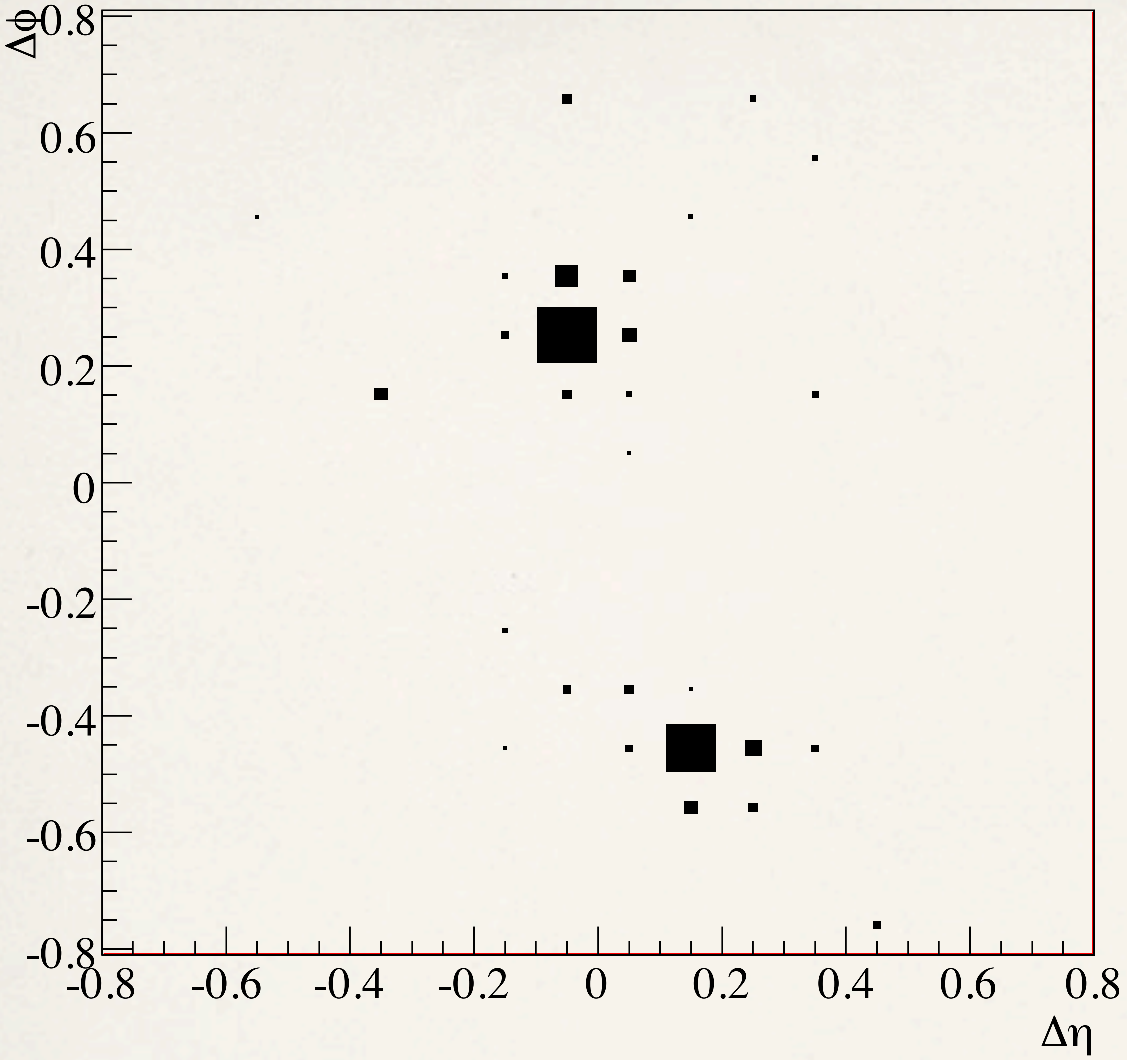
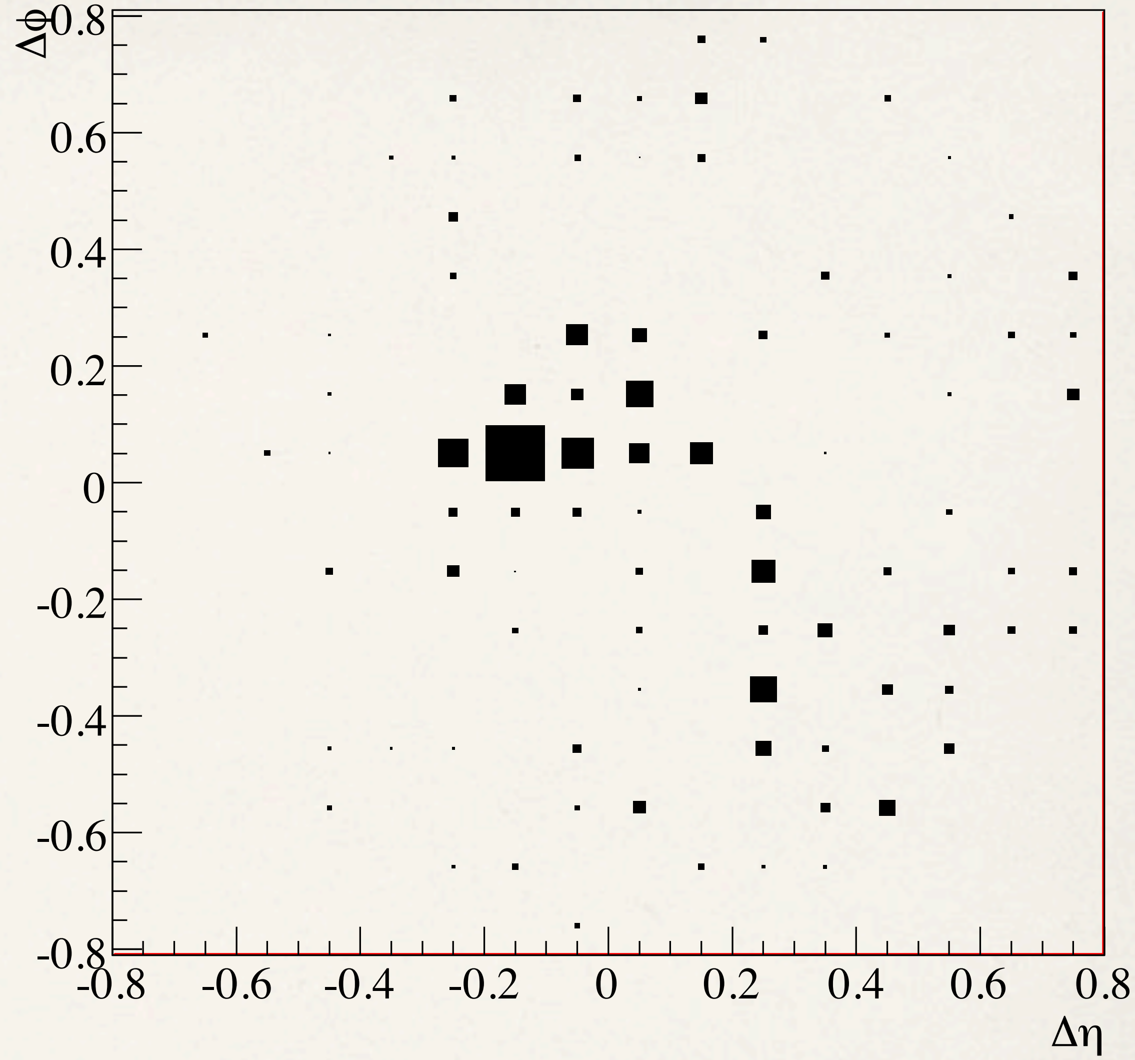
Boosted Heavy Particle



QCD Jet



Using Jet Substructure to Find New Physics at the LHC

David Krohn (Harvard)

Cornell Journal Club 4/8/11

Upcoming Jet Conference

- Boost 2011, Princeton [5/22-5/26]

Goal: Study jets from boosted heavy objects, as well as exotic jets (lepton jets, etc.).

http://boost2011.org

Outline

- Introduction to Jets and Jet Algorithms
- Boosted Heavy Objects and their Identification
  - Examples
- Superstructure
- Conclusions

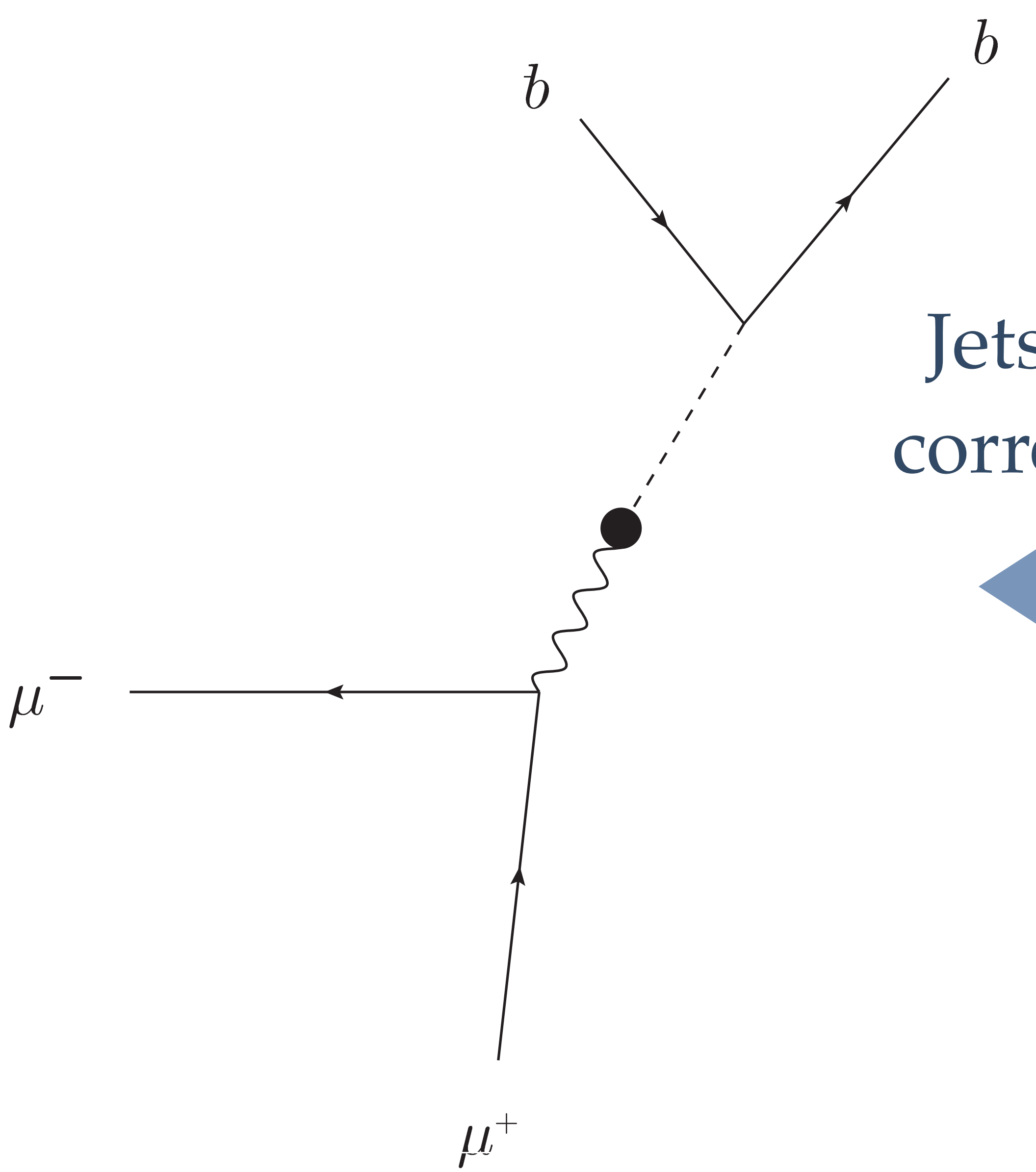
Takeaway

- At the LHC, even particles we normally regard as heavy (t/W/Z/h) can be so energetic their decay products are resolved in a single jet.
- This might seem like a problem, but in fact these jets look very different from the jets of ordinary QCD!
- By looking at the radiation distribution inside a jet we can try to recover some of the original heavy particle's properties (identity, polarization, color structure, etc).
- Remarkably, not only can we recover information on the jet's properties, but in some cases by looking in the boosted region we can actually do better than conventional analyses.

Introduction to Jets and Jet Algorithms

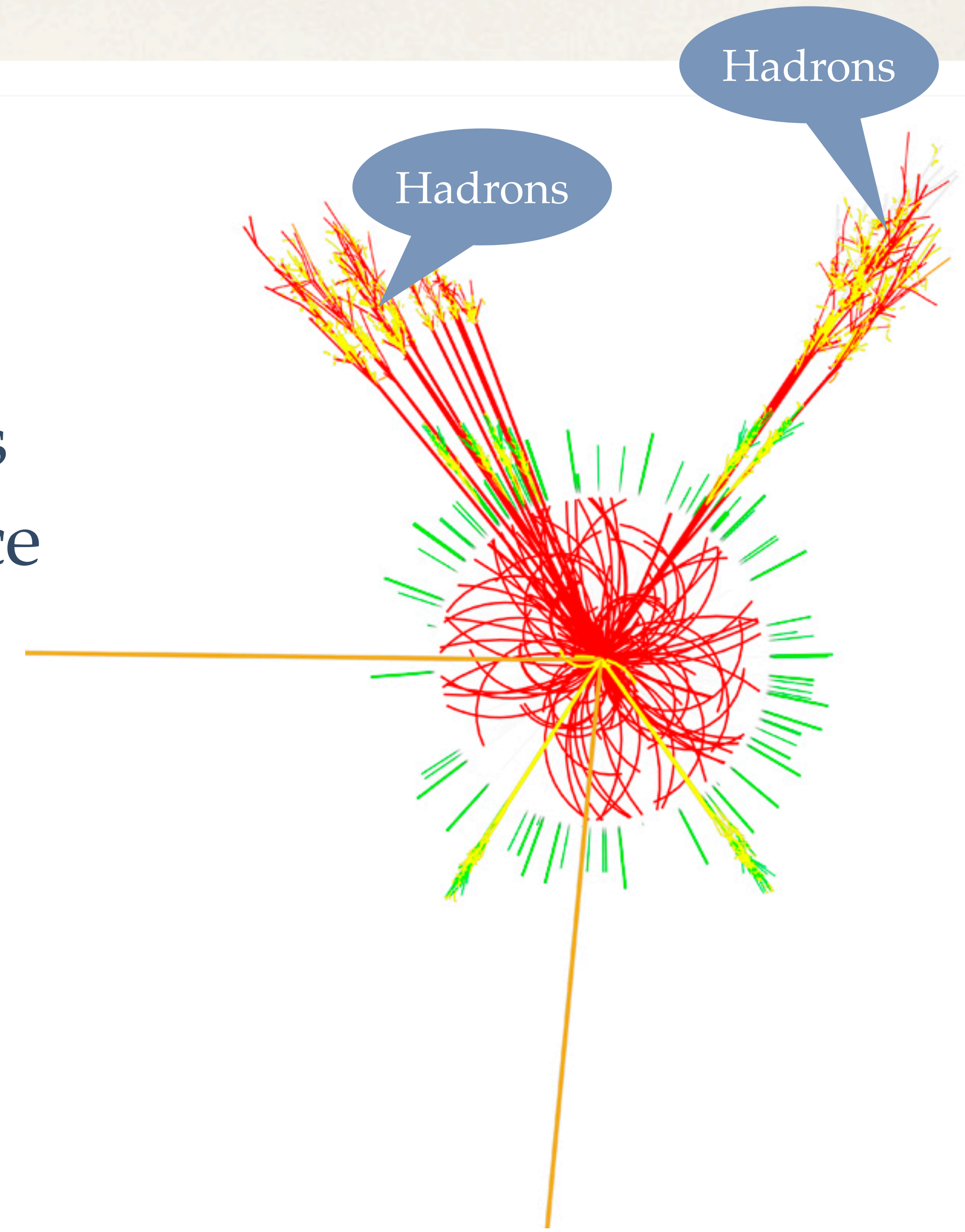
What is a Jet?

- A jet is a kinematical object we construct from collider data.
  - Specifically, jets are collections of hadronic four-vectors used to approximate the kinematics of the hard scattering in a collider event.
- They help us map things we cannot easily calculate (the exact energy distribution in the calorimeter) to things we can (perturbative Feynman amplitudes)



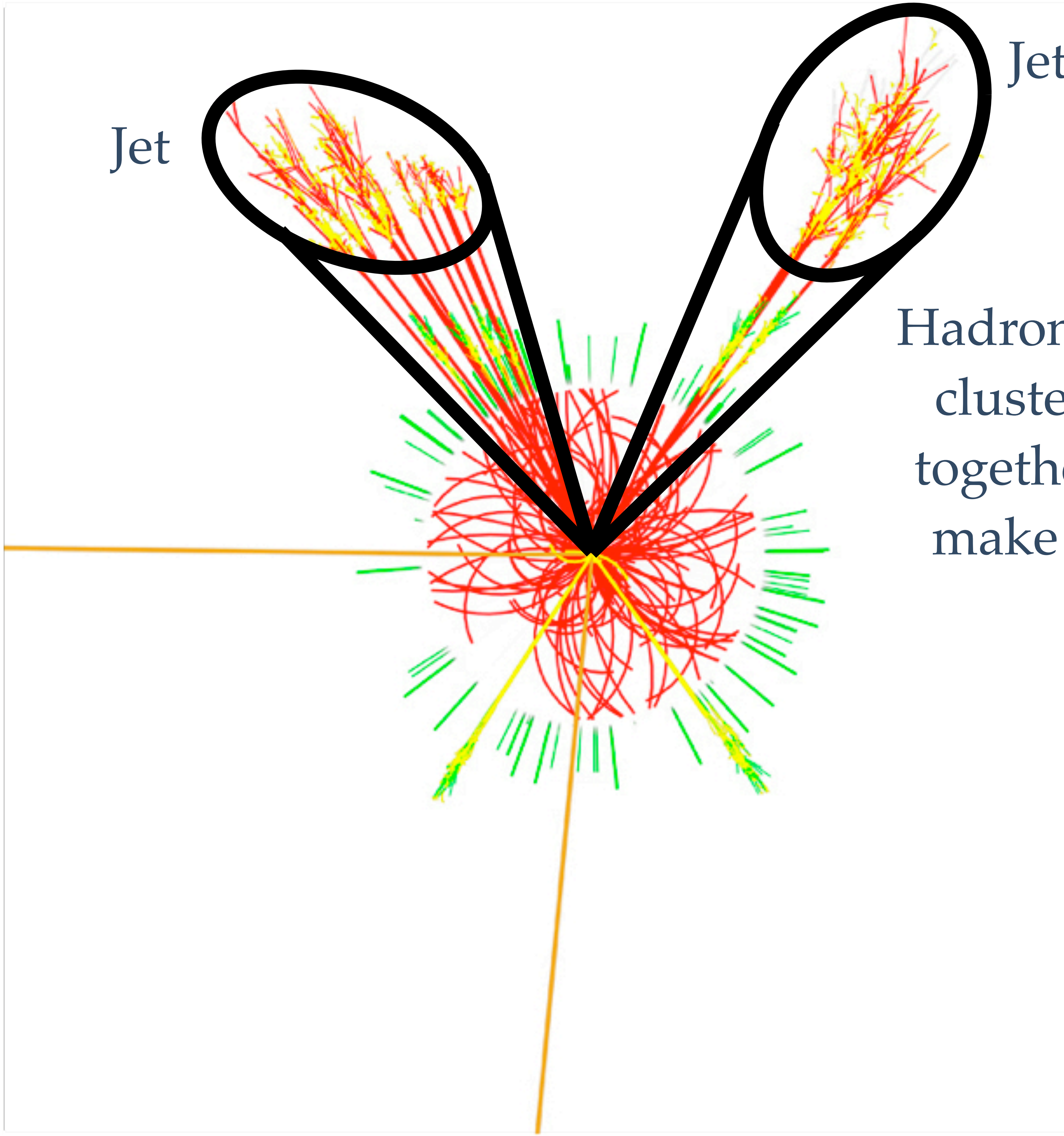
What we calculate

Jets make this correspondence



What we measure

Event picture from http://atlas.ch/photos/events.html



Hadrons are clustered together to make jets

Types of Algorithms

- There are two main classes of jet algorithm
- Sequential recombinations
  - Combine four-momenta one by one
- Cone algorithms
  - Stamp out jets as with a cookie cutter

Focus on these

Sequential Recombination

- Define a distance measure between every pair of four-momenta in an event (jet-jet distances)

$$d_{ij}$$

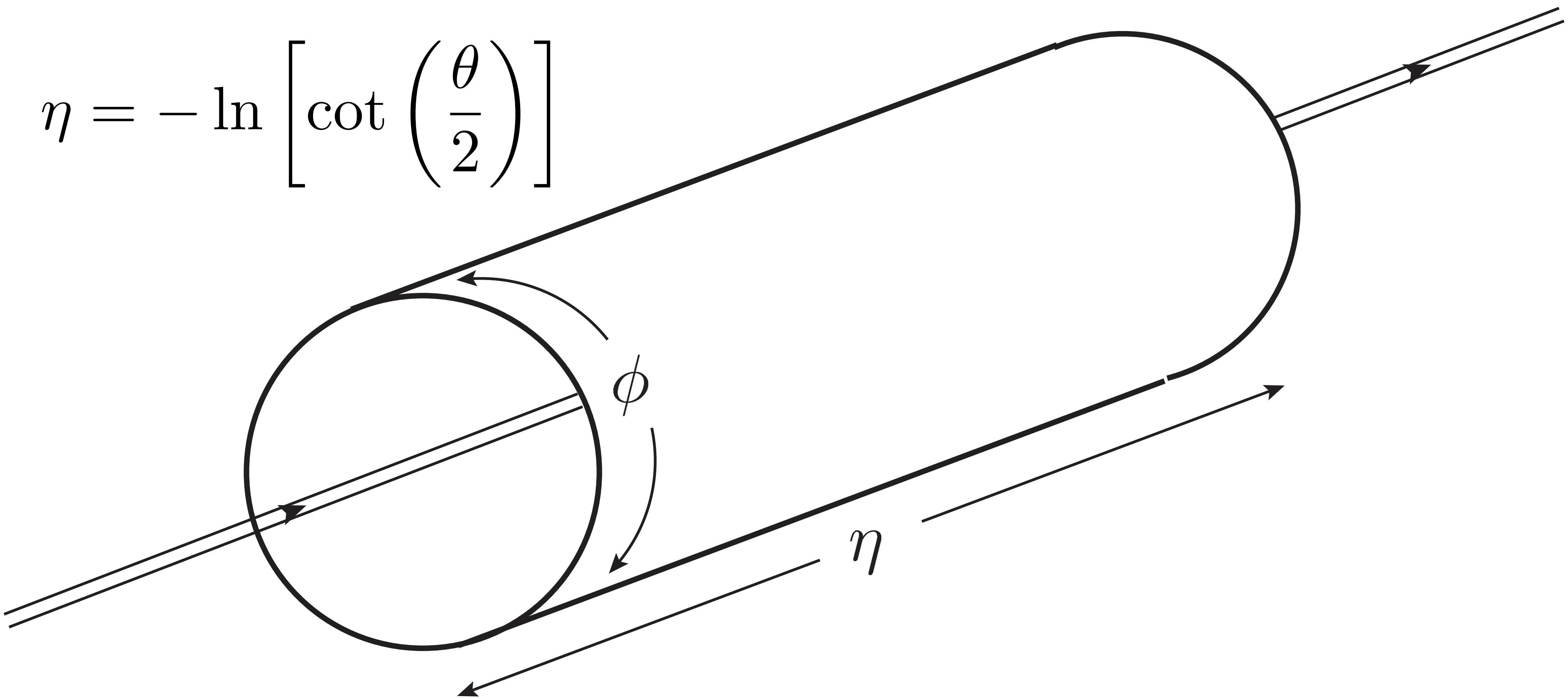
- Define a distance measure for each four-momenta individually (jet-beam distances)

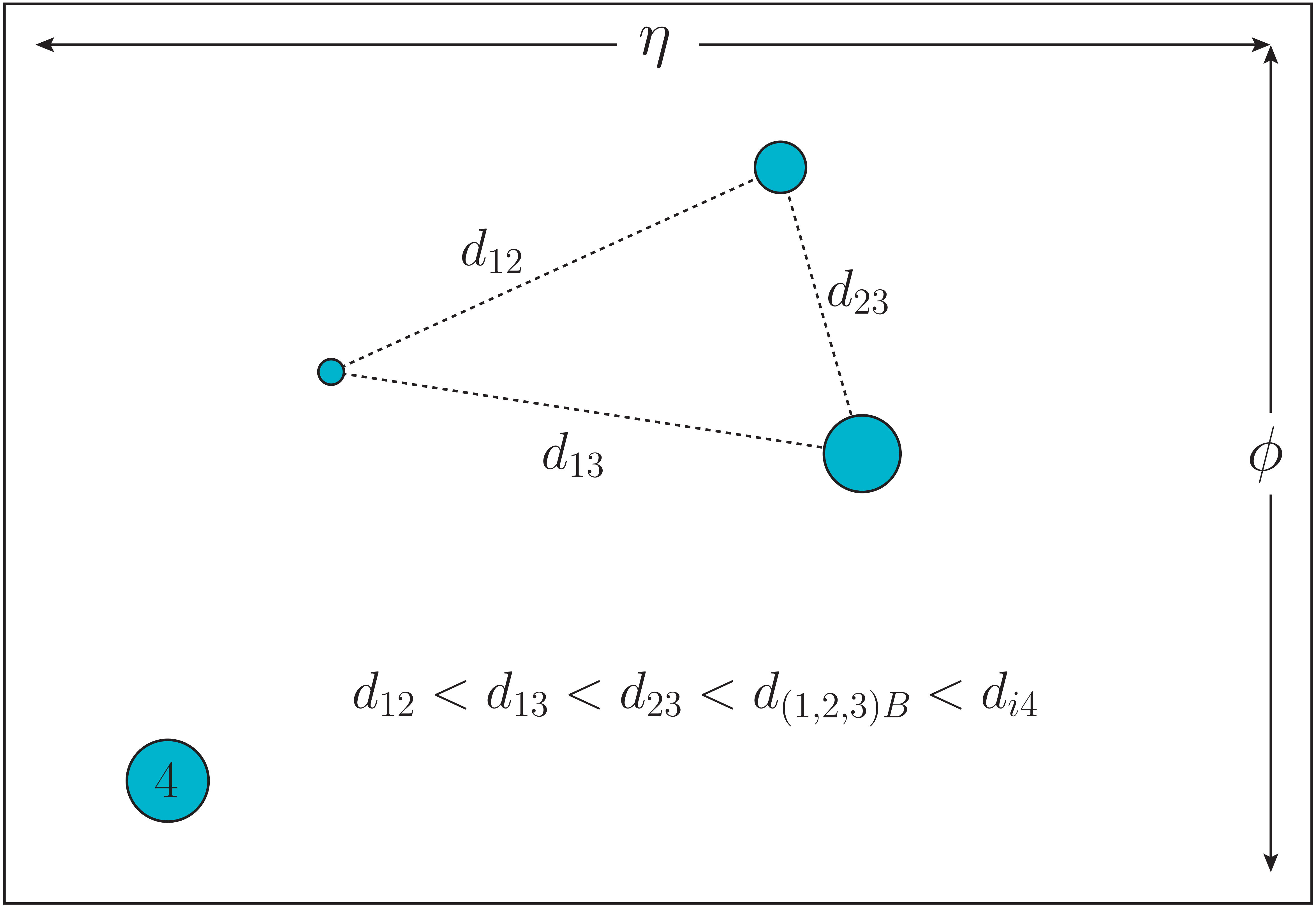
$$d_{iB}$$

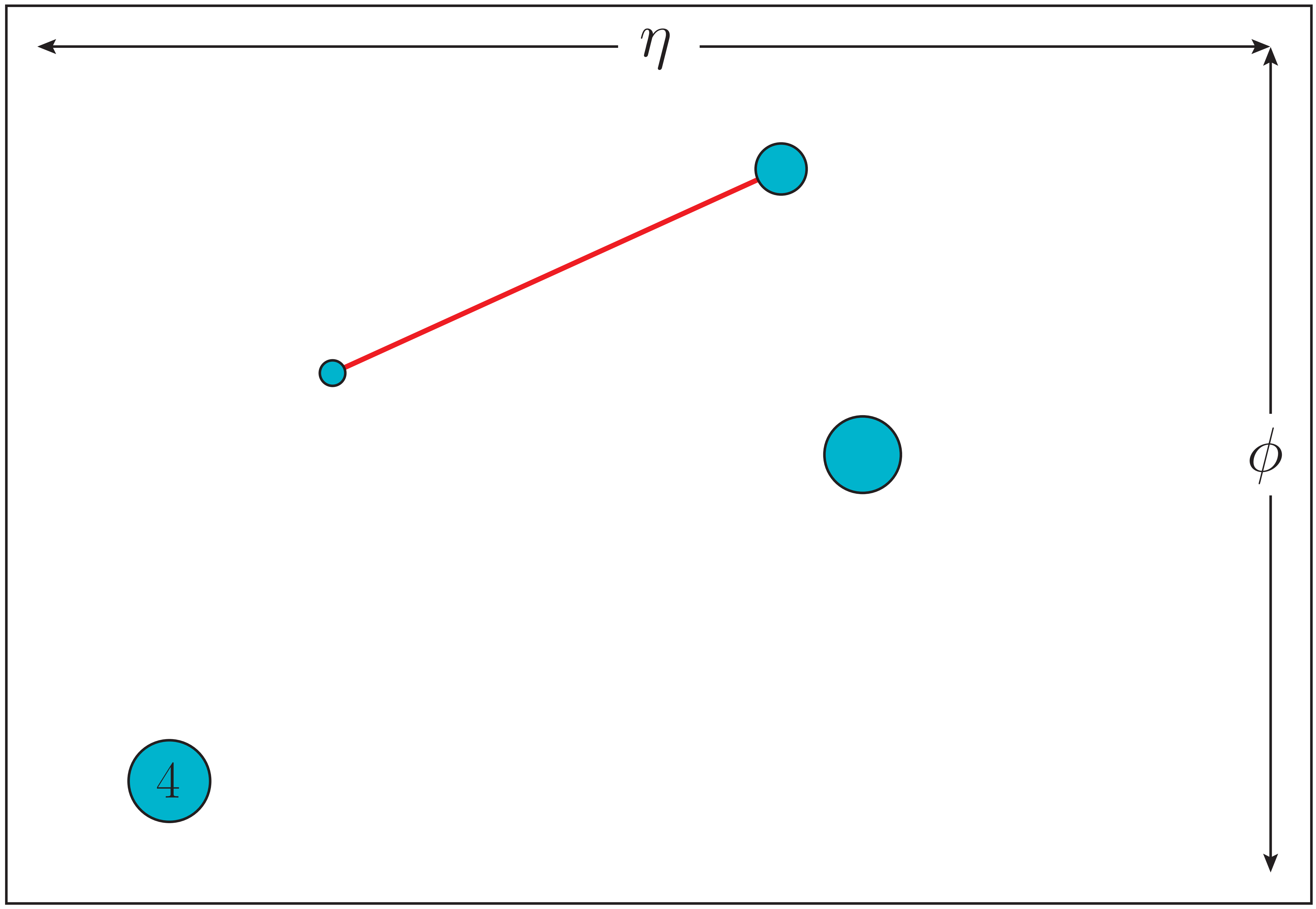
- If smallest distance at any stage in clustering is jet-jet, add together corresponding four-momenta
  - Otherwise take jet with smallest jet-beam distance and set it aside
- Repeat till all jets are set aside

Coordinate System

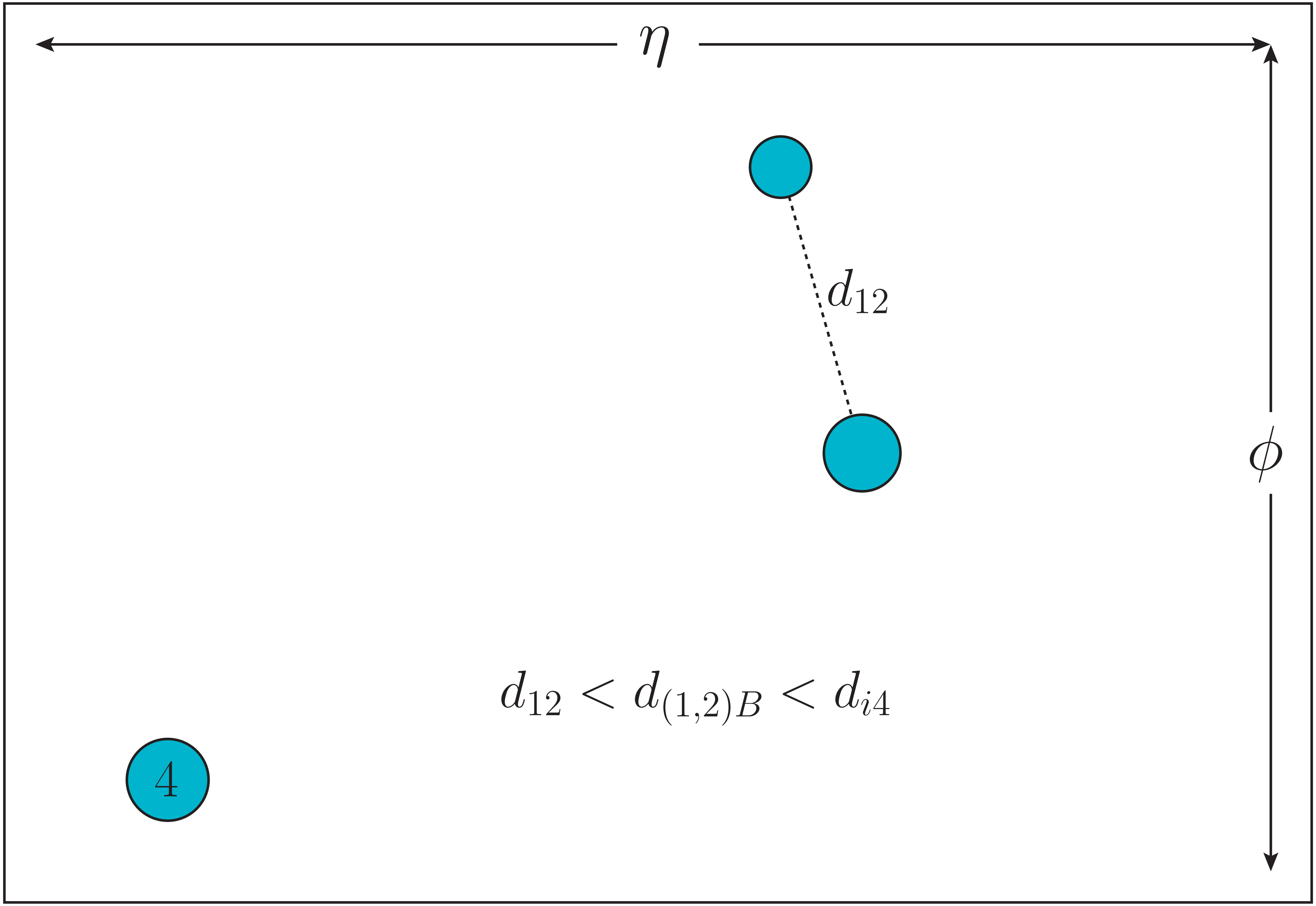
$$\eta = -\ln\left[\cot\left(\frac{\theta}{2}\right)\right]$$



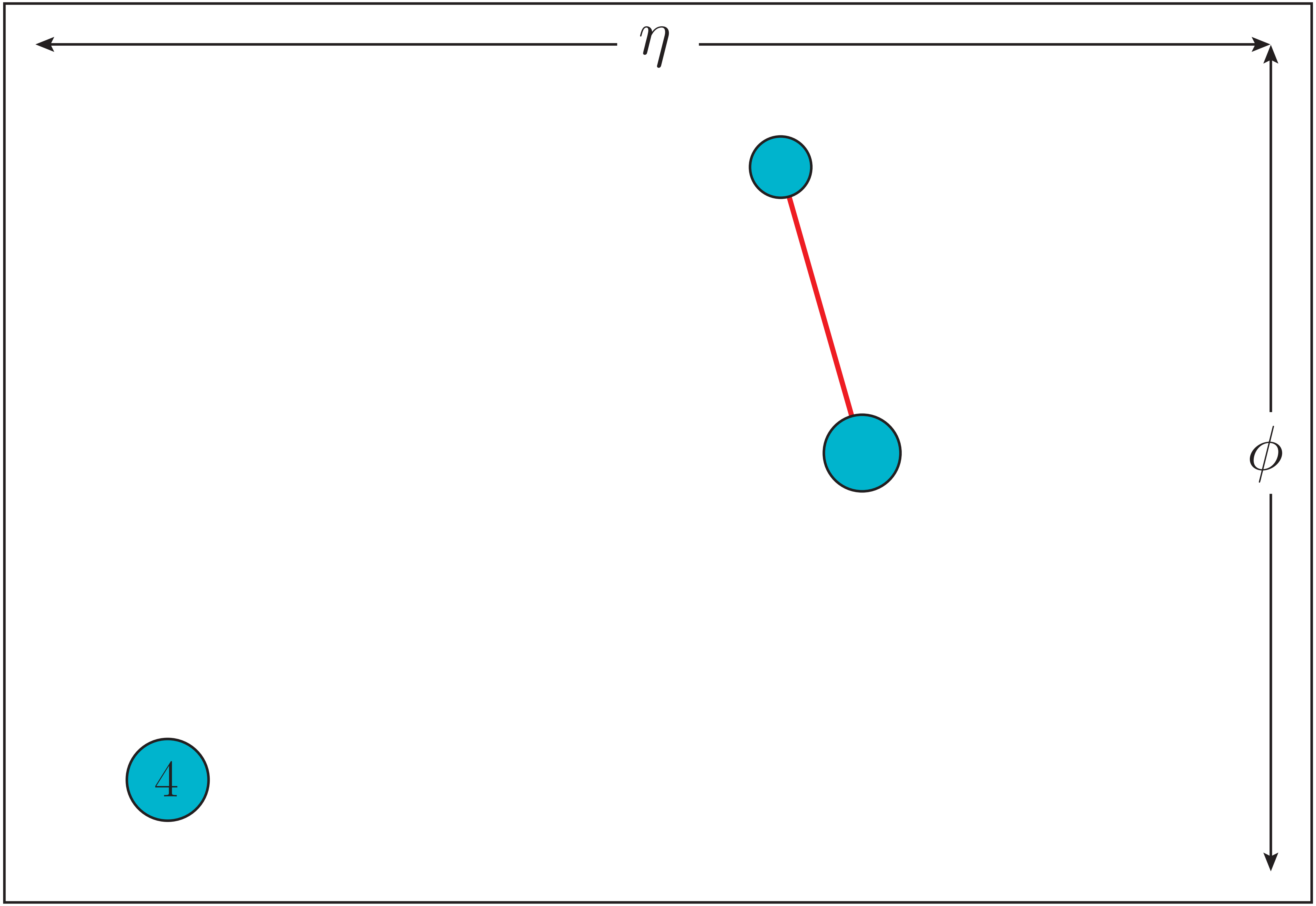

η
d_{12}
d_{23}
d_{13}
ϕ
$d_{12} < d_{13} < d_{23} < d_{(1,2,3)B} < d_{i4}$
4



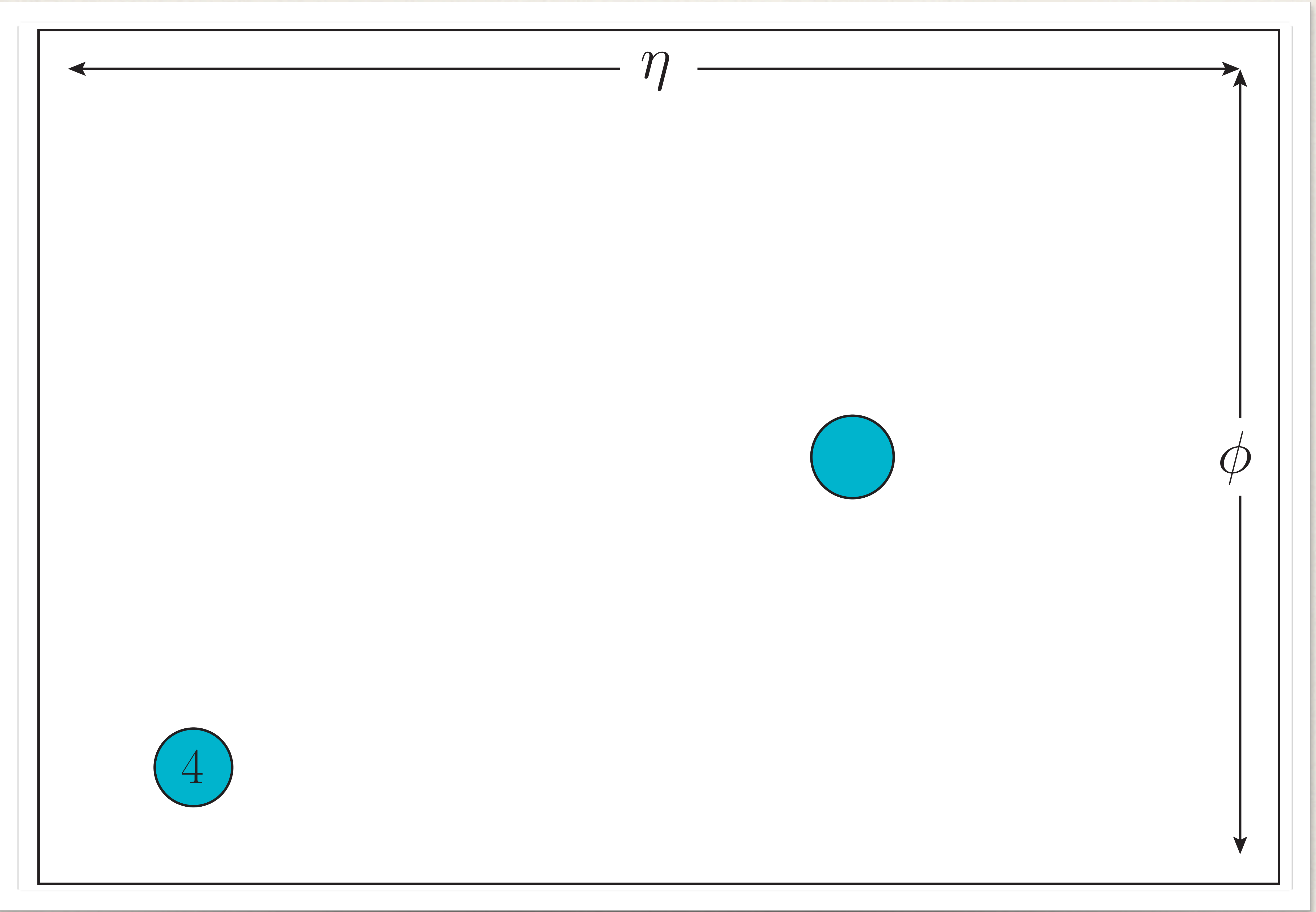
η
ϕ
4



η
ϕ
d_{12}
$d_{12} < d_{(1,2)B} < d_{i4}$
4



η
φ
4



Done!

Standard Recombination Algorithms

- k_T algorithm

$$d_{ij} = \min(p_{Ti}^2, p_{Tj}^2)\left(\frac{\Delta R}{R_0}\right)^2,\ \ d_{iB} = p_{Ti}^2$$

- C/A algorithm

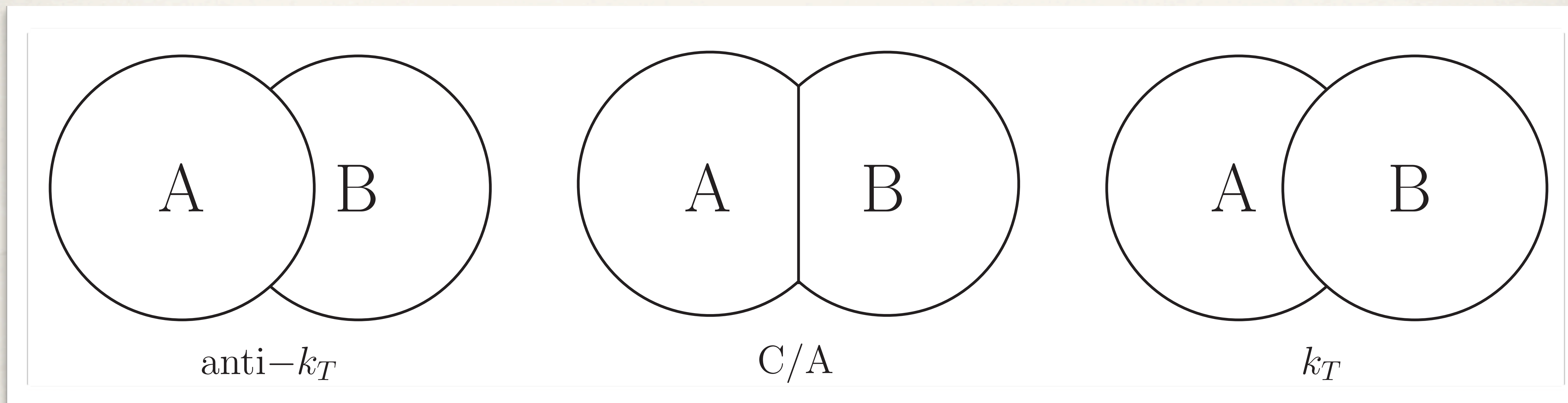
$$d_{ij} = \left(\frac{\Delta R}{R_0}\right)^2,\ \ d_{iB} = 1$$

- anti-k_T algorithm

$$d_{ij} = \min(p_{Ti}^{-2}, p_{Tj}^{-2})\left(\frac{\Delta R}{R_0}\right)^2,\ \ d_{iB} = p_{Ti}^{-2}$$

Approximate Jet Behavior:

$$p_{TA} > p_{TB}$$



Hard to Soft | Near to Far | Soft to Hard

Practical Aside: Fastjet

- Nearly everyone who works with jets uses Fastjet
  - Fastjet is a well written C++ framework for clustering jets
  - Download: http://www.lpthe.jussieu.fr/~salam/fastjet/
- Advantages:
  - Very well documented and supported
  - Easy to cluster and unwind jets
  - Nearly all standard algorithms supported
  - Support for plugins (e.g. pruning/filtering/trimming/VR, etc.)

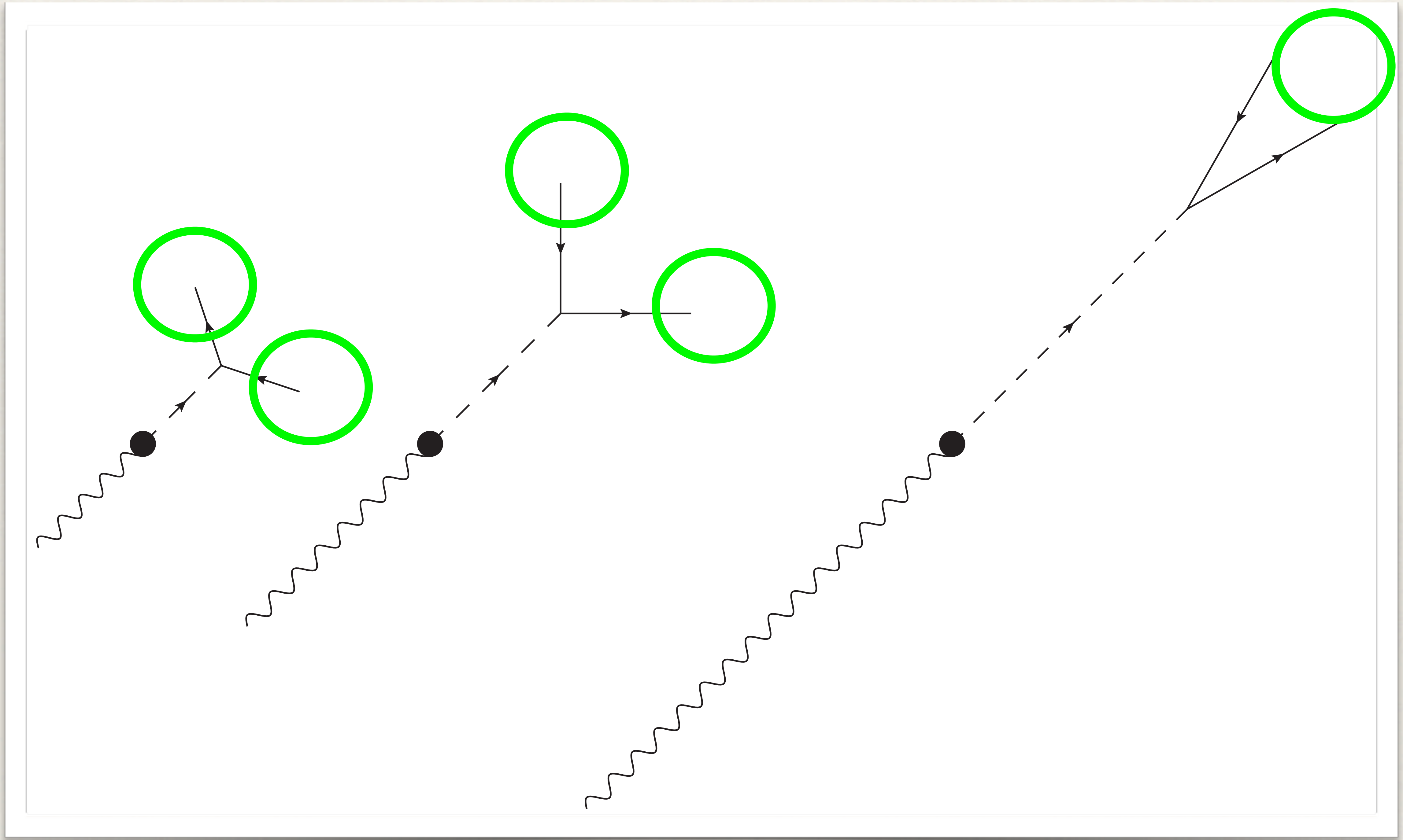
Boosted Heavy Objects and their Identification

Kinematics of Boosted Particles

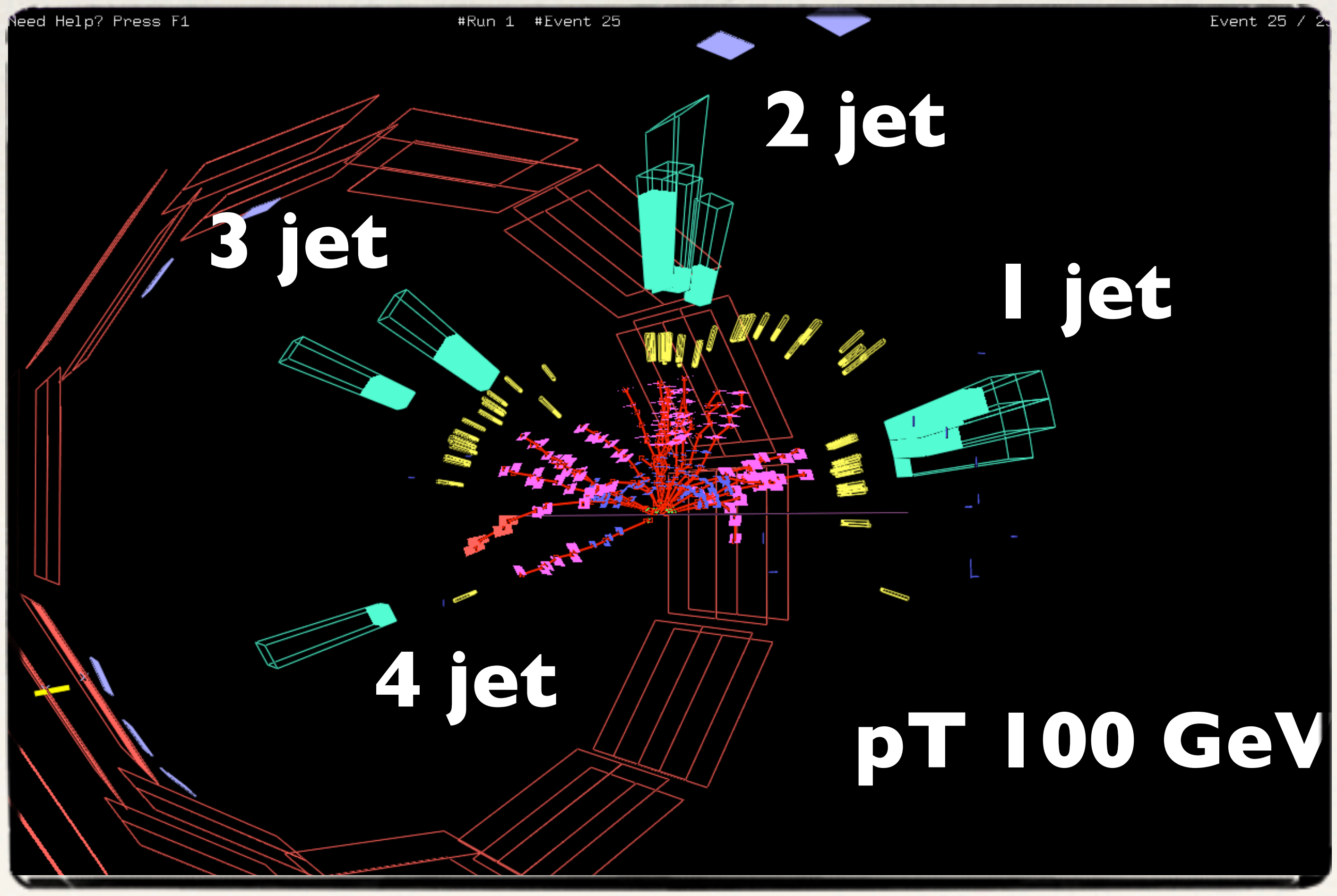
- The cone containing the decay products of a particle scales as

$$R \sim \frac{2m_X}{p_T}$$

- At LHC energies, even the heaviest particles we know of (Top, W, Z, Higgs) can become collimated (roughly $R < 1$).
- When this happens we say that they're "boosted".
- So we find that EW scale particles are clustered as a single jet as soon as their p_T exceeds a few hundred GeV.



Here one can see the effect - as we boost more and more (i.e. go to higher pT), the particles become more collimated.



Unboosted t-tbar pair

Boosted t-tbar pair

All three decay products of the top go into one jet

Figure source: http://www.pha.jhu.edu/groups/particle-theory/seminars/talks/F08/Yumiceva.pdf

What can we say about these jets?

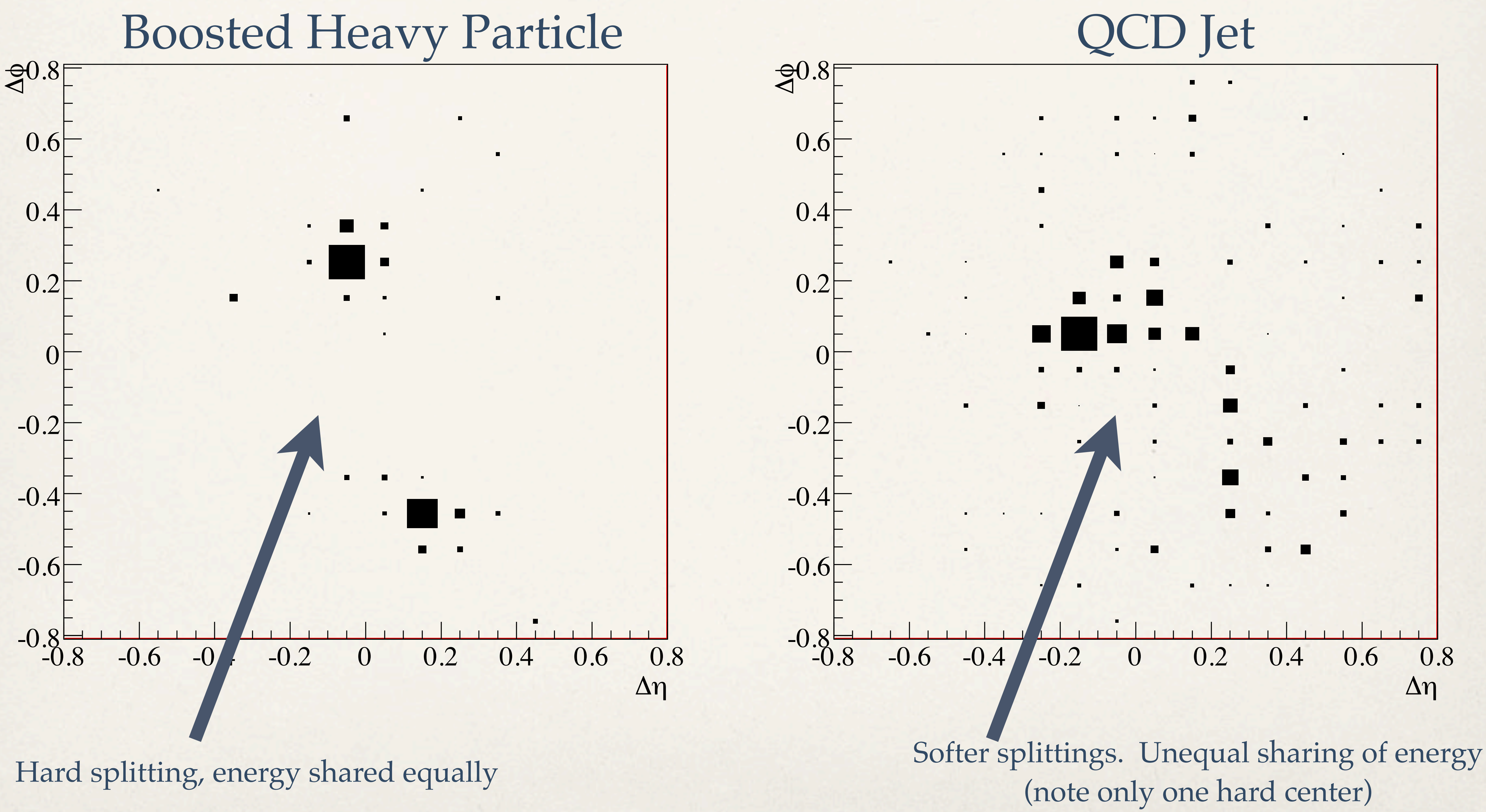
- Internally, QCD jets look really different than the jets of boosted heavy objects.

- If we start with a high energy gluon/quark, it wants to emit soft/collinear gluons:

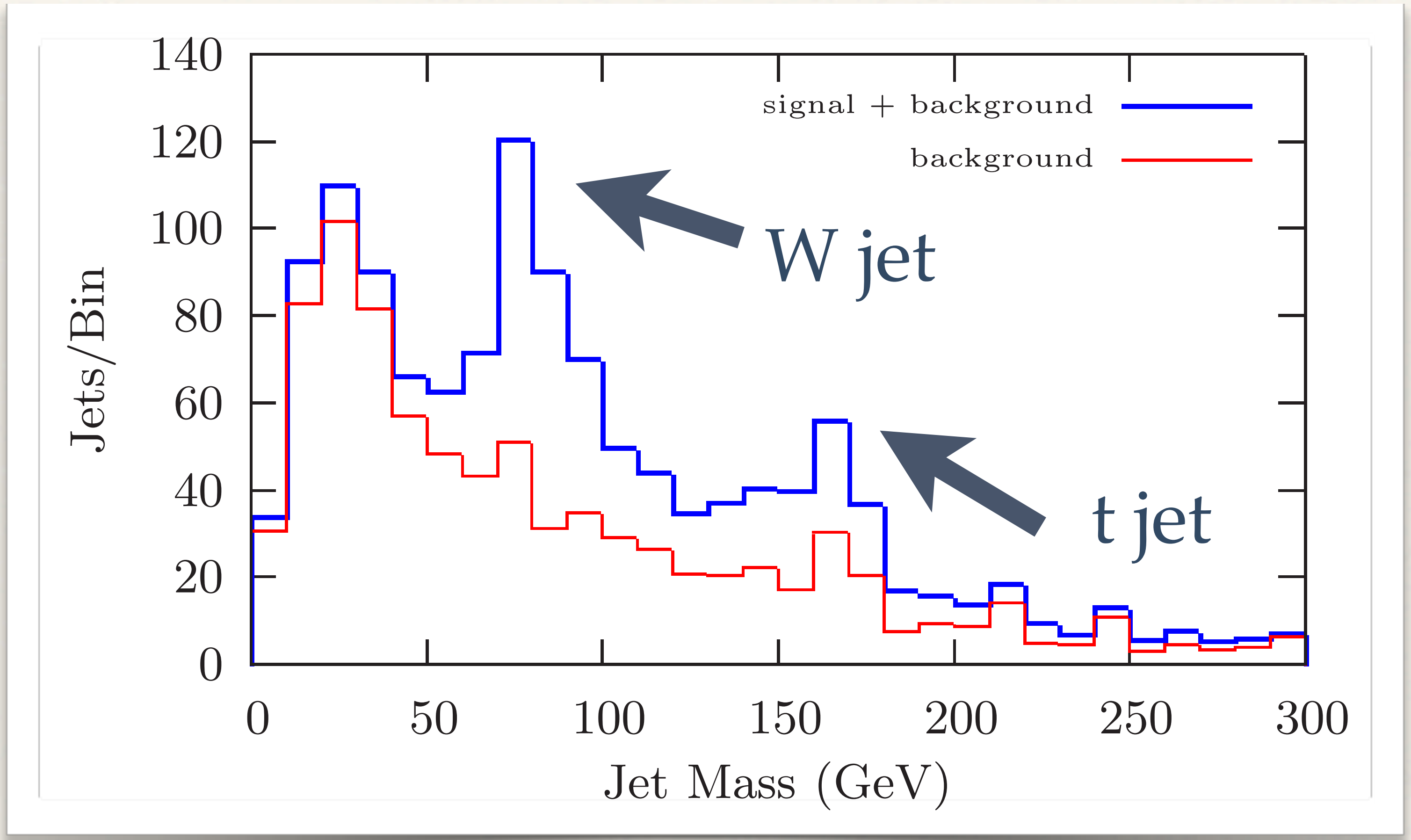
$$P_{q\to qg}(z) = C_F \frac{1+z^2}{1-z},$$
$$P_{g\to gg}(z) = C_A \left[\frac{1-z}{z} + \frac{z}{1-z} + z(1-z)\right]$$
$$P_{g\to q\bar{q}}(z) = T_R \left[z^2 + (1-z)^2\right],$$

Here P(z) measures how much a particle wants to emit another with energy fraction "z" (Altarelli-Parisi splitting fcns.).

- In contrast, a high energy heavy particle (W/Z/t/h) just decays - it has no singularity.



- Moreover, QCD jets have a continuum mass distribution, while the jets of boosted heavy particles have a fixed mass.



- These will form our main tools as we study the jets of boosted objects:

  1. Jet internal radiation distribution

  2. Jet mass

Figure: Skiba, Tucker-Smith, [hep-ph/0701247]

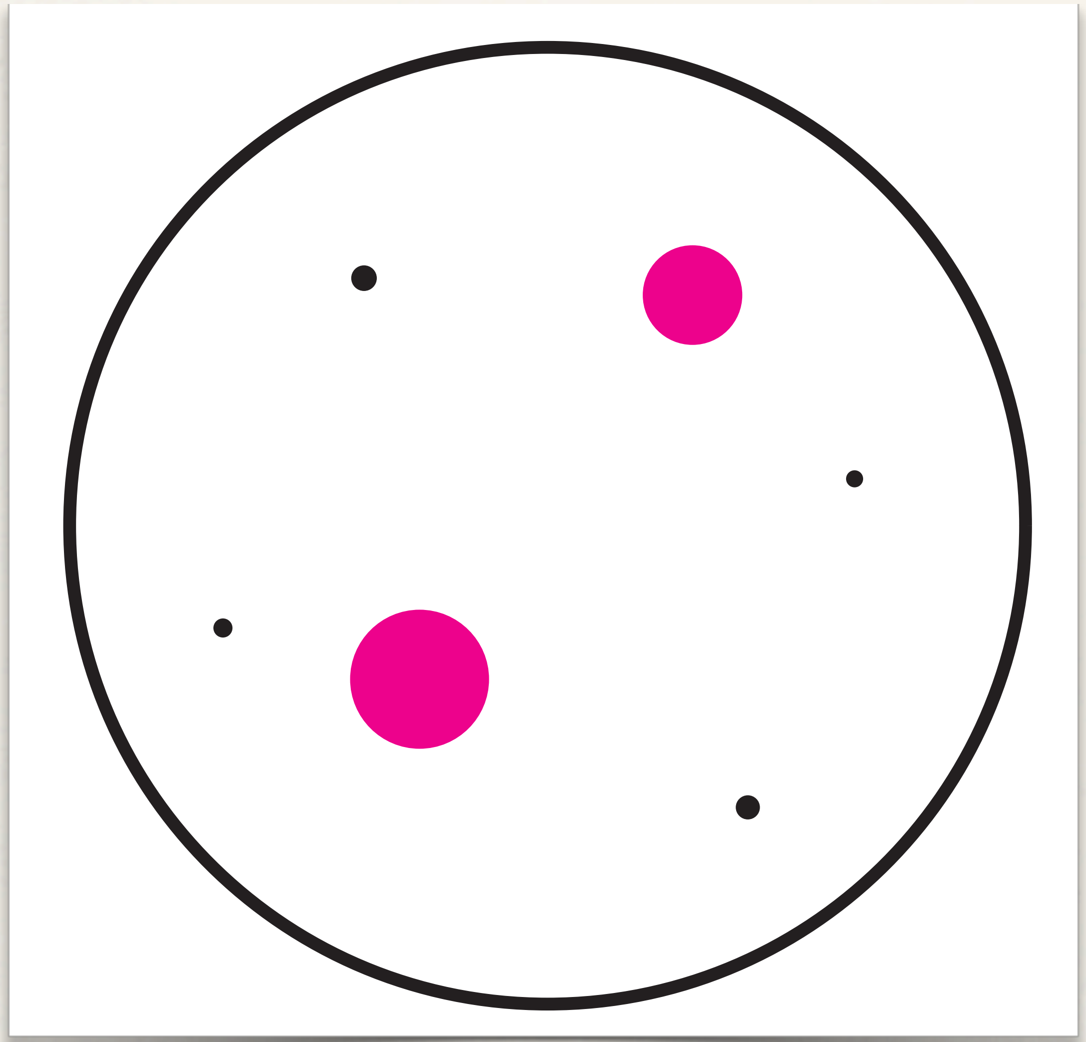
(1) Internal Radiation Distribution

- We can quantify the difference in appearance between QCD jets and the jets of boosted heavy particles.
  - Boosted objects tend to have a many hard prongs
  - QCD jets tend to have one central hard core
- Lots of jet shapes have been proposed. For instance (many more..):
  - Angularity: Almeida, Lee, Perez, Sterman, Sung, Virzi [0807.0234]
    - Measures how pencil-like/pancake-like a jet is
  - N-Subjettiness: Kim [1011.1493], Thaler, Van Tilburg [1011.2268]
    - How `N-prong like' does a jet look?

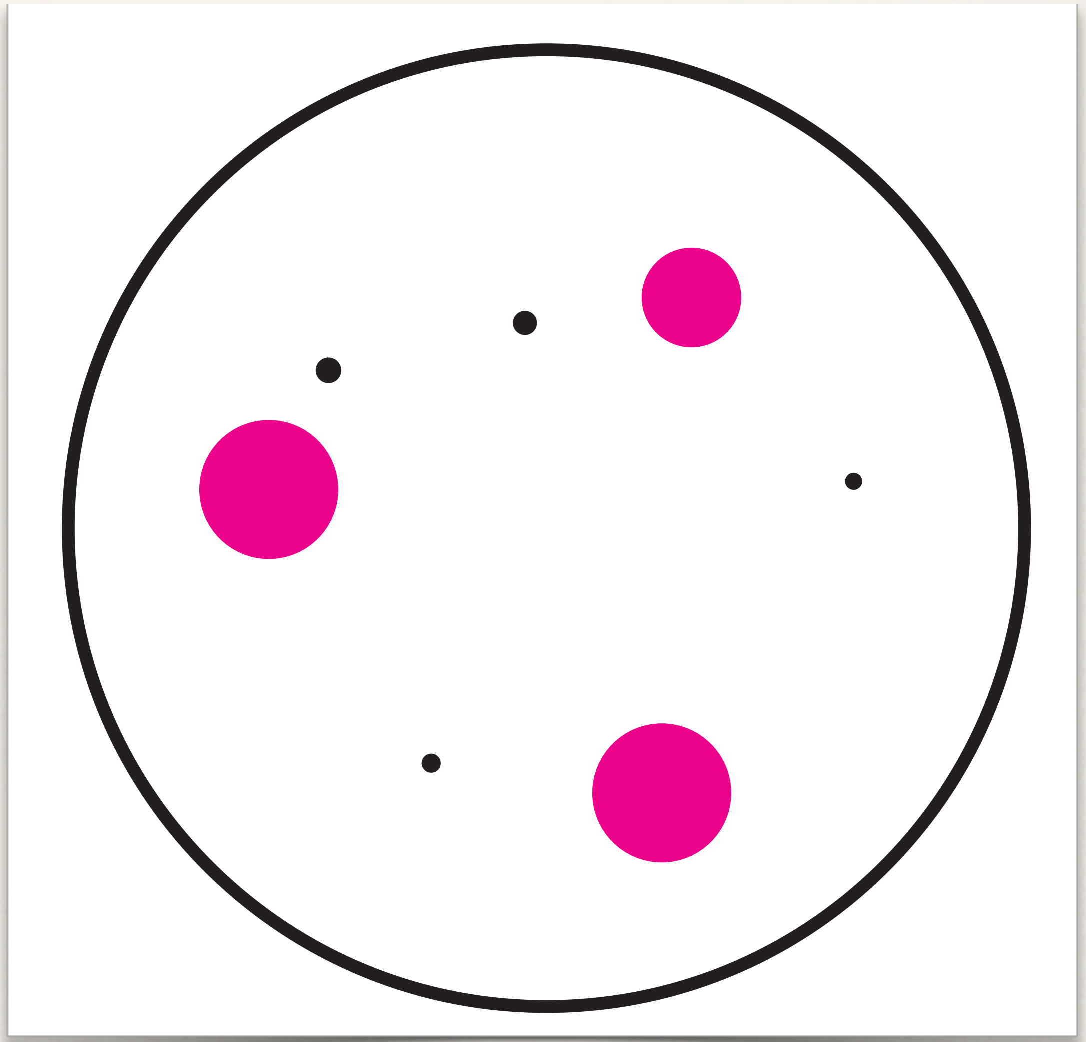
Example: Planar Flow

See Almeida, Lee, Perez, Sterman, Sung, Virzi [0807.0234] and Thaler, Wang [0806.0023]

QCD-like



Top-like



$$I_w^{kl} = \sum_i E_i \frac{p_{i,k}}{E_i} \frac{p_{i,l}}{E_i}$$

λ=Eigenvalues of I

$$\text{Pf} = \frac{4\lambda_1\lambda_2}{(\lambda_1+\lambda_2)^2}$$

(2) Jet Mass

- A lot of work on boosted objects has gone into into improving jet mass reconstruction (i.e. *Jet Topiary*).
- These aim to discard `contamination'
- Techniques on the market:
  - Filtering [Butterworth, Davison, Rubin, Salam: 0802.2470]
  - Pruning [Ellis, Vermillion, Walsh: 0903.5081,0912.0033]
  - Trimming [DK, Thaler, Wang: 0912.1342]

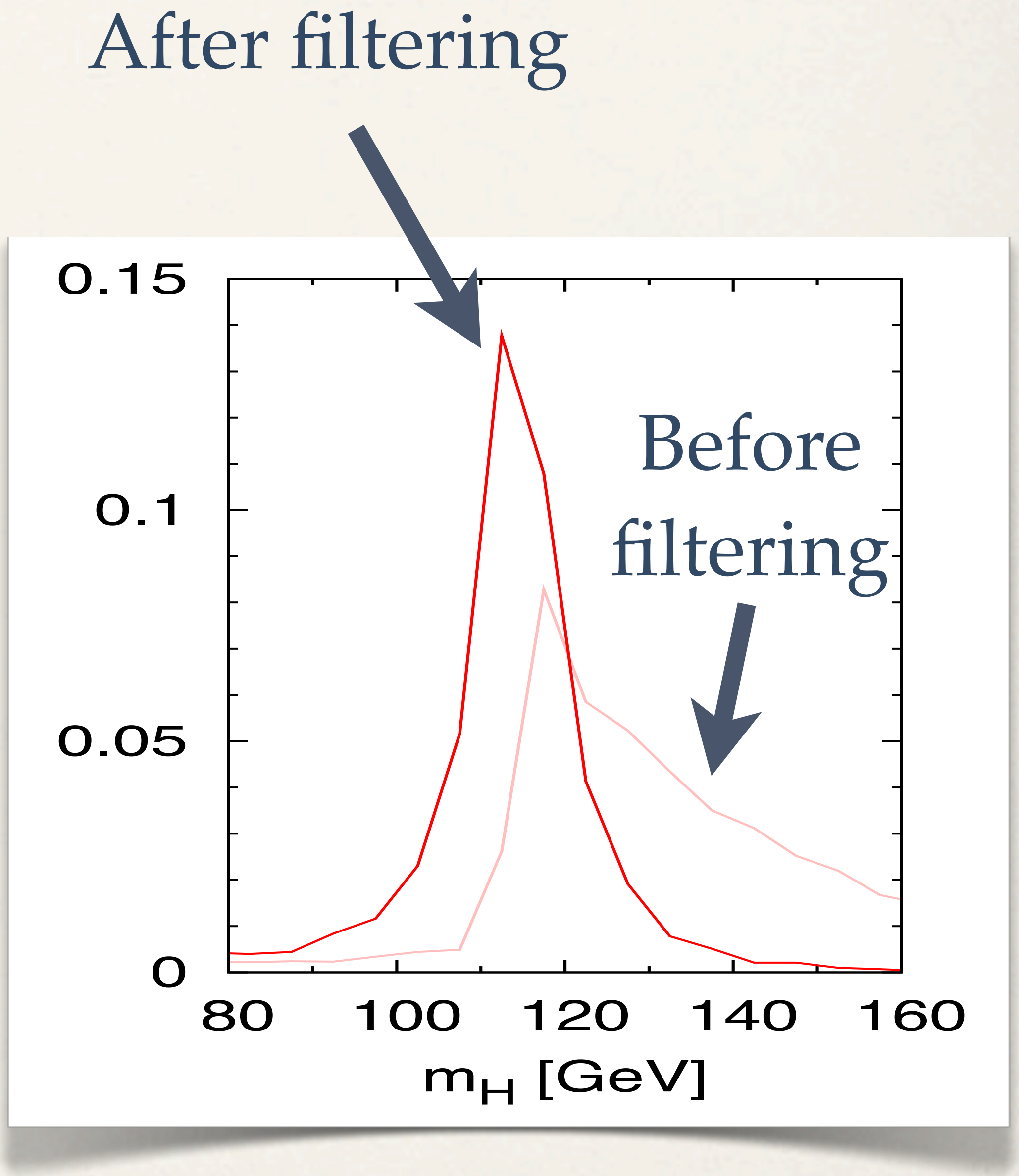


Figure: http://www.lpthe.jussieu.fr/~salam/talks/repo/2009-Princeton-BoostedHiggs.pdf

Example in Depth: Boosted SM Higgs

Boosted Higgs in pp->VH

- Canonical example of jet substructure usage.
- Want to see the Higgs in pp->VH where V is a W/Z decaying leptonically.
    - This is normally considered a challenging signal.

Jet substructure as a new Higgs search channel at the LHC, J. M. Butterworth, A. R. Davison, M. Rubin, G. P. Salam, [arXiv:0802.2470] Phys.Rev.Lett. 100 (2008) 242001

- Basic topology/kinematic requirements:
  - Look for one W/Z's decay products (i.e. two leptons or one lepton and MET)
  - Require pT(j)>200GeV
    - Ensures 2m/pT ~ 1 or less
  - Two b-tags inside the candidate Higgs jet
    - Kills a lot of background, but there is still some from g->b bbar

- Main idea:
  - Use what we know about the radiation pattern to improve the mass measurement of the Higgs jet.
    - This process is called filtering
  - Look for three hard subjets inside the Higgs jet.
    - Two from the b's

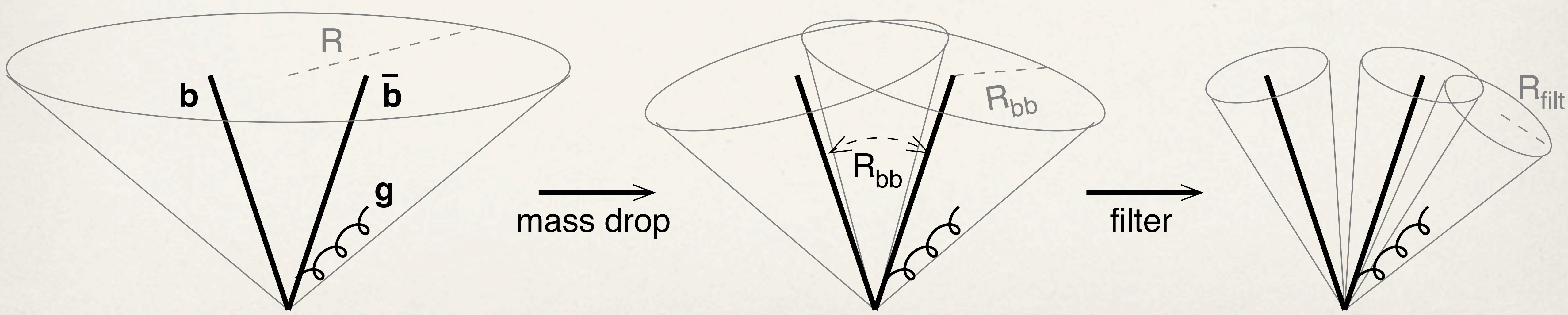


Figure source: Jet substructure as a new Higgs search channel at the LHC, J. M. Butterworth, A. R. Davison, M. Rubin, G. P. Salam, [arXiv:0802.2470] Phys.Rev.Lett. 100 (2008) 242001

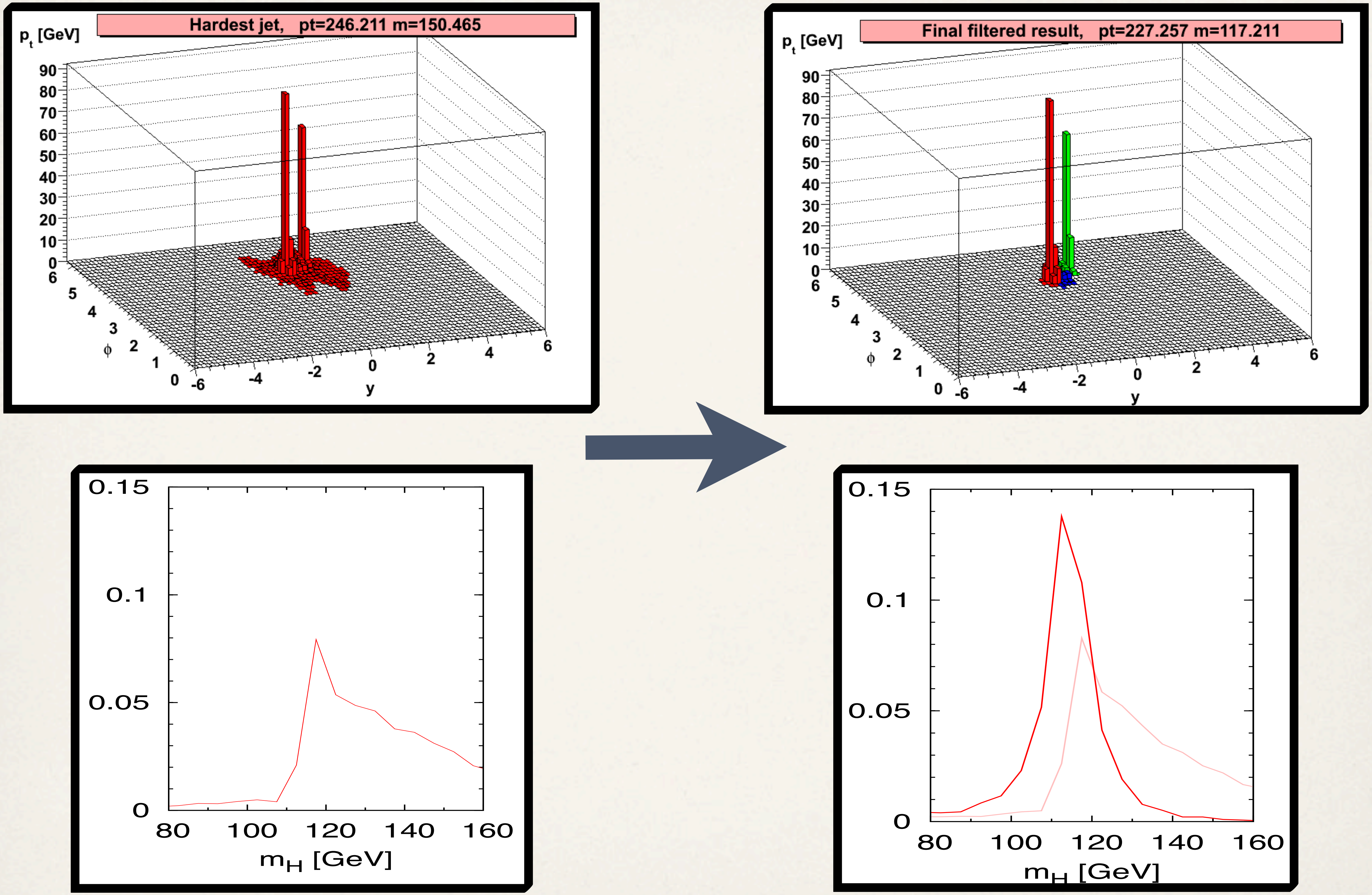


Figure source: http://www.lpthe.jussieu.fr/~salam/talks/repo/2009-boosted-IPMU-merged.pdf

Higgs is now visible in b-bbar with 30 fb^{-1}

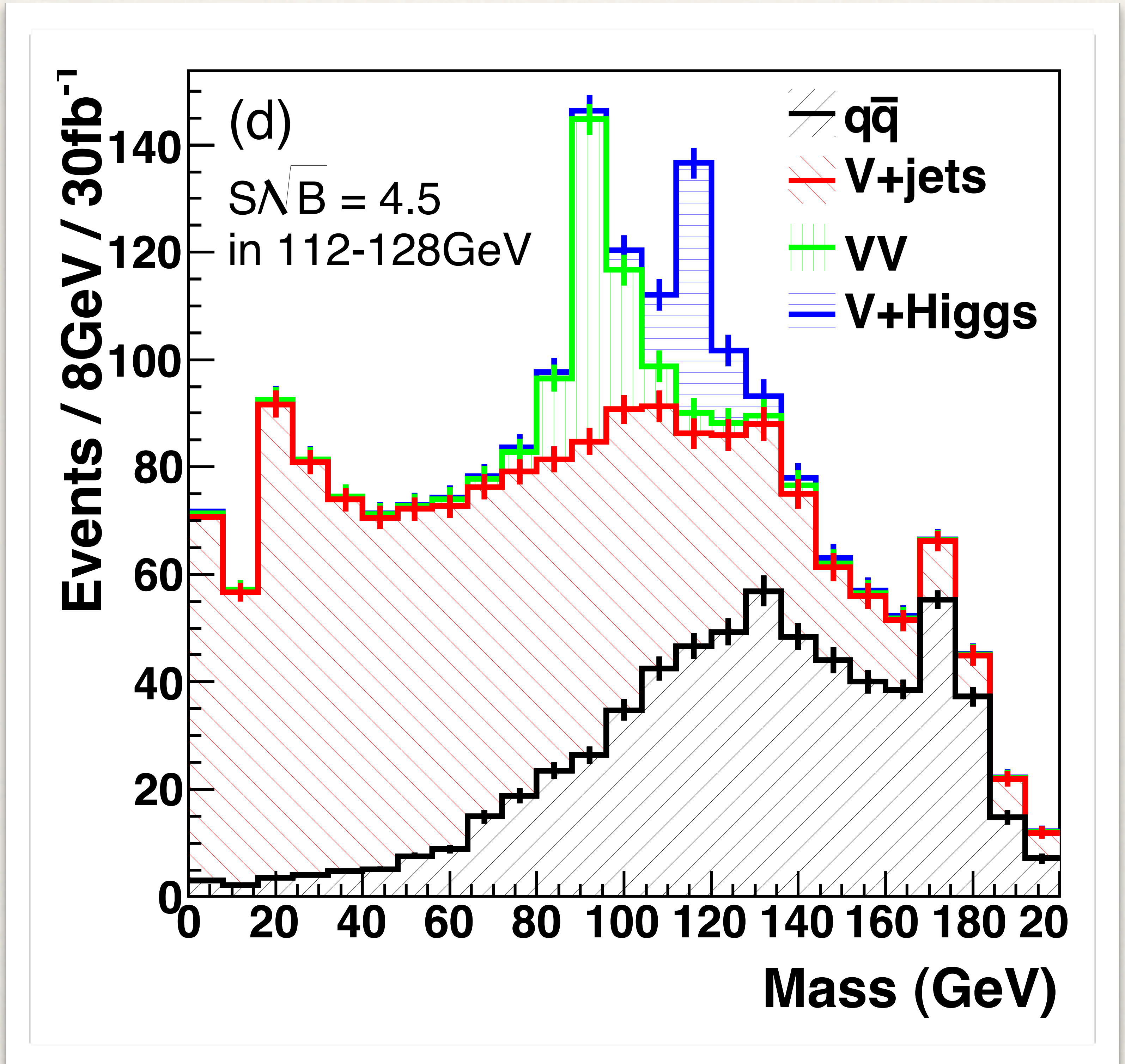


Figure source: Jet substructure as a new Higgs search channel at the LHC, J. M. Butterworth, A. R. Davison, M. Rubin, G. P. Salam, [arXiv:0802.2470] Phys.Rev.Lett. 100 (2008) 242001

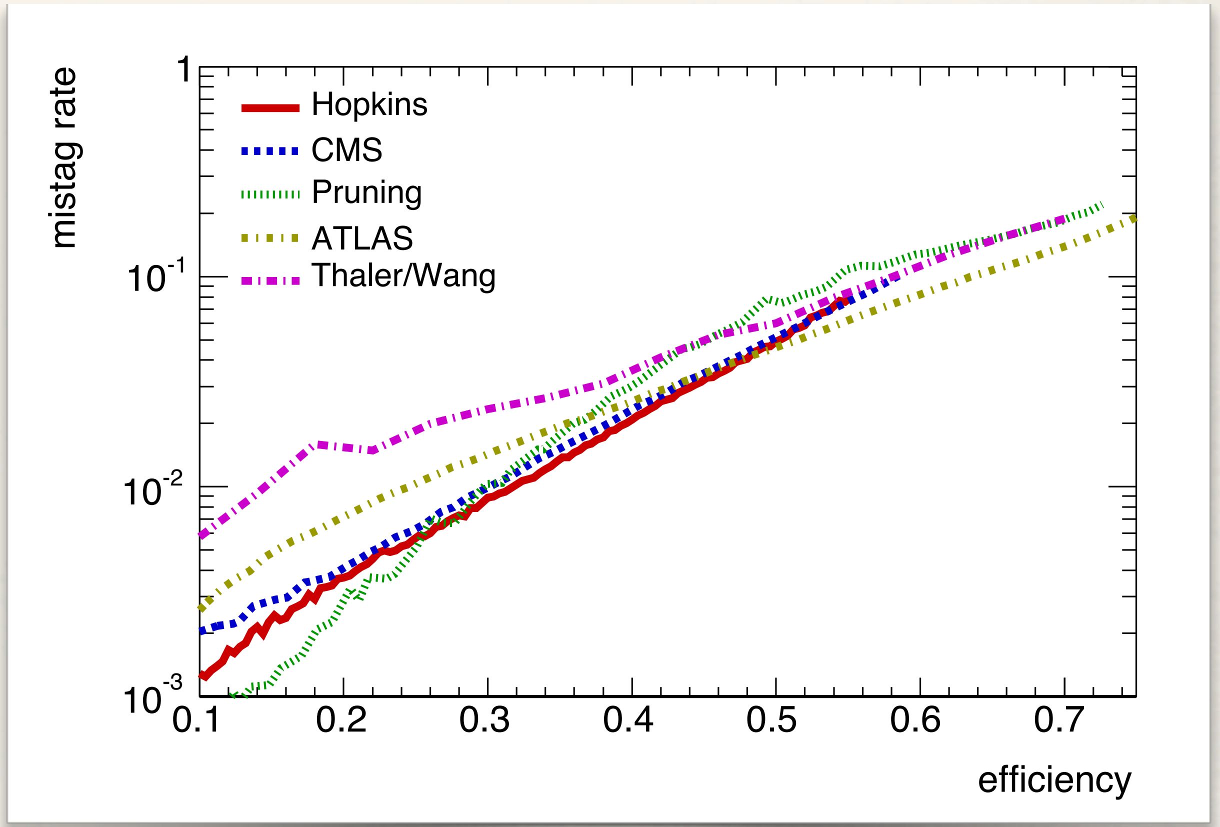
Overview of Other SM Analyses

Higgs

- Another possibility (beyond the aforementioned V+H analysis) is tt +H: Plehn, Salam, Spannowsky [0910.5472]
    - Analysis uses `fat' jets (R=1.5)
        - Similar, but more complicated jet unraveling process
    - Solves many combinatoric issues, improves mass resolution
    - Useful algorithm for finding mildly boosted tops.
        - Known as `HEP' tagger (Heidelberg, Eugene, Paris)

Tops

- Top tagging is now a well studied topic
- Many established tagging techniques:
  - Kaplan, Rehermann, Schwartz, Tweedie [0806.0848]
  - Thaler, Wang [0806.0023]
  - Almeida, Lee, Perez, Sterman, Sung, Virzi [0807.0234]
  - See [1012.5412] for an overview and review



Roughly comparable efficiencies

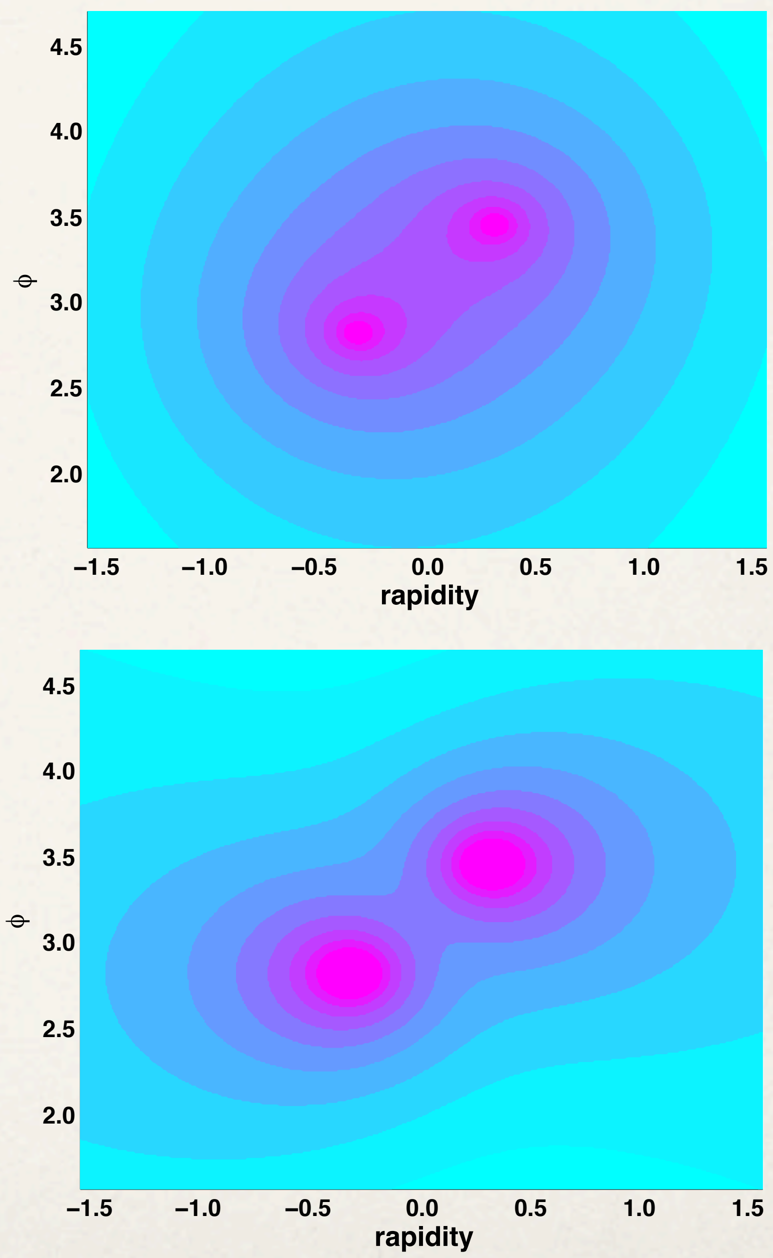
Figure source: 1012.5412

W/Z

- Multivariate methods sensitive to radiation pattern: Cut, Han, Schwartz [1012.2077]: W-Tagging
- Measure W polarization in vector boston fusion to distinguish anomalous Higgs couplings: DK, Han, Wang, Zhu [0911.3656]
- Applications to Z′ physics: Katz, Son, Tweedie [1010.5253]

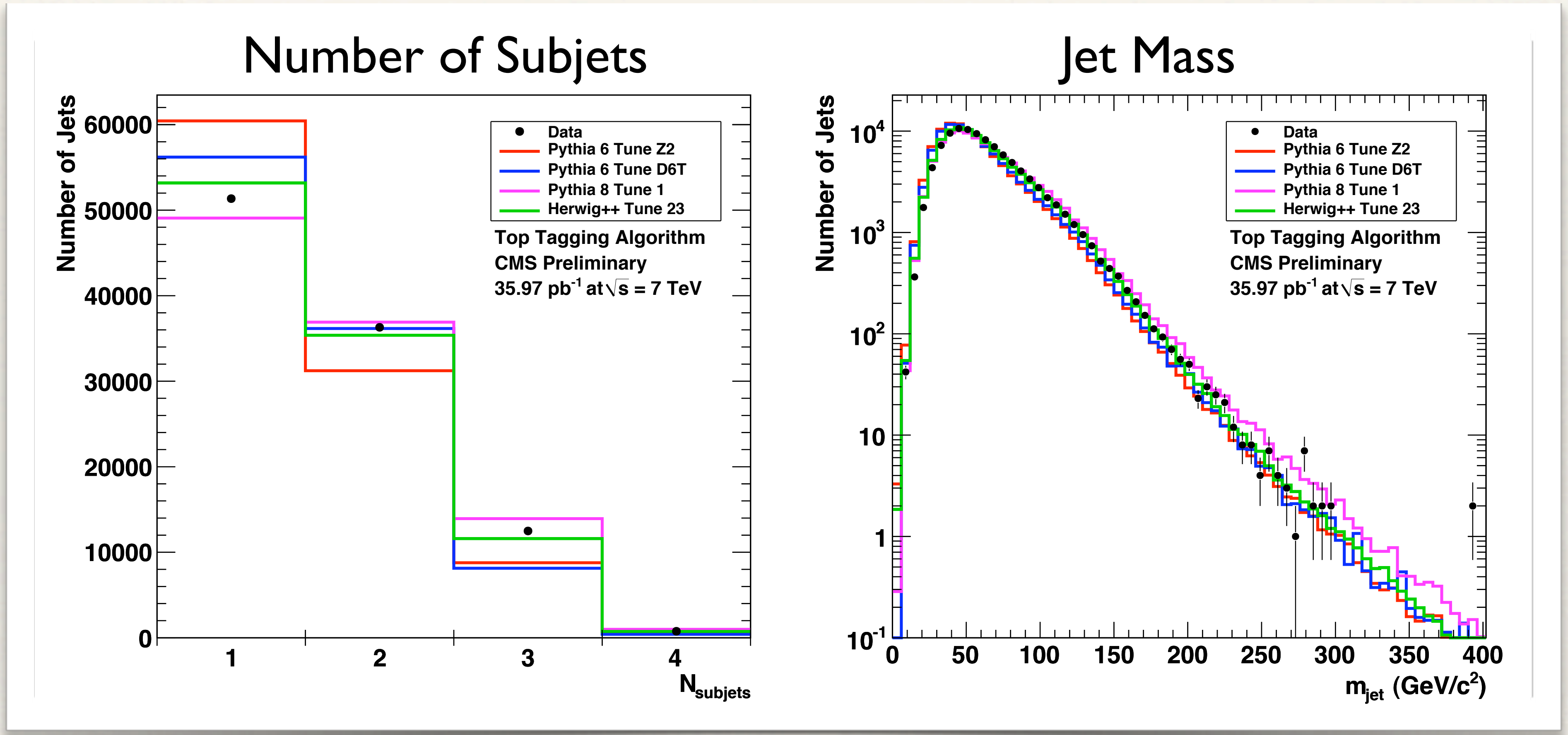
Ever more sophisticated techniques

- Measure hadronic top polarization: DK, Shelton, Wang [0909.3855]
- Measure fit to radiation profile: Almeida, Lee, Perez, Sterman, Sung [1006.2035]
- Top identification taking into account color singlet nature of W: Hook, Jankowiak, Wacker [1102.1012]



Source: 1102.1012

- LHC experiments are going to use these techniques!
- Validation has started using QCD data.
- MC and experiment agree well:



- Herwig++ seems to show the best agreement

Source: http://www.physics.uoregon.edu/~soper/Jets2011/Dolen.pdf

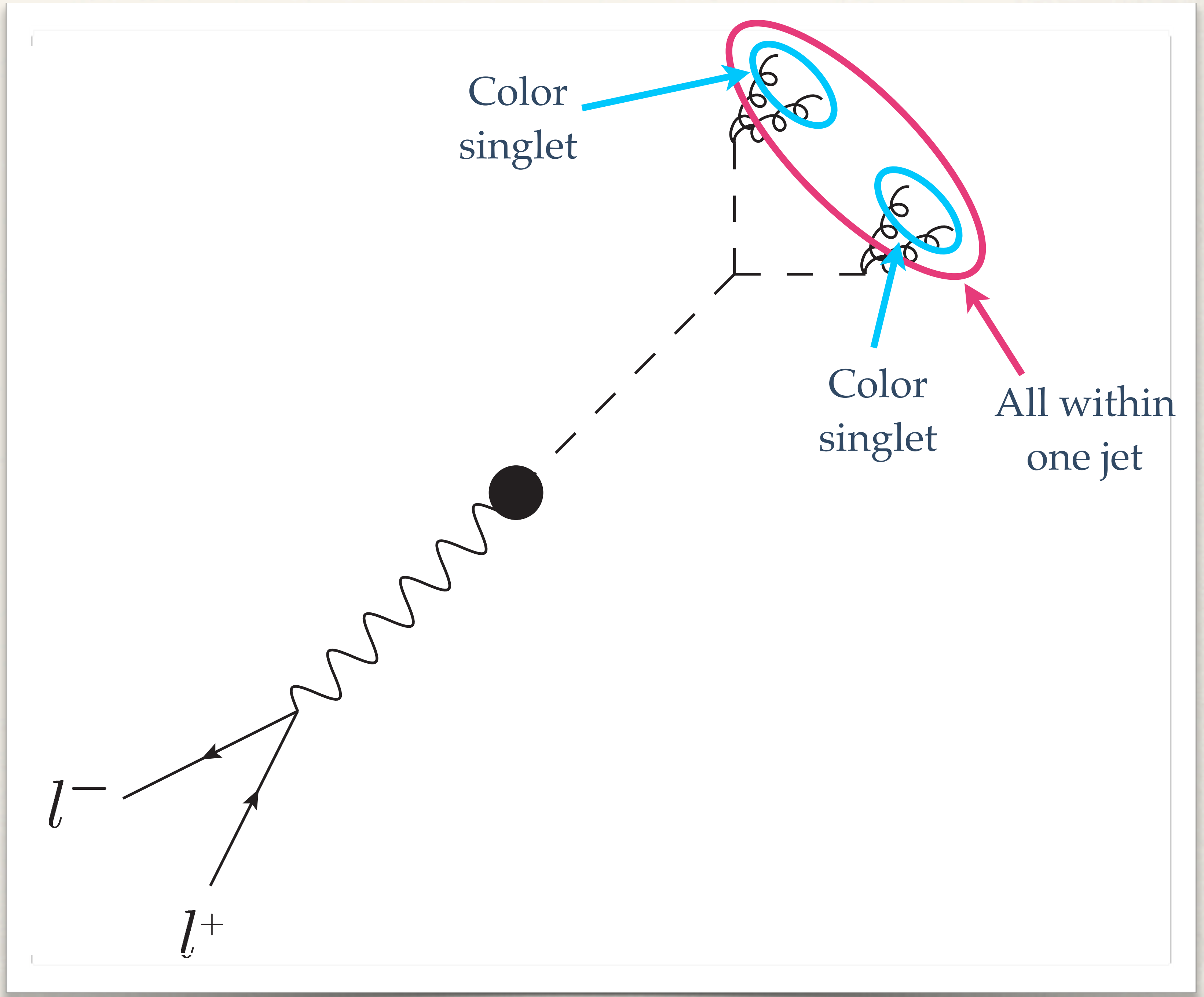
Applications in BSM Physics

Significant Benefits Using Substructure

- TeV scale BSM physics coupling to EW scale particles -> boosted objects!
- BSM events are complicated - jet substructure aids in reconstruction (HEP tagger)
- Substructure can resolve color structure - useful for exotic decays.

Qualitatively New Measurements

- In `Buried Higgs' models the Higgs decays to two scalars, which subsequently decay to gluons
- h>2a>4g
- Substructure methods can look for color singlet nature of scalar: see details in [1006.1650]

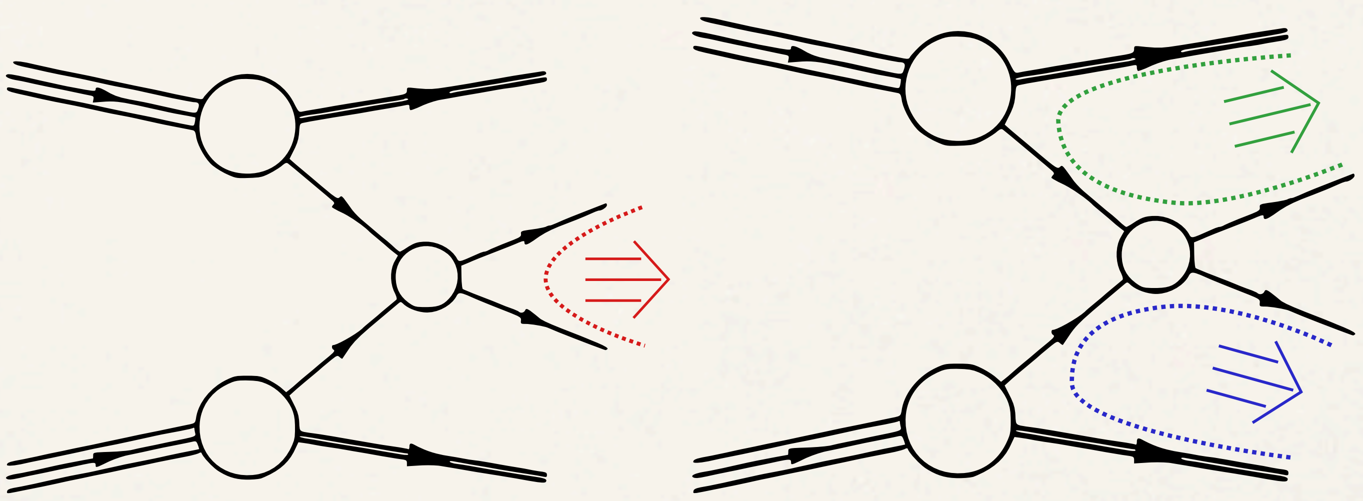


- Some examples (incomplete list):
  - Neutralino decays: Butterworth, Ellis, Raklev, Salam [0906.0728]
  - SUSY Higgses (often result from SUSY cascade decays):
    - Kribs, Martin, Roy, Spannowsky [0912.4731,1006.1656]
  - Exotic Higgses:
    - Chen, Nojiri, Sreethawong: [1006.1151]
    - Falkowski, DK, Shelton, Thalapillil, Wang [1006.1650]
    - Bellazzini, Csaki, Hubisz, Shao [1012.1316]
  - Boosted gluinos from heavy squarks: Fan, DK, Mosteiro, Thalapillil, Wang [1102.0302]

Jet Superstructure

Basic Idea

- The processes giving rise to new-physics particles can have a color structure different than that of the SM background:



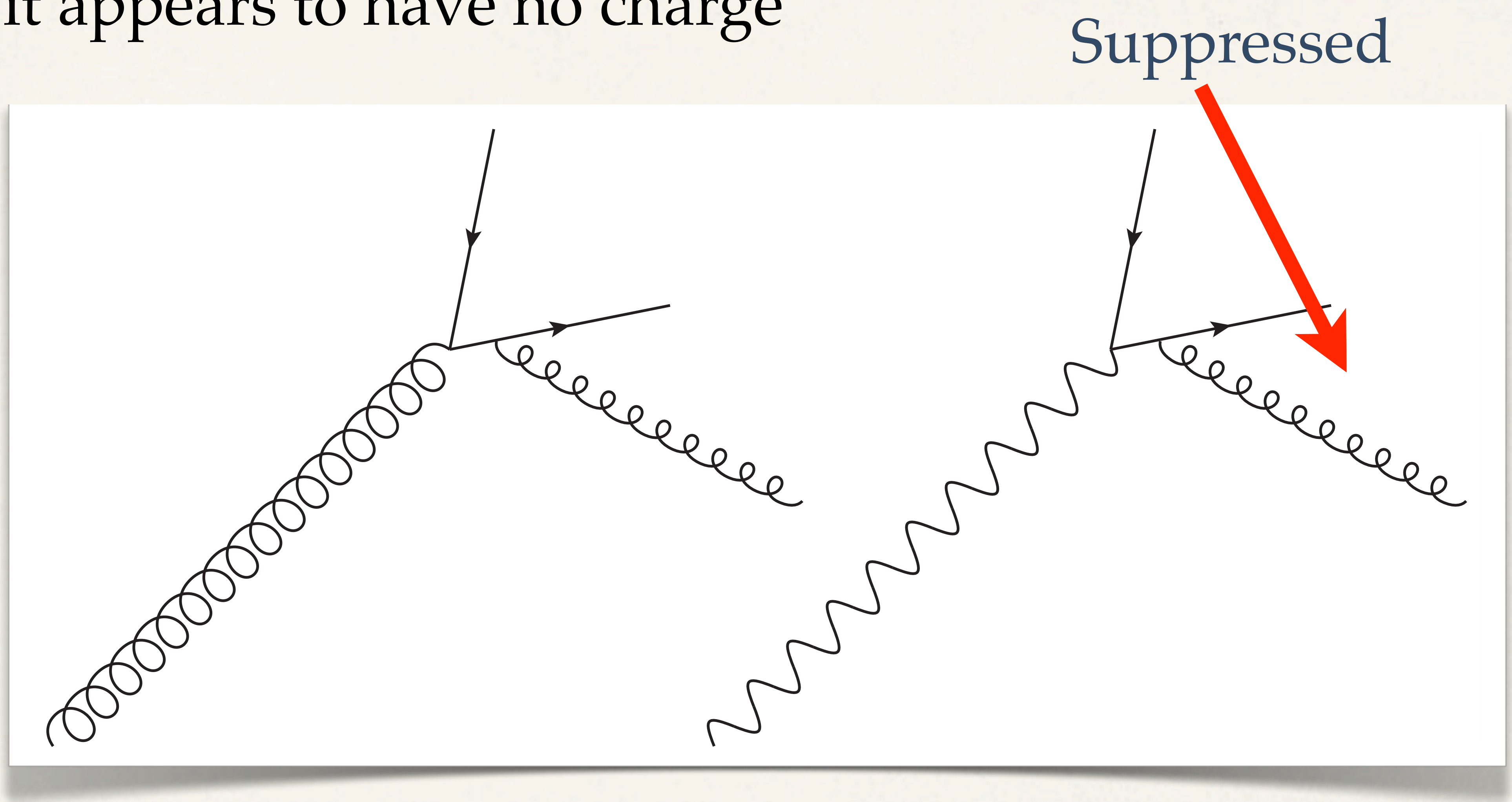
pp->H->bb pp->g->bb

- This will affect a jet's relation with other colored particles in the event.

Figure source: Seeing in Color: Jet Superstructure J. Gallicchio and M. D. Schwartz [1001.5027]

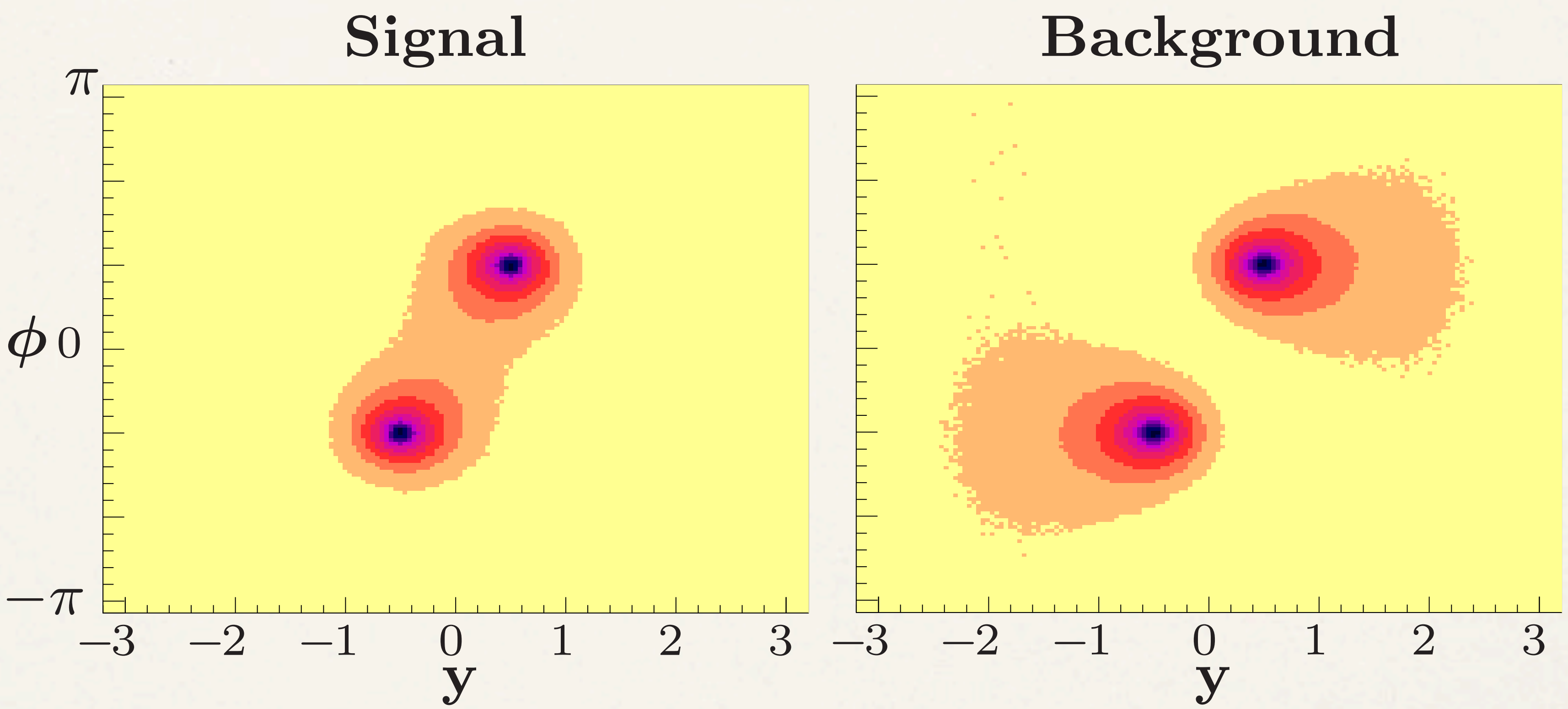
- Wide angle radiation away from a color singlet is suppressed.

- Think of the color dipole picture from E&M: far enough away from a dipole it appears to have no charge



- By looking for the regions of suppressed/enhanced radiation one can hope to infer information about an event's color structure.

- This was originally proposed to reduce the background (g->bb) to Higgs searches (h->bb) where the Higgs is a color singlet and the gluon is an octet:



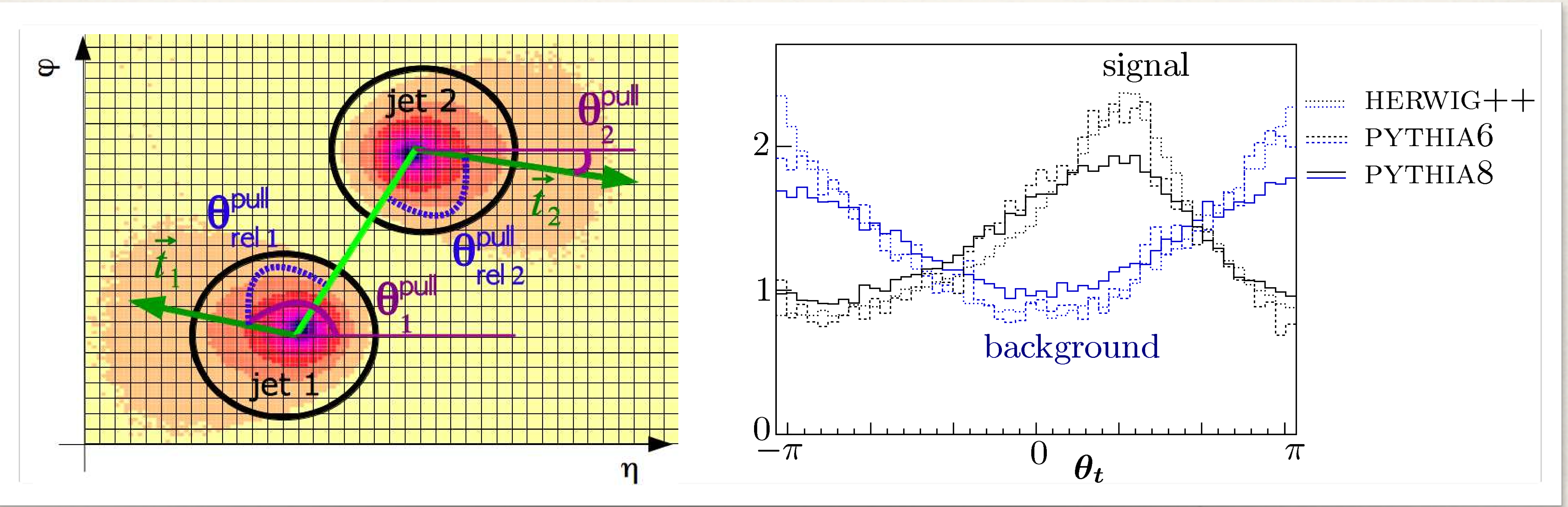
- There are many possible further applications - reducing combinatorics is the most prominent.

Figure source: Seeing in Color: Jet Superstructure J. Gallicchio and M. D. Schwartz [1001.5027]

- The technique has already been applied to data
  - *Measurement of color flow in ttbar events from ppbar collisions at $\sqrt{s} = 1.96$ TeV* [1101.0648]
- They essentially measure whether the radiation points from one jet to the other, or back to the beam.

- Assign a direction to each jet:

$$\vec{t} = \sum_{i \in \text{jet}} \frac{p_T^i \, |r_i|}{p_T^{\text{jet}}} \, \vec{r}_i$$



- “We find that the fraction of uncolored W bosons is 0.56 ± 0.42(stat+syst), in agreement with the standard model.” - Phew!

Conclusions

- At the LHC, particles with EW scale masses are often so energetic (i.e. boosted) their decay products are lumped together into a single jet.
- By looking at the radiation inside jets we find we can identify these particles and learn about their properties (e.g. polarization).
    - This leads to a marked improvement in discover reach.
- Techniques already undergoing validation with LHC data. Promising results!
- See the proceedings of Boost 2010 [1012.5412] for more discussion.

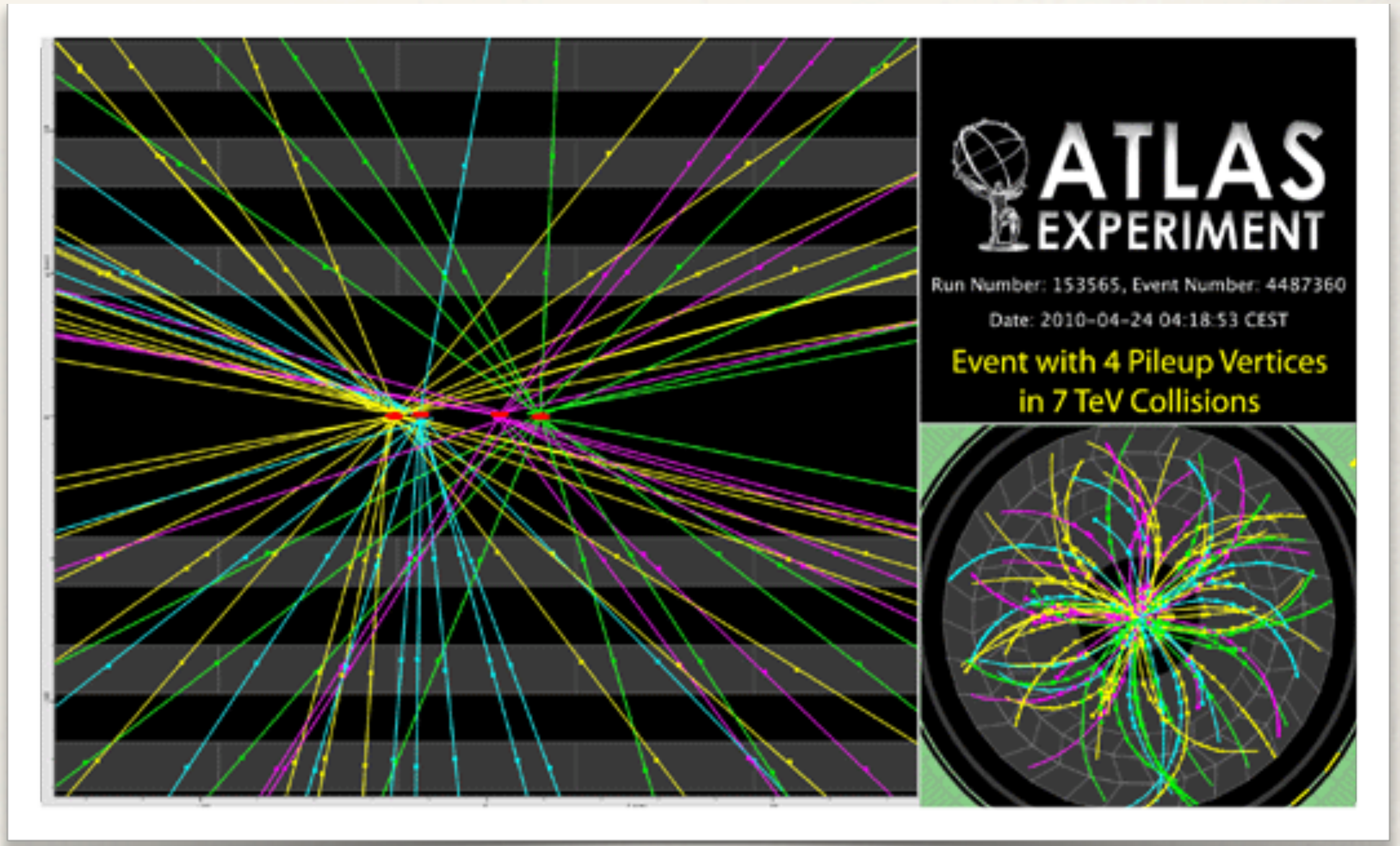
http://boost2011.org

Backup: Jet Trimming

Source: DK, J. Thaler, and L. Wang, JHEP 1002 (2010) 084 [arXiv:0912.1342]

Messy Jets

- The LHC is a messy place.
- Contaminating radiation can always come from ISR and multiple interactions, but at the LHC a major source is pileup.
  - Pileup is when multiple scatterings take place in the same bunch-bunch crossing



Quantifying Contamination

- How much contamination is there?

- Pileup density in GeV/area:

$$\rho \sim \left(1 + \frac{N_{\rm PU}}{4}\right) \times (2 \leftrightarrow 3 \ {\rm GeV})$$

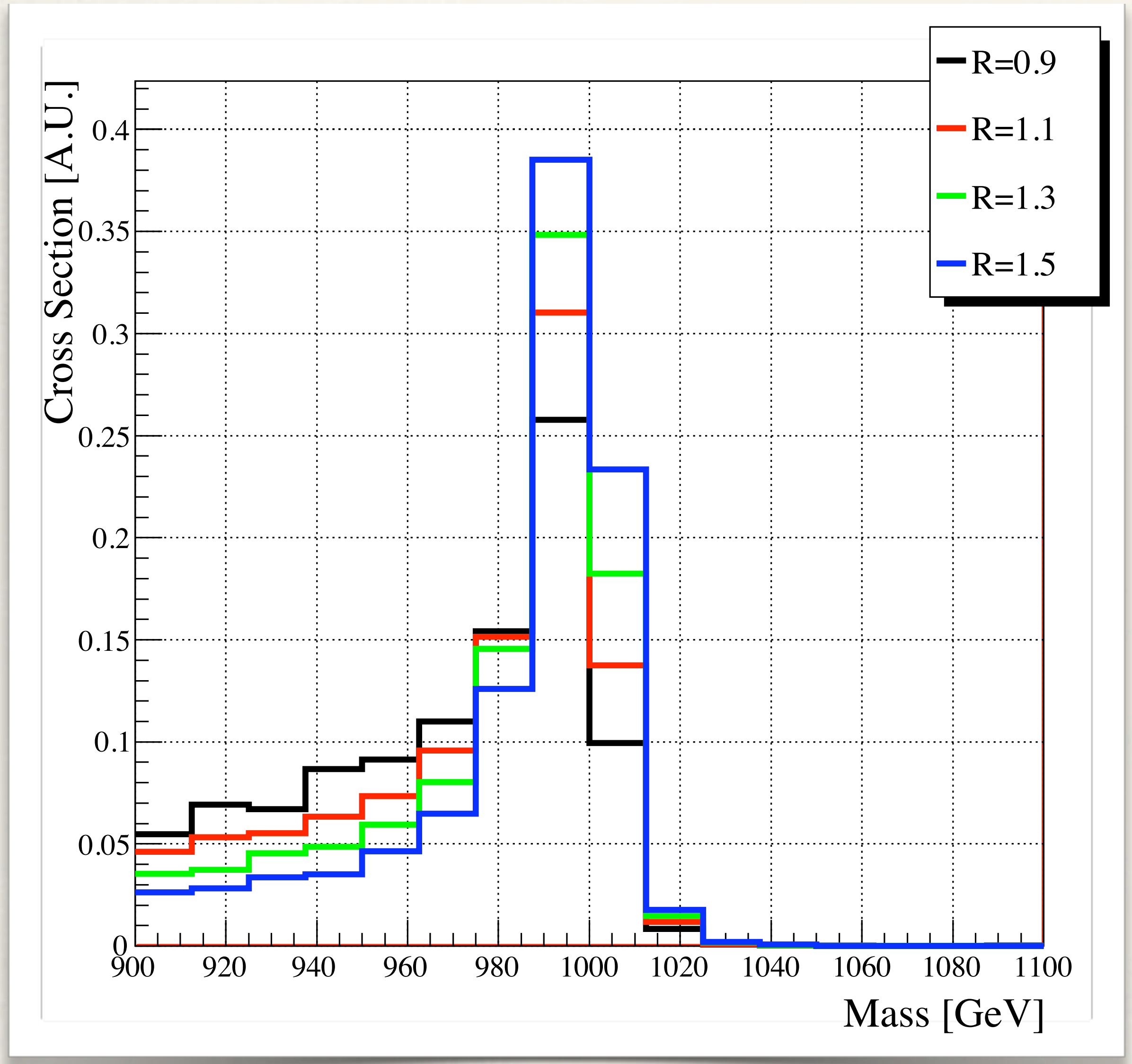
- The number of pileup events per crossing (N_{PU}) depends on the LHC running parameters. Roughly though, at 14 TeV we should start at ~20 and go to ~40.

Source: 1002.4557

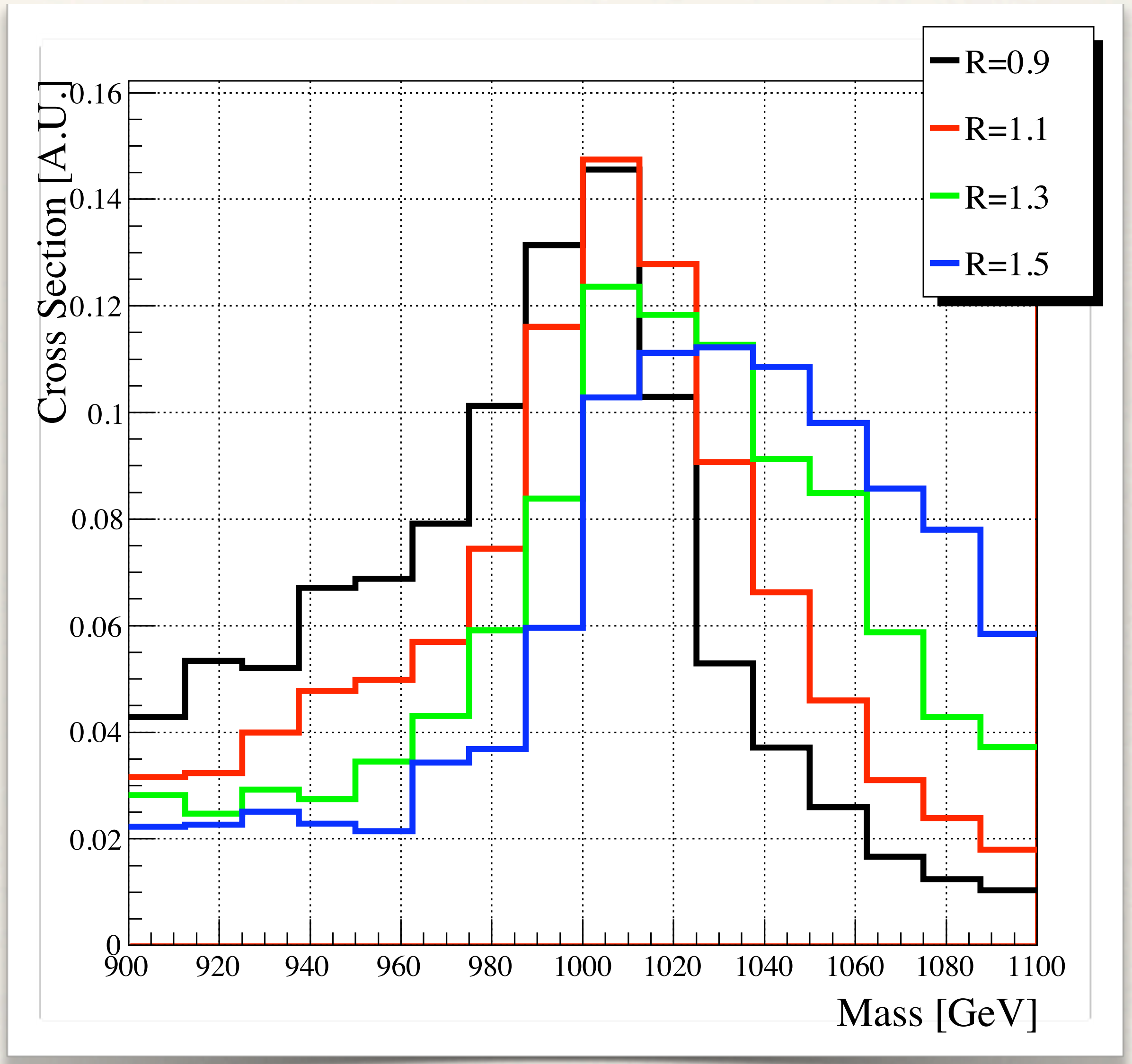
Motivation

- When we cluster jets there's inevitably a tradeoff:
  - Larger cones are less likely to miss radiation
  - But, they're also more susceptible to contamination

Contamination in Resonance Reconstruction



Contamination Off

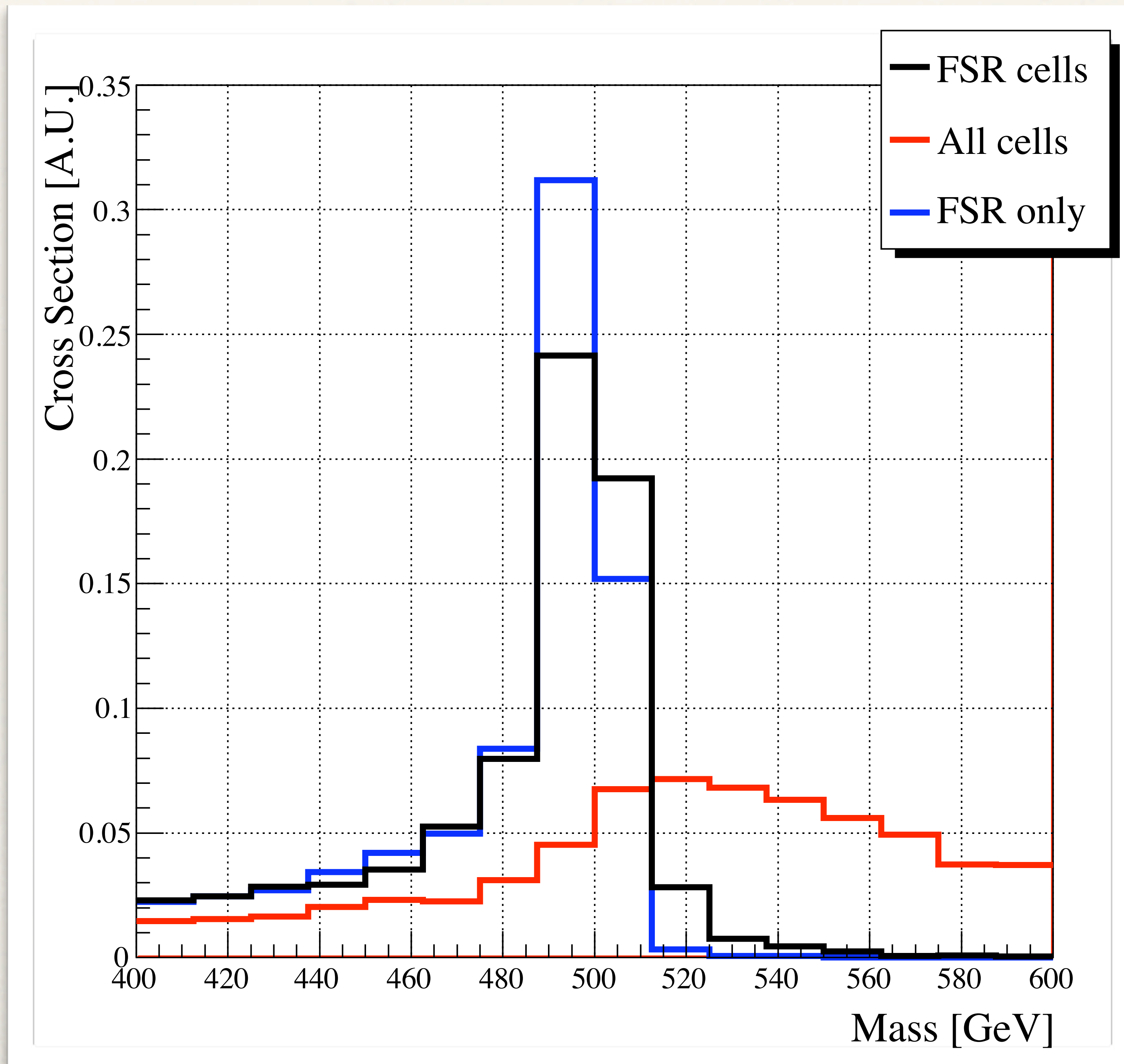


Contamination On

- In *Jet Trimming* we investigated ways to systematically remove jet contamination and improve reconstruction.
- There's a lot of room for reconstruction improvement.
- Irreducible contamination (we can't distinguish radiation in the same cell) is not a problem

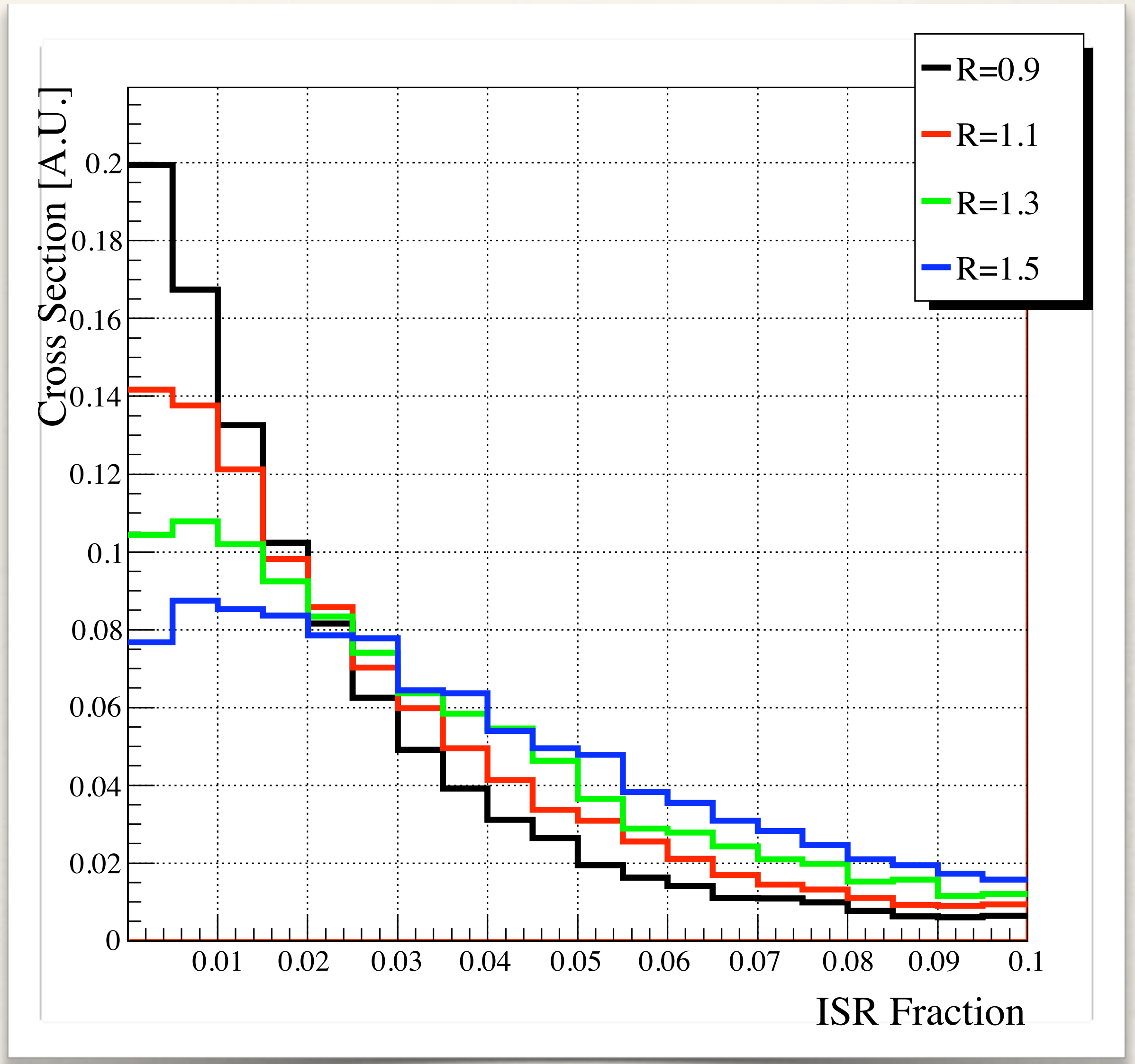
	Improvement	R_0	Γ [GeV]	M [GeV]
	$gg \to \phi \to gg$			
All cells	-	1.2	69	518
FSR cells	309%	1.5	15	501
	$q\bar{q} \to \phi \to q\bar{q}$			
All cells	-	0.8	31	505
FSR cells	189%	1.5	11	501

- If we knew what cells contained significant FSR, then we'd be able to remove everything else and nearly reproduce the distribution without contamination:



Trimming in Practice

- Contamination is usually quite soft (total ~5% of p_T).
- Use this to our advantage by only keeping the hard parts of a jet.



Implementation

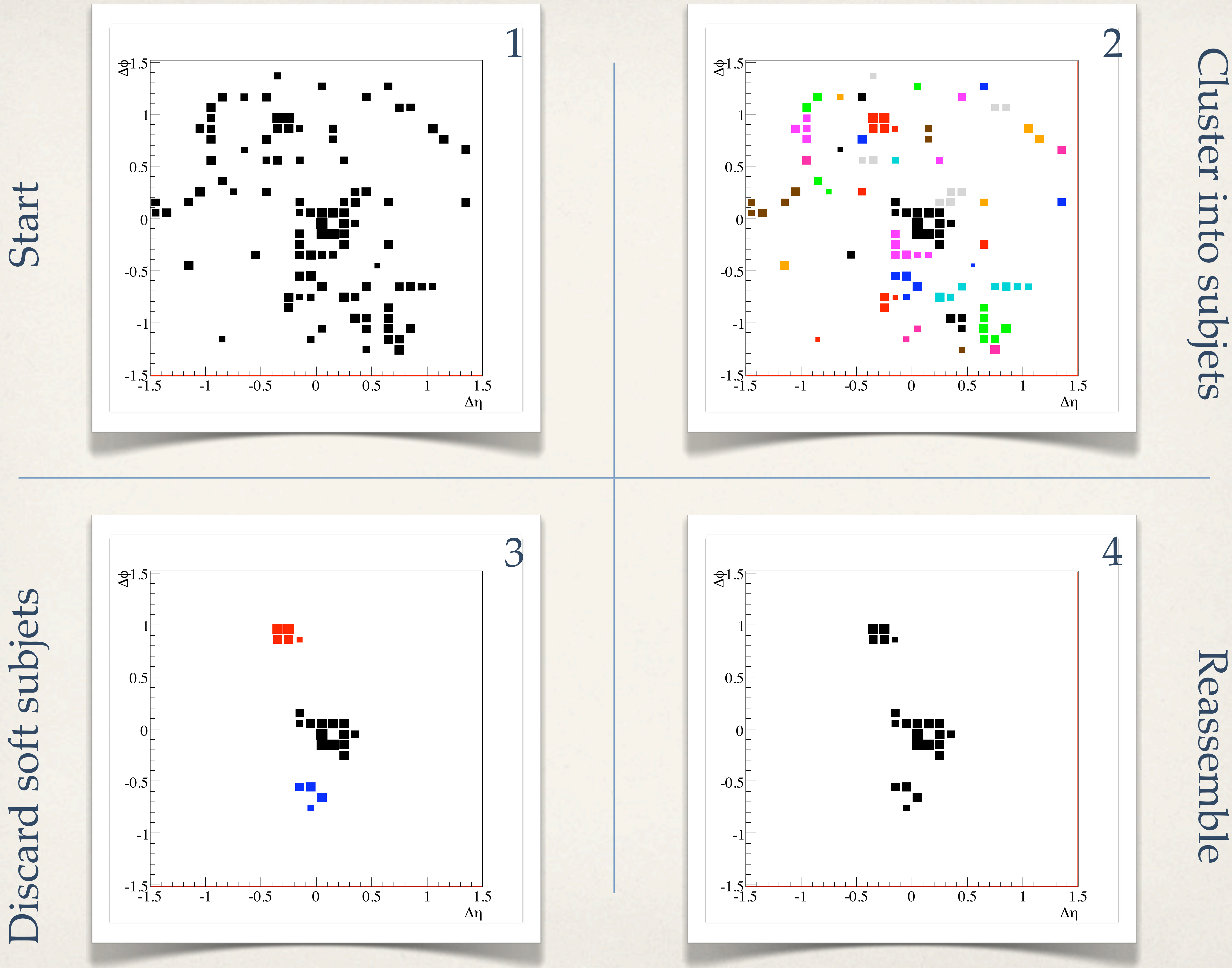
1. Cluster all calorimeter data using any algorithm

2. Take the constituents of each jet and recluster them using another, possibly different, algorithm (we advocate k_T) with smaller radius R_{sub} ($R_{sub} = 0.2$ seems to work well).

3. Discard the subjet i if

$$p_{Ti} < f_{\mathrm{cut}} \cdot \Lambda_{\mathrm{hard}}$$

4. Reassemble the remaining subjets into the trimmed jet

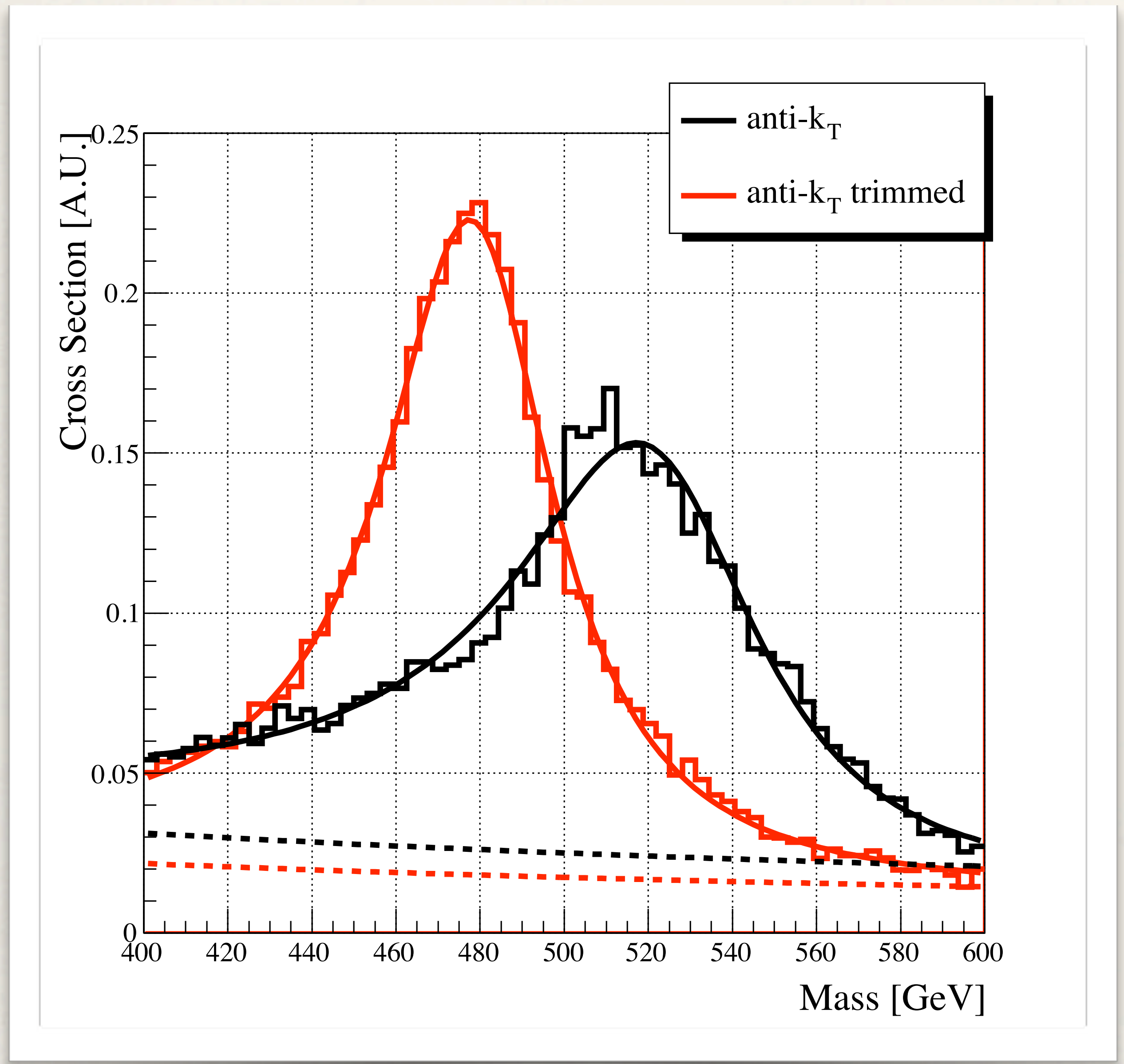

1
Start
2
Cluster into subjets
3
Discard soft subjets
4
Reassemble
Δφ
Δη
1.5
1
0.5
0
-0.5
-1
-1.5

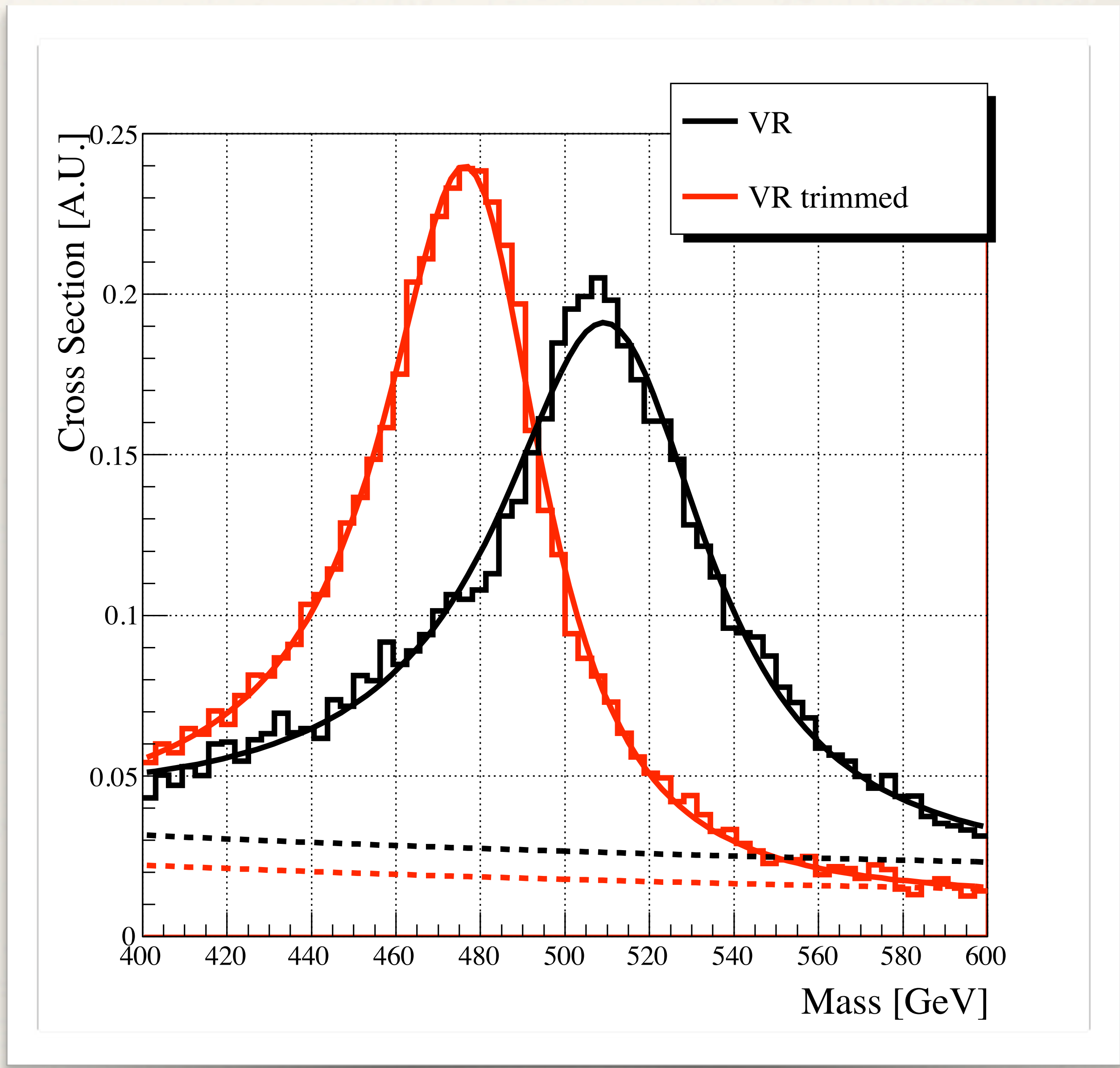
Results

- Find a significant improvement from using trimming to reconstruct a resonance decaying to dijets ($gg \to \phi \to gg$)

	Improvement	$f_{\rm cut}$, $N_{\rm cut}$	$R_{\rm sub}$	R_0, ρ	Γ [GeV]	M [GeV]
anti-k_T	-	-	-	1.0^*	71	522
anti-k_T (N)	40%	5^*	0.2^*	1.5^*	62	499
anti-k_T (f, p_T)	59%	$3 \times 10^{-2*}$	0.2	1.5	52	475
anti-k_T (f, H)	61%	$1 \times 10^{-2*}$	0.2	1.5	50	478
VR	30%	-	-	200^* GeV	62	511
VR (N)	53%	5	0.2	275^* GeV	53	498
VR (f, p_T)	68%	3×10^{-2}	0.2	300^* GeV	49	475
VR (f, H)	73%	1×10^{-2}	0.2	300^* GeV	47	478
Filtering	27%	2	$R_0/2$	1.3^*	61	515

All histograms (those with and without trimming) are made using optimized parameters.





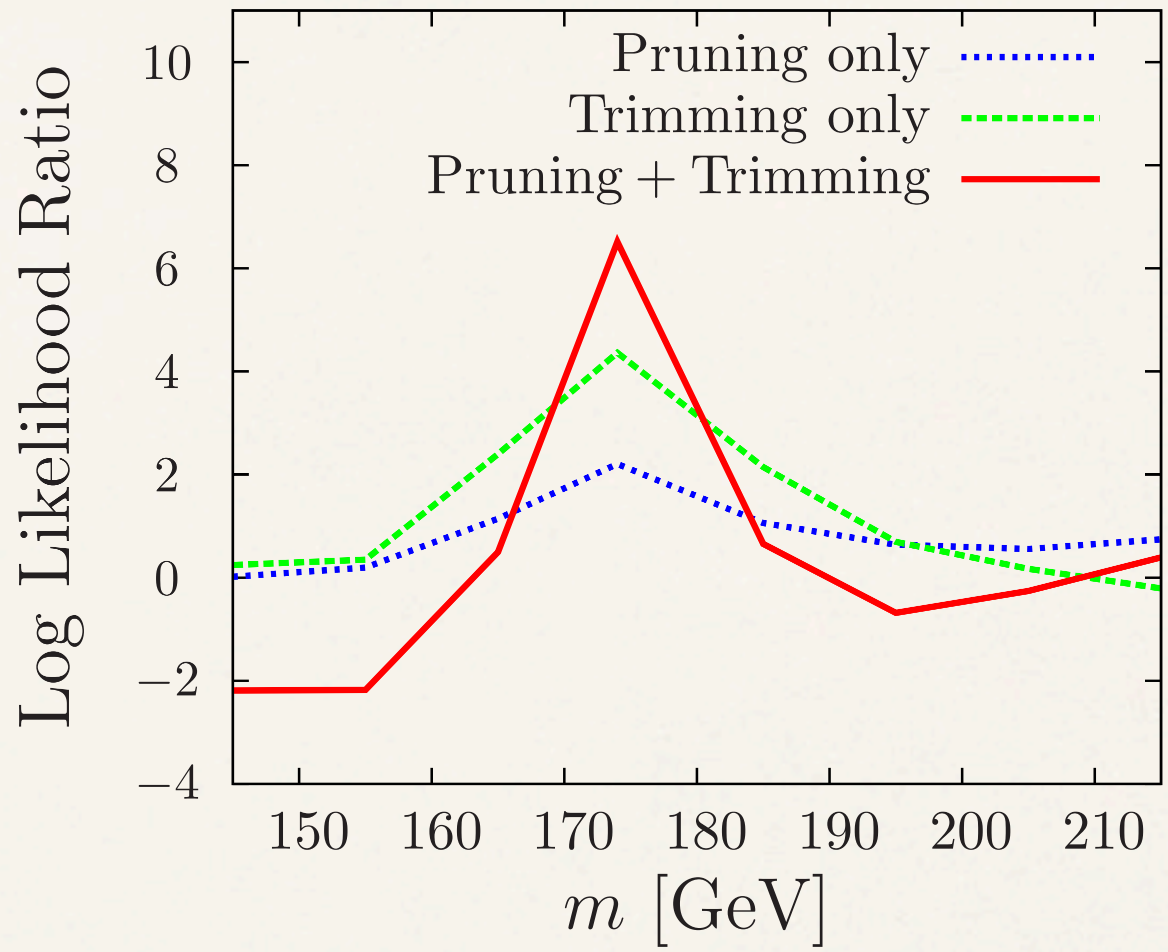
Jet Topiary

- Trimming was designed to clean up boosted “QCD Jets”. There are other approaches focused on cleaning up jets from boosted heavy objects

1. Jet Pruning (Ellis, Vermilion, Walsh): 0903.5081, 0912.0033

2. Filtering (Butterworth, Davison, Rubin, Salam): 0802.2470

- If you combine multiple algorithms together and statistically optimized you can do better:



- Still a lot of room for improvement!

Figure source: Combining subjet algorithms to enhance ZH detection at the LHC, D. E. Soper, M. Spannowsky, [arXiv:1005.0417]