Jet Substructure at the LHC

with the Template Overlap Method

Jose Juknevich

מכון ויצמן למדע
WEIZMANN INSTITUTE OF SCIENCE

Based on a work with L. Almeida, M. Backovic, O. Erdogan,
S. Lee, G. Perez, G. Sterman & J. Winter

CERN BSM-TH Institute

June 28, 2012

Boosted objects

- ➔A heavy object (Z', KK gluon, ..) decays into light ones (t, H, W/Z), which then come out with a boost
- ➔At the LHC, a heavy particle (t, H, W/Z) can be produced at high transverse momentum by recoiling against another energetic object.

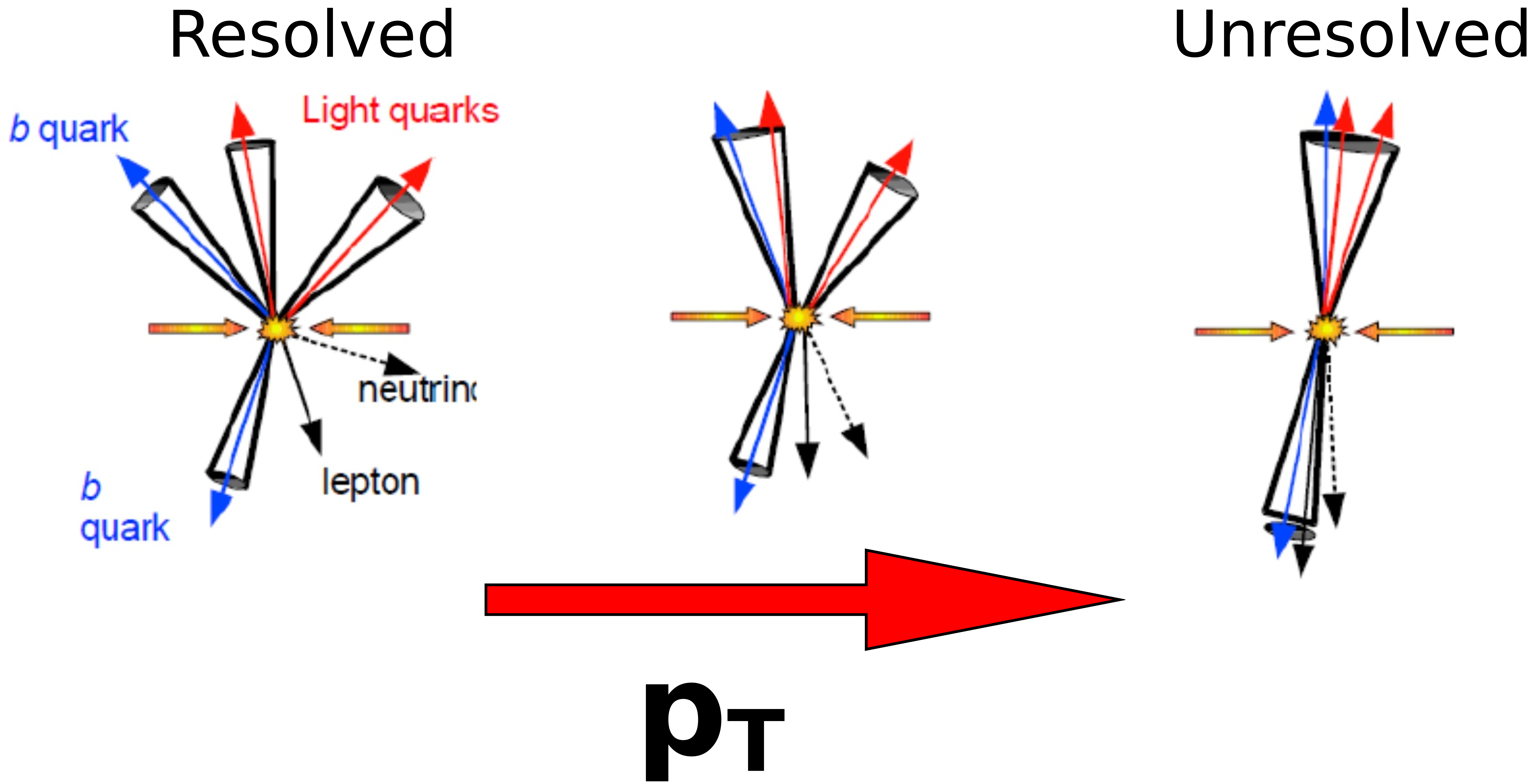
(See Christoph and José's talks)

- ➔Jets with high transverse momentum and high mass are a playground for testing perturbative QCD (pQCD)

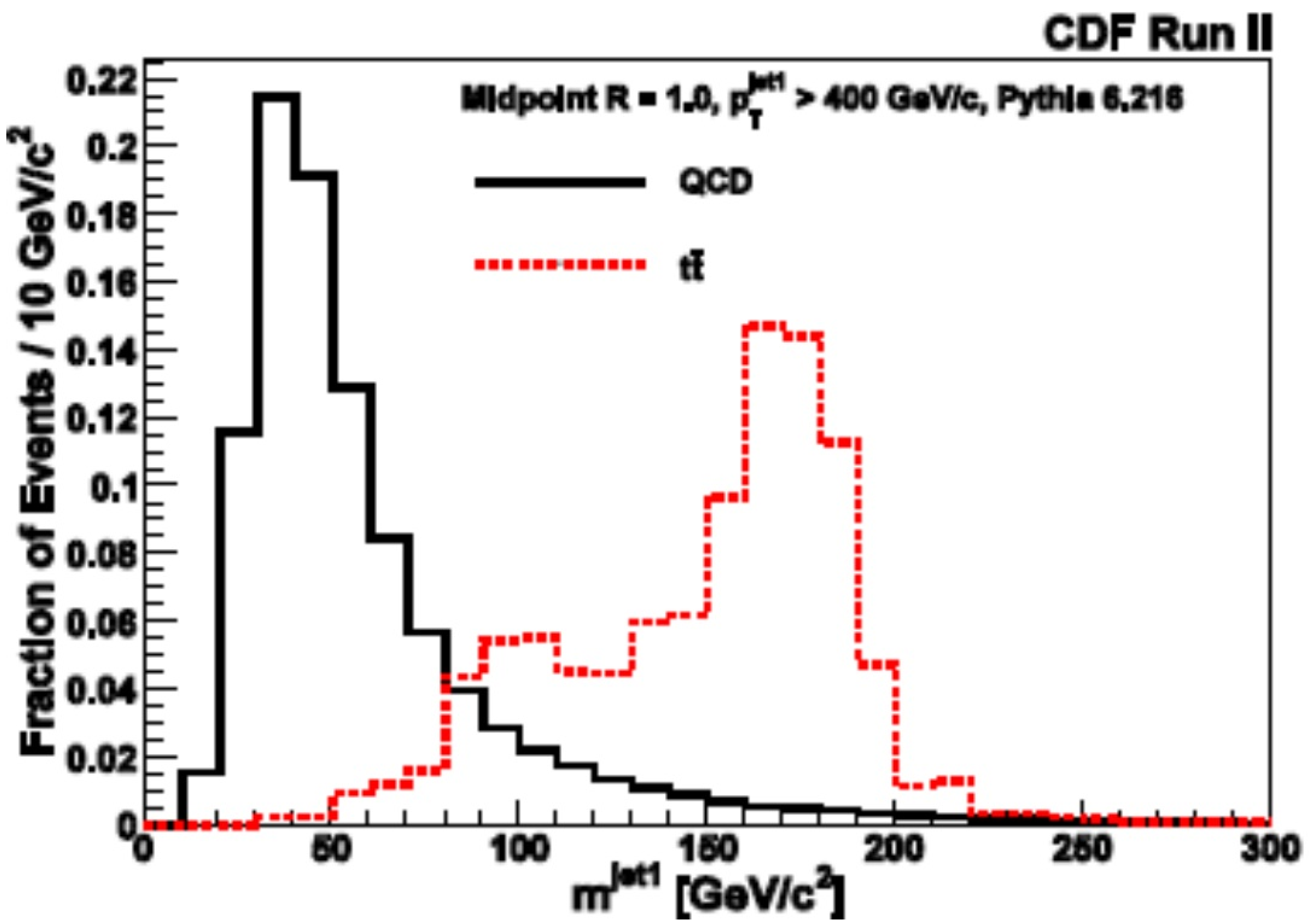
Boosted massive particles

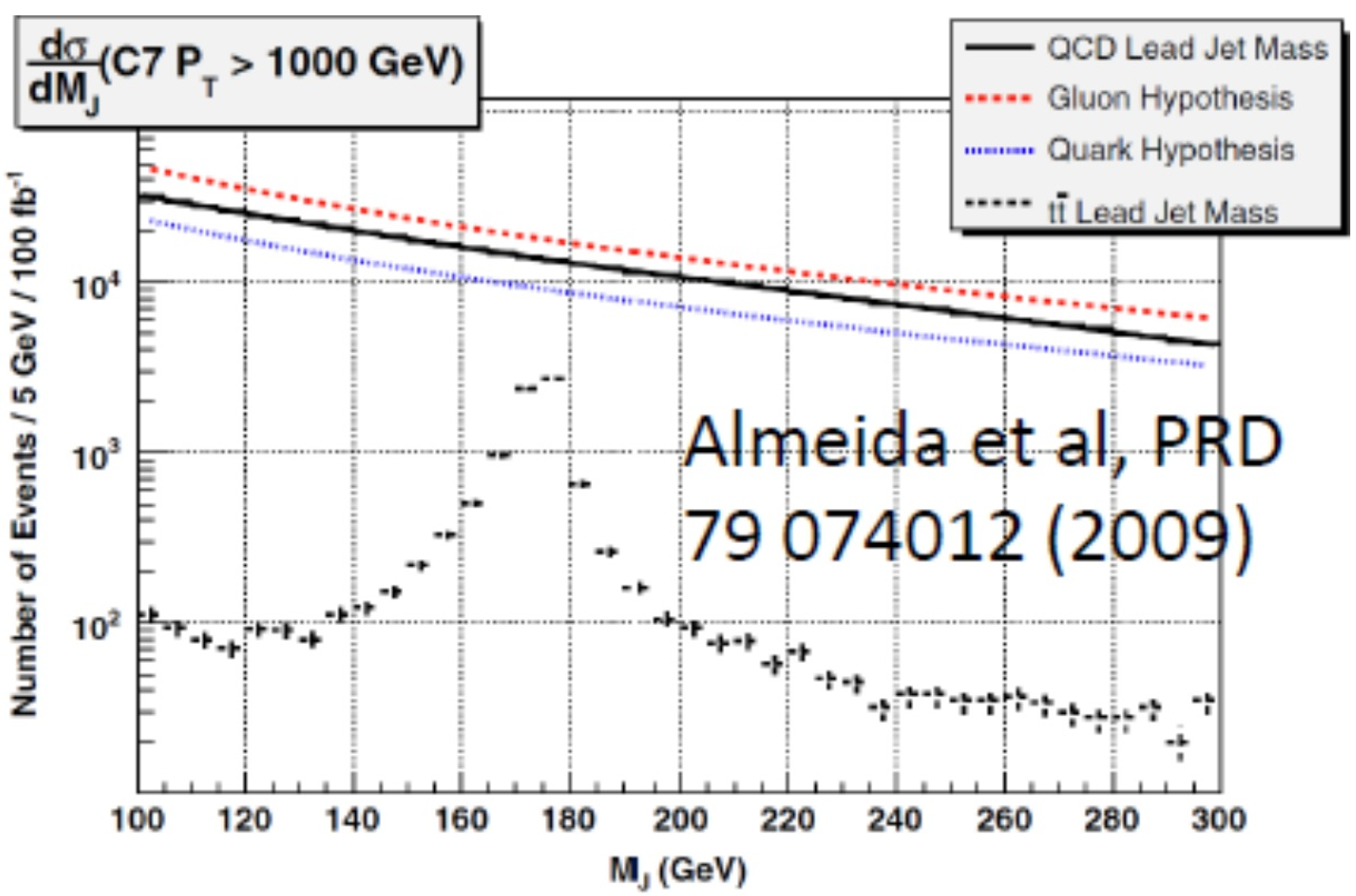
Classical “resolved” algorithms run into problems for highly boosted objects:

- At high enough pT, decay products not resolved by jet algorithms
- Isolation of leptons



But QCD is our enemy...



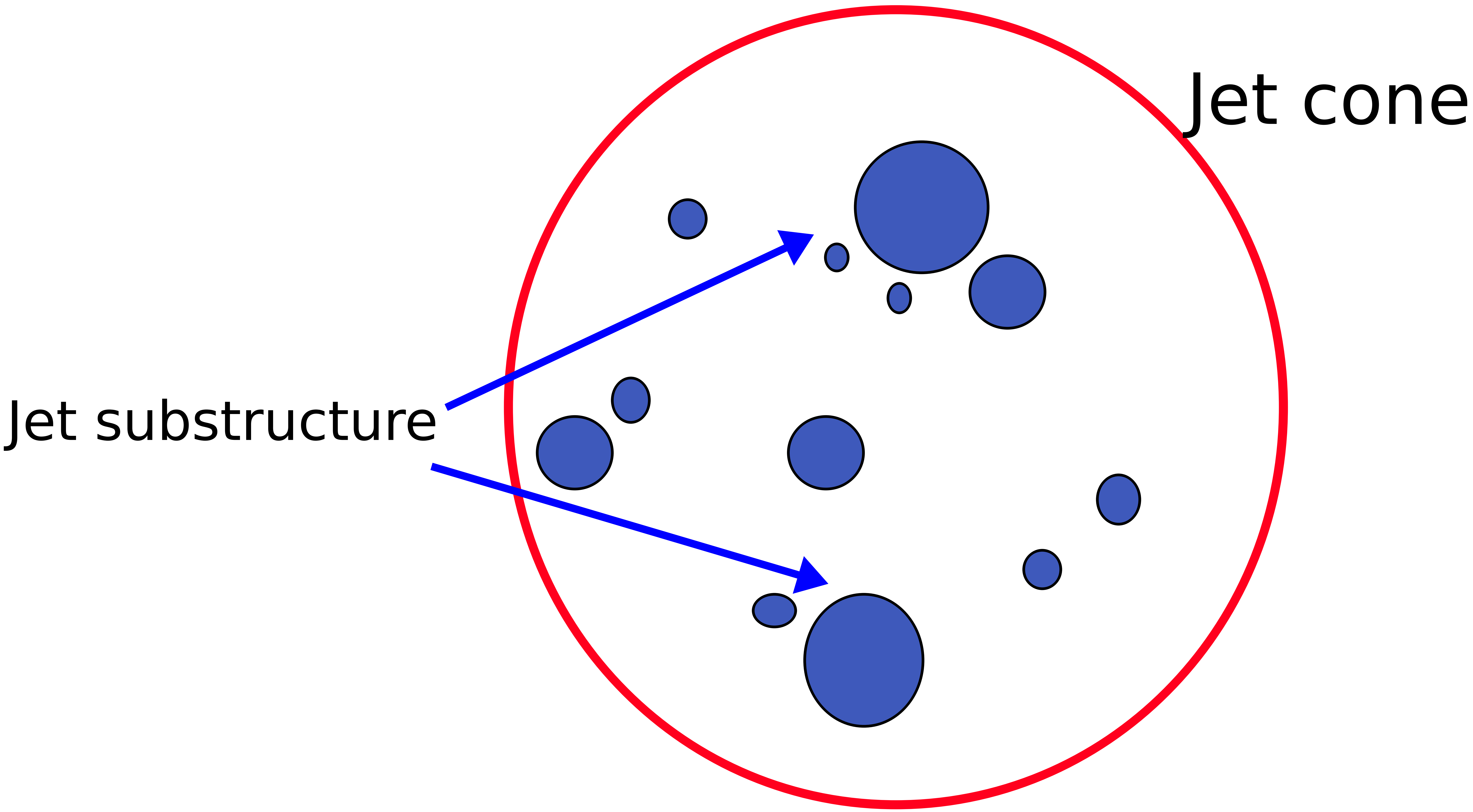


- Simplest observable is the jet mass
- Even such exotic states will coexist with a large tail of the mass distribution of light parton QCD jets

So how do we proceed?

Jet Substructure

Thus, it will generaly be necessary to use jet substructure systematically to identify the states that initiated the jets



Background rejection: basic approaches

Not an exhaustive list

- Subjet/grooming techniques

Filtering (Butterworth et al. 2008)
Pruning (Ellis et al. 2009)
Trimming (Krohn et al. 2010)

y-splitter (Butterworth et al.)
John Hopkins tagger (Kaplan et al.)
HEP Top tagger (Pehn et al.)
tree less approach (Jankowiak et al.)

- Energy flow methods

Template Overlap (Almeida et al. 2010)

Jet shapes (Thaler and Wang) (Almeida et al.)) (Gallicchio, Schwartz)

- Hybrids

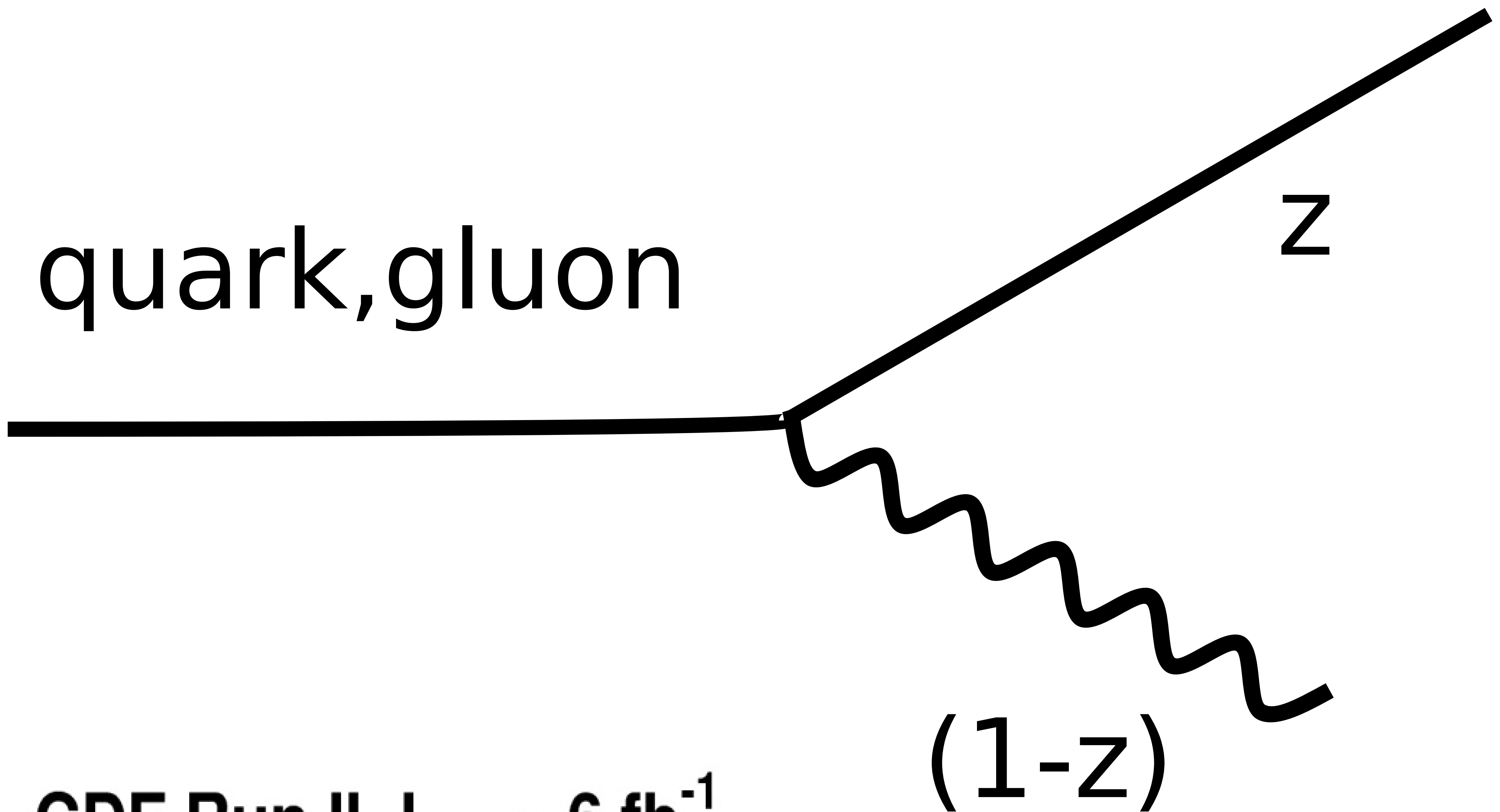
Dipolarity (Hook, Jankowiak,)
N-subjettiness (Thaler, Van Tilburg)

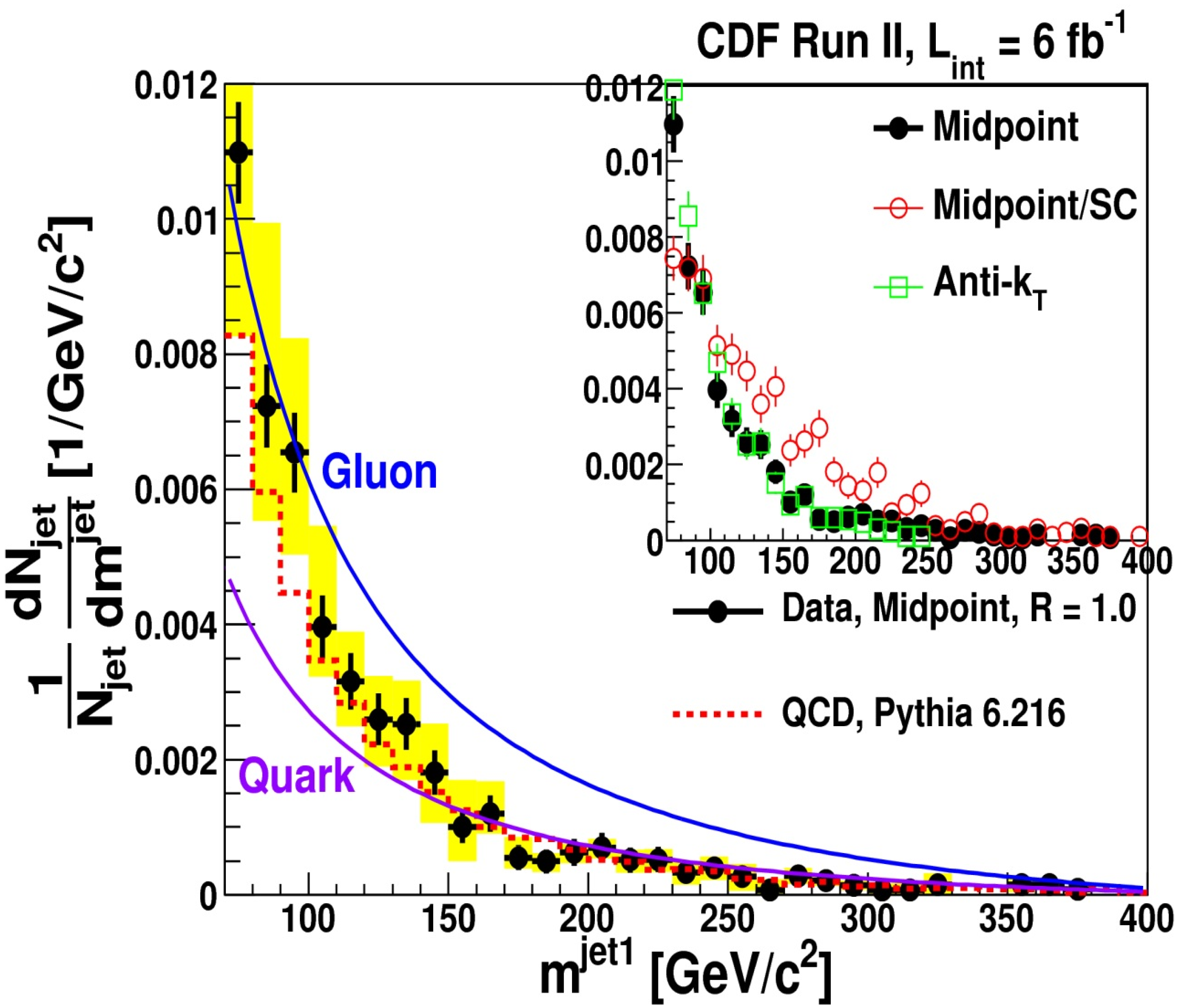
- And more ...

Shower deconstruction (Soper et al., 2011)
Multivariate (Gallicchio et al.)

Warm-up: QCD jet mass distribution

Almeida, Lee, Perez, Sterman, Sung (2008)



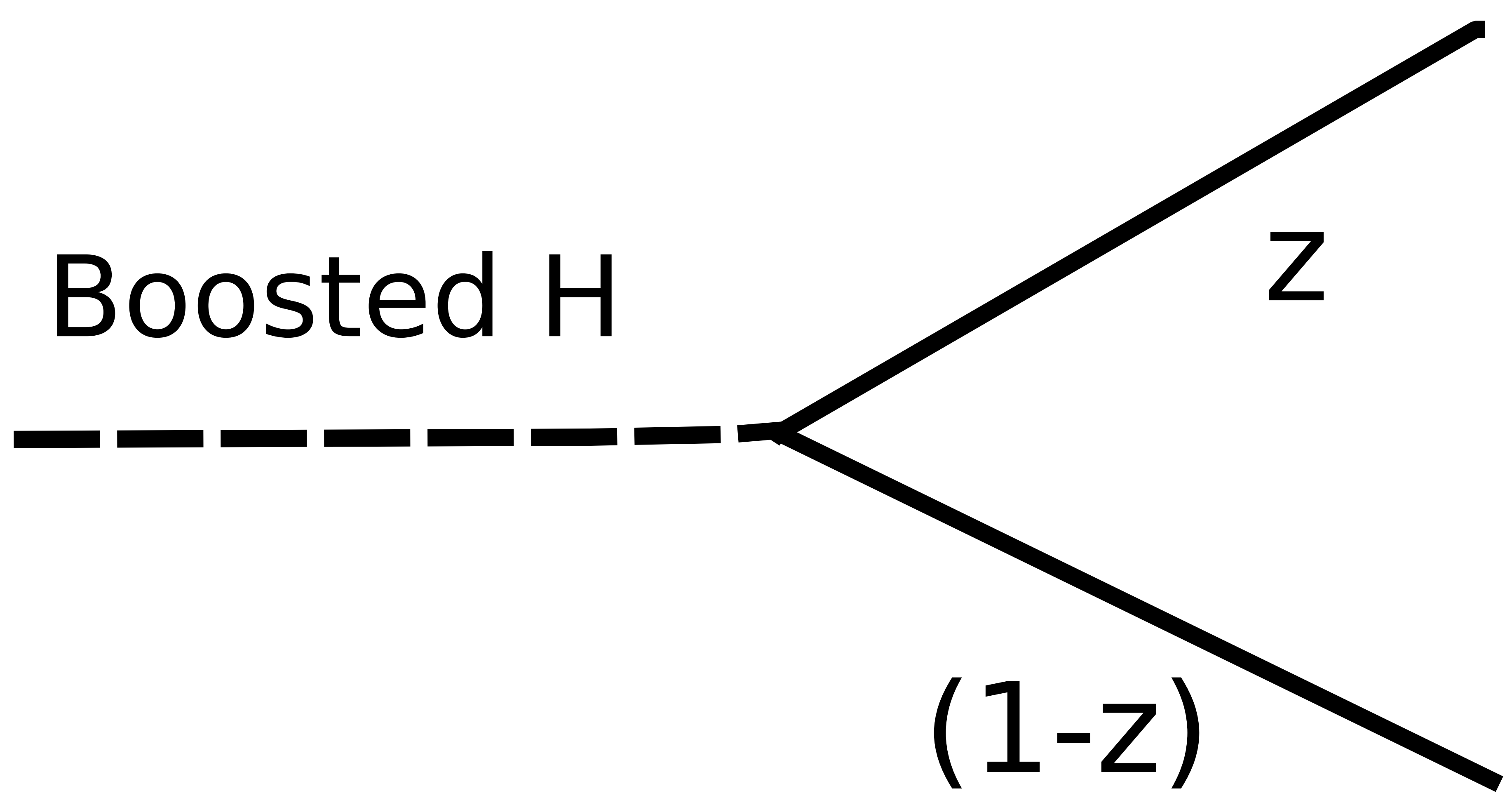


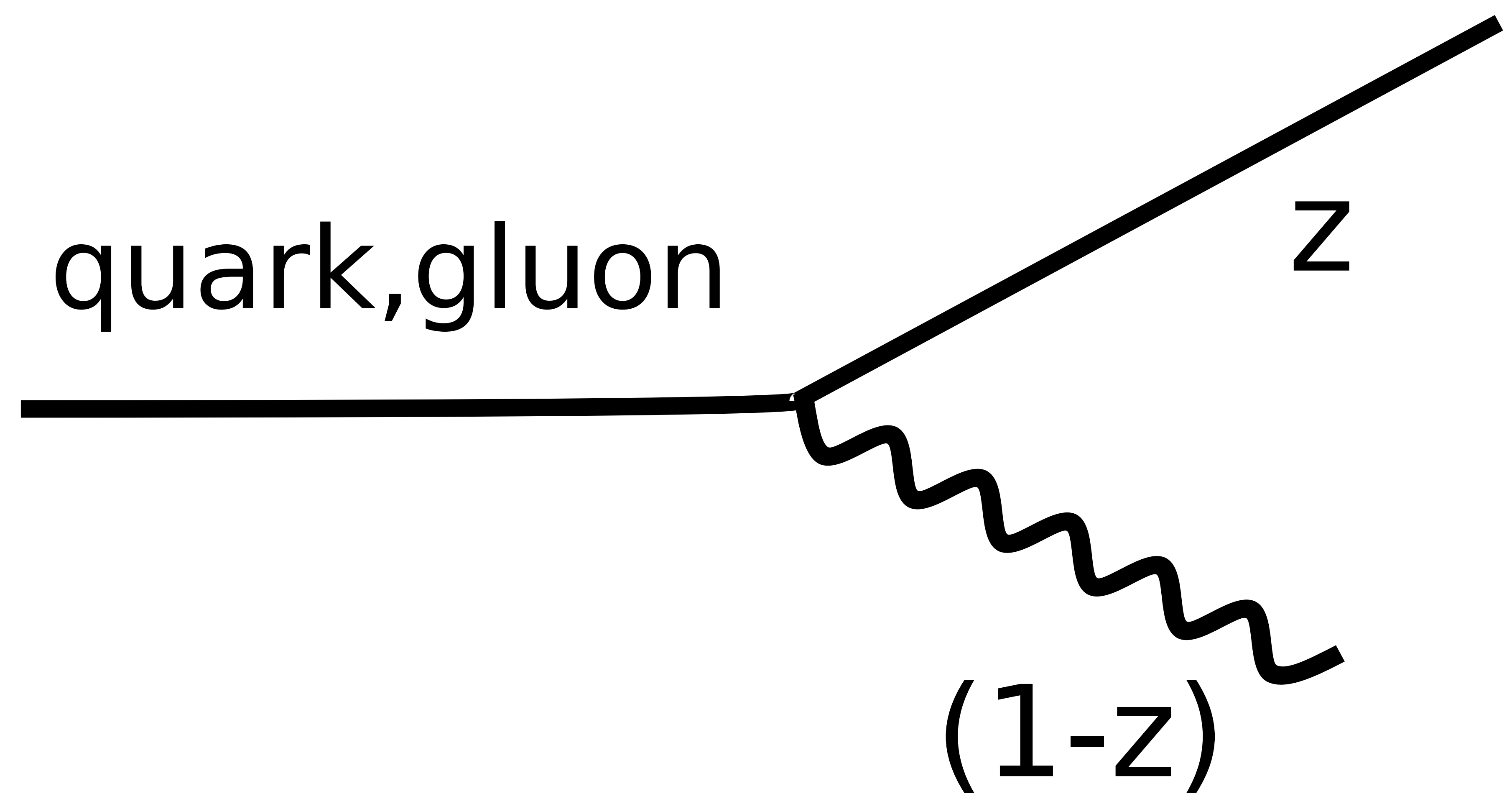
$$\frac{d\sigma}{dm_J^2} \propto \frac{C_F}{m_J^2} \ln\left(\frac{E^2 R^2}{m_J^2}\right)$$

Aceptable agreement with data

Beyond jet mass

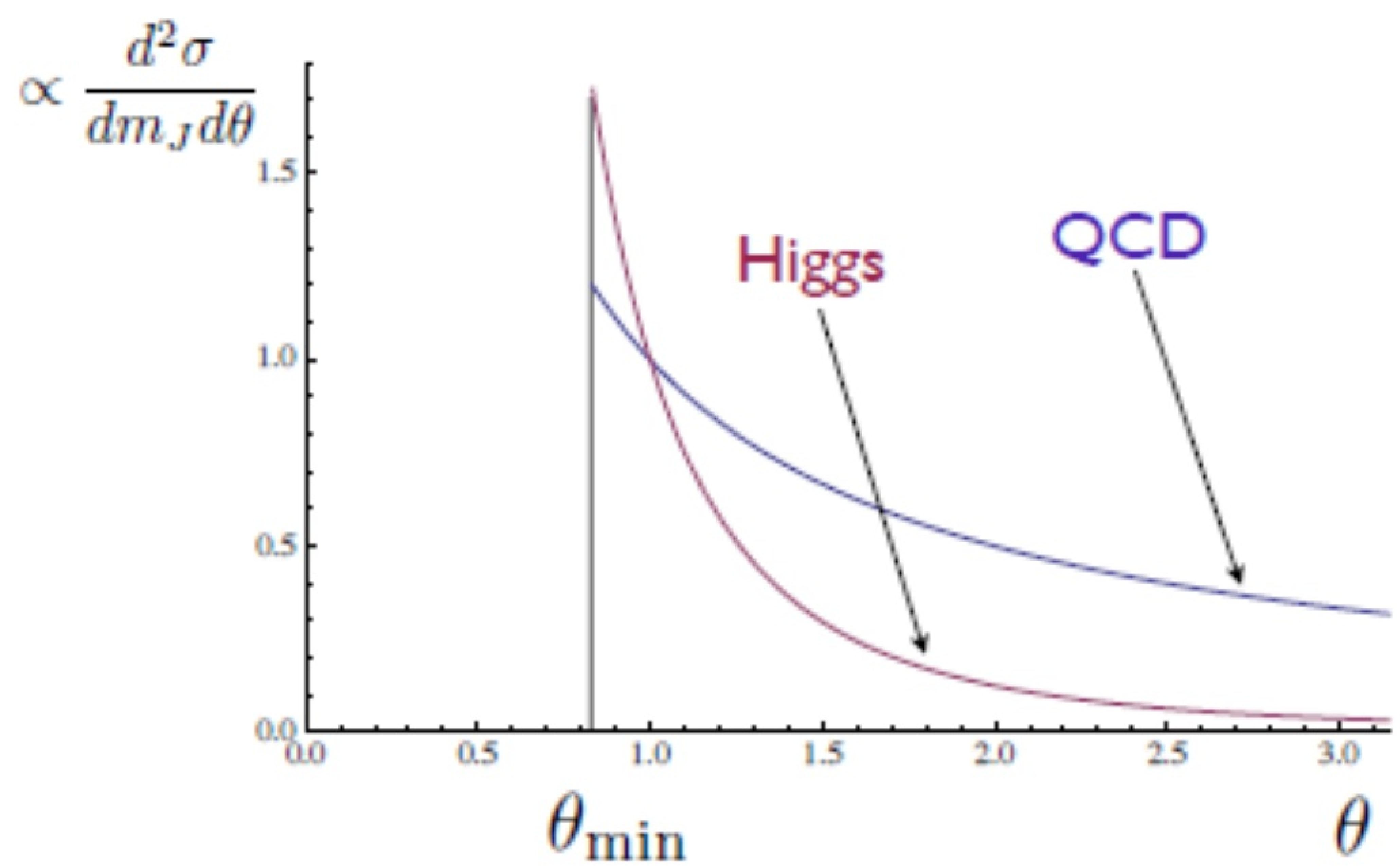
- One more variable to LO (Energy ratio or angle)
- Higgs decays are democratic, sharing energy evenly





QCD $$\frac{d^2\sigma}{dm_J^2 d\theta} \propto \alpha_s \frac{C_F}{m_J^2} \frac{1}{\theta}\bigg|_{R>\theta>\frac{m_J}{E_J}}$$

Higgs $$\frac{d^2\sigma}{dm_J^2 d\theta} \propto \delta(m_J^2 - m_H^2) \frac{1}{\theta^3}\bigg|_{R>\theta>\frac{m_J}{E_J}}$$



Almeida et al. 08

Need to understand energy flow inside jets

Guiding principles

- Particle decays are "democratic", sharing energy evenly
- Angular ordering implies that the majority of QCD radiation emitted by the color-singlet objects is within cones of radius ΔR (See Jason's talk for color radiation from colored objects)
- The spiky hard part of the signal should be insensitive to the diffuse QCD radiation from UE/Pileup

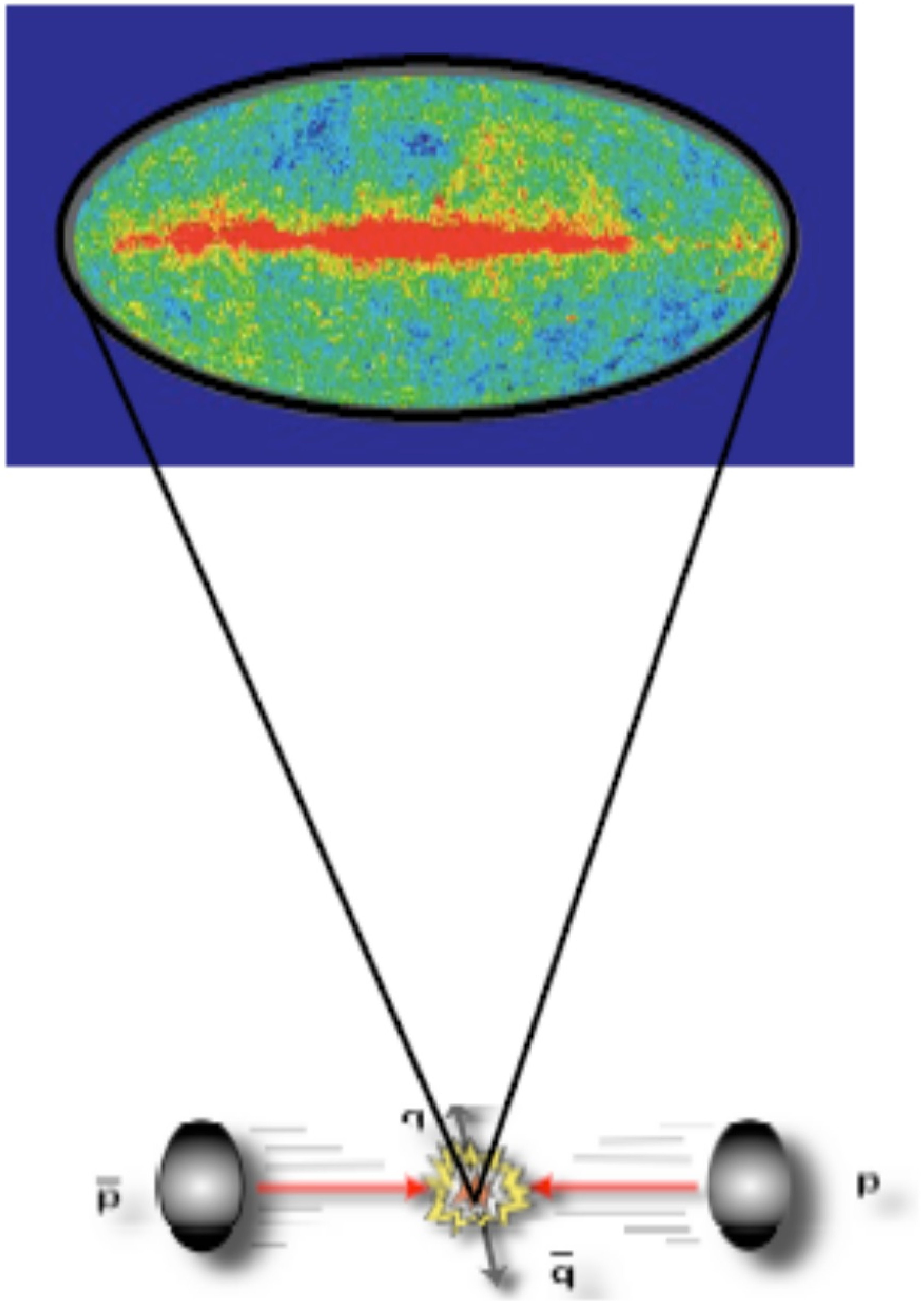
Can we use the partonic decay information, comparing directly to data, to tell us about the substructure of massive jets?

Template Overlap: basic idea

Almeida, Lee, Perez, Sterman, Sung (2010)

- We can characterize the signal via spikes of energy, that we can calculate in perturbation theory: templates

- Build catalog of all partonic boosted decays of our signal
- Search for a configuracion that give us a good match to the current jet.
- The best-matched partonic decay gives us information about the likelihood that the jet is signal or background



Template Overlap Method

Almeida, Lee, Perez, Sterman, Sung (2010)

Functional measure for the quantitative comparison of the energy flow of observed jets at high-pT with the flow of a selected sets (the templates) of partonic states

$$\langle j|f\rangle = \mathscr{F}\left[\frac{dE(j)}{d\Omega}, \frac{dE(f)}{d\Omega}\right]$$

Our Templates will be sets of partonic momenta

$$\sum_{i=1}^{N} p_i = P,\ P^2 = M^2$$

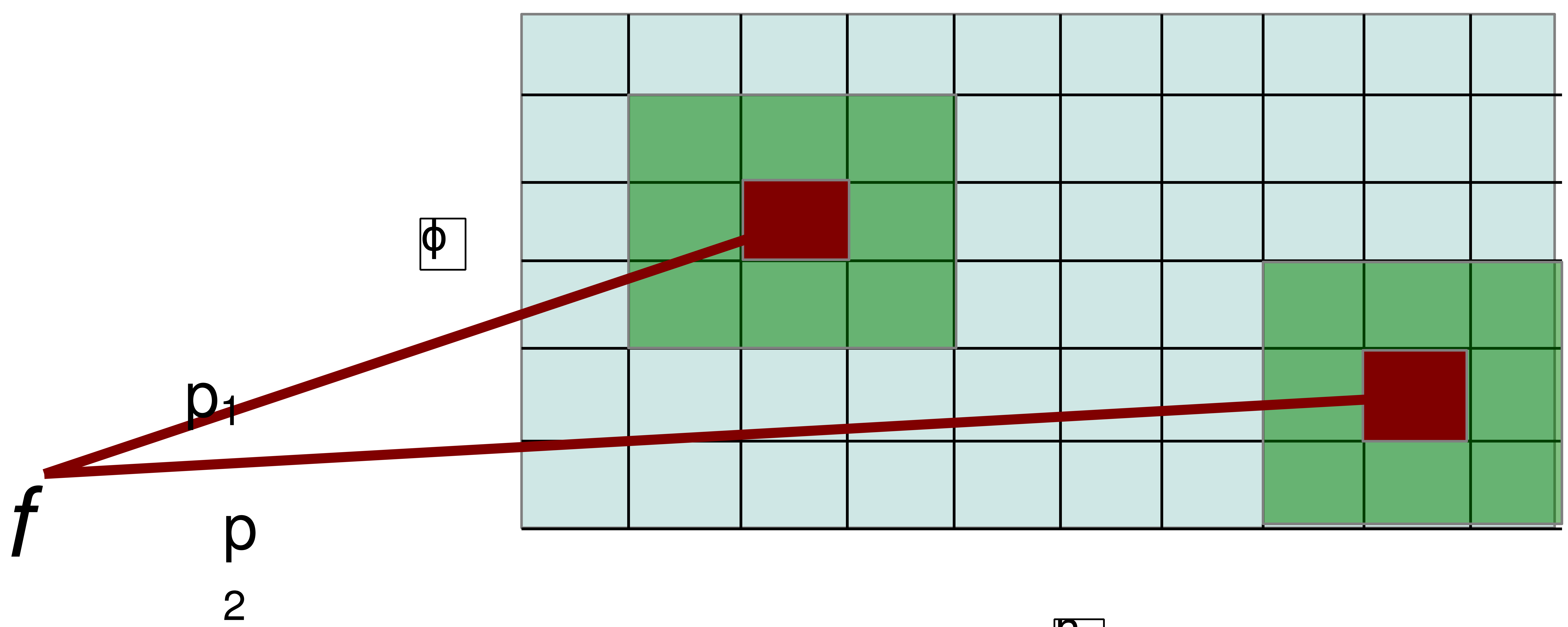
For a given j, we determine the template state f[j] for which the measure is maximized:

$$Ov(j, f) = \max_{\{f\}}\mathcal{F}(j, f)$$

Templates with data discretization

We compute the overlap between data state j and N-body template f from the sum of the energy in the nine cells of state j surrounding and including the cells of template f

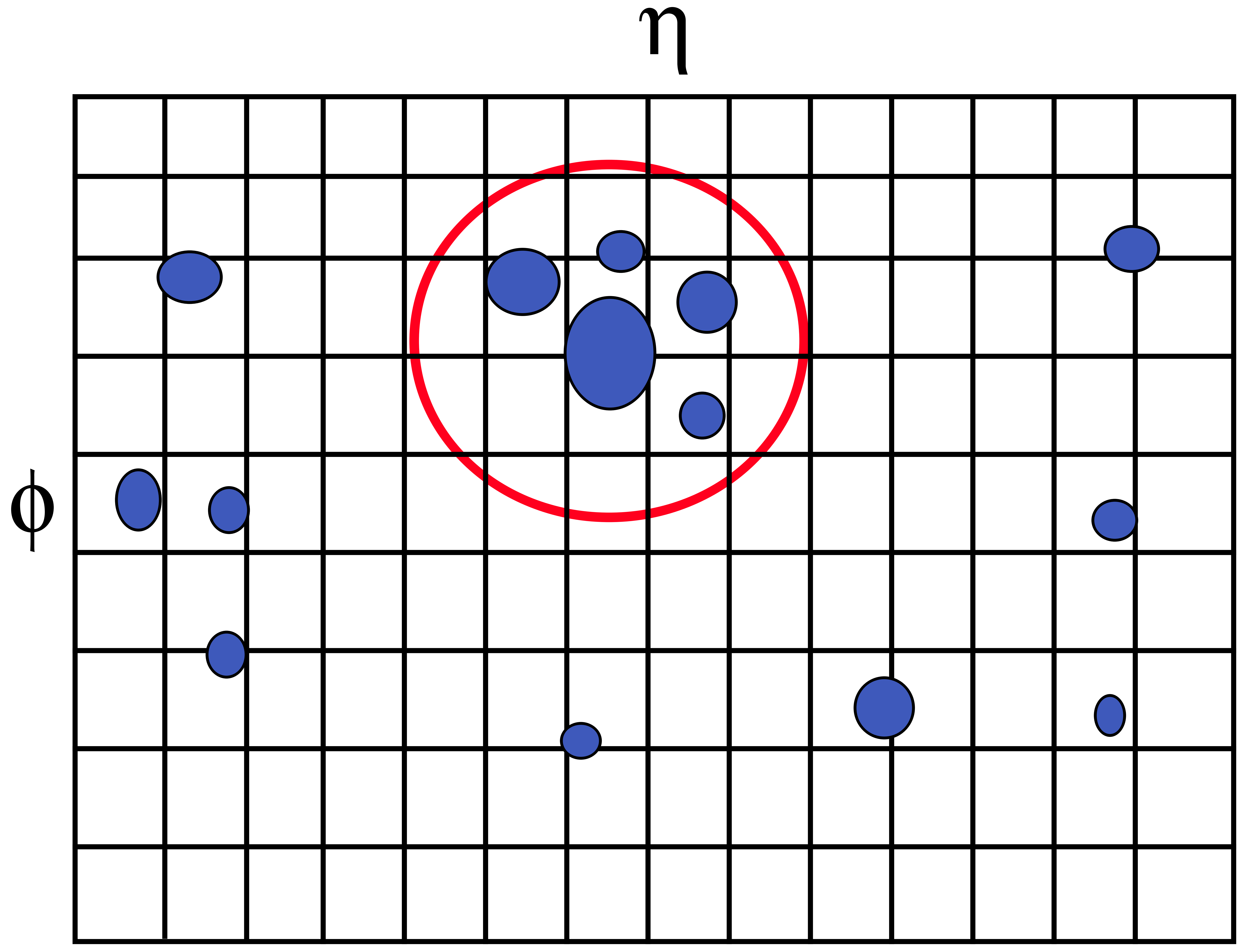
$$Ov_N(j,f) = \max_{\tau_N^{(R)}} \exp\left[- \sum_{a=1}^{N} \frac{1}{2\sigma_a^2} \left(\sum_{k=i_a-1}^{i_a+1} \sum_{l=j_a-1}^{j_a+1} E(k,l) - E(i_a, j_a)^{(f)} \right)^2 \right]$$



Initial jet selection

Build anti-k_T jets with large cone sizes (R=1.0, 1.2, 1.4)

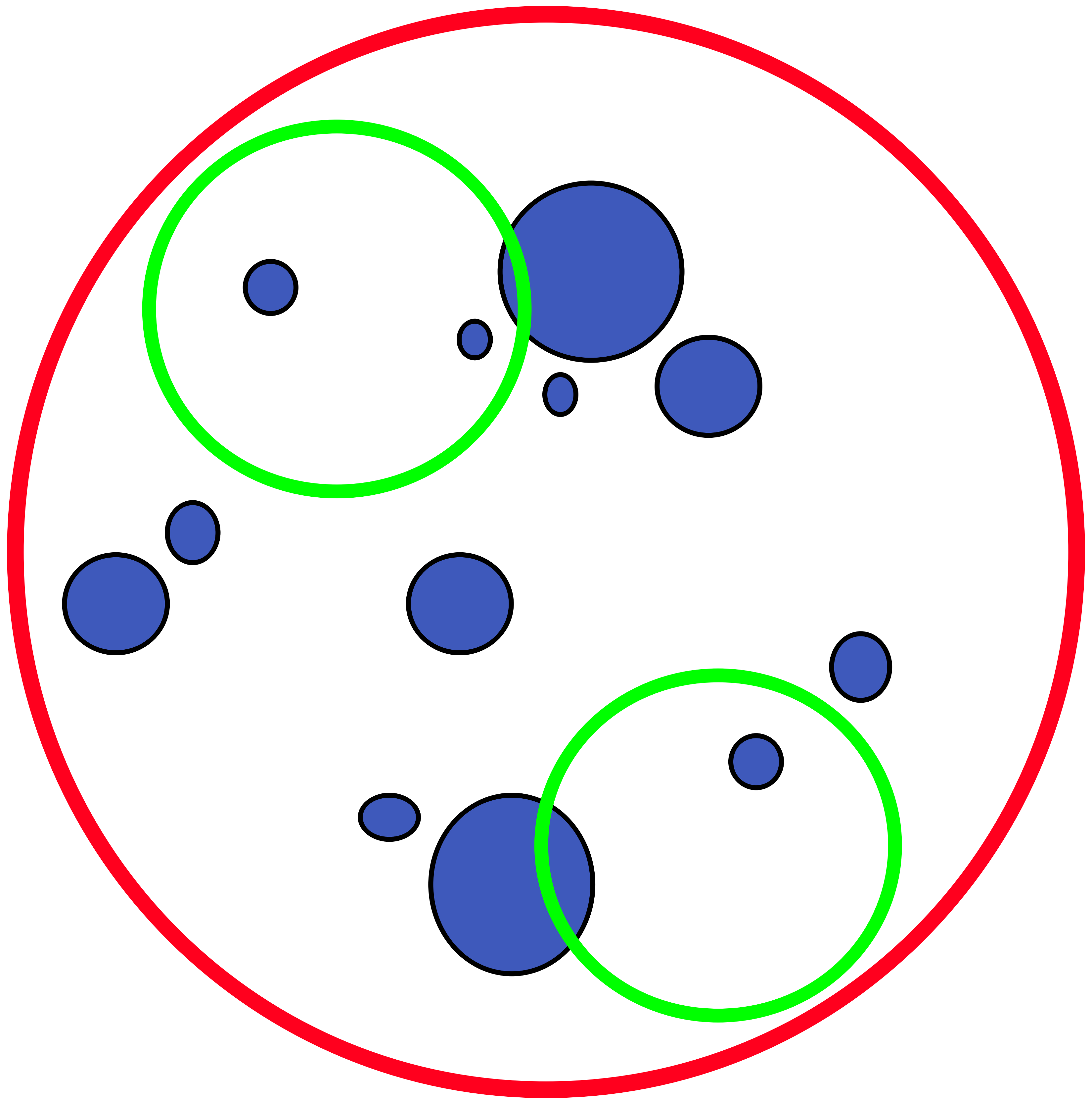
Impose jet selection cuts
pT_{jet} > 200 GeV
|eta| <2.5
....



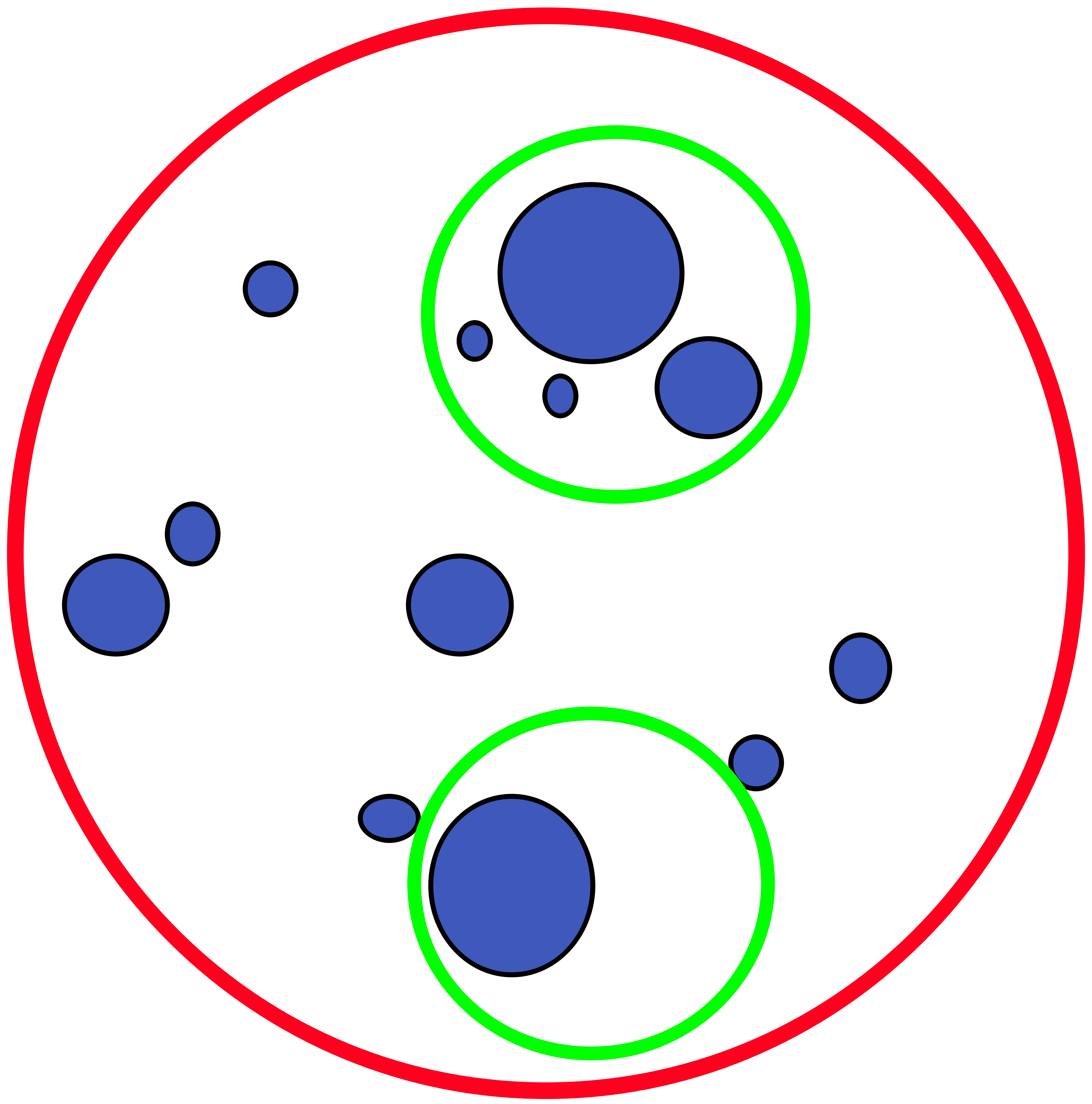
Template Overlap: implementation

Build catalog of all boosted partonic decays of the signal

Check to see if there is a template that matches the current jet



Poor overlap Ov ~ 0



Good overlap Ov ~ 1

$$Ov \sim \exp\left(-\frac{1}{2E_t^2}\sum_N (\sum_{j\in t} E_j - E_t)^2\right)$$

Application of the template method

Highly boosted signal, 3-pronged

top tagging

Mildly boosted signal, 2-pronged

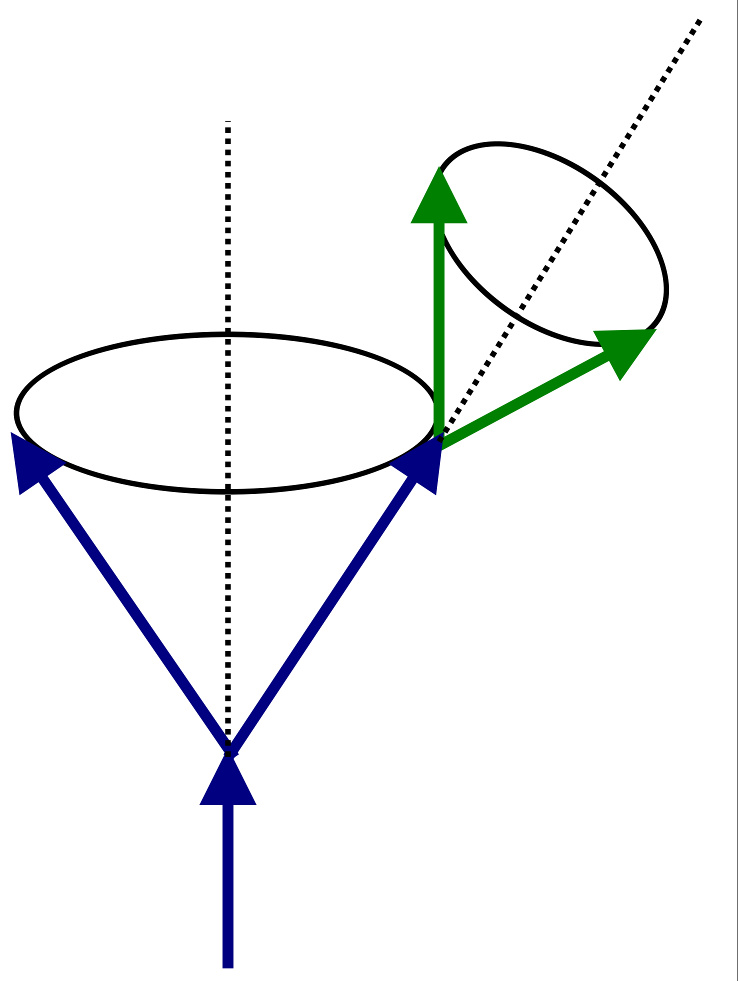
Search for boosted Higgs, in VH channel

Three-particle templates and top decay

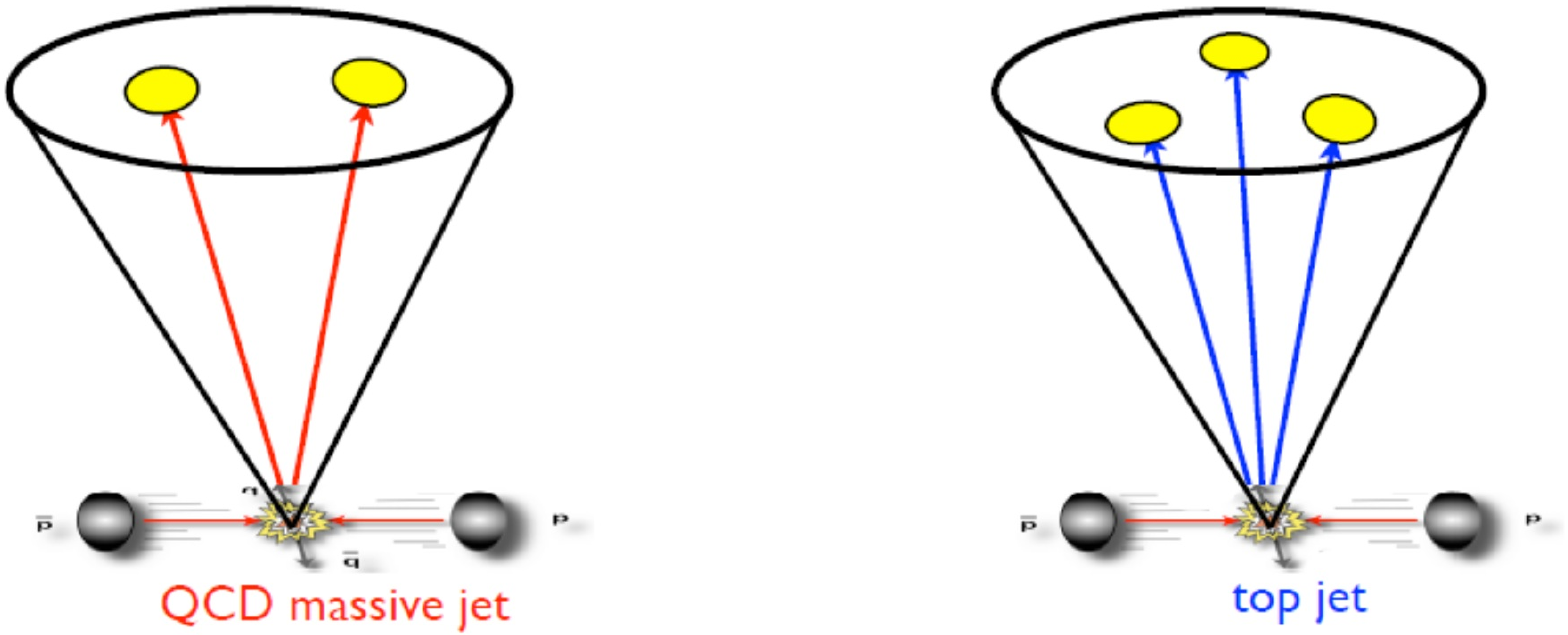
Almeida, Lee, Perez, Sterman, Sung (2010)

- At LO, top decay has a simple three-body kinematics

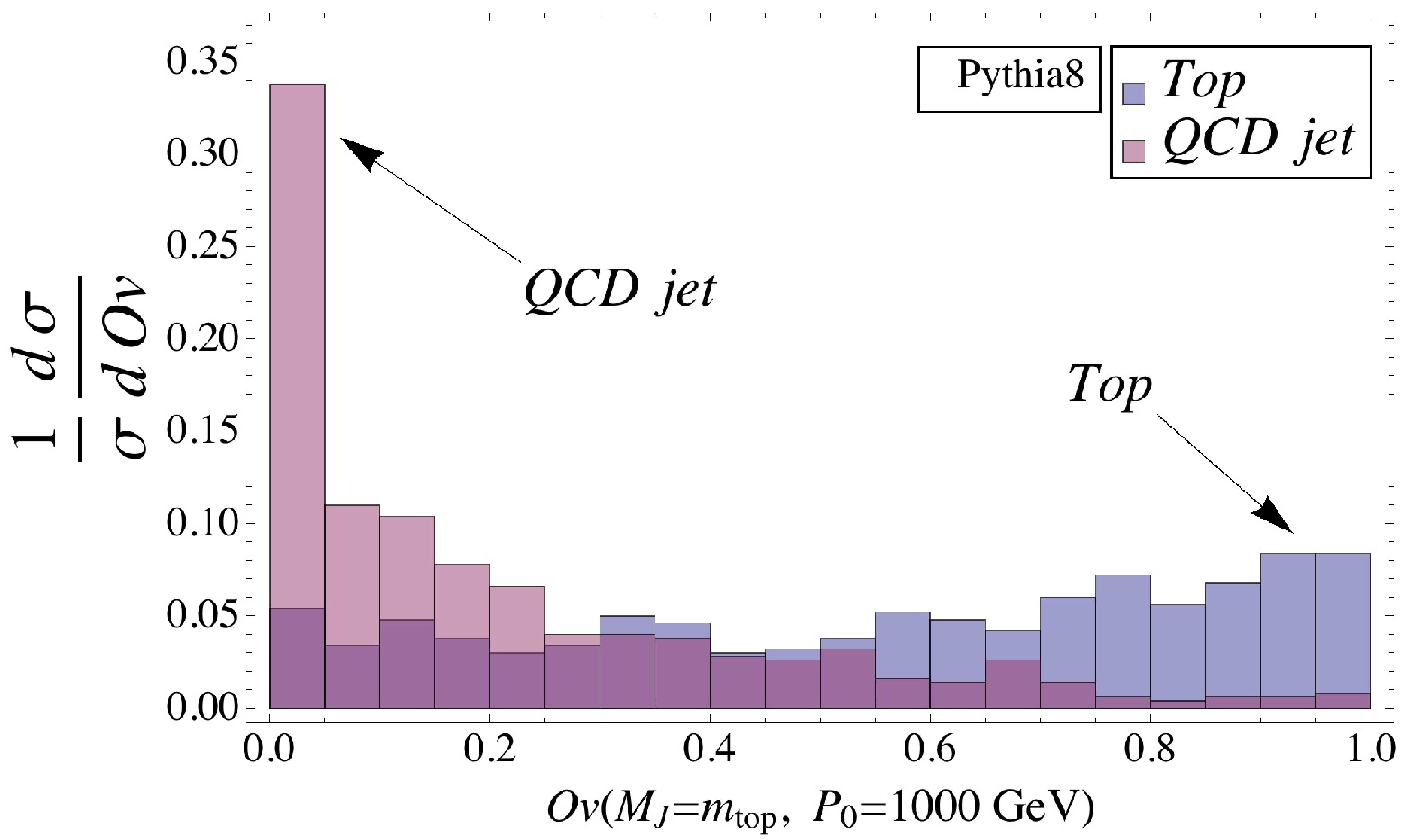
$$t \to b + W \to b + q + \bar{q}. \qquad (p_q + p_{\bar{q}})^2 = M_W^2$$



- While we expect high mass, QCD jets have a two-subjet topology



Three-particle templates and top decays



Jet mass and pT:

160 GeV < m_J < 190 GeV,

950 GeV < E_J < 1050 GeV

Jets found with anti-kt algorithms D=0.5

Can be combined with jet shapes (planar flow, pull) to distinghuish between many three-jet events with large overlap.

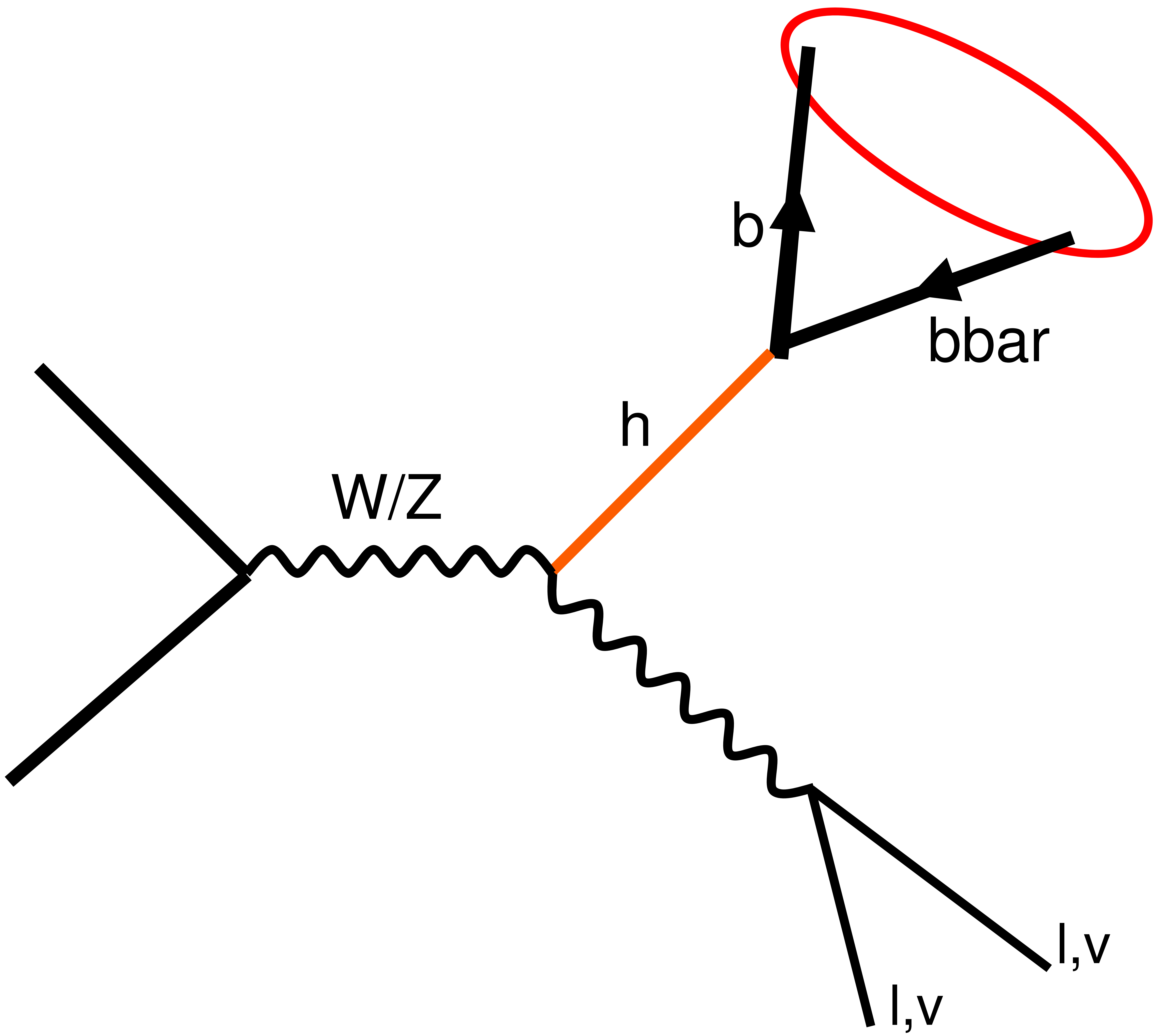
eff. 13 %, fake 1%

Rejection power ~O(10^2)

Boosted Higgs searches

Butterworth, Davison, Rubin, Salam (2008)

Look for associated VH production
Use the leptonic mode of the W/Z to suppress the QCD background



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

$$\Delta R \sim 2m_H/p_T$$

$$p_T^j = 200\,\mathrm{GeV}$$

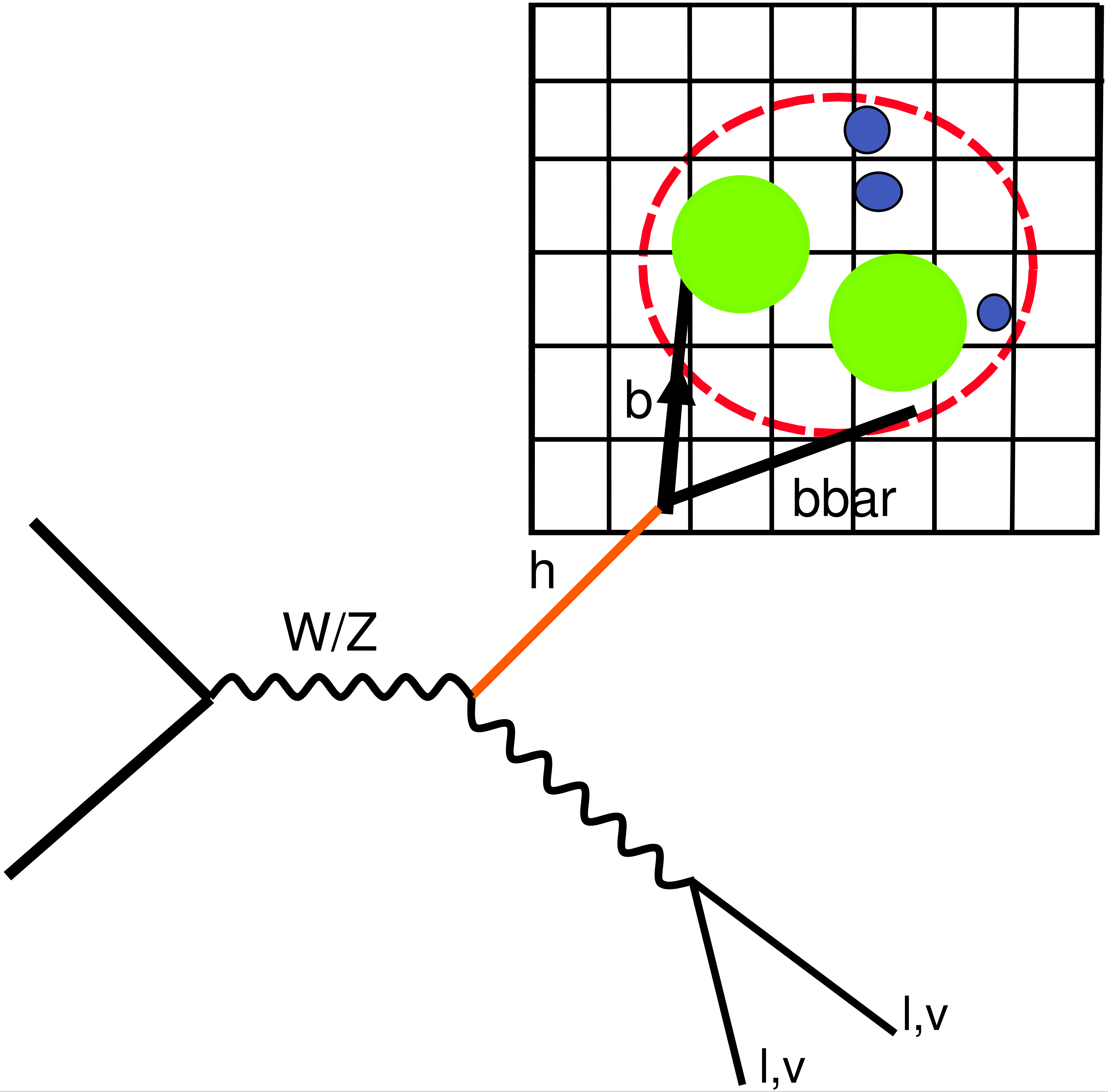
$$\Rightarrow \Delta R \sim 1.2$$

Typical jet size

$$\Delta R \sim 1.2$$

Boosted Higgs searches

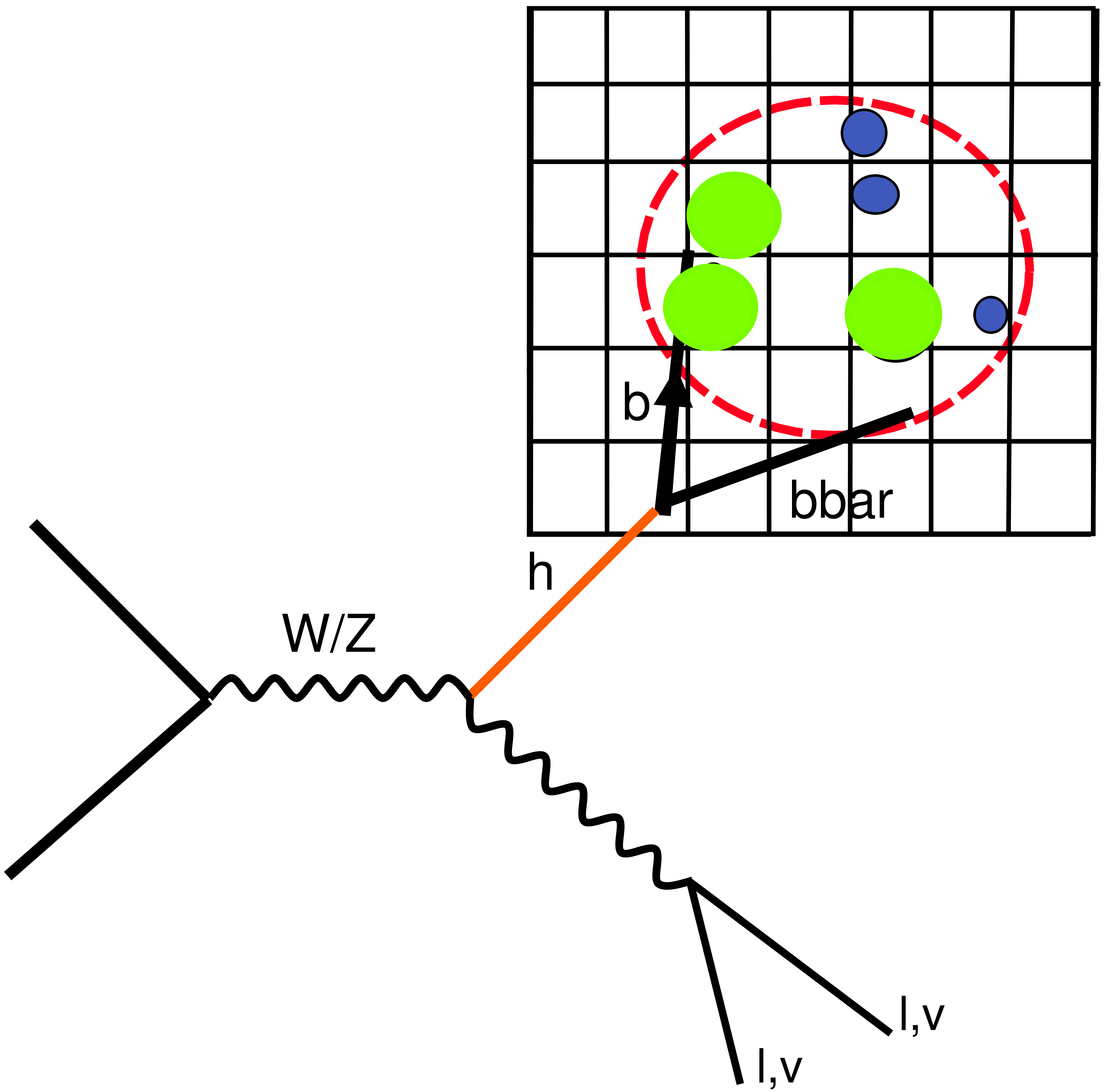
Almeida, Lee, Perez, Sterman, Sung (2010)
Almeida, Erdogan, JJ, Lee, Perez, Sterman (2011)



Use 2 body templates find hard substructure

Boosted Higgs searches

Almeida, Erdogan, JJ, Lee, Perez, Sterman (2011)

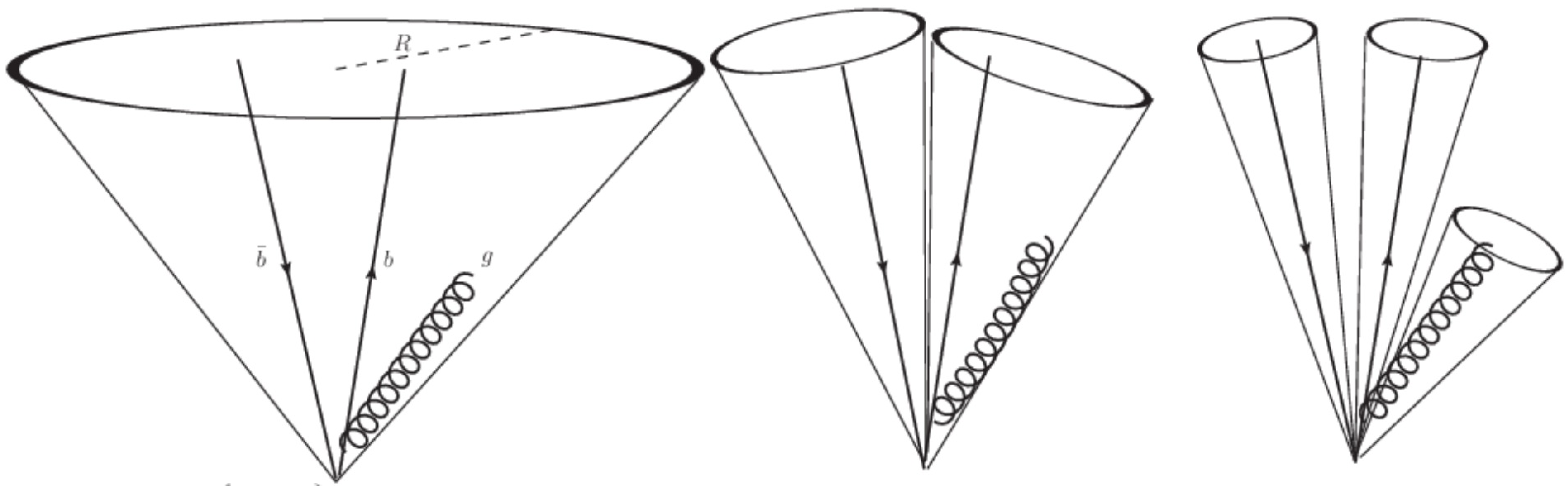


Templates with more than the minimum number of particles allow us to resolve finer details of the substructure

Smaller cone insensitive to incoherent noise (Pileup/UE)

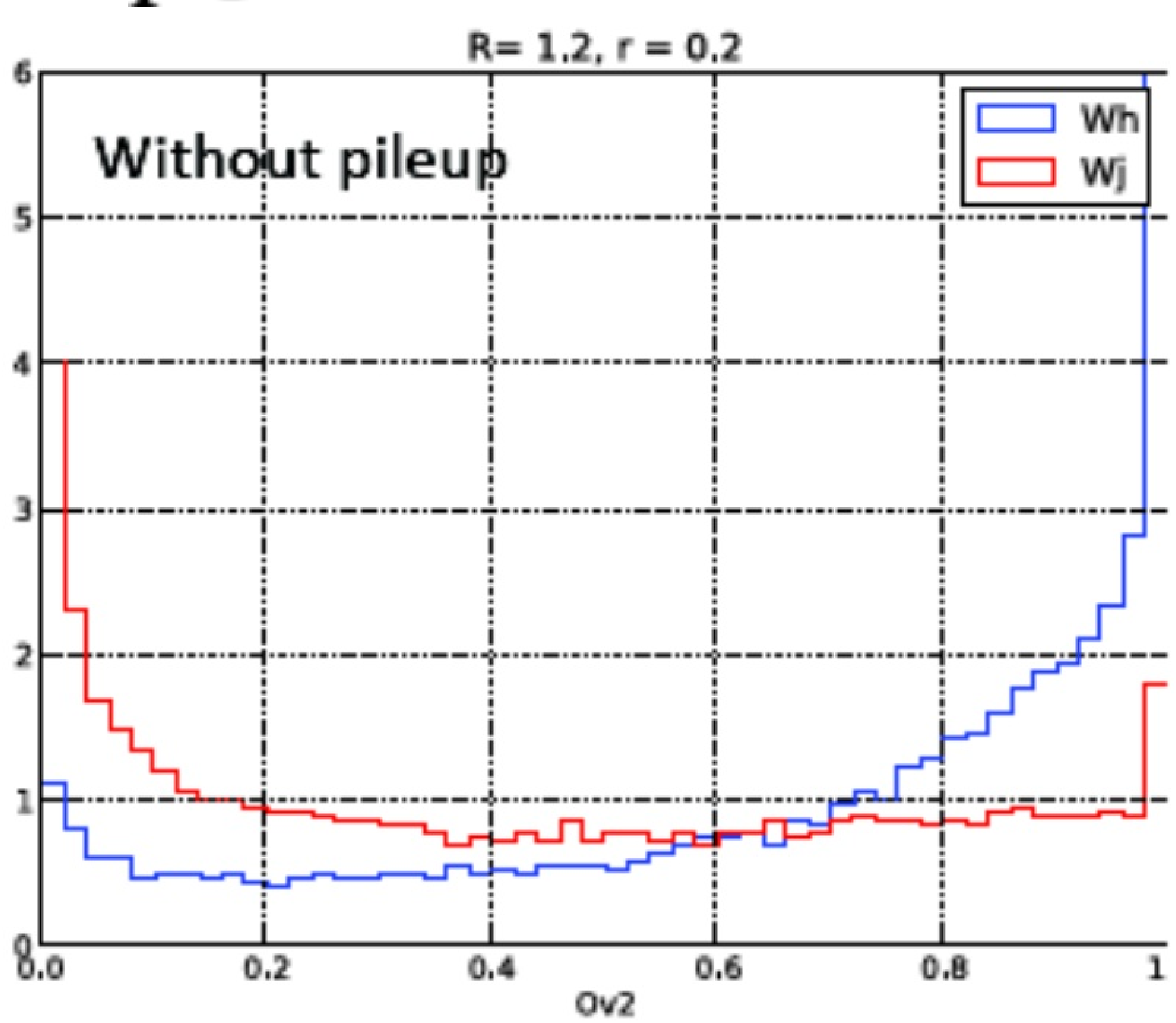
Boosted Higgs searches

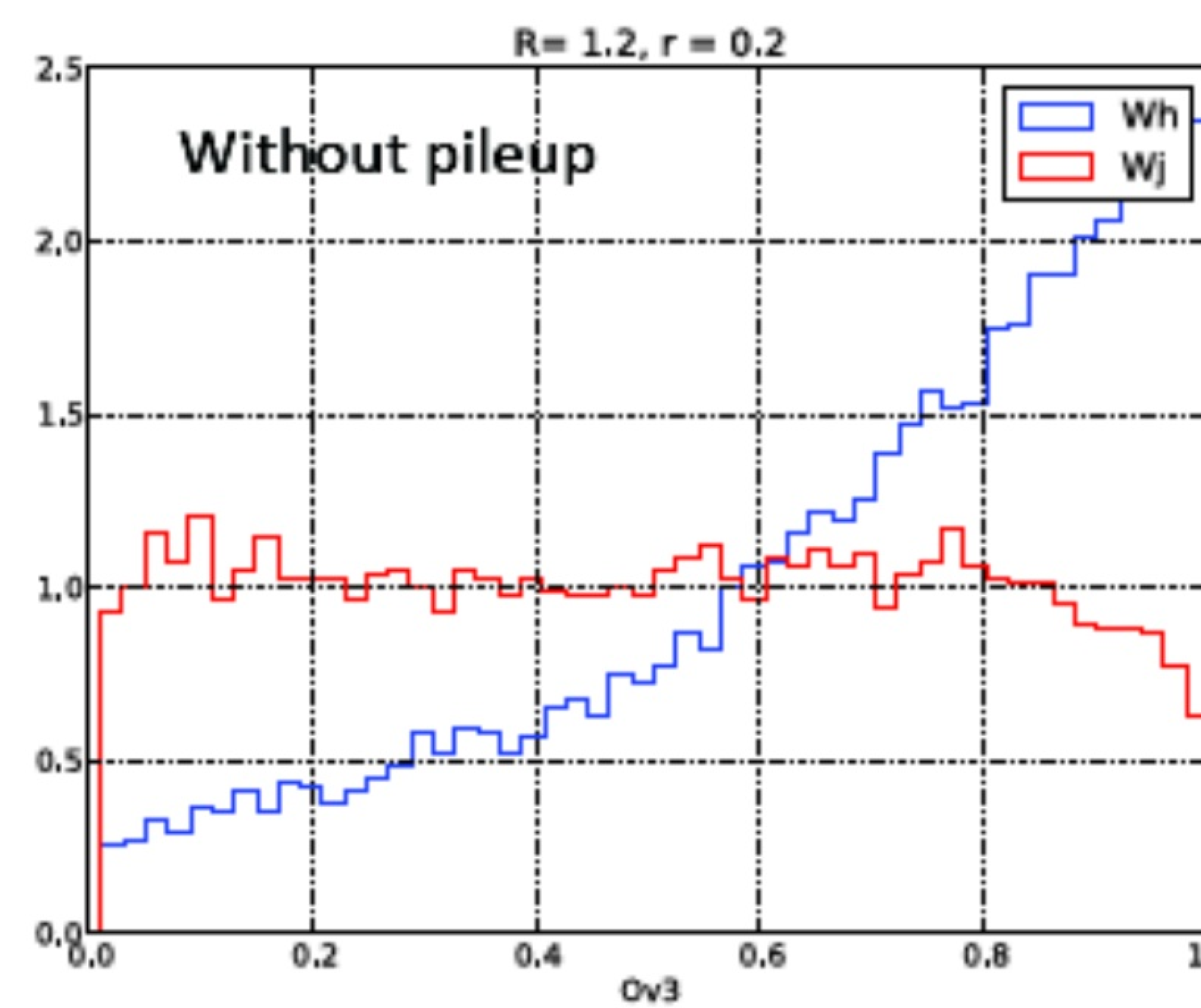
Backovic, JJ, Perez, Winter in preparation



$$|f\rangle = |h\rangle^{(\mathrm{LO})} = |p_1, p_2\rangle \qquad |f\rangle = |h\rangle^{(\mathrm{NLO})} = |p_1, p_2, p_3\rangle$$

$$p_T \sim 200\,\mathrm{GeV} \qquad p_T \sim 200\,\mathrm{GeV}$$



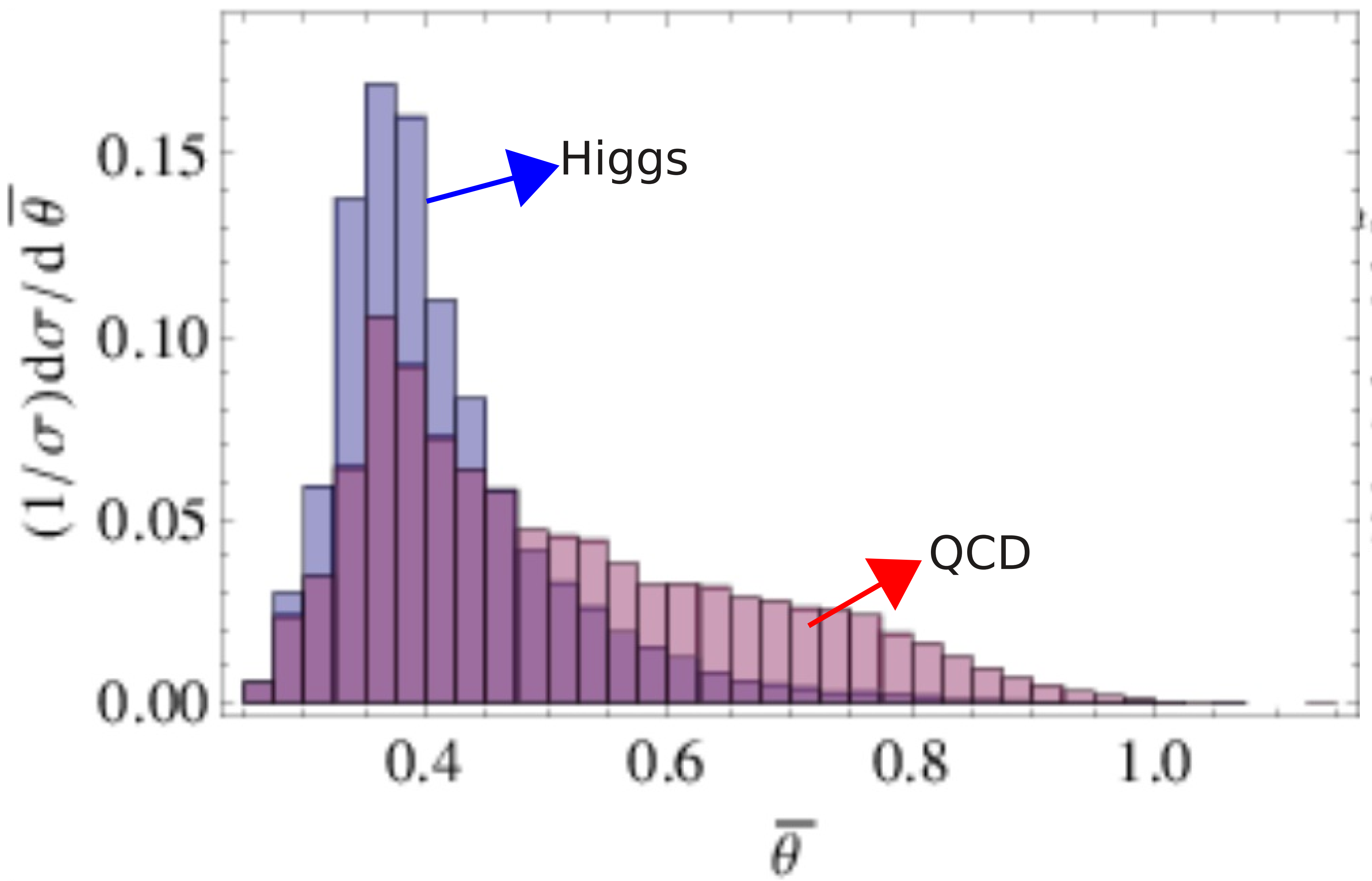


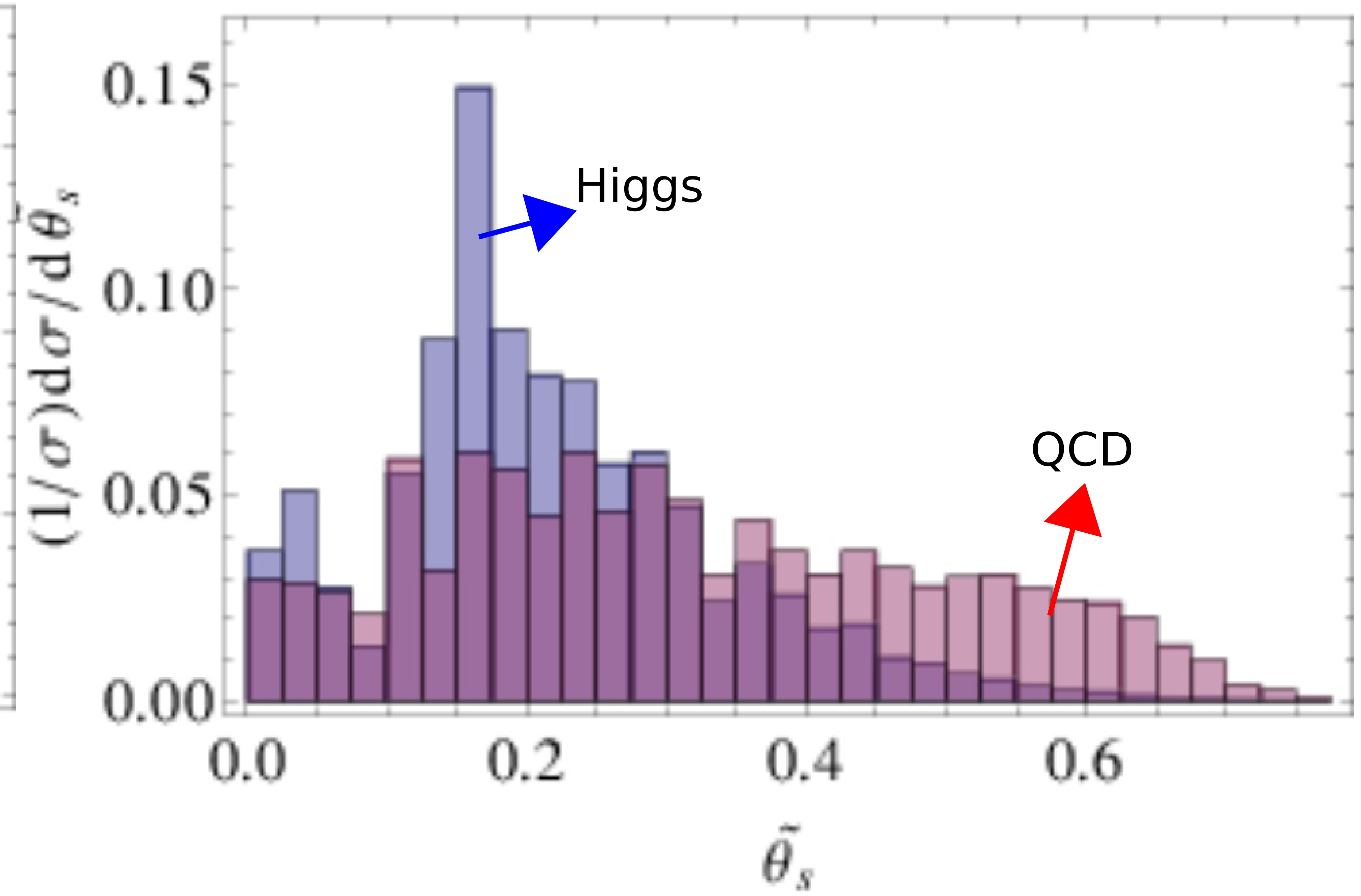
Boosted Higgs searches

We can analyze angular distributions of best-matched templates

Max. Ov: $f[j]$

Kinematical variables $|f>$ ⟷ Jets $|j>$





$$\bar{\theta} = \sum_i \sin\theta_{iJ}$$

$$\theta_s = \min\{\theta_{iJ}\}$$

Preliminary

Rejection power

Backovic, JJ, Perez, Winter in preparation

	WH	WZ	Wt	$Wb\bar{b}$	$t\bar{t}$	$\frac{S}{B}$	$\frac{S}{\sqrt{B}}$
$\sigma[fb]$	8.42	13.08	41.43	72	7200	-	-
K factor	1.3	3.2	2	1.3	1.8	-	-
$15fb^{-1}$	13	23	1	23	124	0.07	1
Fake % (Eff 10%)	-	0.1	2.6	1.6	0.3	0.6	0.8
Fake % (Eff. 20%)	-	0.2	1.3	1.8	0.3	0.5	1.1
Fake %(Eff. 30%)	-	0.3	1.2	1.2	0.4	0.4	1.2

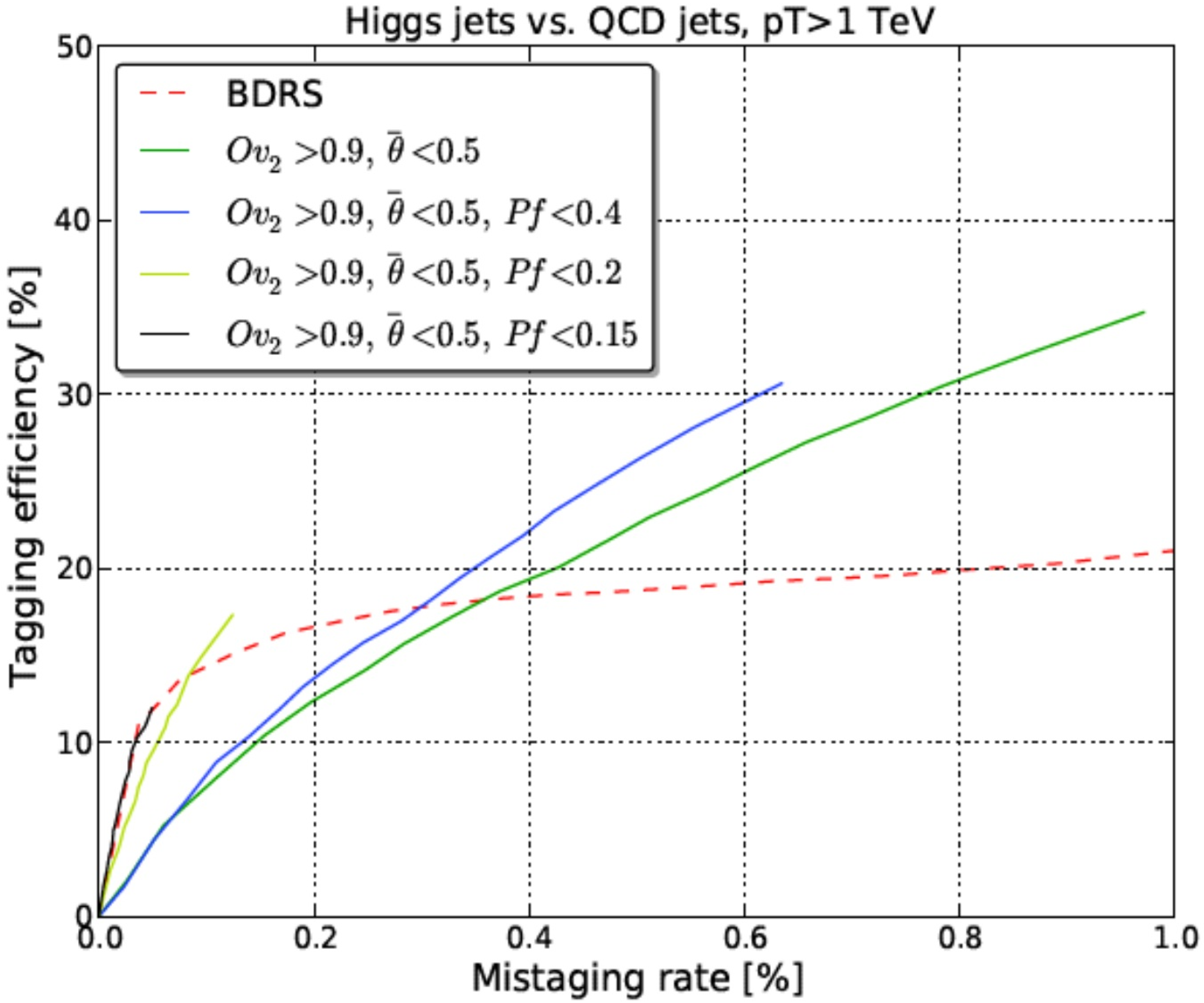
Table 1: Results for $\sqrt{s} = 8\,\text{TeV}$. Jets selected with anti-k_T $R = 1.4$, $p_T^{\text{jet}} > 220\,\text{GeV}$, $|y| < 2.5$, $p_T(\text{lepton} + \text{MET}) > 150\,\text{GeV}$. We also require exactly 2 $b-$tagged jets inside the fatjet, and 0 jets outside. Quoted fake rates correspond to the rejection given by Template Overlap, after the set of minimal kinematic cuts. $S/\sqrt{B}$ values correspond to $15\,\text{fb}^{-1}$ of integrated luminosity.

- LHC 8 TeV; 15 fb^{-1}
- Pythia with "Atlas" tune
- 75% b tag; 1% mistag
- Only HW channel

Summary

- ◆Top/Higgs/W/Z jets play key role in new physics searches
- ◆Boosted configurations can improve searches
- ◆Template Overlap is a general, flexible method for jet substructure
- ◆Code soon to be public. Available upon request.

Comparison of the techniques



BDRS: Butterworth, Davison, Rubin, Salam, PRL 100 (2008)

Fat jet: CA with R = 0.7,
pT j > 1000 GeV

Filtering with R_{filt} = Rbb/2, μ = 0.667, ycut > 0.09