Template Overlap Method

Jose Juknevich

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

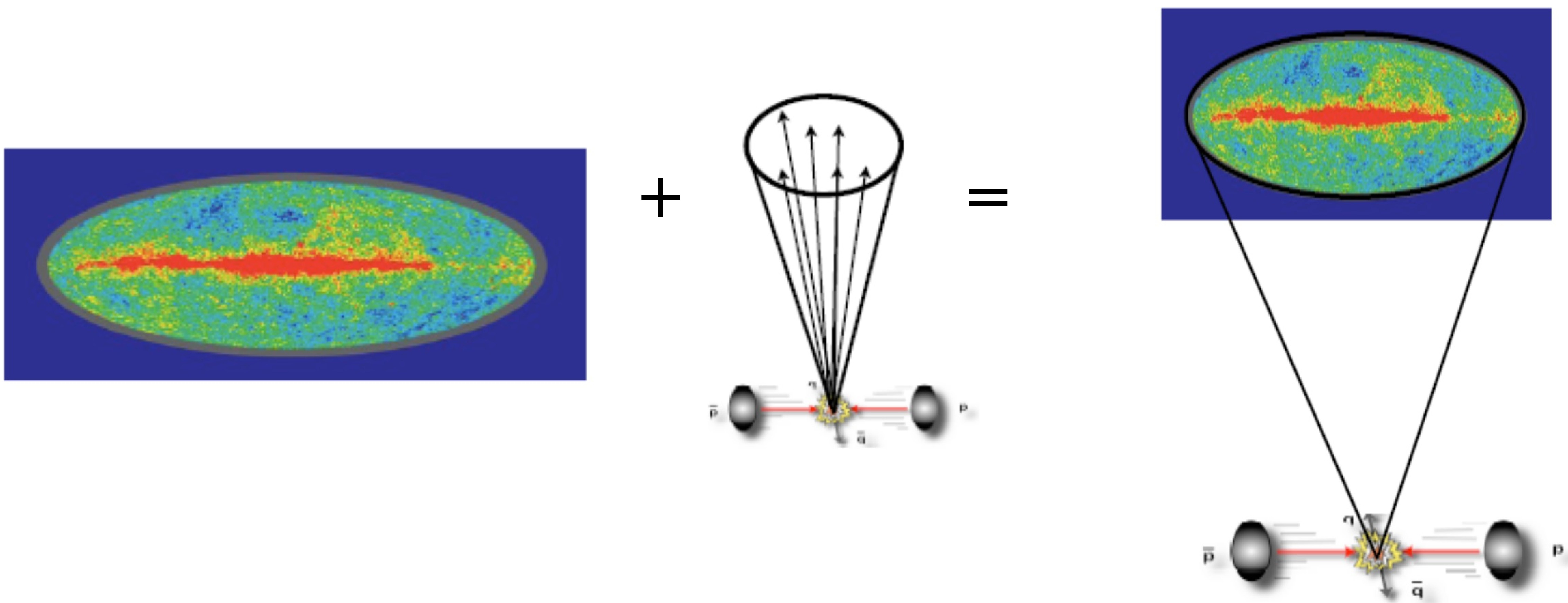
Boost 2012

July 26, 2012

Motivation

Energy flow is a natural language for jet substructure

- Jet cross sections are naturally described in terms of energy correlators
- For QCD, these correlations tend to be strongly peaked around jets



We can characterize the signal with spikes of energy that we can calculate in perturbation theory

Outline

- Template Overlap Method: procedure to discriminate heavy jets using their energy distributions
- Applications: Top and Higgs jet tagging
- The effects of pileup: Why the method can be effective in the presence of pileup?

Template Overlap Method

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

Given a theoretical model f, or template, associate a functional measure to each event j quantifying how well the energy flow of f matches the flow of this event

$$Ov_N(j,f) = \qquad \mathcal{F}(j,f)$$

Prob for event j
to match model f

Matching measure

Template Overlap Method

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

Given a theoretical model f, or template, associate a functional measure to each event j quantifying how well the energy flow of f matches the flow of this event

$$Ov_N(j, f[j]) = \max_{f \in \tau_N^{(R)}} \mathcal{F}(j, f)$$

Prob for event j to match model f

N-particle phase space

Matching measure

We identify the difference in terms of the template configuration $f[j]$ with the closest match to j

Functional measure

- As a measure of the matching we introduce a function that is maximized to 1 for a "perfect" match
  - In practice modeled by a Gaussian in energy differences

$$Ov_N^{(F)}(j,f) = \max_{\tau_N^{(R)}} \ \exp\left[-\frac{1}{2\sigma_E^2} \left(\int d\Omega \ \left[\frac{dE(j)}{d\Omega} - \frac{dE(f)}{d\Omega} \right] F_N(\Omega, f) \right)^2 \right] ,$$

$\frac{dE(f)}{d\Omega}, \ \frac{dE(j)}{d\Omega}$ are the template and jet energy flows, resp.

$F_N(\Omega, f)$ smooth function of template angles; for example, step function around template directions

$\tau_N^{(R)}$ N-particle phase space

Functional measure

- As a measure of the matching we introduce a function that is maximized to 1 for a "perfect" match
    - In practice modeled by a Gaussian in energy differences

$$Ov_N^{(F)}(j,f) = \max_{\tau_N^{(R)}} \exp\left[-\frac{1}{2\sigma_E^2} \left(\int d\Omega \, \left[\frac{dE(j)}{d\Omega} - \frac{dE(f)}{d\Omega} \right] F_N(\Omega, f) \right)^2 \right],$$

Experimental input

Theory input

$\frac{dE(f)}{d\Omega}, \; \frac{dE(j)}{d\Omega}$ are the template and jet energy flows, resp.

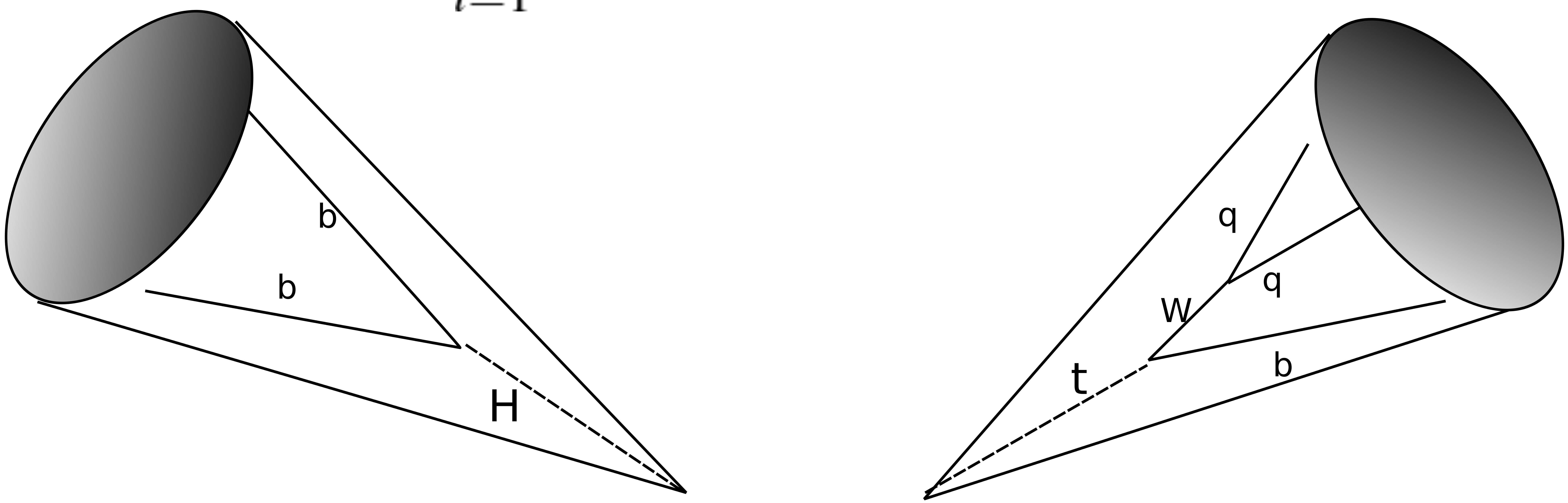
$F_N(\Omega, f)$ smooth function of template angles; for example, step function around template directions

$\tau_N^{(R)}$ N-particle phase space

Template Overlap Method

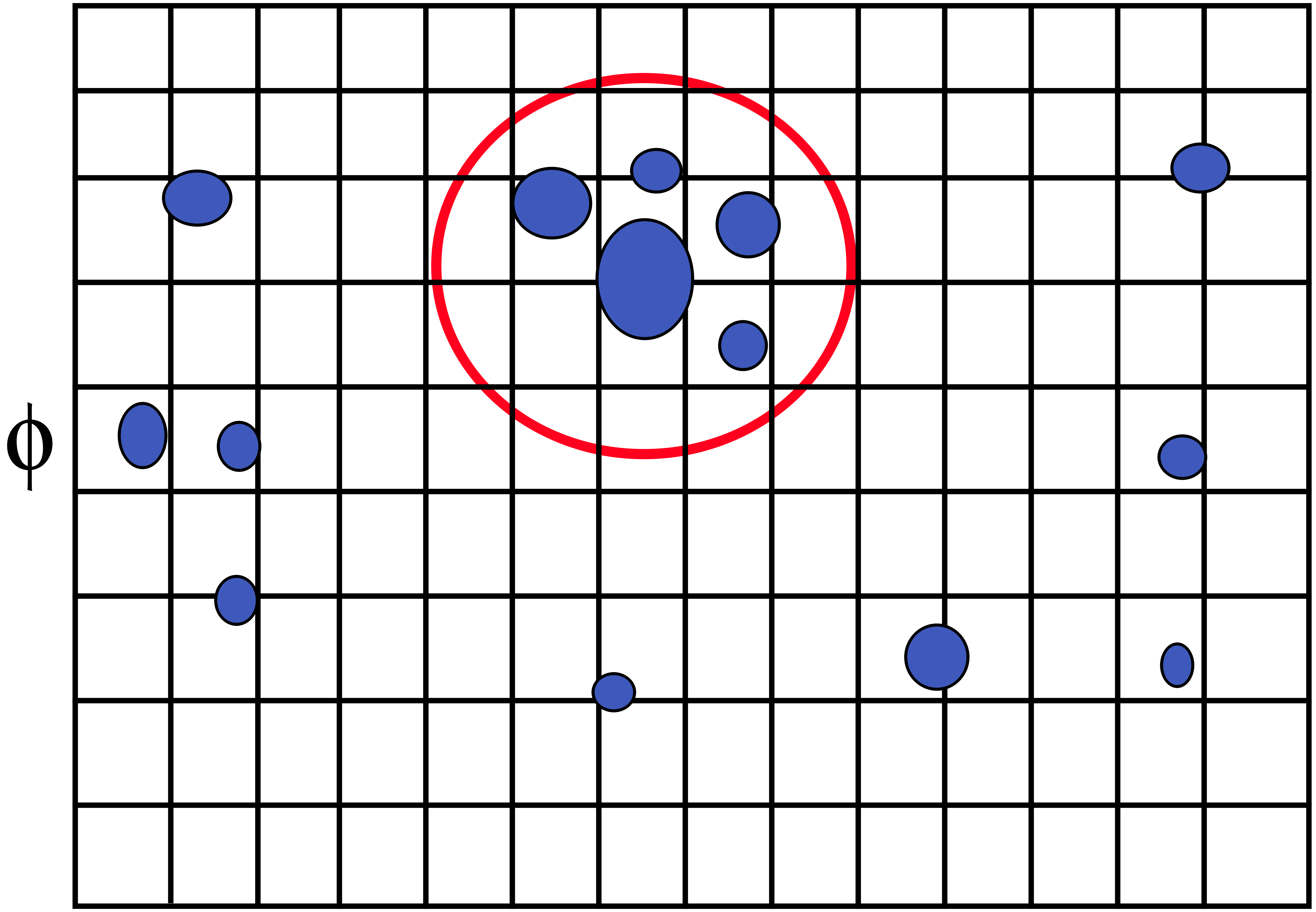
1. Build a catalog of all partonic boosted decays of our signal

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



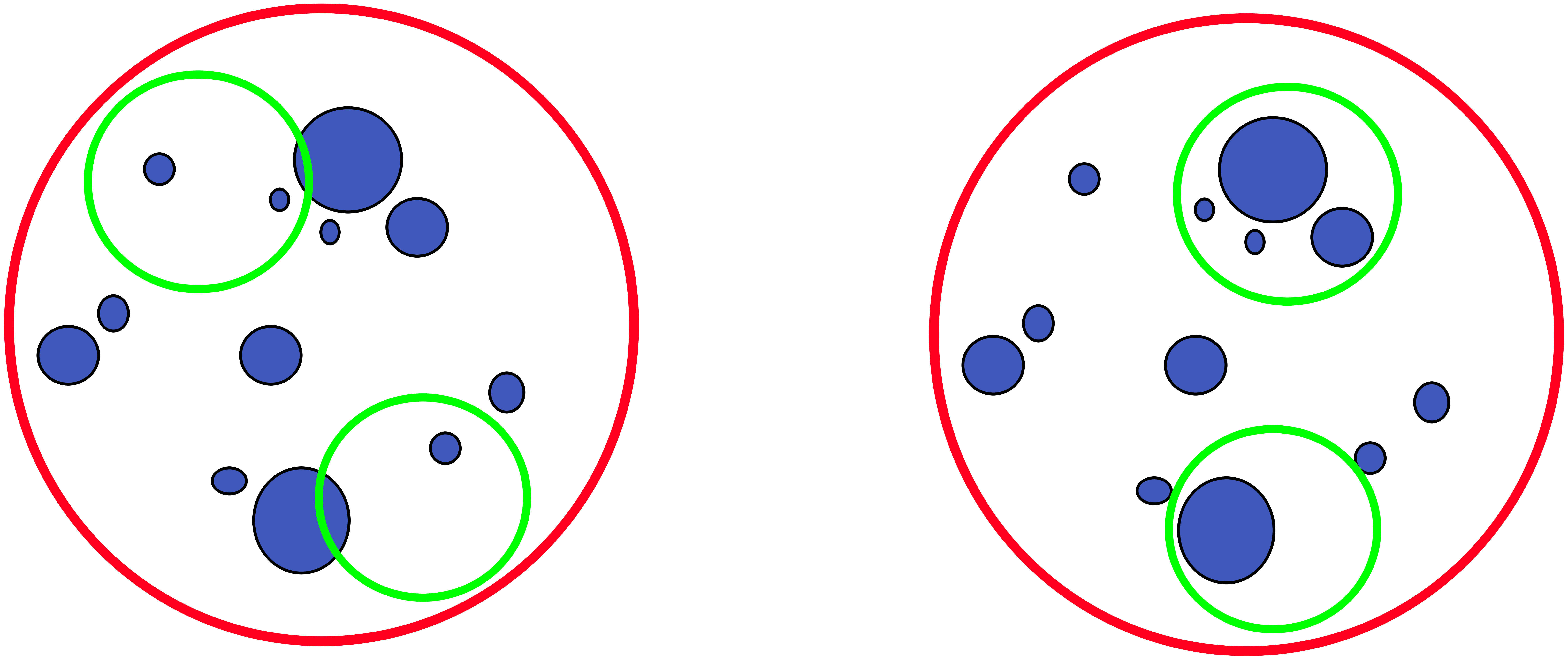
Template Overlap Method

2. Using the anti-kT algorithm cluster the event into fat R=1.4 jets.



Template Overlap Method

3. Search for a configuration (template) that gives a good match to the current jet.



Poor overlap Ov ~ 0

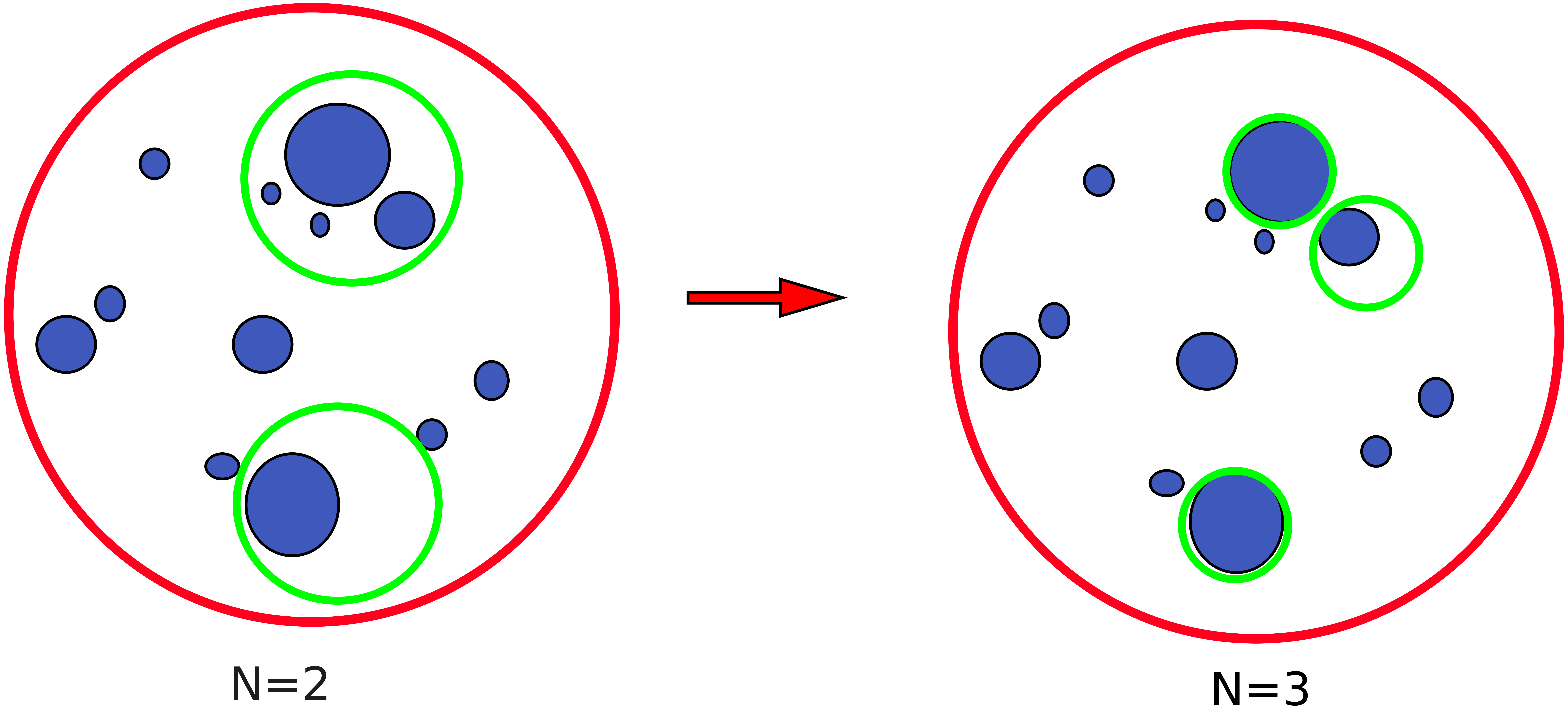
Good overlap Ov ~ 1

$$F(\Omega, f) = \max\left[1 - \frac{\eta^2 + \phi^2}{r^2}, 0\right]$$

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

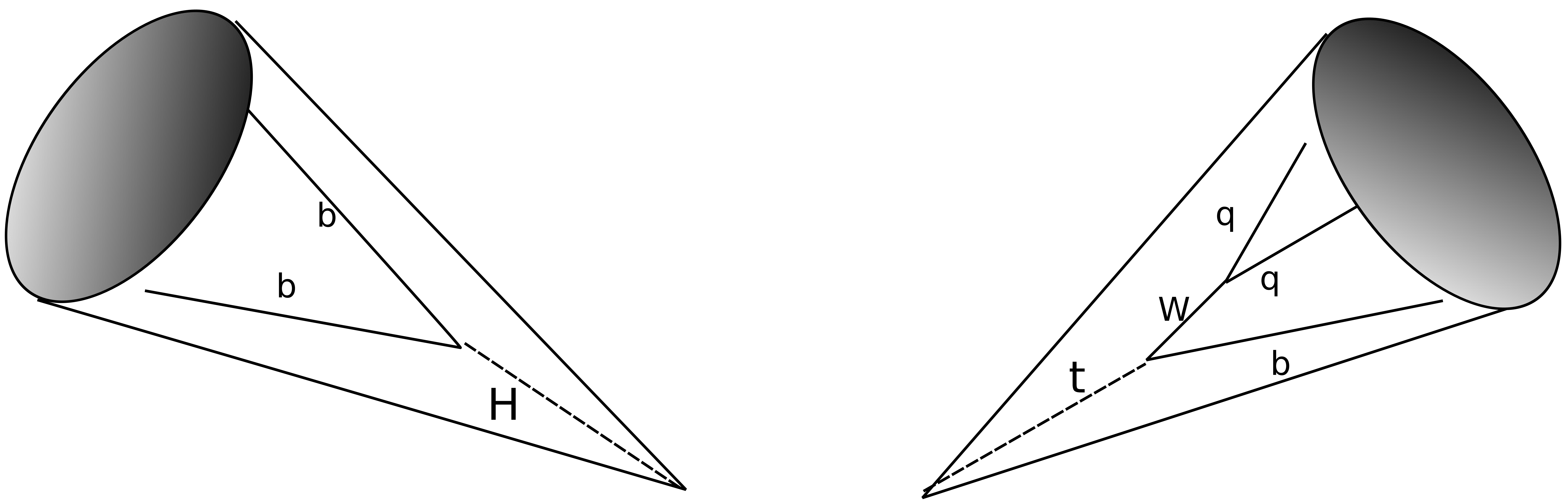
Template Overlap Method

4. If desired, use templates with more than the minimum number of particles to resolve finer details of the substructure.



Template Overlap Method

5. Place limits on the angles of the best matched templates *f*[*j*]
Gives us further information about the likelihood that the event is signal or background.

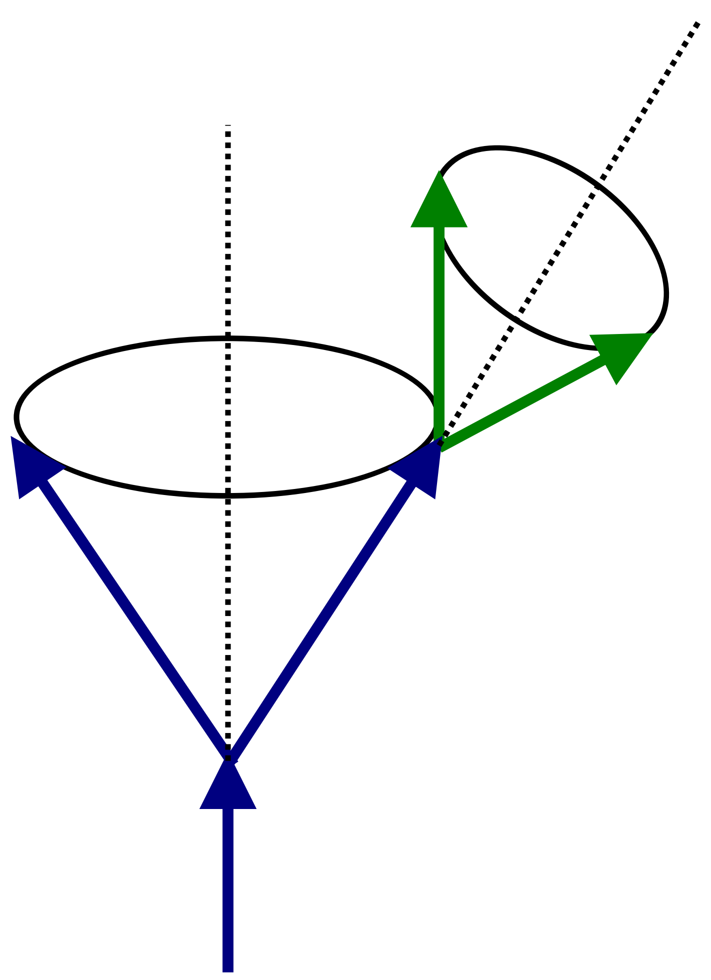


Idea: Calculate additional parameters: planar flow (Pf), angularity (tau_{-2}), ... from the best matched templates *f*[*j*], instead of jet constituents

Applications

Example 1: Boosted top tagging at the LHC

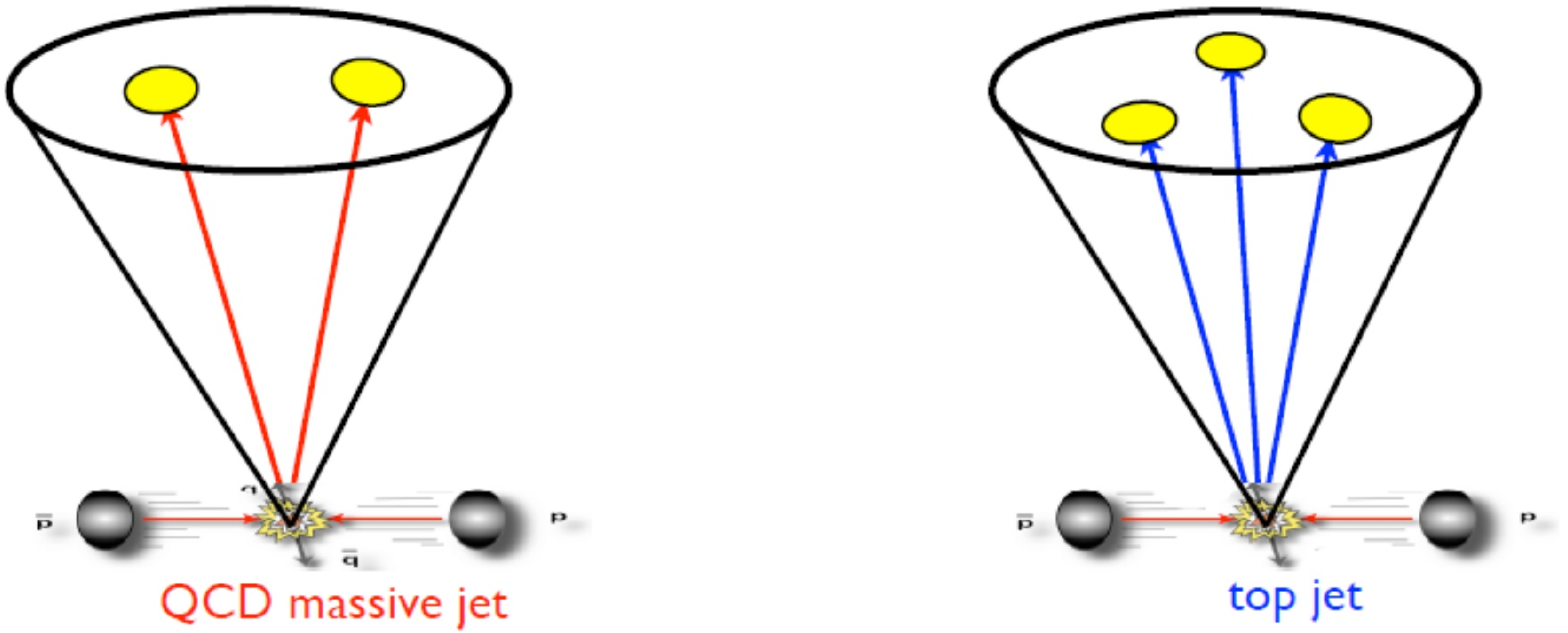
Almeida, Lee, Perez, Sterman, Sung (2010)



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

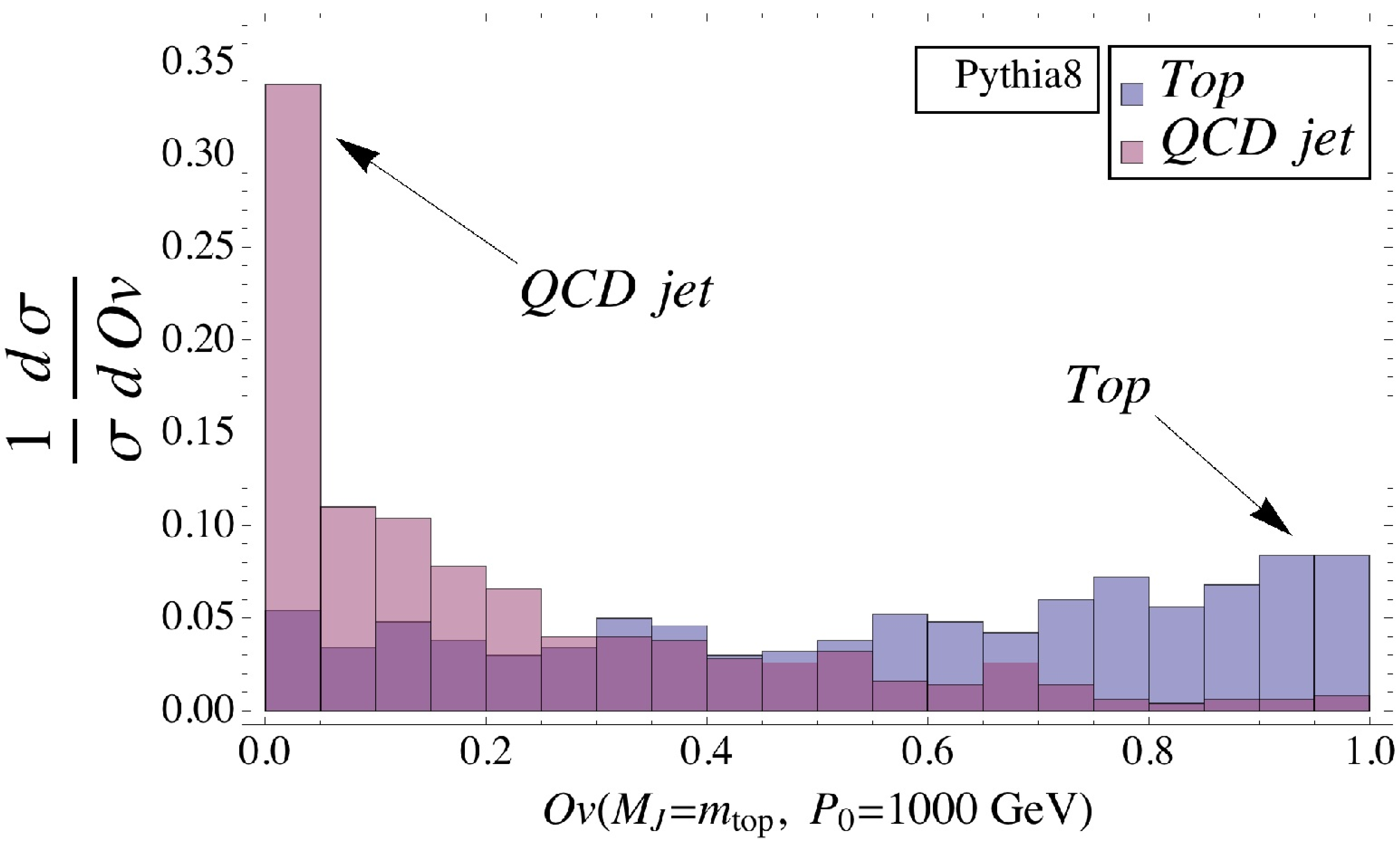
$$t \to b + W \to b + q + \bar{q}. \qquad (p_q + p_{\bar{q}})^2 = M_W^2$$

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



Top jets vs QCD jets

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



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

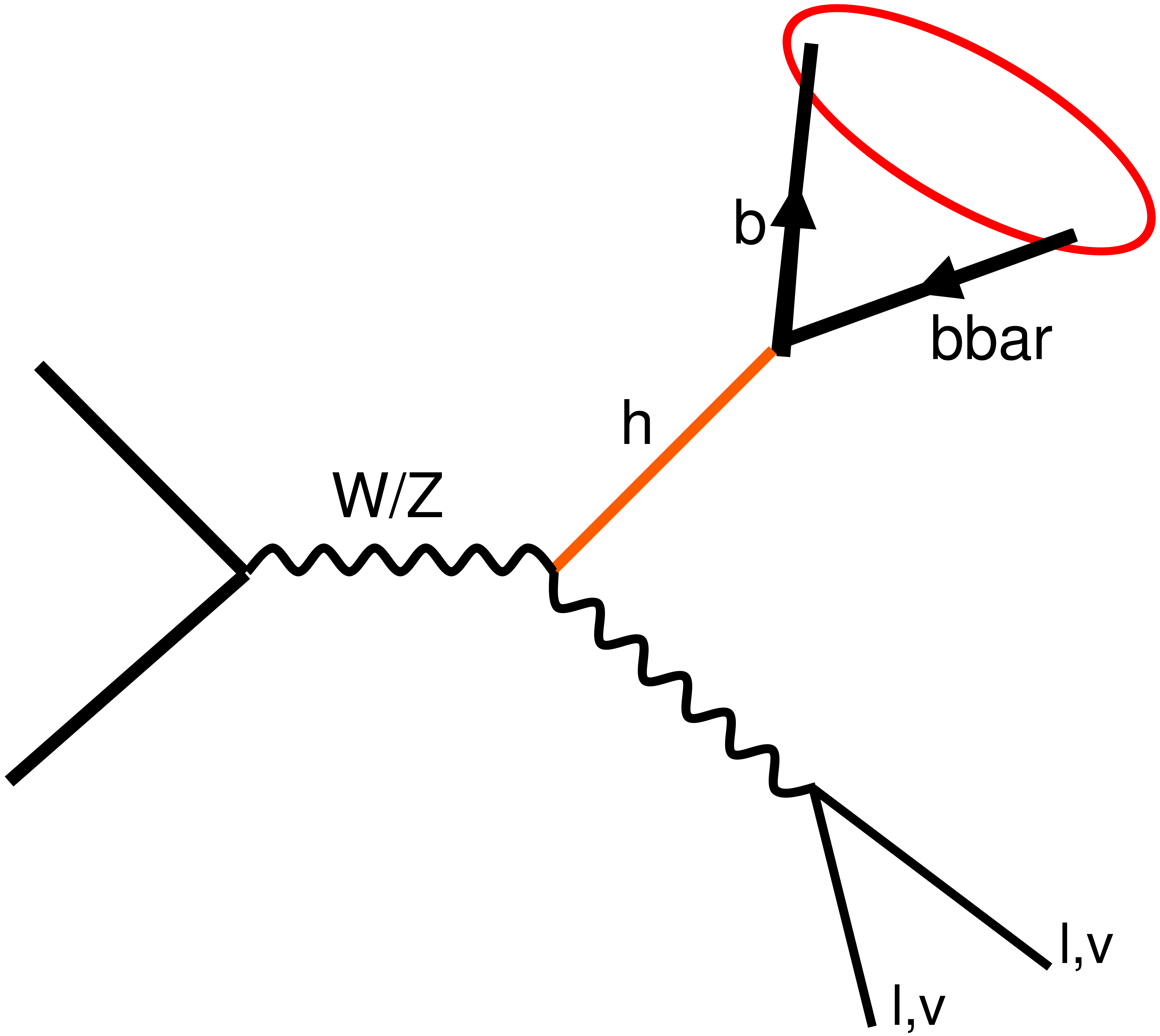
Can be combined with jet shapes (planar flow, pull) to distinghuish between many three-jet events with large overlap.

eff. 10 %, fake 0.02%

Rejection power ~O(10^2)

Example 2: Boosted Higgs Searches

A interesting application is to use the Template Overlap method to look for associated VH production



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

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

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

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

Typical jet size

$$\Delta R \sim 1.2$$

Inspiration (but different) from Butterworth, Davison, Rubin, Salam (2008)

Two-particle templates

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

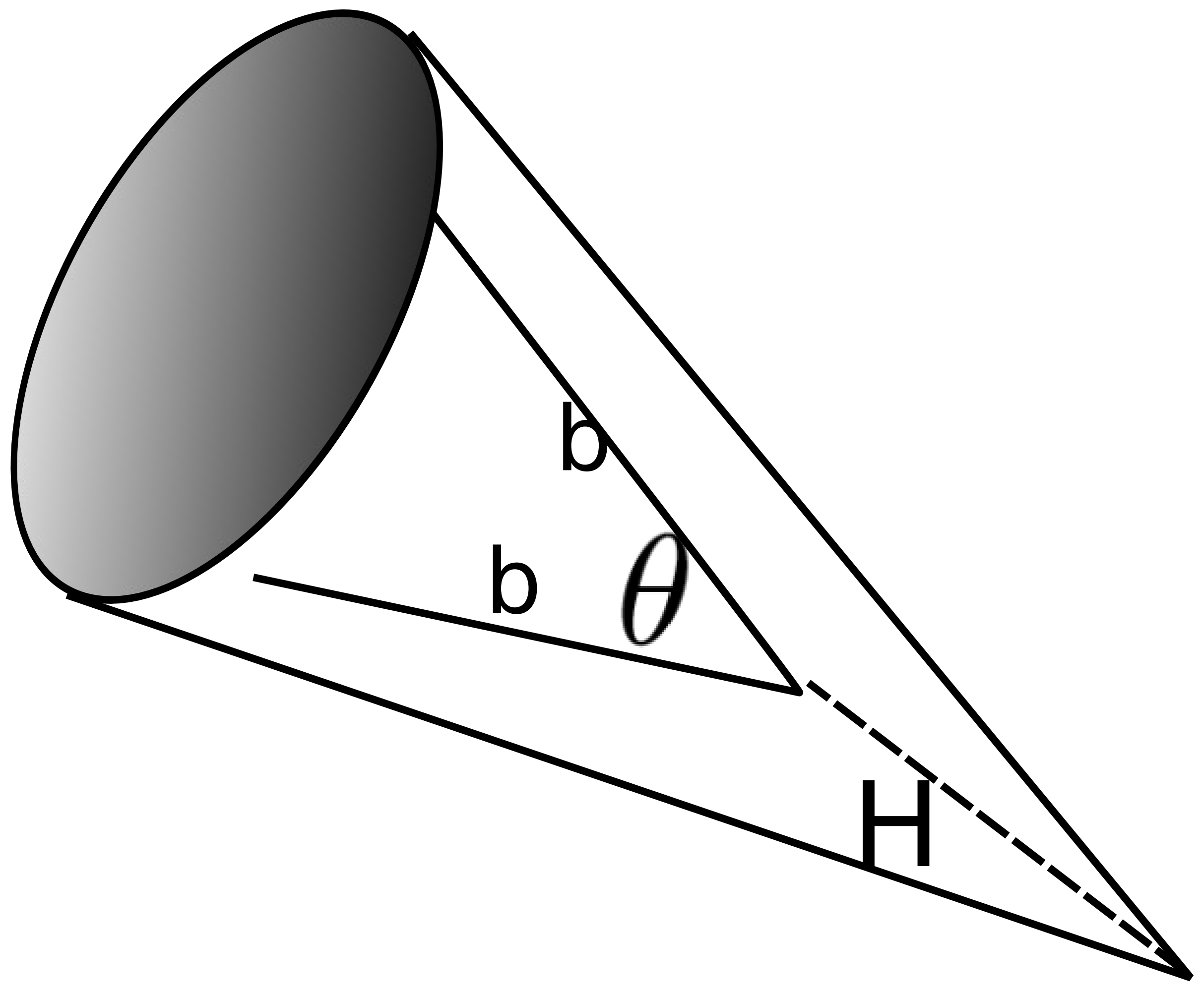
Construct template: two-particle phase space for Higgs decay (easy)

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

Higgs decay are democratic, sharing energy evenly

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

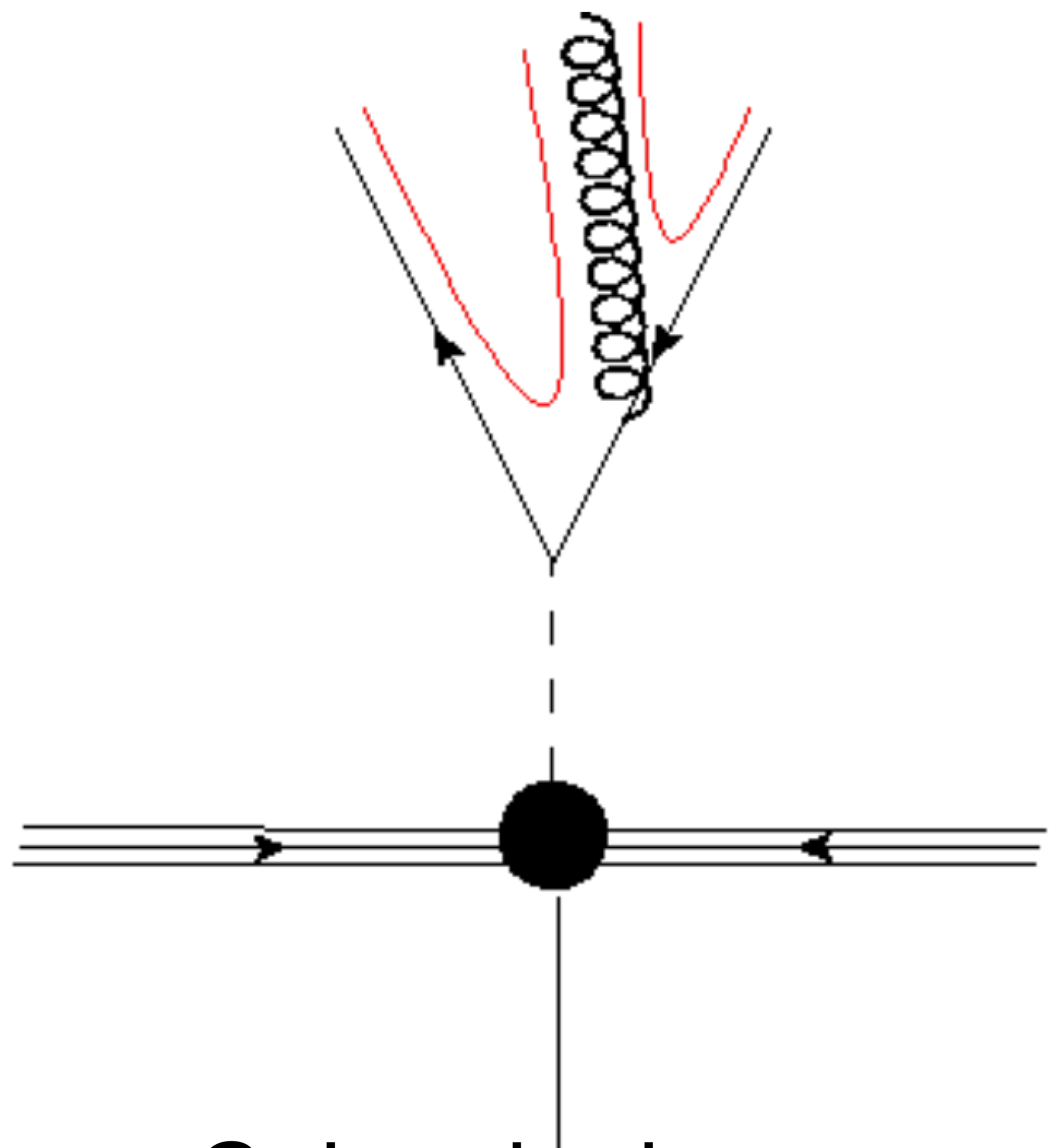
- Higgs $$\frac{d^2\sigma}{dm_J^2 d\theta} \propto \delta(m_J^2 - m_H^2) \frac{1}{\theta^3}\bigg|_{R>\theta>\frac{m_J}{E_J}}$$



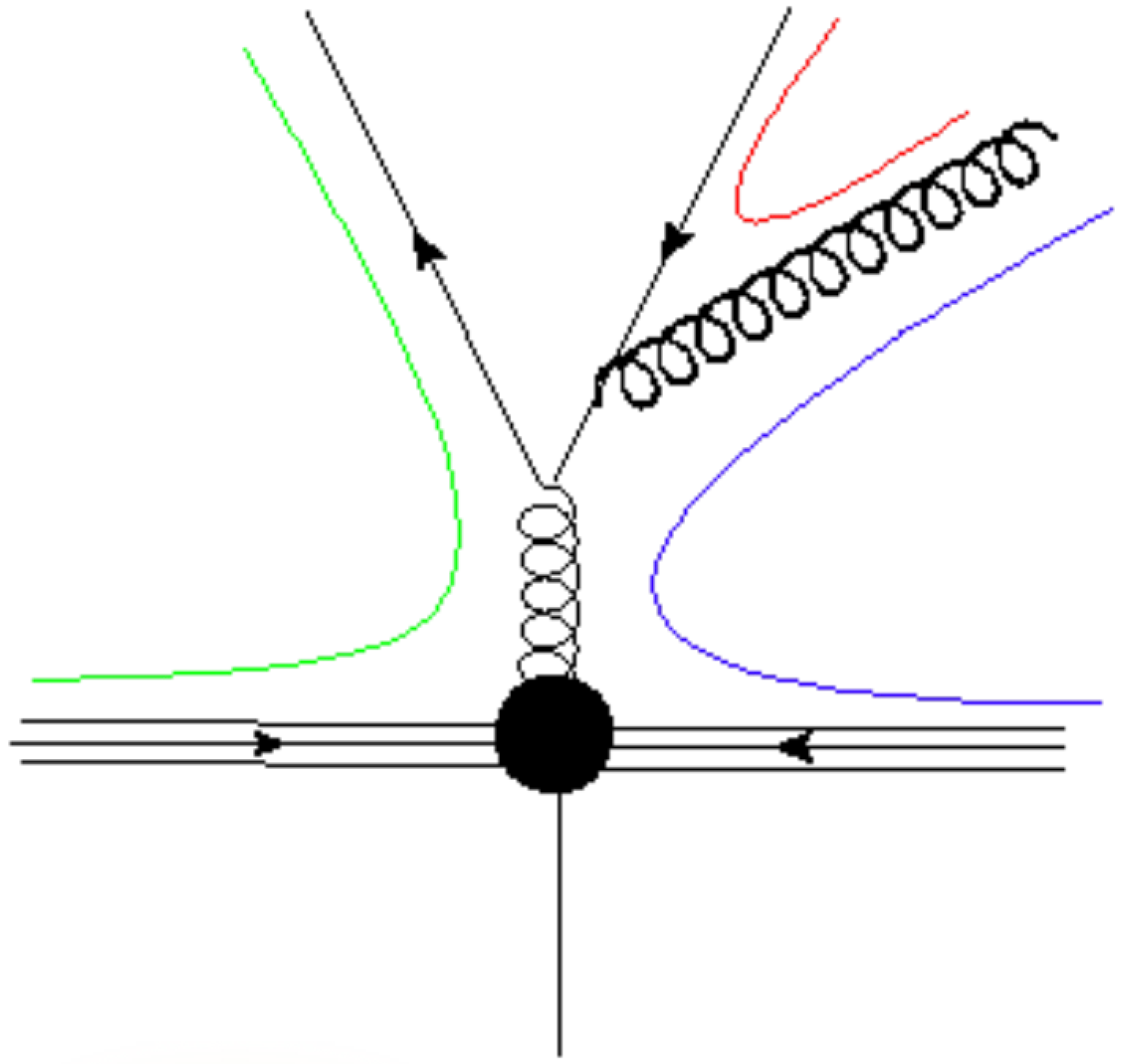
Color flow and three-particle templates

I. Sung (09)
J. Gallicchio and M. Schwartz (10),
K. Black, J. Gallicchio, J. Huith, M. Kagan, M. Schwartz, B. Tweedie (10)

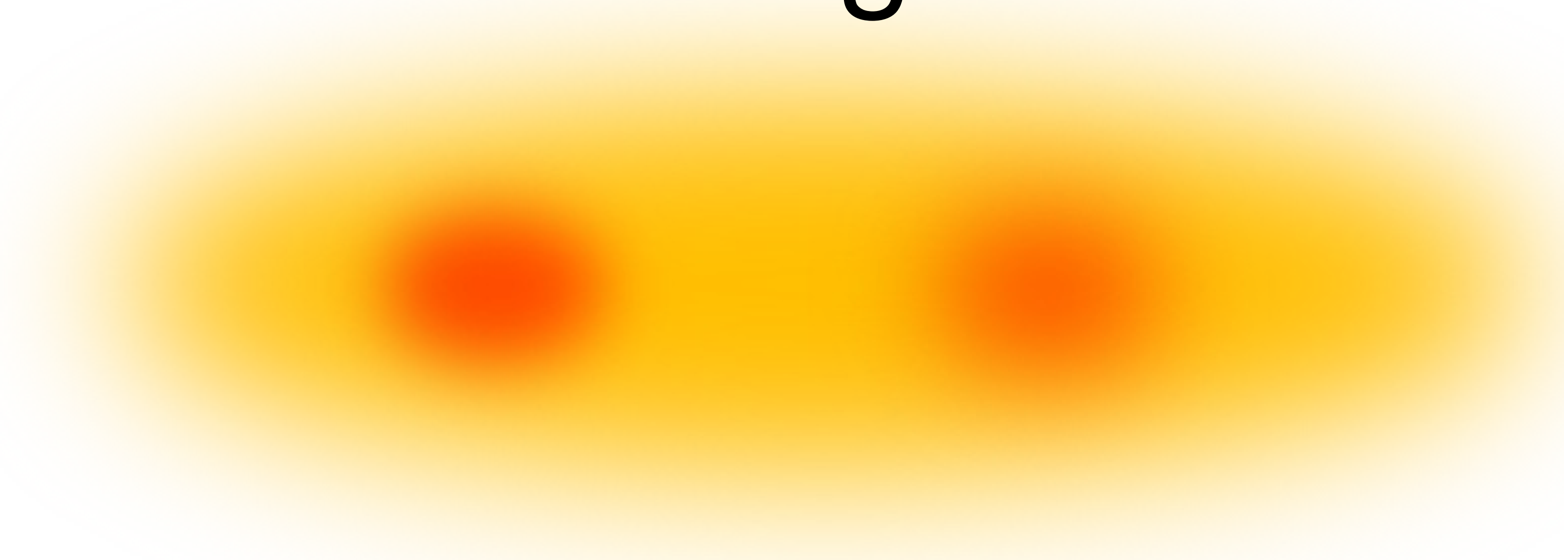
QCD radiation in Higgs decay limited by angular ordering



Color singlet



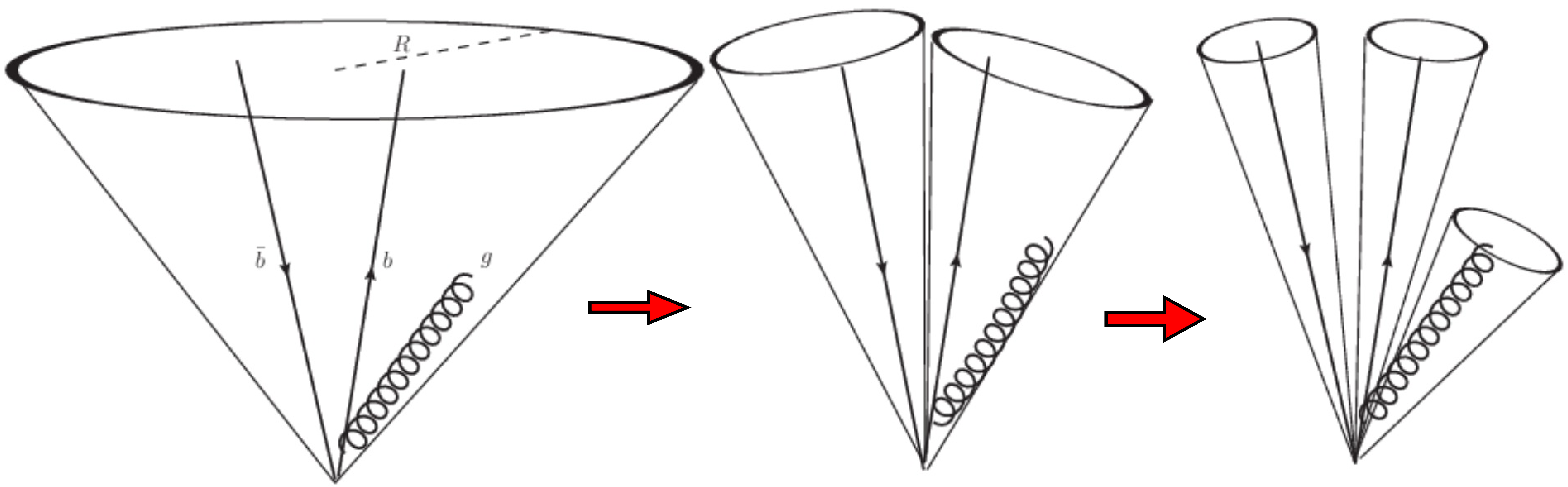
Color octet



Three-particle templates

L. Almeida, O. Erdogan, JJ, S. Lee, G. Perez, & G. Sterman (1112:1957)

- We consider templates with more than the minimum number of particles in the final state
- Allow us to resolve finer details about the substructure

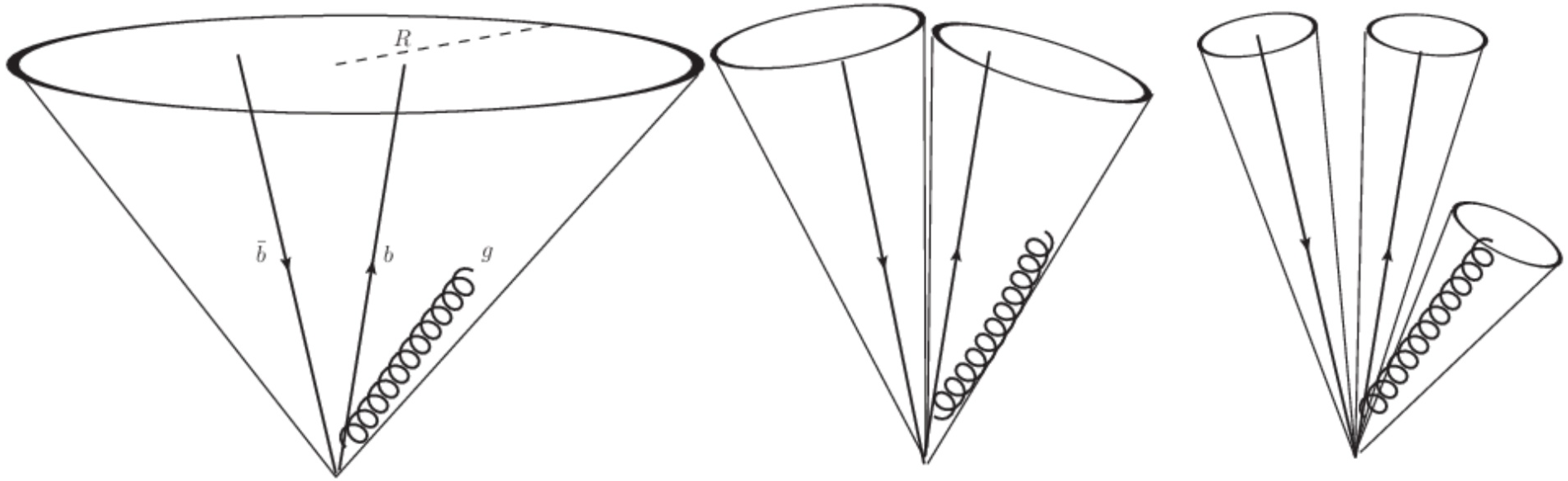


$$|f\rangle = |h\rangle^{(\mathrm{NLO})} = |p_1, p_2, p_3\rangle$$

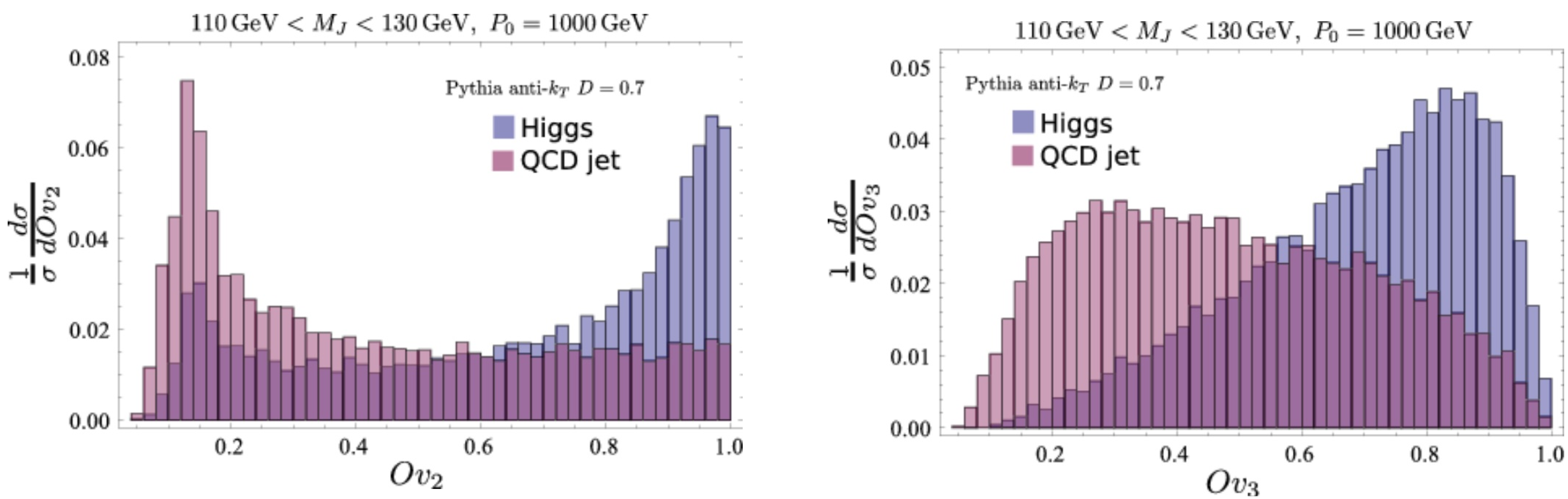
five degrees of freedom

A first look at the observables: Ov2 & Ov3

L. Almeida, O. Erdogan, JJ, S. Lee,
G. Perez, & G. Sterman (1112:1957)



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



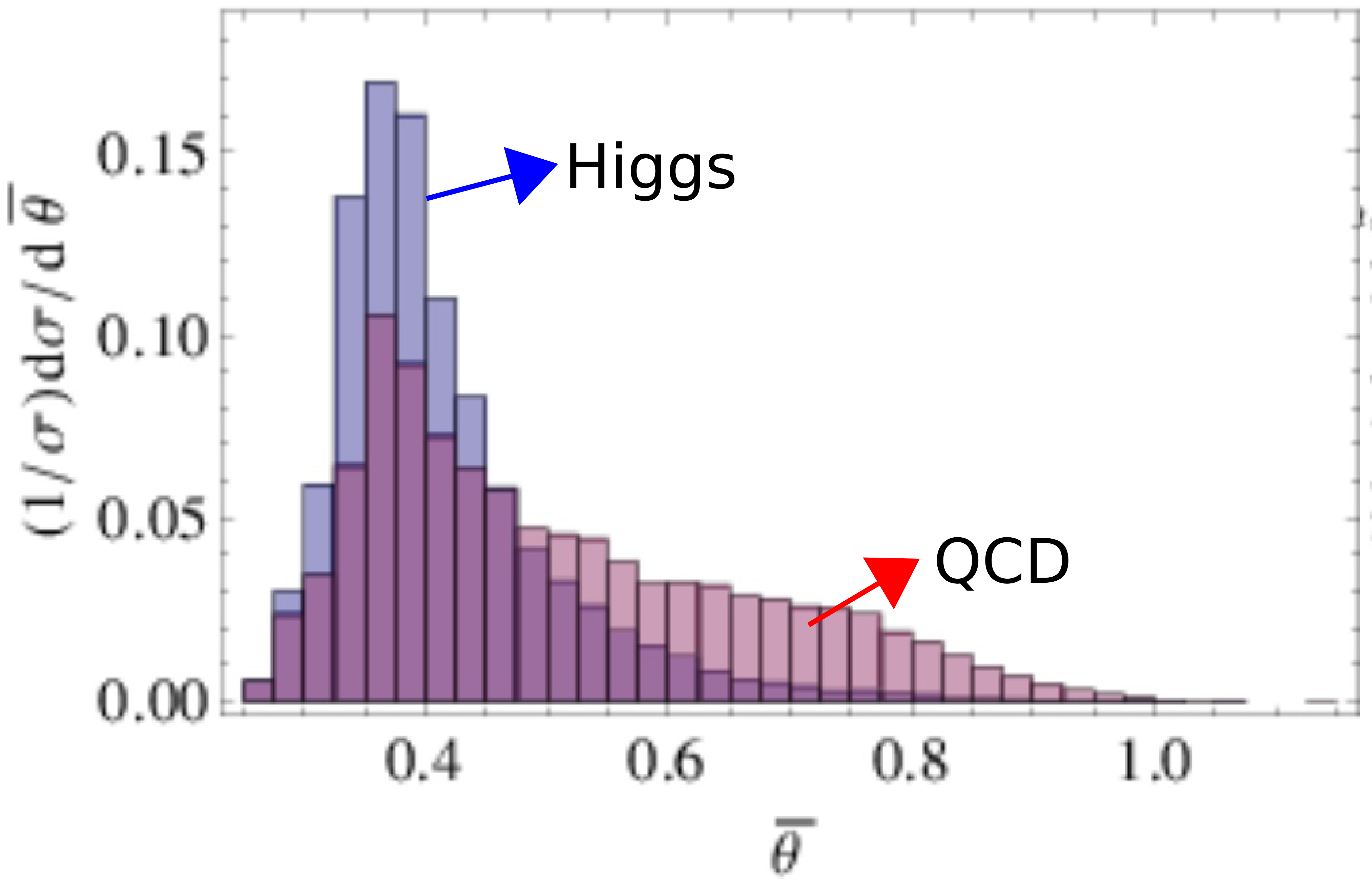
Histograms of template overlap Ov with Higgs jets and QCD jets from Pythia 8, for $R = 0.5$, 950 GeV$\leq P_0 \leq$1050 GeV, 110 GeV$\leq m_J \leq$130 GeV and m_{higgs} = 120 GeV using 2-body templates (Left) and 3-body templates (Right).

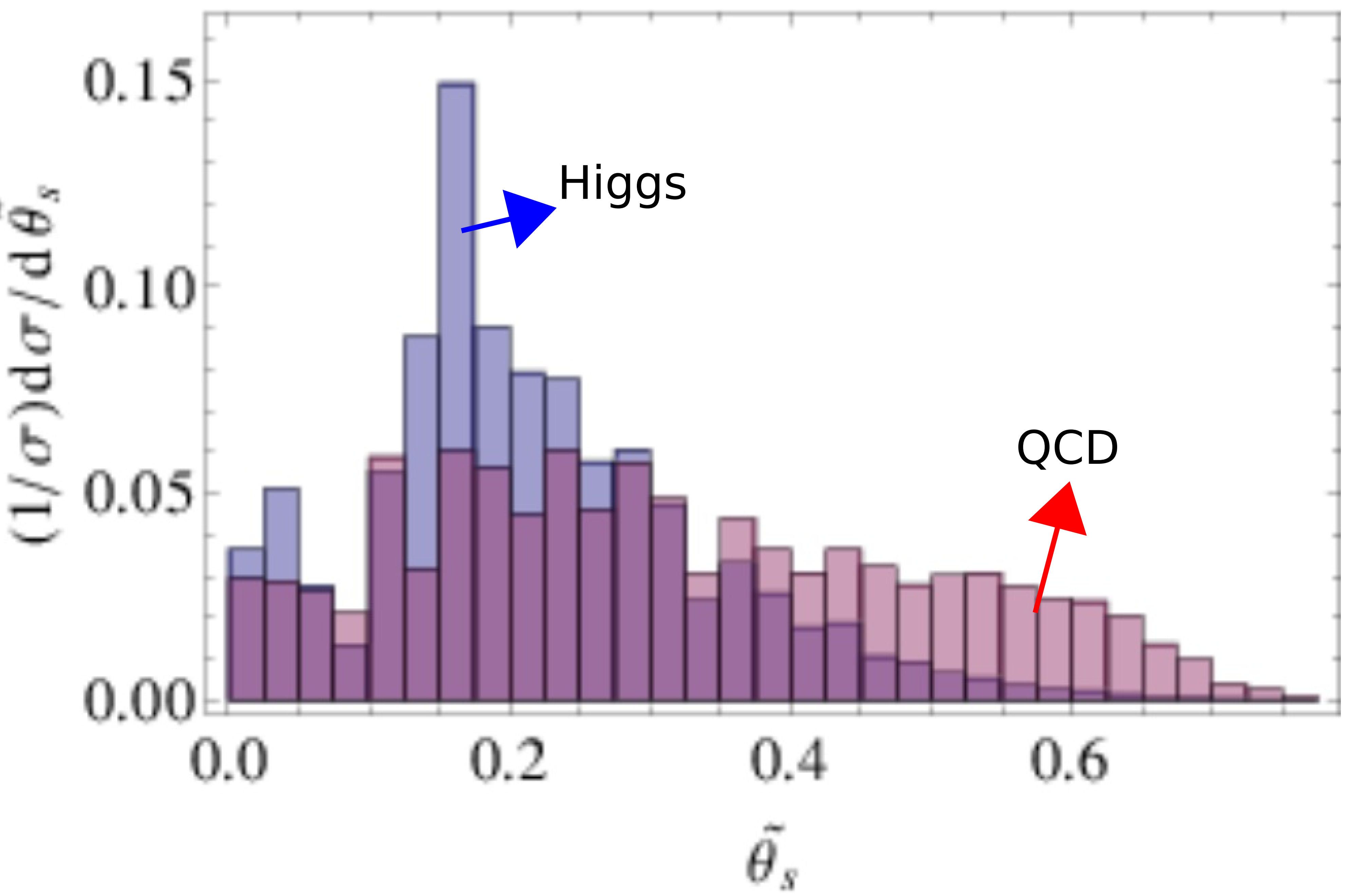
Can do better than that

L. Almeida, O. Erdogan, JJ, S. Lee, G. Perez, & G. Sterman (1112:1957)

We can analyze angular distributions of best-matched templates

Kinematical variables $|f>$ Max. Ov: $f[j]$ Jets $|j>$

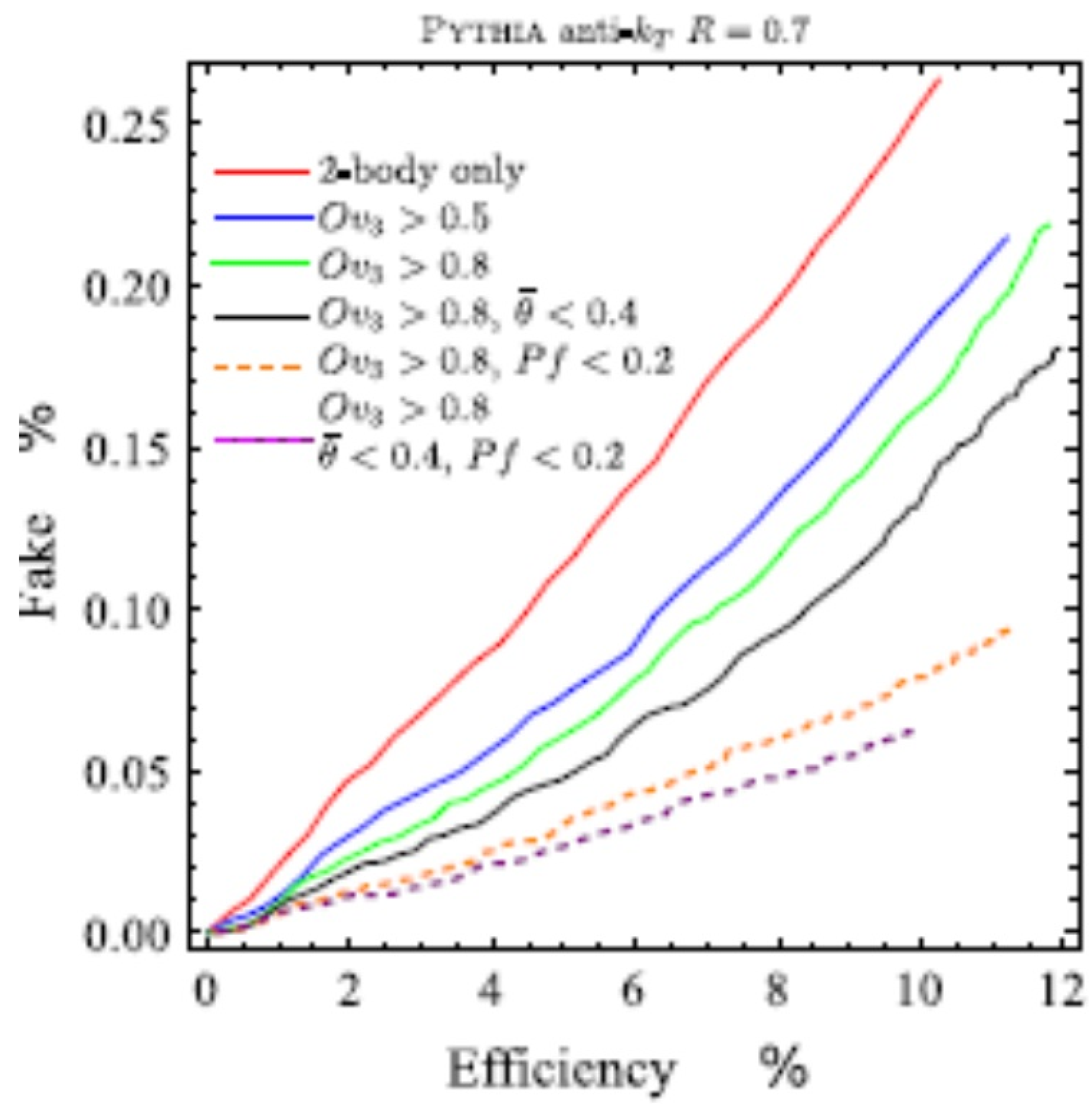




$$\bar{\theta} = \sum_i \sin\theta_{iJ} \qquad \theta_s = \min\{\theta_{iJ}\}$$

Mistag rate vs. Efficiency

L. Almeida, O. Erdogan, JJ, S. Lee,
G. Perez, & G. Sterman (1112:1957)

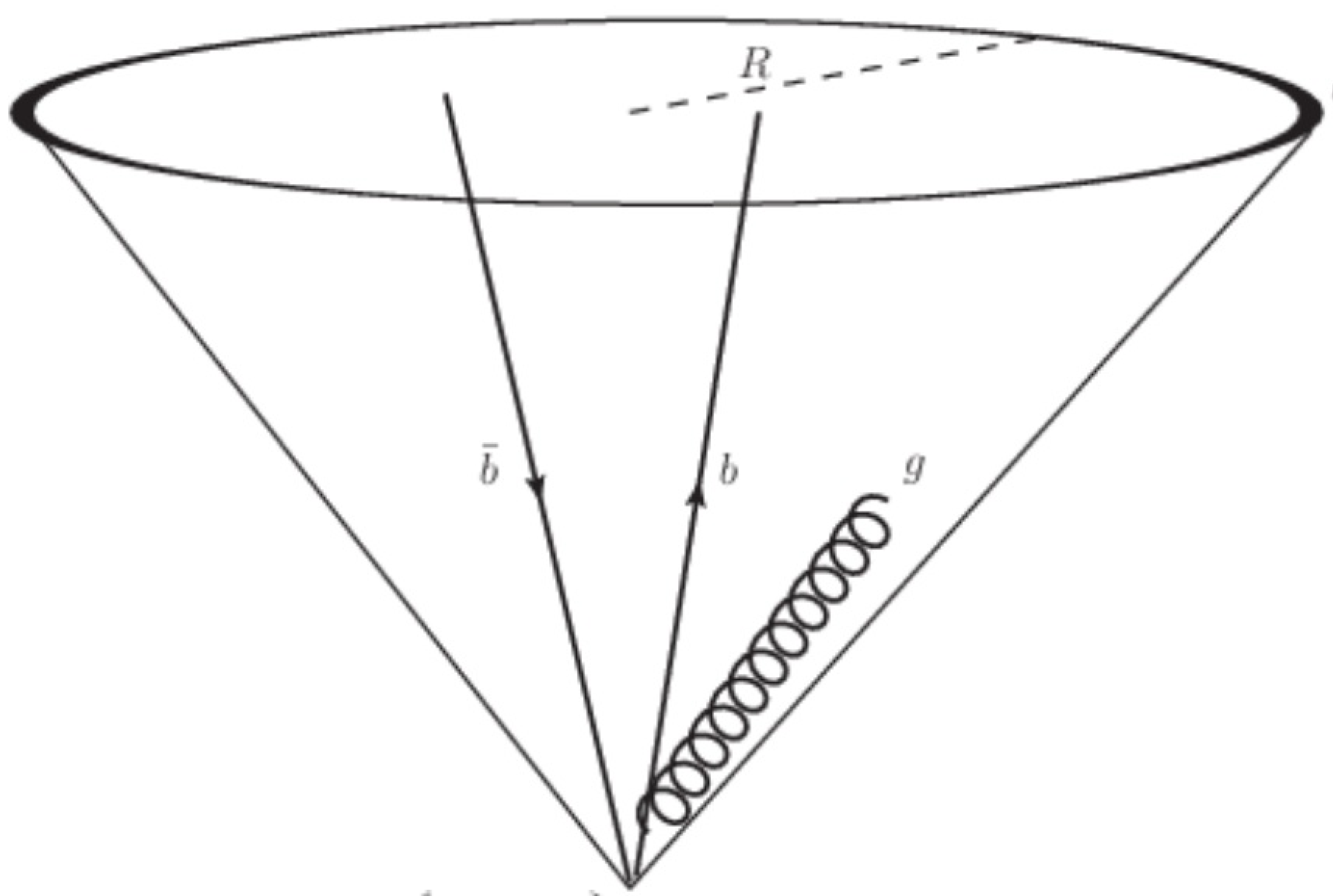


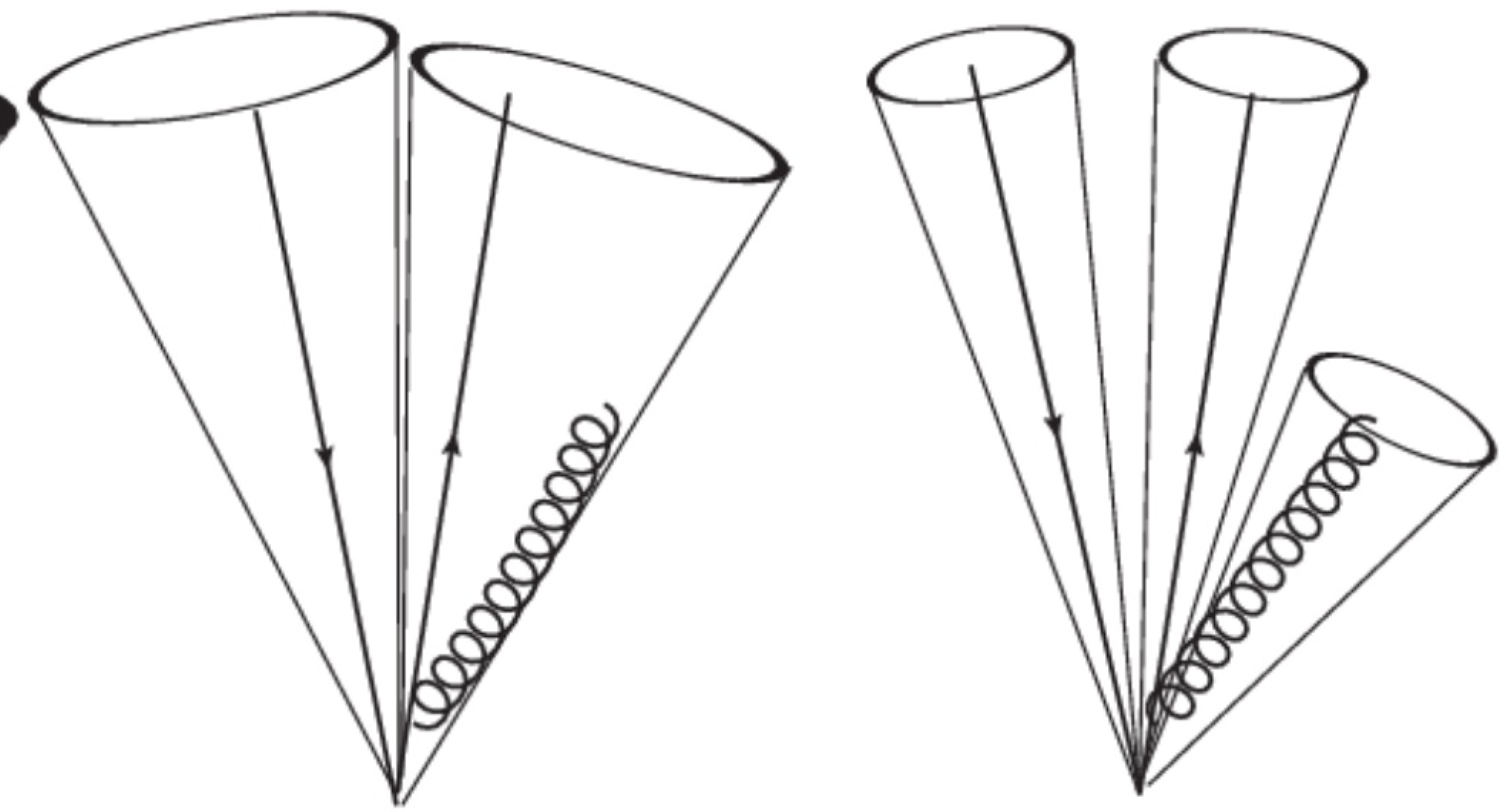
MC	Jet mass cut only		Mass cut + $Ov + \bar{\theta} + Pf$	
	Higgs-jet efficiency [%]	fake rate [%]	Higgs-jet efficiency [%]	fake rate [%]
PYTHIA 8	70	10	10	0.05
MG/ME	70	10	10	0.05
SHERPA	60	10	10	0.05

Table 2: Efficiencies and fake rates for jets with $R = 0.7$ (using anti-k_T: $D = 0.7$), 950 GeV$\leq P_0 \leq$1050 GeV, 110 GeV$\leq m_J \leq$130 GeV and $m_H = 120$ GeV.

Moving down to low pT

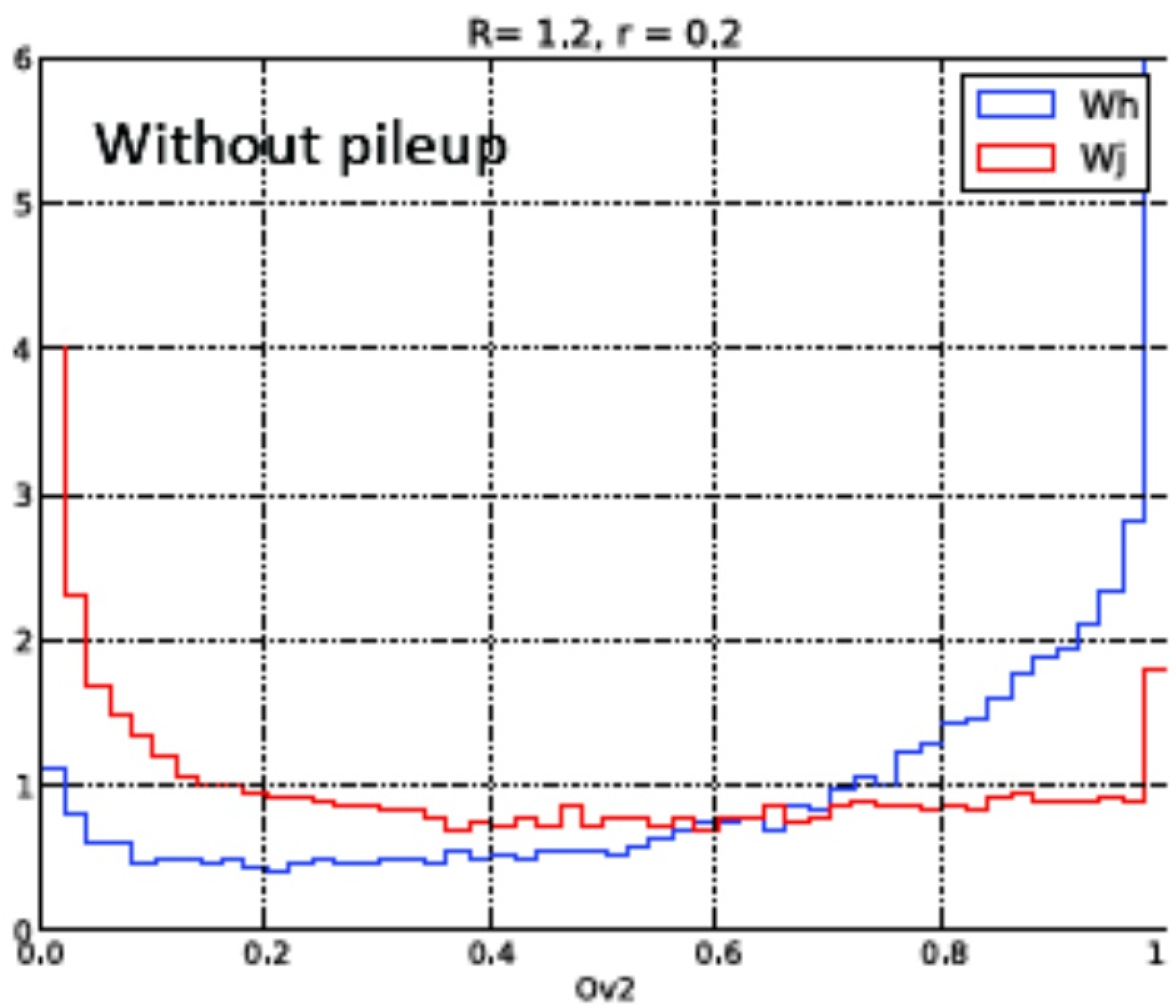
Backovic, JJ, Perez, Winter, in preparation

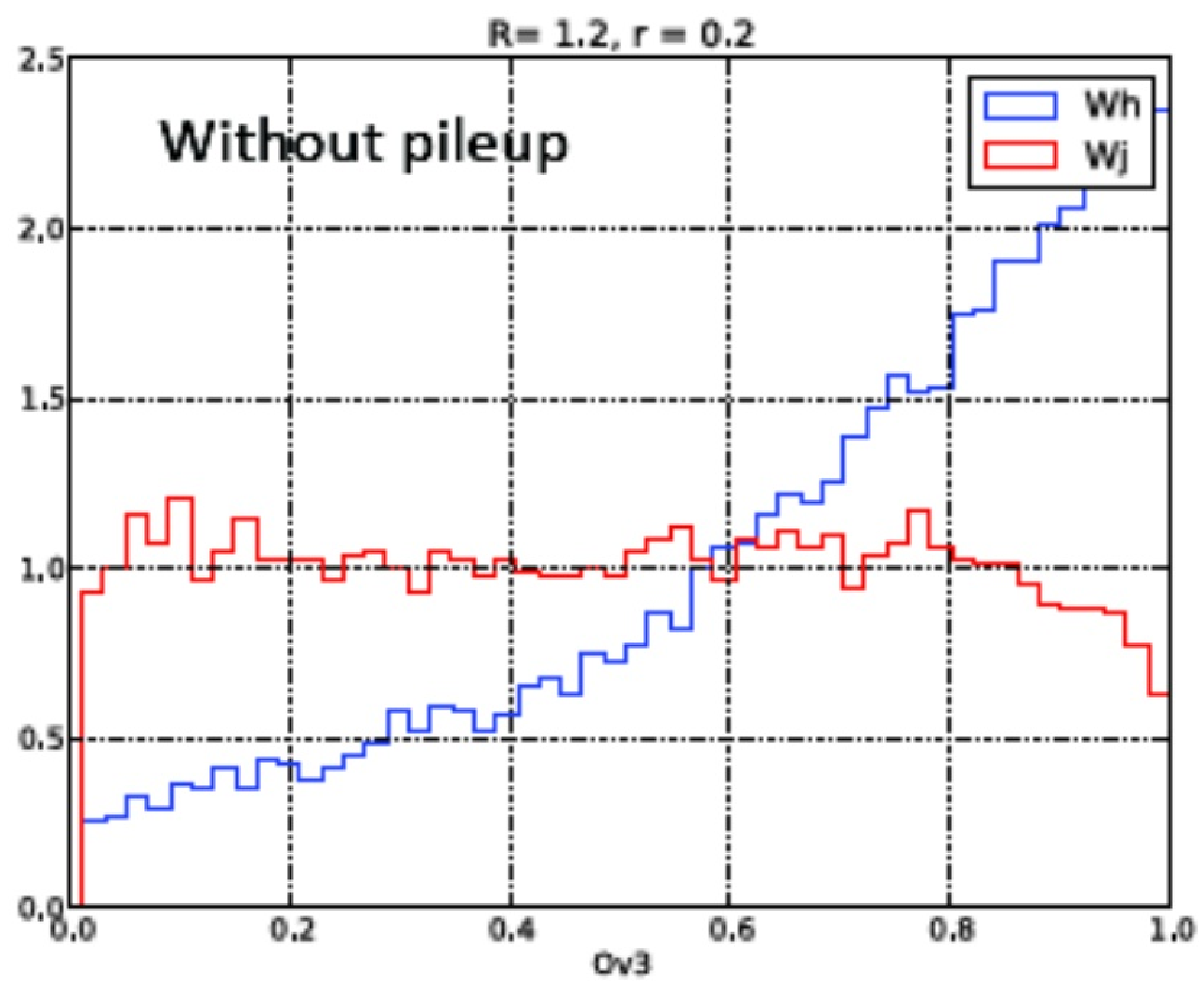




$$|f\rangle = |h\rangle^{(\mathrm{LO})} = |p_1, p_2\rangle \qquad |f\rangle = |h\rangle^{(\mathrm{NLO})} = |p_1, p_2, p_3\rangle$$

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



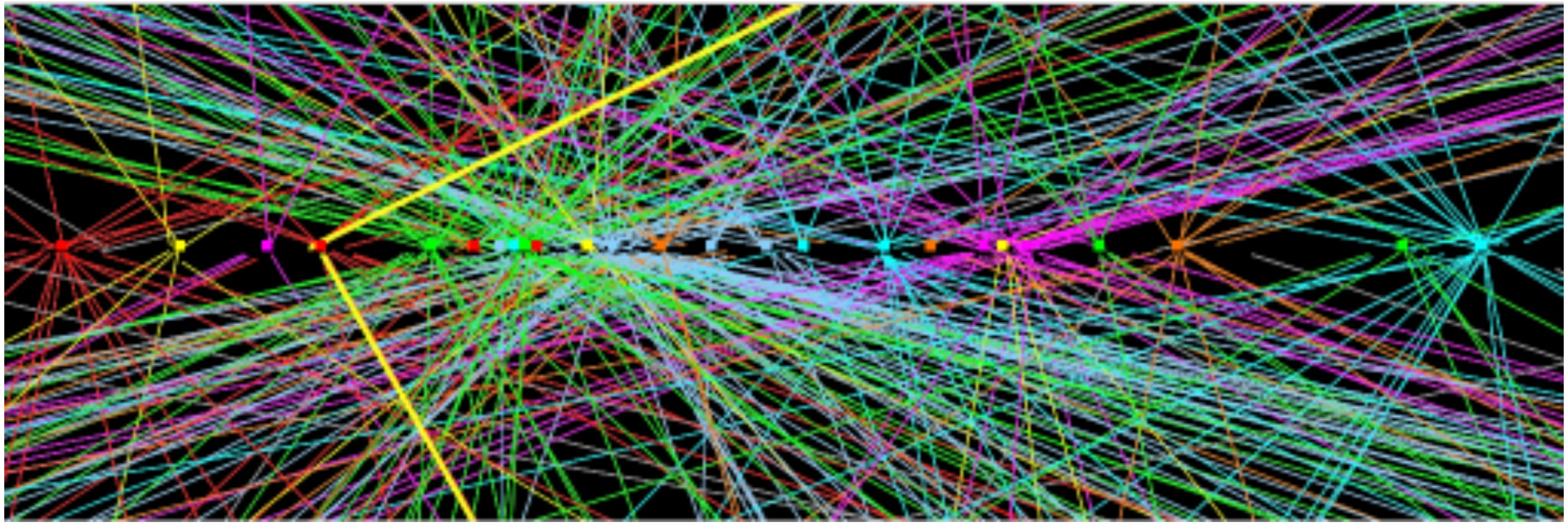


Rejection power

	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.

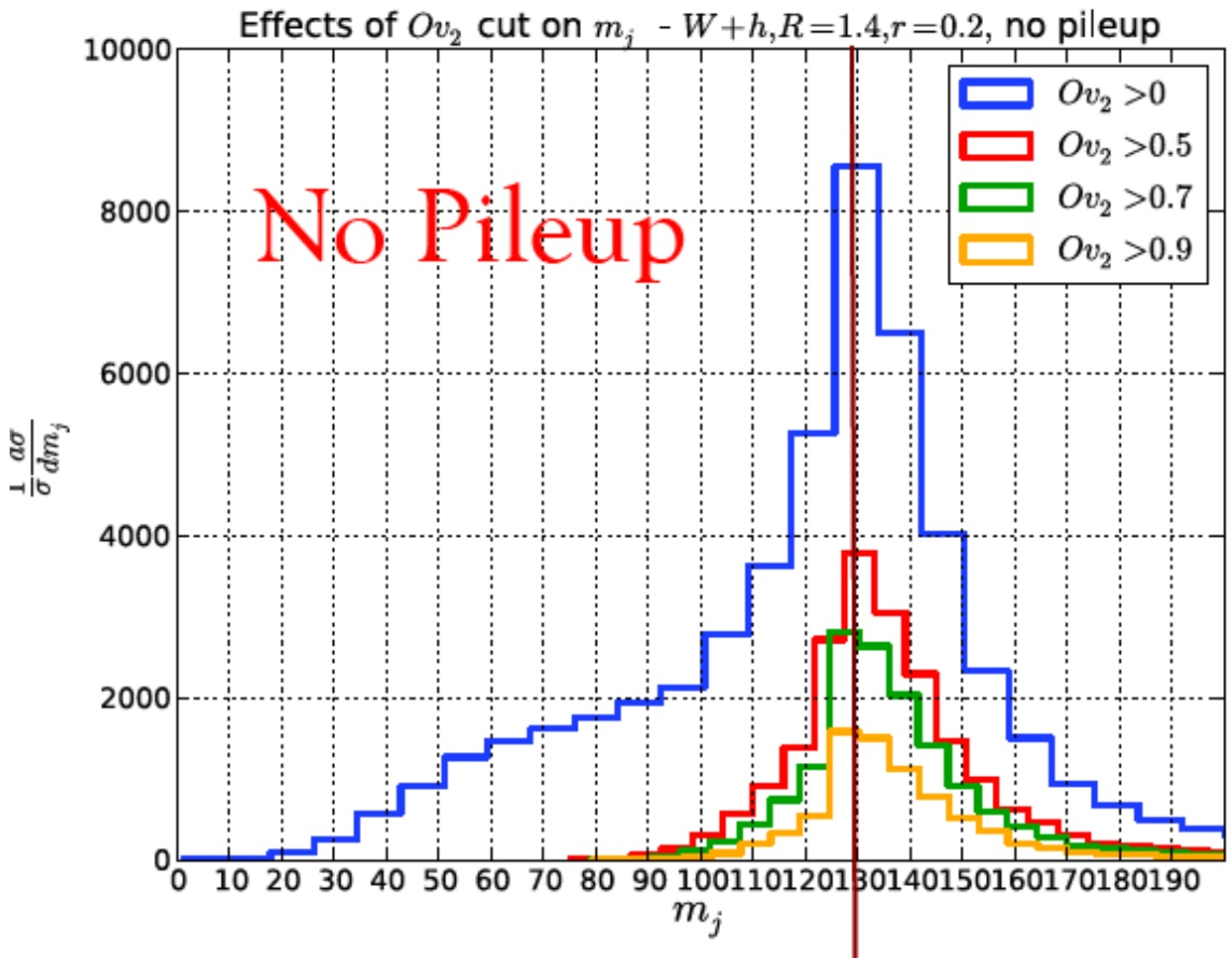
Why the method is useful for pileup rejection?

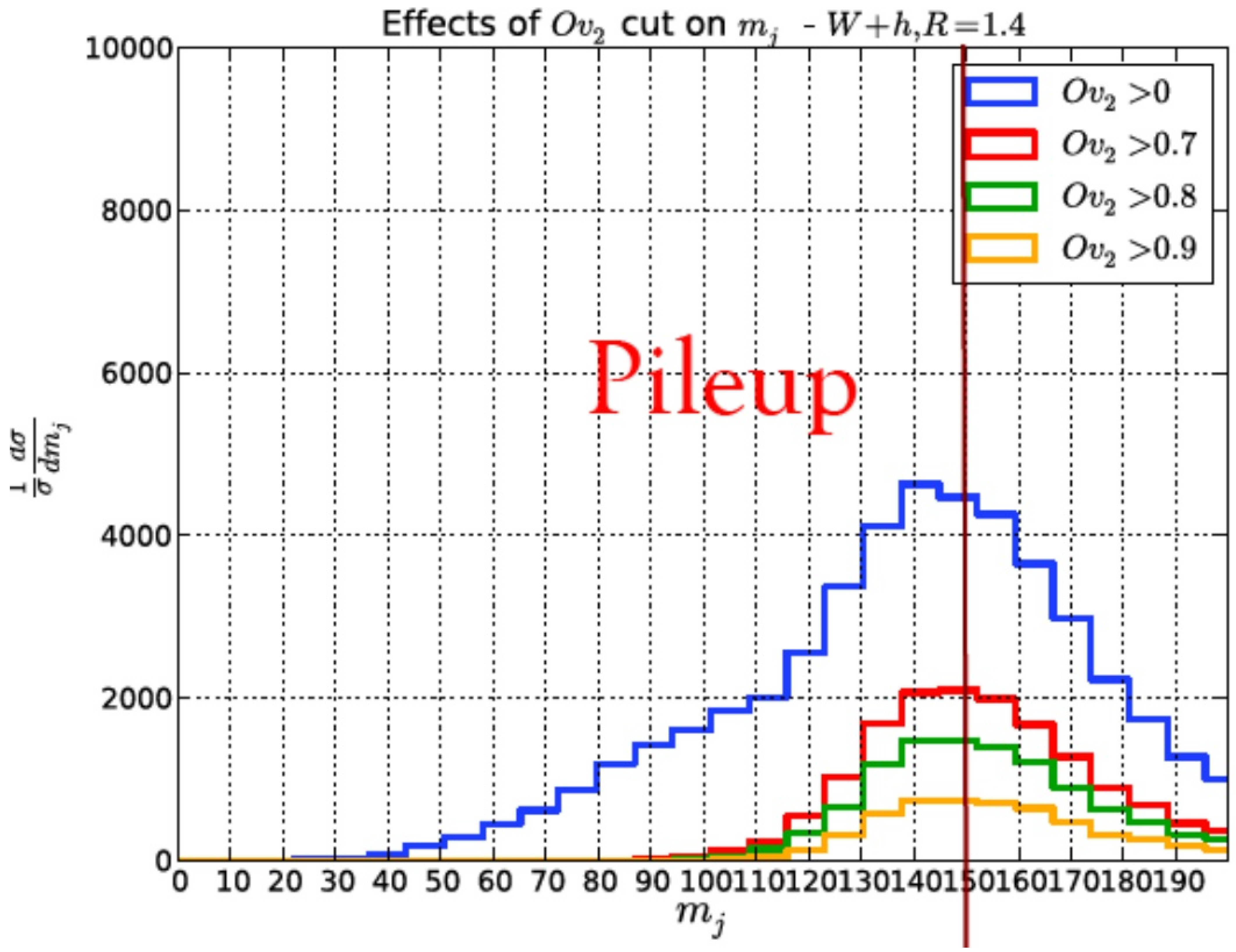


Backovic, JJ, Perez, Winter in preparation

Impact of pileup on jet mass

- pileups/UE results in extra energy inside jets
- How to correct for the contamination in events?

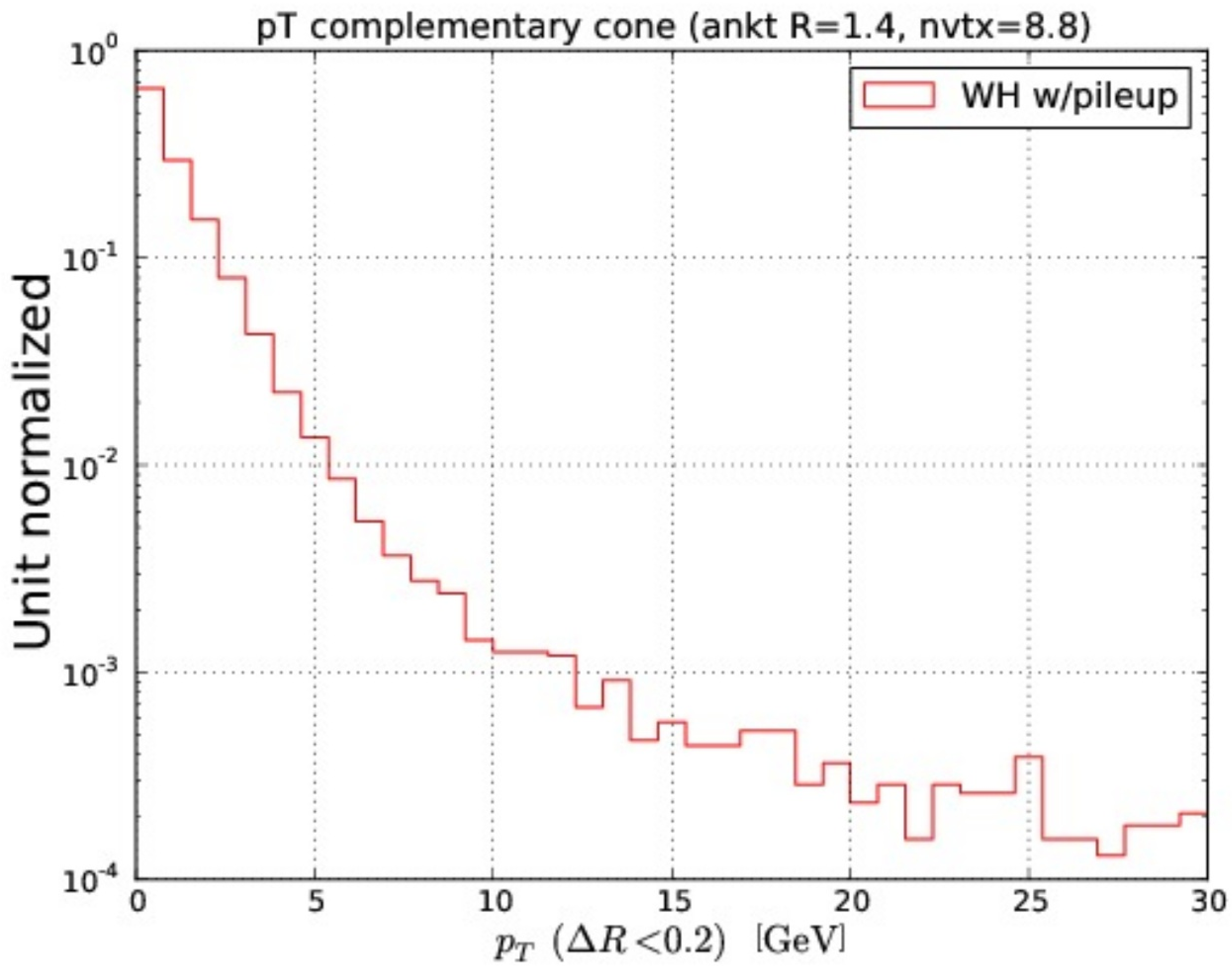




nxvt~9

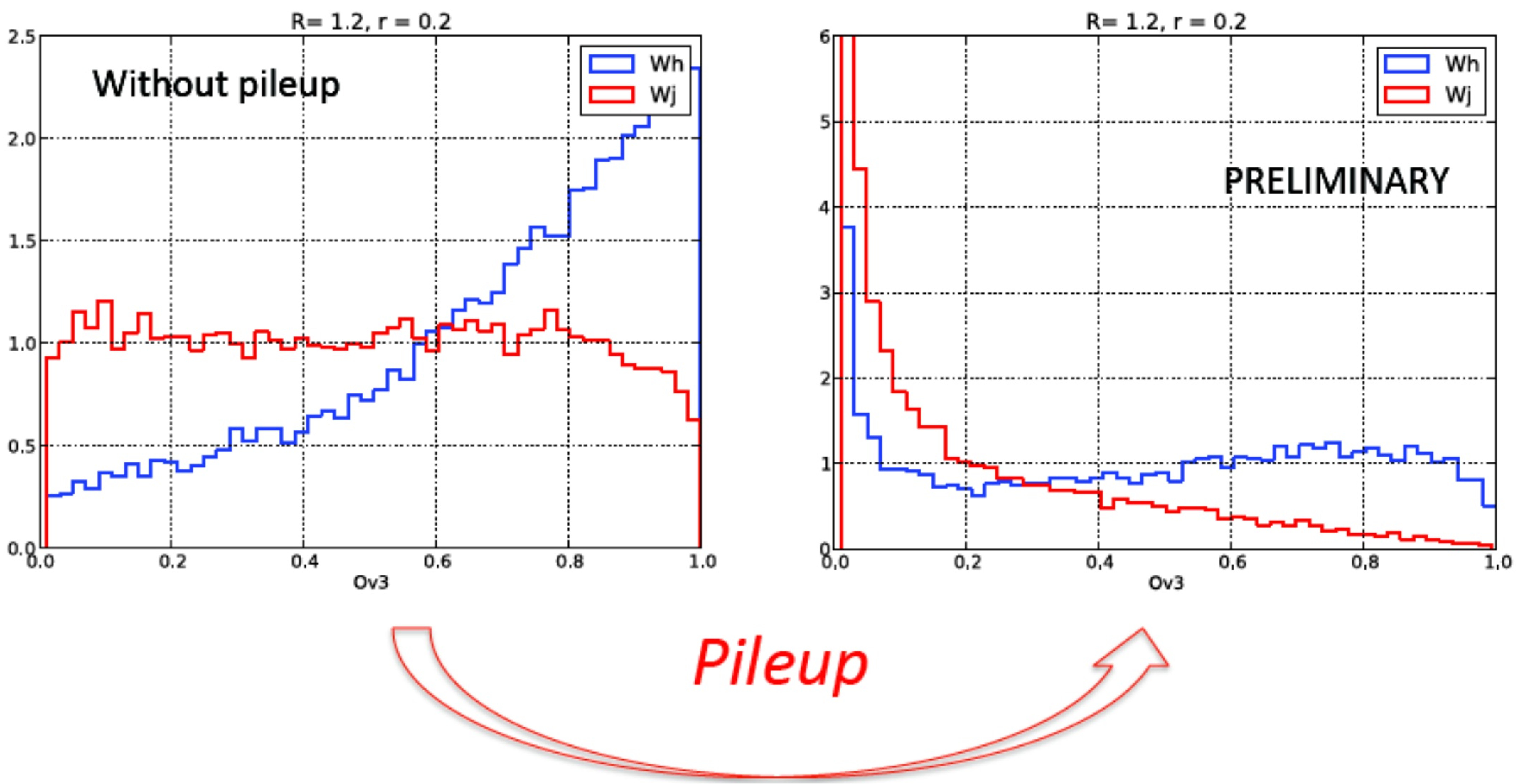
Quantifying impact of pileup

- The template overlap method looks for spikes of enegy inside a large jet
- $\sum$pT(deltaR<0.2) distribution in a complementary cone (140^{o} from the highest pT fat-jet) suggests small effect of the pile up on the sub-jet pT and therefore on the maximum overlap value



Impact of pileup on overlap

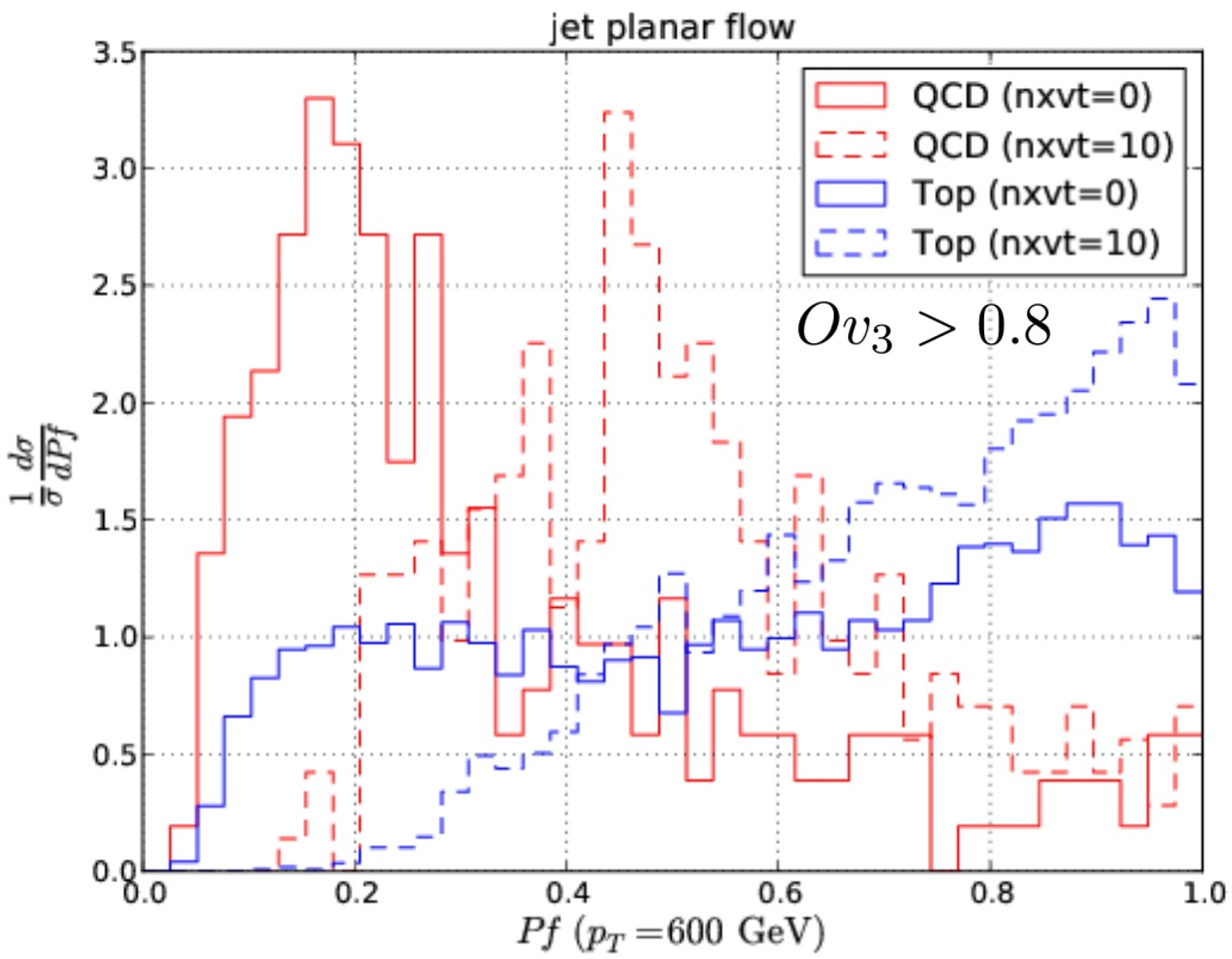
- As expected, pileup decreases the maximum overlap value

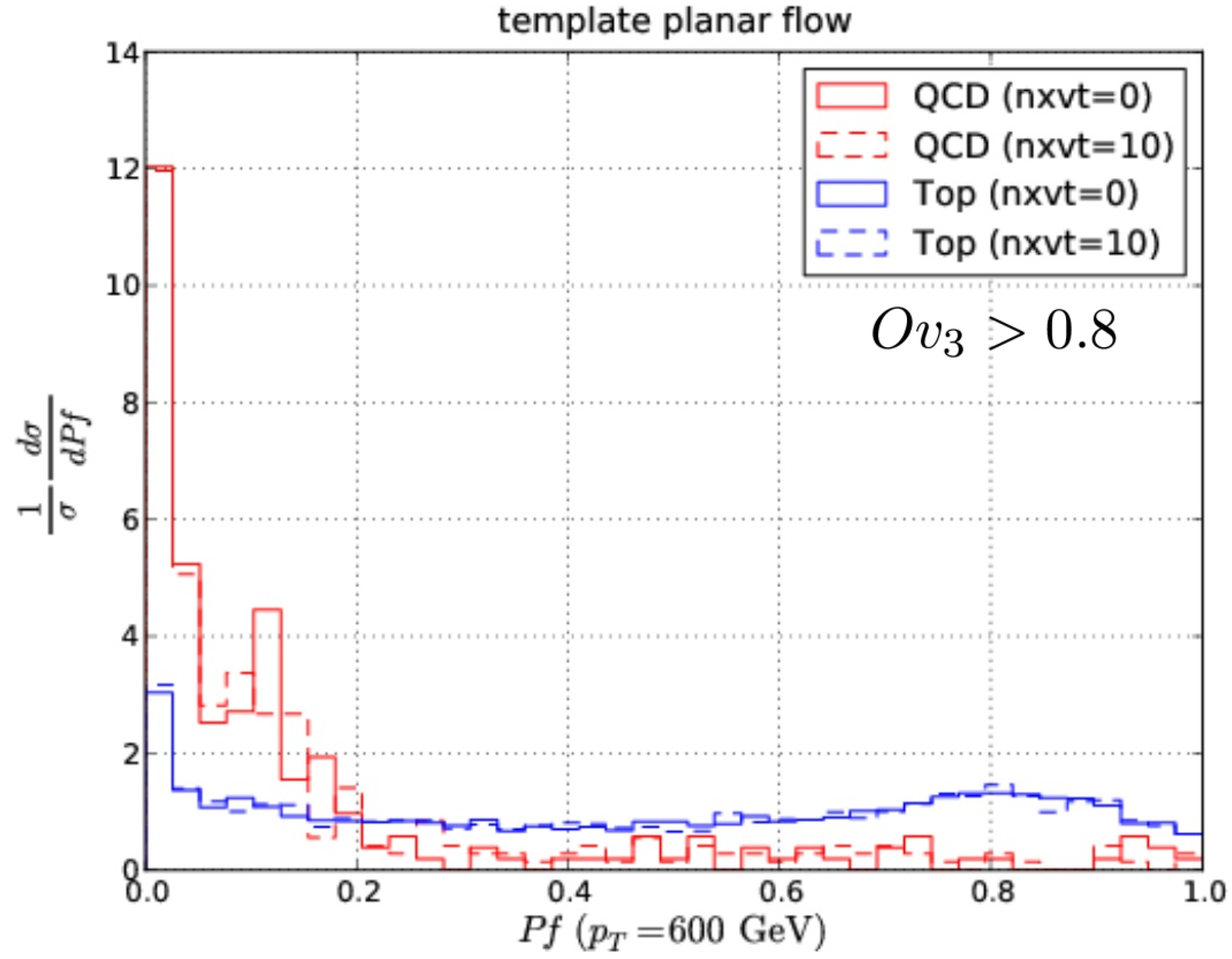


Template "jet shapes" at work

- Pile up yields lots of soft incoherent deposition
- Does not affect the spiky hard part of the signal

Example: Planar Flow $$Pf = \frac{4\,\det(I_\omega)}{\mathrm{tr}(I_\omega)^2}$$

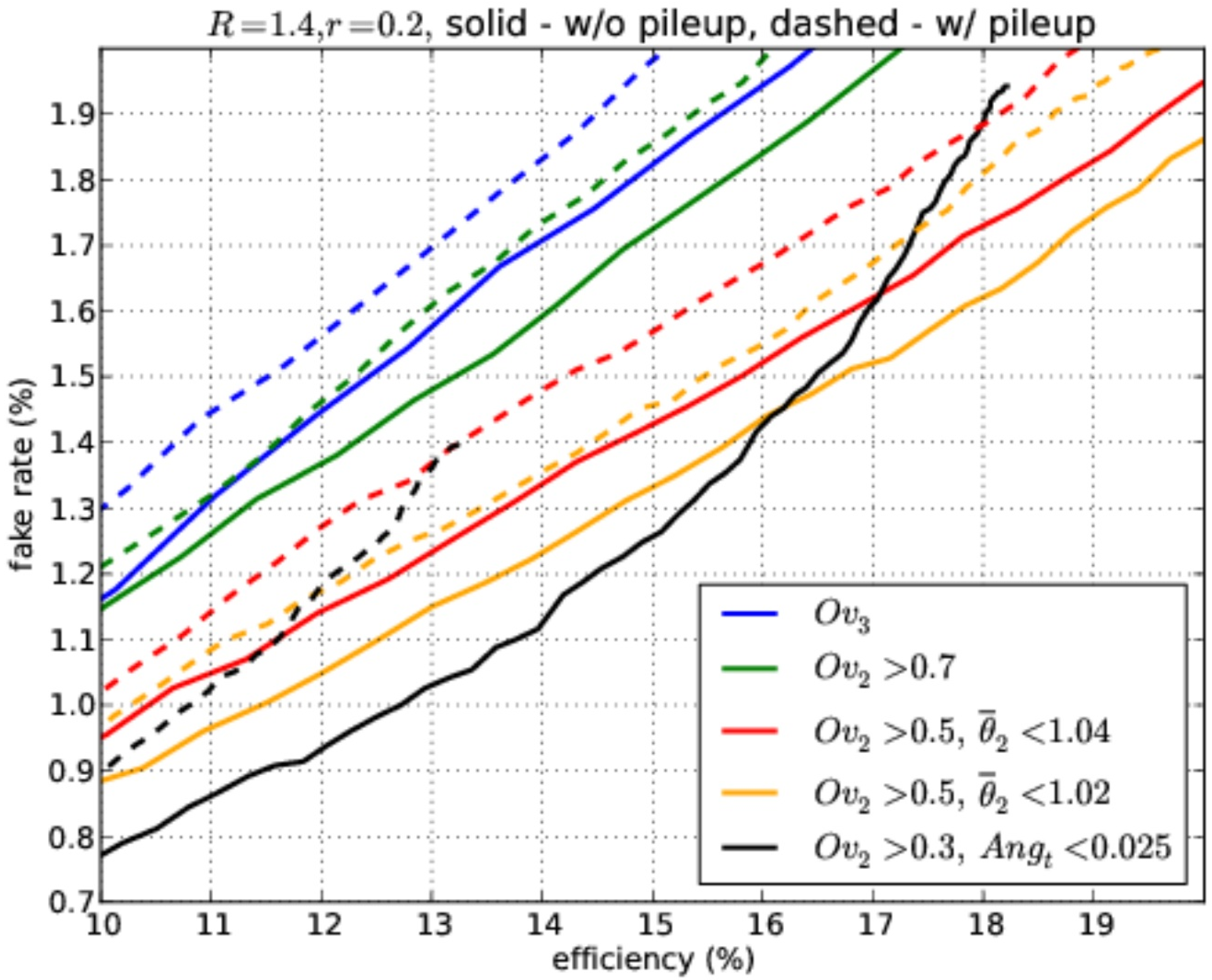




- Jet shapes computed from templates robust to pileup
- Can be used to improve top tagging

Revisiting boosted Higgs

- With pileup without pileup substruction applied
- Effects of pileup not severe (at nxvt ~ 9)



Summary

- The Template Overlap method is versatile enough to work for a range of processes for which theoretical models have been established
- Useful for events where the energy distribution is all that is available
- Flexibility to add both theoretical and experimental information
- Code coming out soon