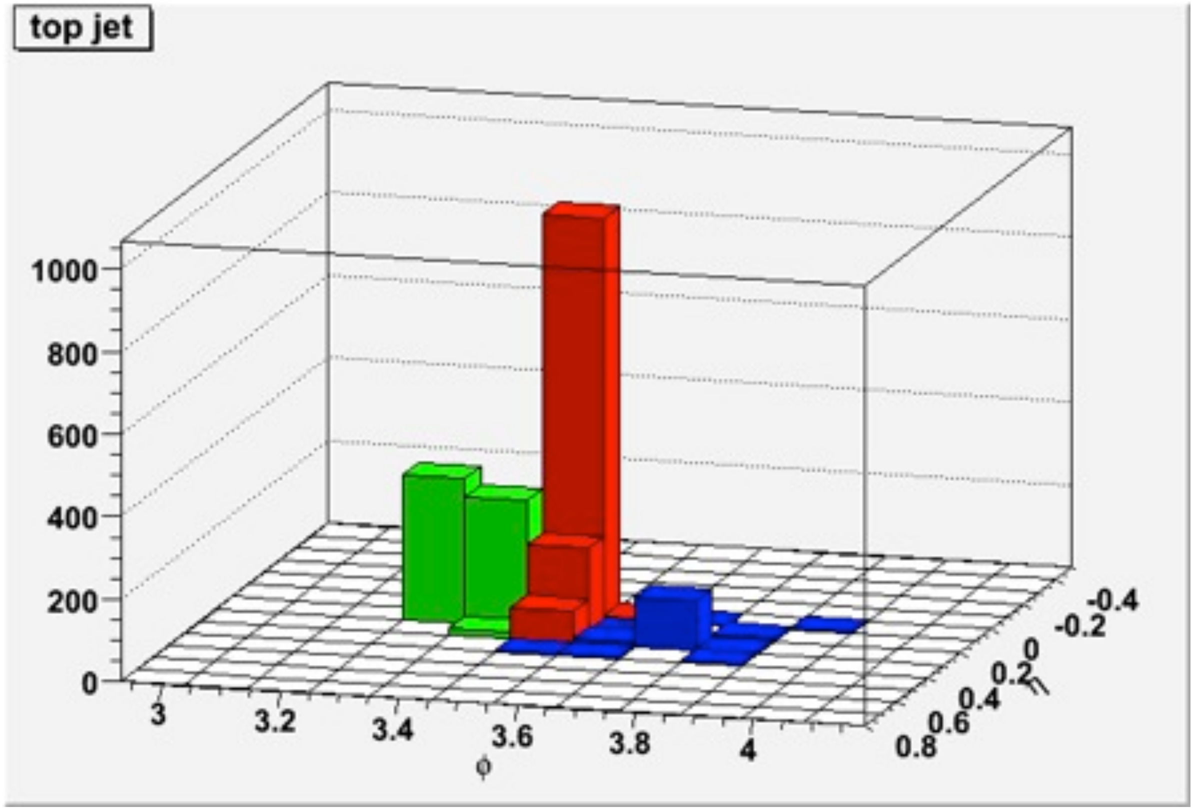




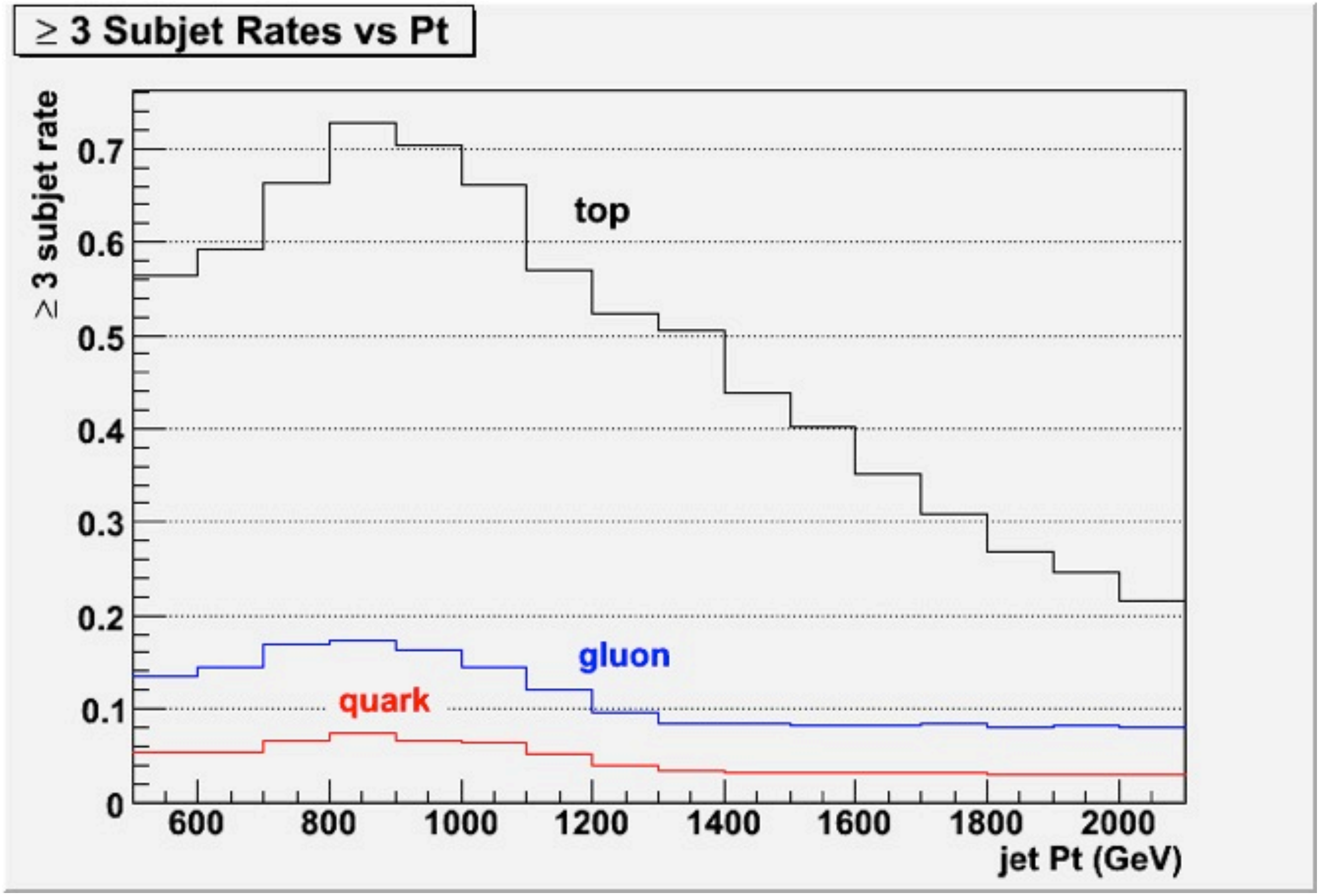


Some 2 TeV Top-Jets





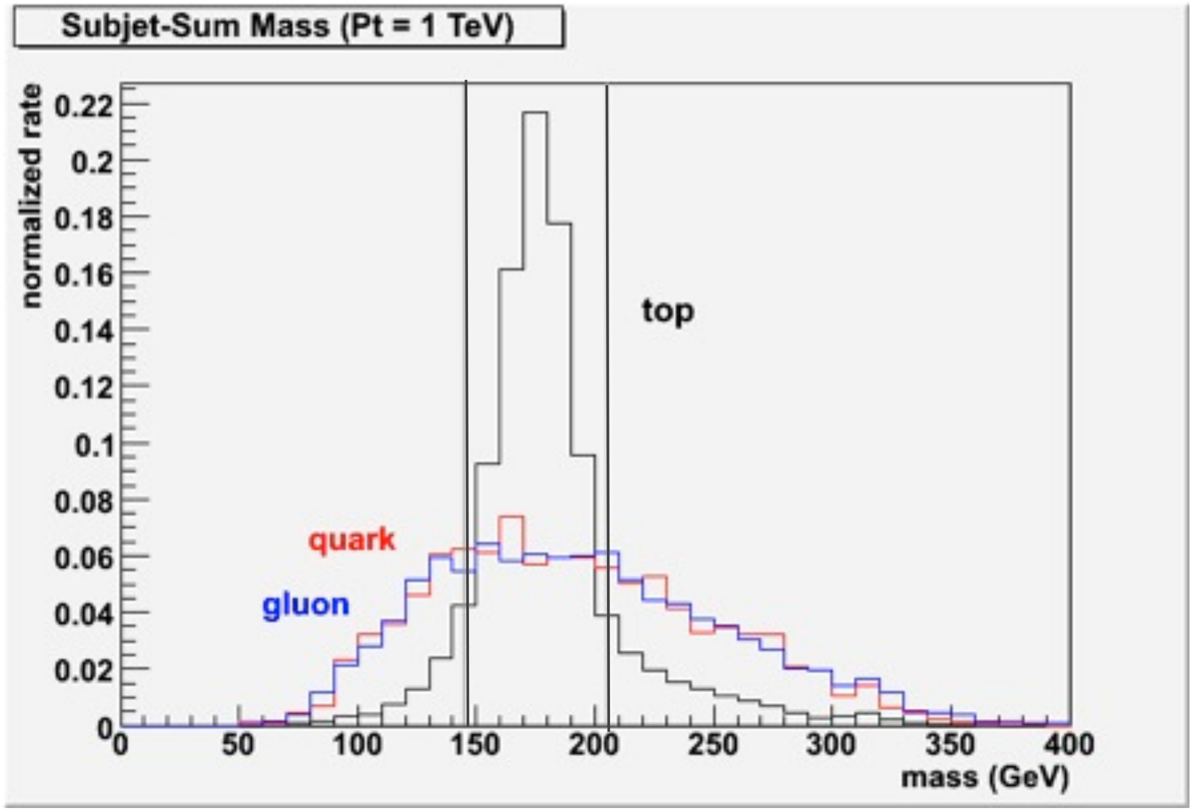
Subjet Rates

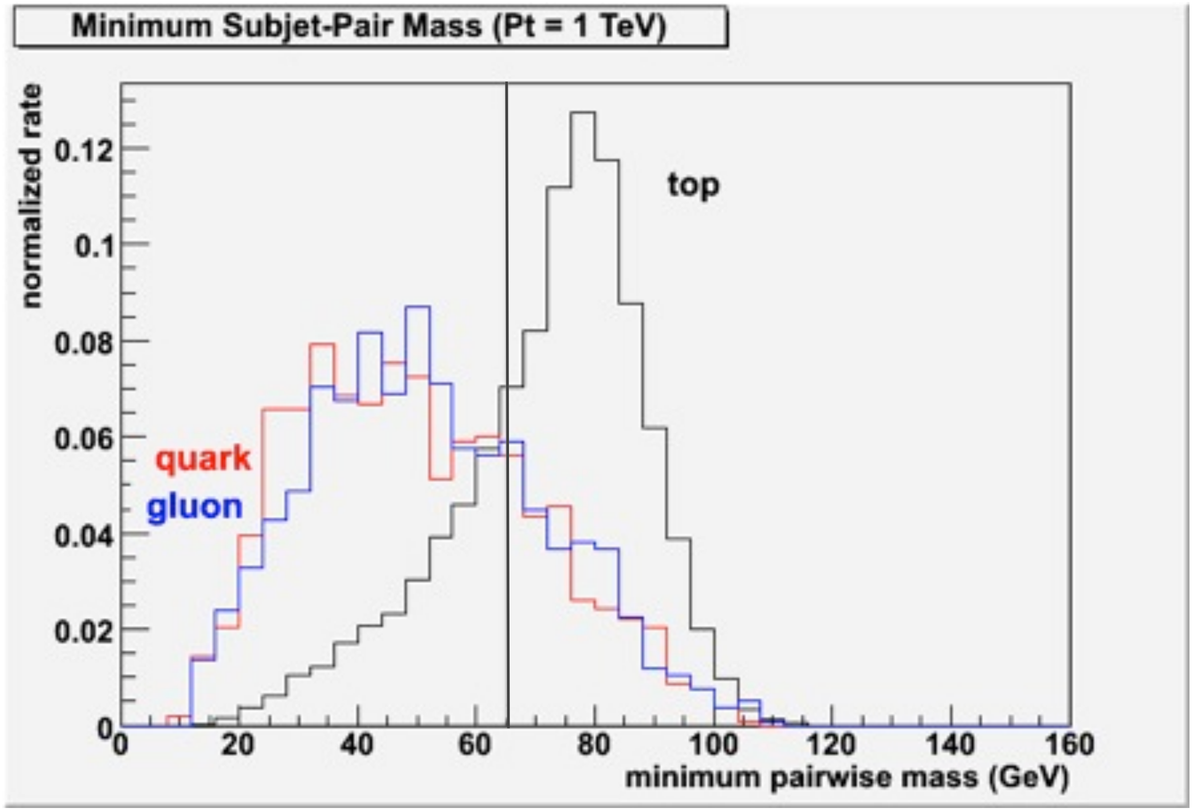


Kinematic Cuts

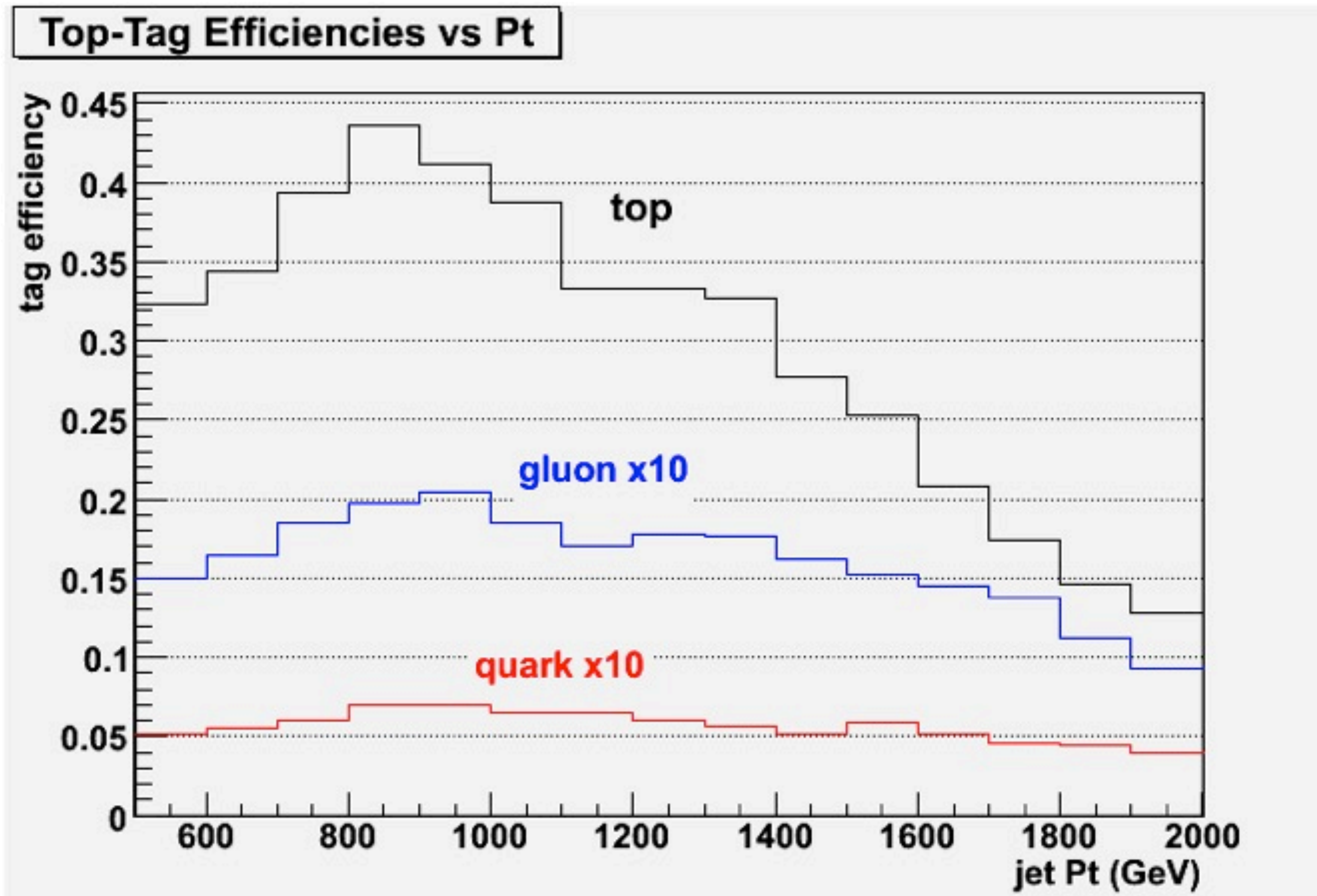
- Top mass
  - Pt < 1 TeV: $m_{1234} = [145, 205]$
  - Pt > 1 TeV: $m_{1234} = [145, Pt/20+155]$
- Minimum pairwise mass
  - $m_{min} > 65$ GeV

Kinematics

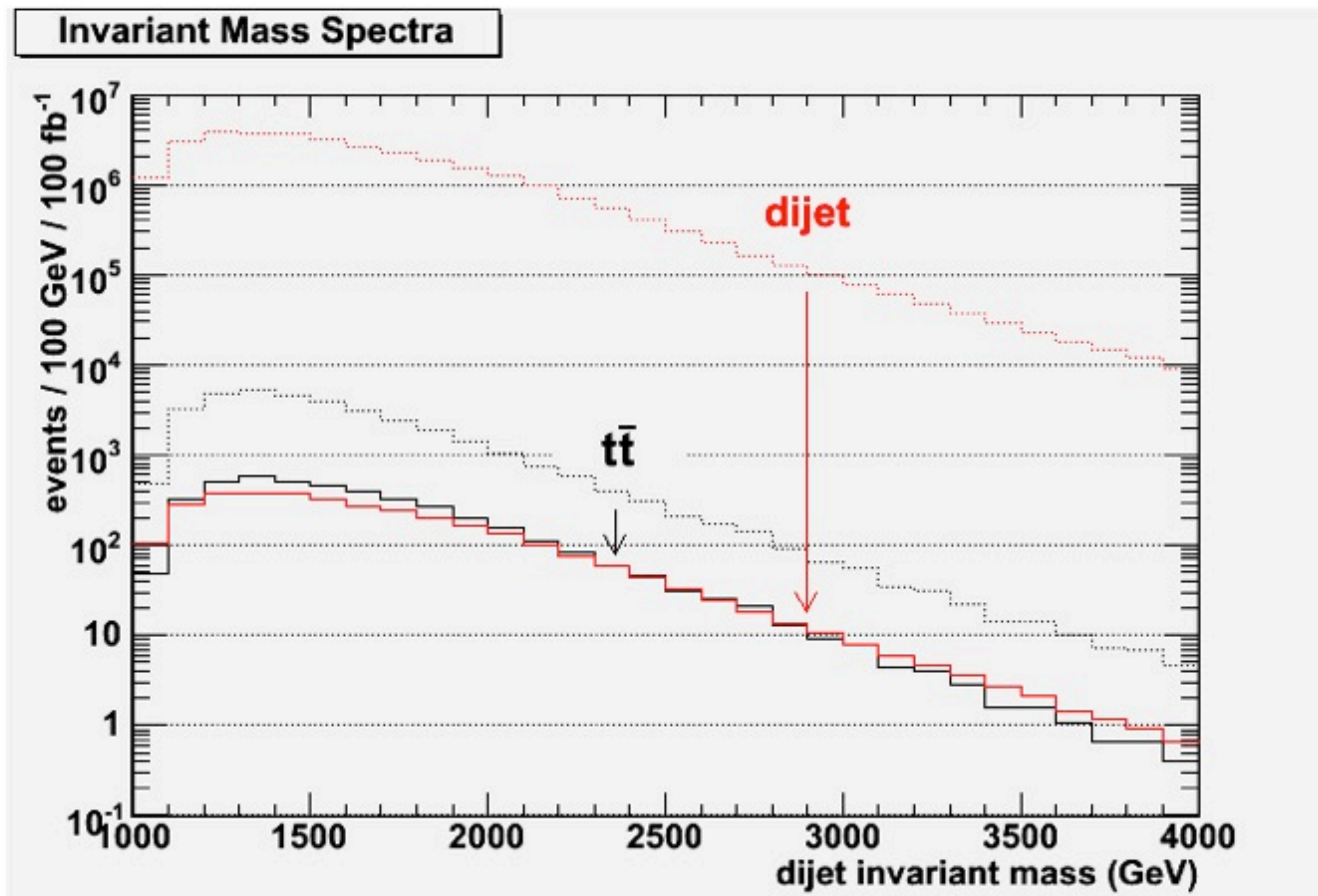

Subjet-Sum Mass (Pt = 1 TeV)
normalized rate
0.22
0.2
0.18
0.16
0.14
0.12
0.1
0.08
0.06
0.04
0.02
0
top
quark
gluon
0
50
100
150
200
250
300
350
400
mass (GeV)


Minimum Subjet-Pair Mass (Pt = 1 TeV)
normalized rate
0.12
0.1
0.08
0.06
0.04
0.02
0
top
quark
gluon
0
20
40
60
80
100
120
140
160
minimum pairwise mass (GeV)

Final Efficiencies



Final Dijet Mass Spectrum



Loose Ends

- Physics
  - PYTHIA = QCD?
- Technology
  - Ideal calorimeter = real calorimeter?

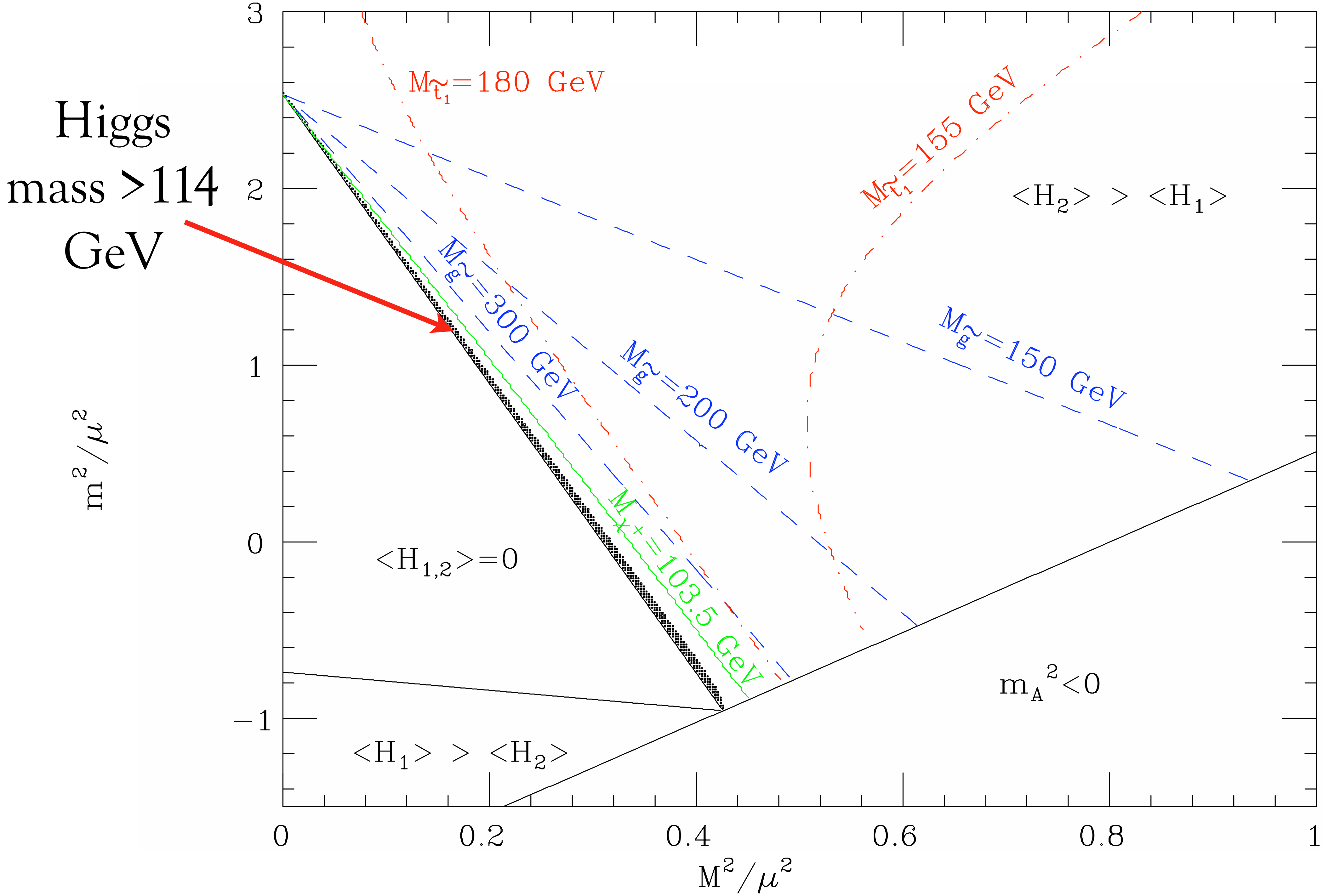
New Higgs Decay

$$h^0 \rightarrow a^0 a^0$$

And then the pseudo-scalars decay.

mSUGRA

Higgs mass >114 GeV



Giudice, Rattazzi '06

Corrections to the Higgs mass

$$\lambda v^2 = m^2_{phys} = 2\mu^2 + m^2_{soft}$$

$$\delta(m_h^2)_{phys} \propto y_t^2 m_t^2 \ln(m_{\tilde{t}}/m_t)$$ grows as a log

$$\delta(m_h^2)_{soft} \propto y_t^2 m_{\tilde{t}}^2 \ln(\Lambda/m_{\tilde{t}})$$ grows as a power

Typically need stop masses near 1 TeV

More Higgses

$$W = \lambda \hat{H}_u \hat{H}_d \hat{S} + \frac{\kappa}{3} \hat{S}^3$$

More Higgses

$$W = \lambda \hat{H}_u \hat{H}_d \hat{S} + \frac{\kappa}{3} \hat{S}^3$$

$$M_{H_d}^2 = -\frac{\lambda^2}{2}\left(s^2 + v^2 \sin^2\beta\right) + \frac{\lambda\kappa}{2} s^2 \tan\beta - \frac{M_Z^2}{2} \cos 2\beta + m_\lambda s \tan\beta \ ,$$

$$M_{H_u}^2 = -\frac{\lambda^2}{2}\left(s^2 + v^2 \cos^2\beta\right) + \frac{\lambda\kappa s^2}{2\tan\beta} + \frac{M_Z^2}{2} \cos 2\beta + \frac{m_\lambda s}{\tan\beta} \ ,$$

$$M_S^2 = -\frac{\lambda^2}{2} v^2 + \frac{\lambda\kappa}{2} v^2 \sin 2\beta - \kappa^2 s^2 + \frac{m_\lambda v^2}{2s} \sin 2\beta + m_\kappa s \ .$$

More Higgses

$$W = \lambda \hat{H}_u \hat{H}_d \hat{S} + \frac{\kappa}{3}\hat{S}^3$$

$$h_v^0, H_v^0, h_s^0 \qquad A_v^0, A_s^0$$

More Higgses

Mixing of pseudo-scalars thus allows:

$$h \to aa \to b\bar{b}b\bar{b}$$

Beyond SUSY

The pseudo-Goldstone boson can couple directly to gluons, without coupling to flavor.

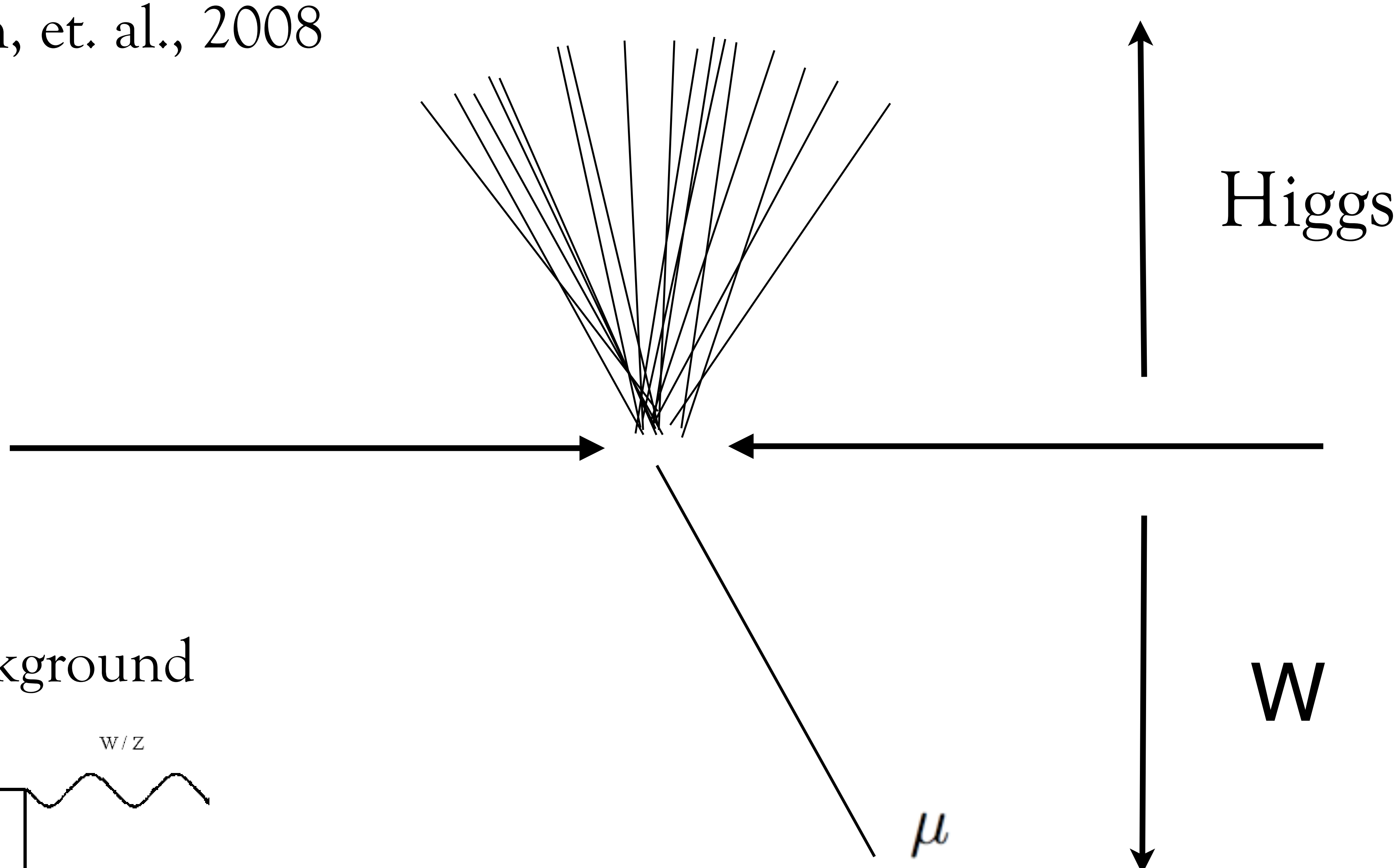
$$\mathcal{L} \sim \frac{v}{f^2} h (\partial_\mu a)^2 + \frac{a}{f} G^{\mu\nu} \tilde{G}_{\mu\nu}$$

Allowing:

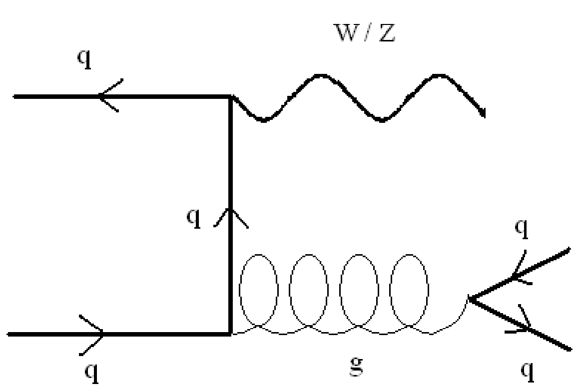
$$h \to aa \to gggg$$

Boosted Higgs

Butterworth, et. al., 2008



background



Boosted Higgs

Apply technique to: $h \to aa \to jjjj$

Modes:

(i) $Zh \to \ell^+\ell^- h$

(ii) $Wh \to \ell \not{E}_T h$

(iii) $Zh \to \not{E}_T h$

Preselection:

- Isolated Lepton ($p_T > 10$ GeV, $p_T^h < 0.2 p_T^l$ w/in $R < 0.4$)
- $\ell^+\ell^-$ seen, $m_{ll} = 91 \pm 10$ GeV and $p_T^{ll} > 200$ GeV **->** (i)
- ℓ seen, $p_T^{l E_T} > 200$ GeV and $m_T^{l E_T} < m_W + 10$ GeV -> (ii)
- $E_T > 200$ GeV -> (iii)

Preselection:

$process$	$Inclusive$	$W/Z\ p_T$	$Dilep$	$Lep + \not{E}_T$	$\not{E}_T$
Wh	0.87	0.047	< 0.001	0.011	< 0.001
Zh	0.74	0.038	0.0018	< 0.001	0.0063
$W + jets$	28600	180	< 0.1	30.8	1.8
$Z + jets$	9300	80.6	3.0	0.5	15.5
$t\bar{t}$	610	54.3	< 0.1	8.4	2.2

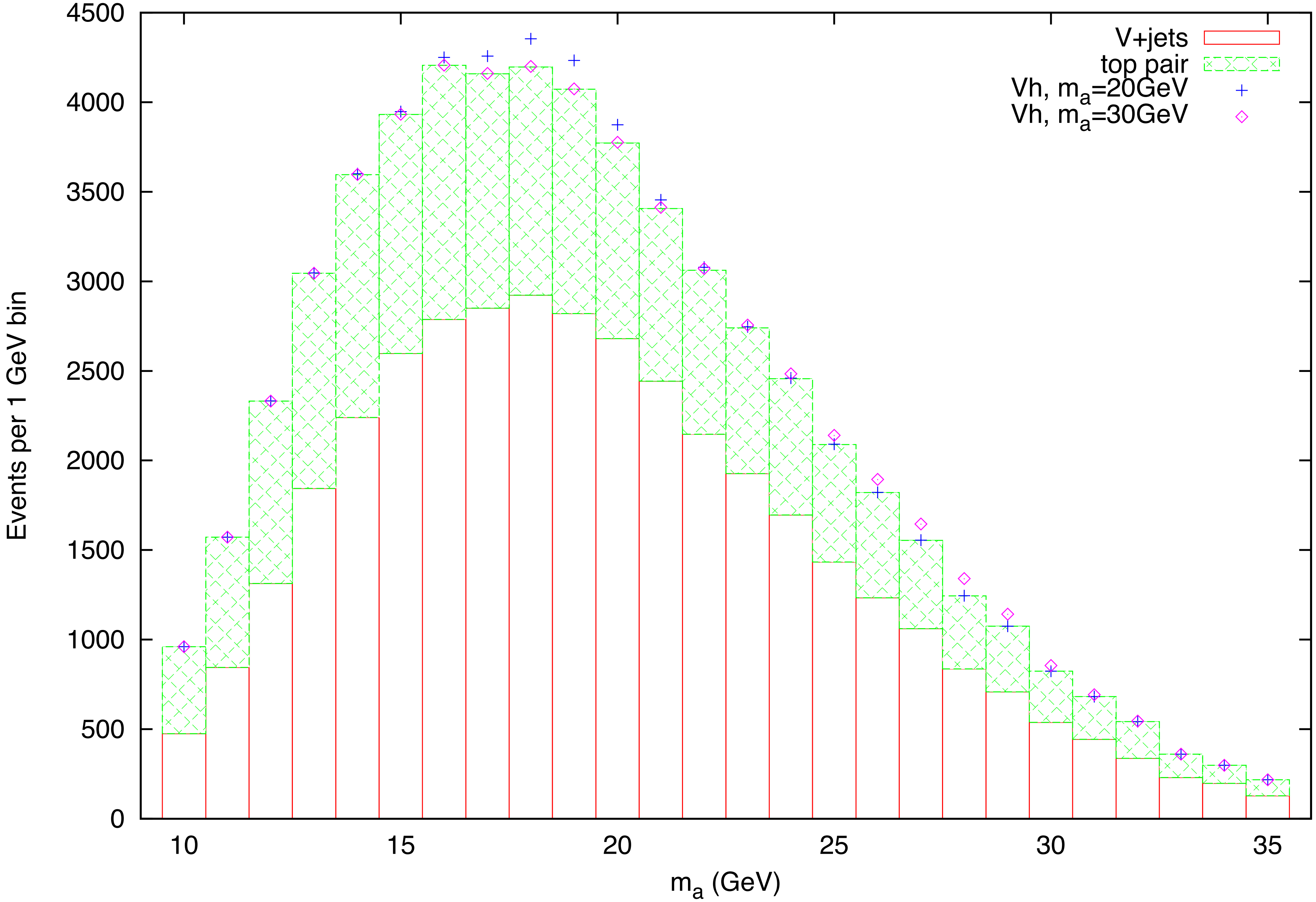
Cross sections in pb at LO (sorry!)

Light scalars

$m_a < 30$ GeV, $m_h = 120$ GeV:

- Require 2+ subjets.
- Combine more than two subjets in various combinations. Require mass of subjets equal to 25%.
- Require total jet mass between 100 and 125 GeV.

Light Scalar: Result



Light Scalar: Result

process	$m_h(i)$	$m_h(ii)$	$m_h(iii)$	$\Delta m_a(i)$	$\Delta m_a(ii)$	$\Delta m_a(iii)$
$V+j$	98.2	990.8	511.1	24.3	244.7	126.3
$t\bar{t}$	0.2	685.9	154.3	< 0.1	164.7	37.2
$4b(15)$	0.72	4.59	2.74	0.49	3.08	1.85
$4b(20)$	0.77	4.83	2.83	0.54	3.33	1.97
$4b(30)$	0.68	4.28	2.52	0.36	2.21	1.31
$4b(40)$	0.38	2.48	1.47	0.06	0.37	0.20
$4g(15)$	0.76	4.50	2.79	0.56	3.32	1.99
$4g(20)$	0.80	4.78	2.85	0.61	3.65	2.17
$4g(30)$	0.81	4.80	2.82	0.46	2.65	1.56
$4g(40)$	0.52	3.13	1.86	0.12	0.60	0.35

Cross sections in pb

Light Scalar: Result

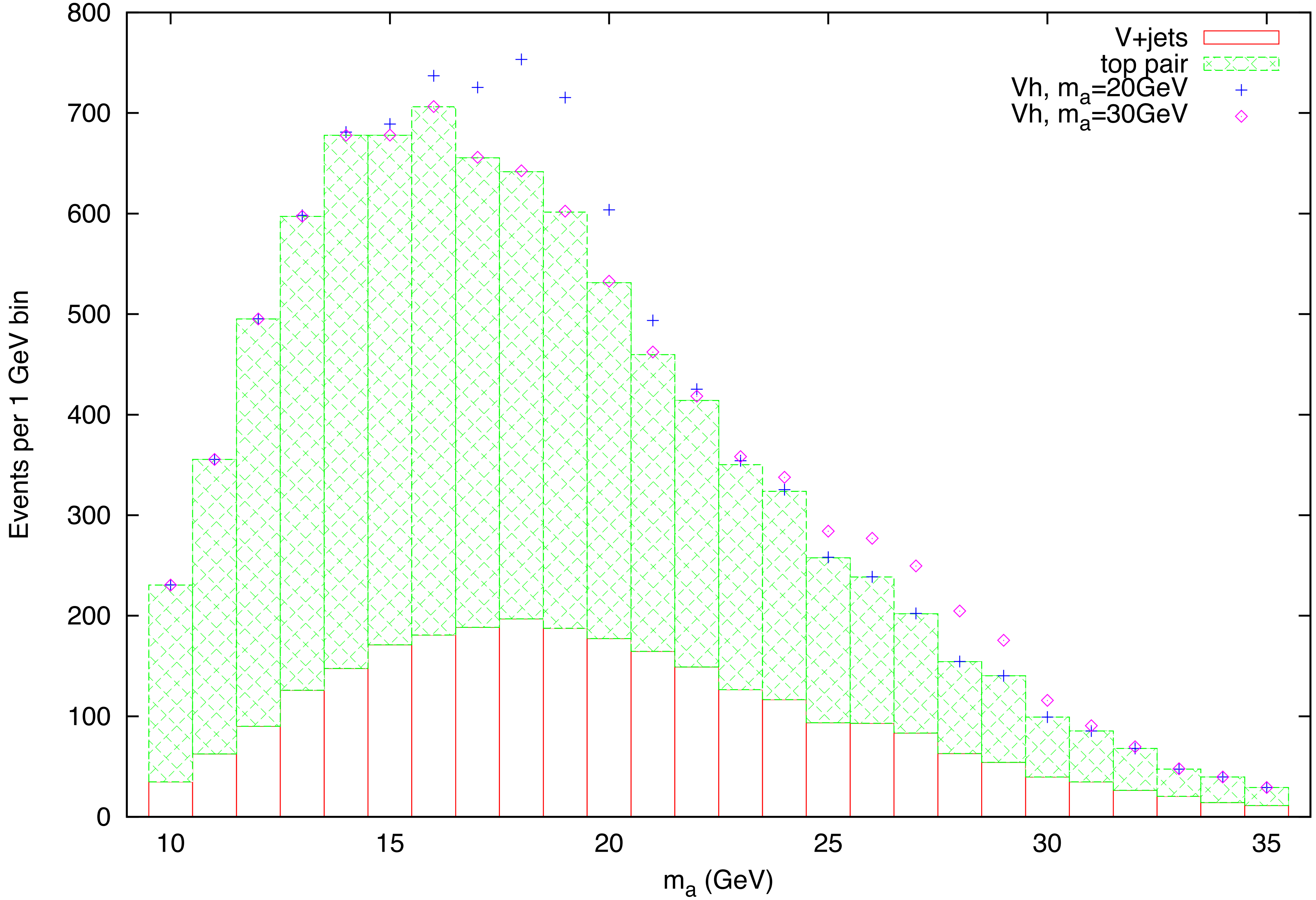
Process	Window	Signal	$V+j$	$t\bar{t}$	s/b	σ
$4b(15)$	$12-17$	479	10780	6320	0.028	3.7
$4b(20)$	$16-22$	536	16500	7310	0.022	3.5
$4b(30)$	$25-31$	317	5810	2800	0.037	3.4
$4b(40)$	$32-40$	34	1130	670	0.019	0.8
$4g(15)$	$12-17$	523	10780	6320	0.031	4.0
$4g(20)$	$16-22$	608	16500	7310	0.025	3.9
$4g(30)$	$25-31$	420	5810	2800	0.049	4.5
$4g(40)$	$32-40$	65	1130	670	0.036	1.5

Number of events and estimated significance with 100 fb^{-1} and 14 TeV.

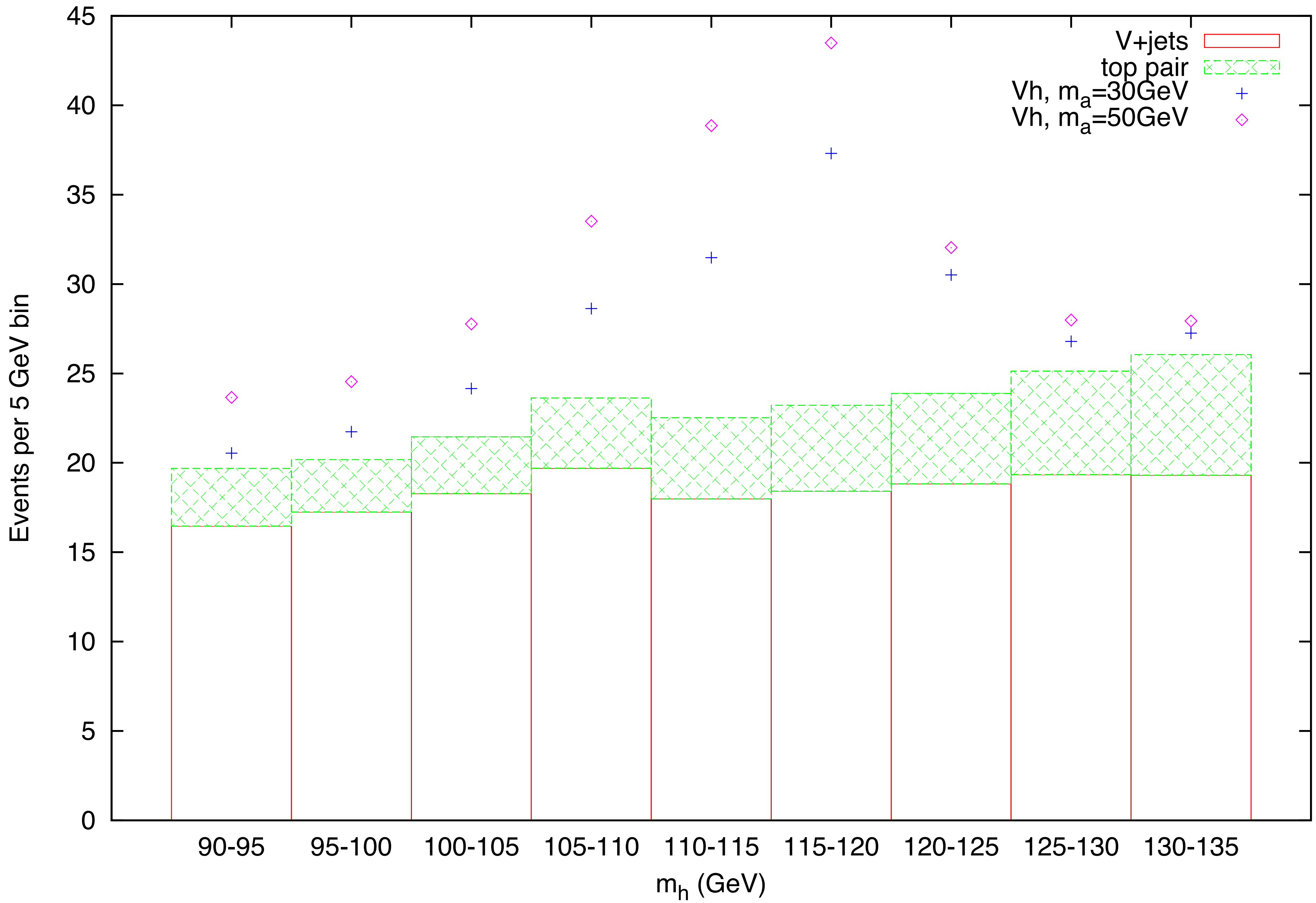
Light Scalar: with b-tags

m_a	s(1 tag)	s(2 tags)	b(1 tag)	b(2 tags)	σ(1 tag)	σ(2tags)
15	414	191	3150	122	7.4	17.3
20	433	167	3600	130	7.2	14.7
30	215	64	1090	39	6.5	10.3
40	21	7	230	11	1.4	2.2

Light Scalar: with b-tags



Heavy Scalars



Heavy Scalars

Cross sections (fb):

process	$m_h(i)$	$m_h(ii)$	$m_h(iii)$	Veto(ii)	Veto(iii)	3+
$V+j$	127.2	1282	658.5	725.3	365.2	138.6
$t\bar{t}$	0.3	833.3	188.3	92.6	11.4	7.86
Vh(15GeV)	0.72	4.47	2.68	3.21	1.92	0.05
Vh(20GeV)	0.76	4.70	2.77	3.40	1.97	0.09
Vh(30GeV)	0.71	4.39	2.59	3.13	1.82	1.14
Vh(40GeV)	0.68	4.35	2.54	3.15	1.80	1.81
Vh(50GeV)	0.70	4.50	2.66	3.26	1.90	1.78

b-tags with 100 fb^{-1}:

m_a(GeV)	*Signal*	s/b	σ
20	1.9	0.02	0.18
30	37.4	0.33	3.49
40	63.1	0.55	5.89
50	61.0	0.53	5.69

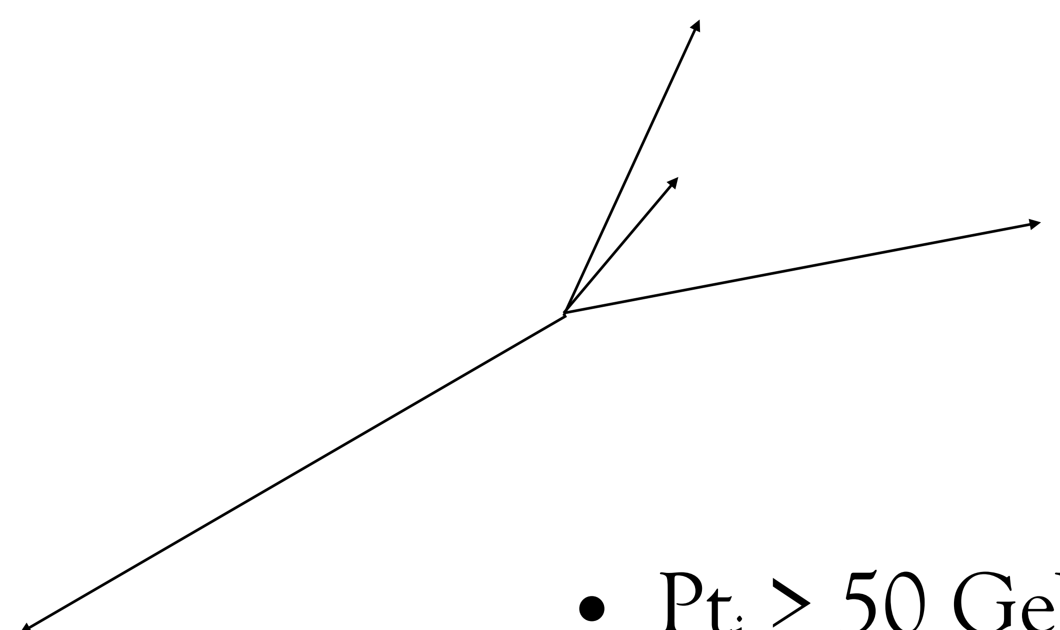
Conclusion

Heavy times require heavy methods!

Extras

MadGraph "Dijet"

"pp > jjjj"

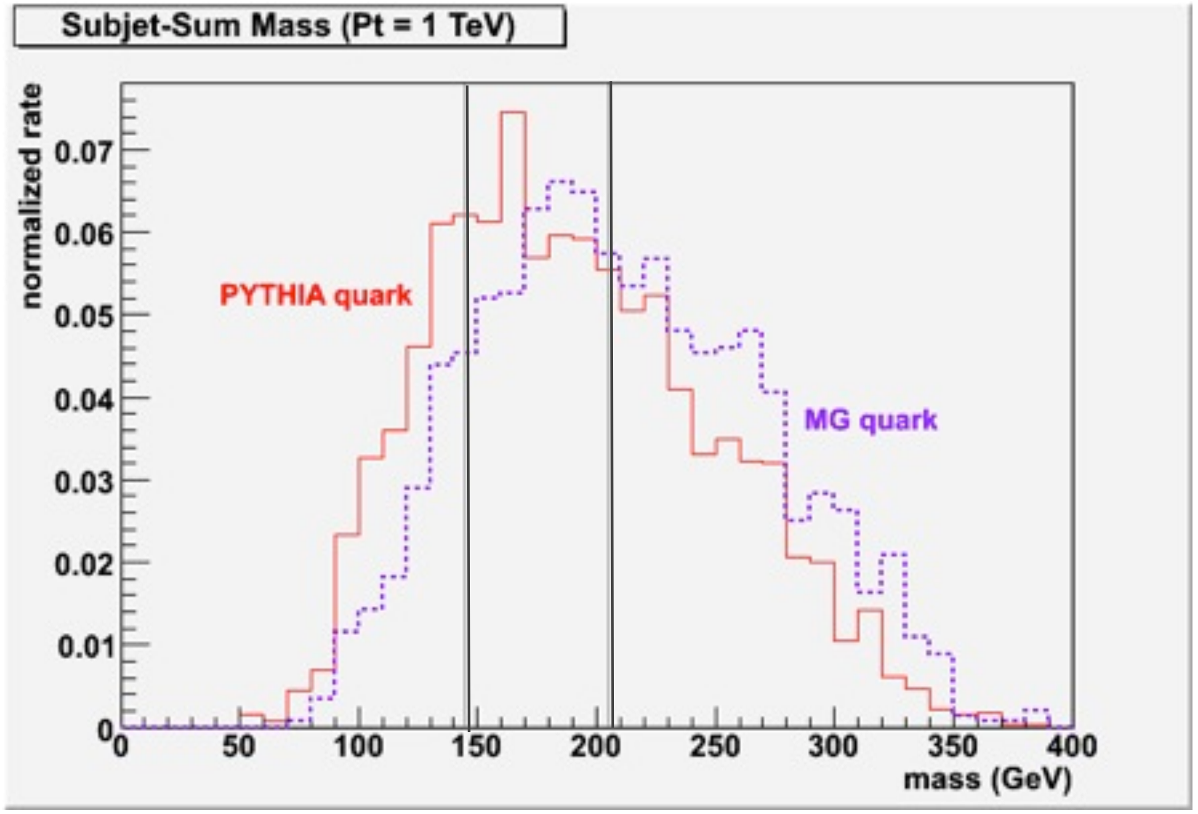


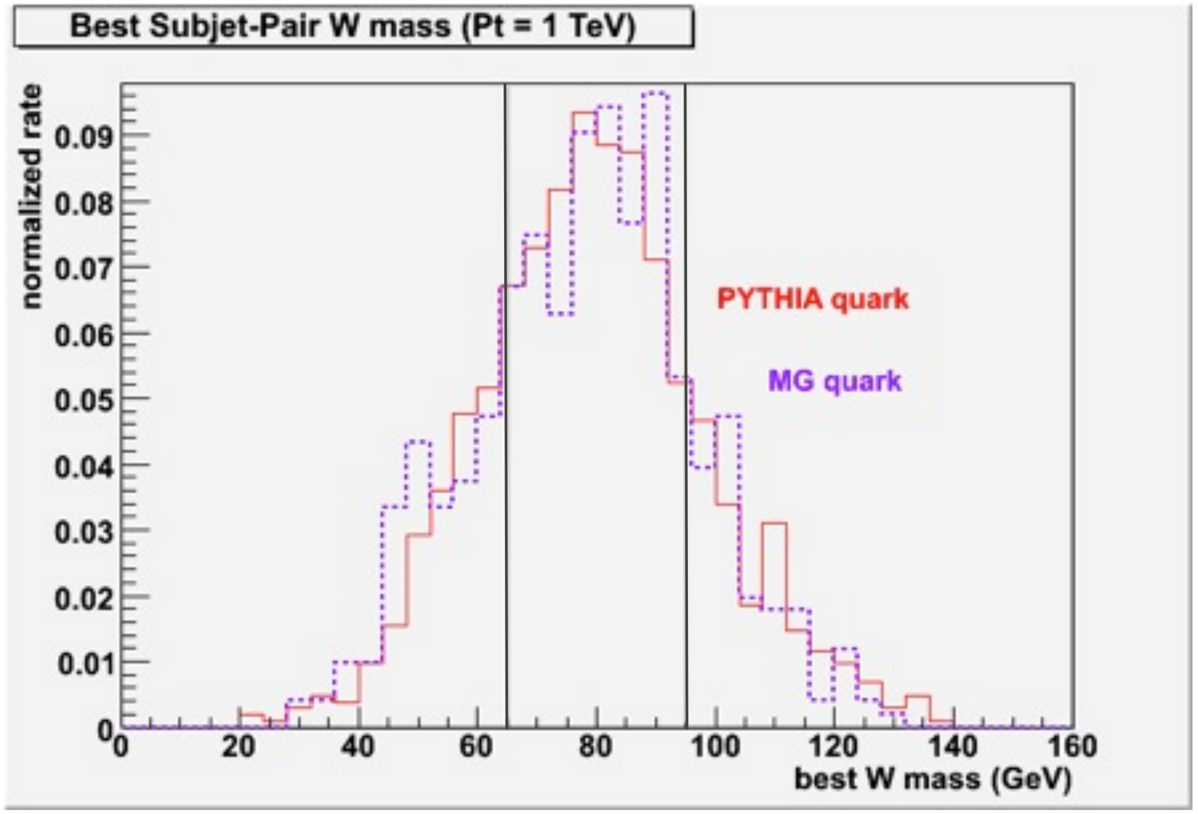
- Pt_j > 50 GeV
- ΔR_{jj} > 0.13
- lead jet Pt = [1, 1.2] TeV
- no parton shower, no hadronization
- α_s rescaled at softer vertices

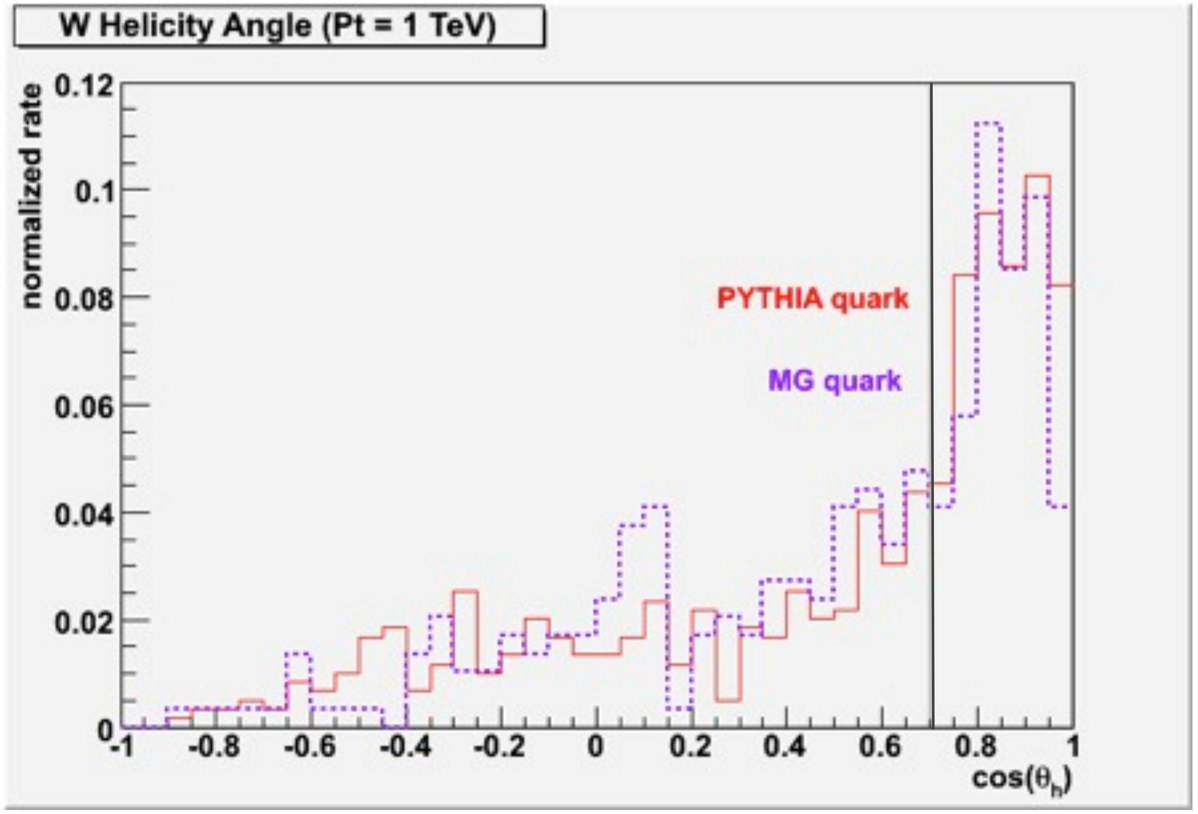
PYTHIA vs MadGraph at 1 TeV

Subjets

- PYTHIA quark: 7%
- MG quark: 9%

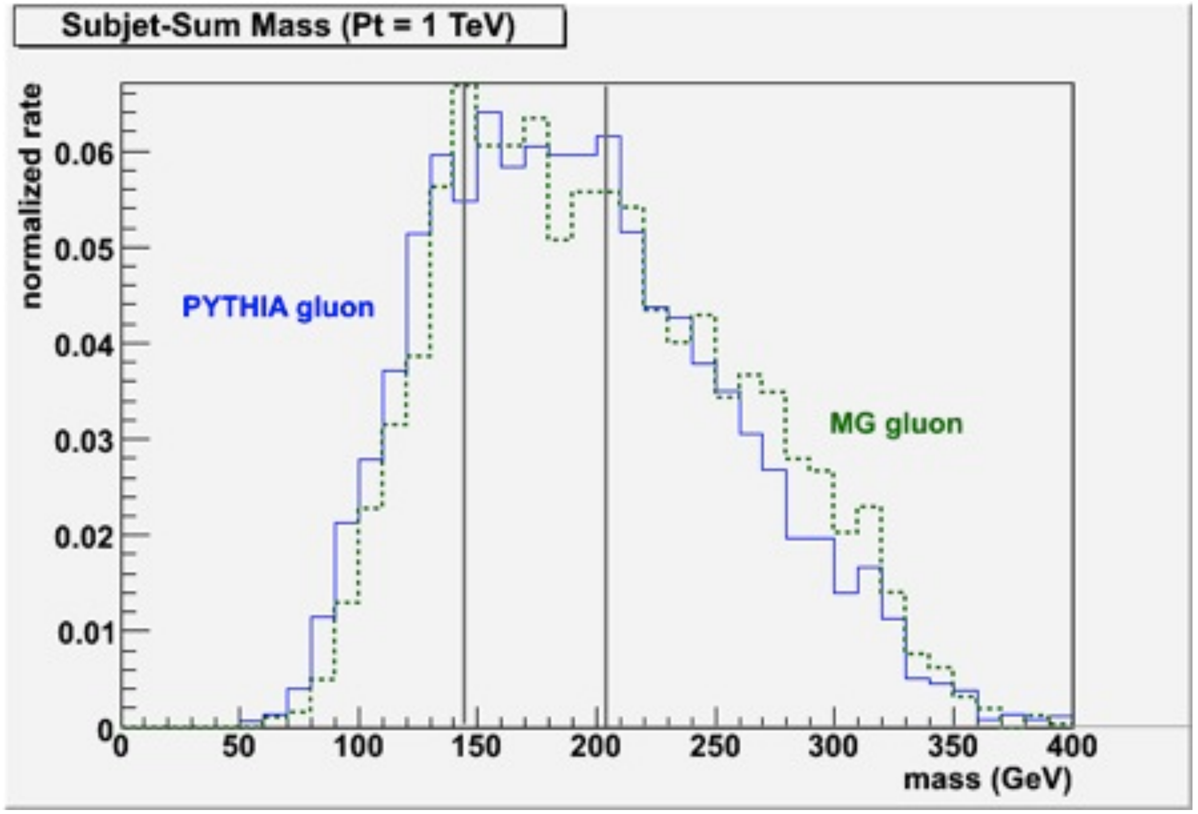


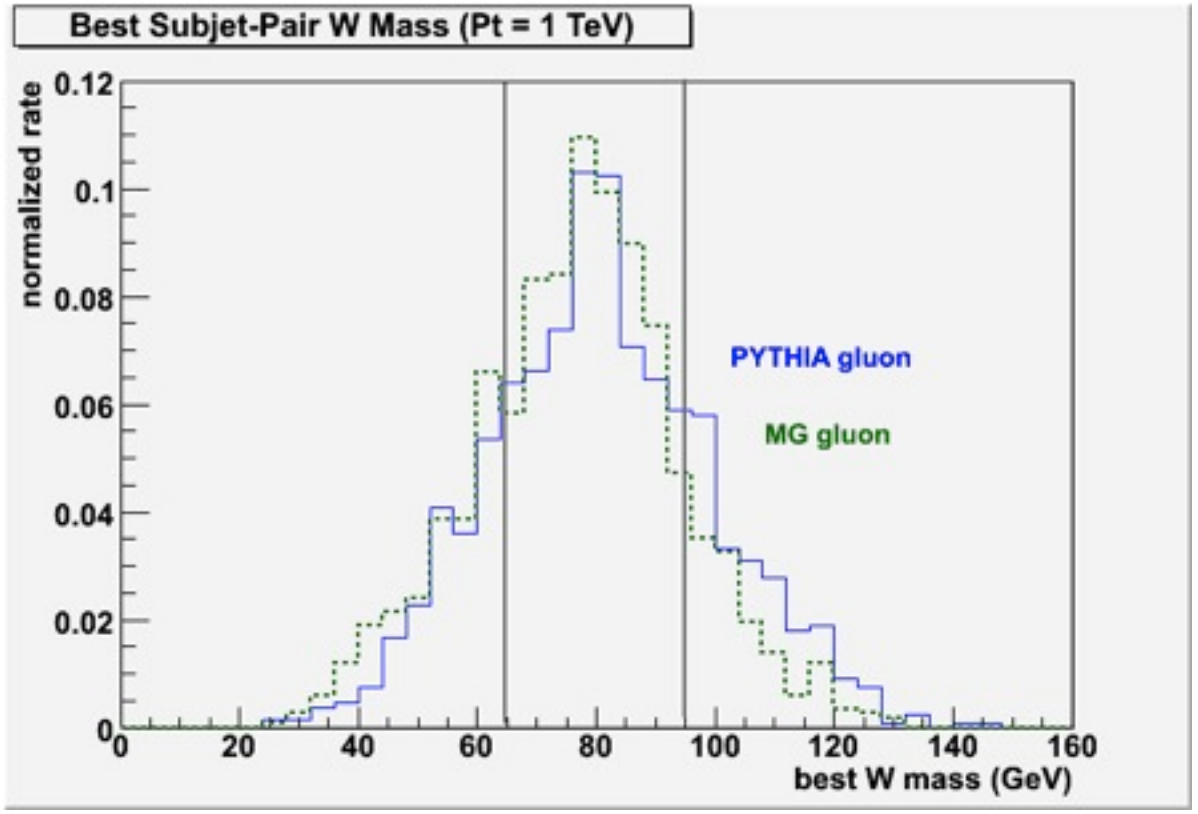


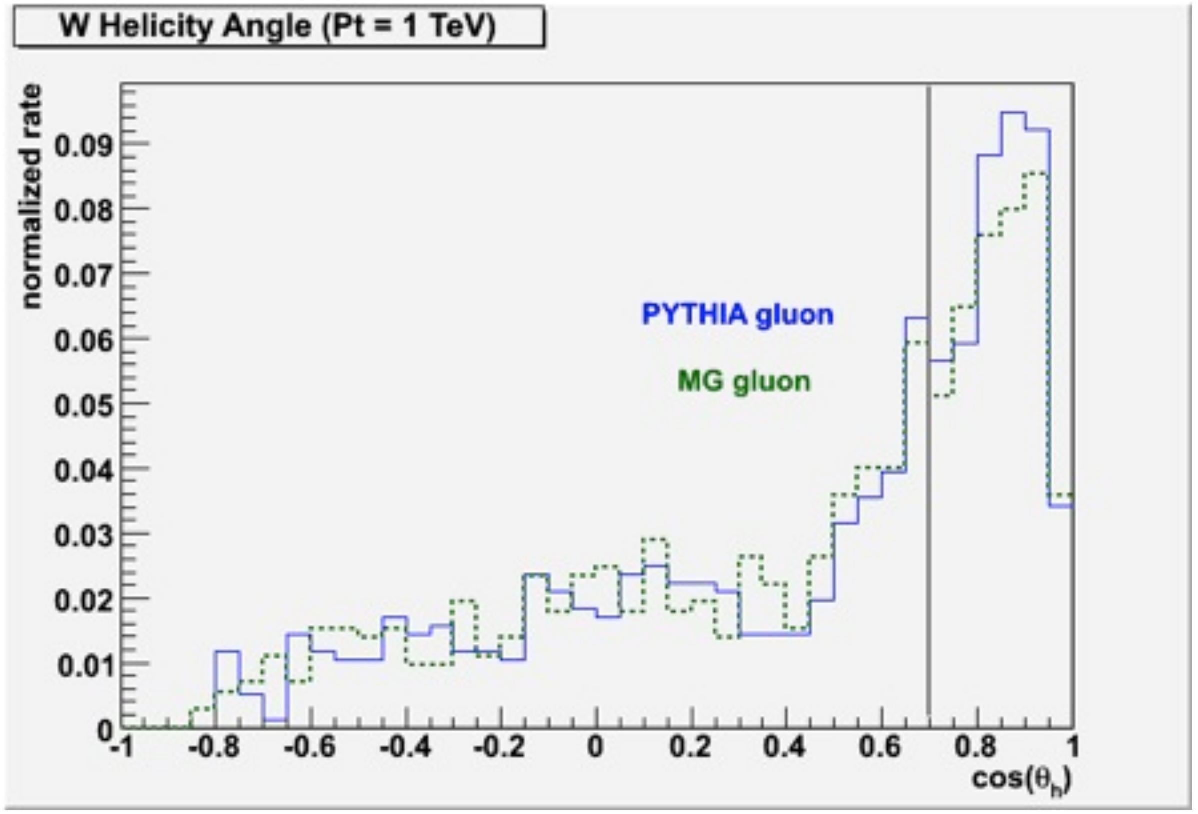


PYTHIA vs MadGraph at 1 TeV

- PYTHIA gluon: 16%
- MG gluon: 22%







... vs HERWIG at 1 TeV

- Same settings as PYTHIA
- Extremely similar story
- Final rates ~50% bigger than PYTHIA

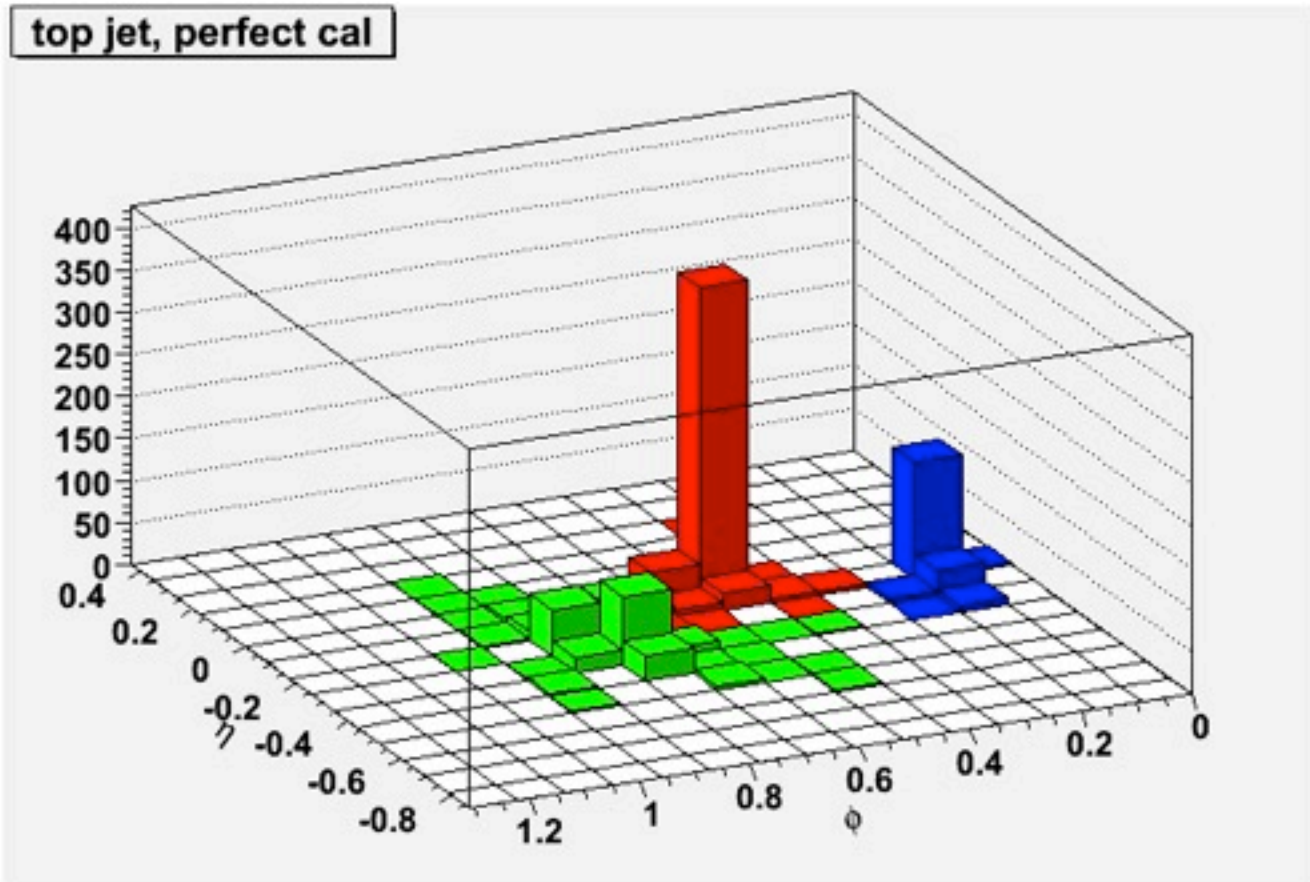
Physics Summary

- PYTHIA vs MG vs HERWIG
  - Rates match at ~50% level
    - Factor of ~2 uncertainty in dijet BG estimate
  - Kinematic distributions extremely similar
- In any event, can probably measure QCD distributions in-situ with sidebands of jet mass

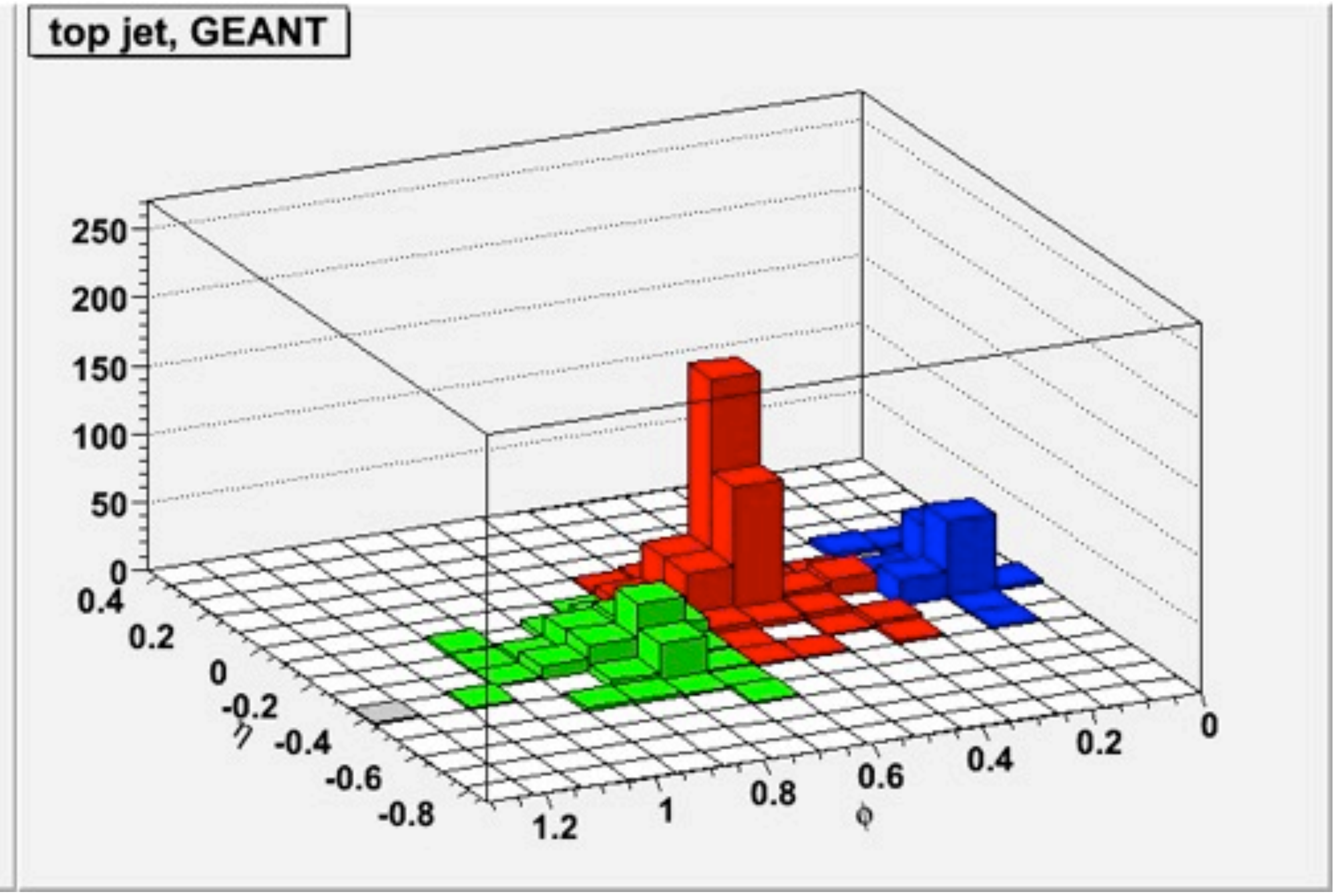
CMS Collaboration

- Sal Rappoccio, Morris Swartz, Petar Maksimovic
- Working on implementation of algorithm in CMS framework
- Proof of concept: 2 TeV *Z*', full detector simulation
  - PYTHIA-based physics
  - Decays to light quarks and tops
  - Pt = 0.5~1 TeV

Us vs Sal

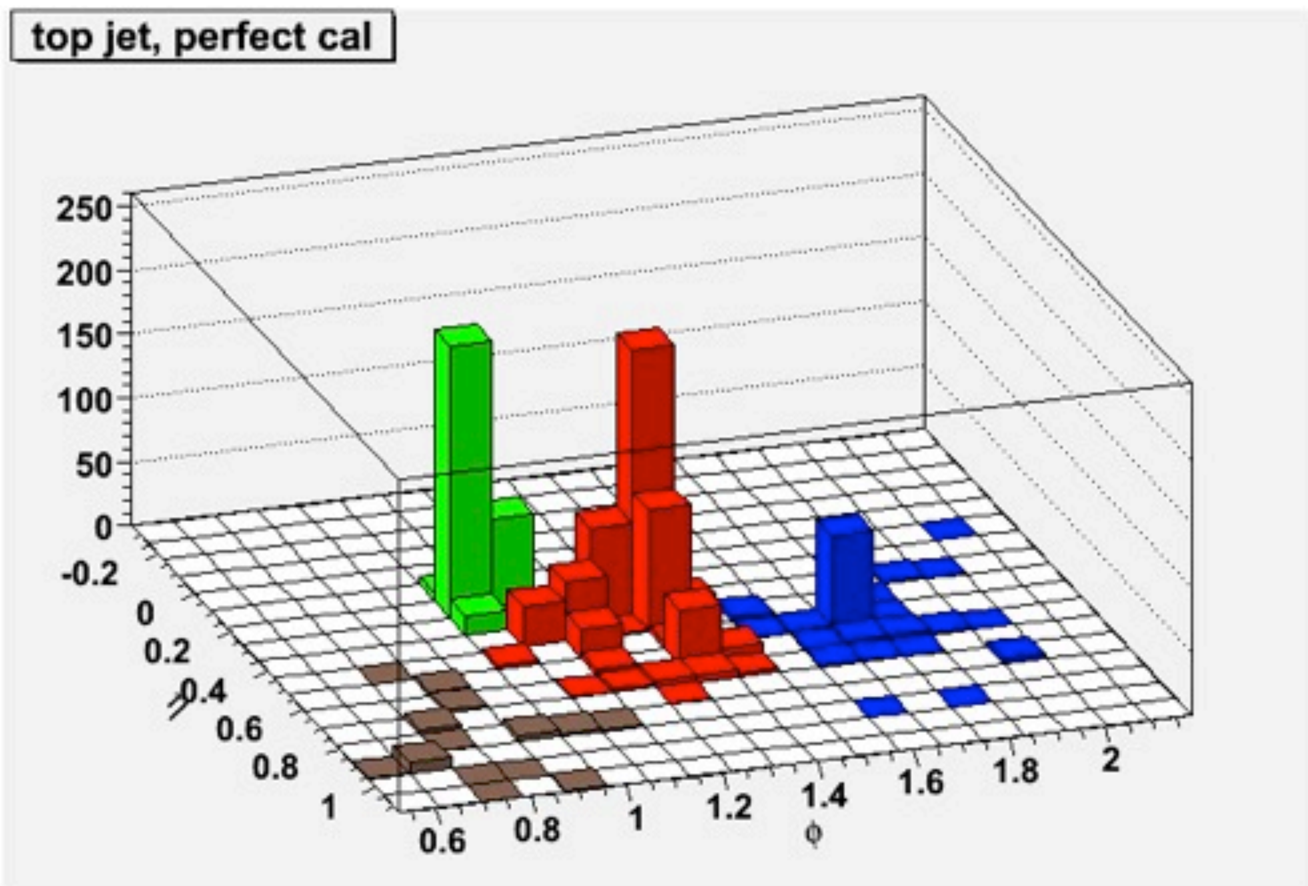


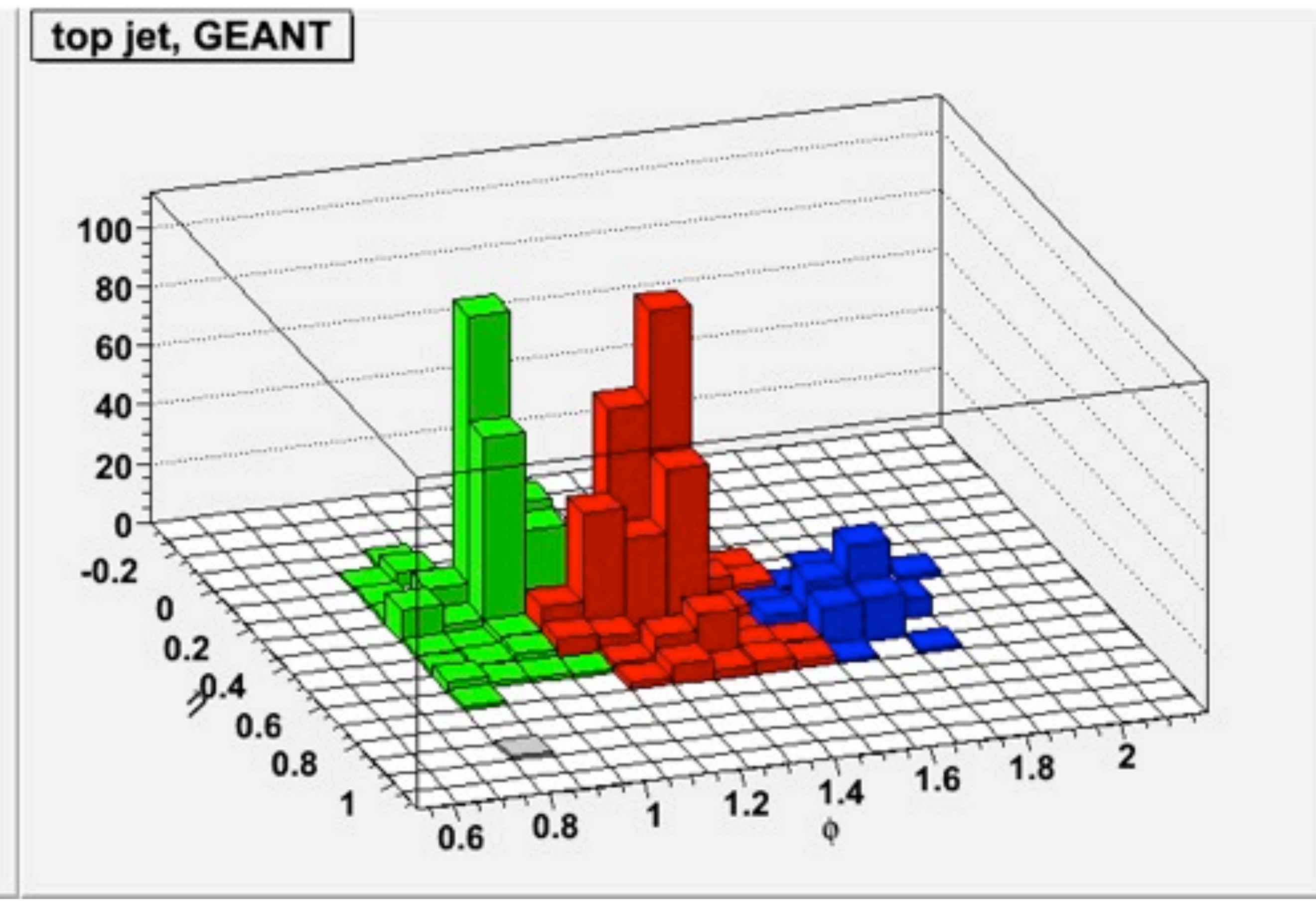
top jet, perfect cal
400
350
300
250
200
150
100
50
0
0.4
0.2
0
-0.2
η
-0.4
-0.6
-0.8
1.2
1
0.8
0.6
0.4
0.2
0
φ



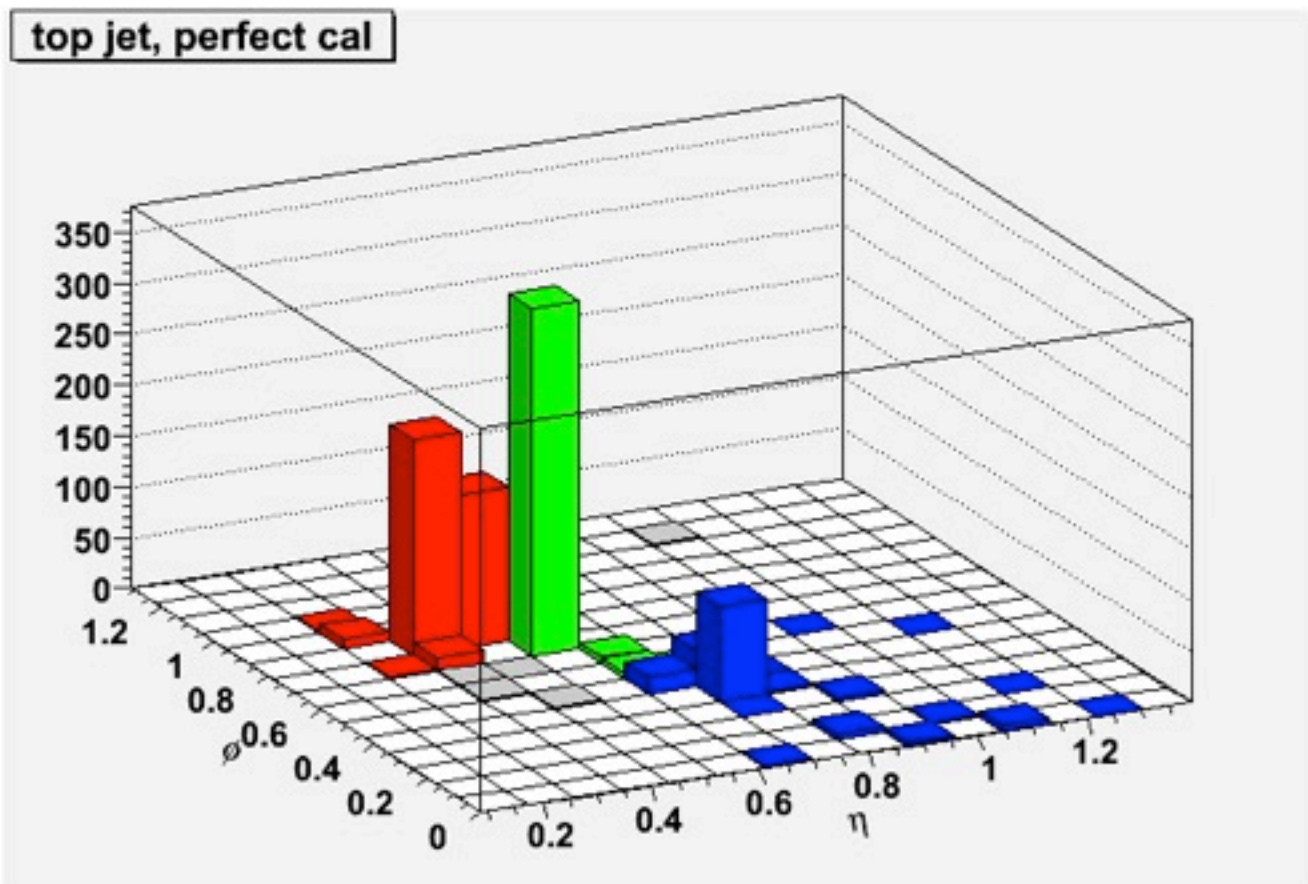
top jet, GEANT
250
200
150
100
50
0
0.4
0.2
0
-0.2
η
-0.4
-0.6
-0.8
1.2
1
0.8
0.6
0.4
0.2
0
φ

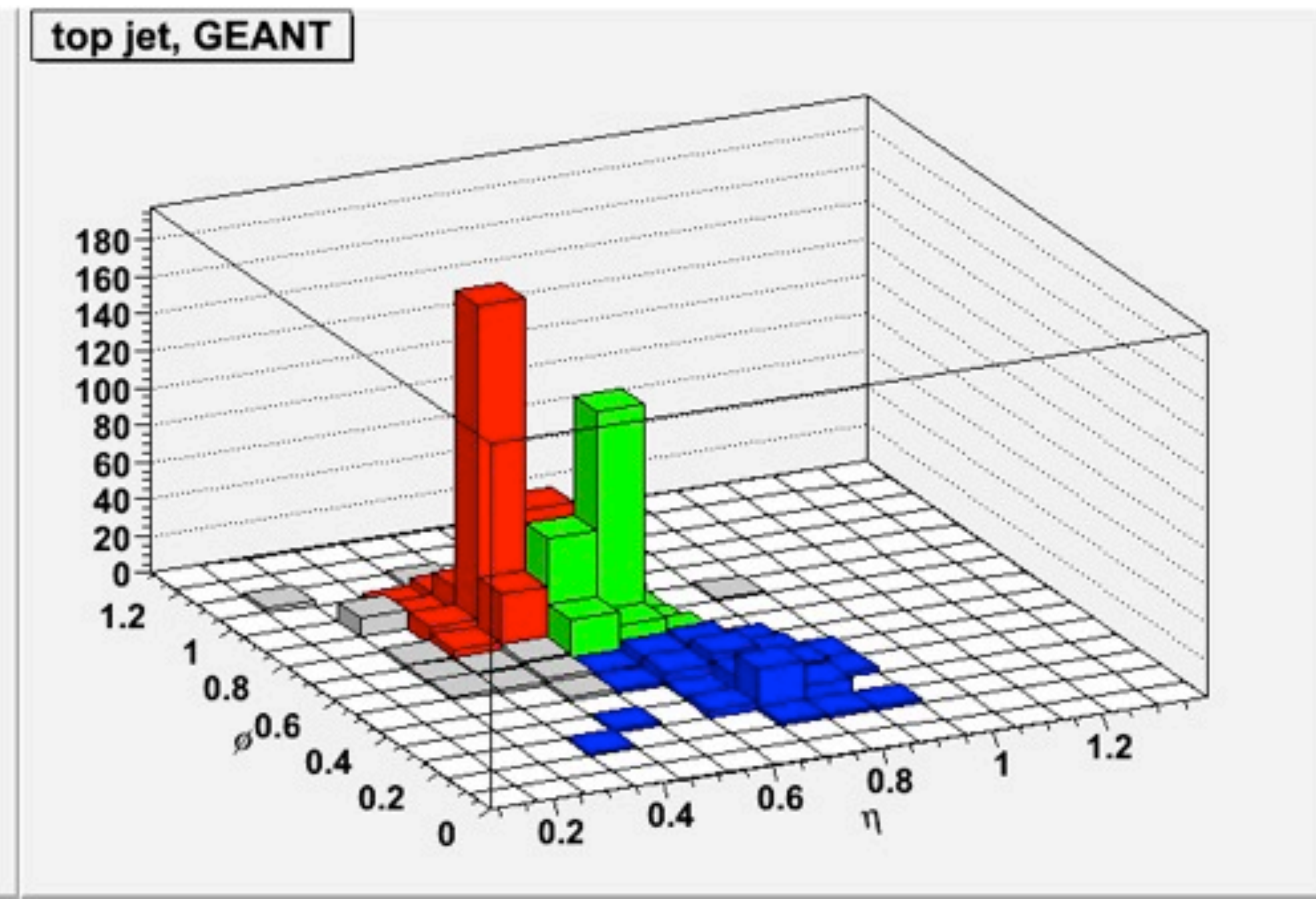
Us vs Sal


top jet, perfect cal
250
200
150
100
50
0
-0.2
0
0.2
0.4
0.6
0.8
1
0.6
0.8
1
1.2
1.4
1.6
1.8
2
ϕ

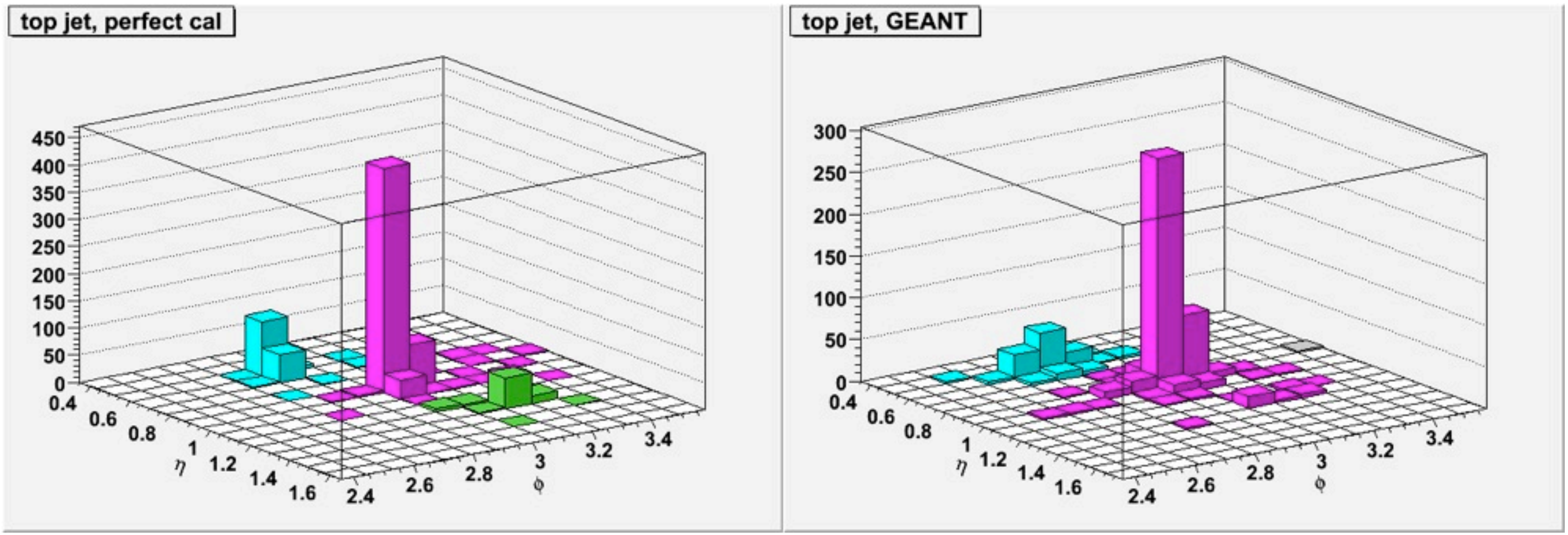

top jet, GEANT
100
80
60
40
20
0
-0.2
0
0.2
0.4
0.6
0.8
1
0.6
0.8
1
1.2
1.4
1.6
1.8
2
ϕ

Us vs Sal


top jet, perfect cal
350
300
250
200
150
100
50
0
1.2
1
0.8
ϕ0.6
0.4
0.2
0
0.2
0.4
0.6
0.8
1
1.2
η


top jet, GEANT
180
160
140
120
100
80
60
40
20
0
1.2
1
0.8
ϕ0.6
0.4
0.2
0
0.2
0.4
0.6
0.8
1
1.2
η

Us vs Sal



top jet, perfect cal
450
400
350
300
250
200
150
100
50
0
0.4
0.6
0.8
1
1.2
1.4
1.6
η
2.4
2.6
2.8
3
3.2
3.4
φ
top jet, GEANT
300
250
200
150
100
50
0
0.4
0.6
0.8
1
1.2
1.4
1.6
η
2.4
2.6
2.8
3
3.2
3.4
φ

Technology Summary

- Bottomline: it still works
- Final efficiencies for t / q (2 TeV Z')
  - Us: 36% / 0.7%
  - Sal: 32% / 1.0%
- Most S/B degradation attributable to energy resolution
- Higher stats / masses in the pipeline
  - How fast do efficiencies fall off?

Future Directions

- ECAL
  - Captures ~10% of jet energy
  - 5x better spatial resolution
- Tracker
  - Sees all charged particles
  - Even better resolution
  - Crowded for individual track ID, but maybe not for tracing Et flow