New jet techniques for at the LHC

Lian-Tao Wang
Princeton University

Early LHC, TRIUMF, January 2011

The importance of hadronic final state:

- "Everywhere" at hadron colliders. $p\ p, \ \mathrm{or}, \ p\bar{p}$ initial state.
- Present in (almost) all new physics signals.
    - Many of them only have hadronic channels.
    - TeV new physics states can decay to SM "heavy" particles, e.g. t, W, Z, often look like a cluster of hadrons.
- Understanding of basic structure of QCD and the properties of new physics has lead to the development of a set of modern tools which significantly enhanced the discovery potential.

Cover two aspects.

- Better characterization of QCD jets
  - Improving jet algorithms.
  - Finding ISR.
- Jet substructure and new physics searches.
  - Boosted tops.
  - Higgs search.

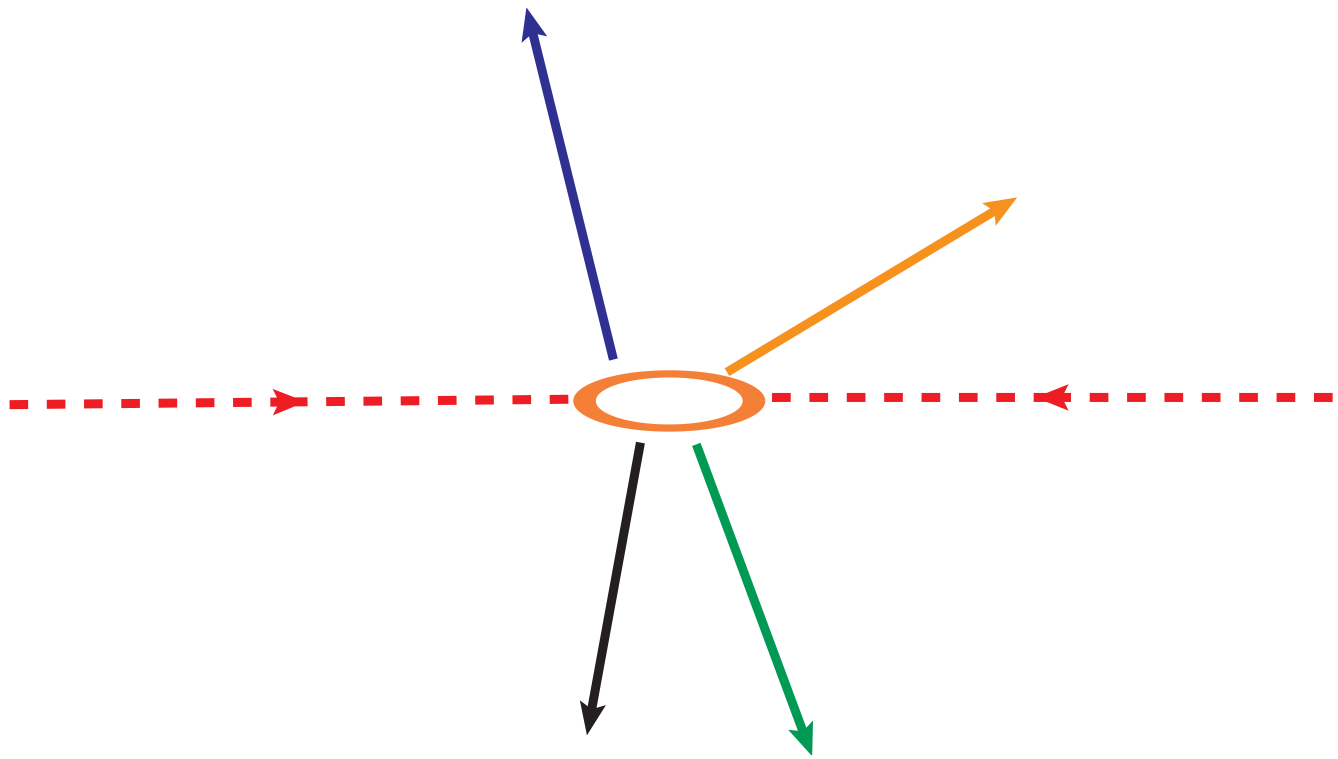
Boston Jet Workshop:
http://jets.physics.harvard.edu/workshop/Main.html

Boost 2011, May, 23-27, Princeton.
http://boost2011.org

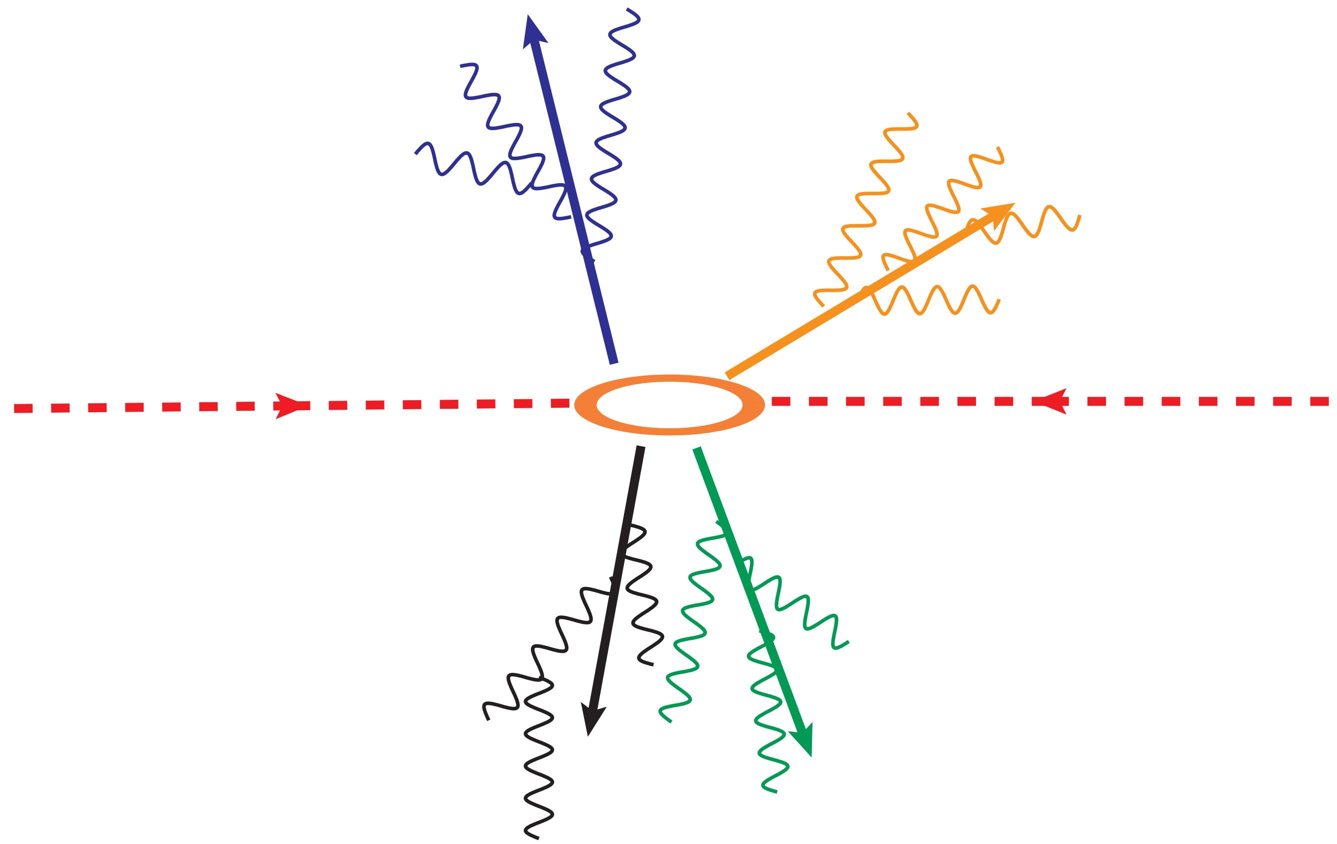
Better QCD jet

Why is it hard?

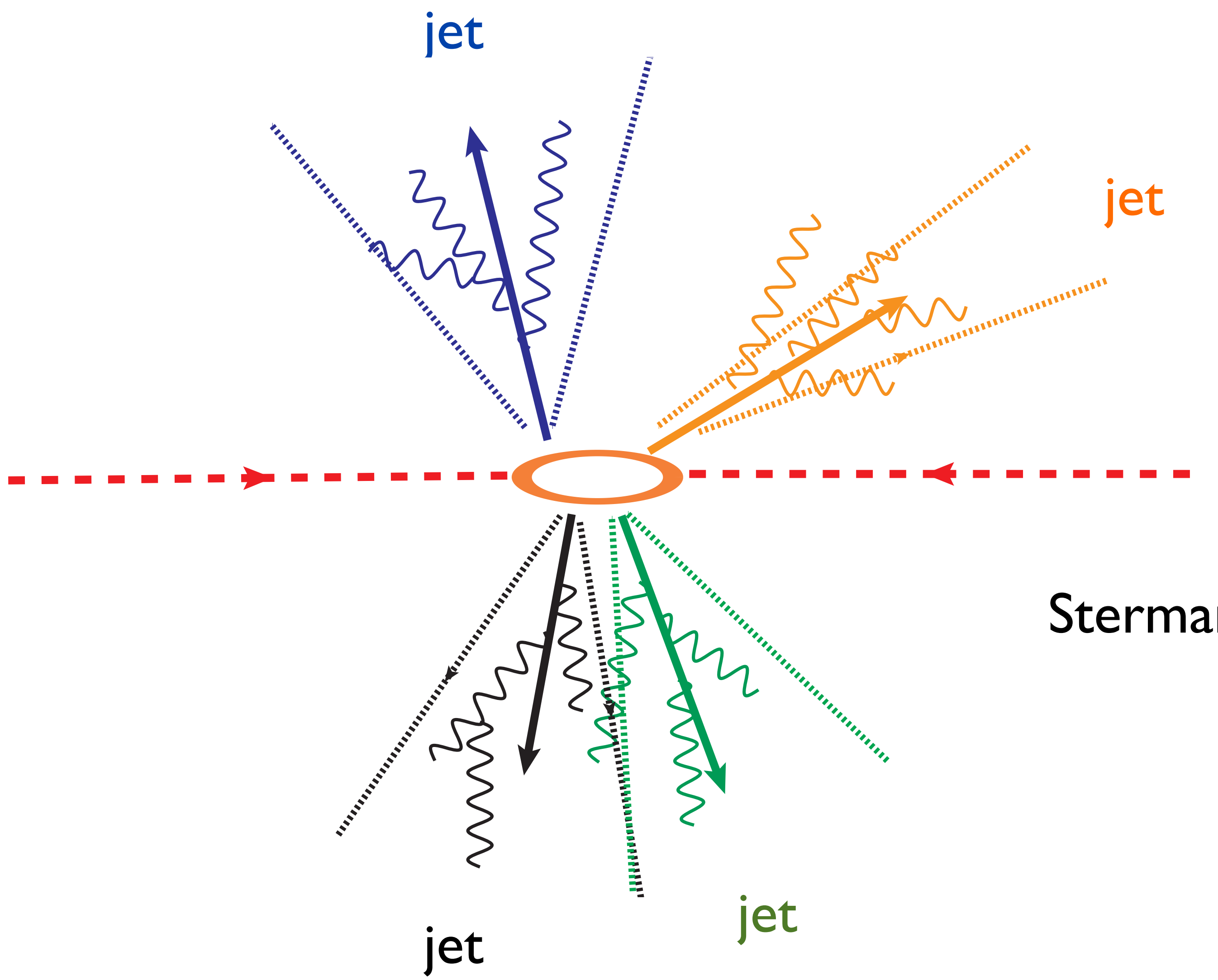
Why is it hard?



Why is it hard?



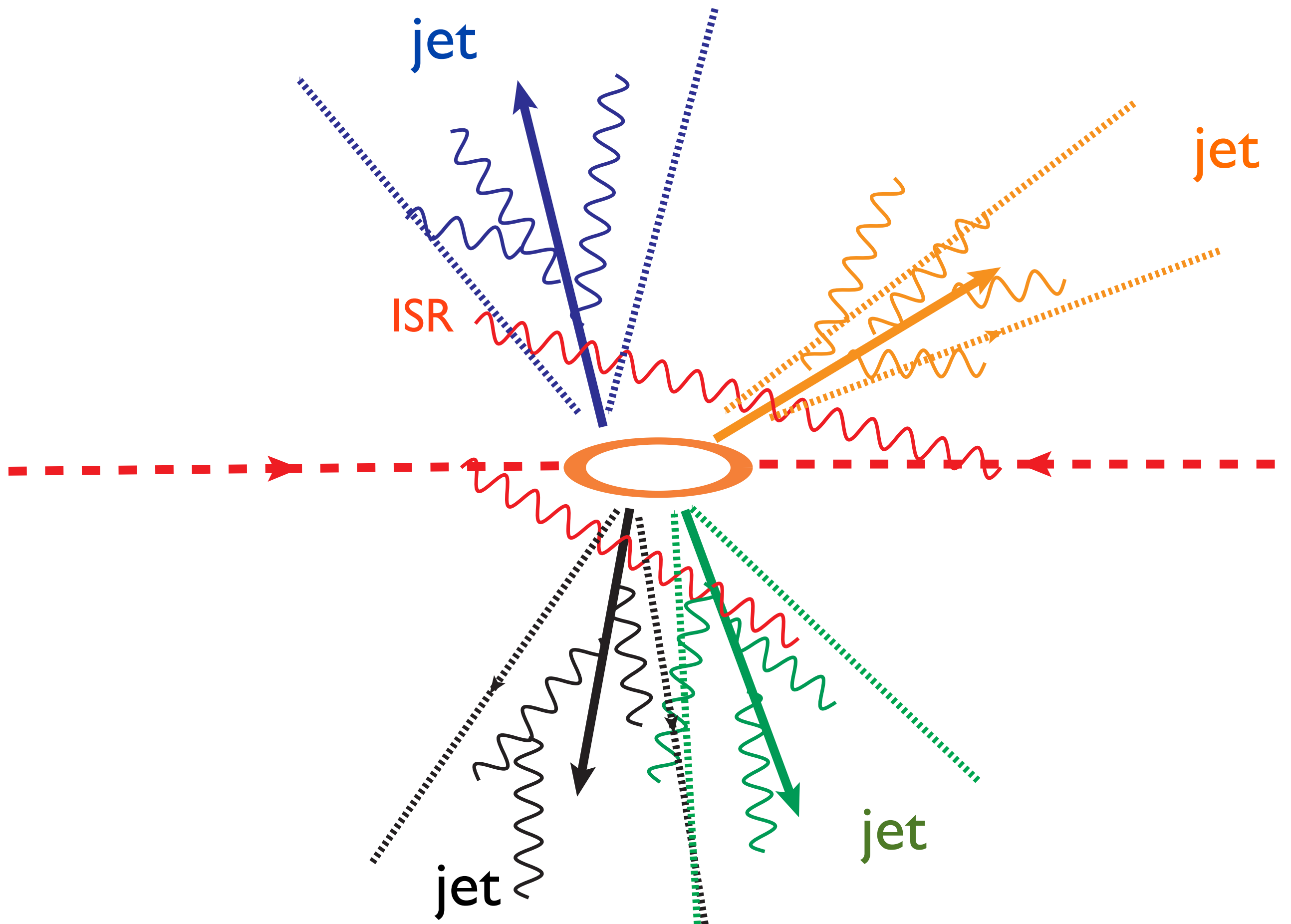
Why is it hard?



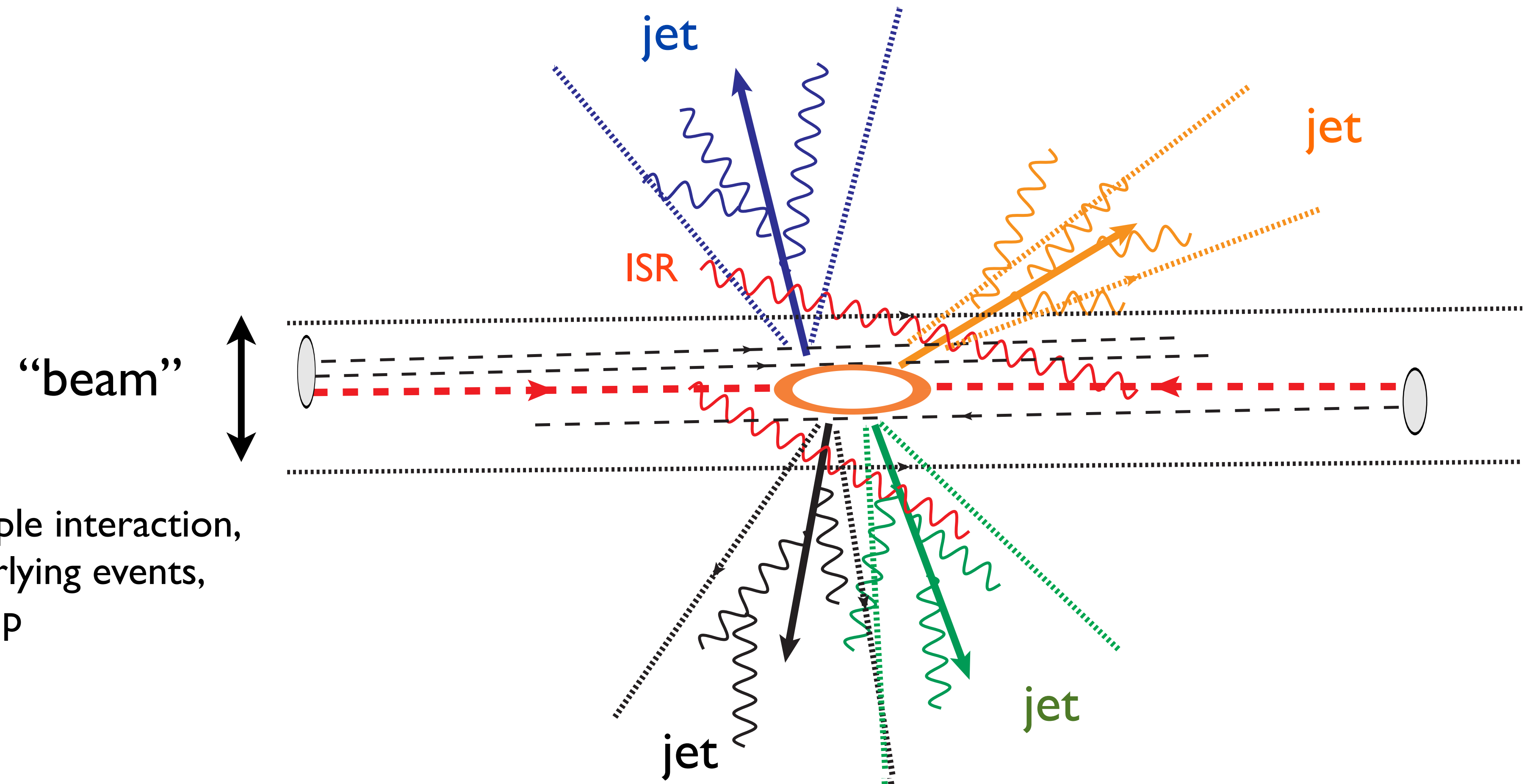
Sterman & Weinberg, PRL, 1977.

- We would like to preserve $p_{\rm jet} \simeq p_{\rm parton}$.

Why is it hard?

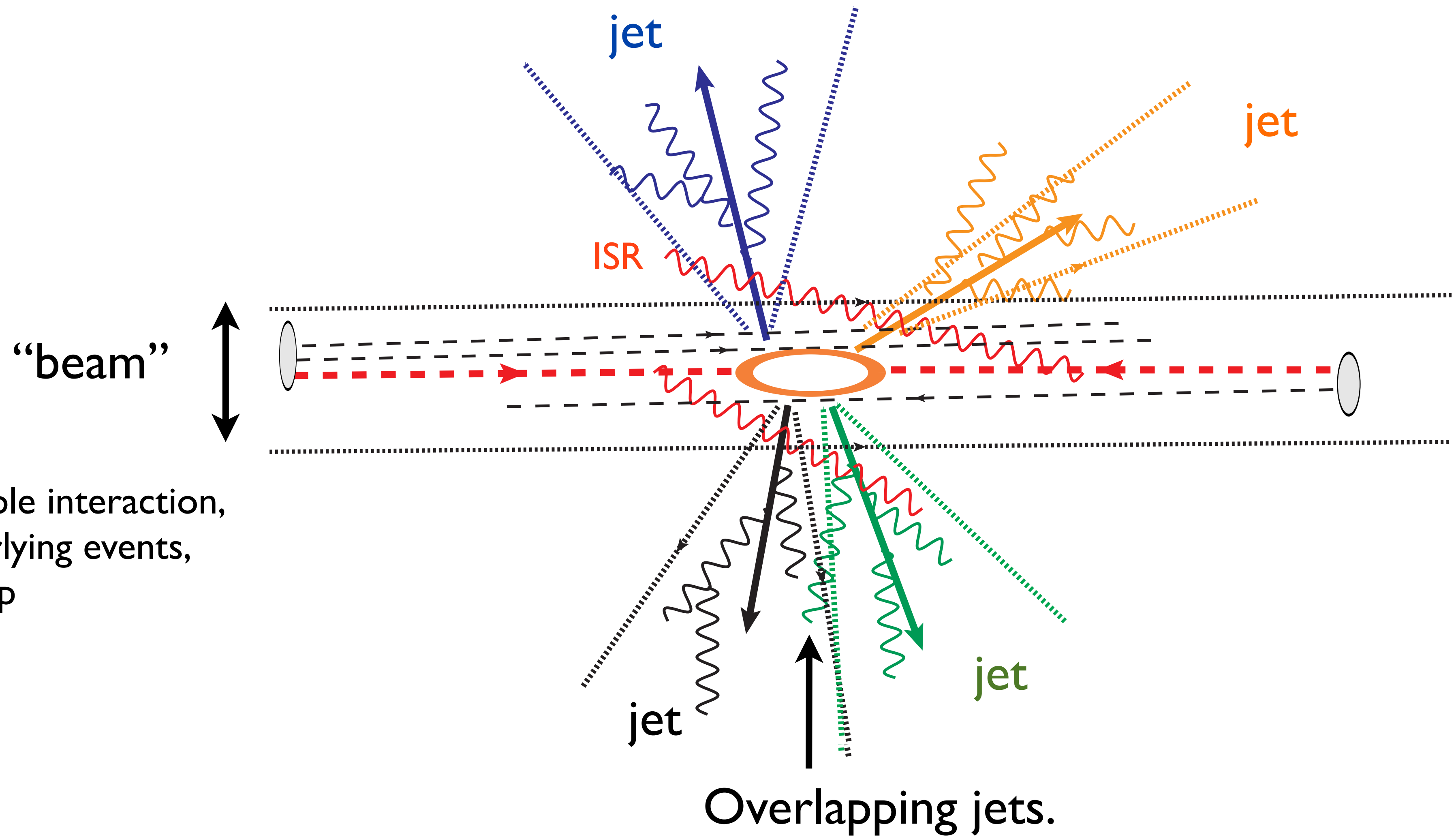


Why is it hard?



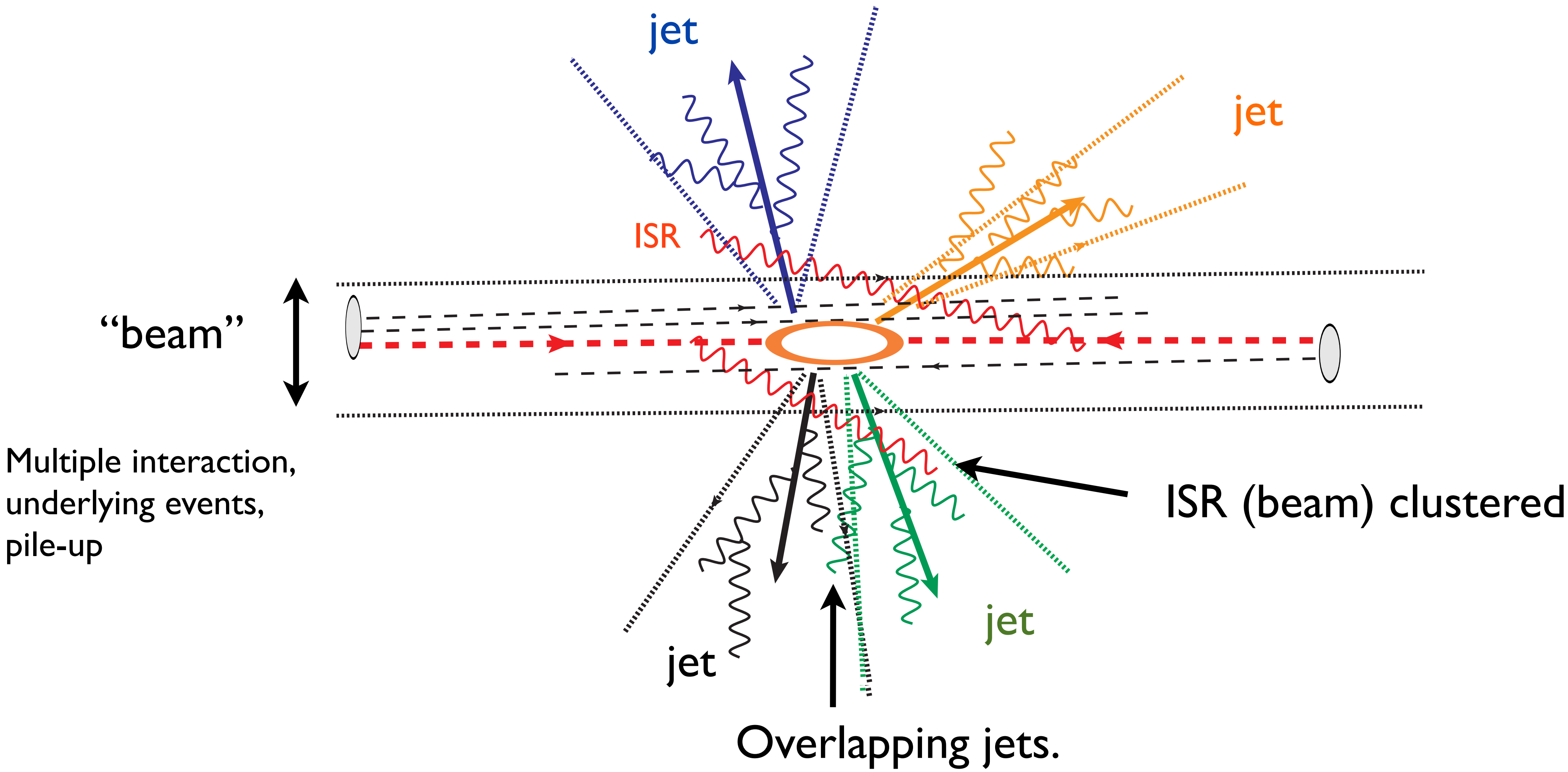
Multiple interaction, underlying events, pile-up

Why is it hard?

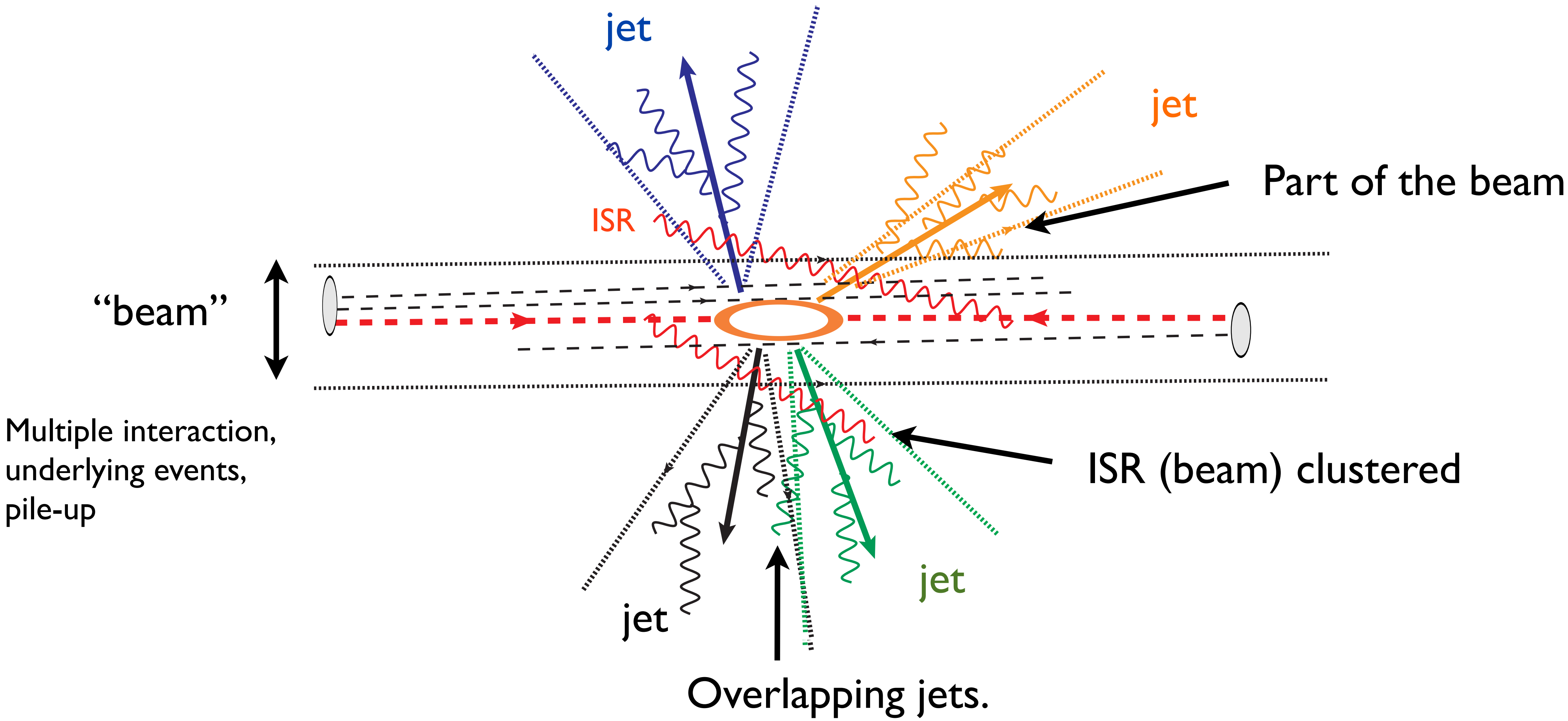


Multiple interaction,
underlying events,
pile-up

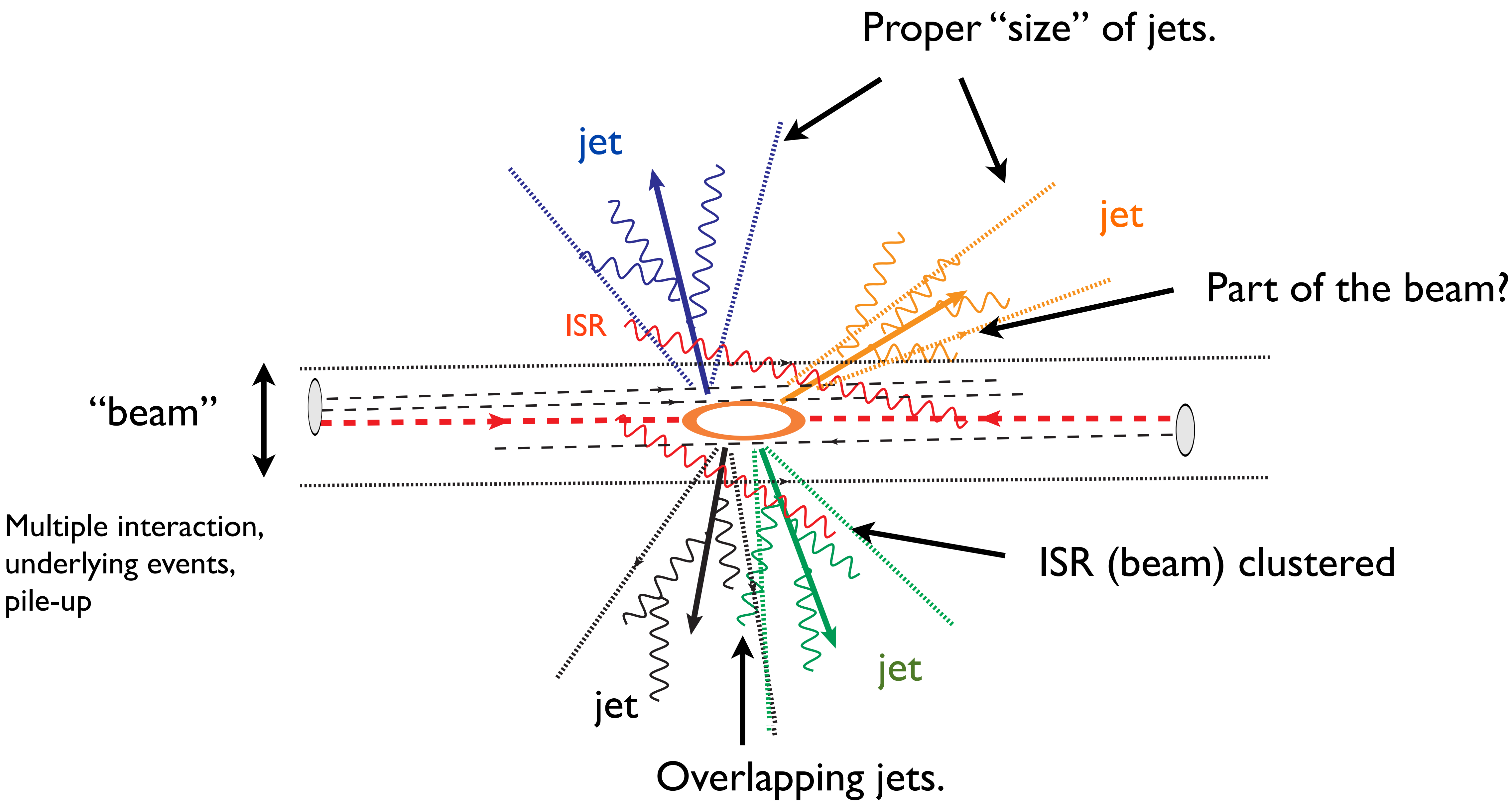
Why is it hard?



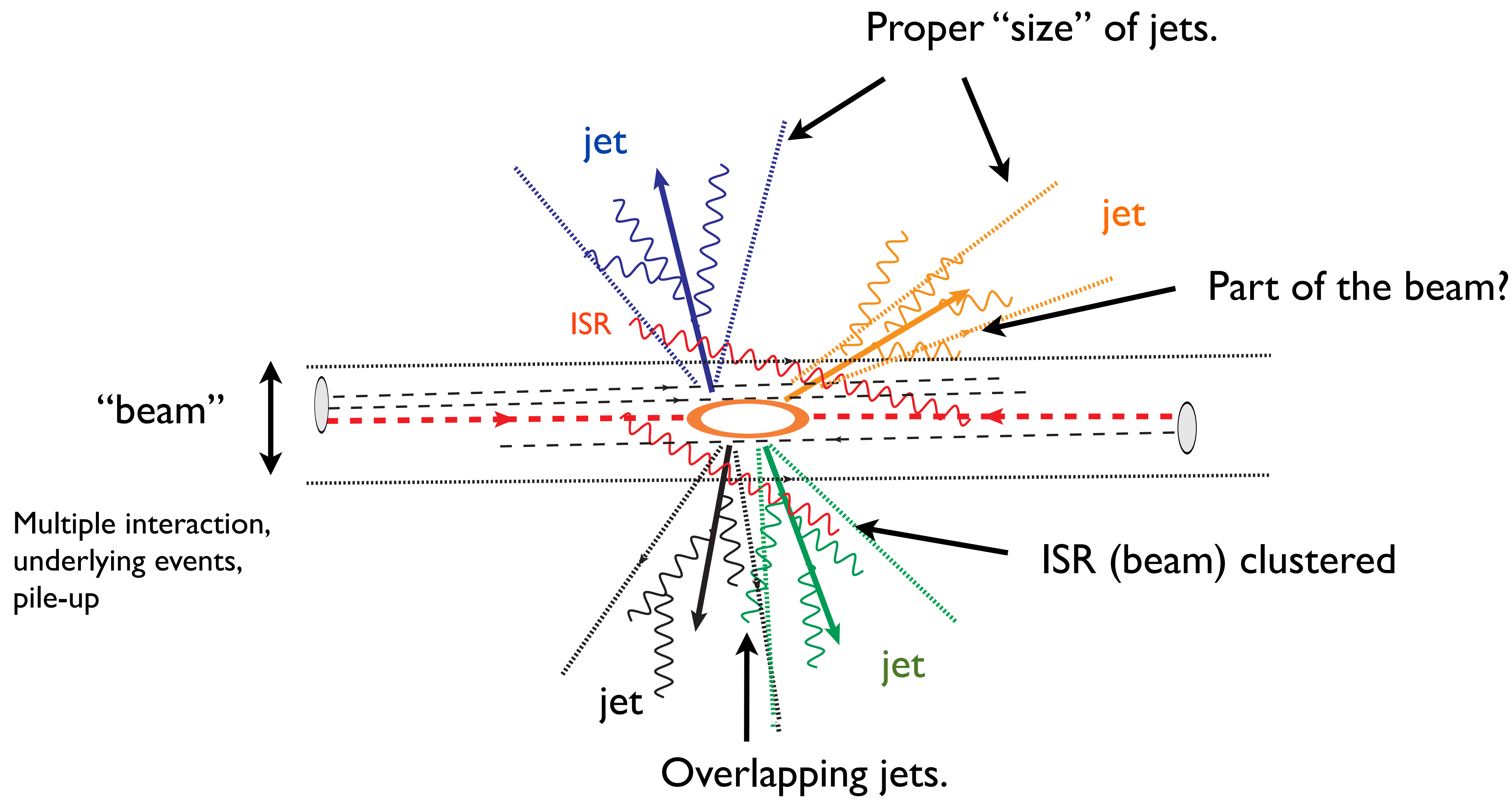
Why is it hard?



Why is it hard?



Why is it hard?



- To best preserve $p_{\rm jet} \simeq p_{\rm parton}$ we would like to:
  - Use "smart" jet shapes.
  - Control "contamination".

Begin with jet algorithm

- An algorithm of clustering together "close by" objects.
- Basic ingredients of a "sequential" jet algorithm.
- Two types of "distances"
  - Jet-jet distance: d_{ij} "when to cluster"
  - Jet-beam distance: d_{iB} "when to stop clustering"
- Pair wise comparison of all distances
  - If smallest distance at any stage in clustering is jet-jet, add together corresponding four-momenta, else take jet with smallest jet-beam distance and set it aside.
  - Repeat till all jets are set aside.

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

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

$$(\Delta R)^2 \equiv (\Delta\eta)^2 + (\Delta\phi)^2$$

Recombination Algorithms

- k_T algorithm

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

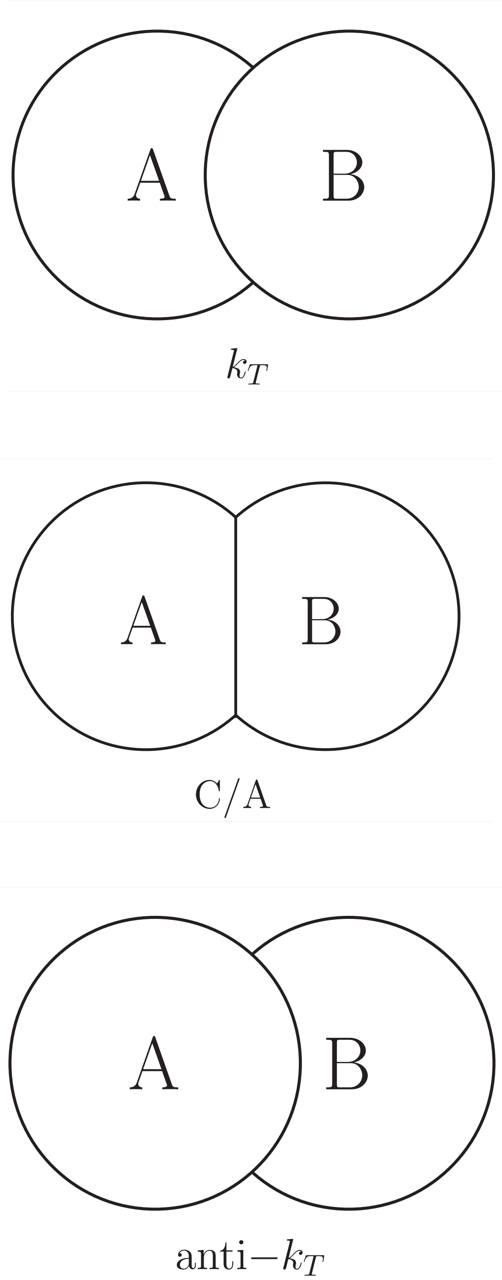
- C/A algorithm

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

- anti-k_T algorithm

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

$$(\Delta R)^2 \equiv (\Delta \eta)^2 + (\Delta \phi)^2$$

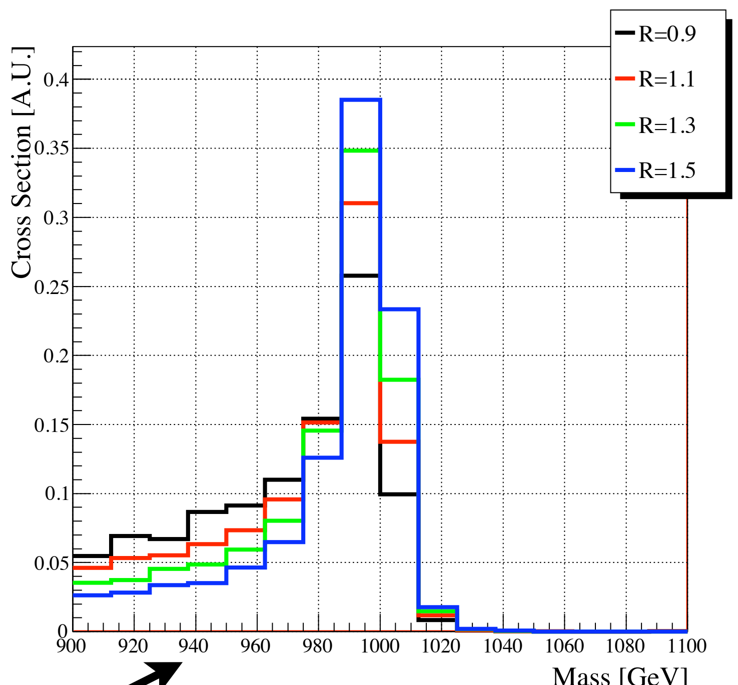


$$p_T^A > p_T^B$$

Jet "trimming"

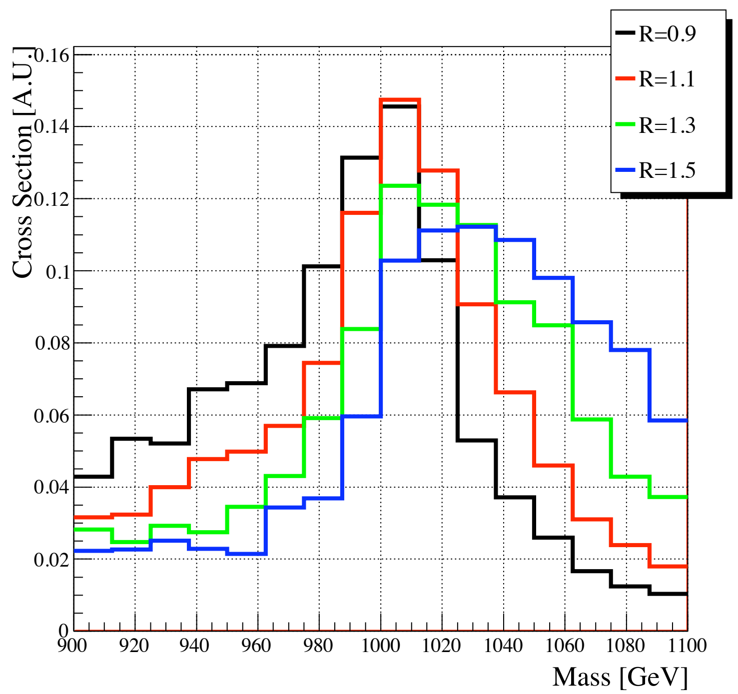
- Effect of the "contamination".
  - Initial state radiation (ISR), multiple interaction (MI), underlying events (UE), pile-up (PU).

$$gg \to \phi \to gg \qquad m_\phi = 1\ \mathrm{TeV}$$



$m^2 = (p_j^1 + p_j^2)^2 \simeq m_\phi^2$

FSR only
No contamination

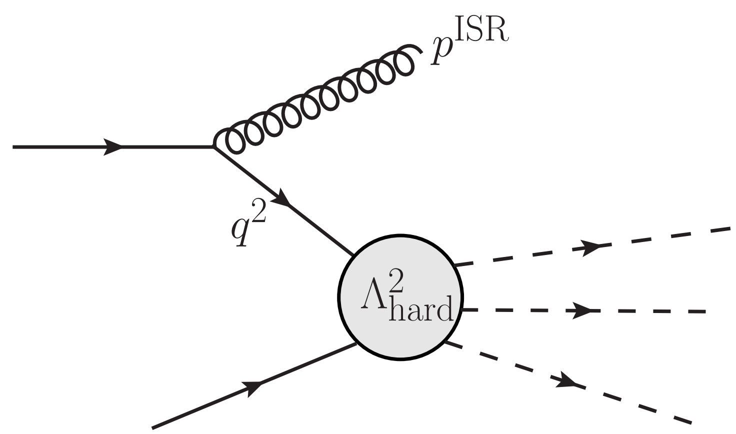


Including ISR, MI, UE, pile-up

Room for improvement!

A closer look at the soft radiations

- ISR scale with the hard collision



$$(p_T^{\mathrm{ISR}})^2_{\max} \simeq |q|^2 \leq \Lambda^2_{\mathrm{hard}}$$

$$\Lambda_{\mathrm{hard}} : \quad \text{hard interaction scale}$$

$$\frac{d\sigma}{d(p_{\mathrm{T}}^{\mathrm{ISR}})^2} \simeq \frac{1}{(p_{\mathrm{T}}^{\mathrm{ISR}})^2} \left(\alpha_s \log\left(\frac{\Lambda^2_{\mathrm{hard}}}{(p_{\mathrm{T}}^{\mathrm{ISR}})^2} \right) + O(\alpha_s^2) \right)$$

$$\langle p_{\mathrm{T}}^{\mathrm{ISR}} \rangle \propto (p_T^{\mathrm{ISR}})_{\max}$$

- MI, UE, and pileup “incoherent”, independent of the hard collision scale.
  - A “universal” soft background. $\delta(p_{\mathrm{T}}^j) \simeq \Lambda_{\mathrm{soft}} \left(\frac{\Delta R^2}{2} + ... \right)$

Jet trimming.

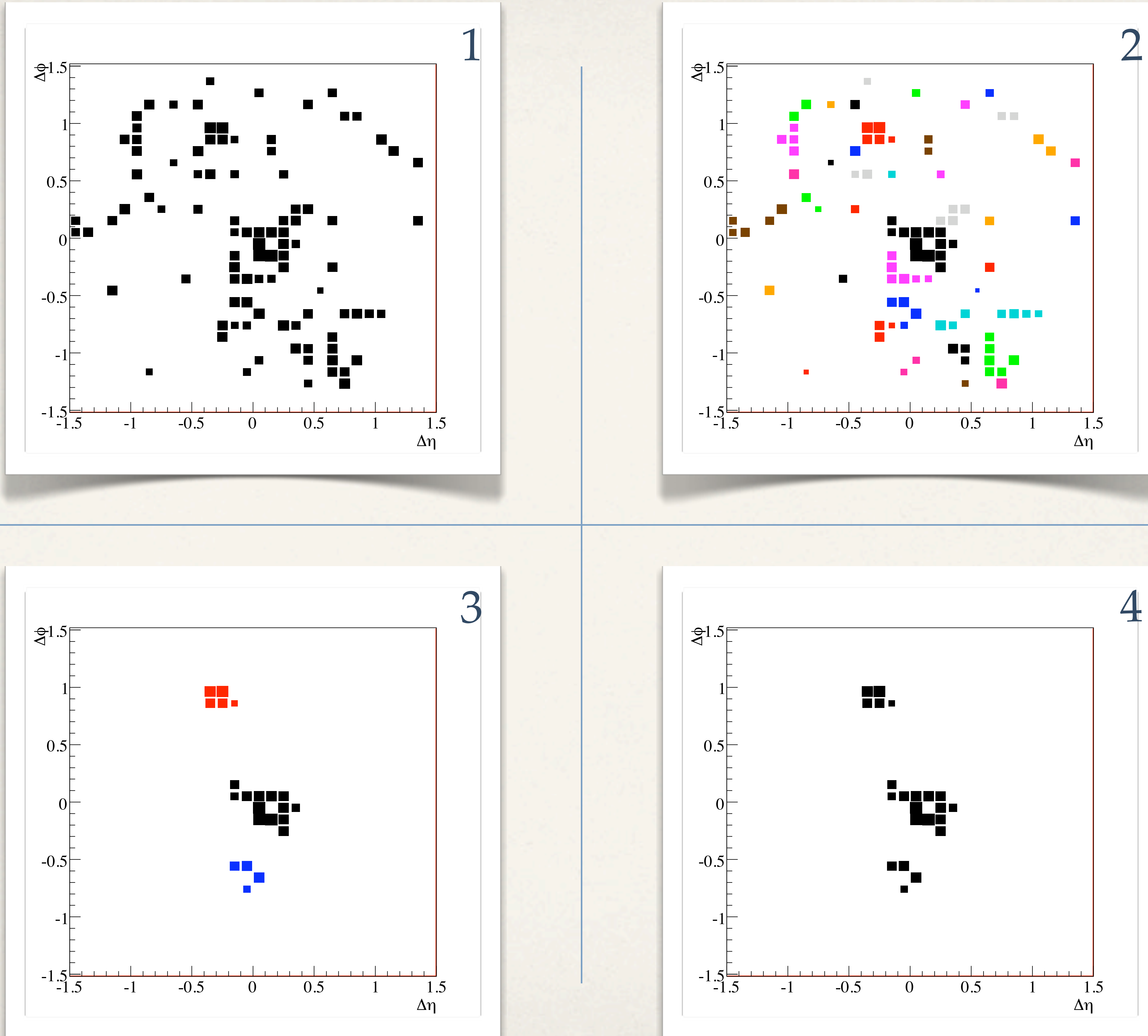
D. Krohn, J. Thaler, LTW, arXiv:0912.1342

- Introducing a "cut" on soft radiation.
  - Discard "stuff" below the cut after jet clustering.
- Our implementation.
  - Cluster all calorimeter data using any algorithm
  - Take the constituents of each jet and recluster with smaller radius Rsub (Rsub = 0.2 seems to work well).
  - Discard the subjet i if $p_{Ti} < f_{\rm cut} \cdot \Lambda_{\rm hard}$ ⟵ ISR argument.
- Best choice of the hard scattering scale and fcut.
  - Process dependent.
  - Can be optimized experimentally.

Related but different "jet grooming" approaches:
Filtering: J. Butterworth, A. Davison, M. Rubin, G. Salam, arXiv:0802.2470
Pruning: S. Ellis, C. Vermilion, J. Walsh, arXiv:0903.5081

Start

Cluster into subjets

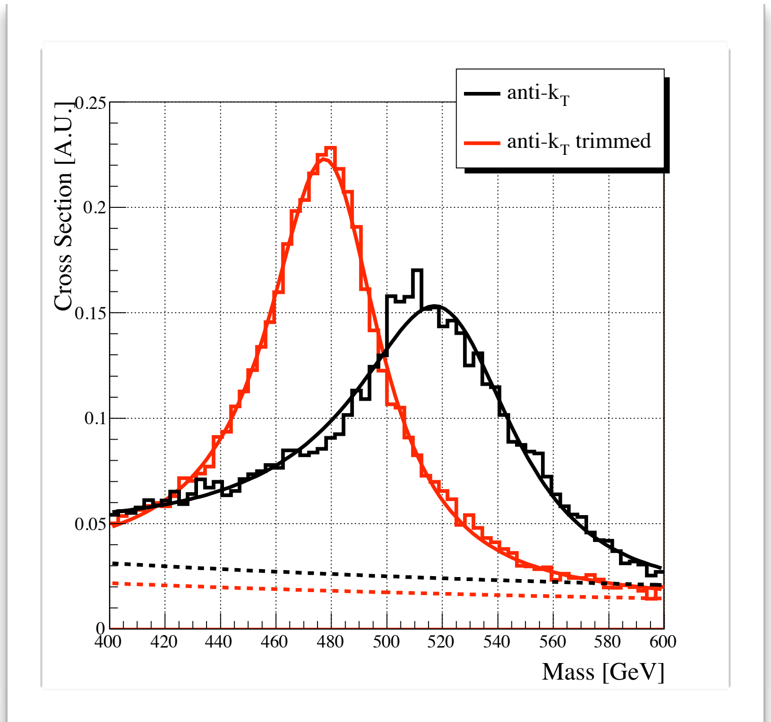


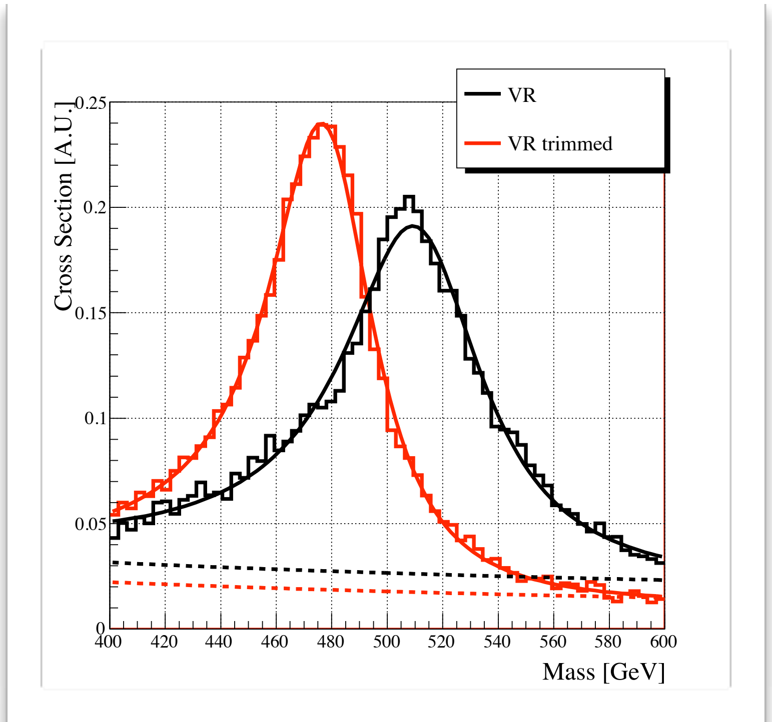
Discard soft subjets

Reassemble

Simple test case: di-jet resonance

$$gg \rightarrow \phi \rightarrow gg$$





	Improvement	f_{cut}, N_{cut}	R_{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

- We provide plugins fully compatible with Fastjet.

  http://jthaler.net/jets/VR_Jets.html

  http://jthaler.net/jets/Jet_Trimming.html

Finding ISR jet.

D. Krohn, L. Randall, LTW, arXiv:1108.0810

- Looking for the different jet.

- **Tag**
  - Take three hardest jets. Look for those

    1. Distinguished in pT

OR

    2. Distinguished in rapidity

OR

    3. Distinguished in m/pT

- **Check**
  - Require the candidate ISR jet

    1. Not be central

AND

    2. Remain somewhat isolated in rapidity

  - And, require that the implicit FSR jets be

    1. Close in pT

Simple kinematical tagger works well.

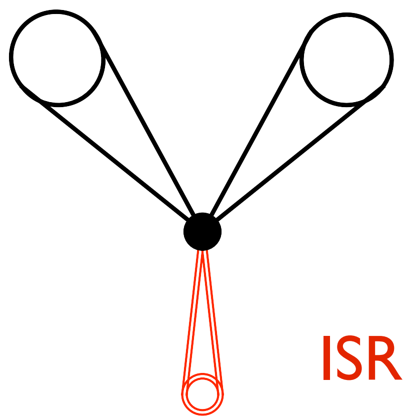
Spectrum		Efficiencies [%]	
$m_{\tilde{q}}/m_{\tilde{g}}$	m_{LSP}	Trigger	Mistag
500 GeV	100 GeV	42	15
500 GeV	450 GeV	42	12
1 TeV	100 GeV	41	11
1 TeV	950 GeV	41	9
500 GeV	100 GeV	13	22
500 GeV	400 GeV	15	10
1 TeV	100 GeV	12	25
1 TeV	900 GeV	16	8

- Further developments underway.
  - Asymmetric toplogy?
  - More inclusive?

Many potential applications.

- Reducing combinatorics.
  - SUSY decay chain, ttbar, ...
- ISR could be the main component of the signal.
  - Squeezed SUSY spectrum, ...
- Measuring mass
  - ISR spectrum is proportional to the scale of hard interaction.
  - Even more directly:

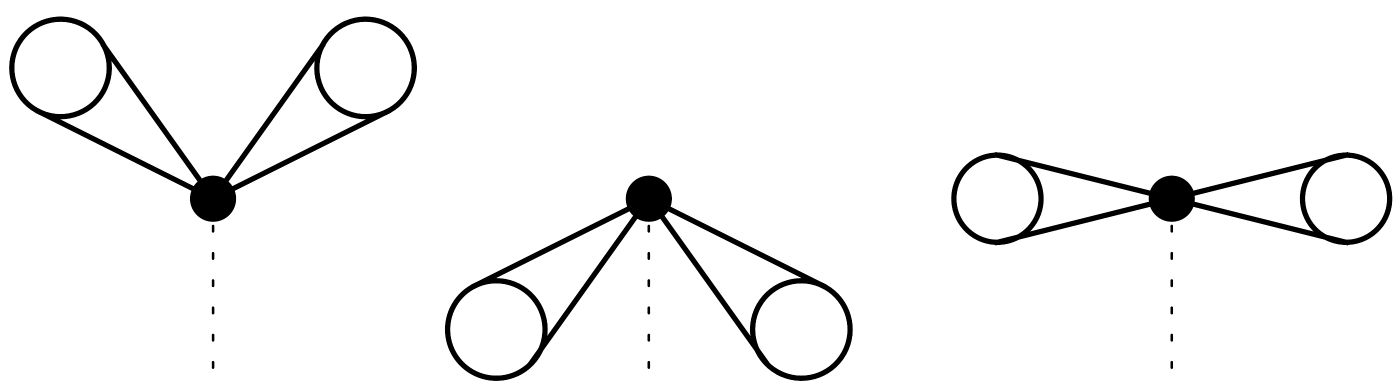
Mass measurement



A system with invariant mass M_0
including visible+invisible

For any FSR with $p_{\bar{T}_i} = \vec{p}_{T_i} \cdot \hat{p}_T^{\rm ISR}$, and assuming $M_{\rm test}$,

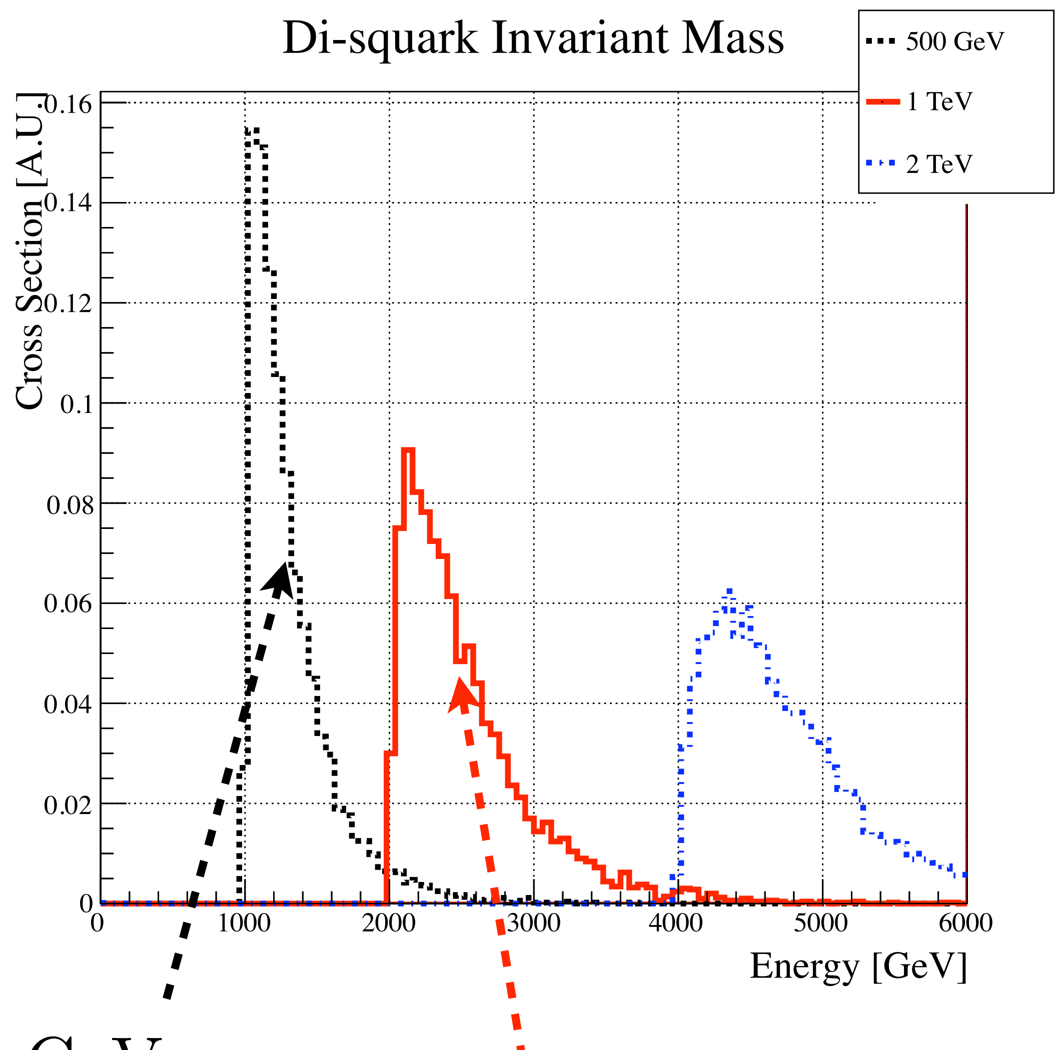
$$p_{\bar{T}_i} \to \frac{p_T^{\rm ISR}}{M_{\rm test}} E_i + \sqrt{1 + \left(p_T^{\rm ISR}/M_{\rm test}\right)^2}\, p_{\bar{T}_i}.$$



under boost | over boost | $M_{\rm test}=M_0$

Example: squark pair production

- Produced near threshold, $M_0 \sim 2m_{\tilde{q}}$



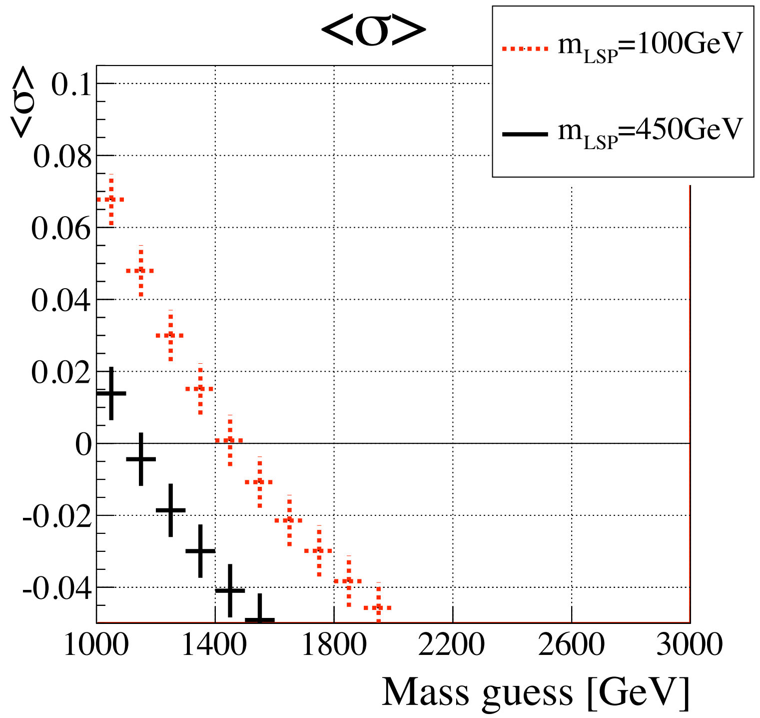
$m_{\tilde{q}} = 500$ GeV

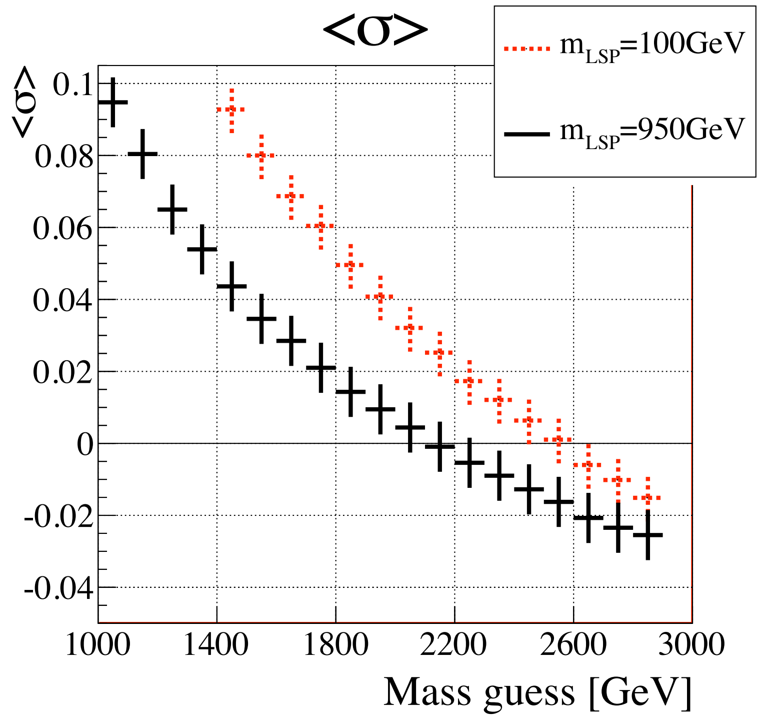
$M_0 \sim 1.2$ TeV

$m_{\tilde{q}} = 1$ TeV

$M_0 \sim 2.5$ TeV

Examples: squark pair production





$$m_{\tilde{q}} = 500\ \text{GeV}$$
$$M_0 \sim 1.2\ \text{TeV}$$

$$m_{\tilde{q}} = 1\ \text{TeV}$$
$$M_0 \sim 2.5\ \text{TeV}$$

$$\sigma = +1\ (-1)\ \text{if} \sum_i p_{\bar{T}_i} > 0\ (< 0),\ \langle\sigma\rangle = \sum_{j=1}^{N} \sigma_j / N$$

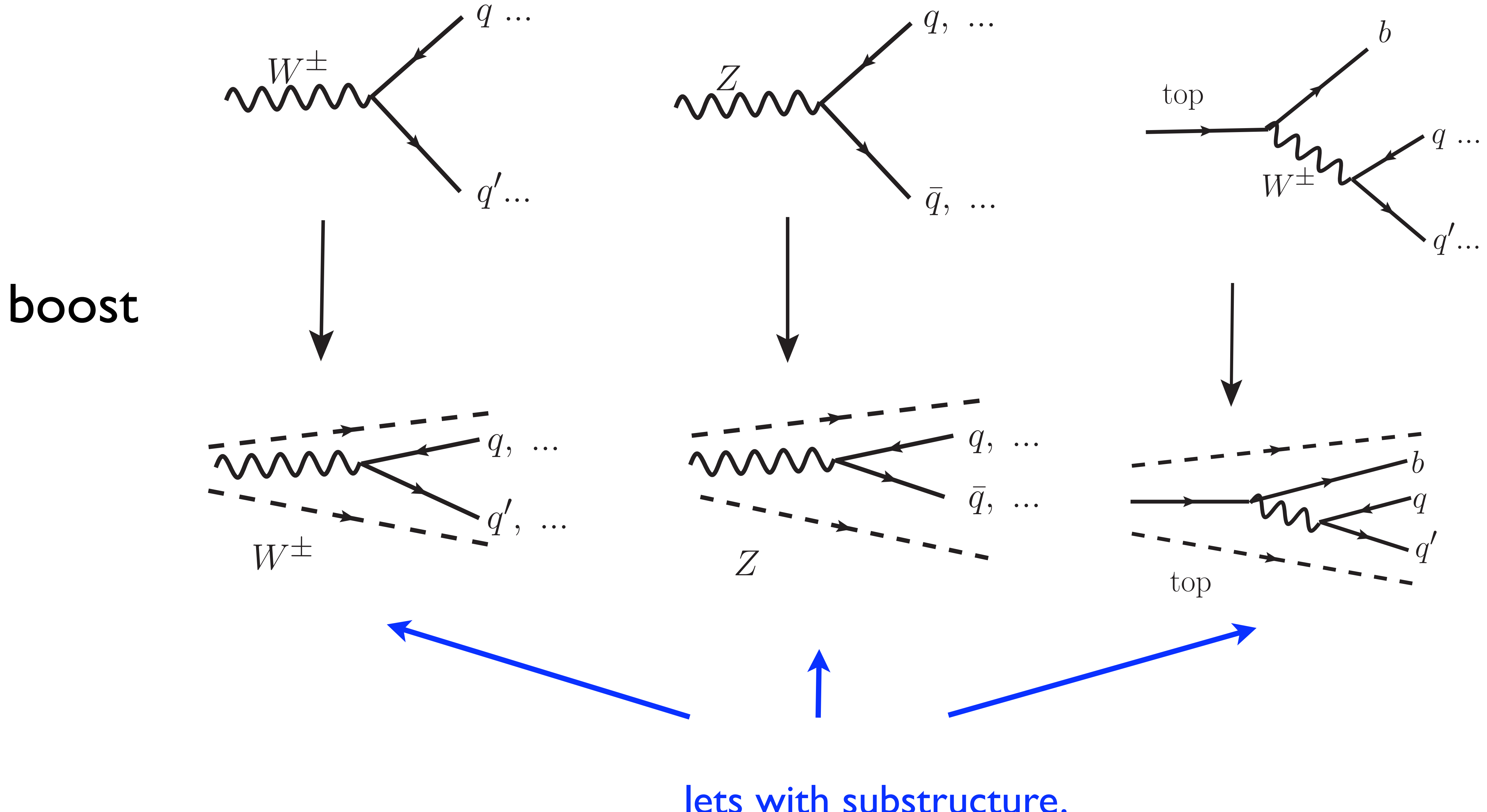
QCD Jets: current and future

- Well established the new improved jet algorithms will be instrumental in new physics discovery.
  - Optimization.
  - LHC experimental groups are testing them.
- More flexible, more dynamical.
- Jet tagging.
  - ISR. q vs g, charge?
- Better theoretical understanding.
  - Factorization ....

Jet substructure, and applications in new physics searches.

Jet substructure.

- When produced at TeV-scale energies, they have a large boost.

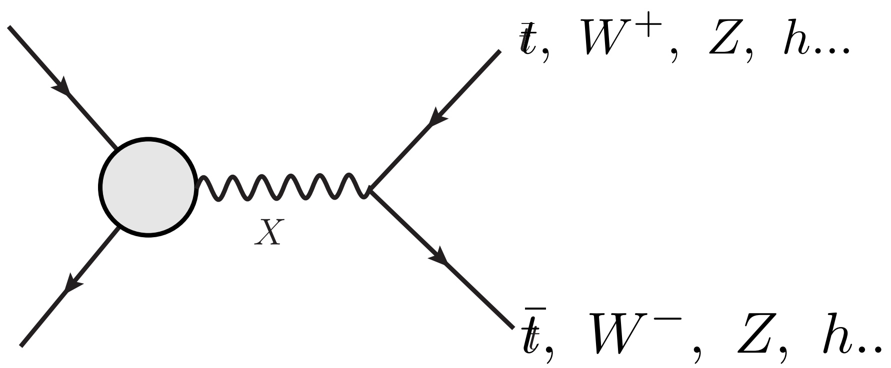


Challenge: distinguishing them from QCD jets (q and g).

When to consider substructure

- Have to consider the boosted objects.

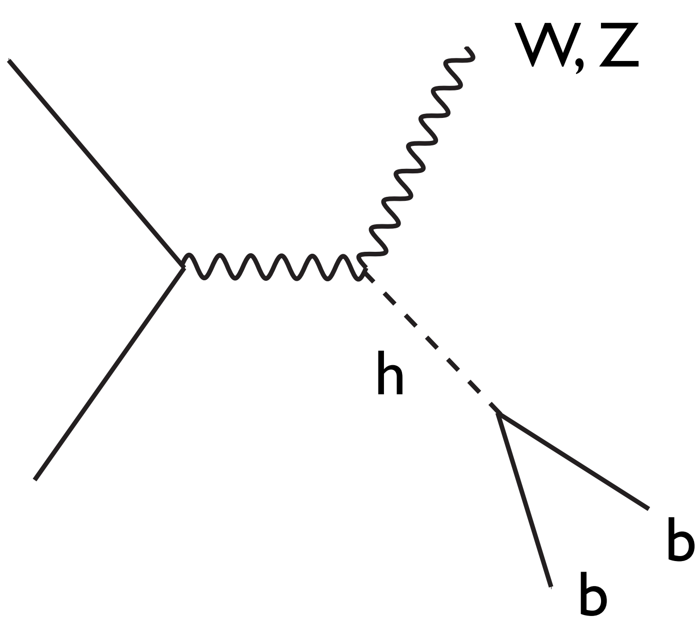
e.g.



For example, boost tops
Brooijmans; Lillie, Randall, LTW; Thaler, LTW;
D. Kaplan, K. Reherman, M. Schwartz, B. Tweedie;
L. Almeida, S. Lee, G. Perez, G. Sterman, I. Sung, J. Virzi
...

- It is beneficial to consider the boosted objects.

e.g.



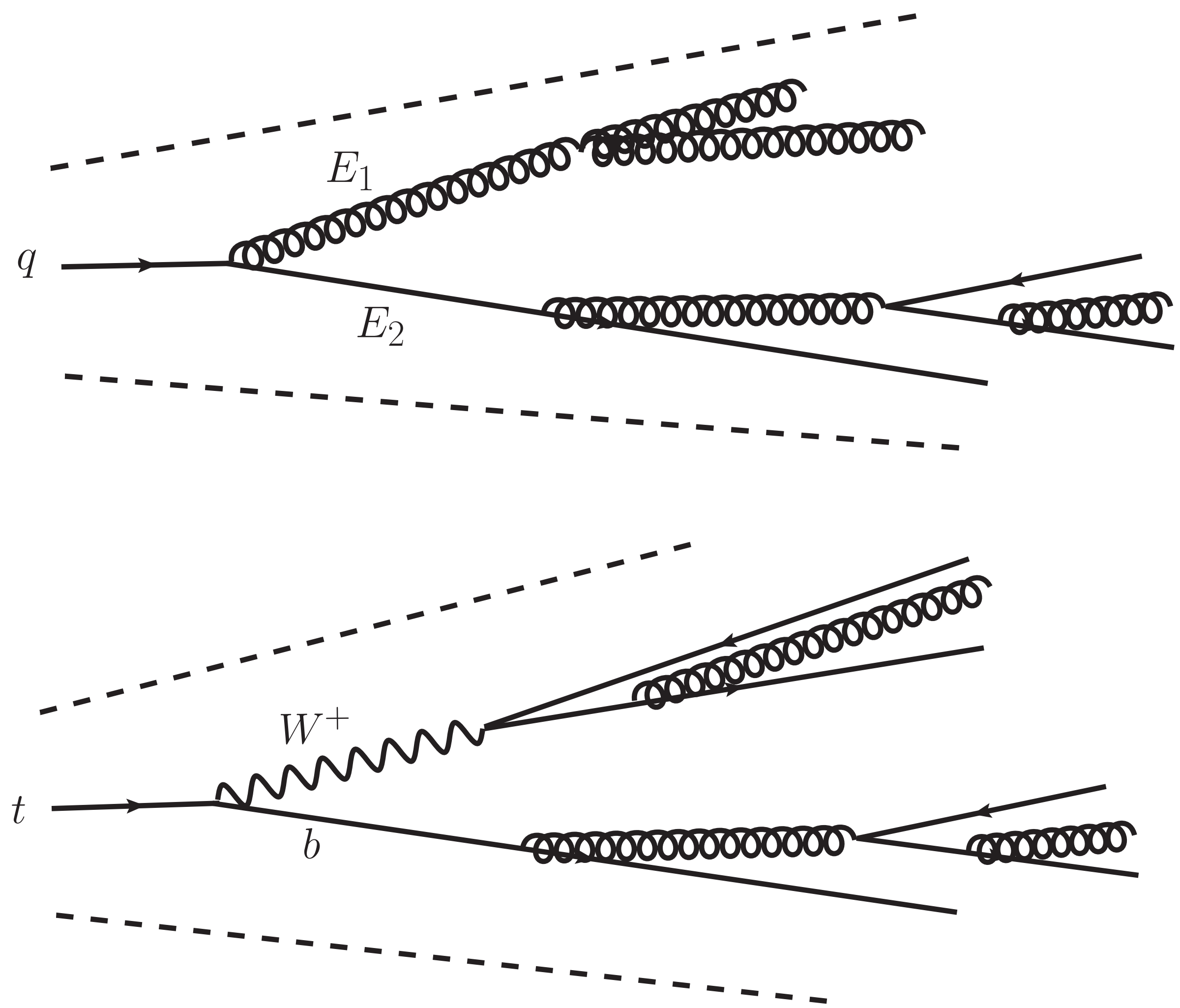
Lower combinatorics,
SM background boost differently.

Butterworth, Davidson, Bubin, Salam

For a summary of recent developments: C. Vermilion, 1001.1335

Example: boosted top tagging at the LHC

- Fully collimated tops look like QCD jets.

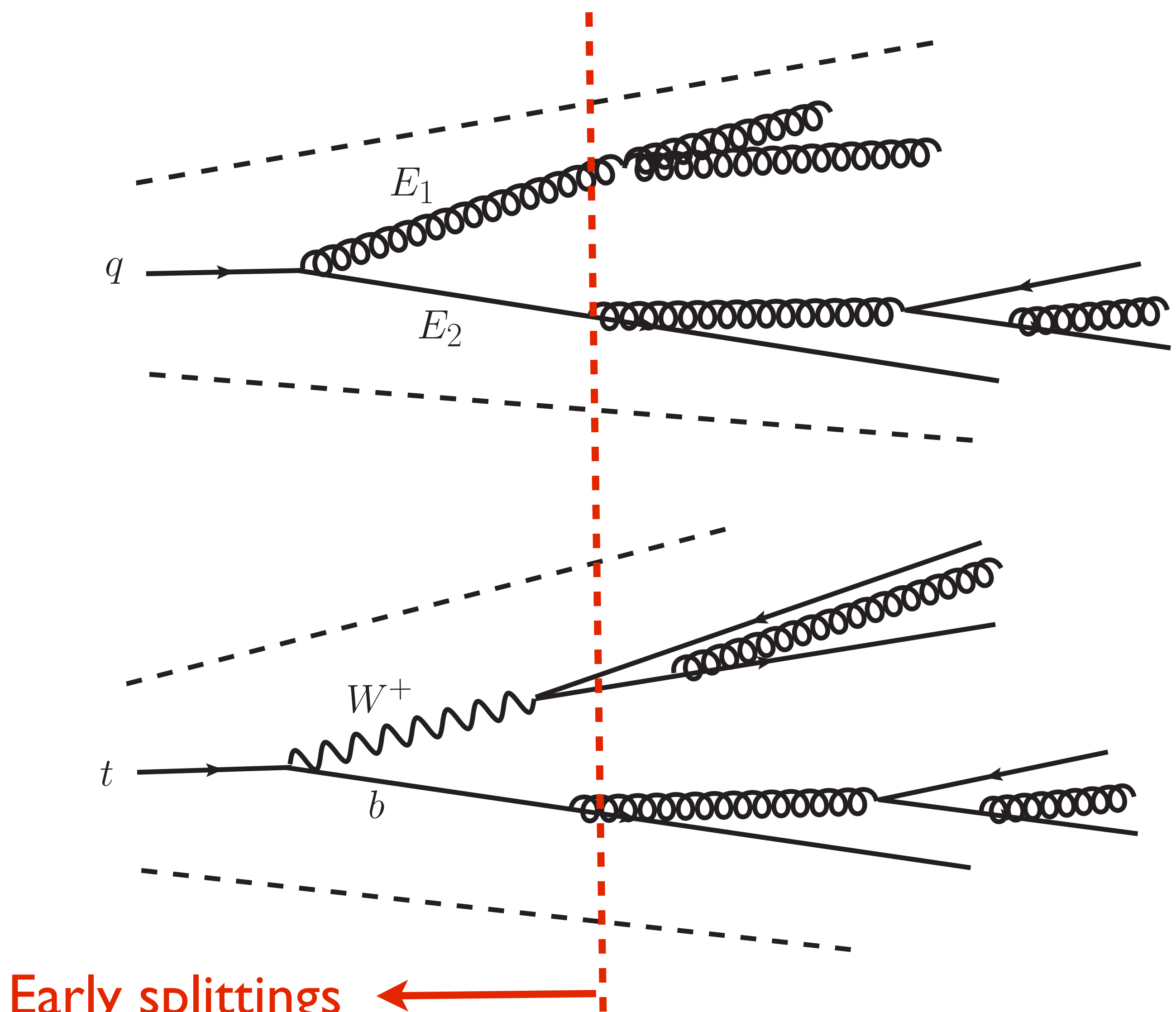


Example: boosted top tagging at the LHC

- Fully collimated tops look like QCD jets.

Basic distinction:

- QCD: radiation.
- Top decay: $t \to bW(\to qq')$ 3 hard objects.



Zooming in near the first splitting

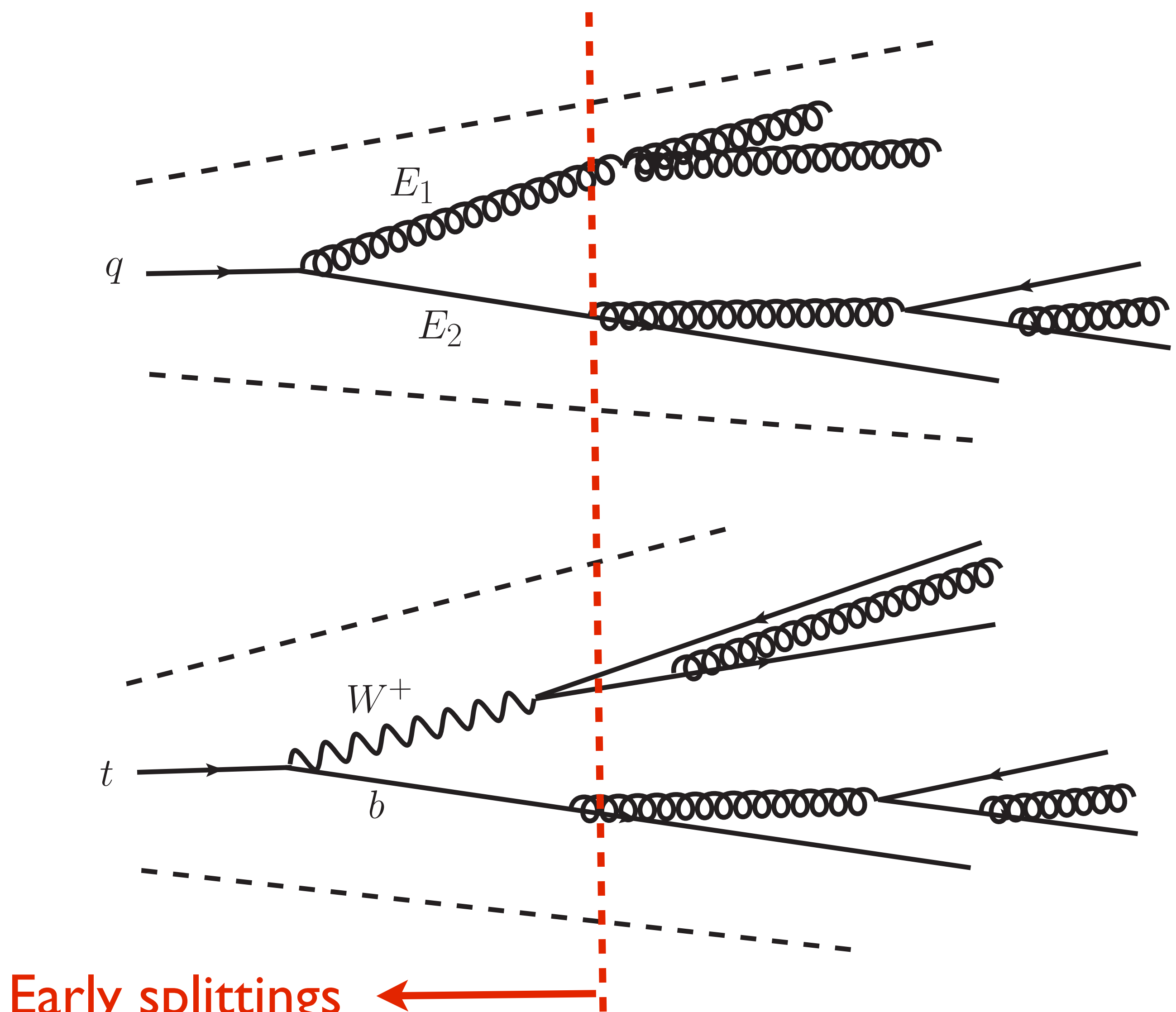
QCD. Soft radiation: $z = \frac{\mathrm{Min}(E_1, E_2)}{E_1 + E_2} \to 0$

Top. Decay: $z = \frac{\mathrm{Min}(E_\mathrm{W}, E_b)}{E_\mathrm{W} + E_b} \to \text{finite}$

Example: boosted top tagging at the LHC

- Fully collimated tops look like QCD jets.

Basic distinction:
- QCD: radiation.
- Top decay: $t \to bW(\to qq')$ 3 hard objects.



Zooming in near the first splitting

QCD. Soft radiation: $z = \dfrac{\text{Min}(E_1, E_2)}{E_1 + E_2} \to 0$

Jet mass: $d\sigma \propto \dfrac{1}{m_{\text{jet}}^2}$

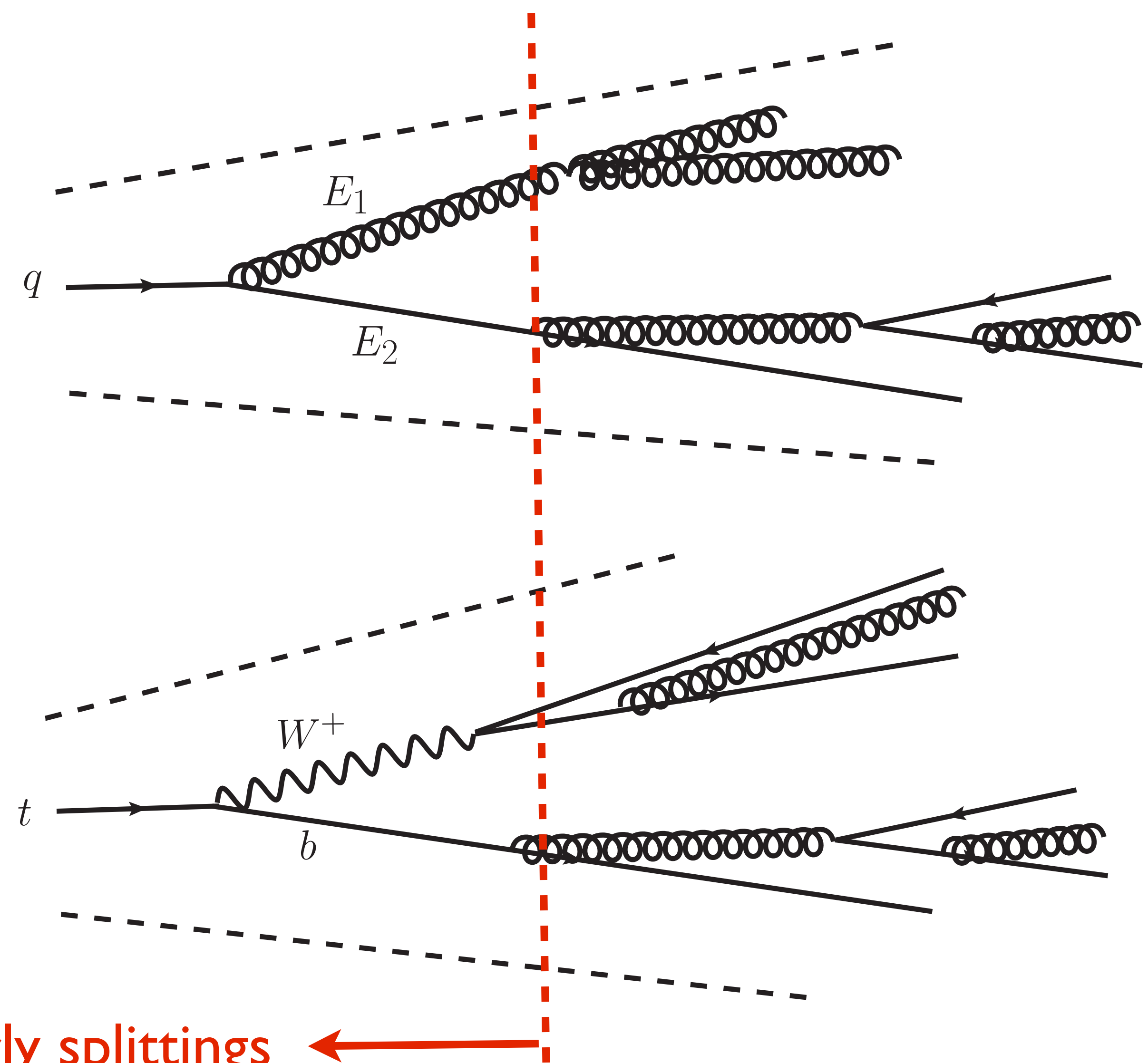
Top. Decay: $z = \dfrac{\text{Min}(E_W, E_b)}{E_W + E_b} \to \text{finite}$

Jet mass: $m_{\text{jet}} \simeq m_{\text{top}}$

Example: boosted top tagging at the LHC

- Fully collimated tops look like QCD jets.

Basic distinction:
- QCD: radiation.
- Top decay: $t \to bW(\to qq')$ 3 hard objects.



Zooming in near the first splitting

QCD. Soft radiation: $z = \frac{\mathrm{Min}(E_1, E_2)}{E_1 + E_2} \to 0$

Jet mass: $d\sigma \propto \frac{1}{m_{\mathrm{jet}}^2}$

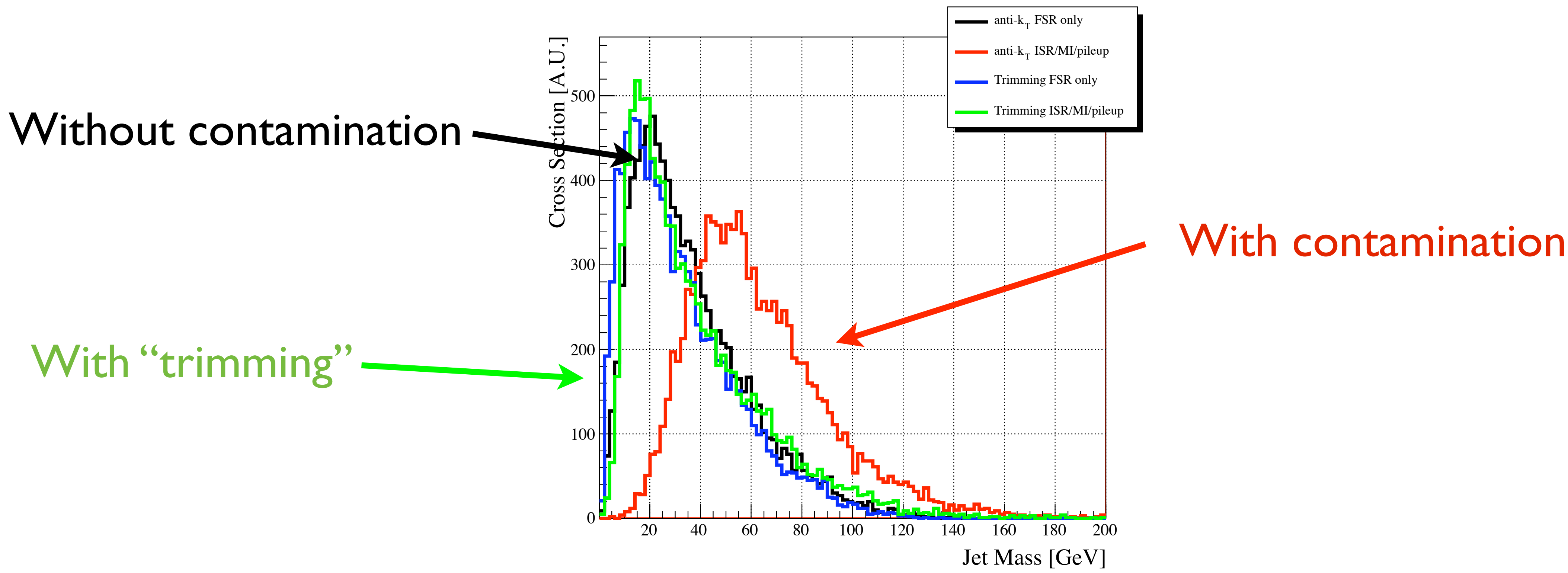
Top. Decay: $z = \frac{\mathrm{Min}(E_{\mathrm{W}}, E_b)}{E_{\mathrm{W}} + E_b} \to \mathrm{finite}$

Jet mass: $m_{\mathrm{jet}} \simeq m_{\mathrm{top}}$

Early splittings

microscope: jet substructure variables

Help from new jet algorithm



More faithful (smaller) jet mass for the background.

- Effect of radiation contamination on the jet mass

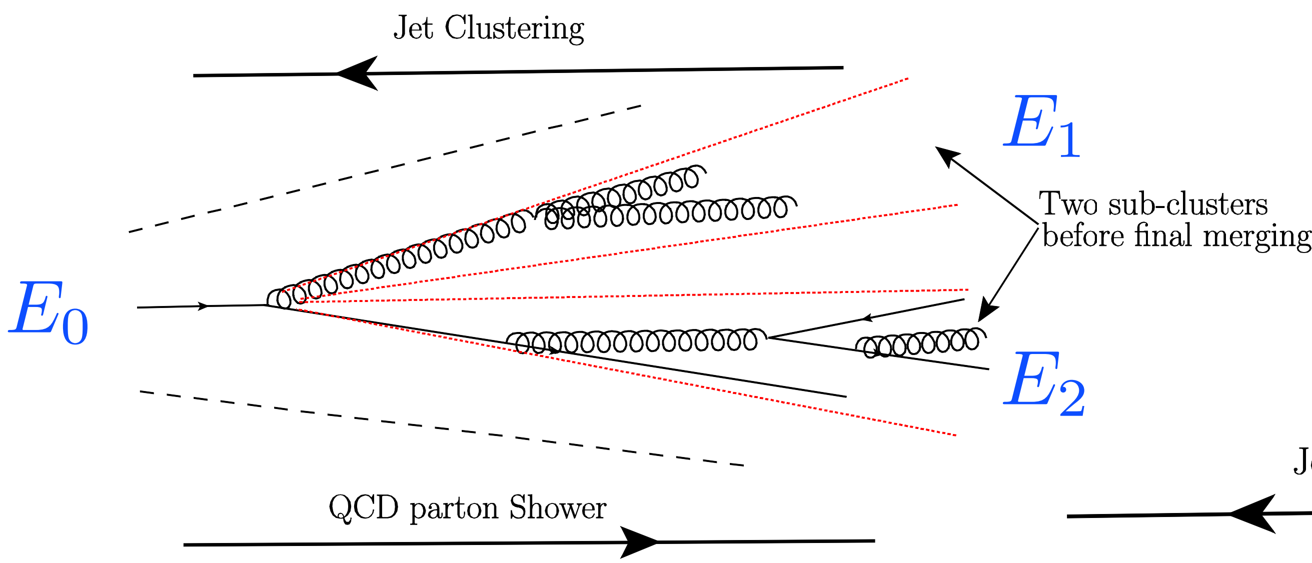
$$\langle \delta M^2 \rangle \simeq (\Lambda_{\rm soft} + p_{\rm T}^{\rm ISR}) p_{\rm T}^{j} \left(\frac{(\Delta R)^4}{4} + ... \right)$$

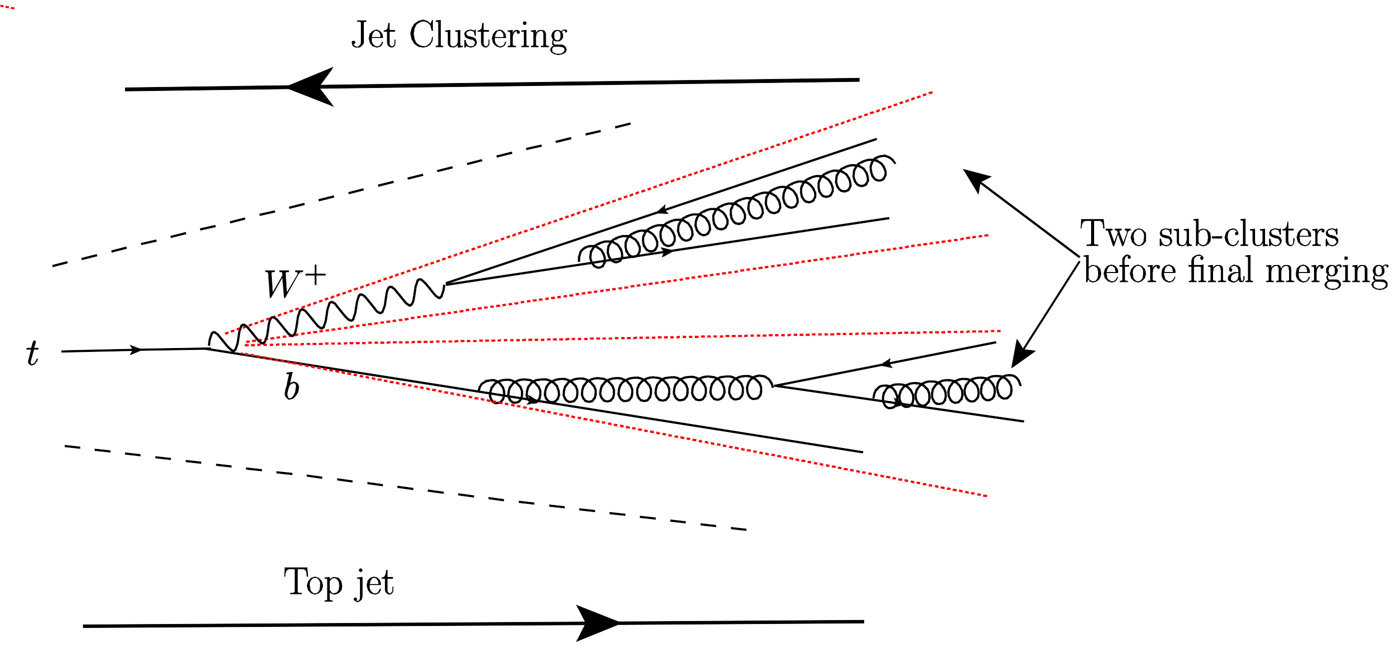
- Trimming gives large improvement by reducing effective jet size significantly.

Substructure, z-finding

$z \to 0$ for QCD jet,

z finite for top jet.



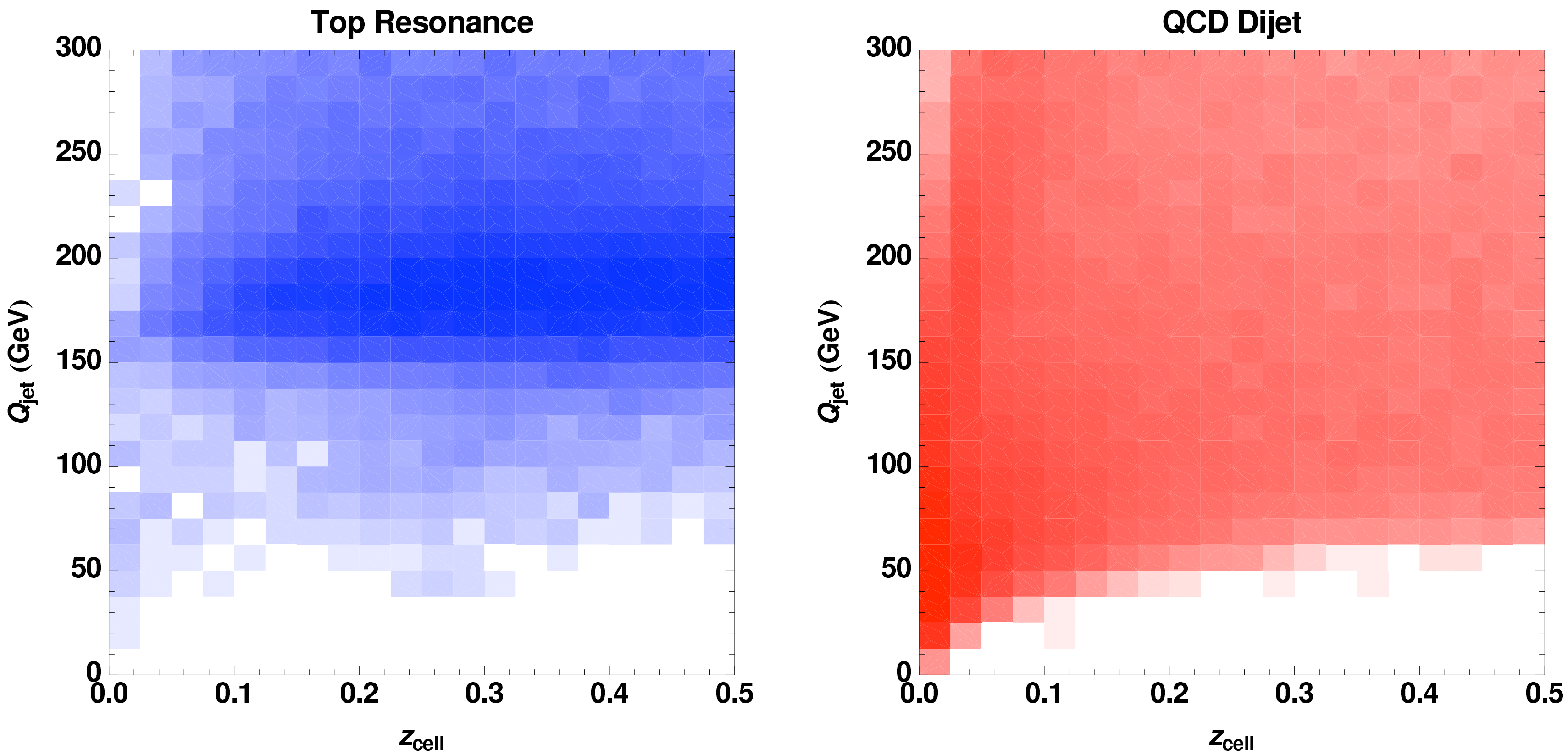


$$z = \frac{\min(E_1, E_2)}{E_0}$$

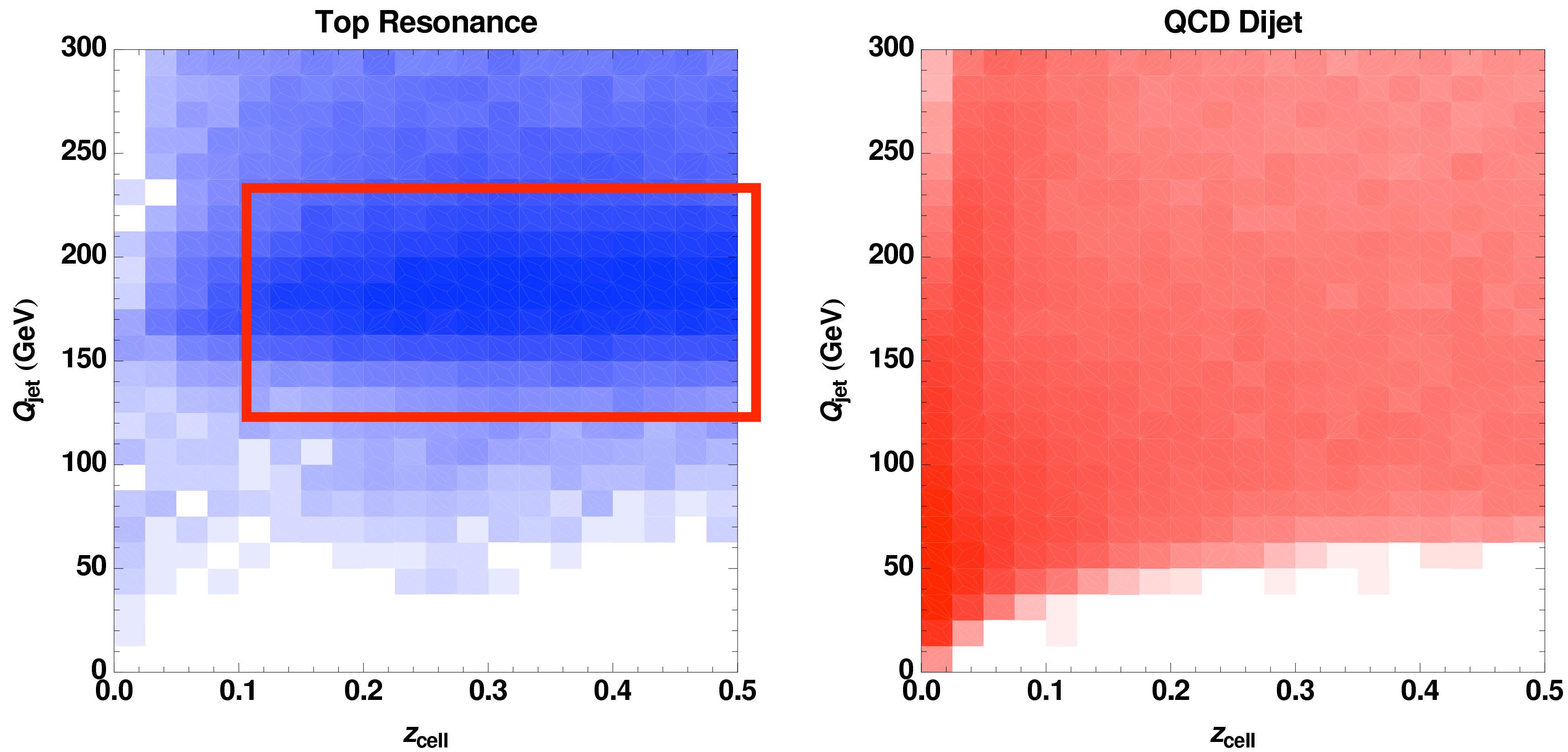
Q : jet mass

- Jet clustering history is approximately the inverse of parton shower.

Top jets vs QCD jets



Top jets vs QCD jets



- Combined cuts on jet mass and z can enhance further the signal with respect to the background.

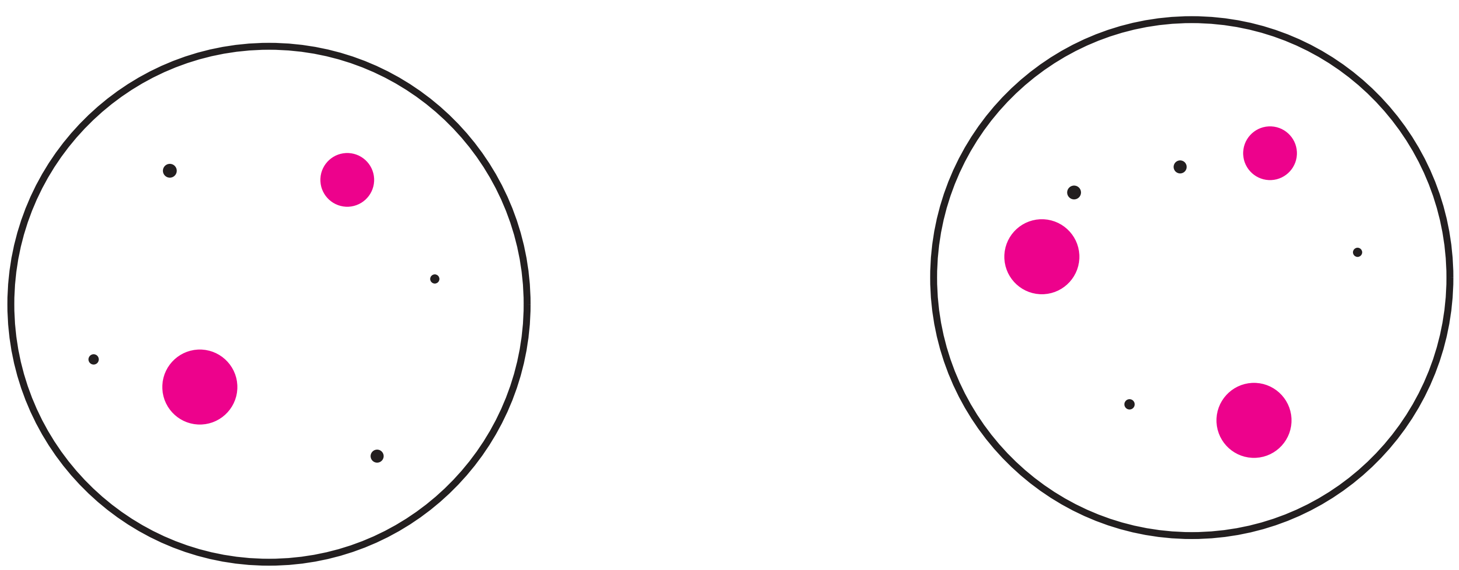
More jet shape variables.

- Top decay is more like 3-body. Span a "plane" perpendicular to the jet axis.
  - Transverse sphericity, or "planar flow"

$$S^{\perp ij} = \frac{\displaystyle\sum_{\alpha\in\text{jet}} \frac{\vec{p}_\alpha^{\perp i}\vec{p}_\alpha^{\perp j}}{|\vec{p}_\alpha^{\perp}|}}{\displaystyle\sum_{\alpha\in\text{jet}} |\vec{p}_\alpha^{\perp}|}.$$

$\vec{p}_\alpha^{\perp i}$: w.r.t. jet axis, $i = 1, 2$

J. Thaler and LTW, arXiv:0806.0023.

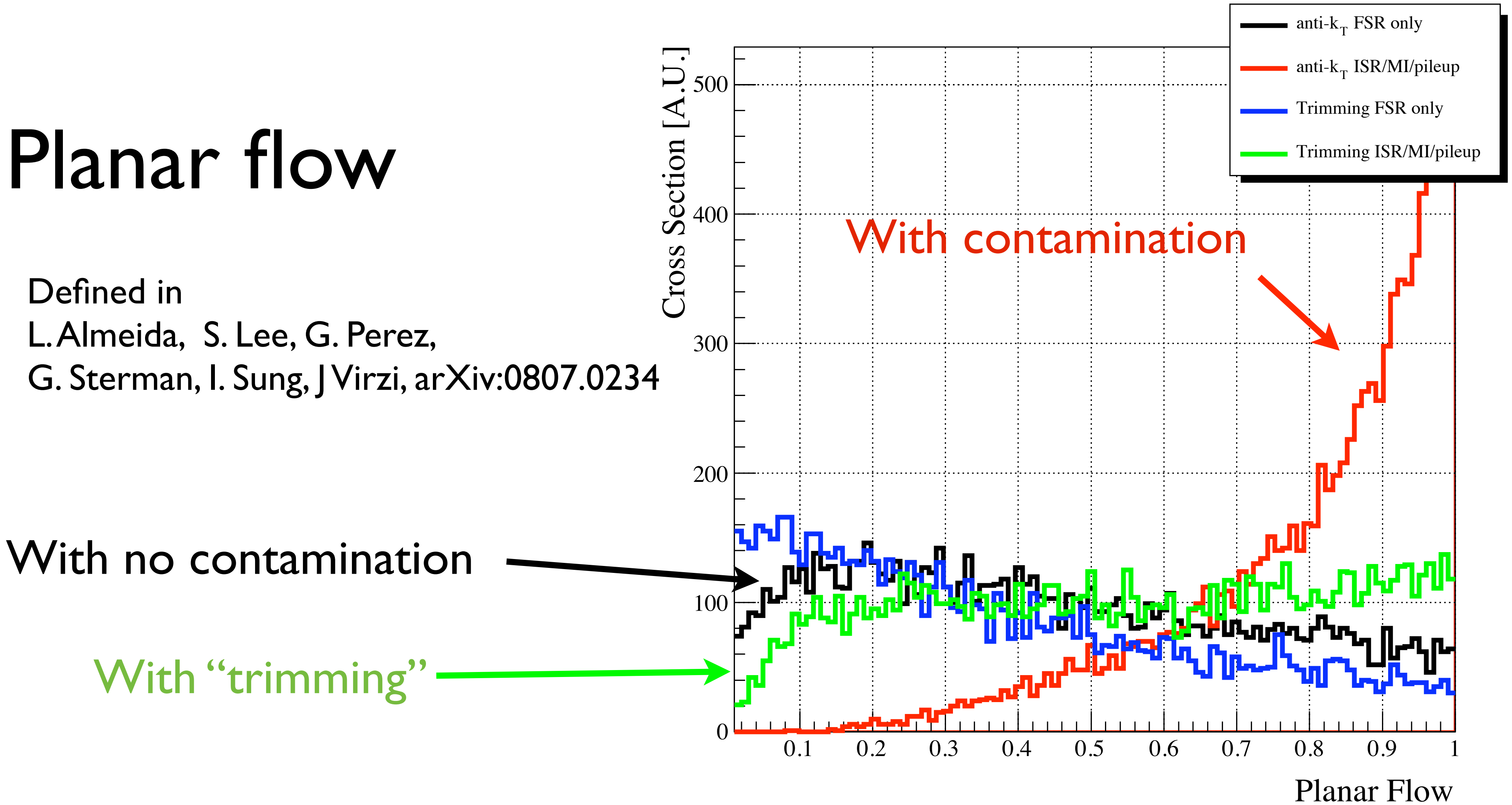


$\det S^{\perp} = 0$ $\qquad\qquad$ $\det S^{\perp} \neq 0$

Better reconstruction of the jet shape

Planar flow

Defined in
L. Almeida, S. Lee, G. Perez,
G. Sterman, I. Sung, J Virzi, arXiv:0807.0234



- Can be used to further improve top tagging. An additional factor of several possible.

- Interesting to compare with improved QCD calculation, using modern technologies such as SCET.

Hiding Higgs.

- Alternative decay channels can dramatically change Higgs search strategy.

$$h \to aa \to 4\tau,\ \ 4b,\ \ \bar{b}b\bar{\tau}\tau$$

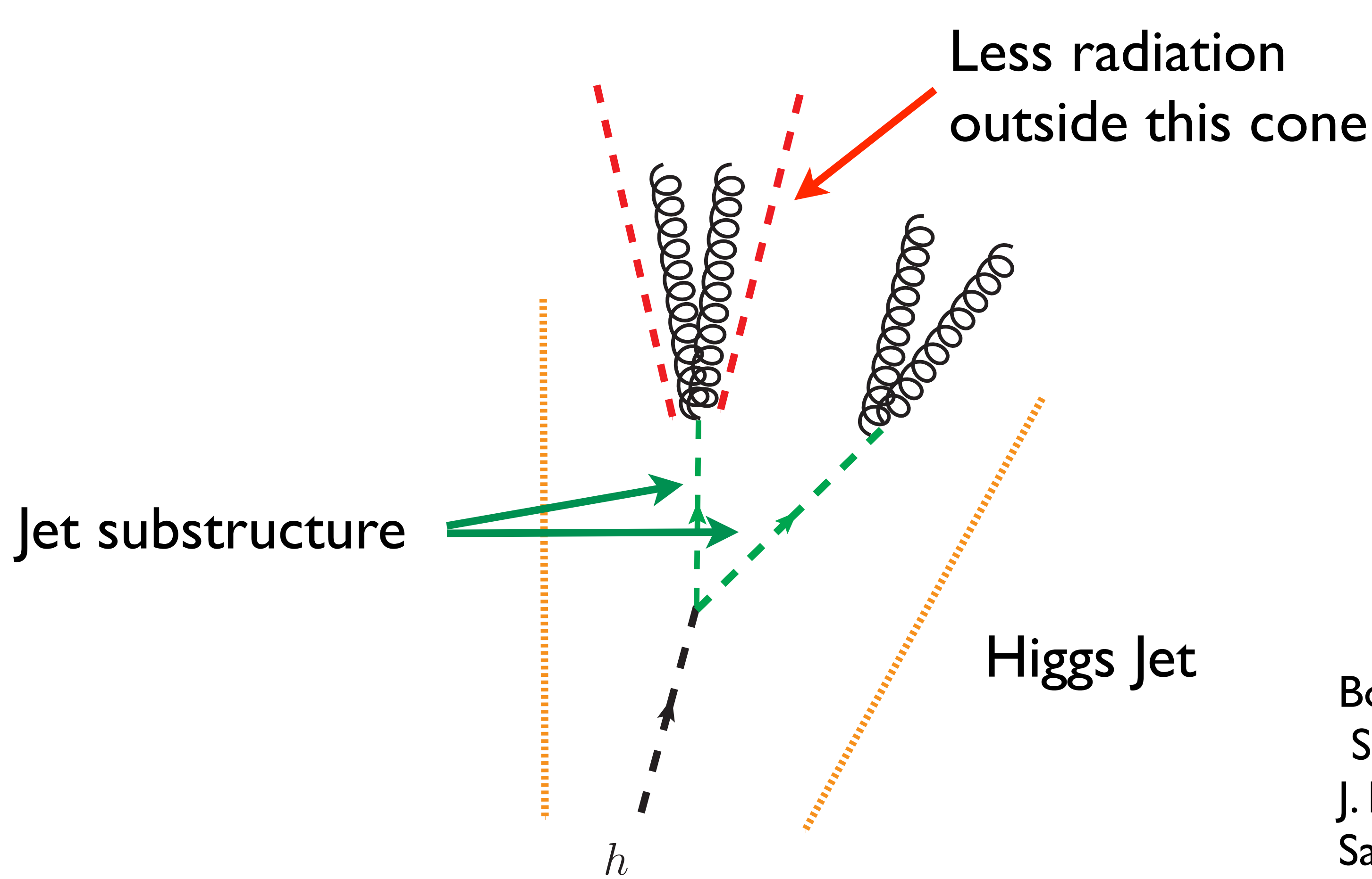
For example:
P. Graham, A. Pierce, J. Wacker, hep-ph/0605162
M. Carena, T. Han, G. Huang, C. Wagner, arXiv:0712.2466

$$h \to aa \to c\bar{c}c\bar{c},\ \ \text{“charmful”?}$$

$$h \to aa \to gggg,\ \ \text{“buried”!}$$

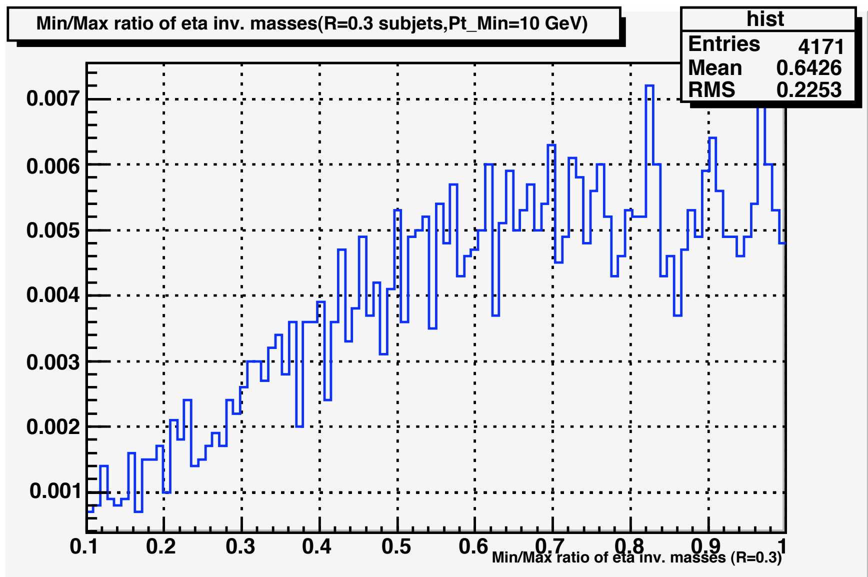
For example:
B. Bellazzini, C. Csaki, A. Falkowski, A. Weiler, arXiv:0910.3210, arXiv:0906.3026

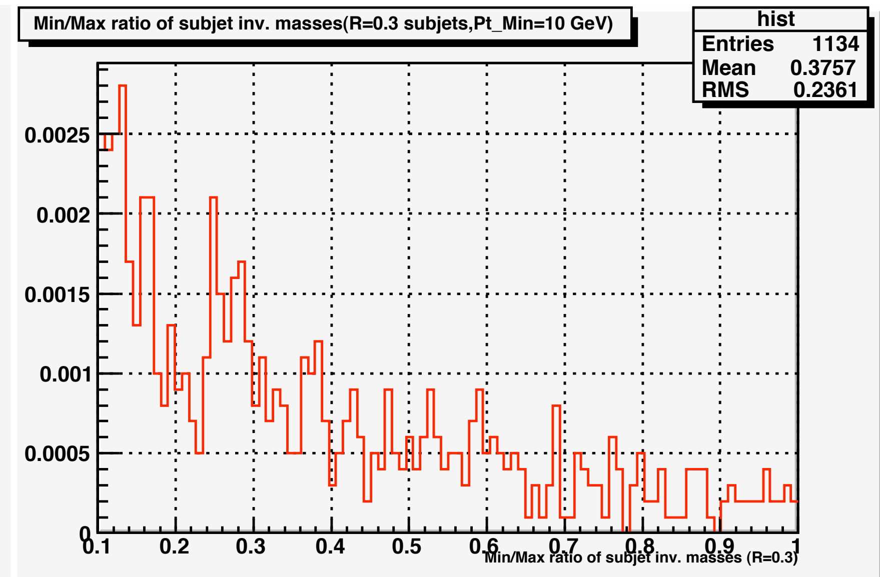
- Why can new jet technology help?

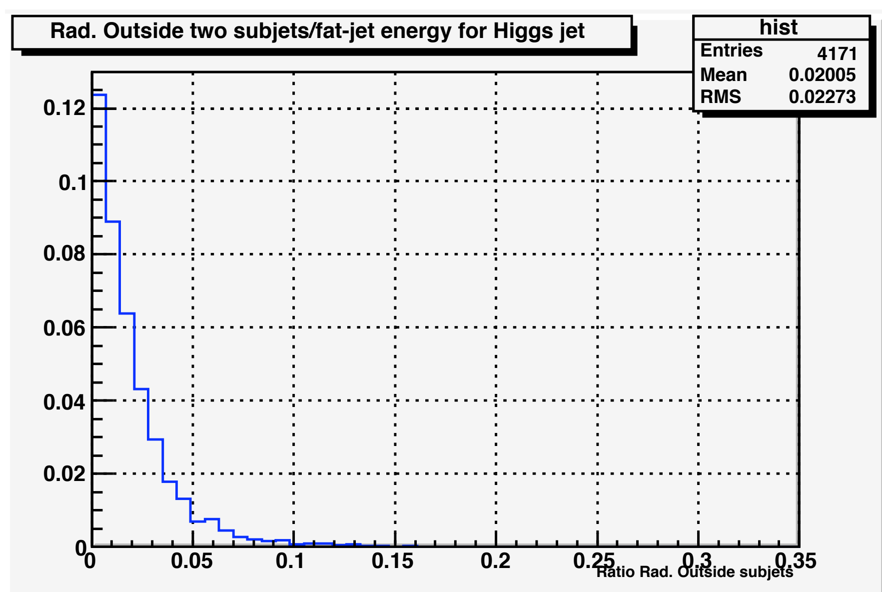


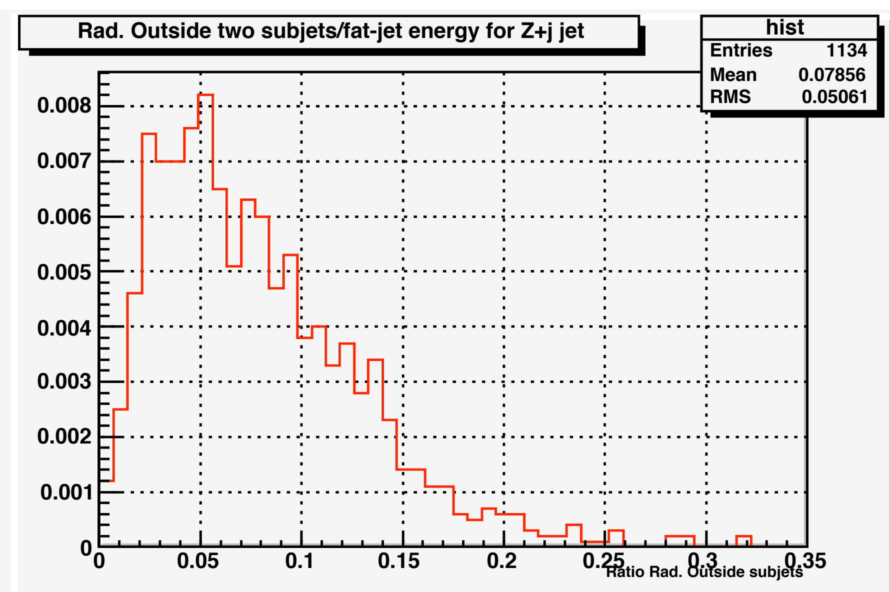
Boosted Higgs, studied in the context of SM-like Higgs by
J. Butterworth, A. Davidson, M. Rubin, G. Salam, arXiv:0802.2470

Some preliminary results.









Higgs + Z signal

Z+jet background

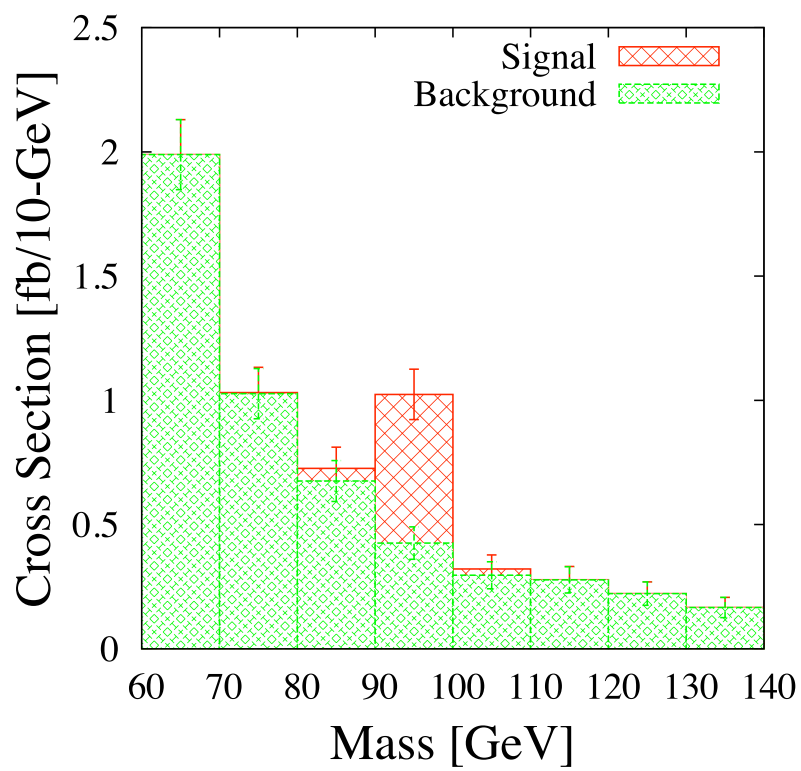
Encouraging results.

	σ_{sig} (fb)	σ_{bg} (fb)	S/B	$S/\sqrt{B}$
$p_T(j) > 200$ GeV	16	30000	0.00052	0.9
subjet mass	12	19000	0.00062	0.9
Higgs window	7.1	400	0.018	3.6
$\alpha > 0.7$	4.1	140	0.030	3.5
$\beta < 0.005$, $p_T^{\min} = 1$ GeV	0.67	0.74	0.90	7.8
$\beta < 0.005$, $p_T^{\min} = 5$ GeV	2.9	2.6	0.11	5.7

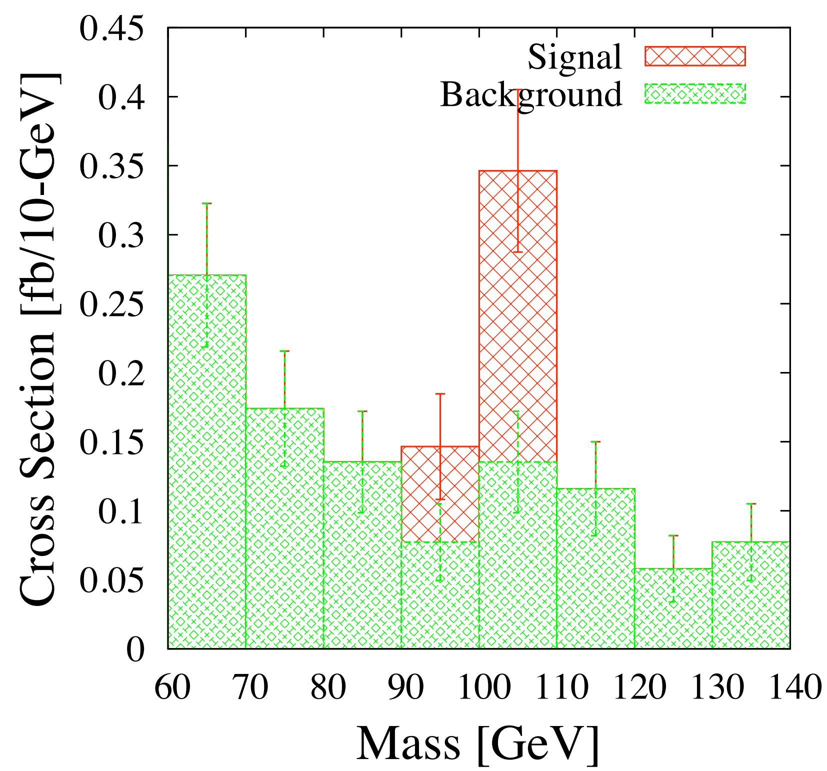
sub-jet mass balance → ($\alpha > 0.7$)

radiation pattern → ($\beta < 0.005$, $p_T^{\min} = 5$ GeV)

$> 5\sigma$ at 100 fb^{-1}



W/Z+h



ttbar+h

A. Falkowski, D. Krohn, J. Shelton, A. Thalapillil, and LTW, arXiv:1006.1650

Substructures: current and future

- Boosted (almost) everything already.
- Rather simple techniques. Currently,
  - Combine, optimize, test.
  - LHC experimental groups have started using them.
- New substructure (jet shape).
  - Color flow...
- Better theoretical (QCD) understanding of the substructure.
  - Particularly for the QCD jets.
  - SCET....

Conclusions

- Better handles on the hadronic final states are instrumental for discovery at the LHC.
- Based on consideration of QCD radiation, we proposed a set of carefully constructed new jet algorithms and substructure variables.
  - Much improved performance, jet mass, jet shape, etc.
  - They can significantly enhance new physics signals in many important new physics channels.
    - Boosted or "slow" hadronic tops, WW scattering, Higgs search, heavy squark...
- A promising direction. Stay tuned.

ISR tagging

- Pick: distinct PT

$$\frac{\max(p_{T_i}, p_{T_j})}{\min(p_{T_i}, p_{T_j})} > 2, \ \ i \neq j$$

- Or, distinct rapidity

$$|y_i - y_j| > 1.5, i \neq j$$

- Or

$$\frac{\max(\Delta_i, \Delta_j)}{\min(\Delta_i, \Delta_j)} > 1.5, \ \ \Delta_i = m_i / p_{T_i}$$

- And, not central

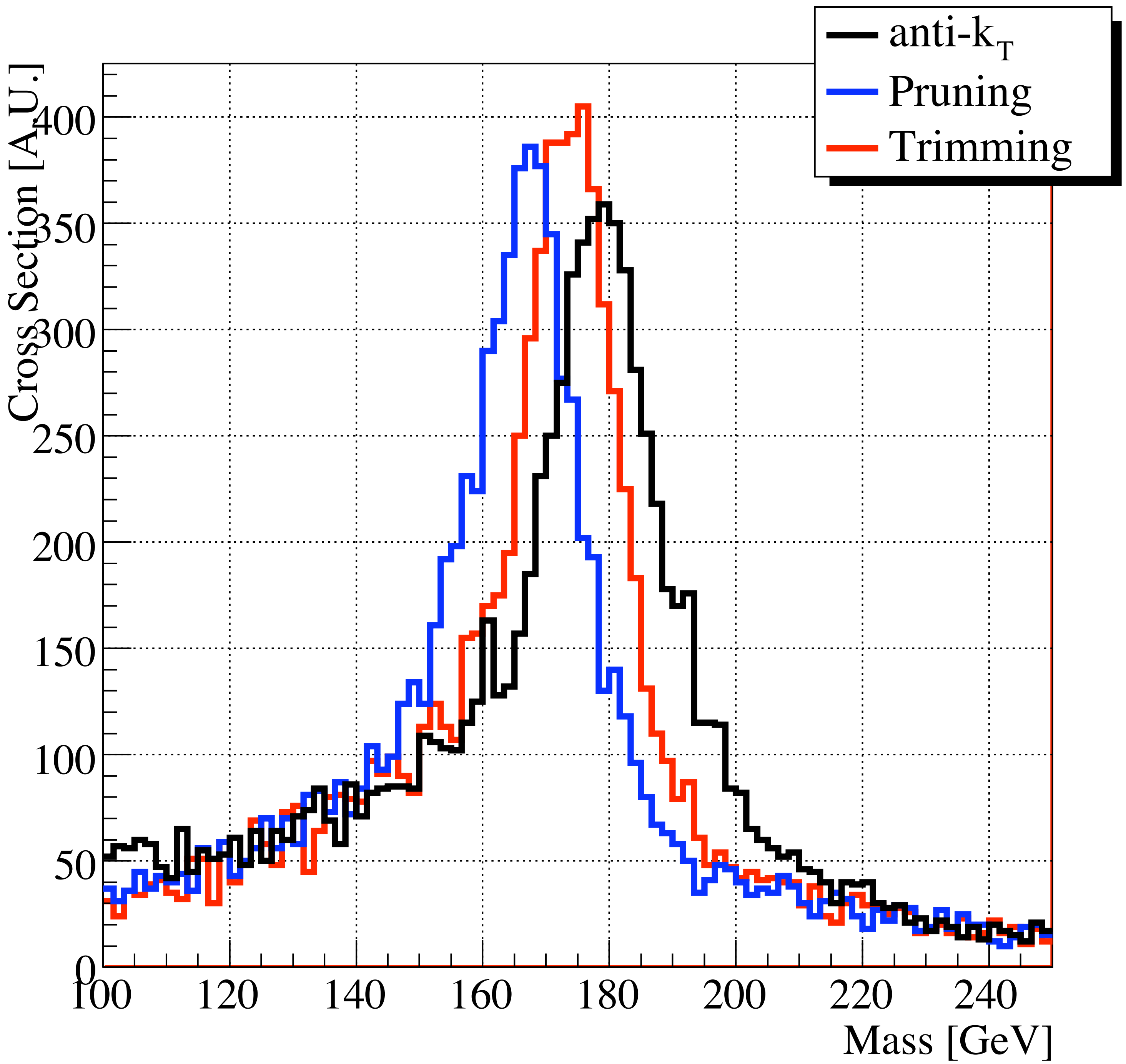
$$|y_i| > 1$$

- Separated from others

$$|y_i - y_j| > 0.5$$

- FSRs must be similar

$$\frac{p_{T_j}}{p_{T_k}} < \rho + \frac{0.5}{1-\alpha}$$
$$p_{T_{j(k)}} = \max(\min)\{p_{T_l}, l \neq i\},$$
$$\rho = 2(3) \text{ for } N_f = 2(4),$$
$$\alpha = \frac{\min(p_{T_i}, \not{E}_T)}{\max(p_{T_i}, \not{E}_T)}$$



anti-k_T
Pruning
Trimming
Cross Section [A.U.]
Mass [GeV]
0
50
100
150
200
250
300
350
400
100
120
140
160
180
200
220
240

Why is it possible to gain?

- MI, UE, and pile-up are incoherent soft background. They can be effectively removed with a cut on soft radiation.

- Both FSR (want to keep) and ISR (want to discard) have soft radiation, but
  - ISR: $d\sigma \propto \frac{dp_{\mathrm{T}}^{\mathrm{ISR}}}{p_{\mathrm{T}}^{\mathrm{ISR}}}$
  - FSR is controlled by both collinear and soft singularities:
    $$d\sigma \propto \frac{d(\Delta R)}{\Delta R} \times \frac{dp_{\mathrm{T}}^{\mathrm{FSR}}}{p_{\mathrm{T}}^{\mathrm{FSR}}}$$
  - Therefore, a soft cut relative to the jet energy flow could enhance FSR relative to ISR.

Planar Flow

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

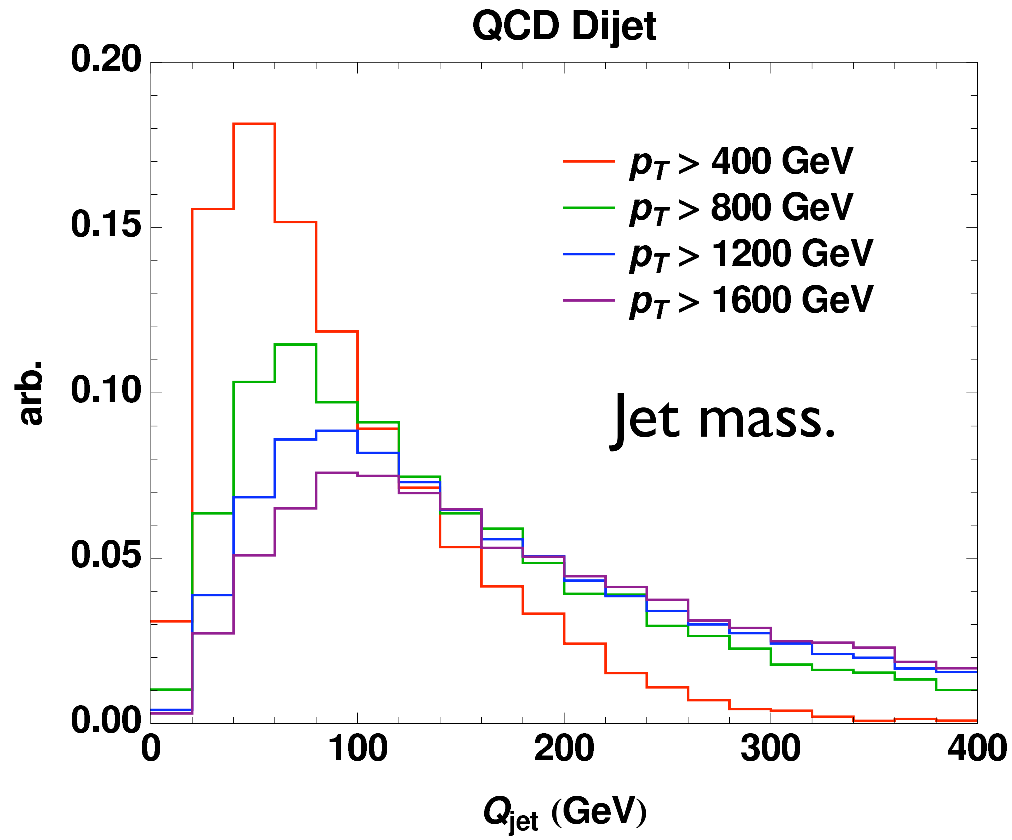
$\lambda_1, \; \lambda_2 : \;$ 2 eigenvalues of I_w^{kl}

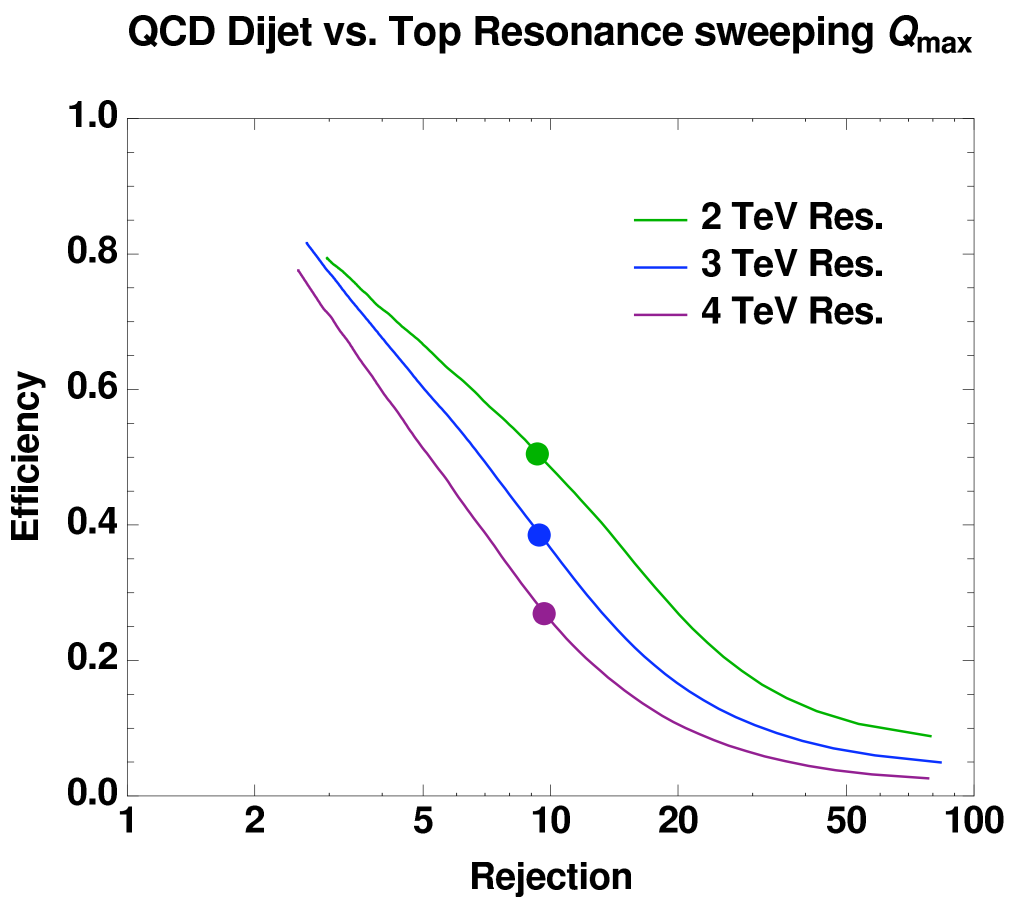
$$\mathrm{Pf} = \frac{4\lambda_1\lambda_2}{(\lambda_1 + \lambda_2)^2}$$

Top tagging: jet mass

- QCD jets also have mass.

$$\begin{aligned}\langle M^2 \rangle &\simeq \int \frac{d\theta^2}{\theta^2} \int dz\ p_{\mathrm{T}}^2 z(1-z)\theta^2\ \frac{\alpha_s(p_{\mathrm{T}})}{2\pi} P(z)\Theta(\Delta R - \theta) \\ &\simeq C\frac{\alpha_s}{\pi} p_{\mathrm{T}}^2 (\Delta R)^2\end{aligned}$$



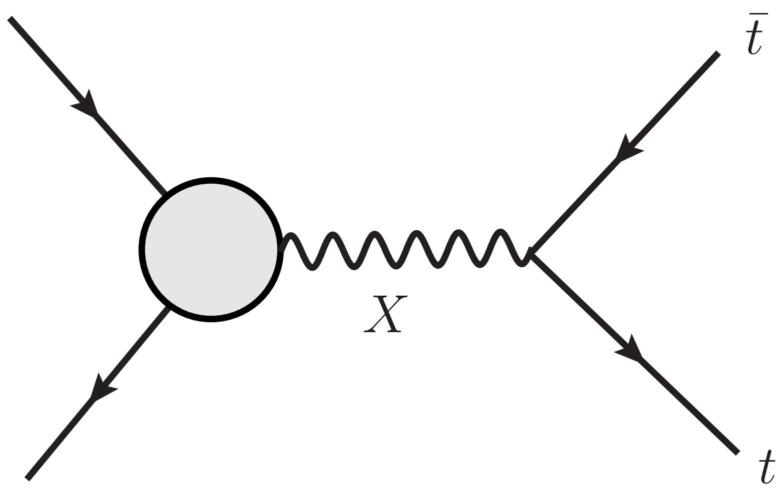


Using jet mass only.

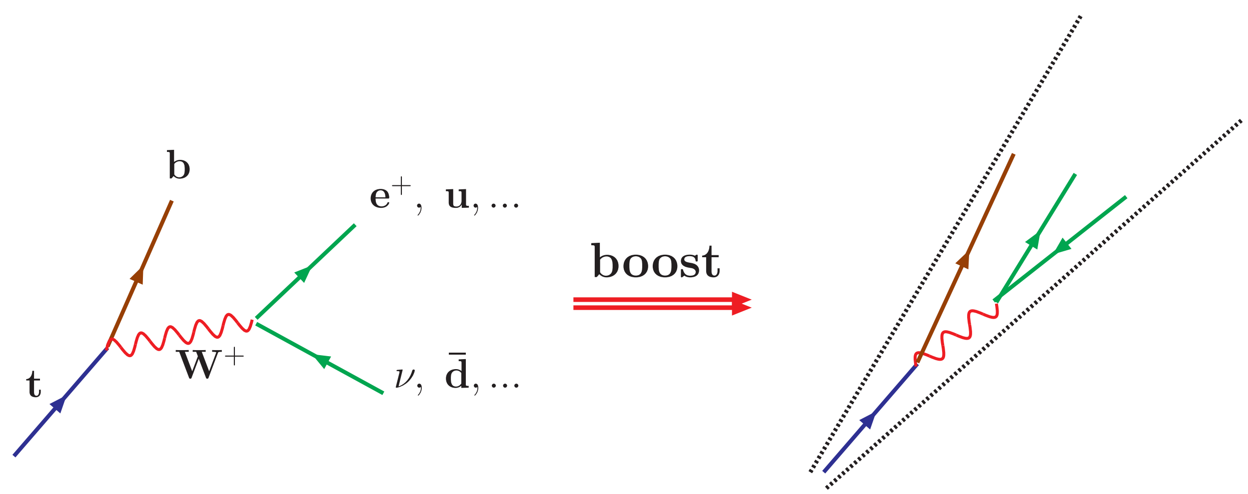
Useful. Additional variable?

Boosted top is also hard to identify.

- Heavy resonance decay.



$$E_{\rm top} \simeq \frac{M_X}{2}$$



← No isolated objects

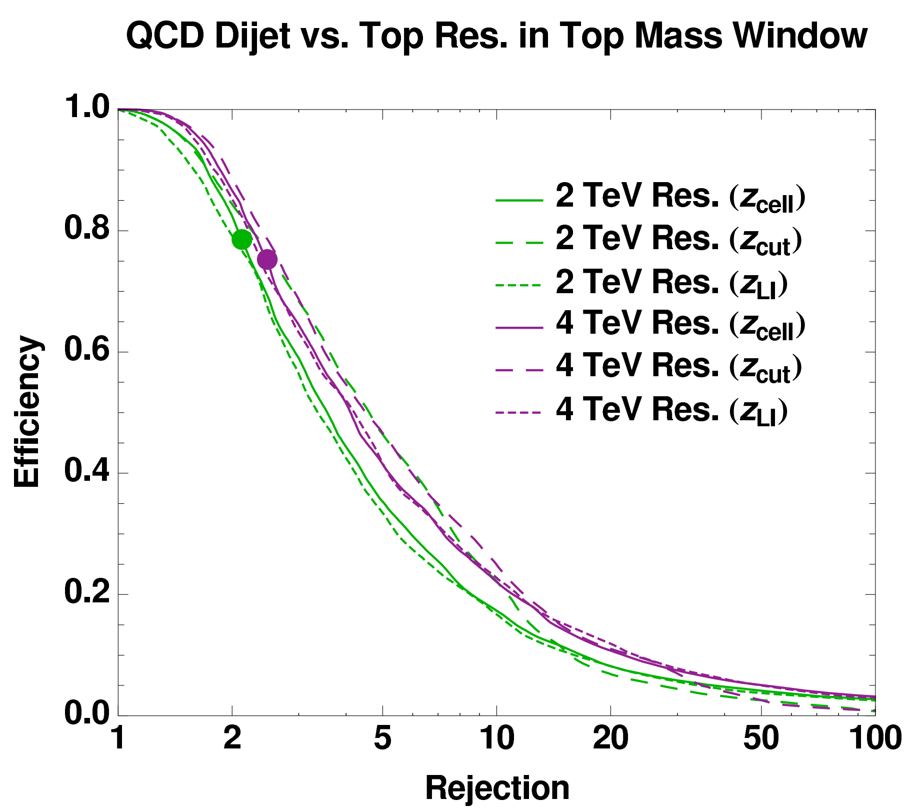
- For $m_{t\bar{t}} > 3$ TeV, $> 90\%$ events with at least one top fully collimated.
- Large fraction of events "2-object"-like. QCD $b\bar{b}$, jj background.
- A few % with lepton isolation

B. Lillie, L. Randall, and LTW, hep-ph/0701166

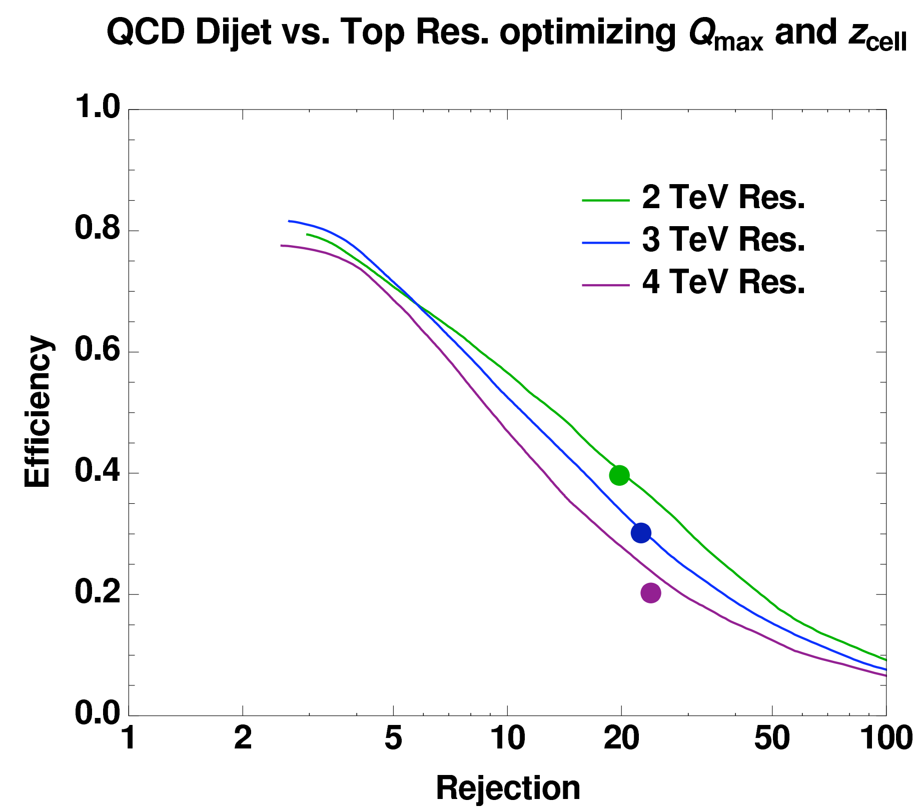
L. Almeida, S. Lee, G. Perez, I. Sung, J. Virzi, arXiv:0810.0934

Top tagging efficiency

J. Thaler and LTW, arXiv:0806.0023.



Performance of different z variables.



Combined cuts

- z-variable gives an additional about factor of 2 enhancement in performance.
- Together with jet mass, an enhancement of 100 of S/B is possible.

Related studies:

D. Kaplan, K. Reherman, M. Schwartz, B. Tweedie, arXiv: 0806.0848.

L. Almeida, S. Lee, G. Perez, G. Sterman, I. Sung, J. Virzi, arXiv:0807.0243

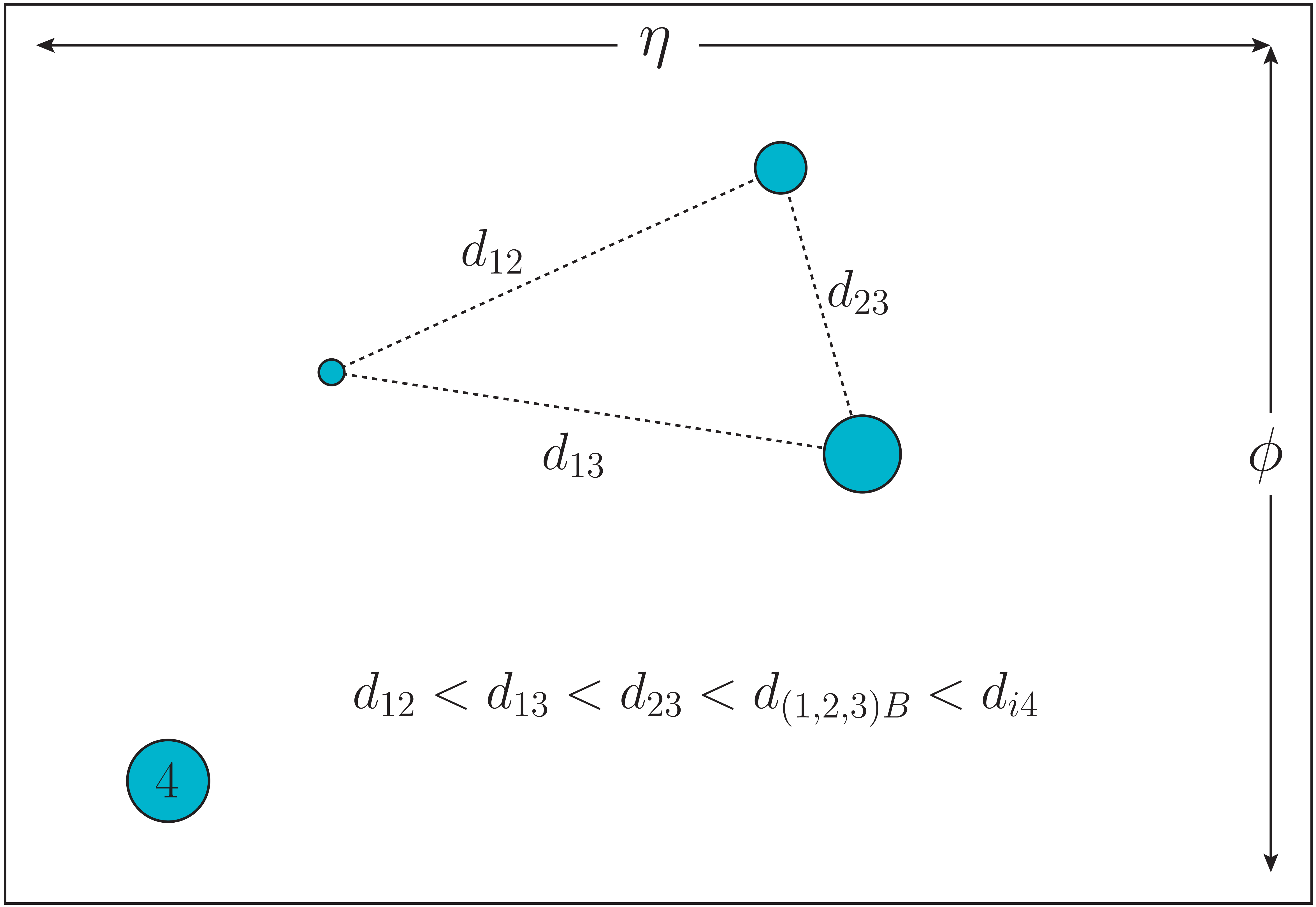
Gustaaf H. Brooijmans, arXiv:0802.3715; CMS, CMS PAS JME-09-001

Boosted tops.

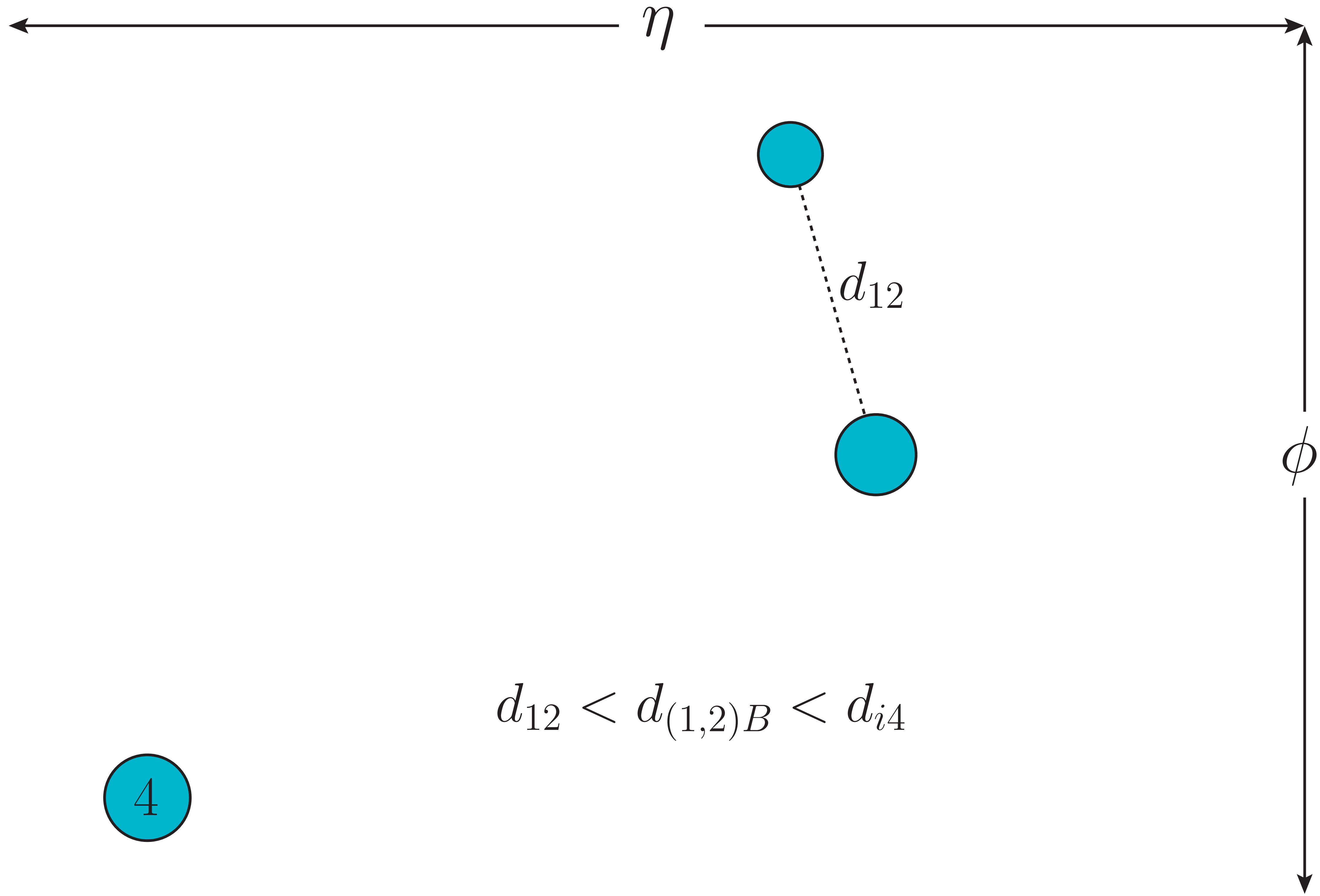
- Tops are interesting!
  - Top plays an important role in electroweak symmetry breaking.
  - Top generically couples to heavy new resonances which is an important part of TeV new physics.
  - Examples.
    - Composite top couples strongly to other composite resonances.

      Many examples.
      K. Agashe, A. Delgado, M. May, R. Sundrum, hep-ph/0308036
      M. Carena, B. Panes, A. Medina, N. Shah, C. Wagner, arXiv:0706.1281, 0712.0095
    - New heavy scalars couple like Higgs.

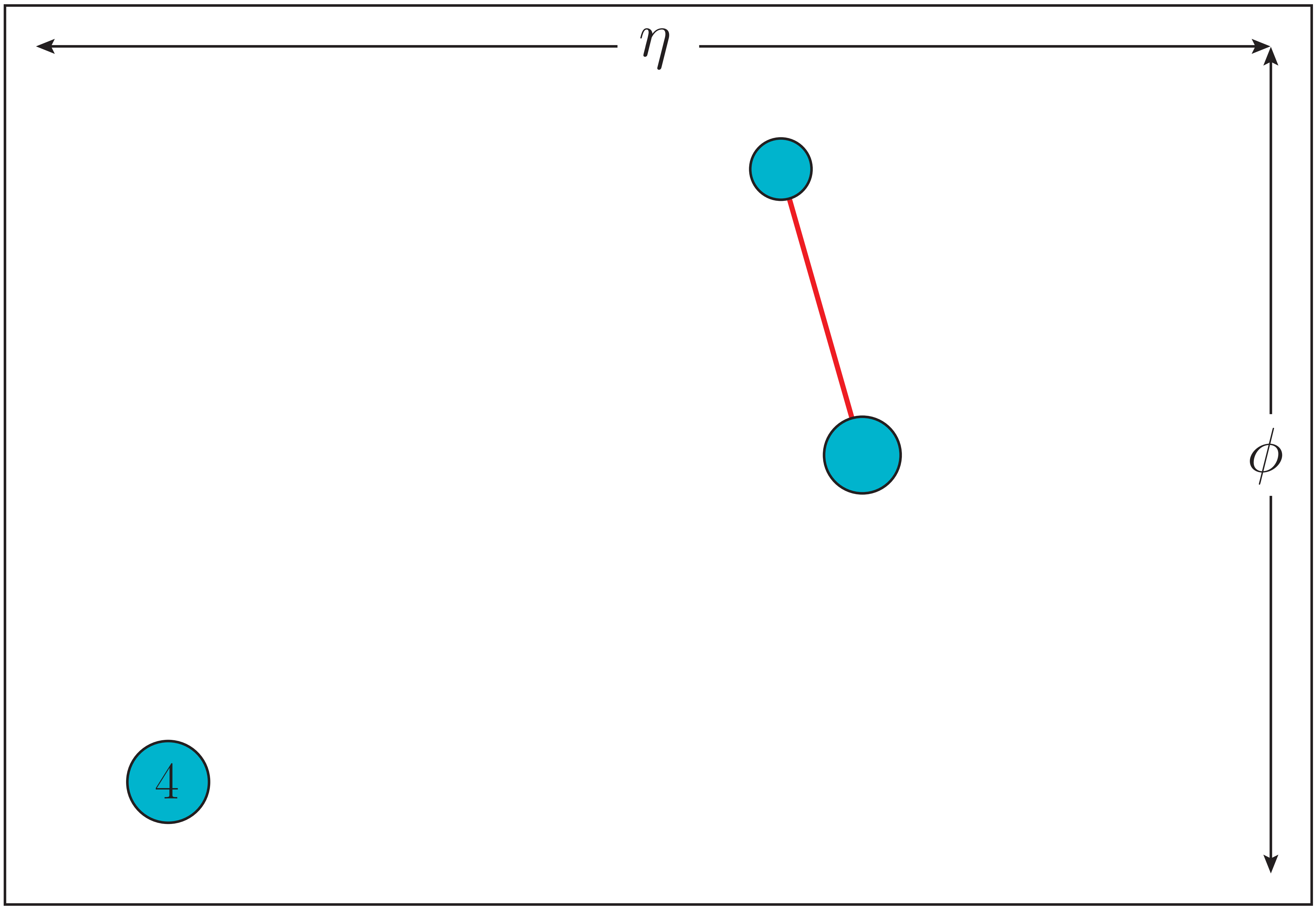
      For example: A. Manohar and M. Wise, hep-ph/0606172
  - A good example of subjet techniques.

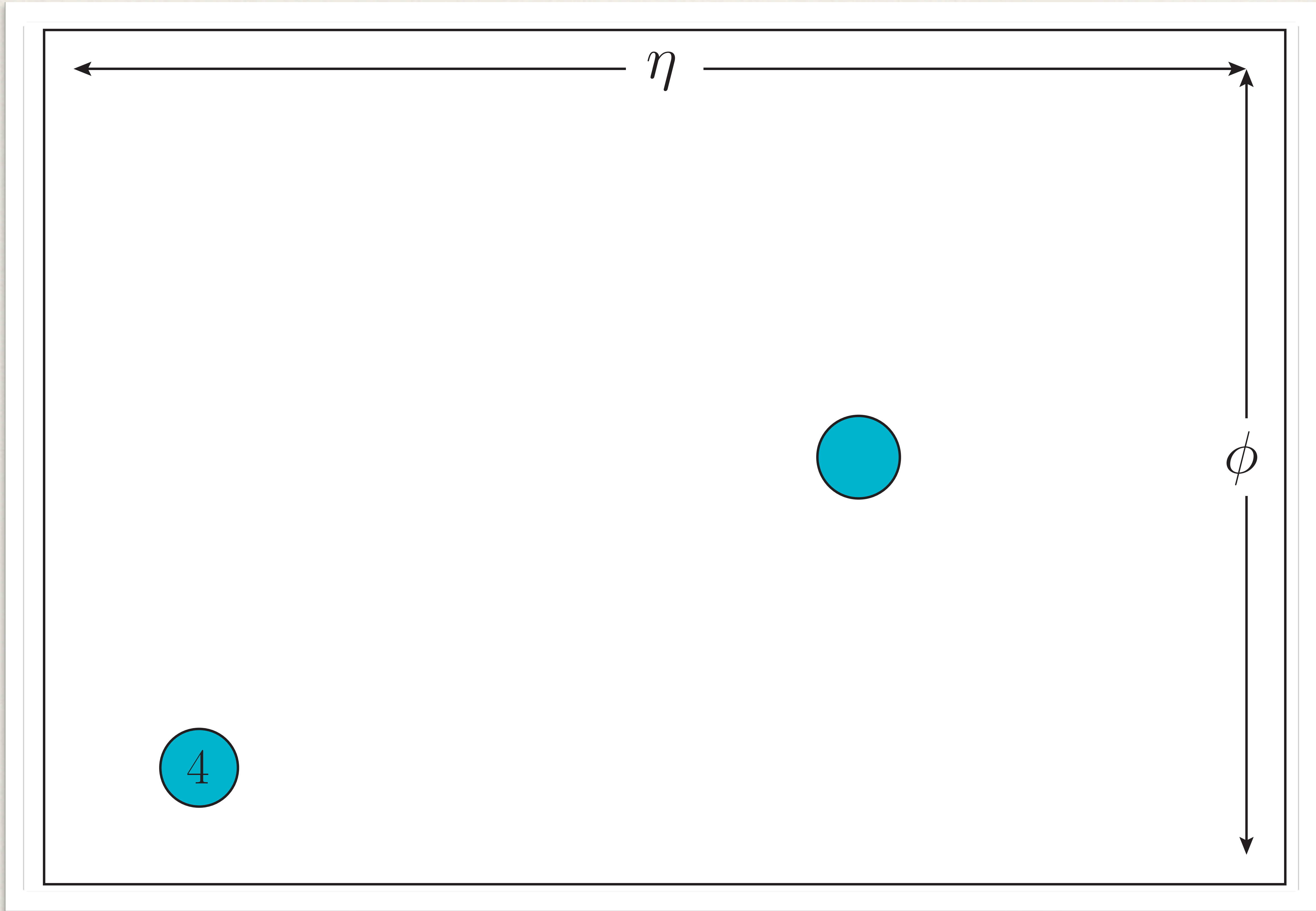


$$d_{12} < d_{13} < d_{23} < d_{(1,2,3)B} < d_{i4}$$

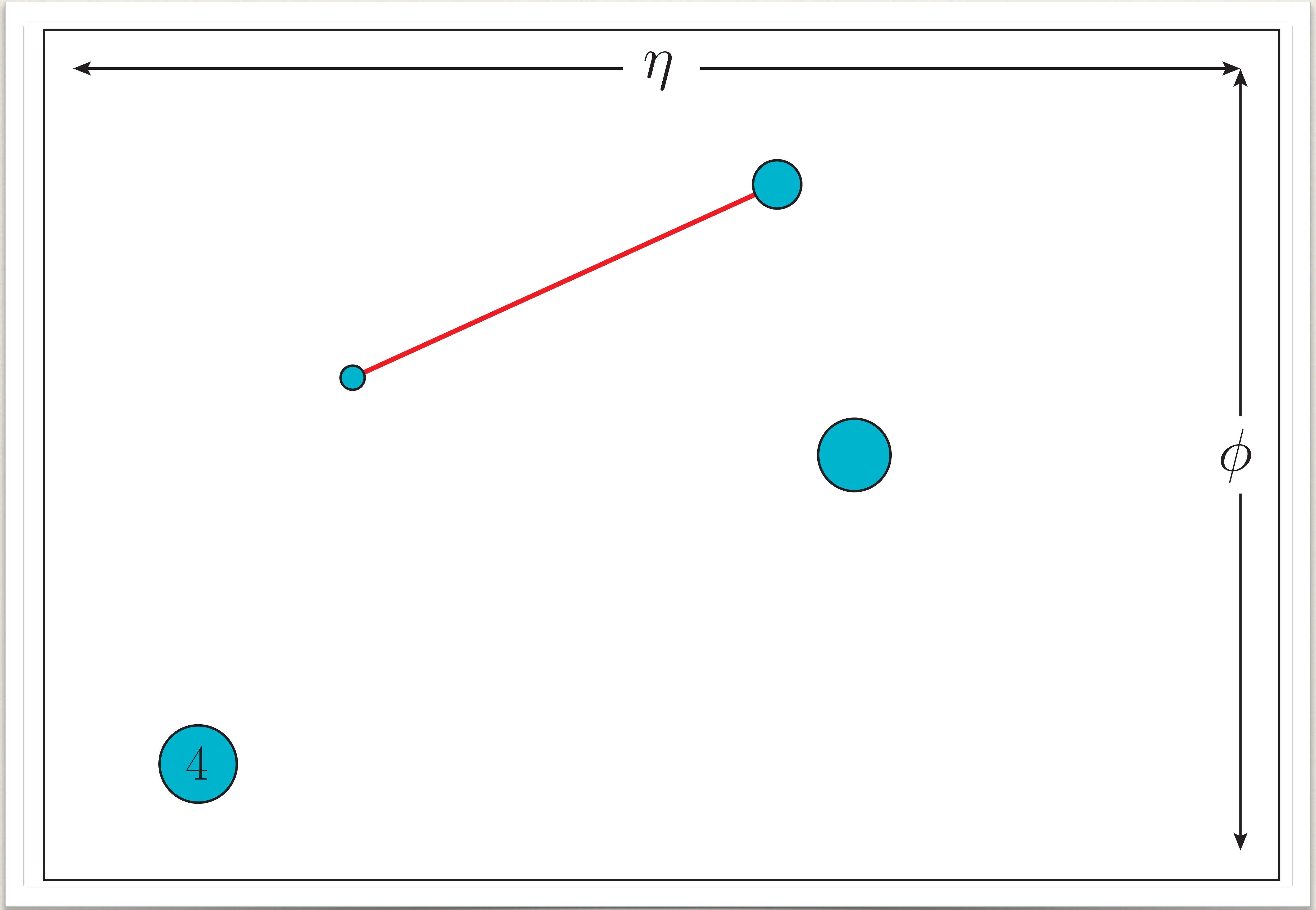


$$d_{12} < d_{(1,2)B} < d_{i4}$$


η
φ
4



Done!



η
ϕ
4