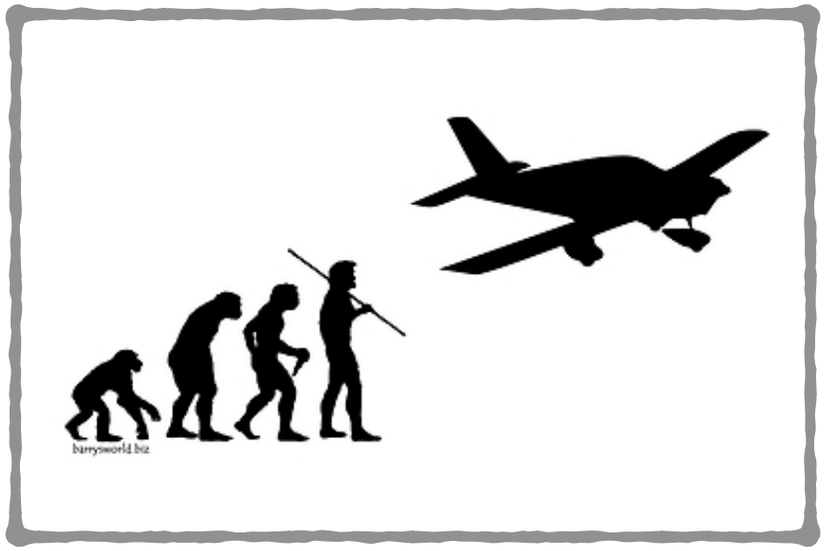
the evolution of jets at the LHC:

jets, subjets and particles



Nhan Tran
Fermi National Accelerator Laboratory

Fermilab Wine and Cheese Seminar
April 25, 2014

outline

introduction: jets at the LHC

a jet substructure (r)evolution

case studies

high mass resonances, $H \rightarrow W^+W^-$, WW scattering

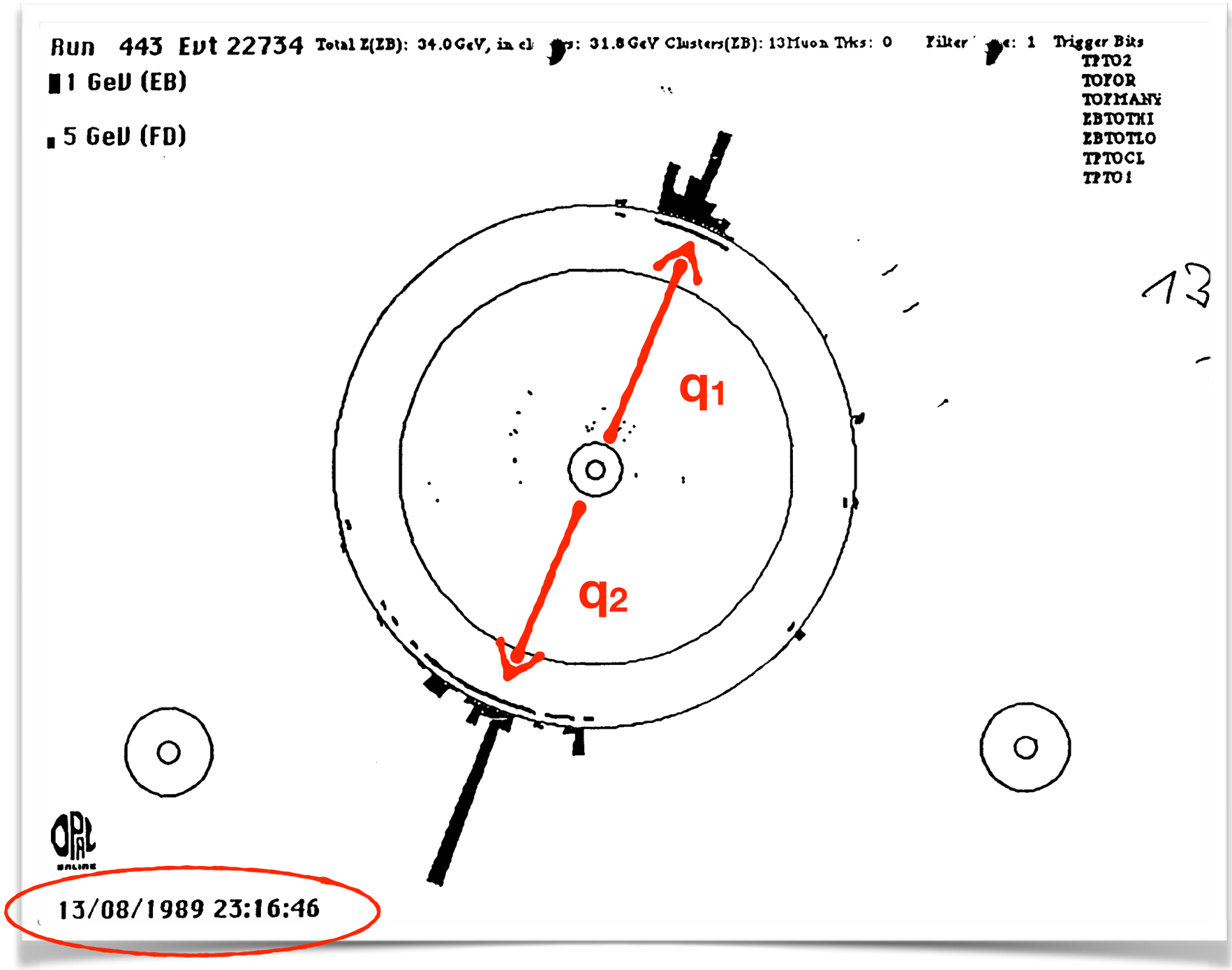
through the looking glass

increasing pileup, new handles

disclaimer: Slides may be a little CMS-biased, but both CMS&ATLAS doing great work

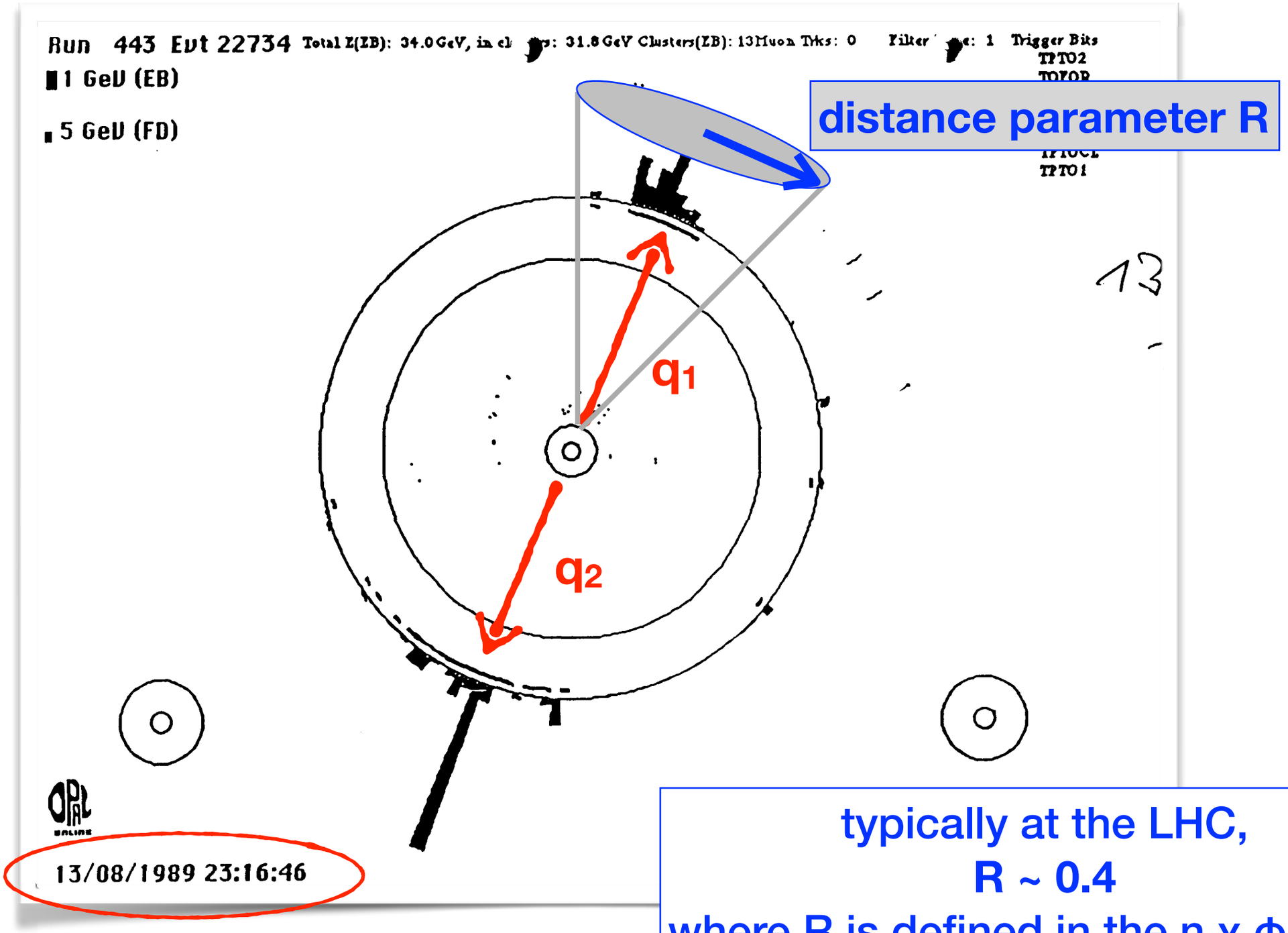
jets as four-vectors

The first hadronic Z event at OPAL



jets as four-vectors

The first hadronic Z event at OPAL



typically at the LHC,
R ~ 0.4
where R is defined in the η χ φ space

jet substructure

the way we treat jets is evolving

a jet is no longer a four-vector

we are pulling them apart, looking at their structure, shapes and properties of subjets

information is used to classify new kinds of jets

wide array of applications

necessary for the physics program at the LHC

LHC era jets

…"necessity is the mother of invention"

more energy → boost

as the center-of-mass energy of the LHC increases, objects will be produced with more transverse momentum

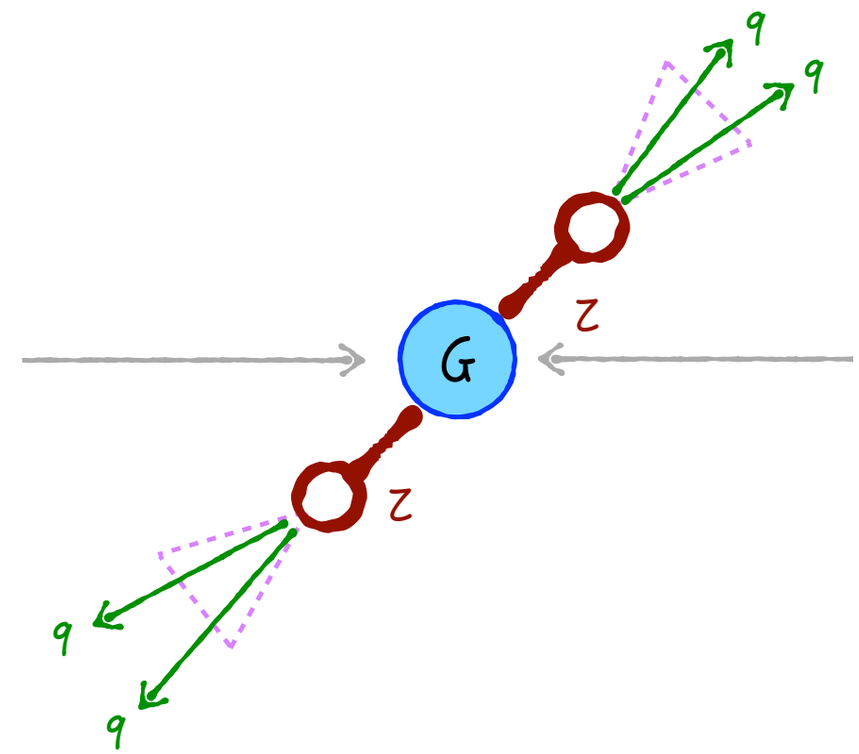
more luminosity → pileup

as the instantaneous luminosity of the LHC increases, we will be inundated with additional pp interactions on top of our process of interest

at the energy frontier

increased reach to new energy regimes

SM particles are produced with more p_T



why quarks and gluons?

boosting statistics

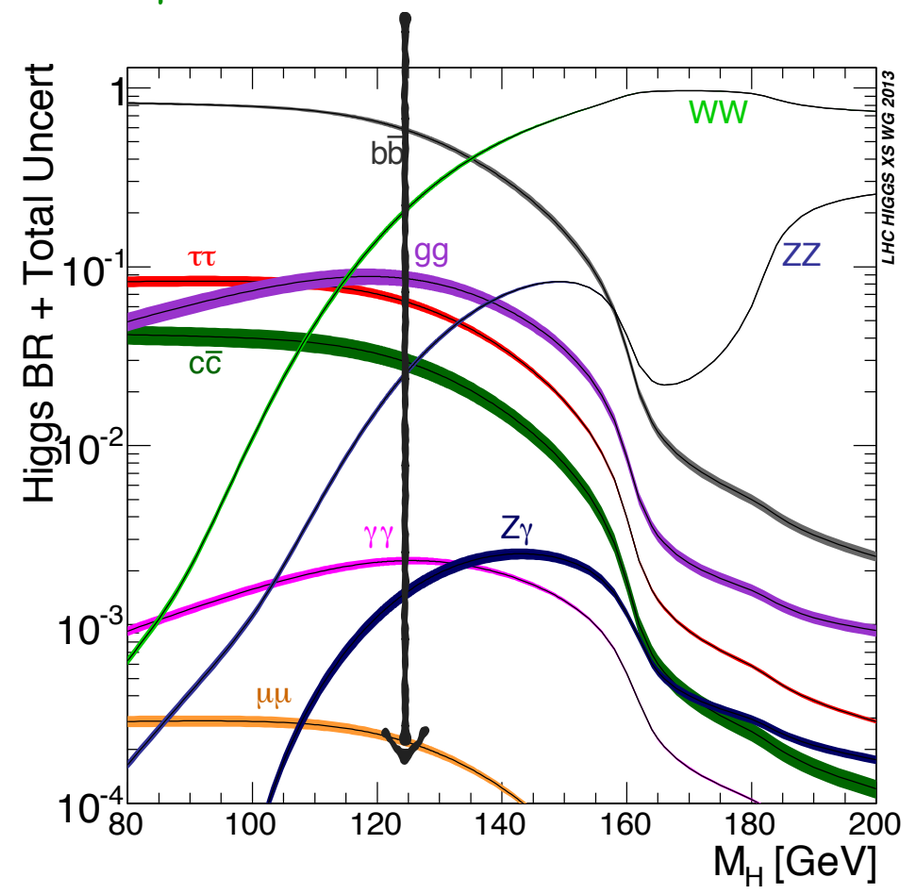
H→bb = 57%

W→qq ~ 67%, W→lν = 22%

(t→Wb ~ 100%), l = e,μ

Z→qq ~ 60%, Z→ll ~ 6%

jet backgrounds are manageable at high pT



pileup

LHC Pileup scenarios…

past:

2011, $\langle n_{PU} \rangle \sim$ **10**

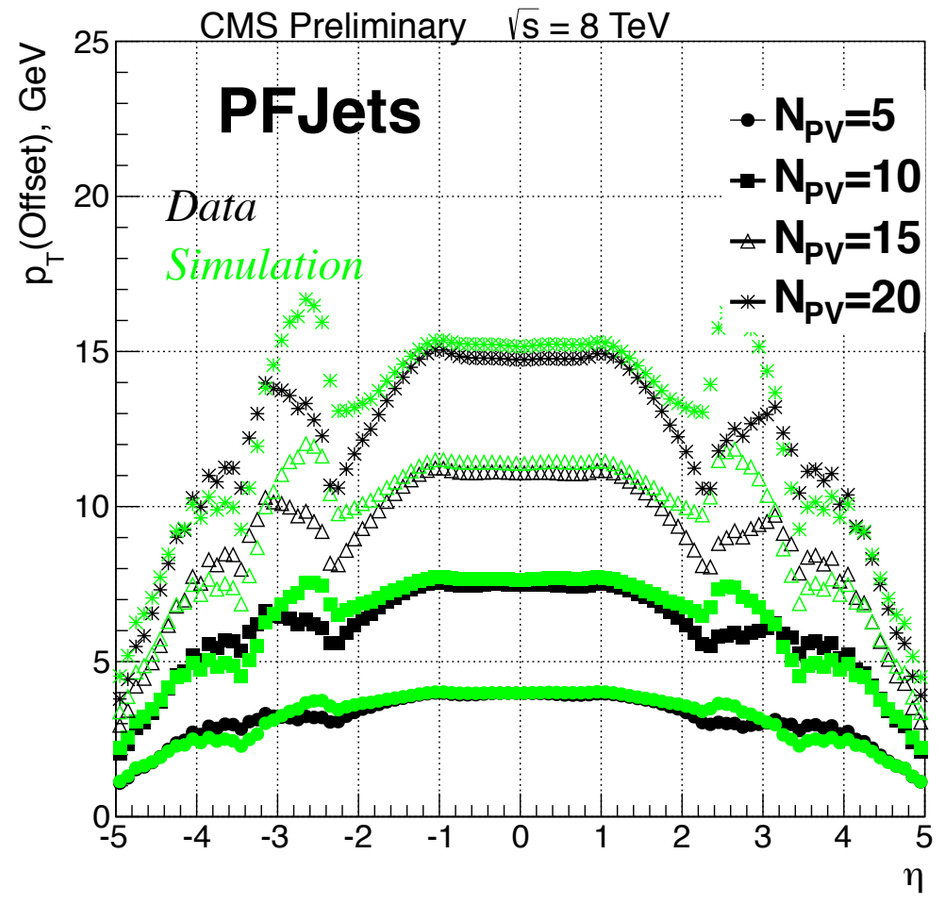
2012, $\langle n_{PU} \rangle \sim$ **20**

future:

~2015, $\langle n_{PU} \rangle \sim$ **20-40** (25/50 ns) ?

~2018, $\langle n_{PU} \rangle \sim$ **80** (25 ns) ?

~2023, $\langle n_{PU} \rangle \sim$ **140** (25 ns) ?

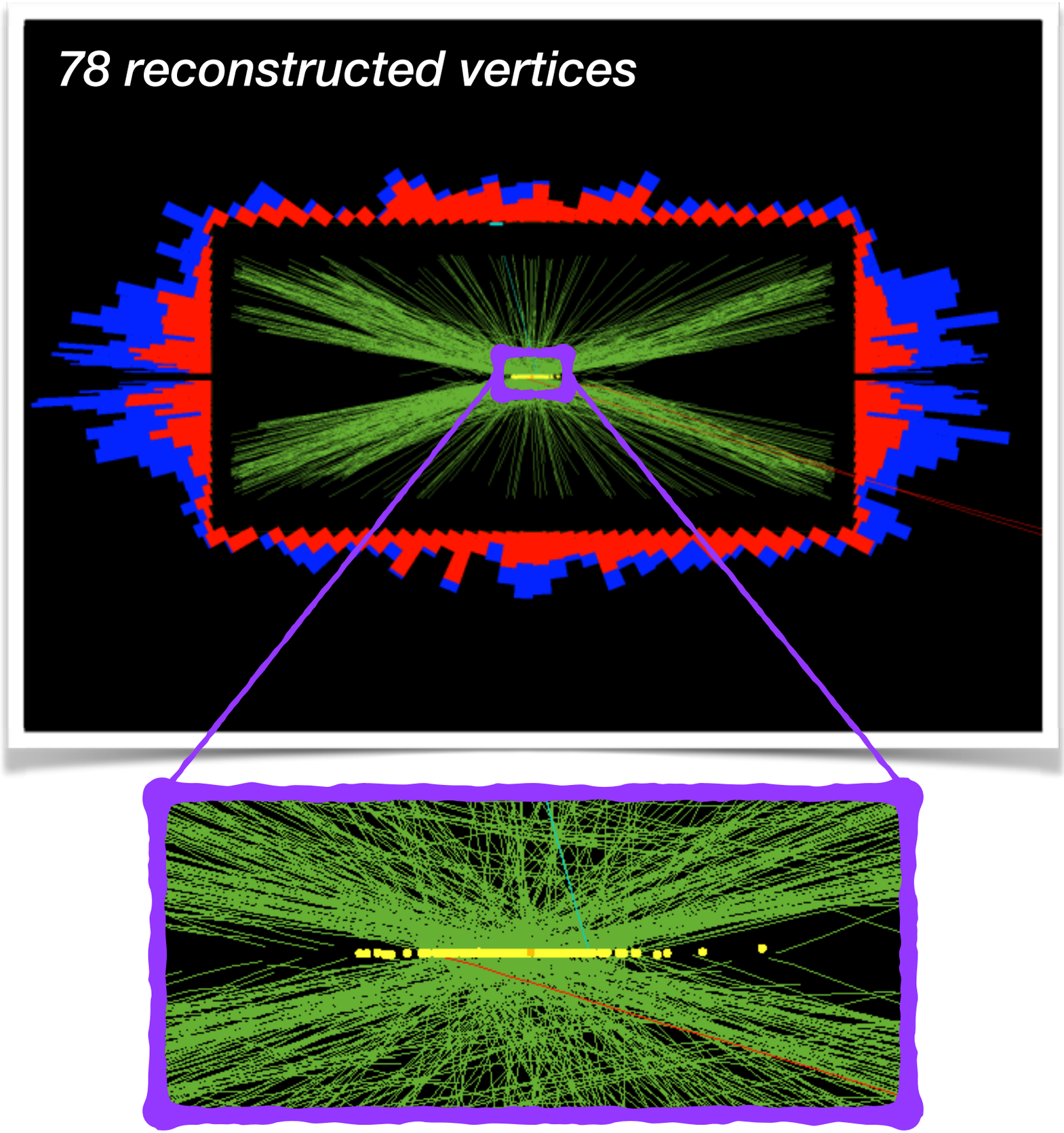


A pileup interaction adds on average ~1 GeV/unit area to the event

(unit area in $\Delta\eta \times \Delta\phi$)

@ $\langle n_{PU} \rangle$ = 40, a 30 GeV jet has additional 30 GeV of pileup energy

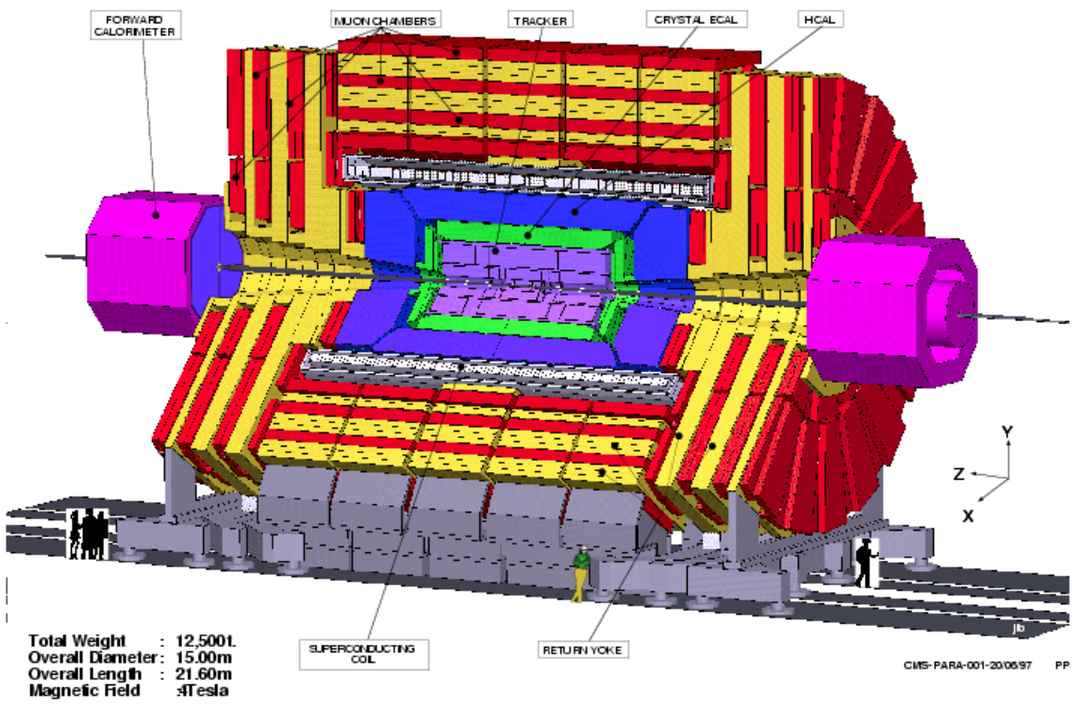
high PU event at the LHC

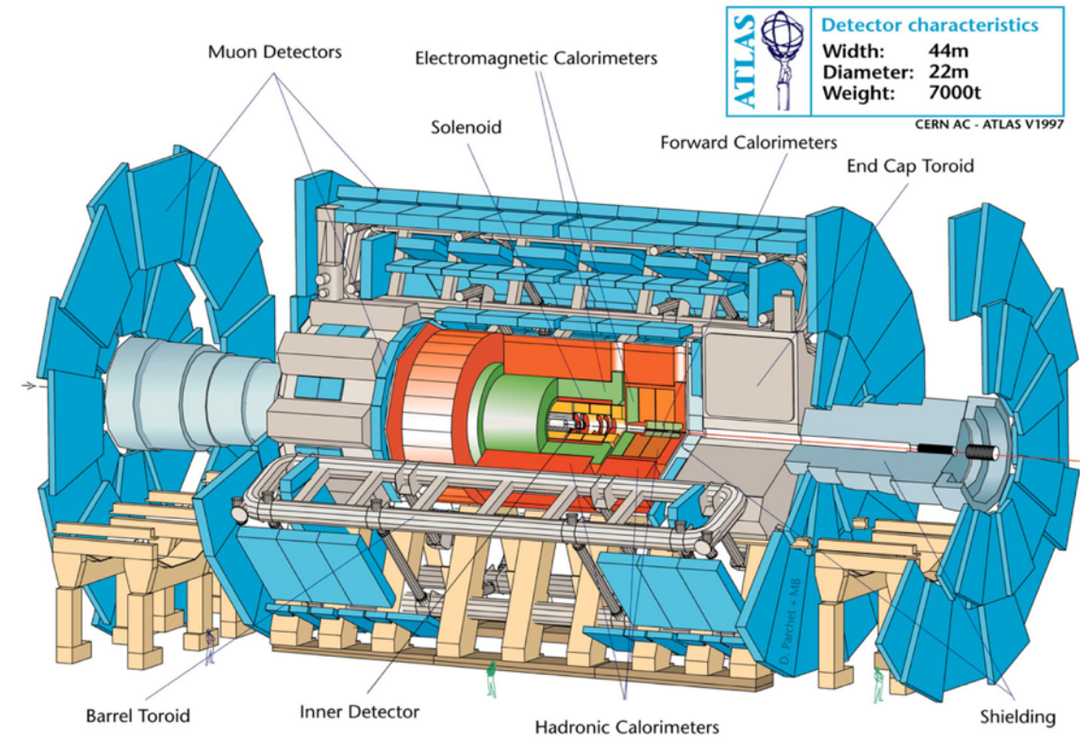


a jet (r)evolution:

these ain't your mama's jets

inputs to jets

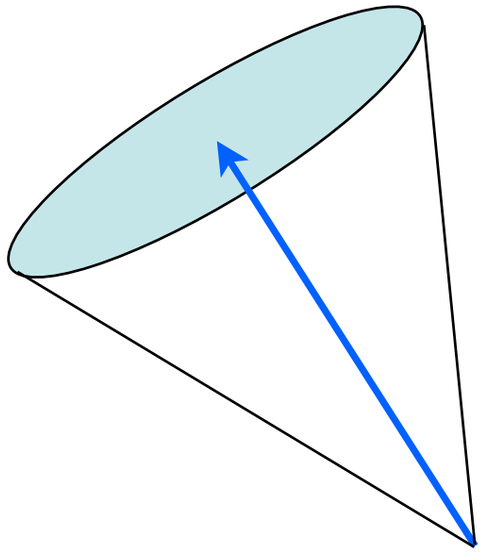




Particle flow algorithm uses high granularity ECAL and all-silicon tracker to provide inputs of **neutral** and **charged hadrons**, **electrons**, **muons** and **photons**

High granularity, depth-segmented EM and hadronic calorimeter, inputs from **3D topological clusters** with good energy resolution

a jet (r)evolution

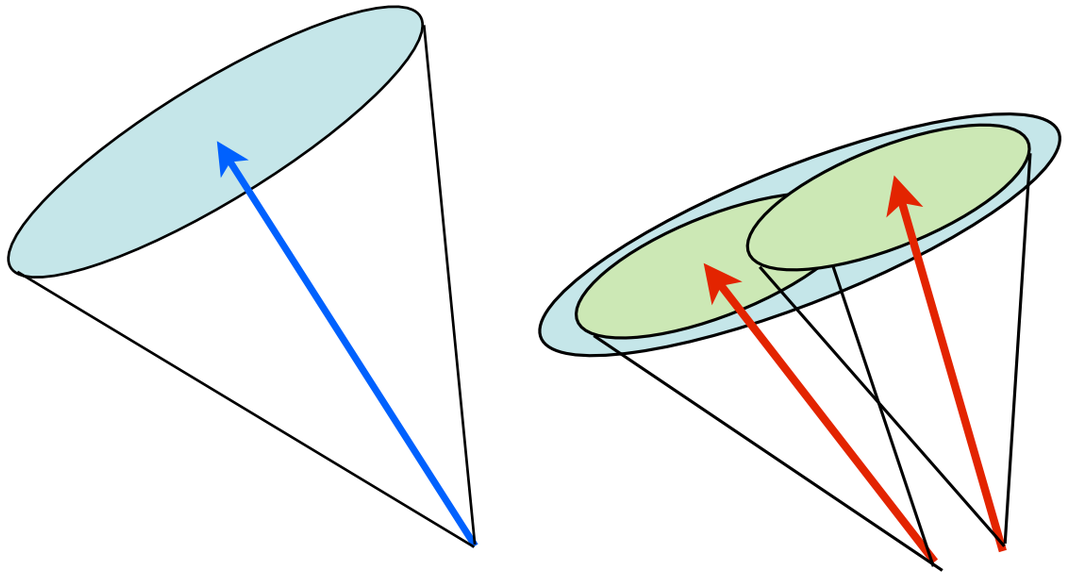


udsg/c/b

{Y,ϕ,p_T}
+
{tracking}

“flavor”-tagging:
b-tagging
c-tagging
uds-tagging

a jet (r)evolution



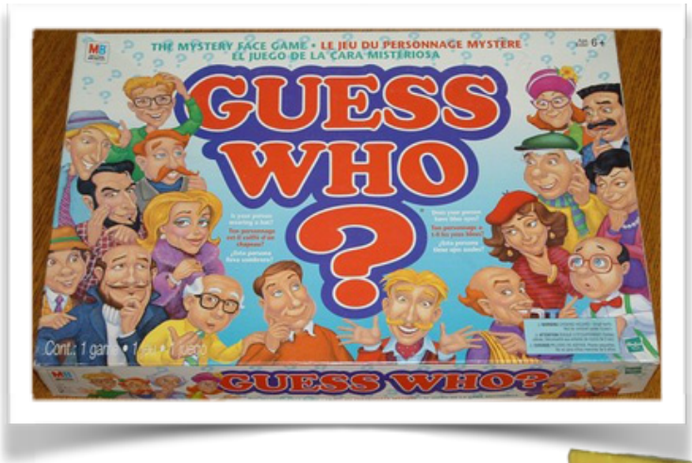
u/ds/g/c/b/W/Z/H/t/pu

quantum numbers:
color charge (quarks vs. gluons)
electric charge
spin

{Y,ϕ,p_T}
+
{tracking}
+
{m,shapes,subjets}

"flavor"-tagging:
b-tagging
c-tagging
u/ds-tagging
top-tagging
W/Z/H-tagging
pileup-tagging

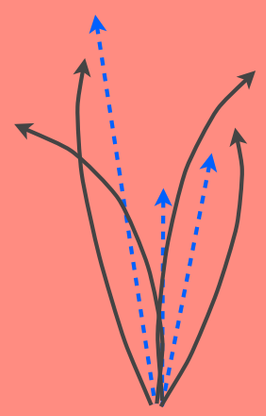
identifying features



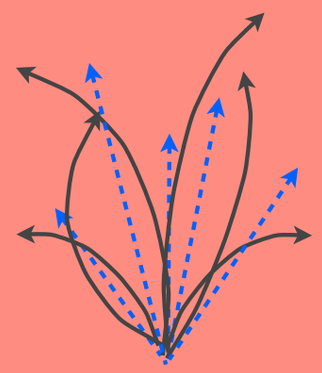
{gender, ethnicity, age}
+
{hair color, headwear}
+
{glasses, facial hair}

{Y,ϕ,p_T}
+
{tracking}
+
{m,shapes,subjets}

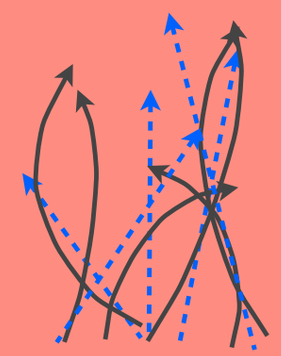
guess who?



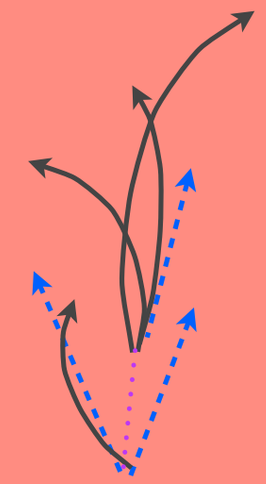
u,*d* or *s* jet



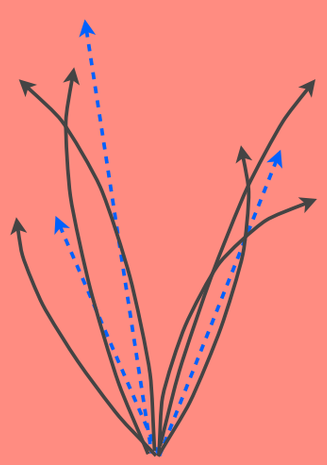
gluon jet



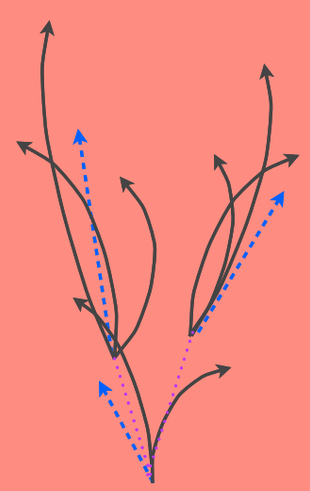
pileup jet



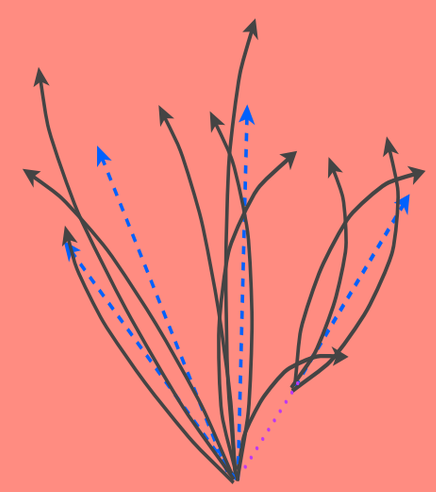
c or *b* jet



W or *Z* jet



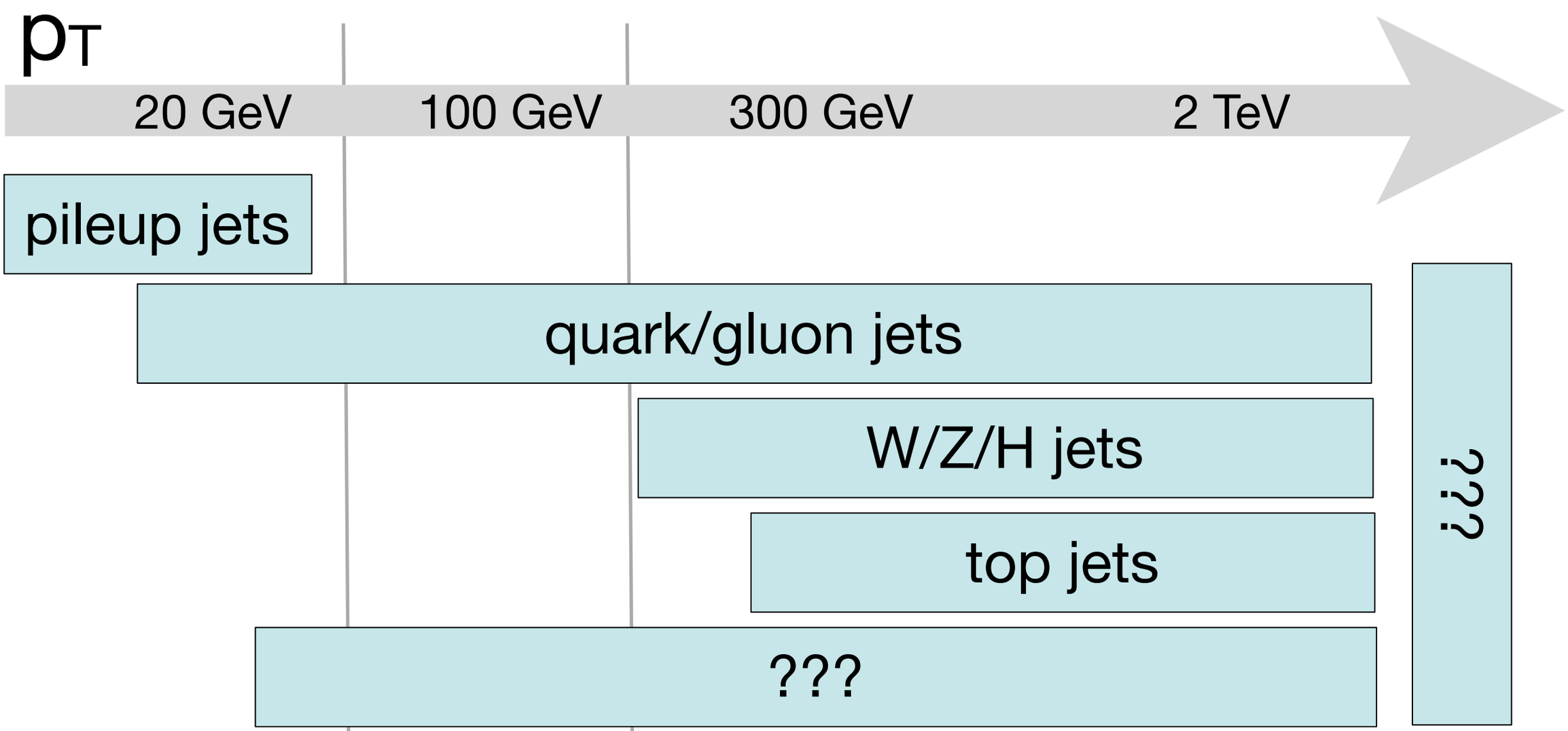
Higgs jet



top jet

?

jet phase space



lower pT regime

moderate pT regime

high pT regime

jet vetoes
VBF jets
high quark-jet multiplicities??

physics at the kinematic limit
resonances searches
boosted V/H/top

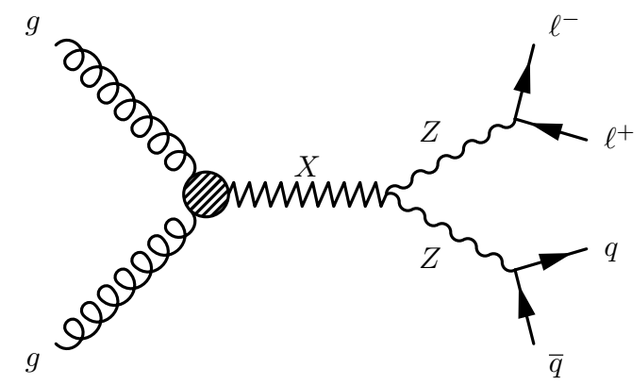
as a general rule,
the boosted regime:
$\Delta R \sim 2m/p_T$

case studies

some physics cases

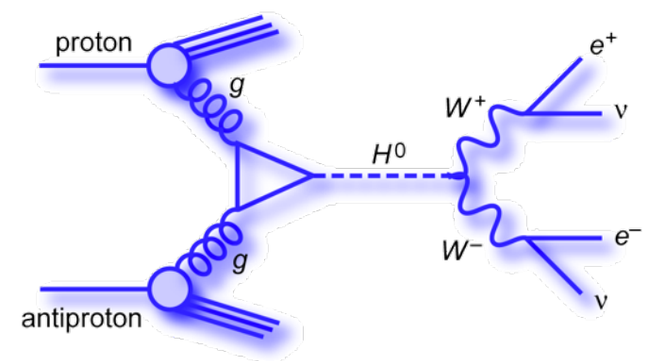
high mass resonances

the *classic* cases…



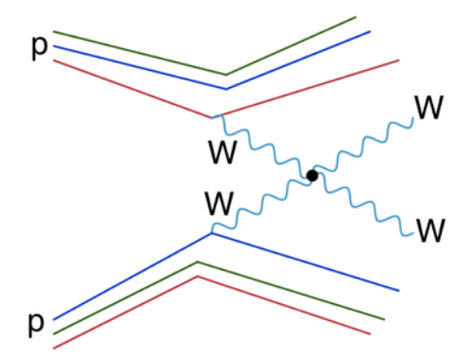
gg → H → W^+W^- → lνlν

***boosting* above the background or low p_T jet counting**



WW scattering, WWjj → lνJjj

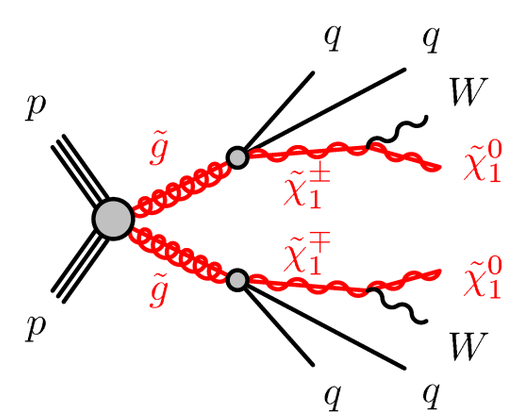
a W tagged jet + VBF jets



SUSY with low MET

in the bonus material

high quark jet multiplicities and little or no missing energy



case #1:

high mass resonances

high mass resonances

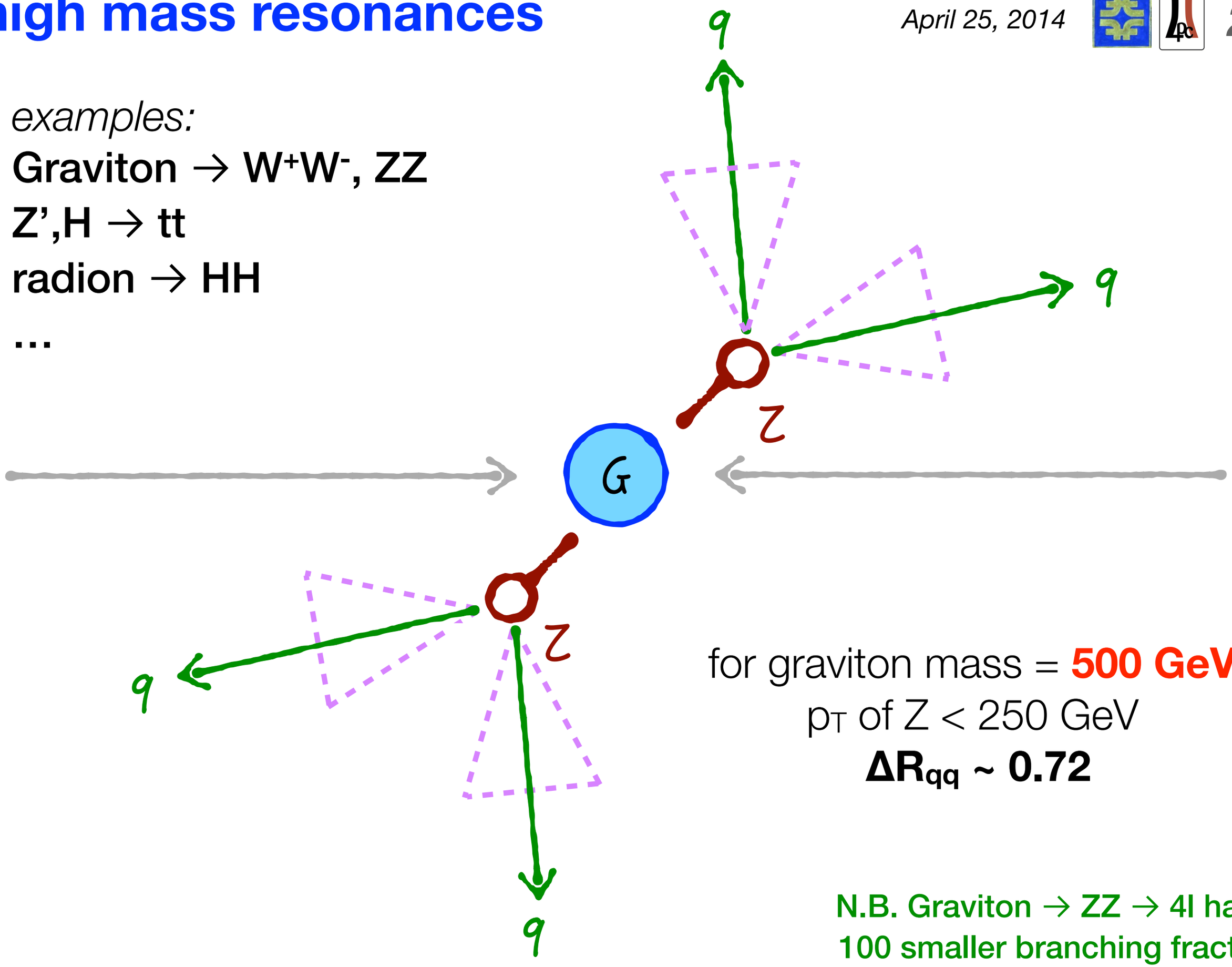
examples:

Graviton → W^+W^-, ZZ

Z',H → tt

radion → HH

...



for graviton mass = **500 GeV**

p_T of Z < 250 GeV

ΔR_{qq} ~ 0.72

N.B. Graviton → ZZ → 4l has a 100 smaller branching fraction

high mass resonances

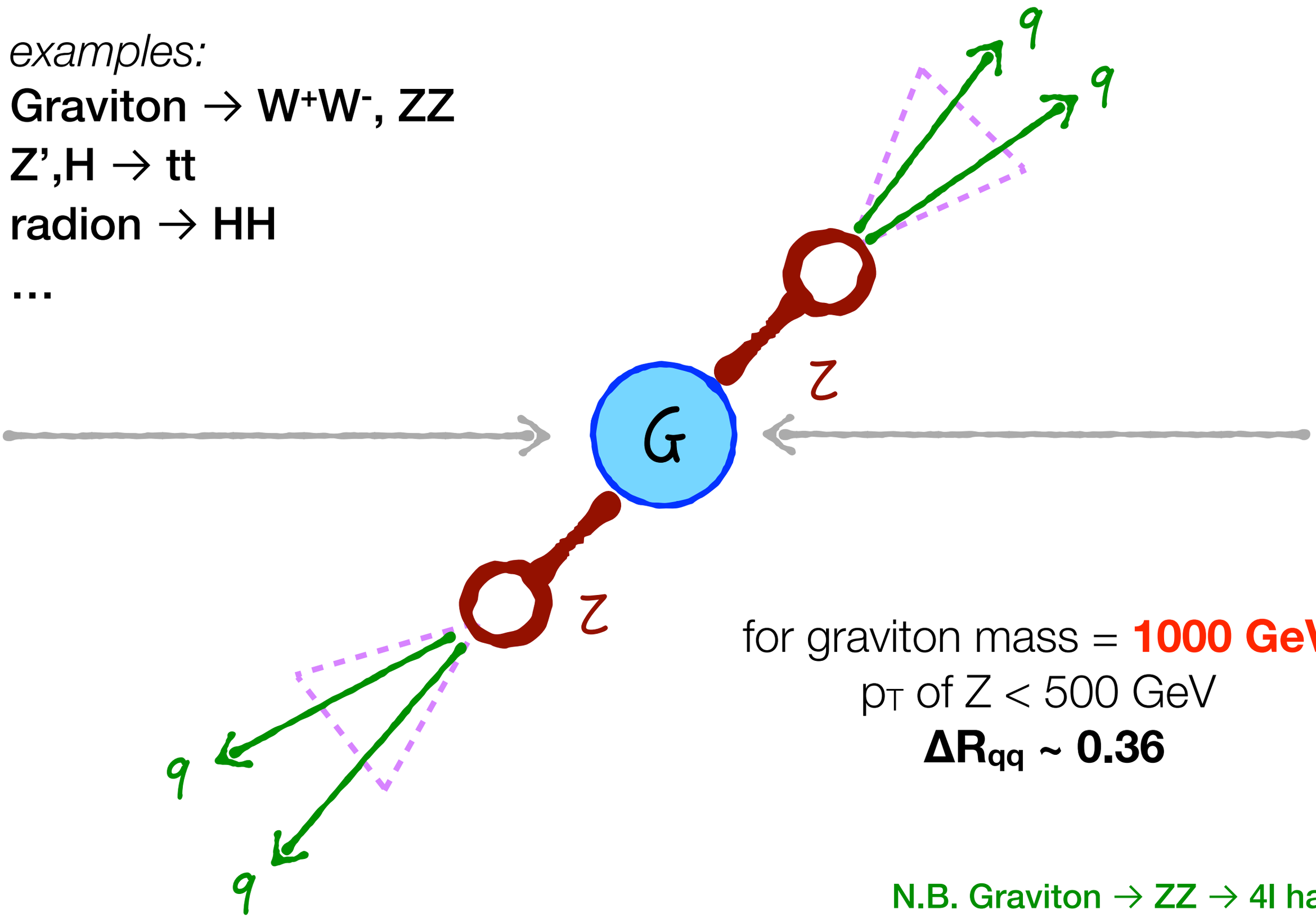
examples:

Graviton → W^+W^-, ZZ

Z',H → tt

radion → HH

...



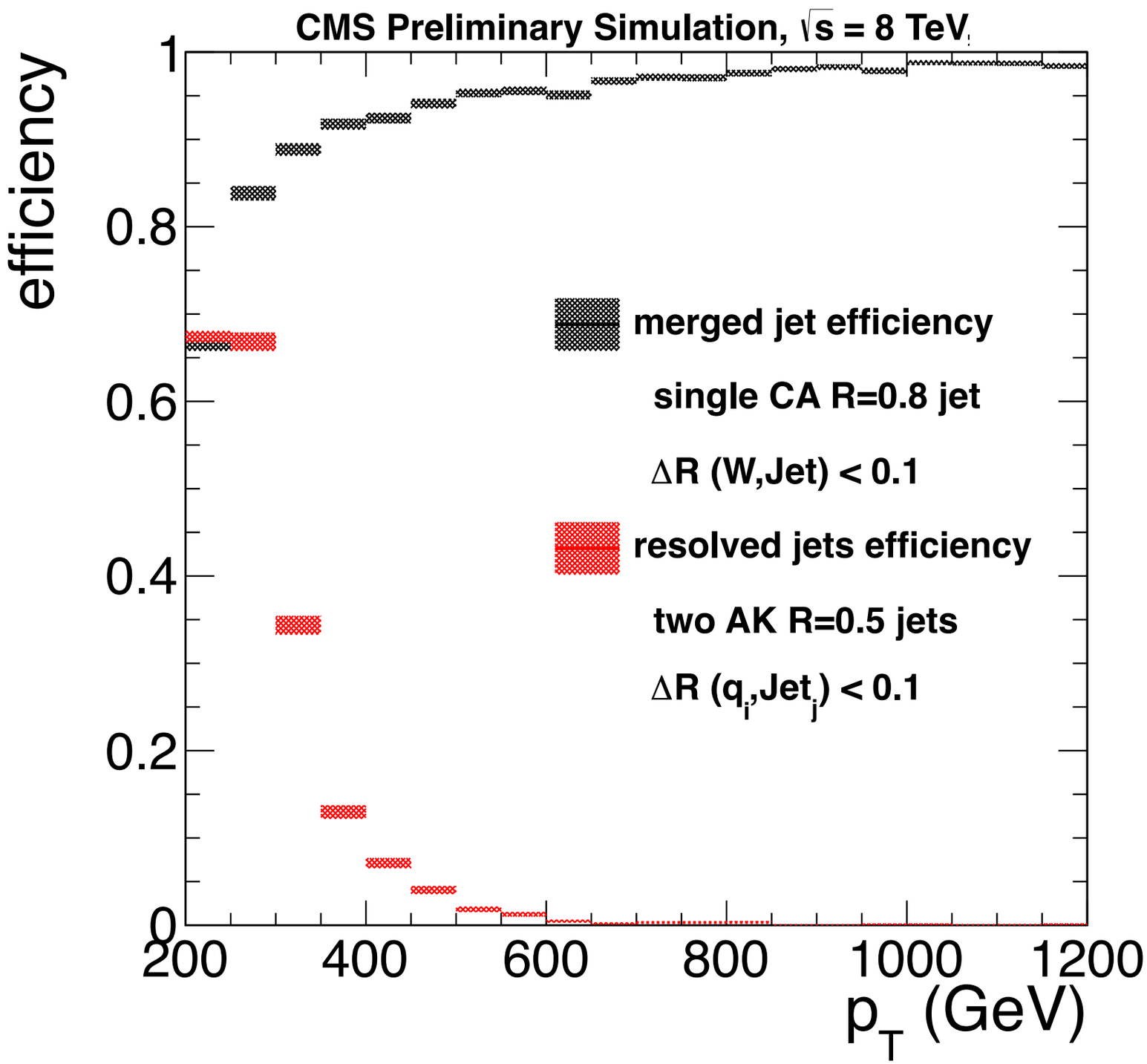
for graviton mass = **1000 GeV**

p_T of Z < 500 GeV

ΔR_{qq} ~ 0.36

N.B. Graviton → ZZ → 4l has a 100 smaller branching fraction

Efficiency for clustering W → qq as a function of p_T of W



When the W boson has p_T > 400 GeV, it nearly always fits inside a cone of R = 0.8

When the W boson has p_T > 400 GeV, it cannot be constructed with 2 R=0.5 jets

tagging W and Z jets

separating a W or Z jet from a high p_T quark or gluon jet

1. jet mass
2. W and Z's have 2 prong structure

tagging W and Z jets

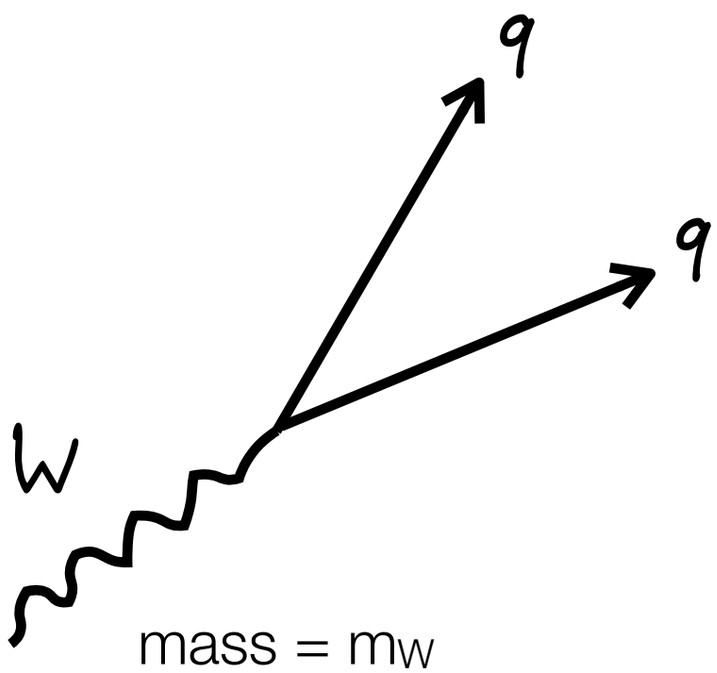
separating a **W** or **Z jet** from a high p_T **quark** or **gluon** jet

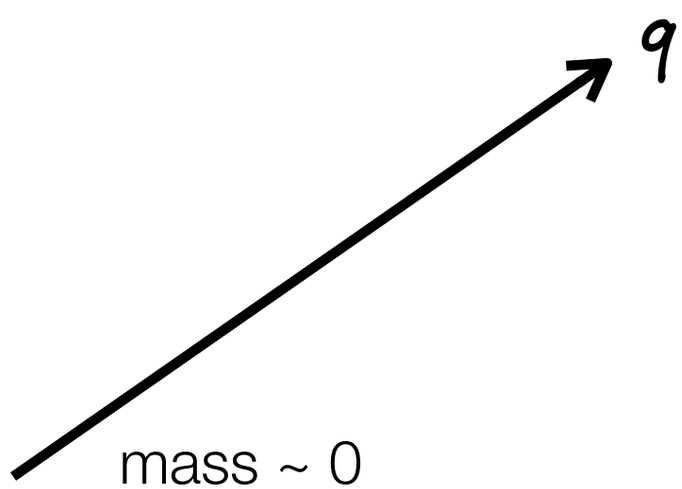
1. **jet mass: four-vector sum of particles in a jet**
2. W and Z's have 2 prong structure

jet mass

mass is a very useful jet shape observable

at parton level, these are pretty easy to tell apart

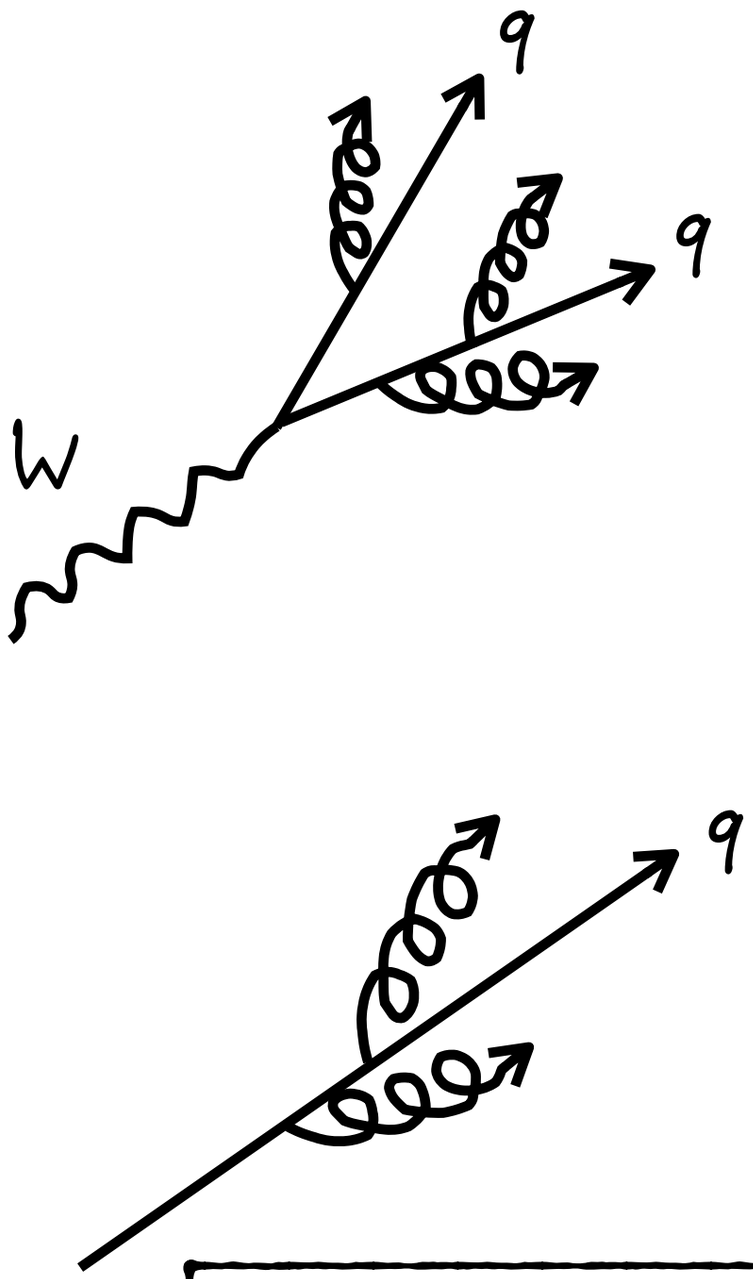


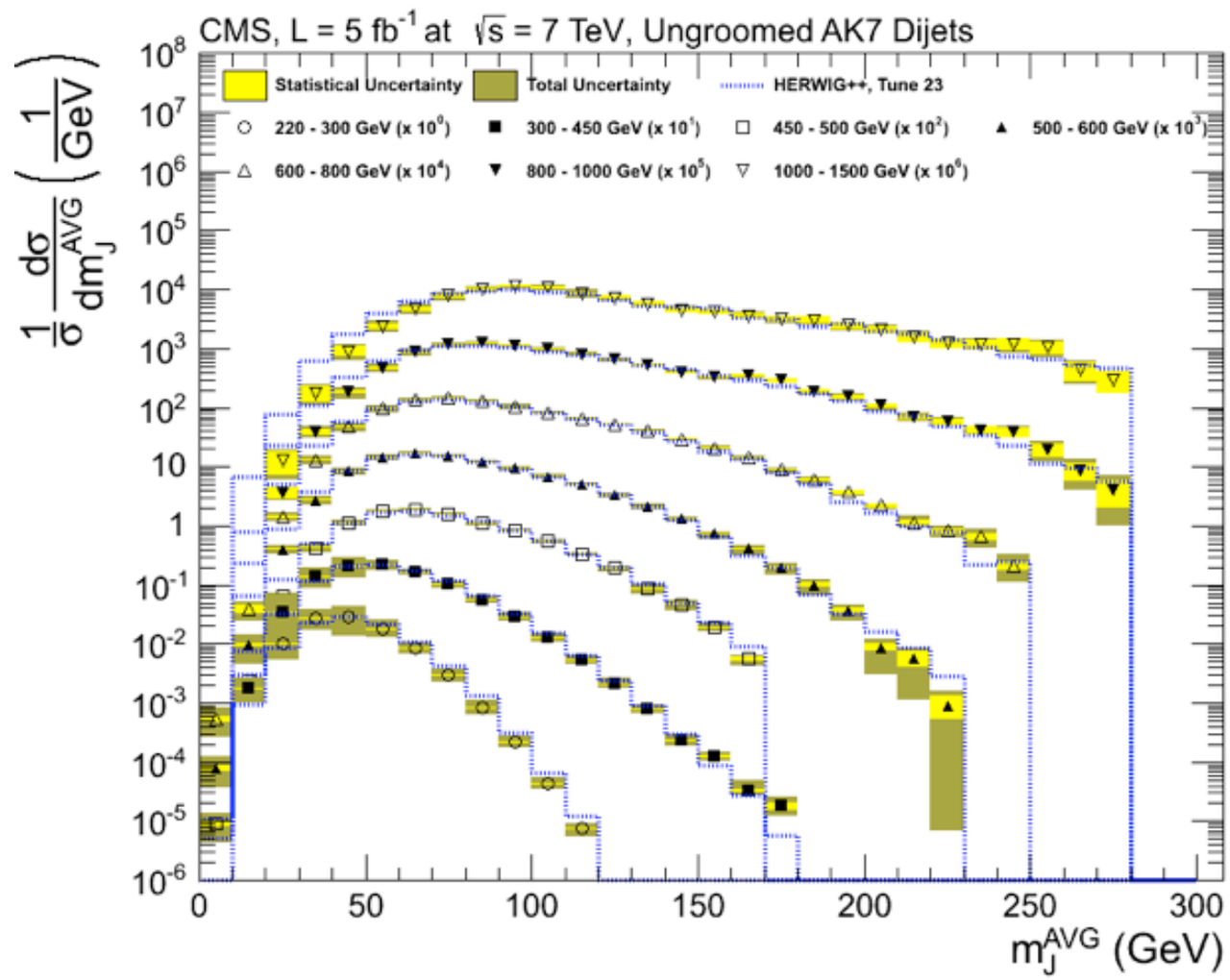


jet mass

JHEP 05 (2013) 090

but jet mass is a perturbative quantity

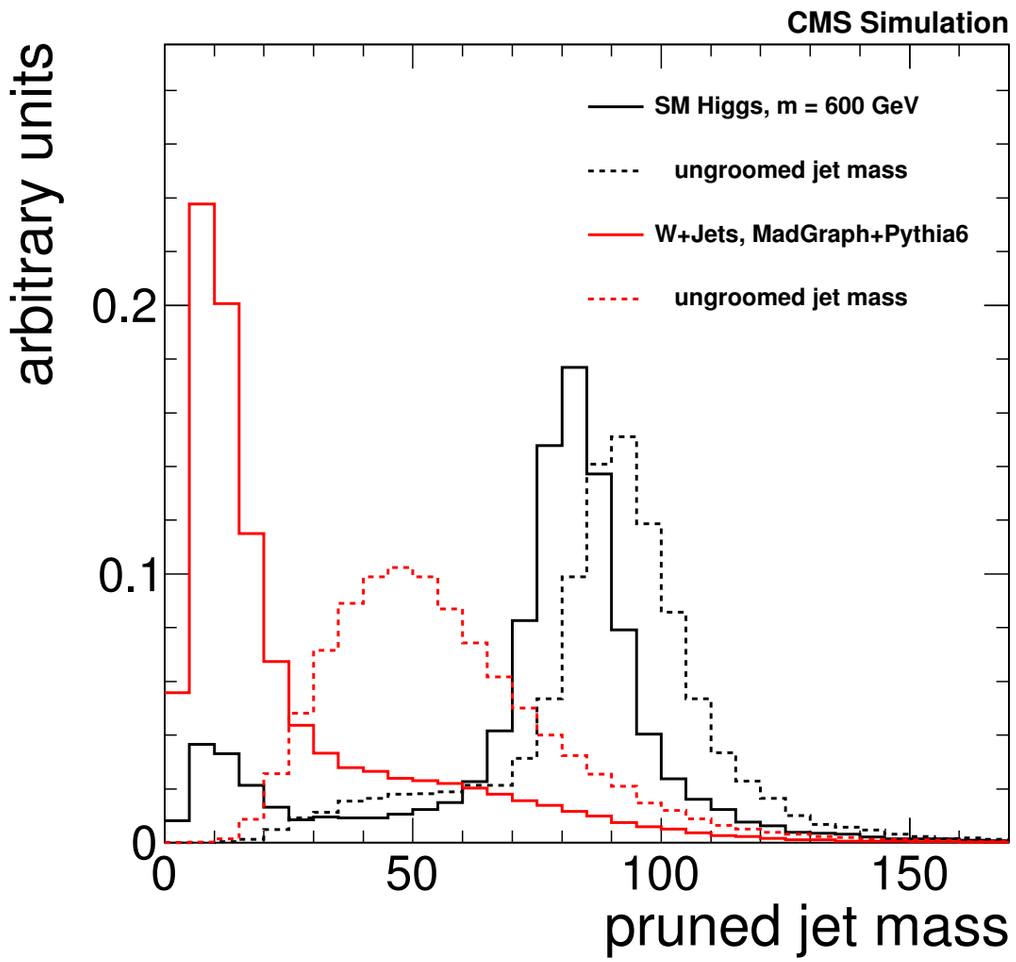


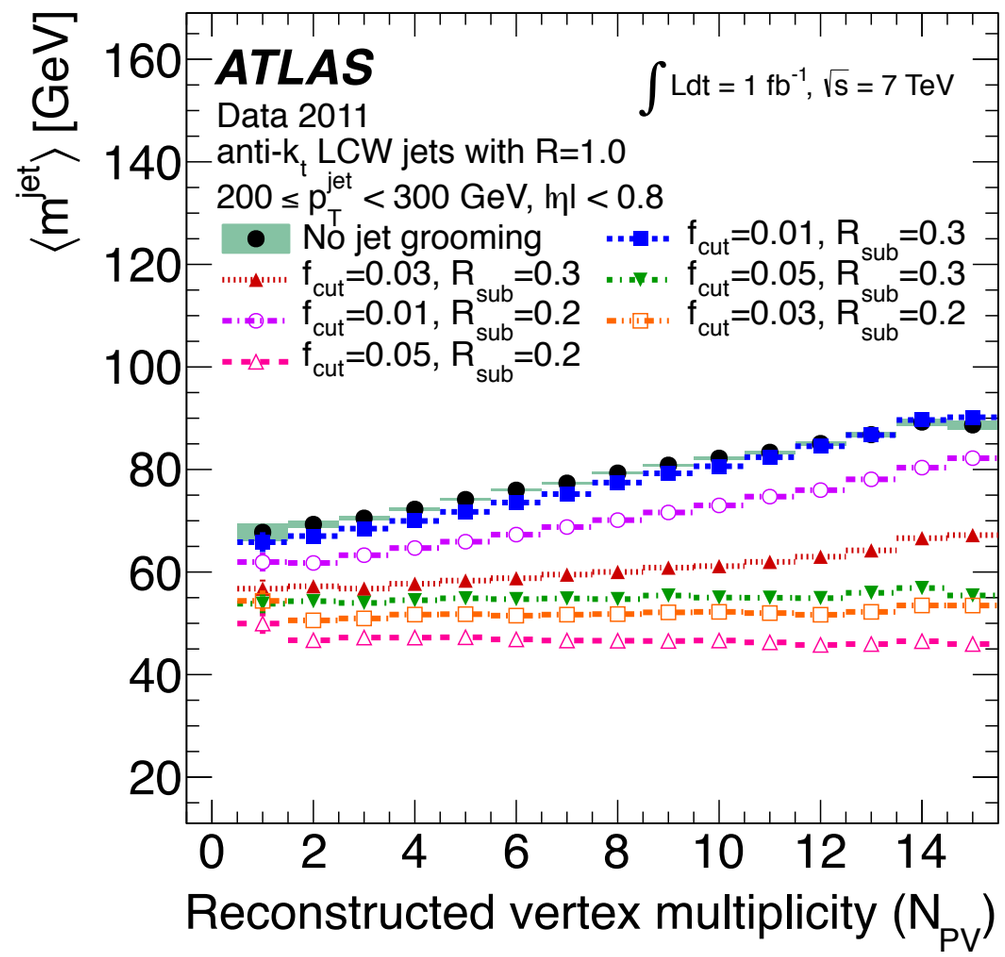


$$\langle M^2 \rangle \simeq \left.\begin{array}{ll}\text{quarks:} & 0.16 \\ \text{gluons:} & 0.37\end{array}\right\} \times \alpha_s p_t^2 R^2$$

groomed mass

CMS-PAS-HIG-13-008
JHEP 1309 (2013) 076



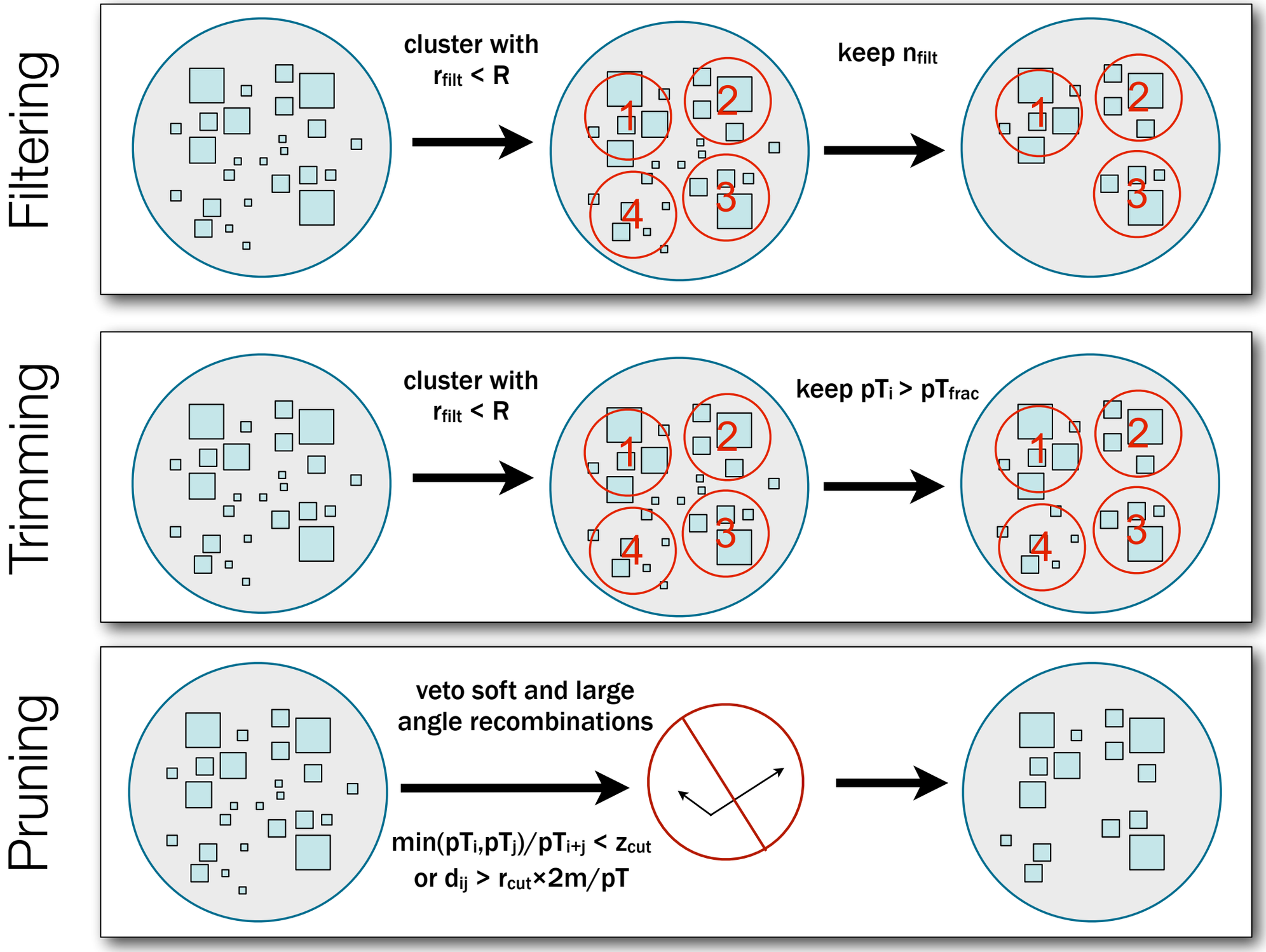


Grooming tends to push the jet mass scale of the **background to lower values** while **preserving the hard scale of the heavy resonance**

Grooming techniques are also vital in **reducing the pileup dependence** of the jet mass

groomed mass

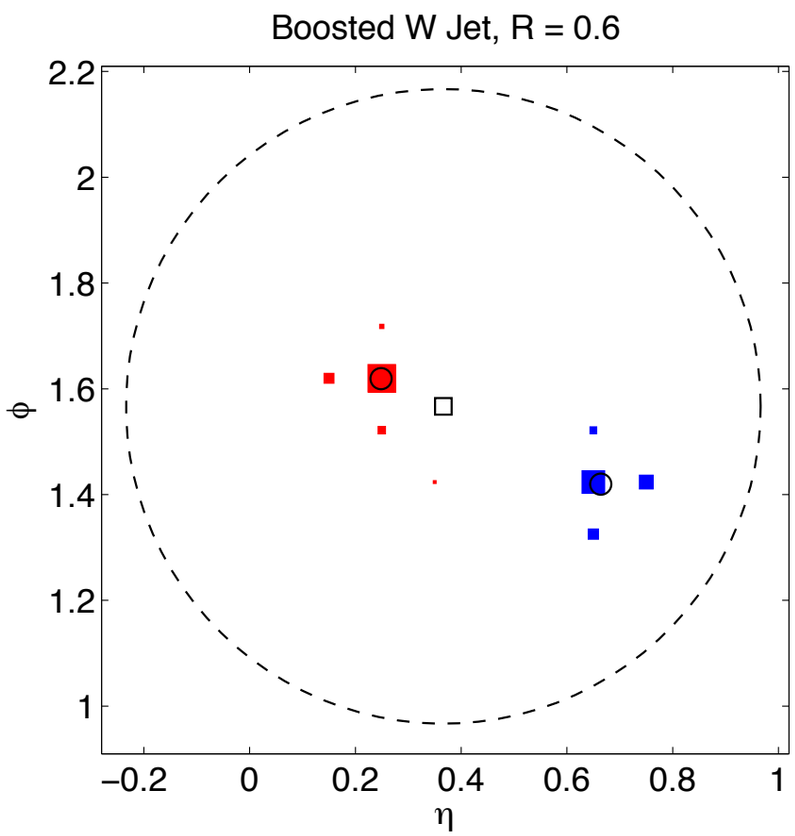
3 classic grooming cases…

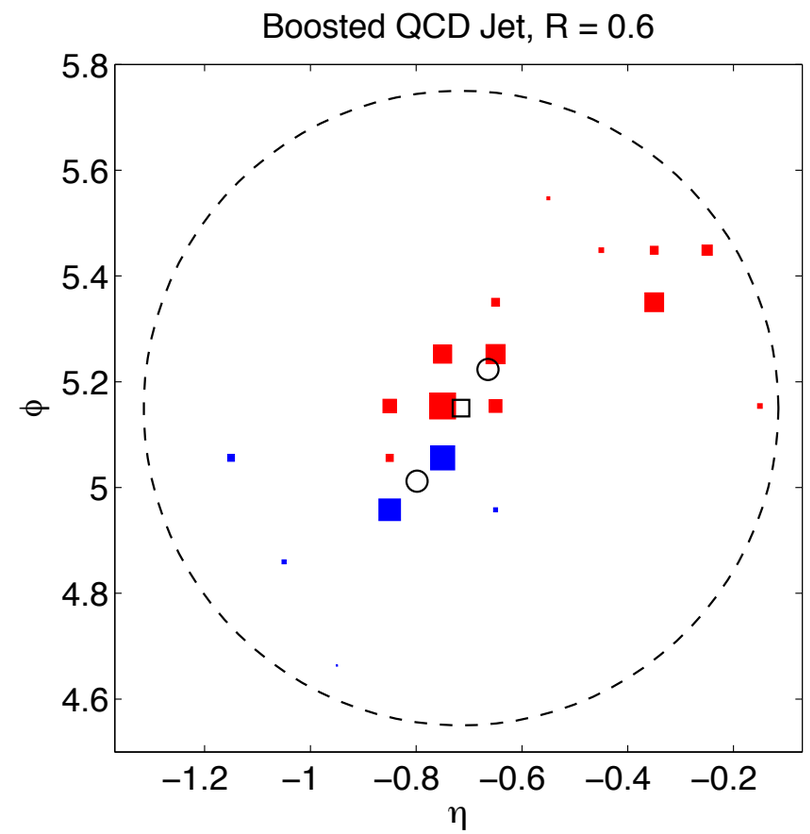


tagging W and Z jets

separating a W or Z jet from a high p_T quark or gluon jet

1. jet mass
2. W and Z's have 2 prong structure





example: N-subjettiness

J. Thaler, K. Van Tilburg, arXiv:1011.2268

generalizing subjets…

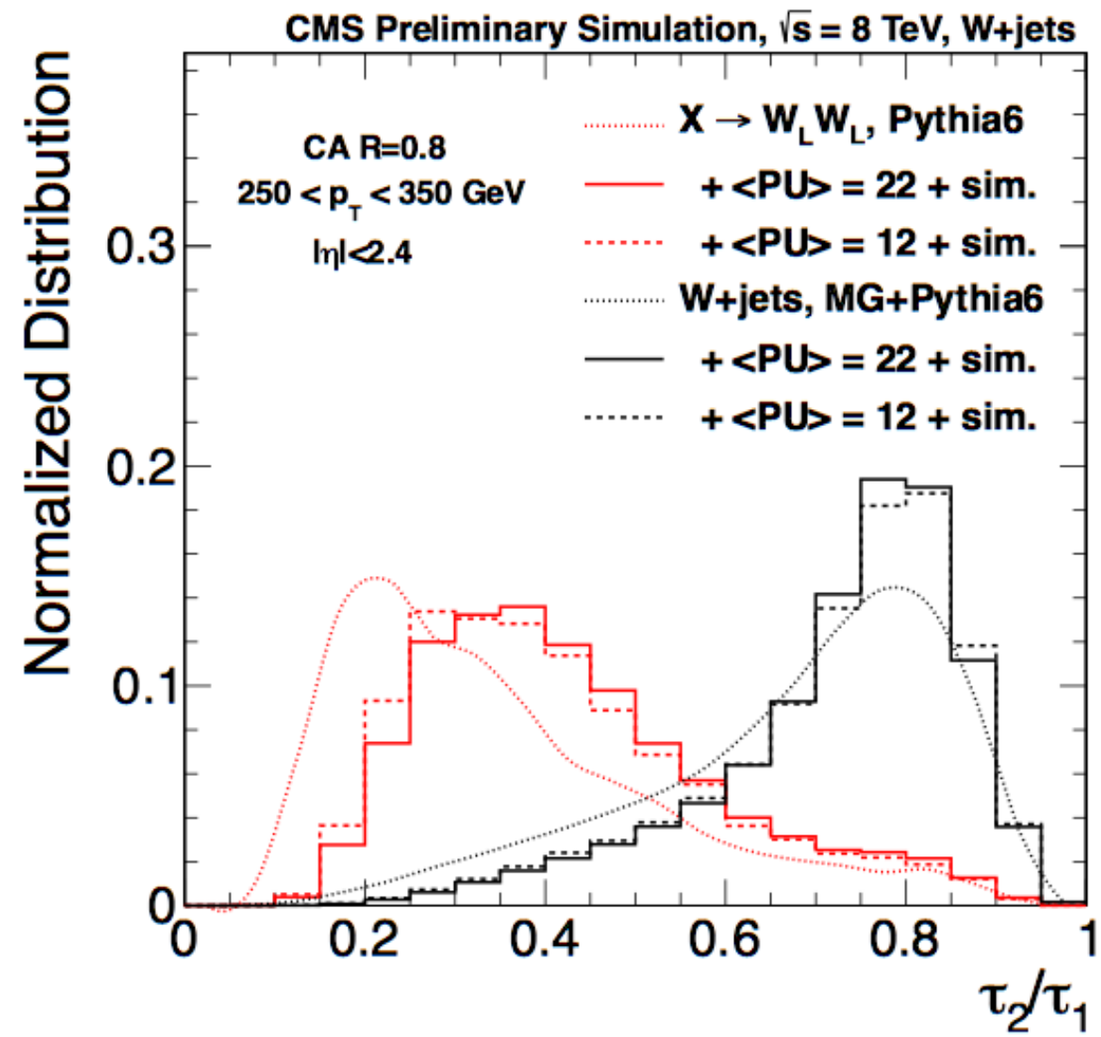
N-subjettiness: a measure of how consistent a jet is with having N subjets, τ_N

$$\tau_N = \frac{1}{d_0} \sum_k p_{T,k} \min\{\Delta R_{1,k}, \Delta R_{2,k}, \cdots, \Delta R_{N,k}\}$$

As $\tau_N \to 0$, jet is more consistent with having N subjets

e.g. $\tau 2 \to 0$, more like a W jet
e.g. $\tau 1 \to 0$, more like a quark jet

Ratios are typically used:
τ_2/τ_1 for separating **W jets** from **quark and gluon jets**



high mass resonances

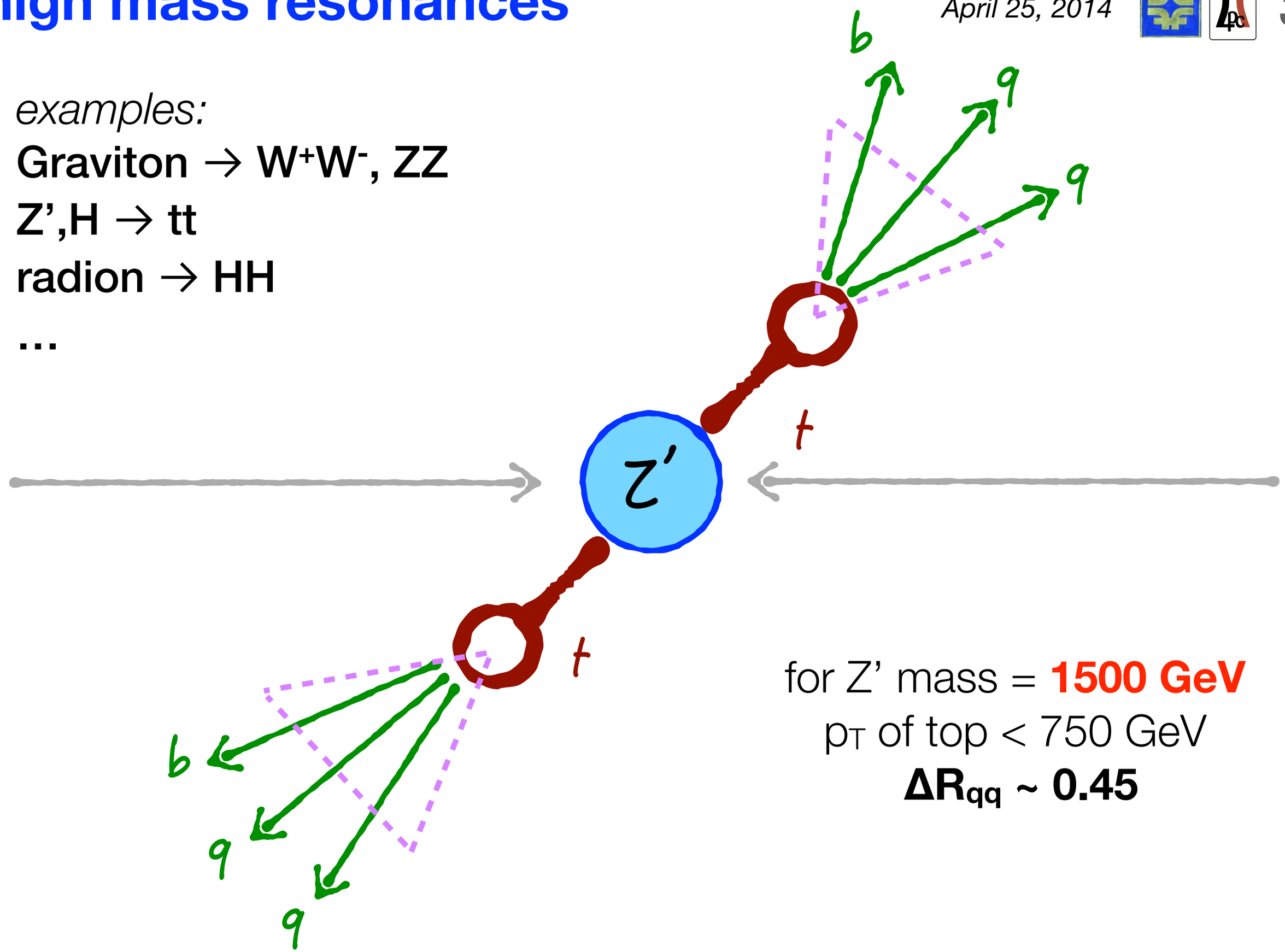
examples:

Graviton → W^+W^-, ZZ

Z',H → tt

radion → HH

...



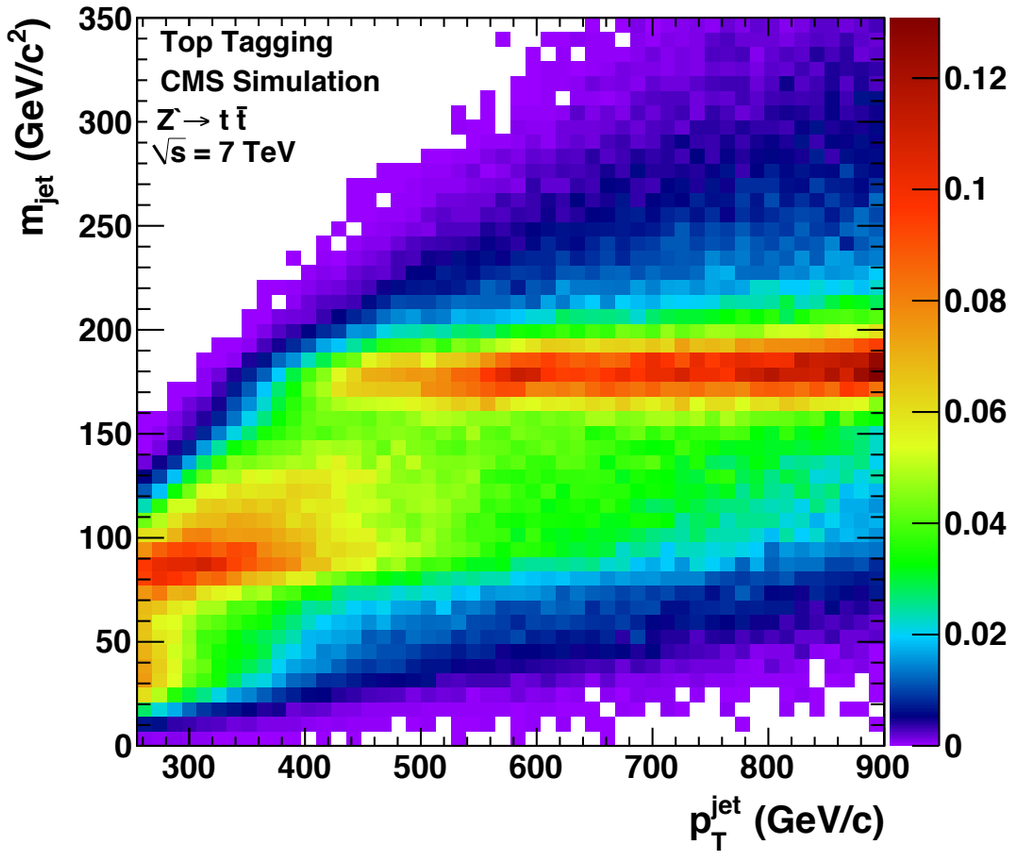
for Z' mass = **1500 GeV**

p_T of top < 750 GeV

ΔR_{qq} ~ 0.45

boosted tops

Tops can be more complicated objects, tagging is not only via ***observables***, but ***algorithms*** too



Top taggers

Examples of single observables

kT splitting scales

N-subjettiness τ_3/τ_2

Decluster the fat jet with particular requirements and look at invariant mass and other kinematic requirements

CMS top tagger (a variant of JHU top tagger [1])

HEP top tagger [2] - larger cones (R=1.5) targeted for moderate p_T

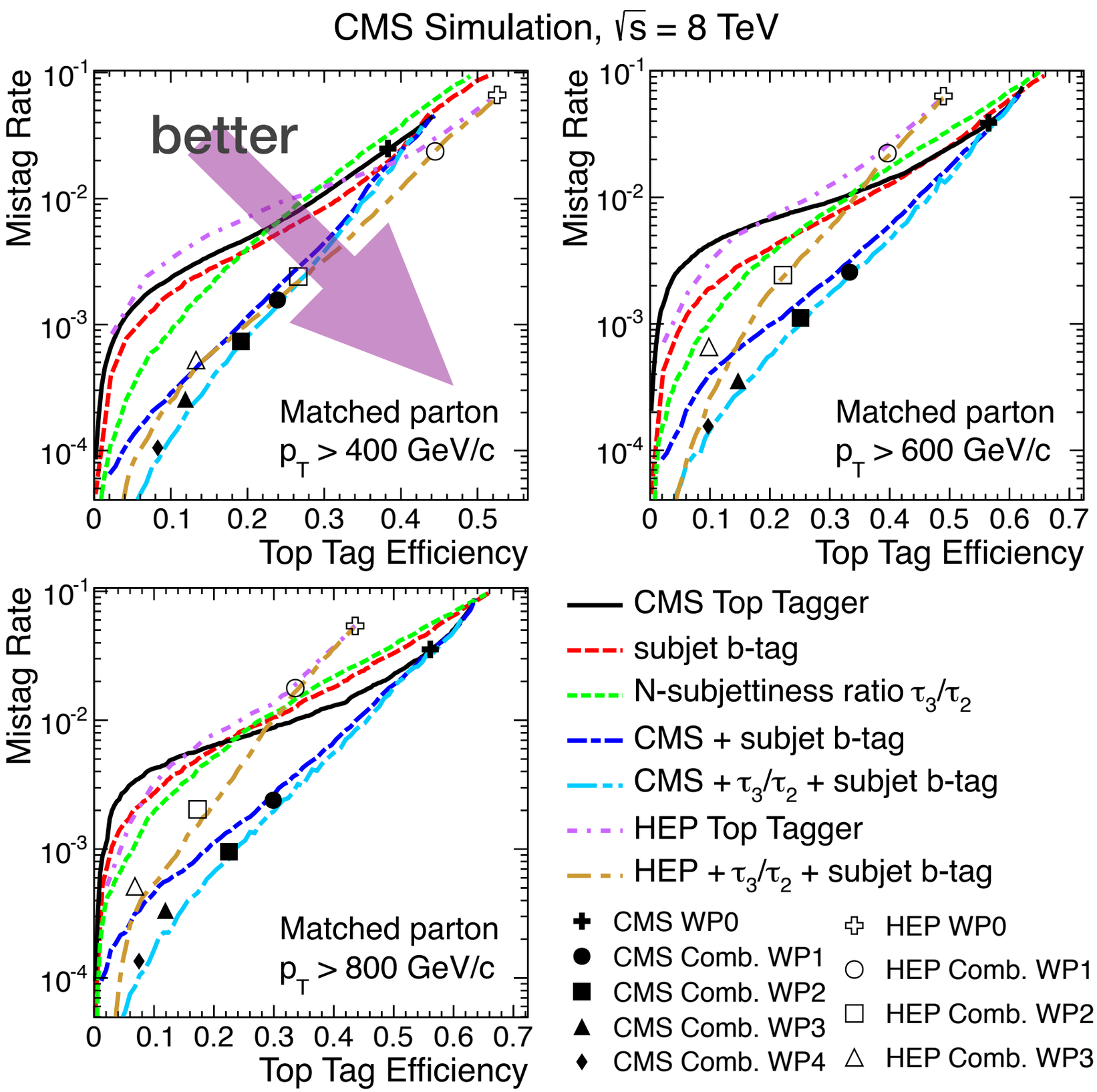
Template top tagger [3] - compares jet energy profile to O(300k) templates

Shower deconstruction [4] - matrix element for a single jet

Kaplan et al., PRL 101/142001 (2008)
Plehn, Spannowsky et. al, JHEP 1010:078,2010
Perez et al., Phys. Rev. D82 054034, 2010
Soper, Spannoswsky, PhysRevD.87.054012

top performance

CMS PAS JME-13-007
CMS-PAS BTV-13-001



Methods for b-tagging each subjet in the jet have been developed

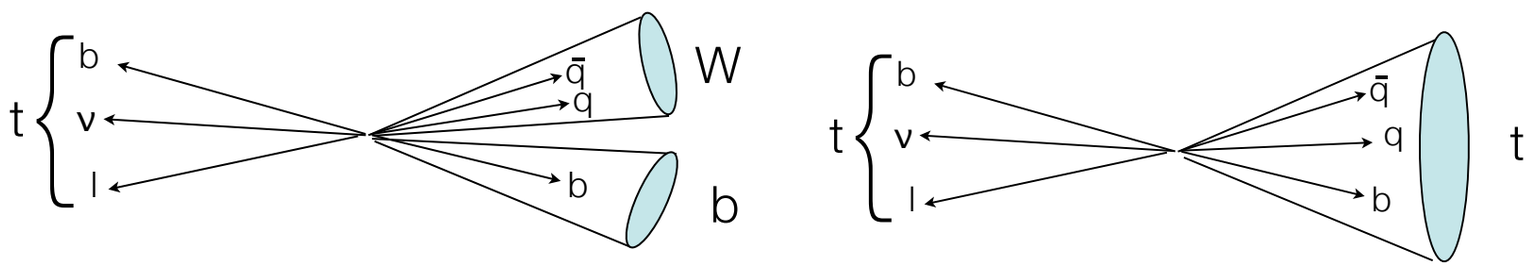
Great for tops and Higgs!

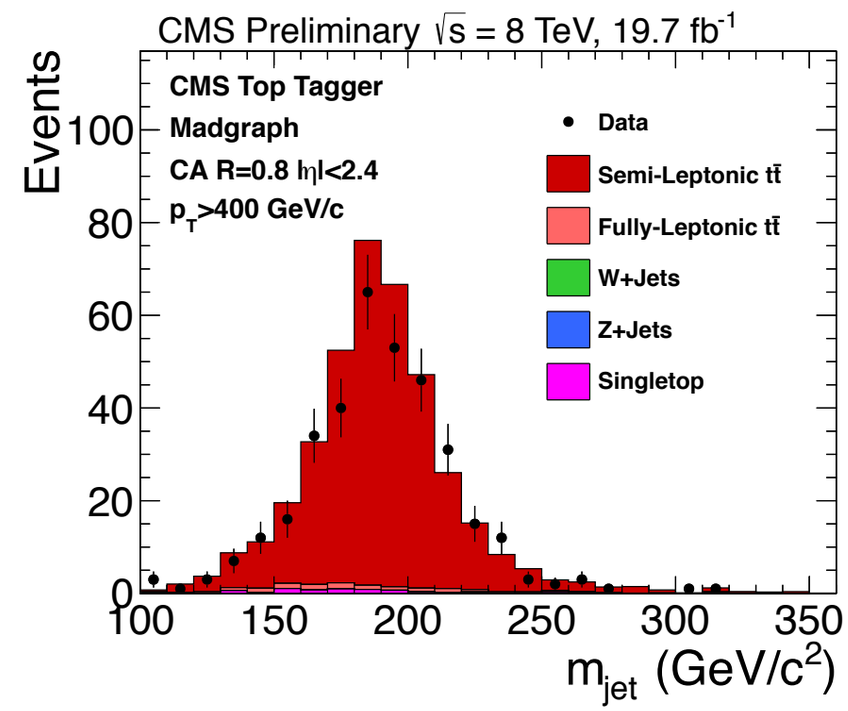
(more info in backup)

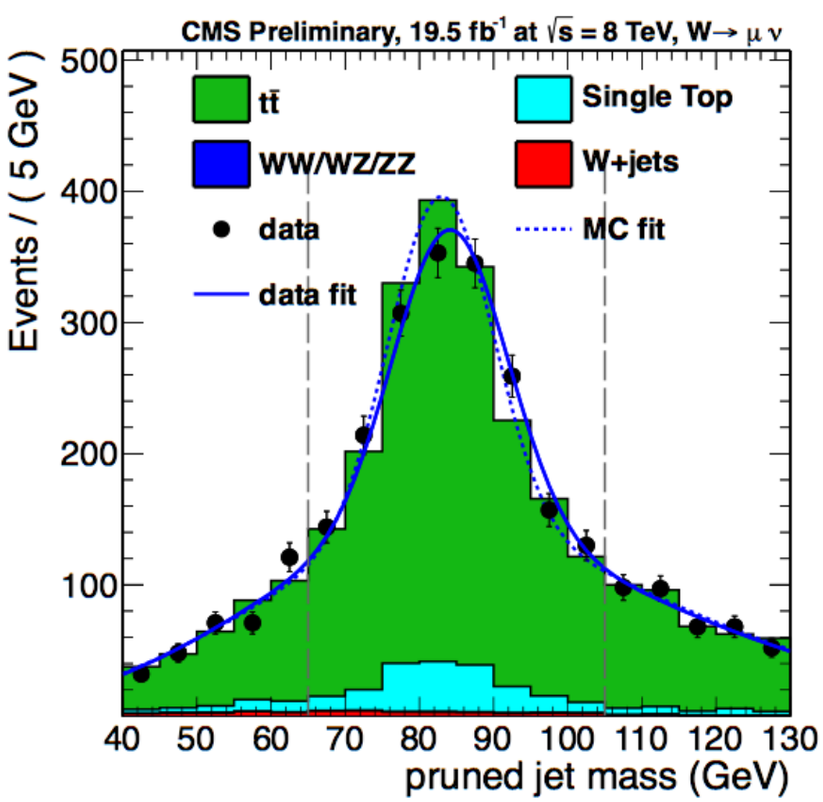
standard candles

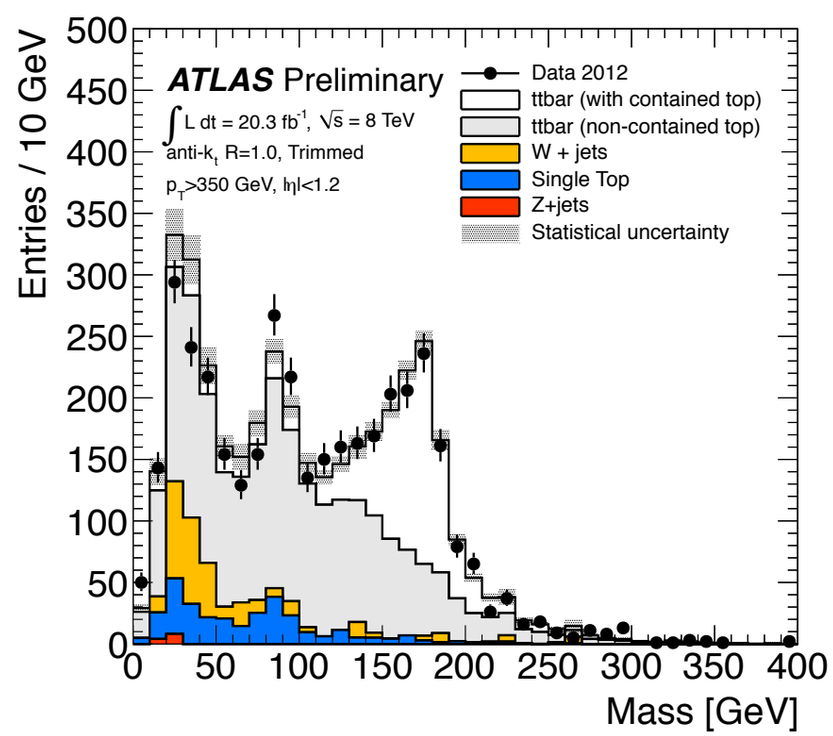
CMS PAS JME-13-006

Semi-leptonic $t\bar{t}$ events are important for validating tagging techniques of heavy objects







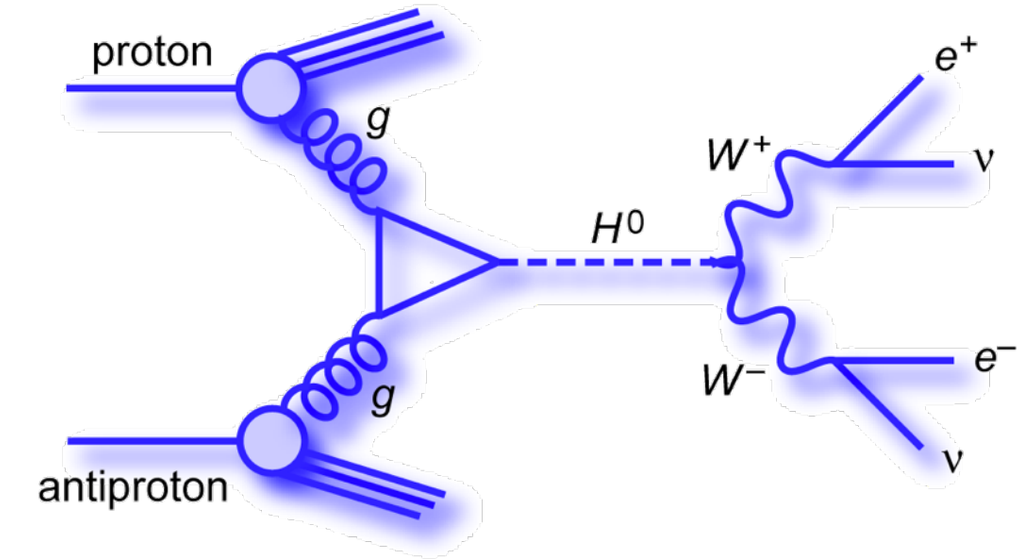


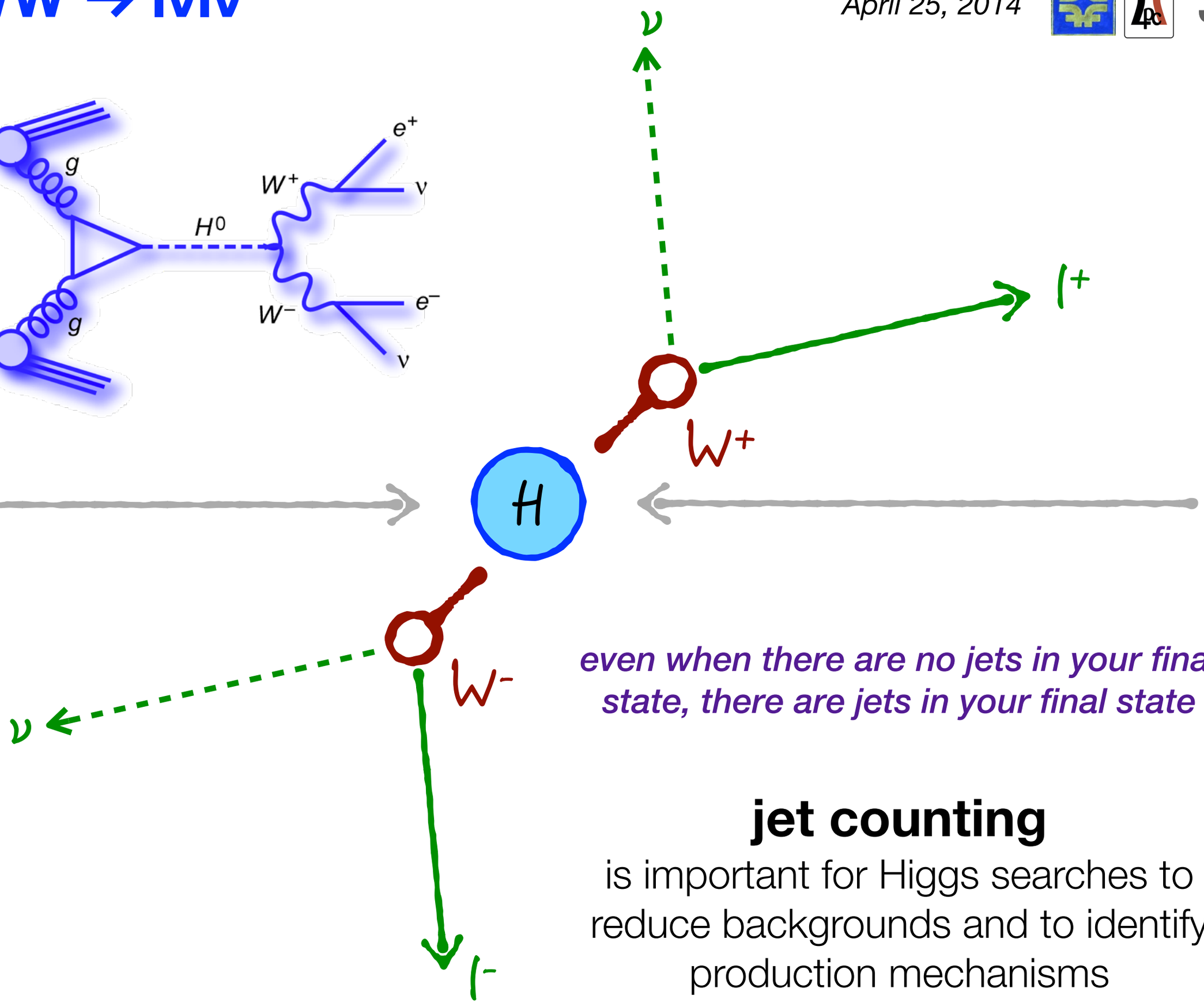
Using events to do evaluate tag-and-probe efficiency **scale factors** and **mass scale/resolution** measurements

case #2:

$gg \rightarrow H \rightarrow W^+W^- \rightarrow l\nu l\nu$

H → WW → lνlν





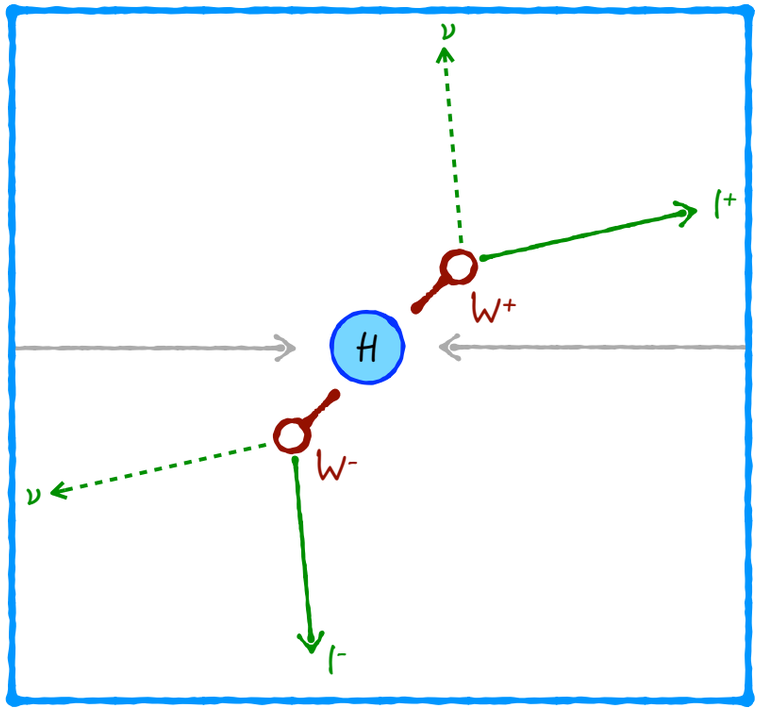
even when there are no jets in your final state, there are jets in your final state

jet counting

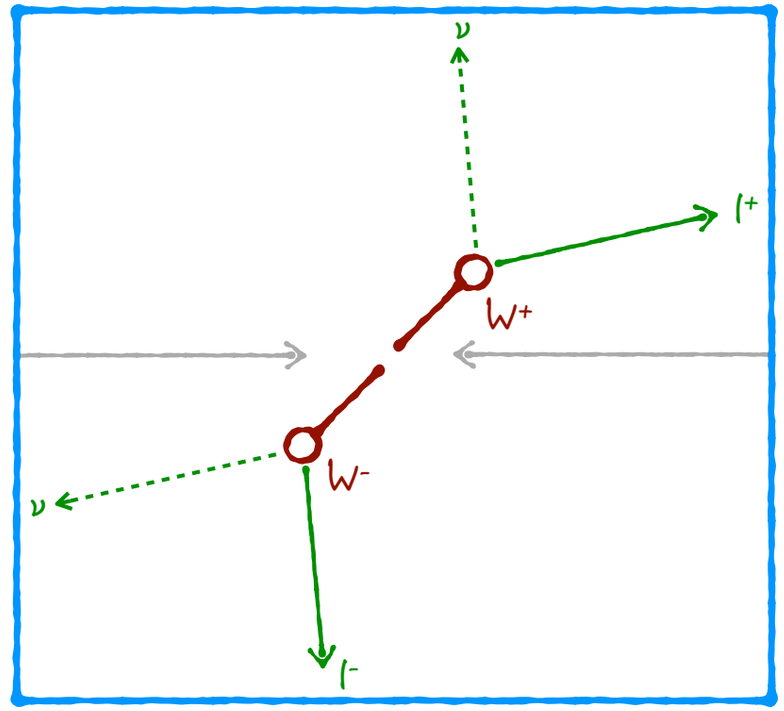
is important for Higgs searches to reduce backgrounds and to identify production mechanisms

jet counting

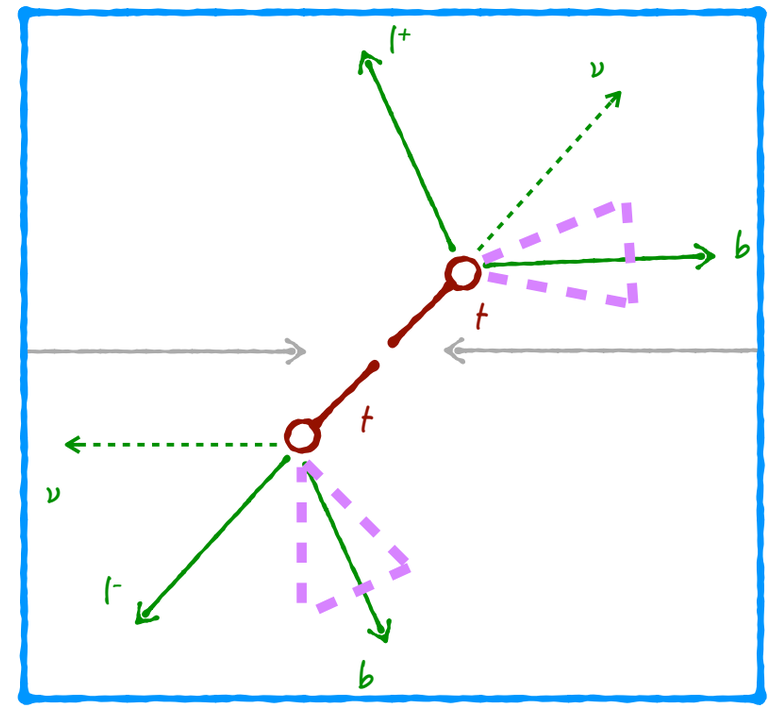
signal



bkg: WW



bkg: ttbar



by categorizing the event as having 0 or 1 or more additional jets, the ttbar background can be greatly reduced

the additional jet from ttbar can have low p_T

jet counting at low p_T

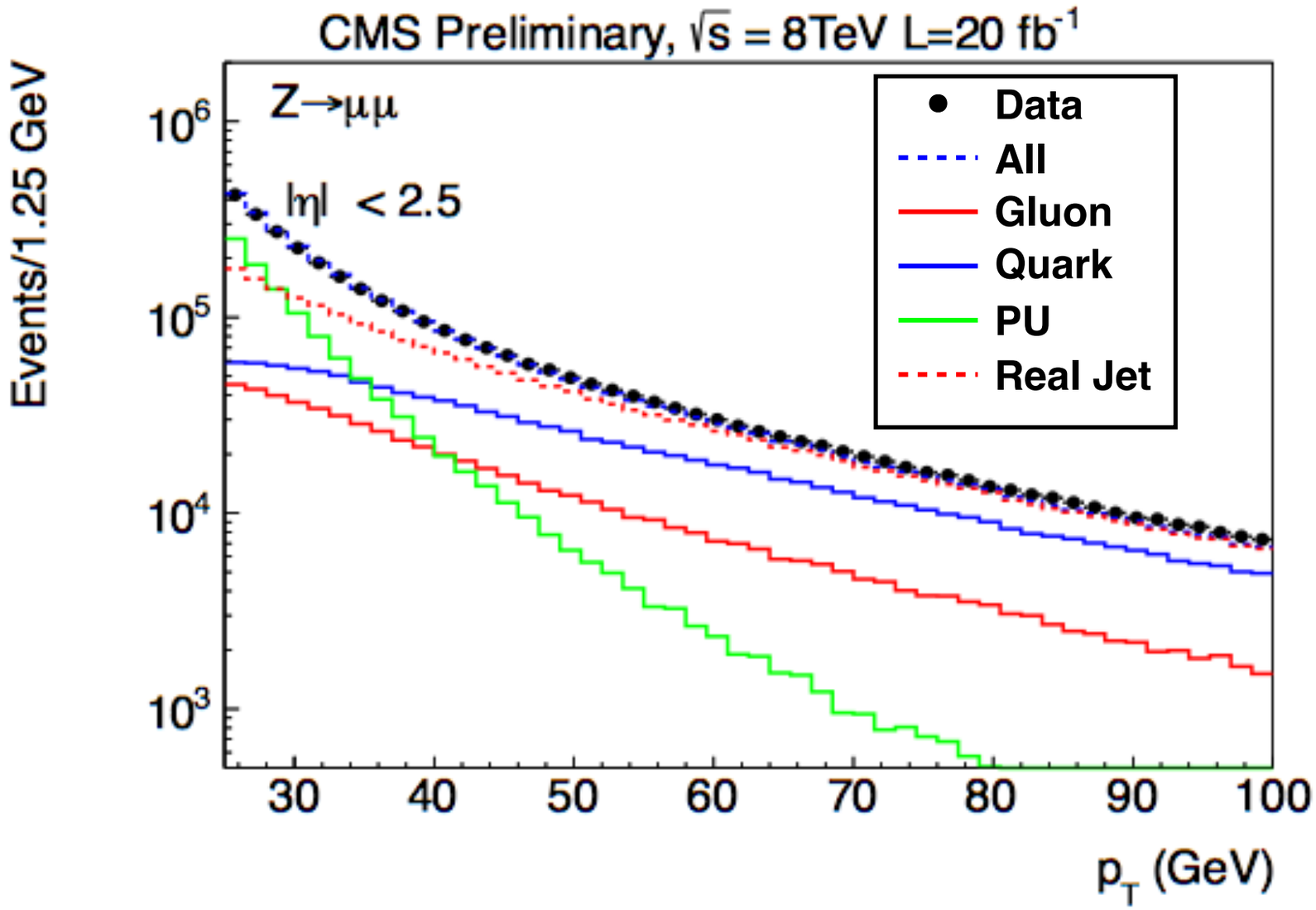
The low p_T regime is dominated by **quark**, **gluon**, and **pileup jets**.

jet counting at low p_T

The low p_T regime is dominated by **quark**, **gluon**, and **pileup jets**.

jet counting at low p_T

The low p_T regime is dominated by **quark**, **gluon**, and **pileup jets**.

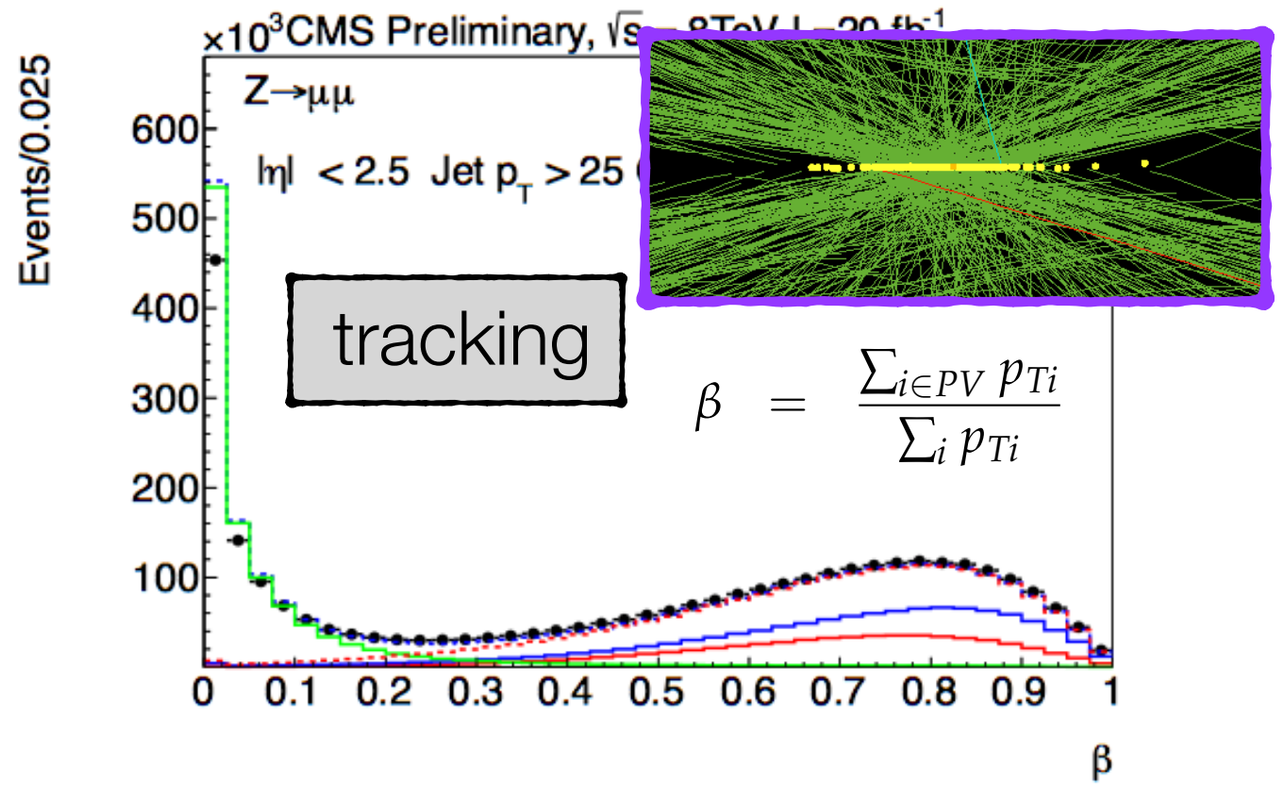


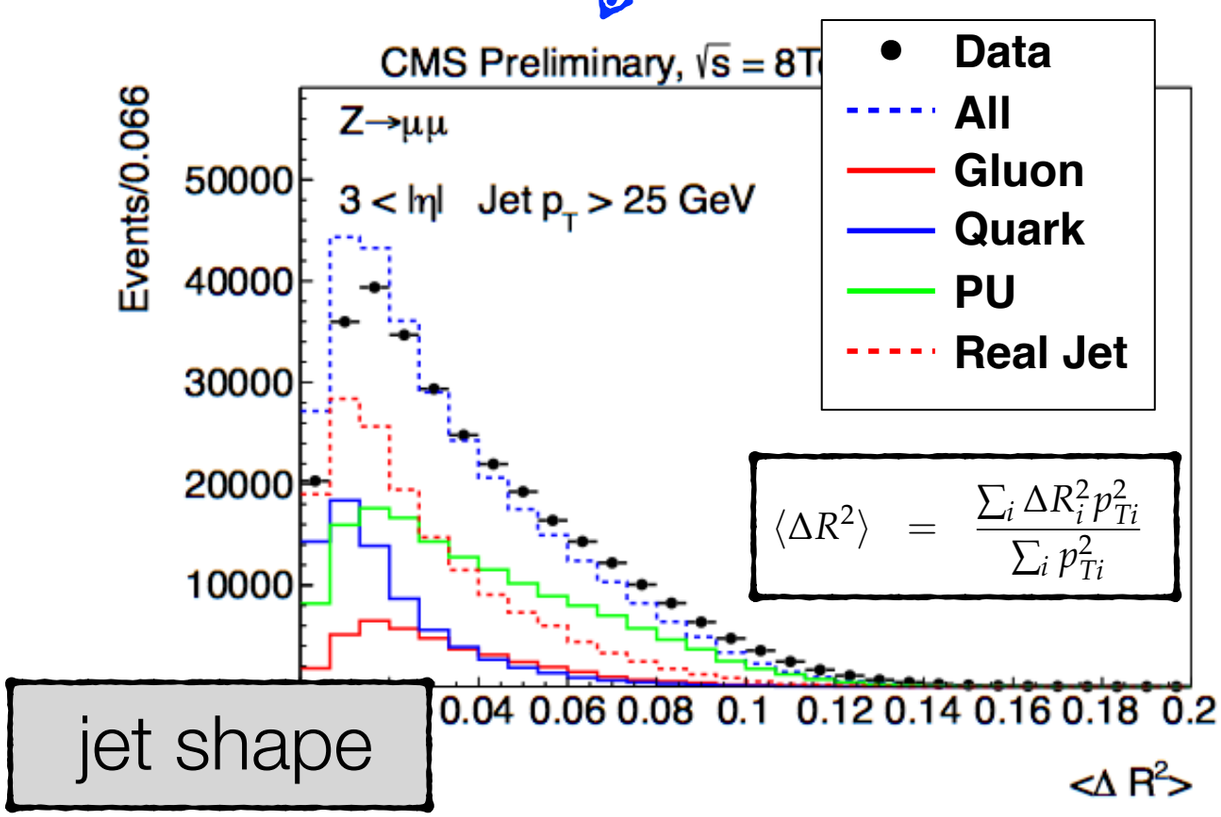
Using all information, can do a better removing them:
track-based observables and jet shapes

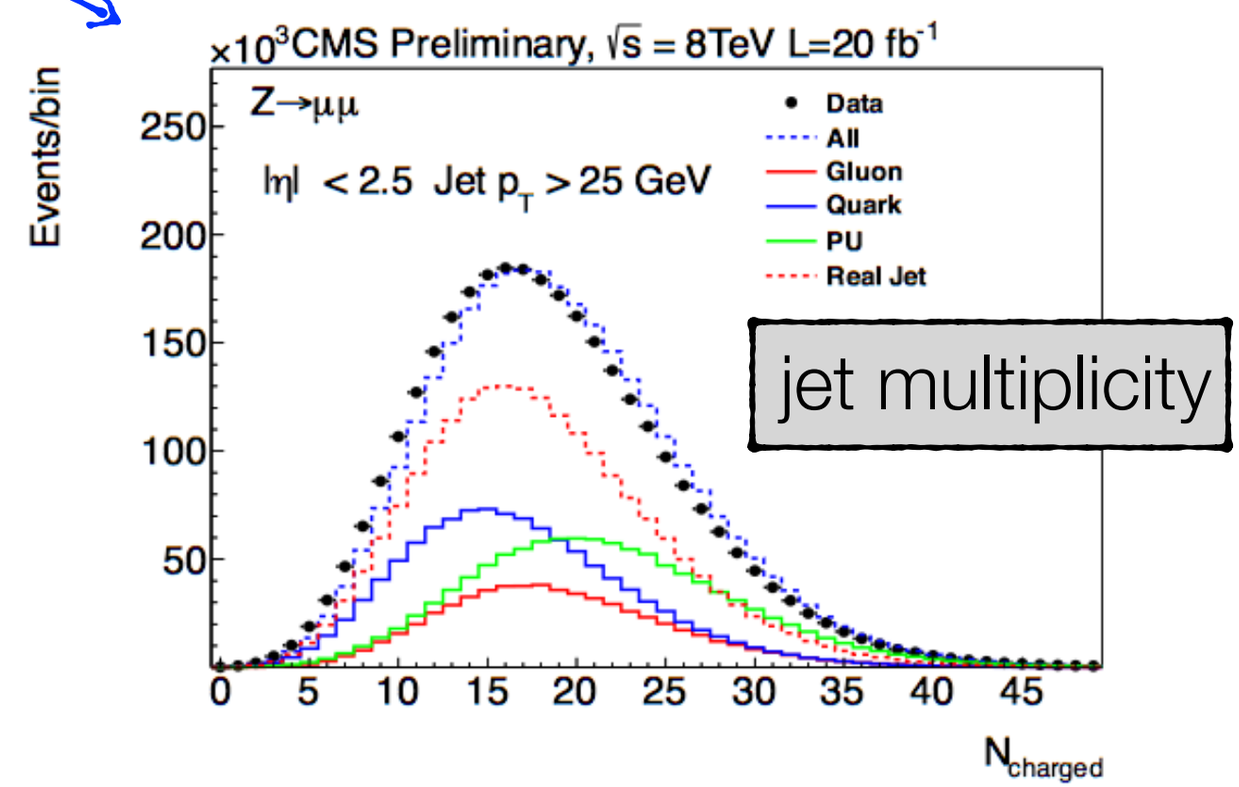
jet counting at low p_T

combine all information to build discriminants for quark, gluon and pileup jets

useful even without tracker!

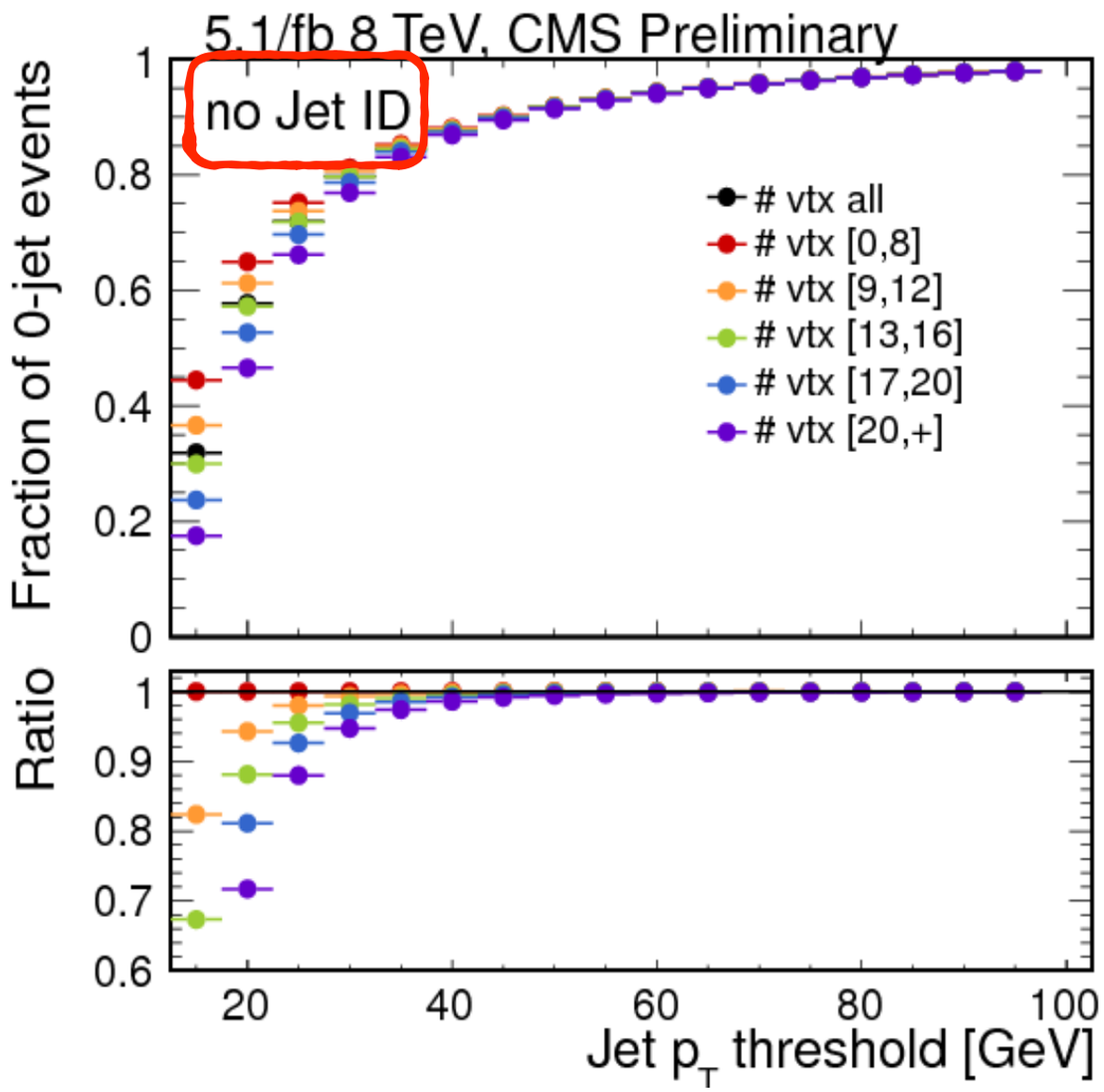


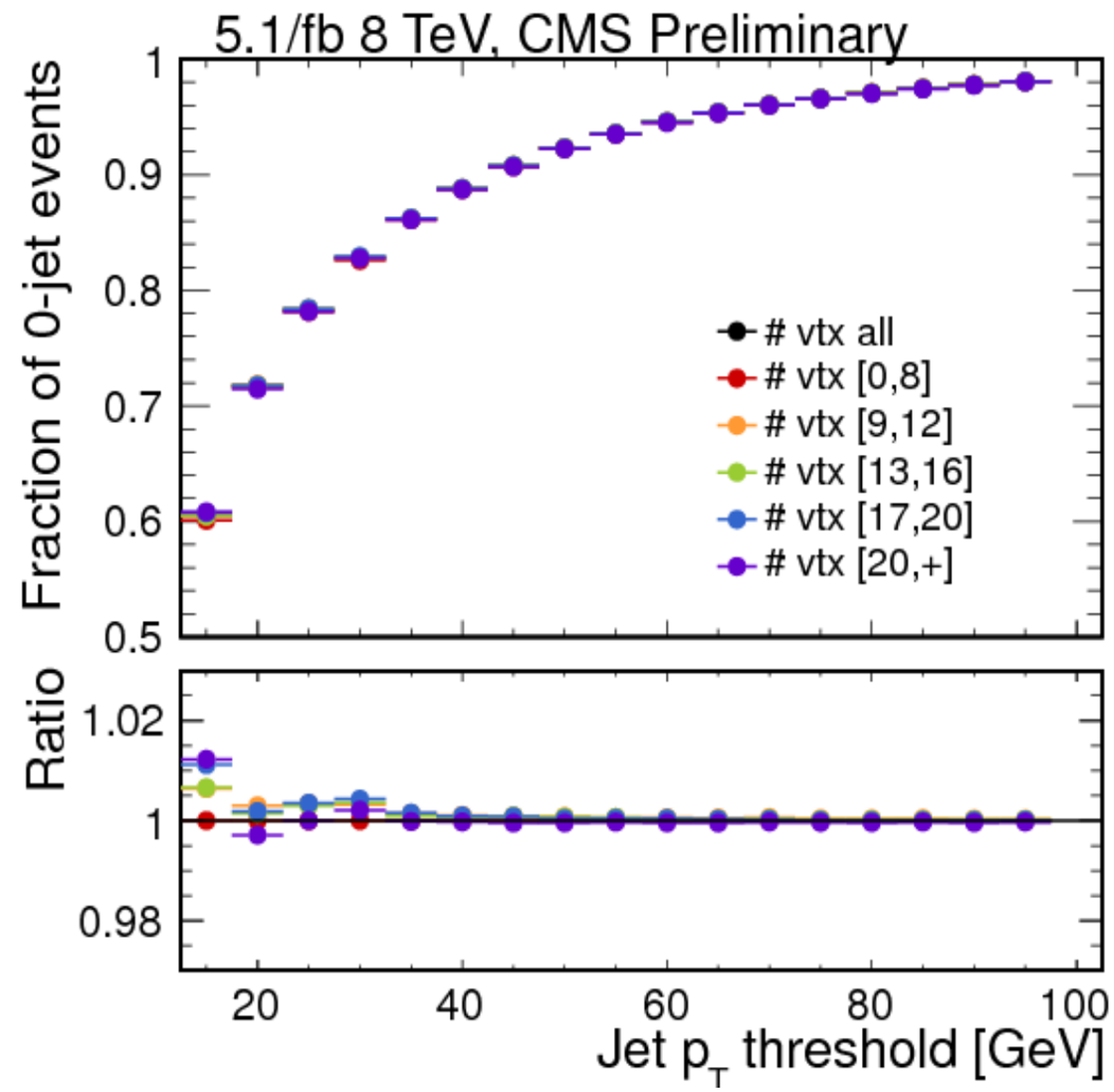




jet counting

jet counting stabilizes with pileup jet ID applied



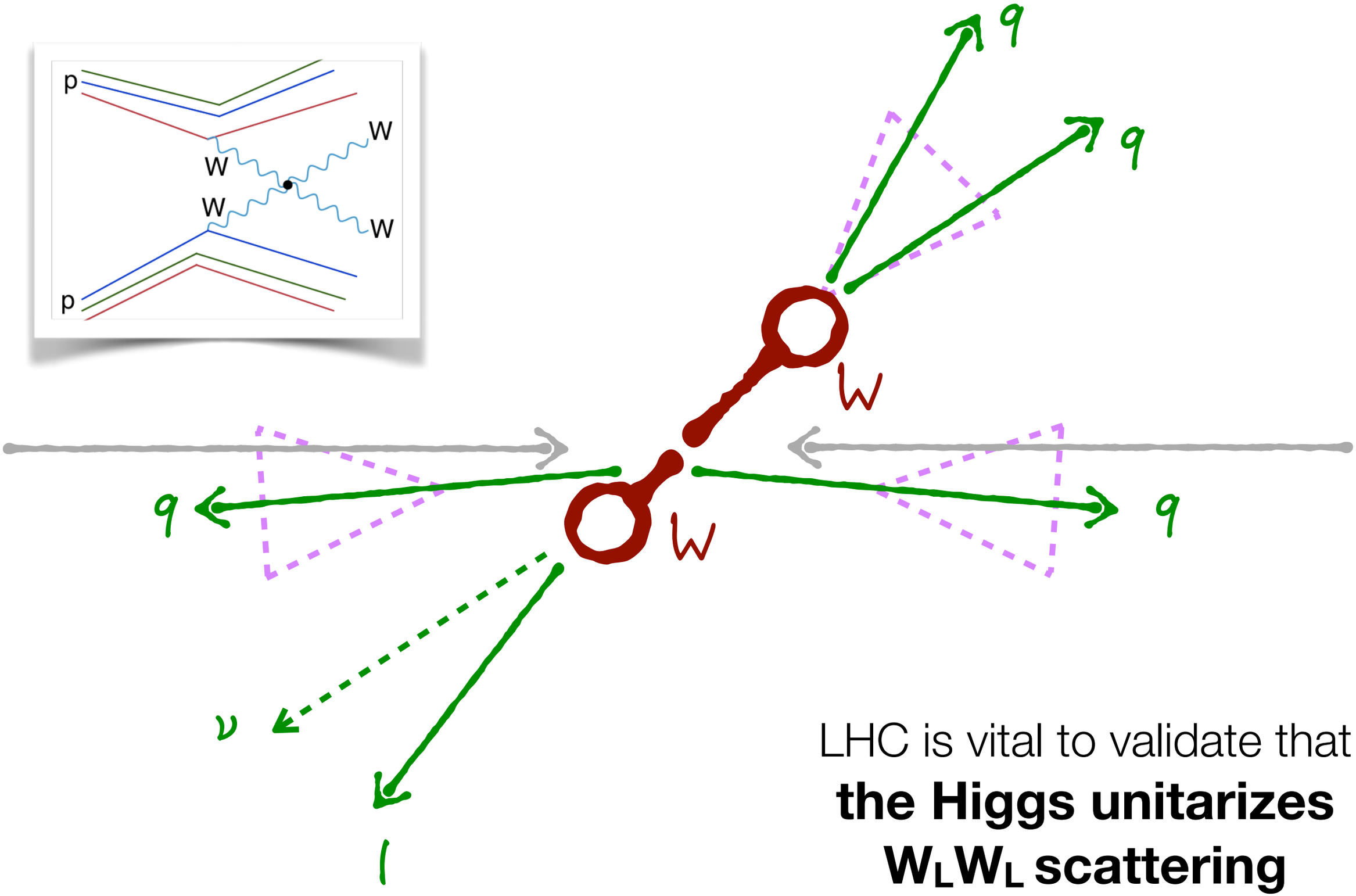


even when there are no jets in your final state, there are jets in your final state

case #3:

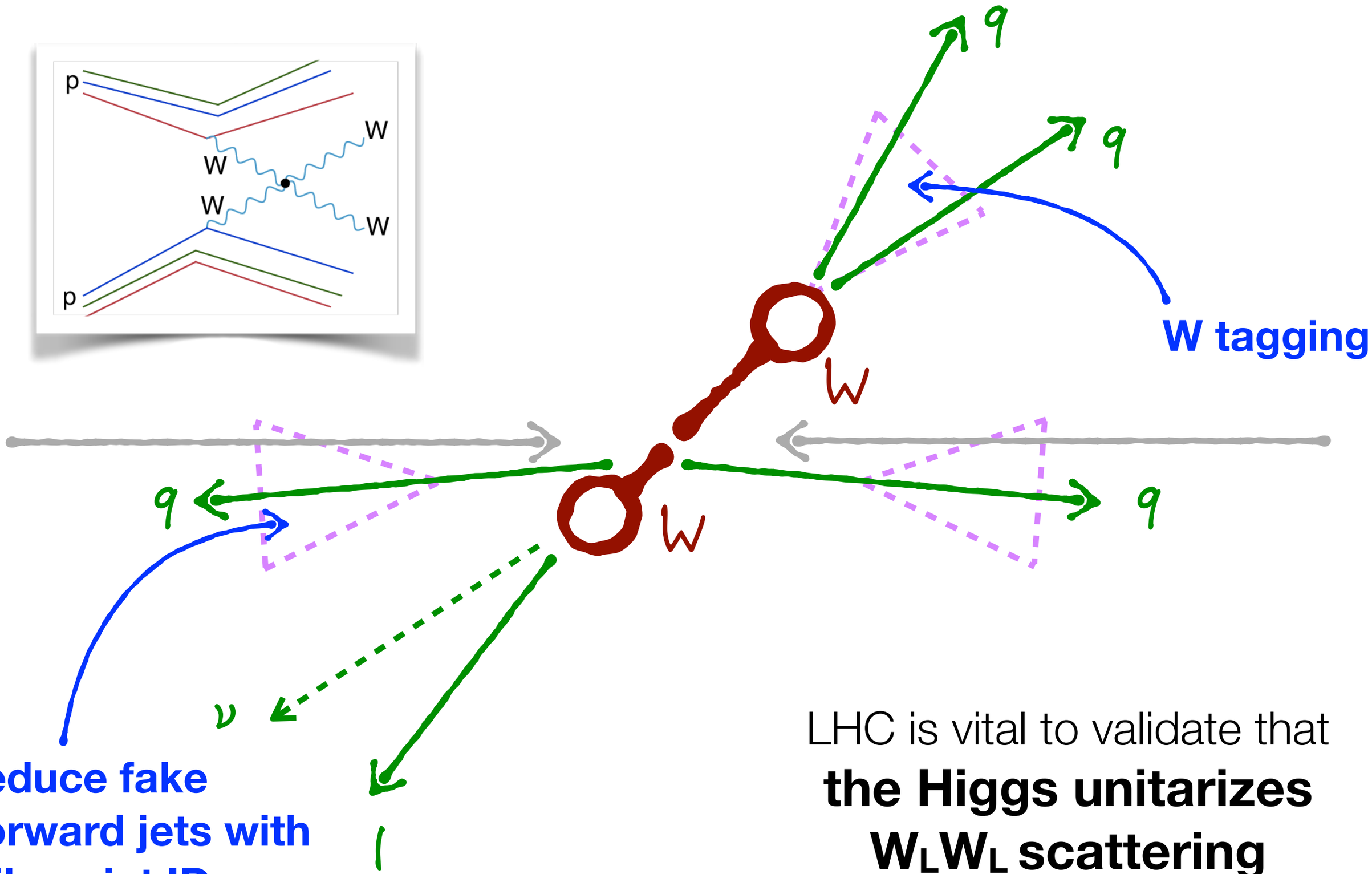
WW scattering

H → WW → lνJjj



LHC is vital to validate that
the Higgs unitarizes
$W_L W_L$ scattering
at high mass scales

H → WW → lνJjj



reduce fake forward jets with pileup jet ID

rate reduced by a factor of 2 per jet

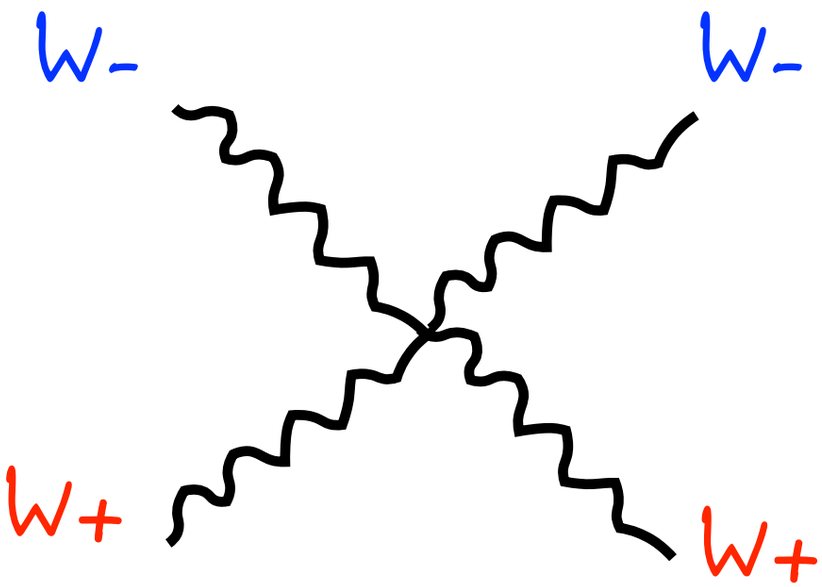
LHC is vital to validate that

the Higgs unitarizes $W_L W_L$ scattering

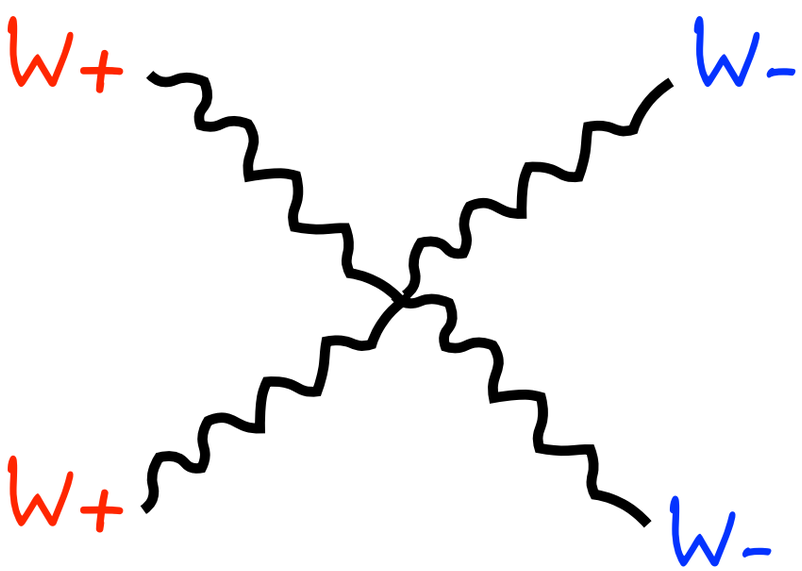
at high mass scales

jet charge

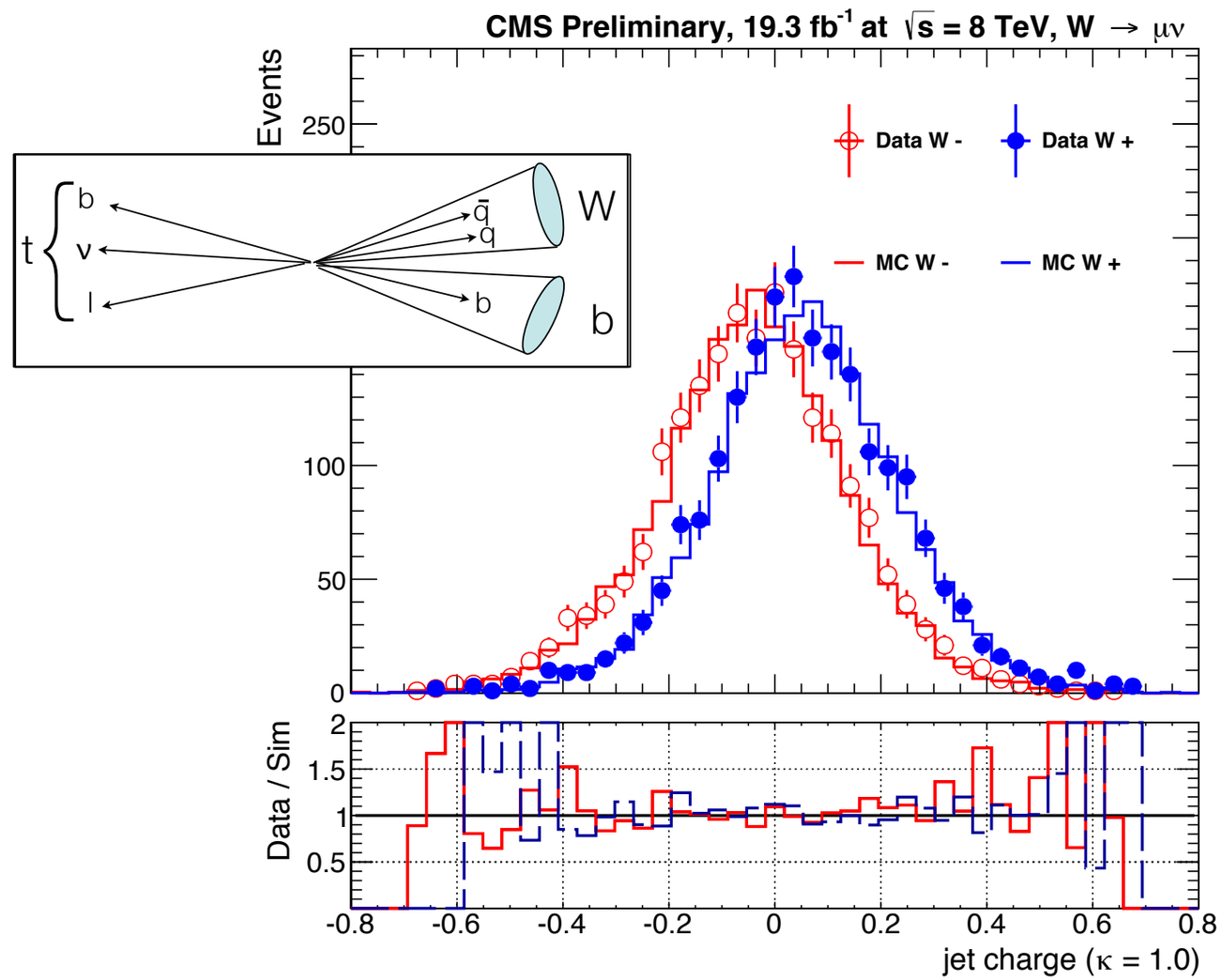
opposite sign



same sign



Jet charge measured for first time at LHC
discrimination for **W+** vs. **W-**
also for **up** and **down** quark (in backup)
Interesting possible applications!



THROUGH
THE LOOKING
GLASS
LEWIS CARROLL

We are breaking down jets to their smallest pieces

Re-evaluate what is possible…

the *future*
Matrix Element Method

removing pileup one particle at a time

snowmass white paper...

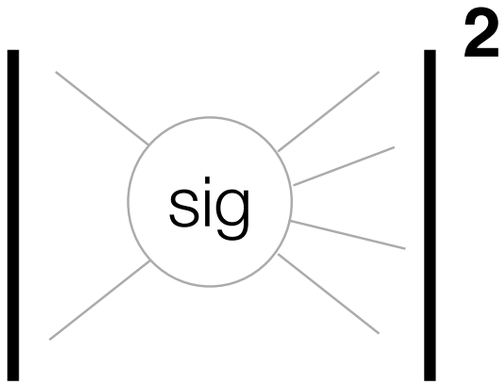
The Matrix Element Method: Past, Present, and Future

James S. Gainer,[1] Joseph Lykken,[2] Konstantin T. Matchev,[1] Stephen Mrenna,[3] and Myeonghun Park[4]

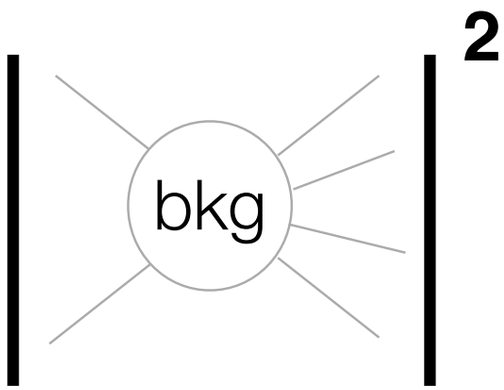
[1] *Physics Department, University of Florida, Gainesville, FL 32611, USA*
[2] *Theoretical Physics Department, Fermilab, Batavia, IL 60510, USA*
[3] *SSE Group, Computing Division, Fermilab, Batavia, IL 60510, USA*
[4] *Theory Division, Physics Department, CERN, CH-1211 Geneva 23, Switzerland*

(Dated: July 26, 2013)

P(sig) ~



P(bkg) ~



the present. We therefore speculate about future developments involving the MEM, without regard to (current) computational limitations.

- **NLO/ Parton Showers**– The extension of the MEM to take into account additional radiation and/or other NLO corrections has already been considered [7, 15, 20]. Work has also gone into extending the MEM to include parton showers [21]. Such approaches, potentially extended beyond NLO, will leverage the more complete higher order calculations of the future.
- **Jet Substructure**– As outlined in Eq. 1, the MEM is only using the four momentum of a jet. However, other properties of the jet, e.g. substructure [22] could give additional information about the hard process parton with which a jet should be associated, thereby increasing the sensitivity of the analysis.
- **Underlying Event, Hadronization, Etc.**– Additional information about hadronization, the underlying event, and other topics could also be used to calculate a more complete likelihood for events in hadron colliders.
- **Detector Resolution**– Ultimately, instead of using a single transfer function for a detector, a separate transfer function could be utilized for each detector element involved in the reconstruction of an event. Ideally the time dependence of the element response, as well as its correlations with other detector elements would be included.

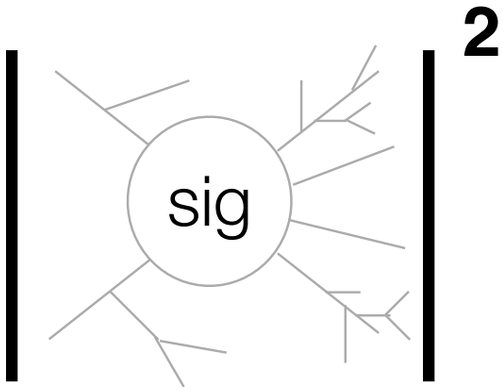
matrix element method

snowmass white paper...

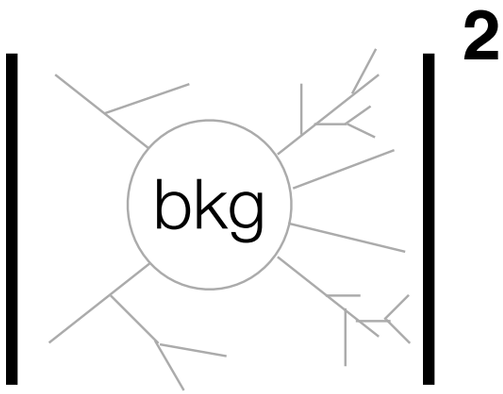
The Matrix Element Method: Past, Present, and Future

James S. Gainer,[1] Joseph Lykken,[2] Konstantin T. Matchev,[1] Stephen Mrenna,[3] and Myeonghun Park[4]
[1] *Physics Department, University of Florida, Gainesville, FL 32611, USA*
[2] *Theoretical Physics Department, Fermilab, Batavia, IL 60510, USA*
[3] *SSE Group, Computing Division, Fermilab, Batavia, IL 60510, USA*
[4] *Theory Division, Physics Department, CERN, CH-1211 Geneva 23, Switzerland*
(Dated: July 26, 2013)

P(sig) ~



P(bkg) ~



extend the MEM to the particle level

the present. We therefore speculate about future developments involving the MEM, without regard to (current) computational limitations.

- **NLO/ Parton Showers**– The extension of the MEM to take into account additional radiation and/or other NLO corrections has already been considered [7, 15, 20]. Work has also gone into extending the MEM to include parton showers [21]. Such approaches, potentially extended beyond NLO, will leverage the more complete higher order calculations of the future.
- **Jet Substructure**– As outlined in Eq. 1, the MEM is only using the four momentum of a jet. However, other properties of the jet, e.g. substructure [22] could give additional information about the hard process parton with which a jet should be associated, thereby increasing the sensitivity of the analysis.
- **Underlying Event, Hadronization, Etc.**– Additional information about hadronization, the underlying event, and other topics could also be used to calculate a more complete likelihood for events in hadron colliders.
- **Detector Resolution**– Ultimately, instead of using a single transfer function for a detector, a separate transfer function could be utilized for each detector element involved in the reconstruction of an event. Ideally the time dependence of the element response, as well as its correlations with other detector elements would be included.

shower deconstruction

"matrix element for a single jet"

top tagging application

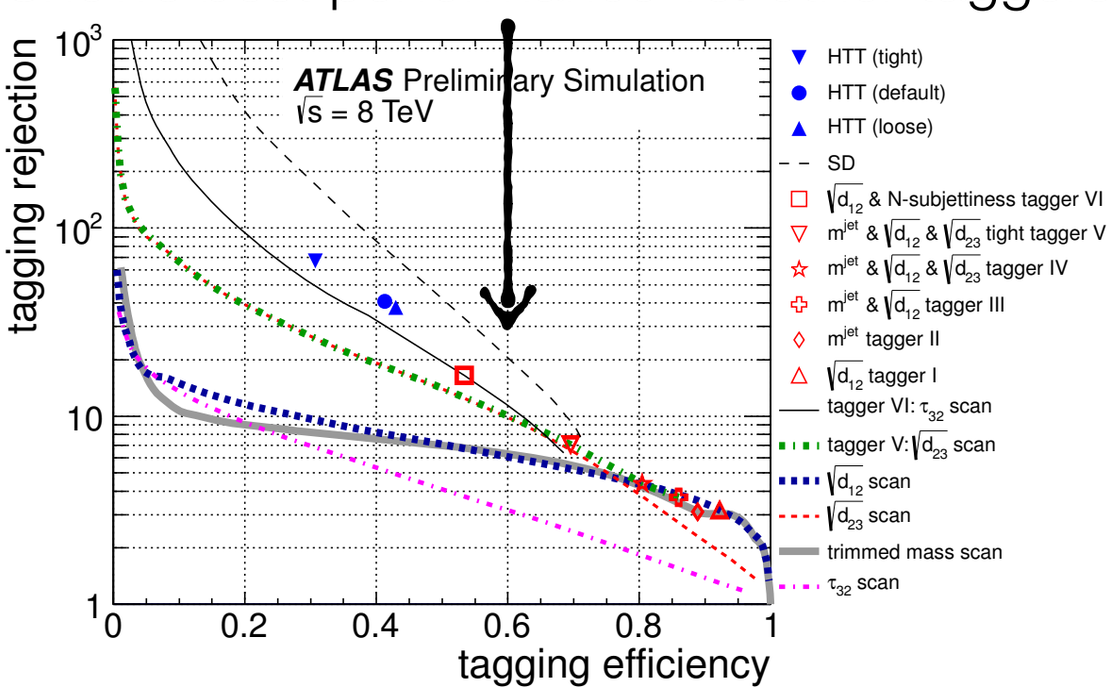
Break jet into microjets

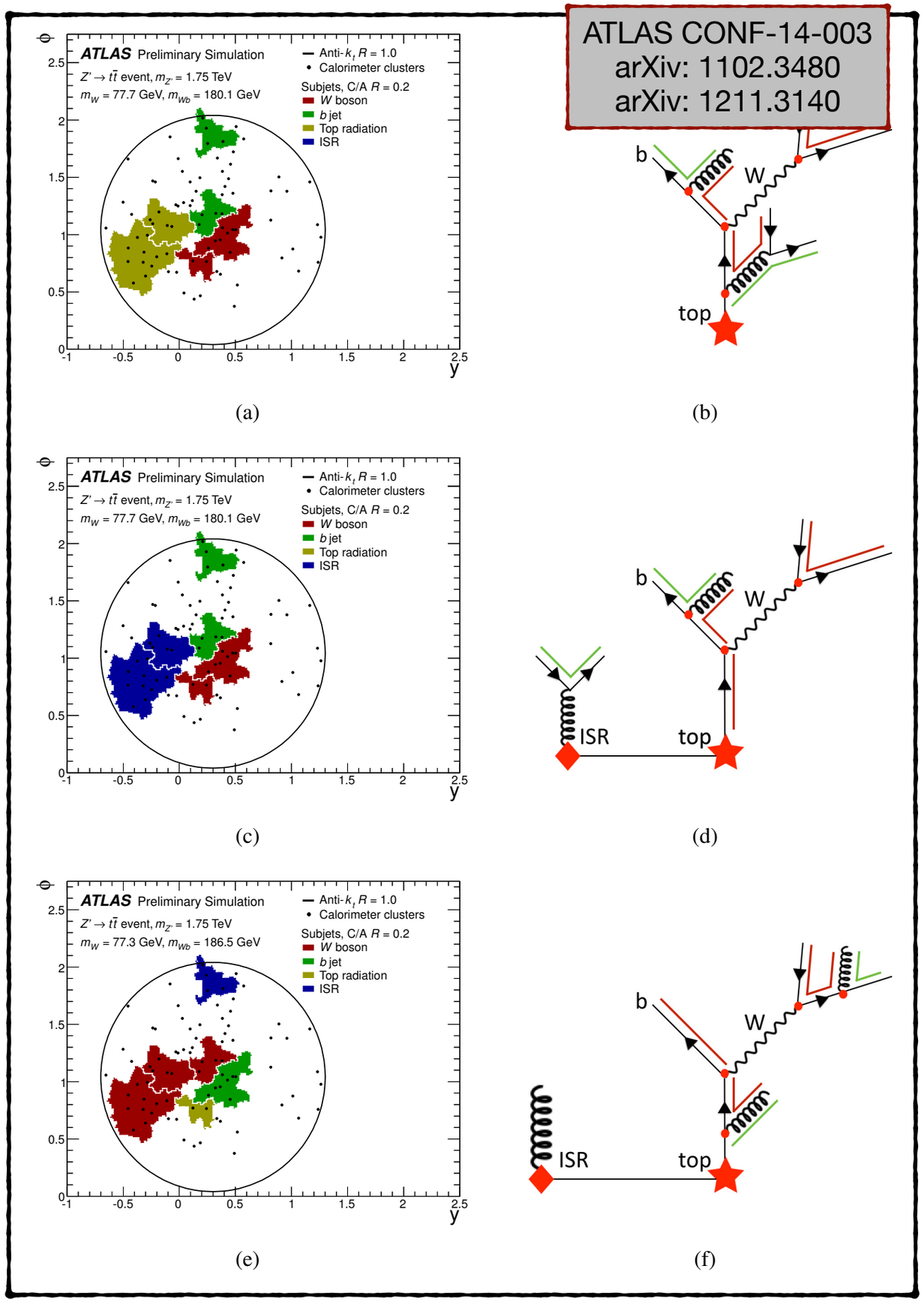
P(sig) = Σ | sig(N microjets) |^2

P(bkg) = Σ | bkg(N microjets) |^2

sum over 1000's of parton shower histories

Shows best performance vs. other taggers



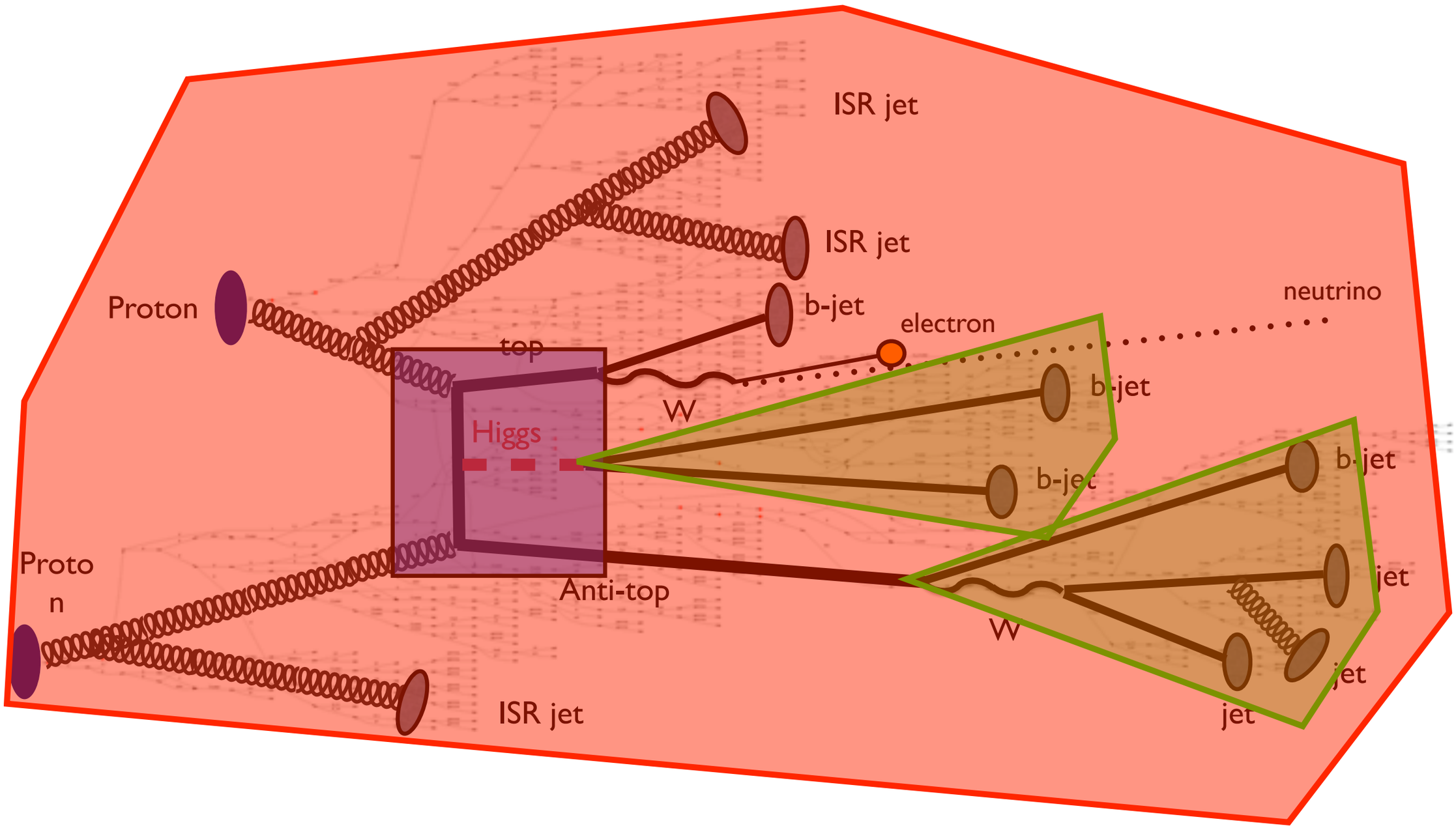


event deconstruction

arXiv: 1402.1189

D. Soper, M. Spannowsky

shower deconstruction -> event deconstruction in 2013



pileup

ONLY ALAN WAS PREPARED TO ACKNOWLEDGE THE ELEPHANT IN THE ROOM.
harrop

current pileup mitigation

arXiv:1111.6097
arXiv:1309.4777

average subtraction: subtract median value of pileup from the jet

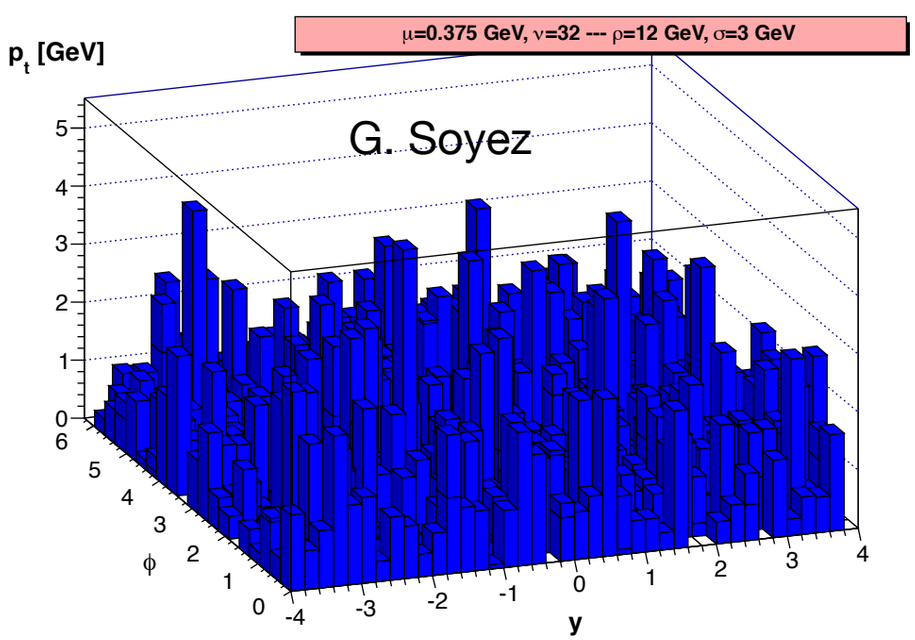
problem: pileup is fluctuating/lumpy, smears out jet properties

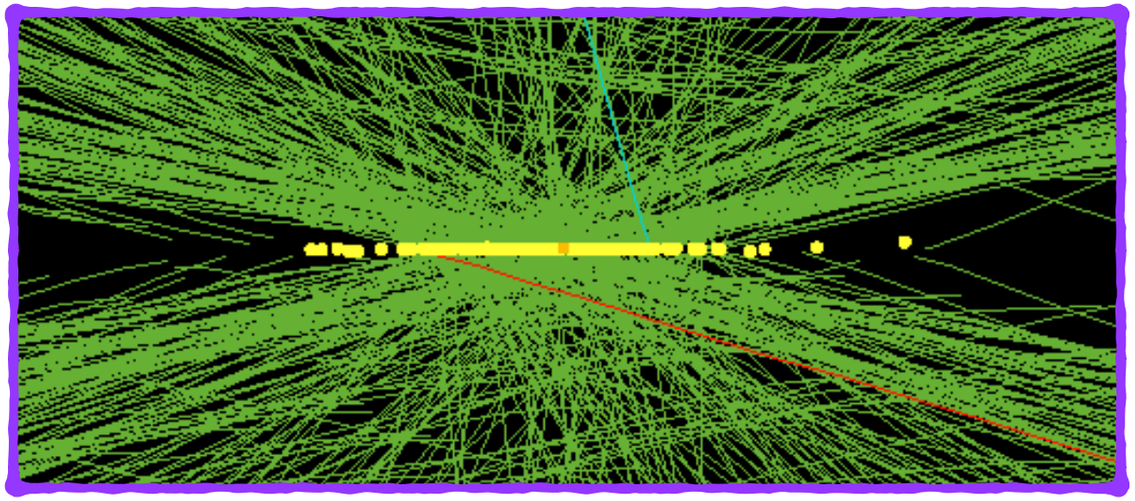
tracking information: remove PU charged particles (CHS)

problem: says nothing about neutral hadrons and photons

grooming (discussed on slide 26)

problem: usually for very high p_T jets, knows nothing of tracking





PileUp Per Particle Id

D. Bertolini, P. Harris, M. Low, N.T.
work in progress

tracking/vertexing information

global event information

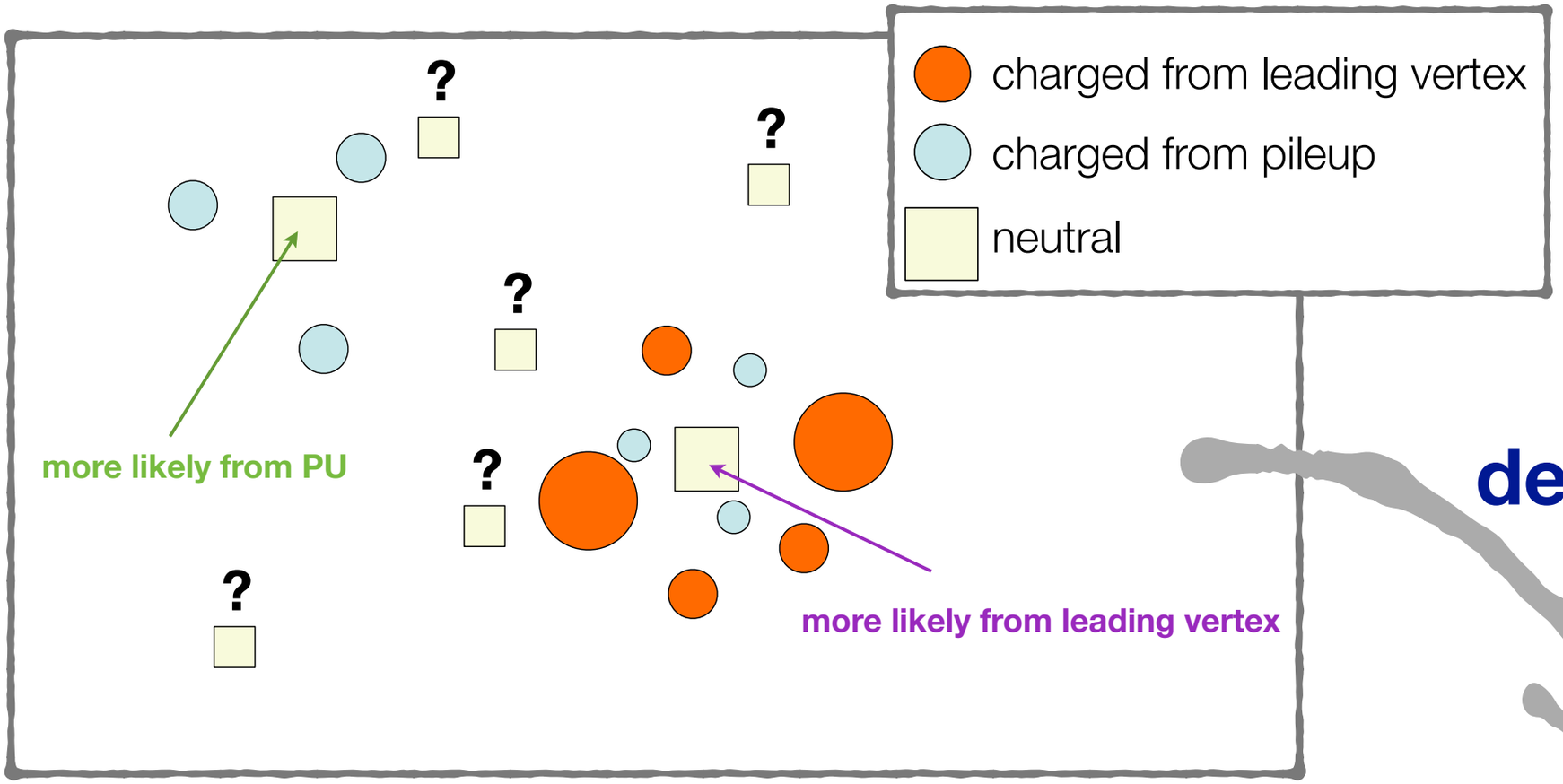
precision timing

grooming/local shape

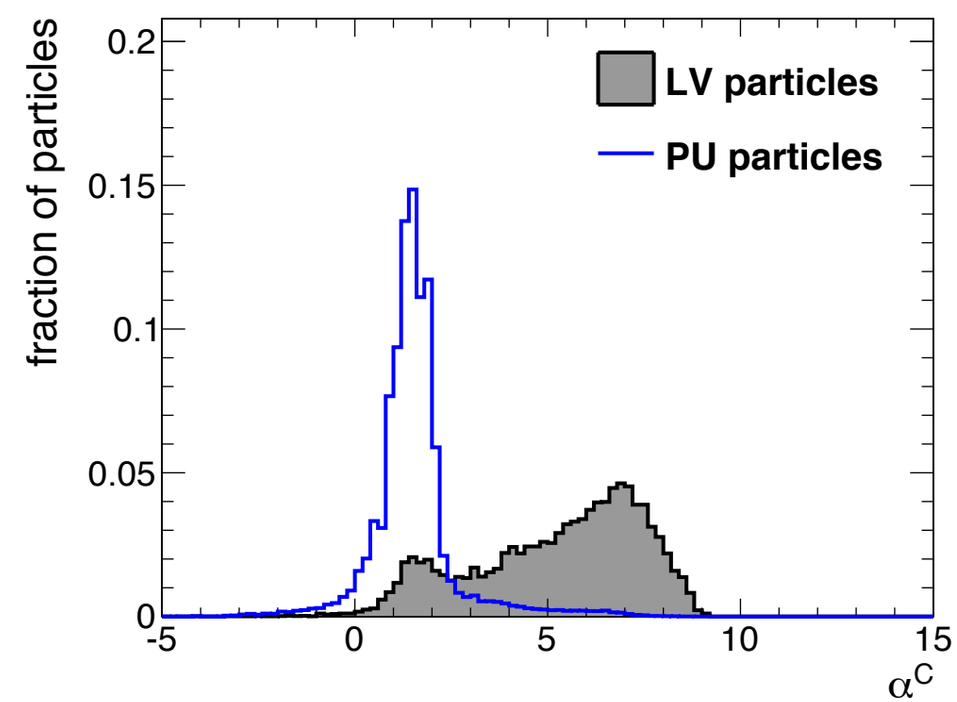
depth segmentation

Define on a ***per particle*** basis, ***before jet clustering***, a weight for ***how likely*** a particle (or jet constituent) is to be from pileup or the leading vertex, then rescale each particle four momentum by that weight

PileUp Per Particle Id



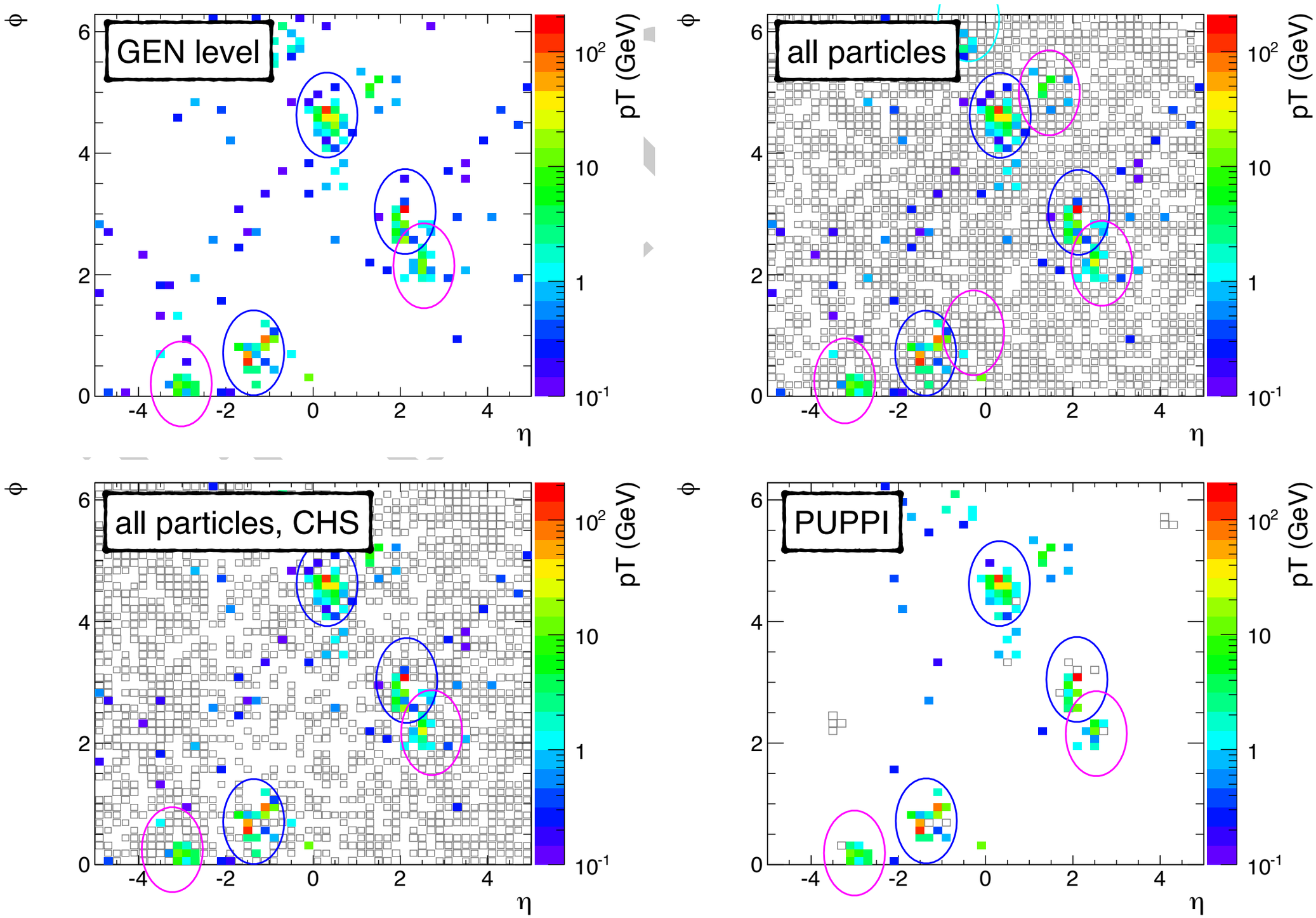
define some metric



N.B. Particle level studies assuming perfect tracking for $|\eta| < 2.5$

PileUp Per Particle Id

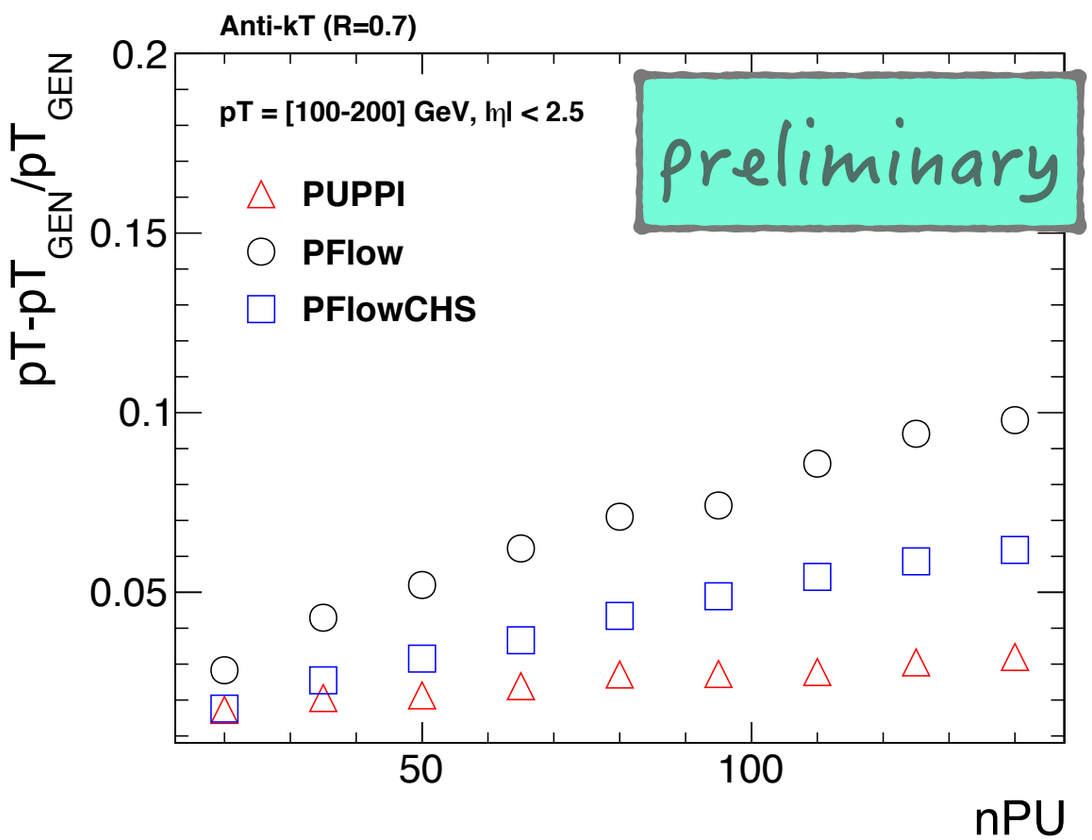
colored cells = process of interest
black cells = pileup

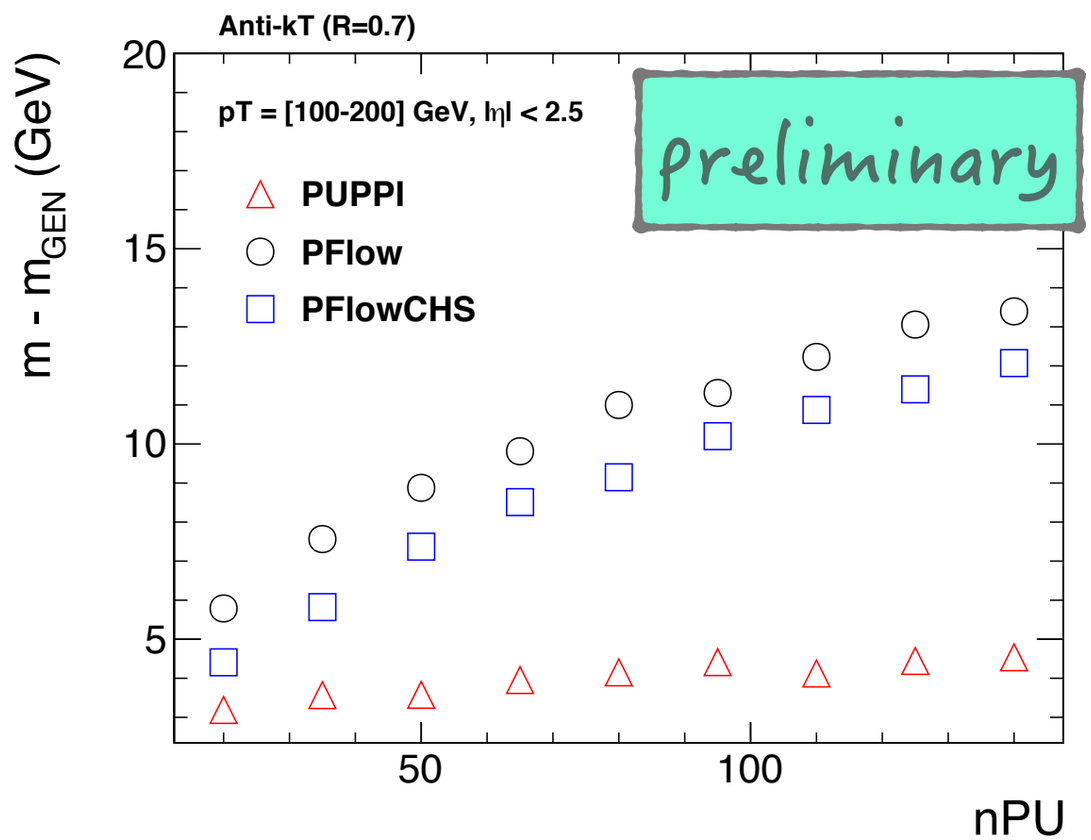


N.B. Particle level studies assuming perfect tracking for |η| < 2.5

PileUp Per Particle Id

drastic improvement in pT and mass resolutions





N.B. Particle level studies assuming perfect tracking for |η| < 2.5

summarizing

a team effort

such a **rapid development** of jet technology is the responsibility of many

a community of physicists dedicated to these issues

experimenters and **theorists** working hard to **calculate, model, measure and implement**

incredibly **quick turnaround**, ideas can be proposed and measurements made public in ~1 year

jets, subjets and particles

the way we treat jets is evolving

a jet is no longer a four-vector

we are pulling them apart, looking at their structure, shapes and properties of subjets

without these methods, many physics analyses would need much more data or might not even be possible

the future is now, classify each particle in an event

jets simply as a way to organize the event for YOUR analysis

bonus material

case #4:

low MET SUSY

reorganizing the event

analyses not bound a standard jet size, R~0.4

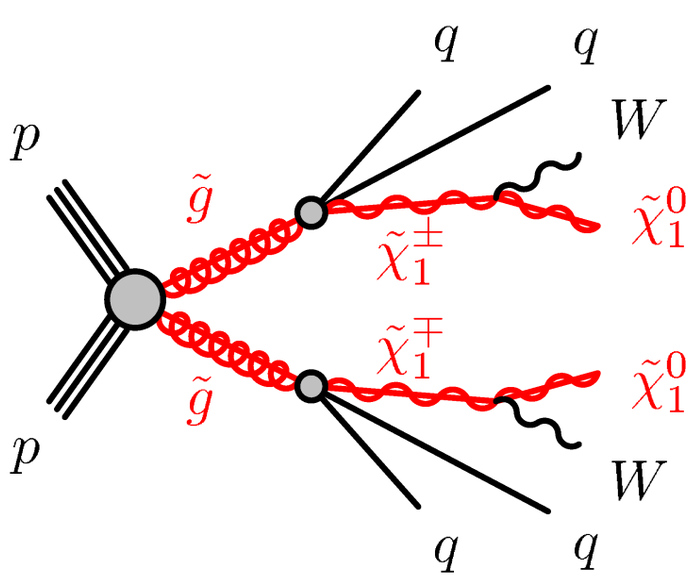
standard boosted case:

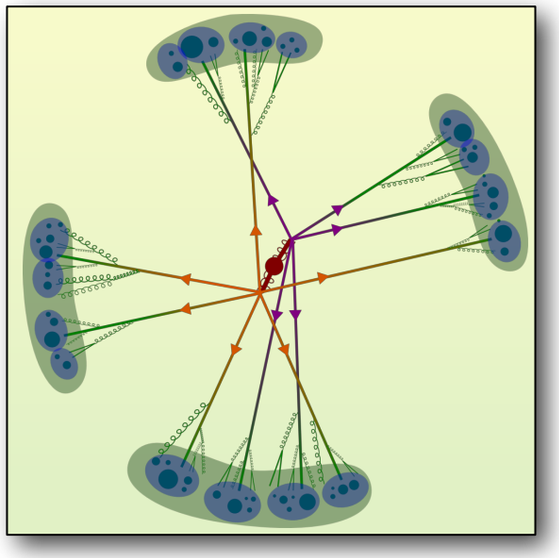
for boosted signals, it's often easier to have a larger clustering parameter R to **improve acceptance**

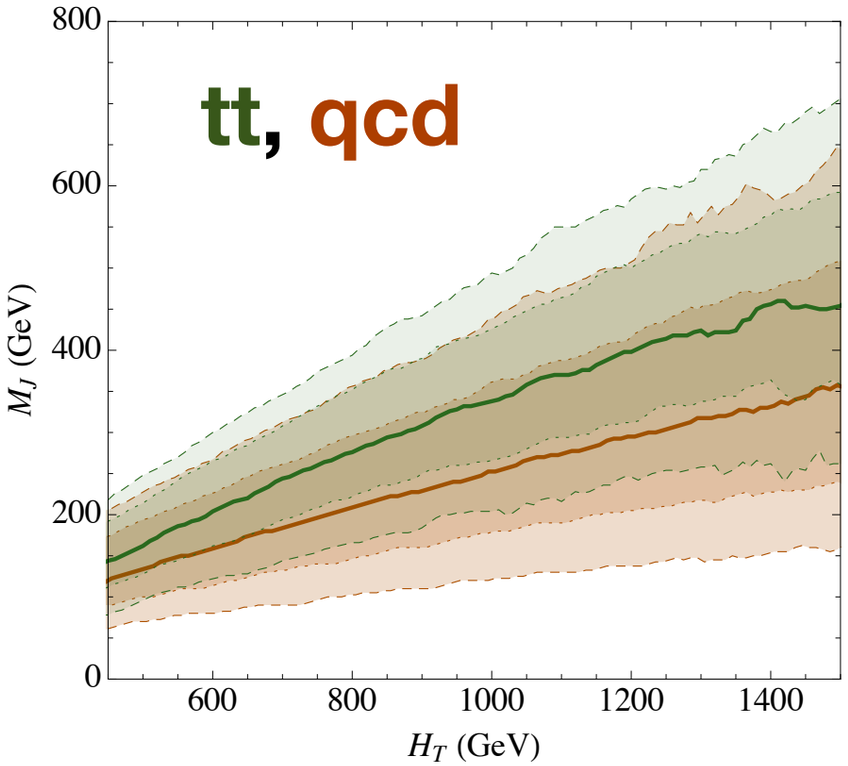
also can use techniques **for moderate to no boost**

e.g. tops with p_T = 150 GeV using R=1.5 jets

e.g. high parton multiplicity SUSY topologies with little MET

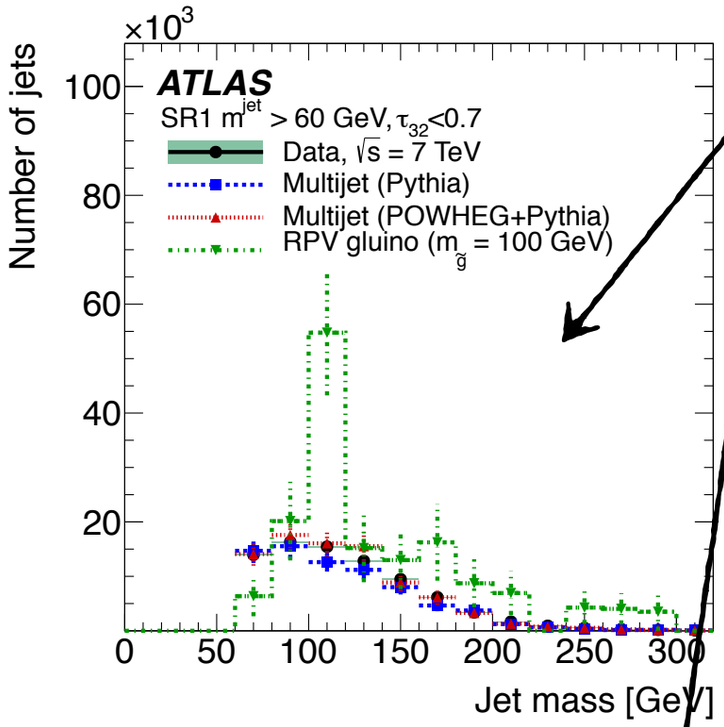




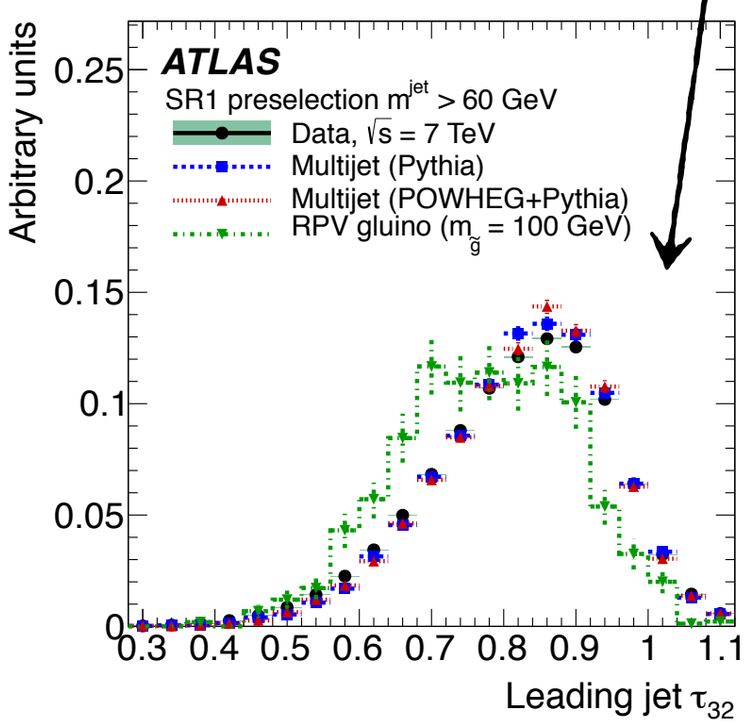


low MET SUSY

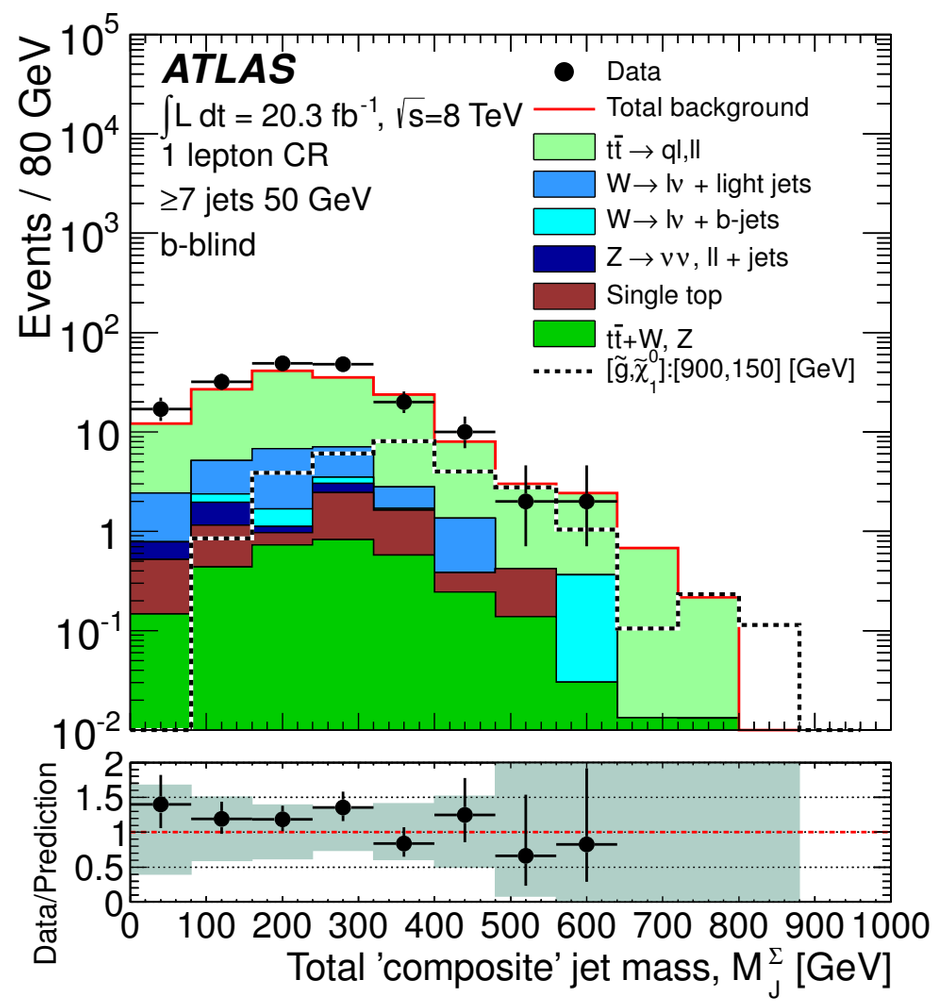
CERN-PH-EP-2012-281
ATLAS-CONF-2013-091

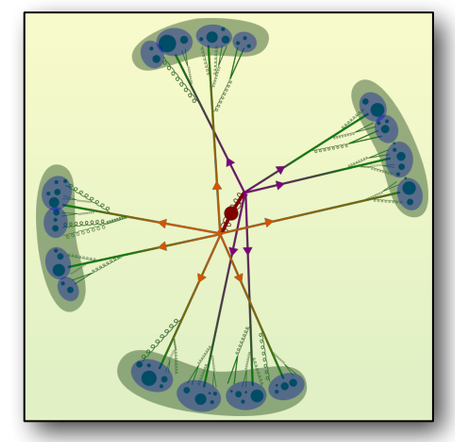


Use of substructure objects in non-SM boosted objects too! High mass RPV gluino decaying to qqq

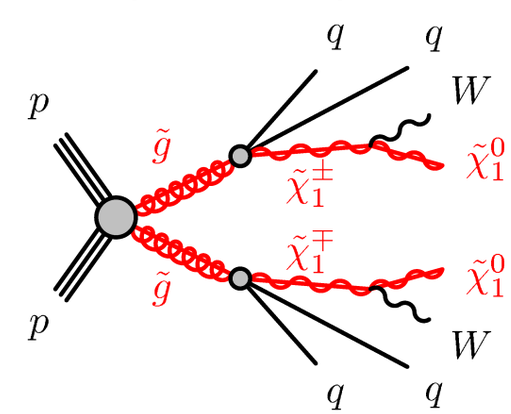


Interesting new ideas to use jet substructure observables for high jet multiplicity final states with low MET





example interpretation



additional material

observables

pT,Y,ϕ + tracking

mass

4-vector sum of jet constituents

highly sensitive to soft QCD and pileup; **grooming** can be used to mitigate these dependencies

shapes

several classes: declustering/reclustering, generalized jet shapes and energy flow, statistical interpretation (Qjets), jet charge

algorithms

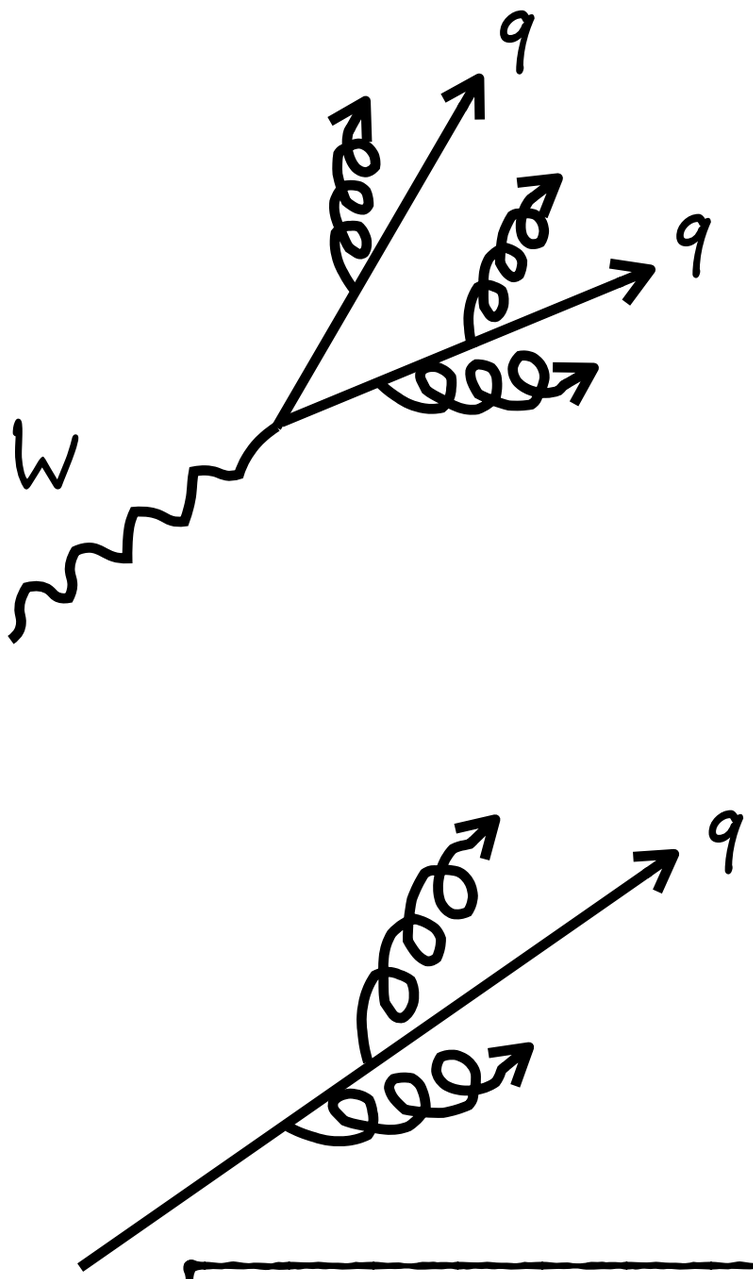
some combination of cuts on mass, shapes, tracking

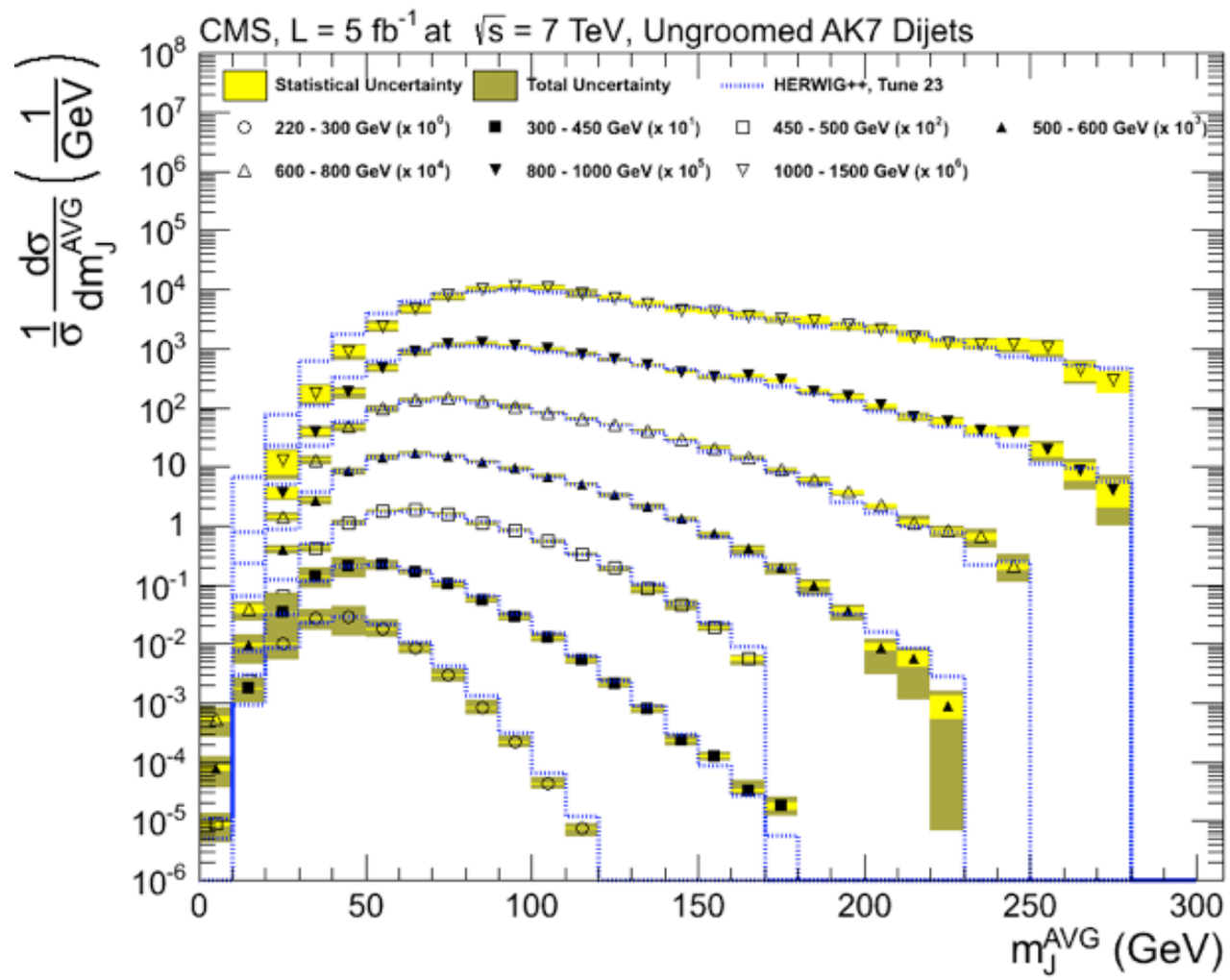
most typical in **top tagging**

jet mass

JHEP 05 (2013) 090

but jet mass is a perturbative quantity





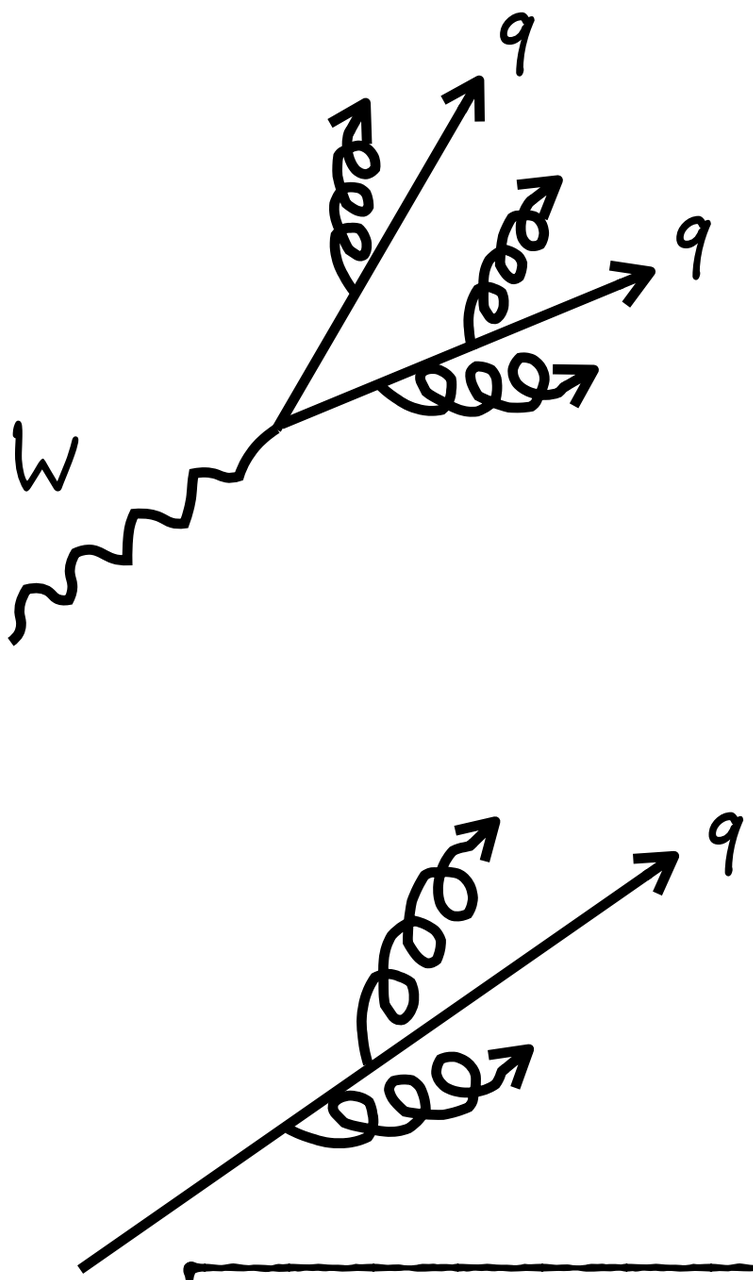
$$\langle M^2 \rangle \simeq \left.\begin{array}{ll}\text{quarks:} & 0.16 \\ \text{gluons:} & 0.37\end{array}\right\} \times \alpha_s p_t^2 R^2$$

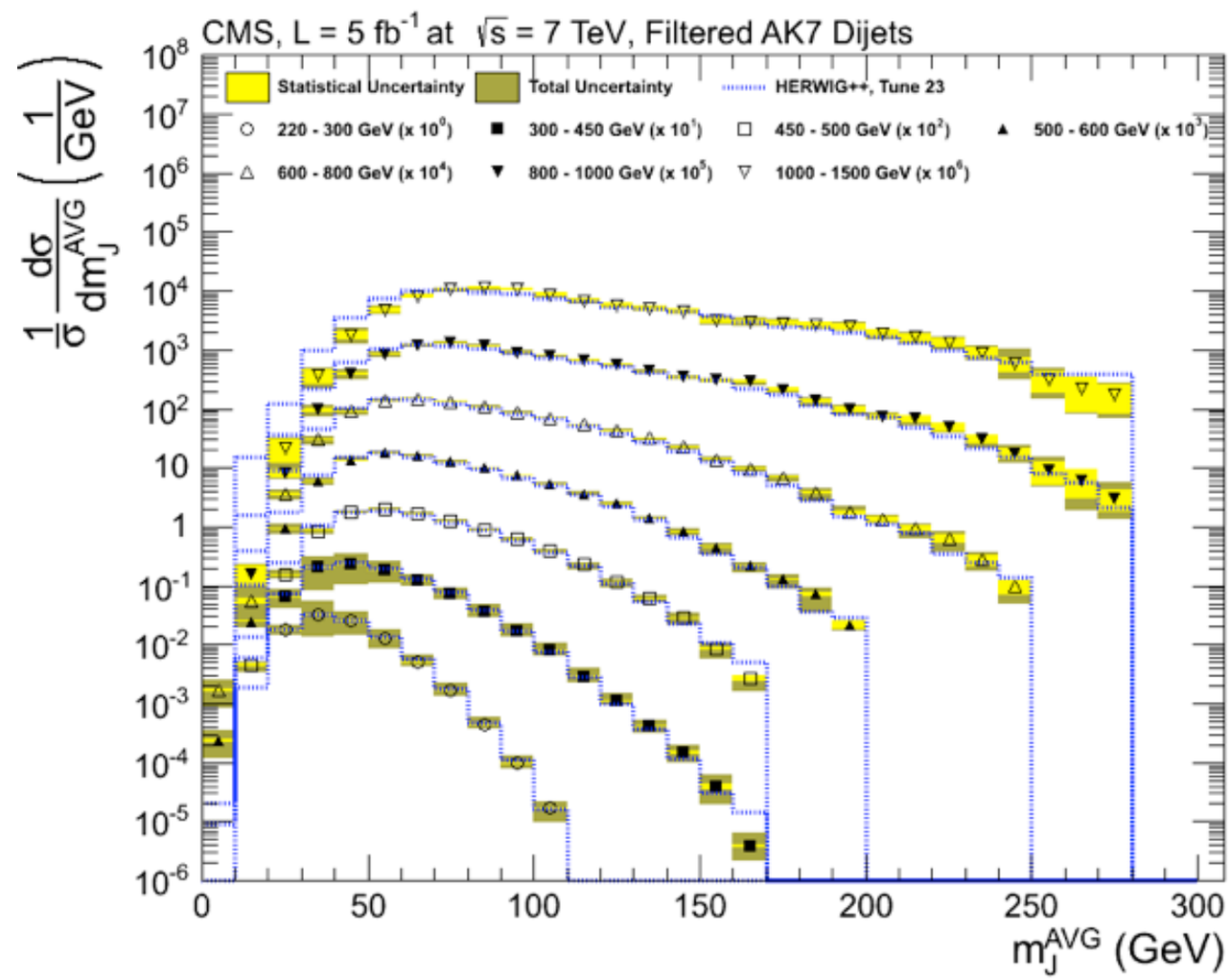
plain jet mass

jet mass

JHEP 05 (2013) 090

but jet mass is a perturbative quantity





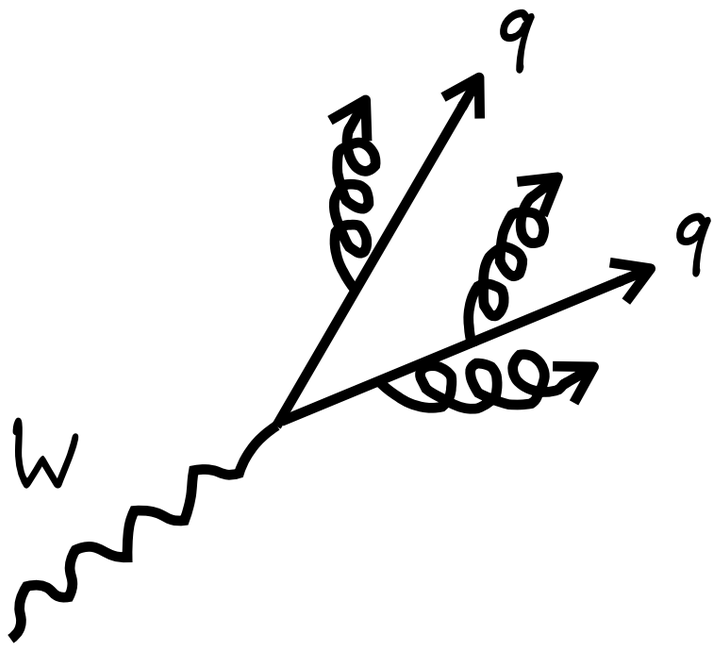
$$\langle M^2 \rangle \simeq \left.\begin{array}{ll}\text{quarks:} & 0.16 \\ \text{gluons:} & 0.37\end{array}\right\} \times \alpha_s p_t^2 R^2$$

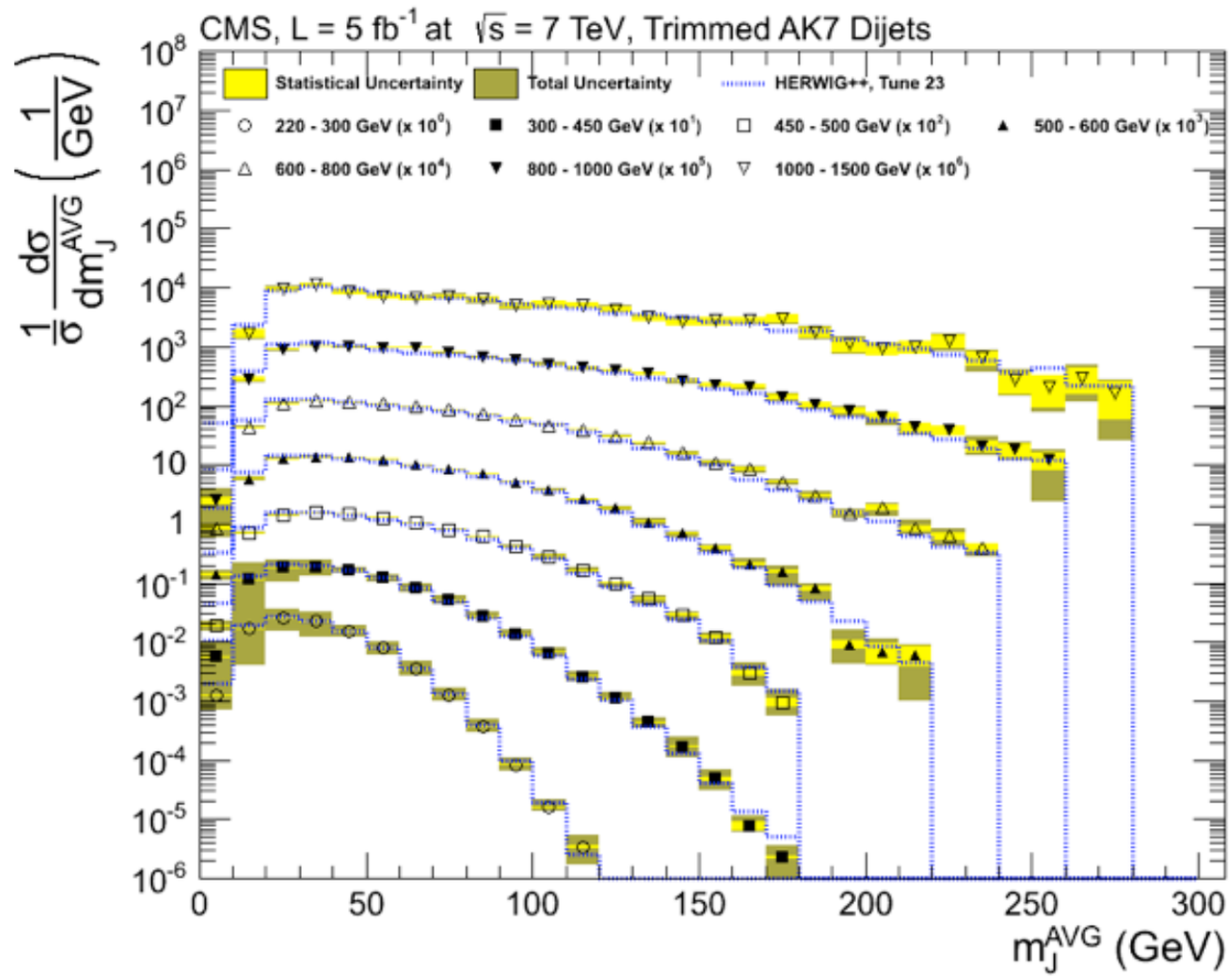
filtered jet mass

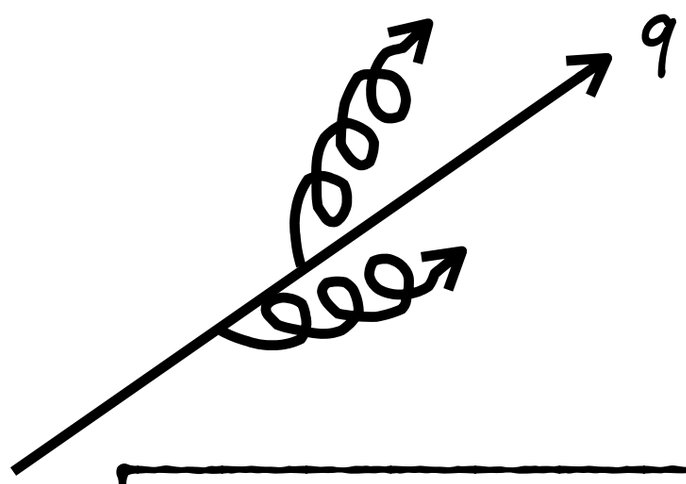
jet mass

JHEP 05 (2013) 090

but jet mass is a perturbative quantity







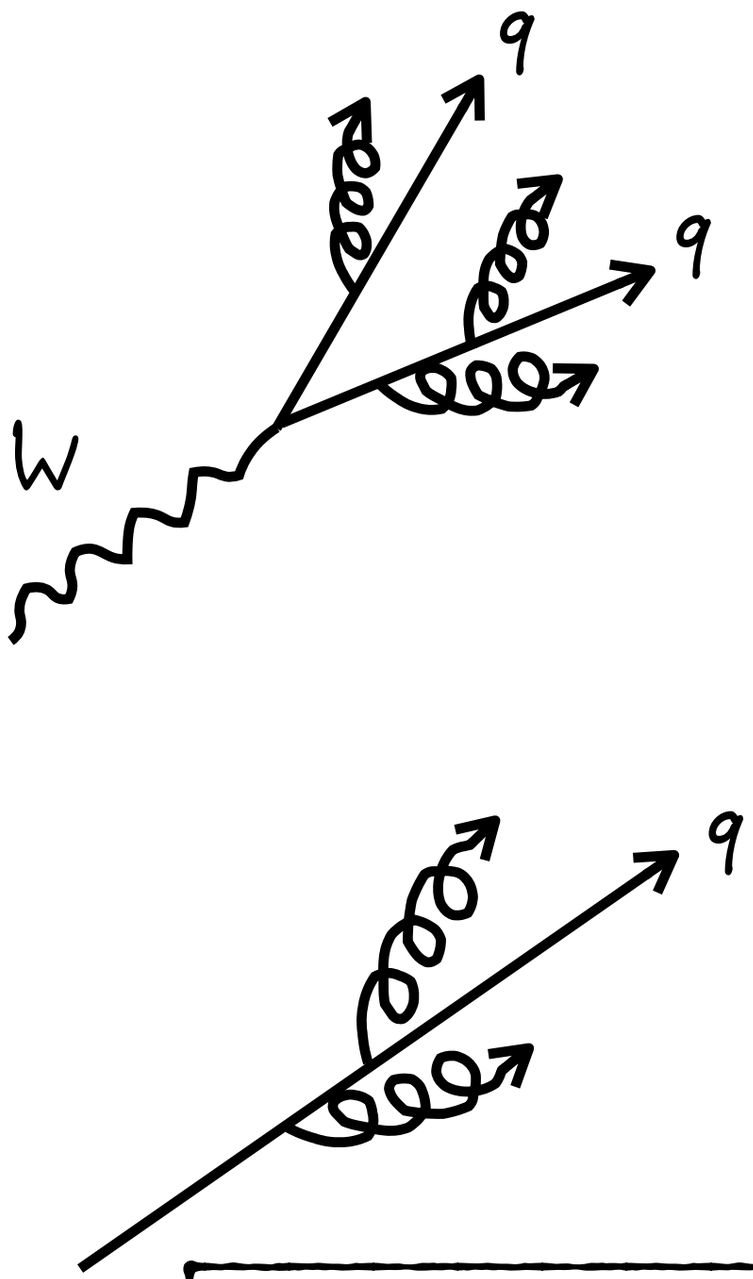
$$\langle M^2 \rangle \simeq \left.\begin{array}{ll}\text{quarks:} & 0.16 \\ \text{gluons:} & 0.37\end{array}\right\} \times \alpha_s p_t^2 R^2$$

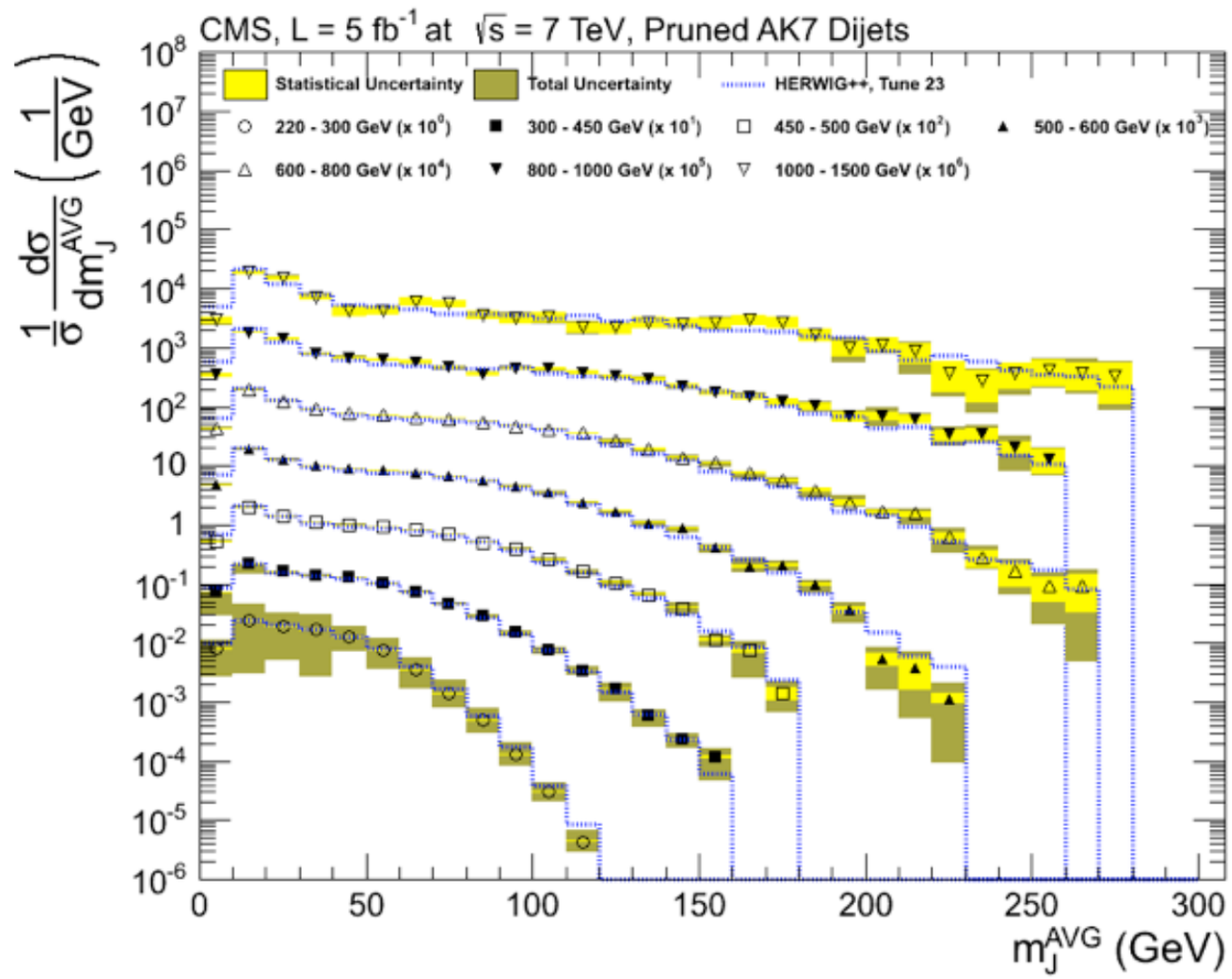
trimmed jet mass

jet mass

JHEP 05 (2013) 090

but jet mass is a perturbative quantity





$$\langle M^2 \rangle \simeq \left.\begin{array}{ll}\text{quarks:} & 0.16 \\ \text{gluons:} & 0.37\end{array}\right\} \times \alpha_s p_t^2 R^2$$

pruned jet mass

tools:

shapes

Declustering and reclustering

For modern sequential recombination jet algorithms, the jet has a history — a series of 2-to-1 combinations

examples: **mass drop** - (m_{sj1}/m) , $\sqrt{d_{12}}$ - first splitting scale of kT algorithm

Generalized jet shapes

some simple questions: *How wide is a jet? How prong-y is a jet? How asymmetric is a jet? How stable is a jet?*

N-subjettiness (τ_N) [1], how consistent a jet is with having N subjets, ratios are typically used, e.g. τ_2/τ_1 for W-jets, τ_3/τ_2 for top jets

energy correlation functions [2], axis-less version of N-subjettiness

Qjets [3], Exploiting the "quantum" nature of jets

jet charge [4], an oldy but a goody

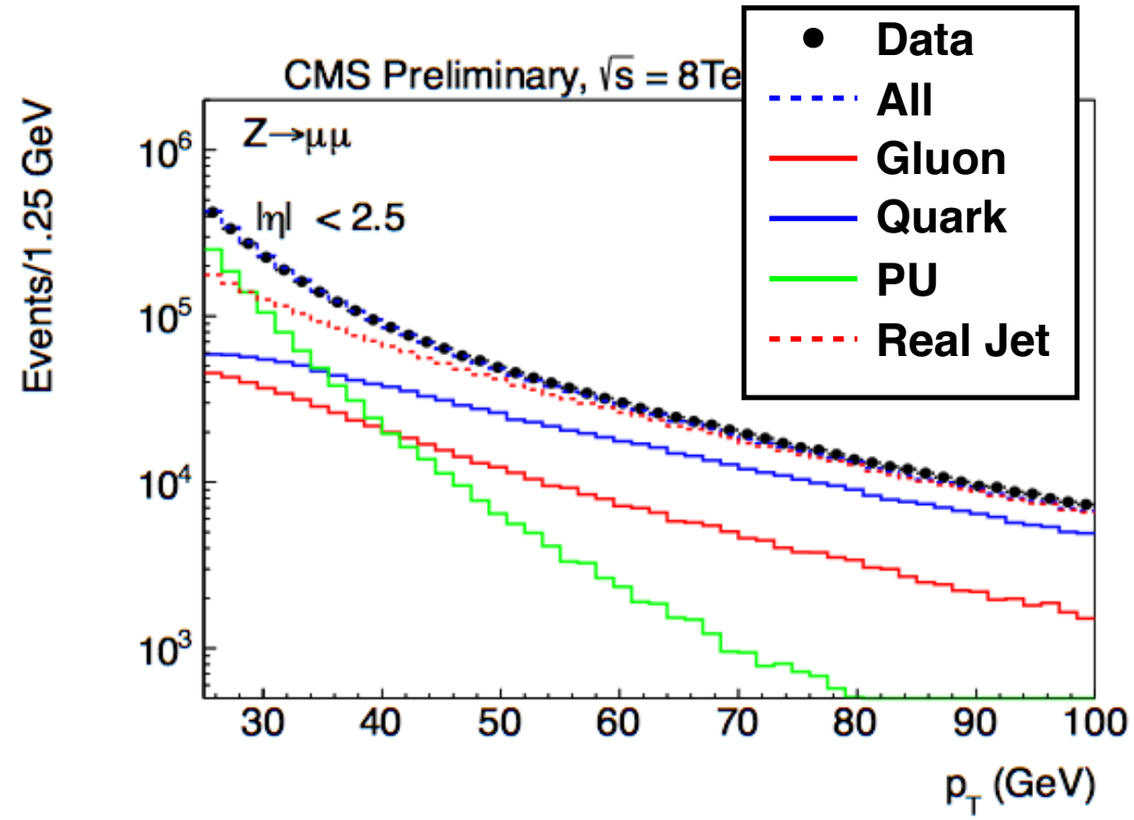
jet width; pT_D; r-cores; planar flow...

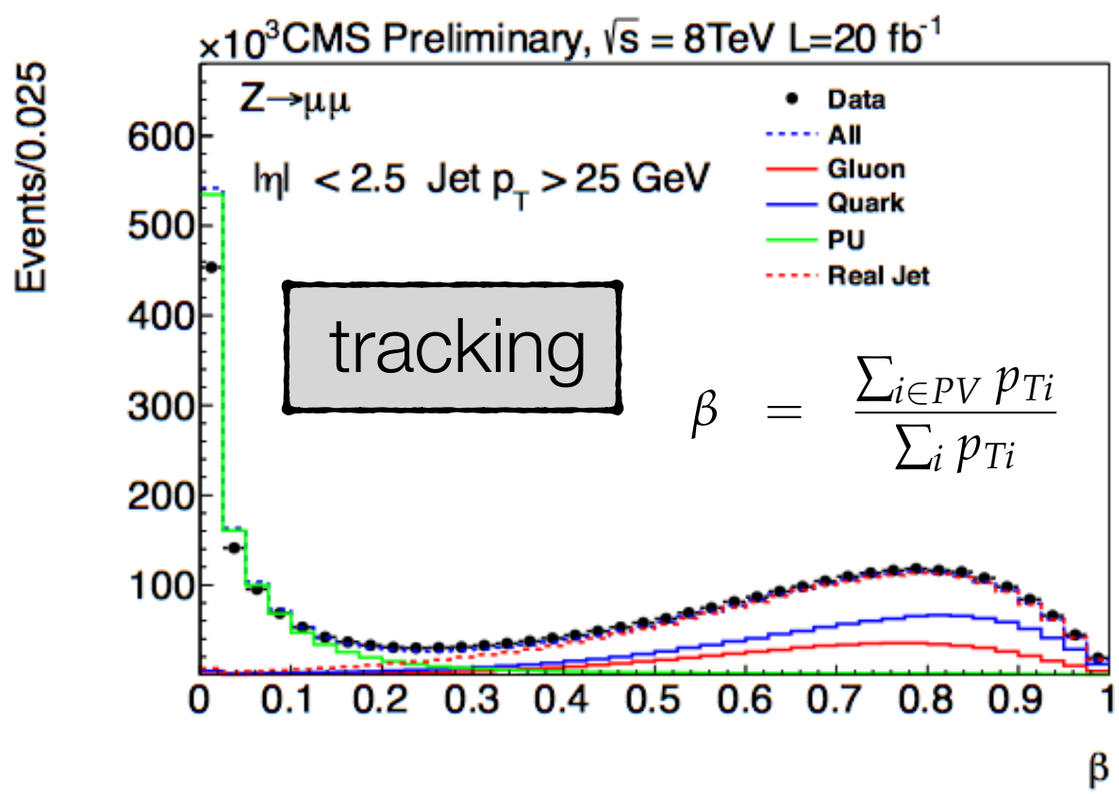
[1] Thaler, Van Tilburg, JHEP 1103:015,2011
[2] Larkoski, Salam, Thaler, arXiv:1305.0007
[3] Ellis et al., PRL 108, 182003 (2012)
[4] Krohn et al., Phys. Rev. Lett. 110 (2013) 21200

jet counting at low p_T

The low pT regime is dominated by **quark**, **gluon**, and **pileup jets**.

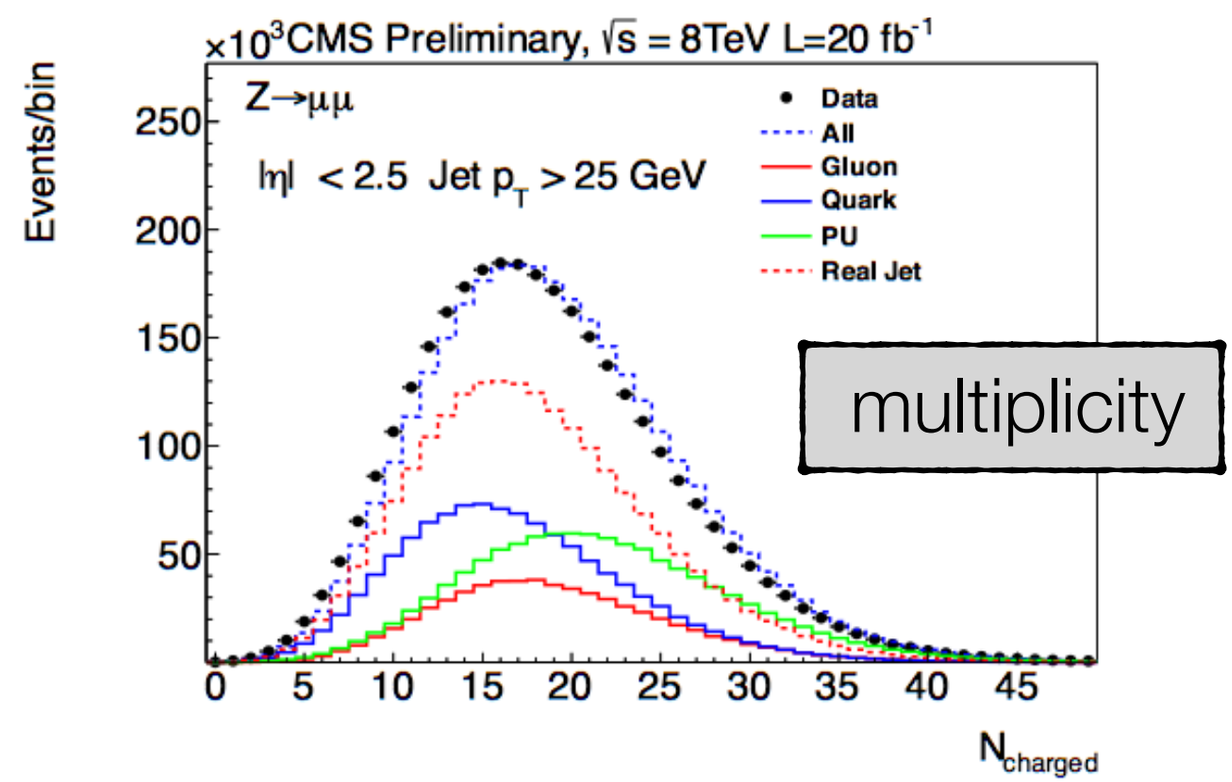
Using all information, can do a better job discriminating them:
track-based observables and jet shapes





tracking

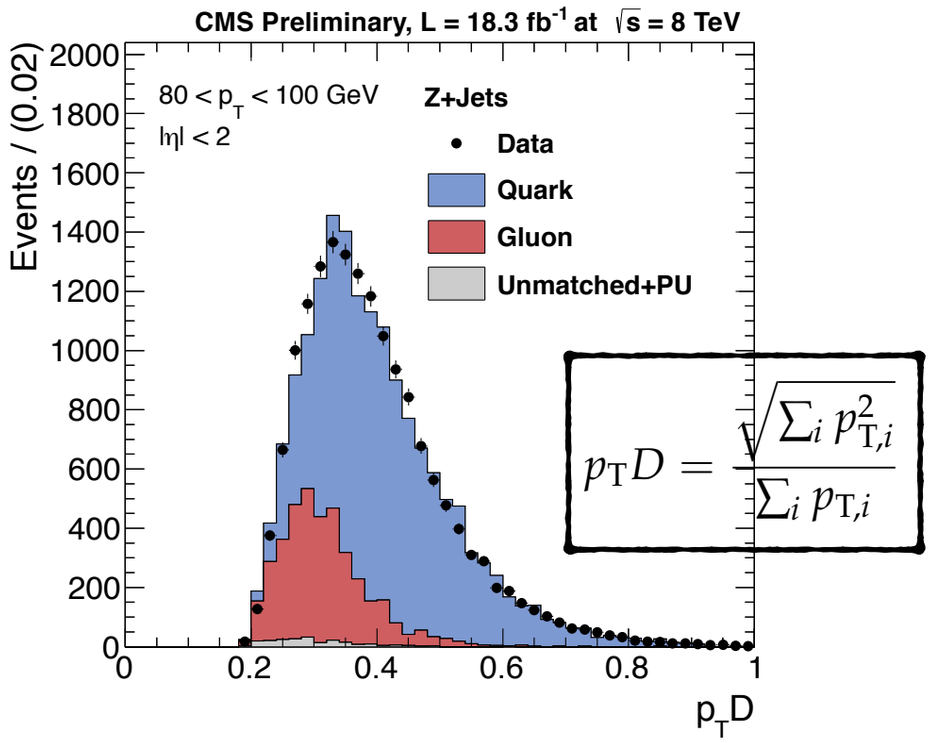
$$\beta = \frac{\sum_{i \in PV} p_{Ti}}{\sum_i p_{Ti}}$$



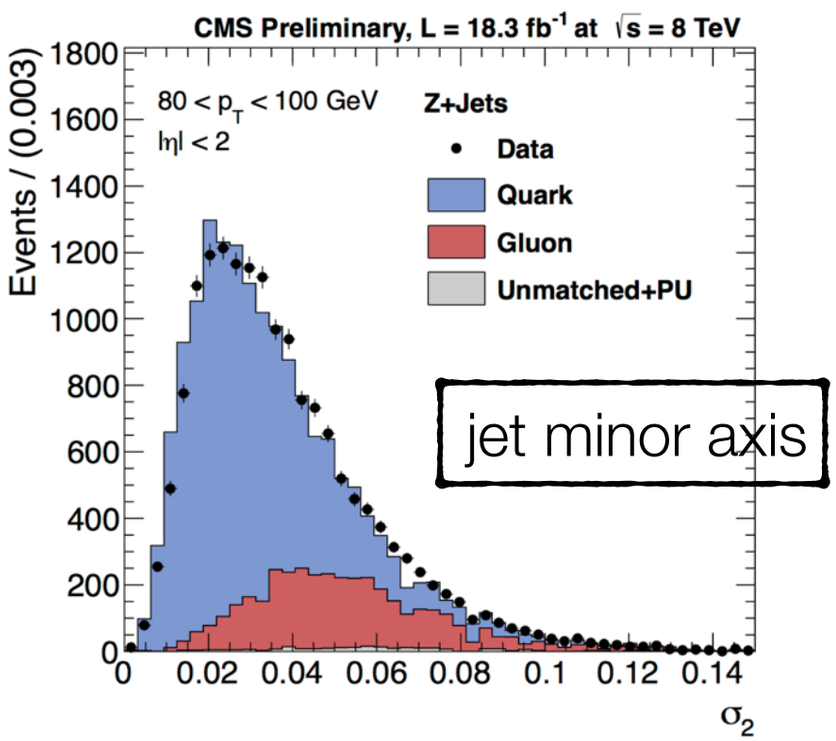
multiplicity

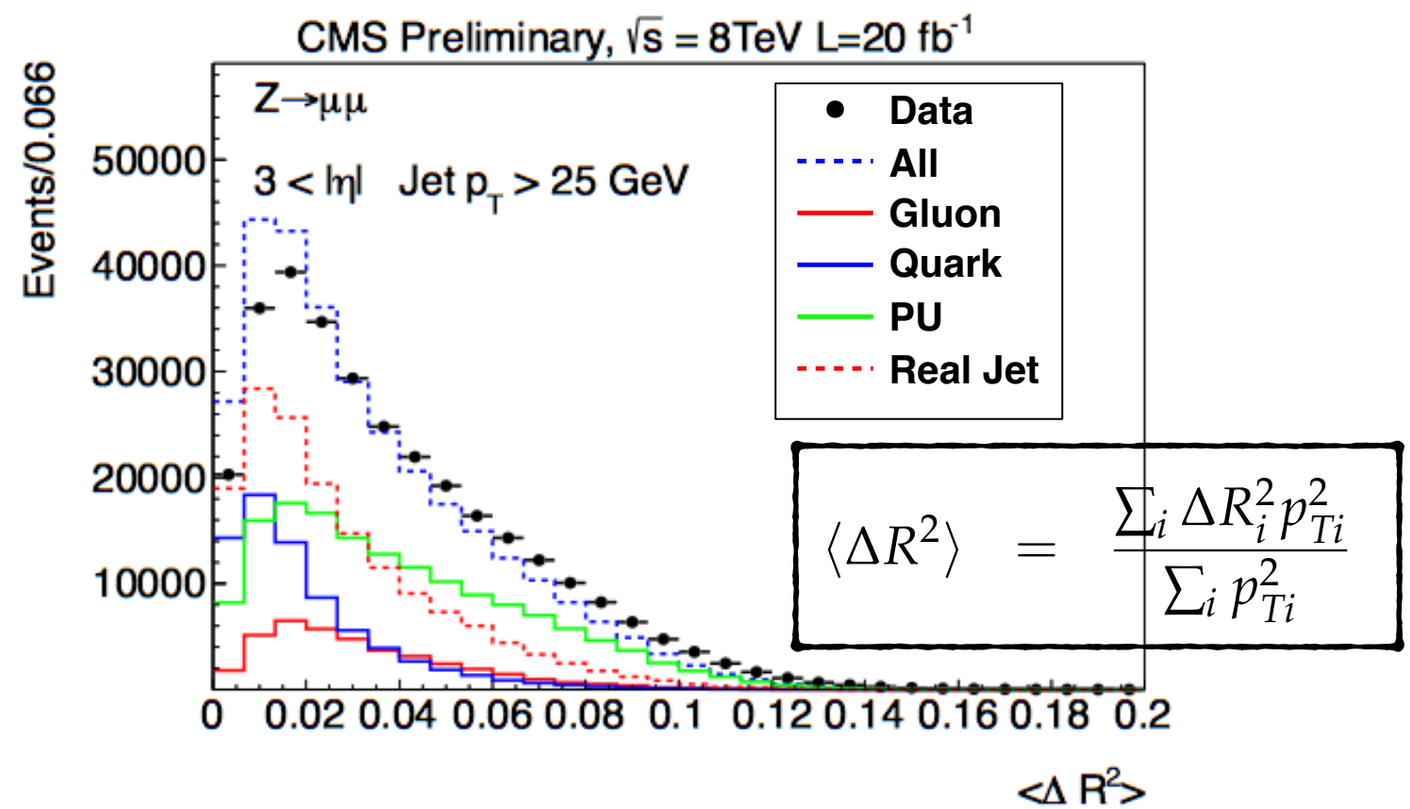
jets at low p_T

shapes



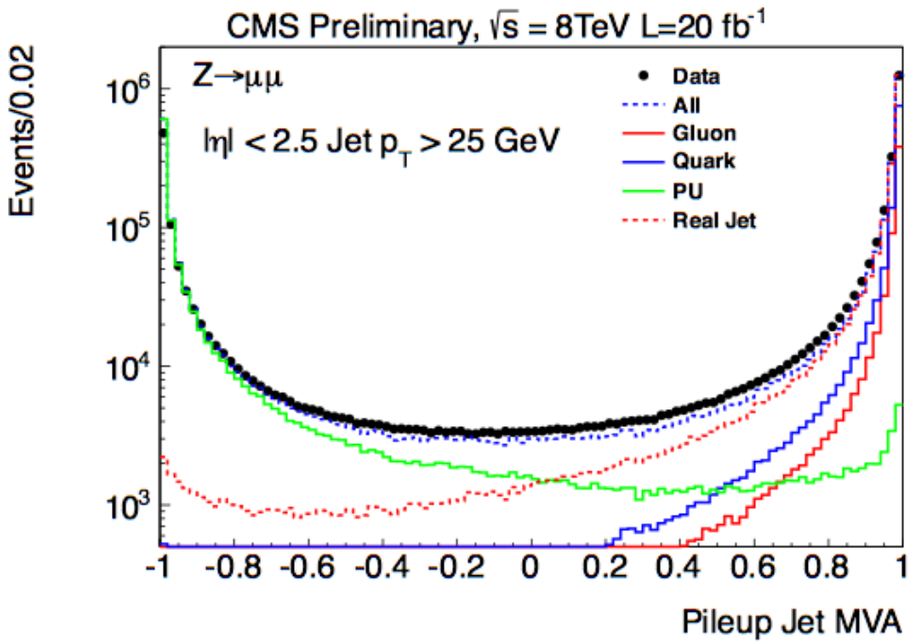
A variety of jet shapes are used for discriminating quarks and gluons, primarily trying to highlight the higher color charge of the gluon w.r.t. the quark.

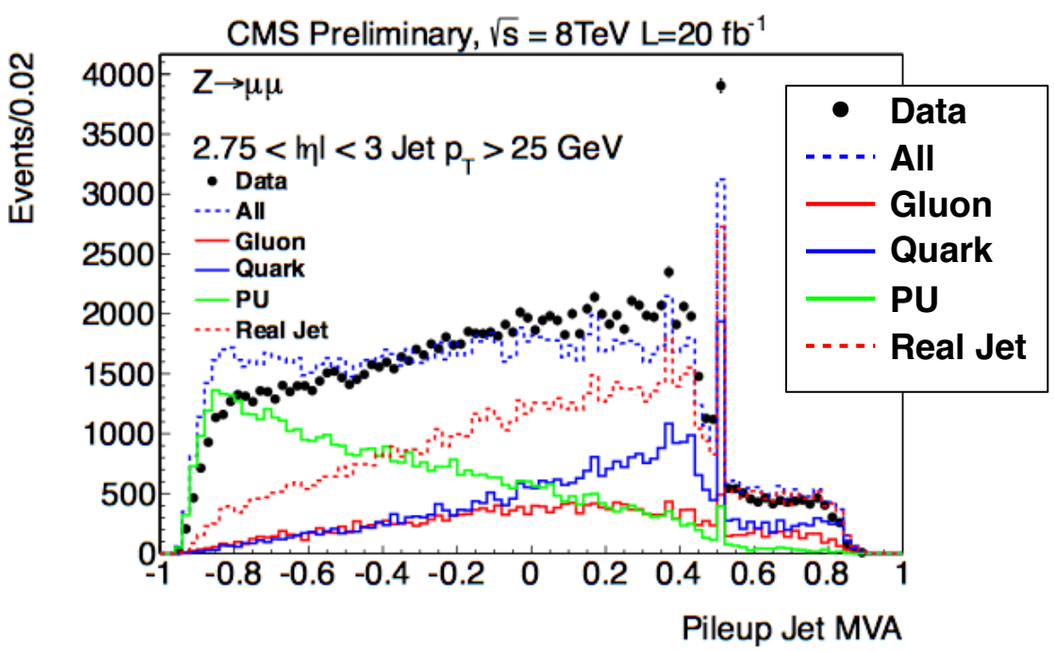




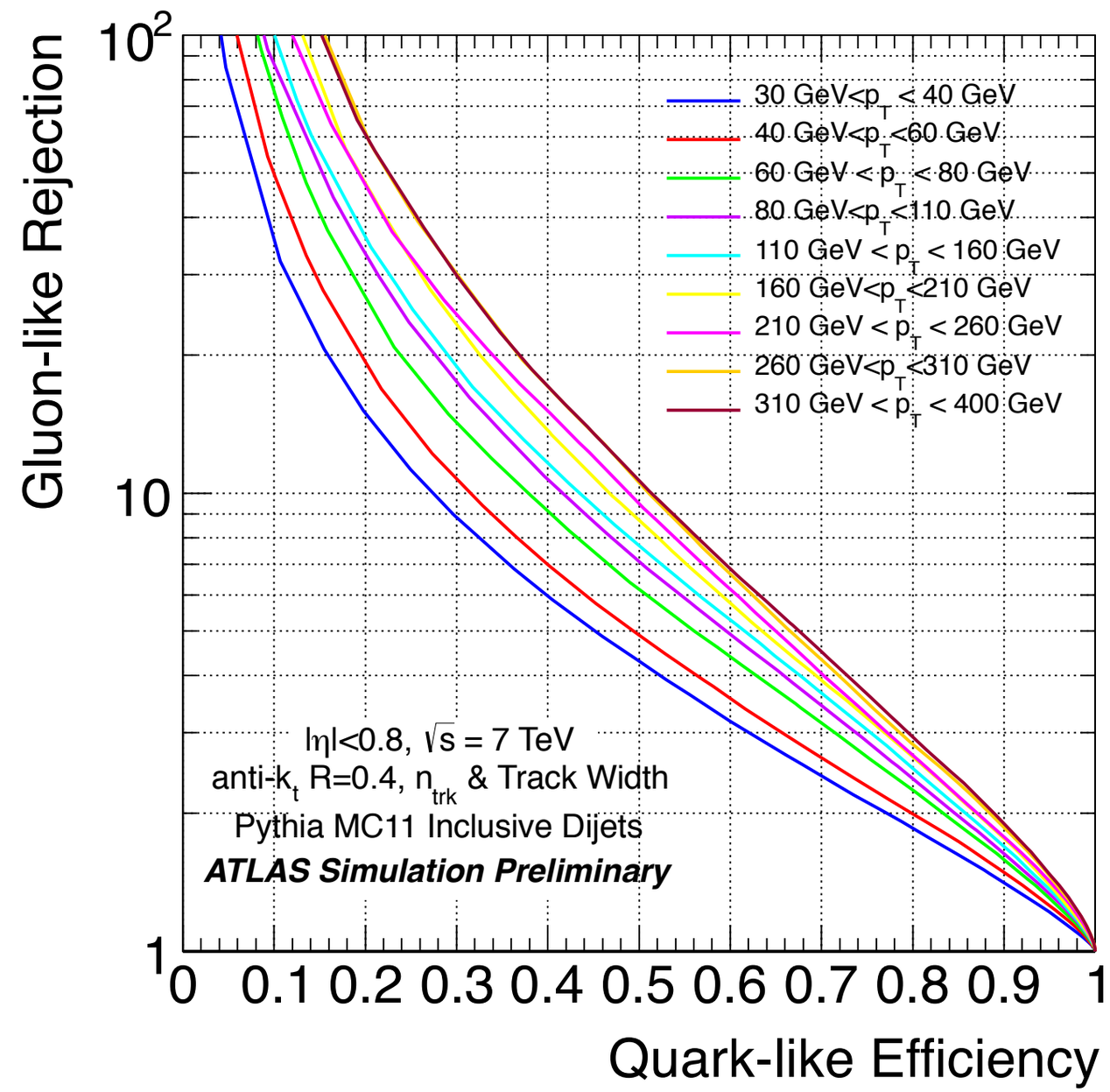
low p_T taggers

Pileup jet ID puts 13 variables into a BDT discriminant





Quark/gluon discriminant puts 2 (4) variables into a likelihood for ATLAS (CMS)



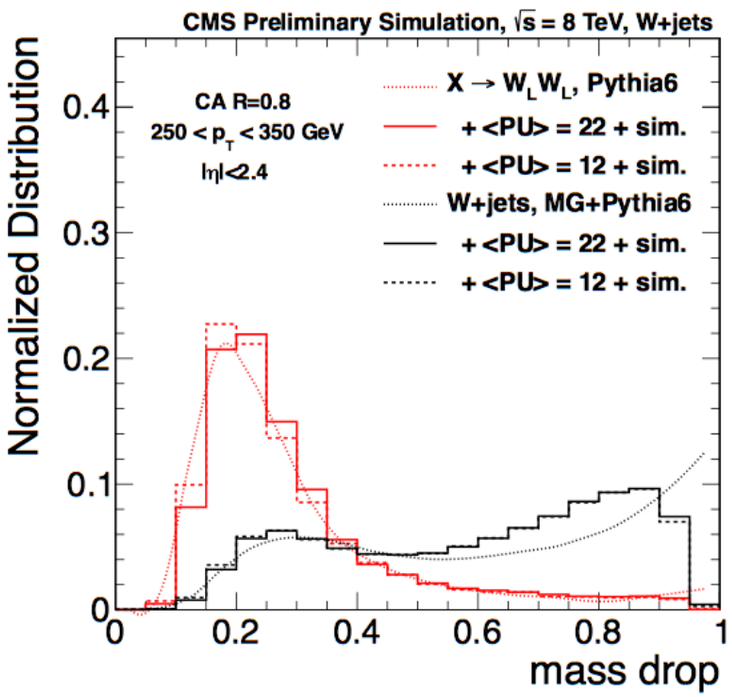
shapes for boosted objects

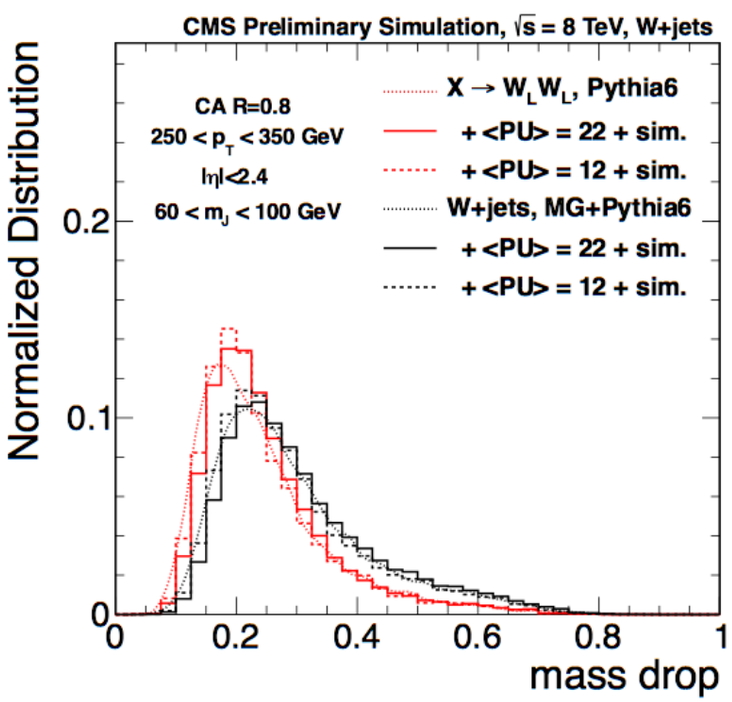
beware the correlations!

CMS PAS JME-13-006

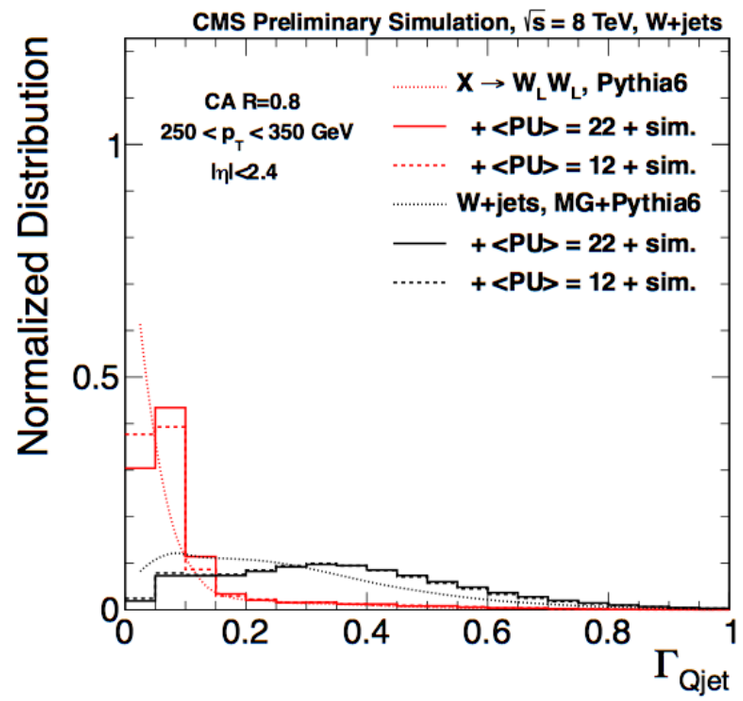
pruned mass cut

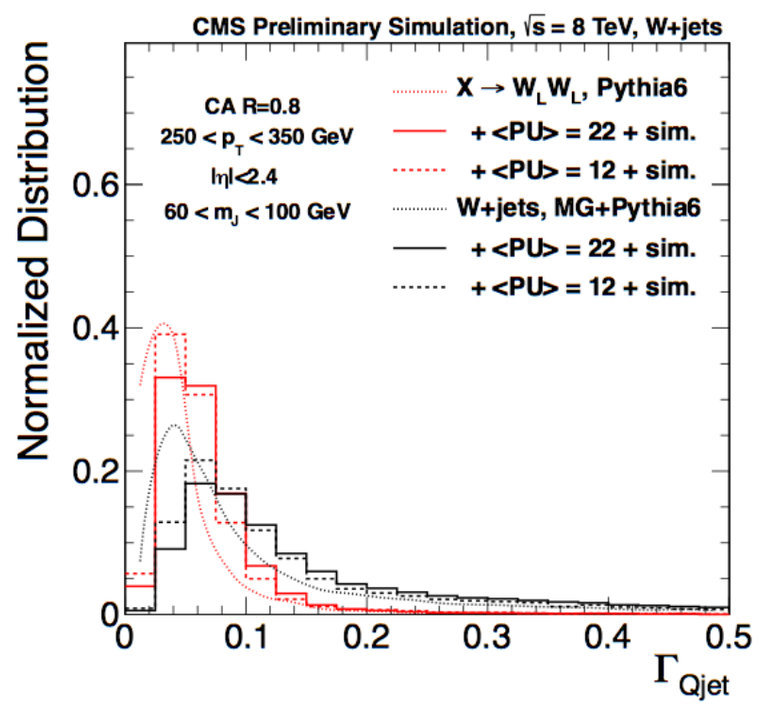
mass drop



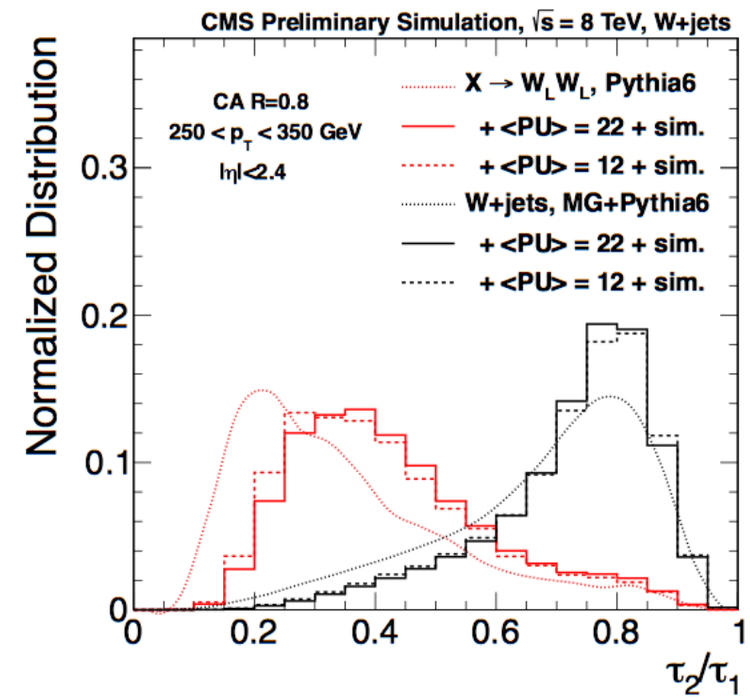


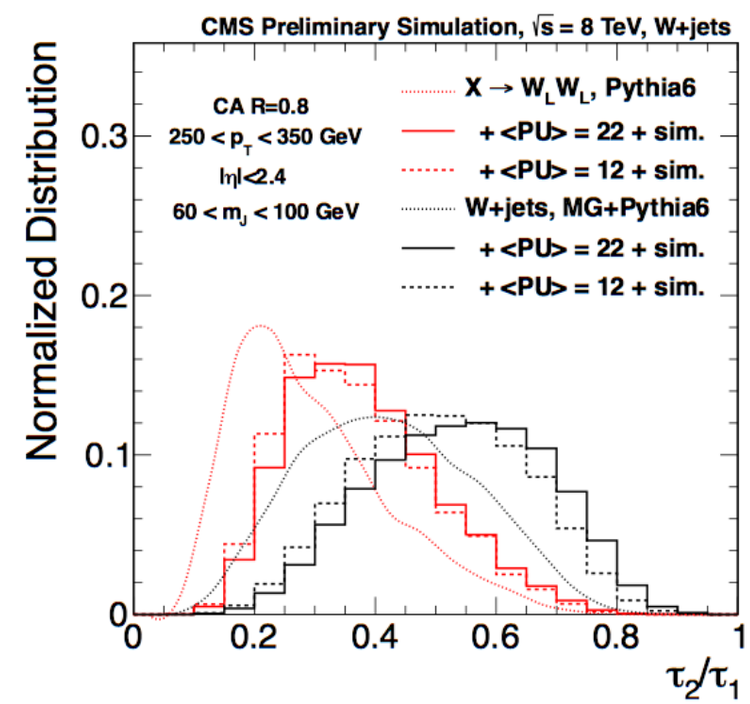
Qjet volatility





τ_2/τ_1





mass + shapes

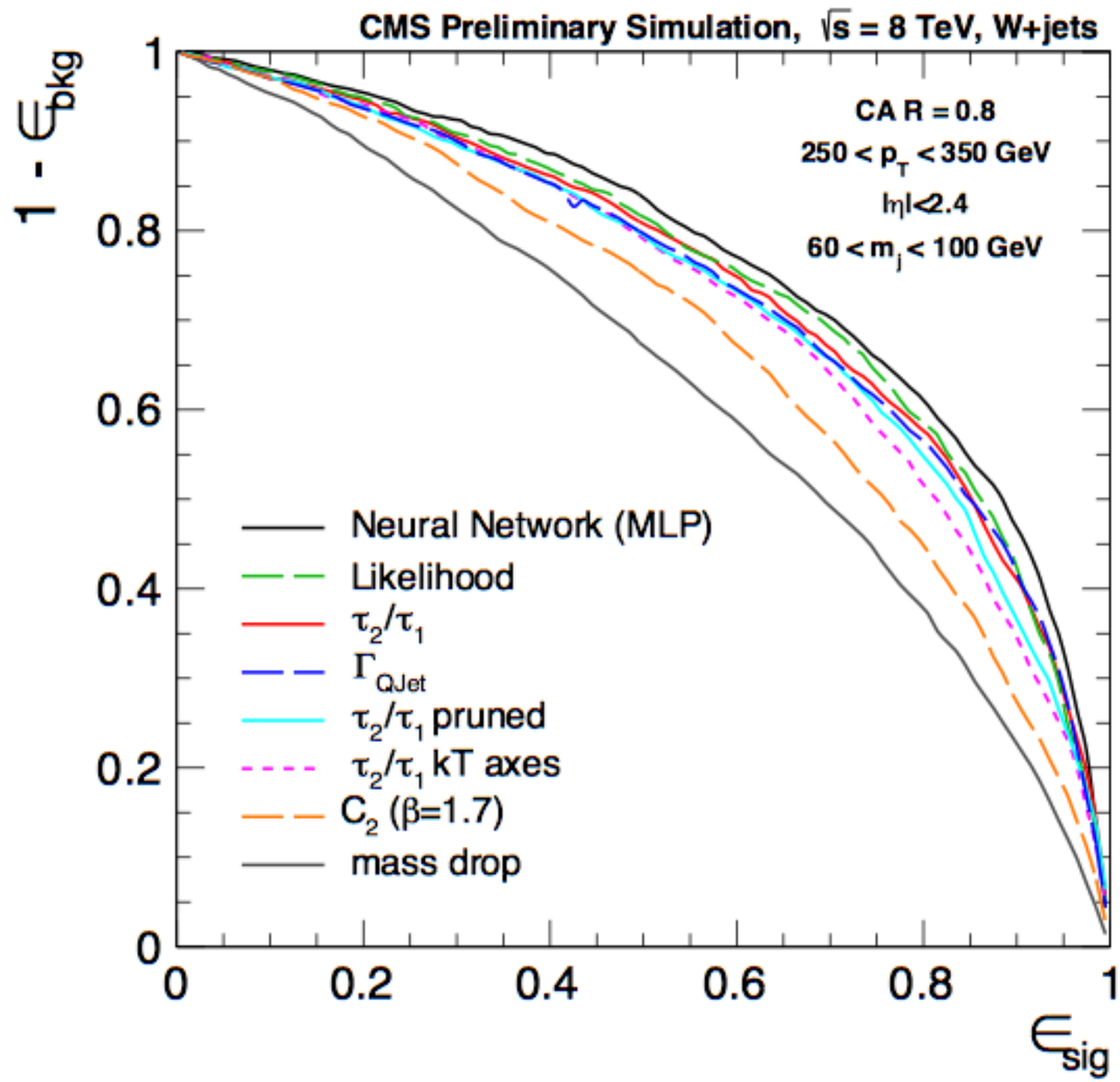
CMS PAS JME-13-006

After pruned mass cut...

Compare the performance of several different observables

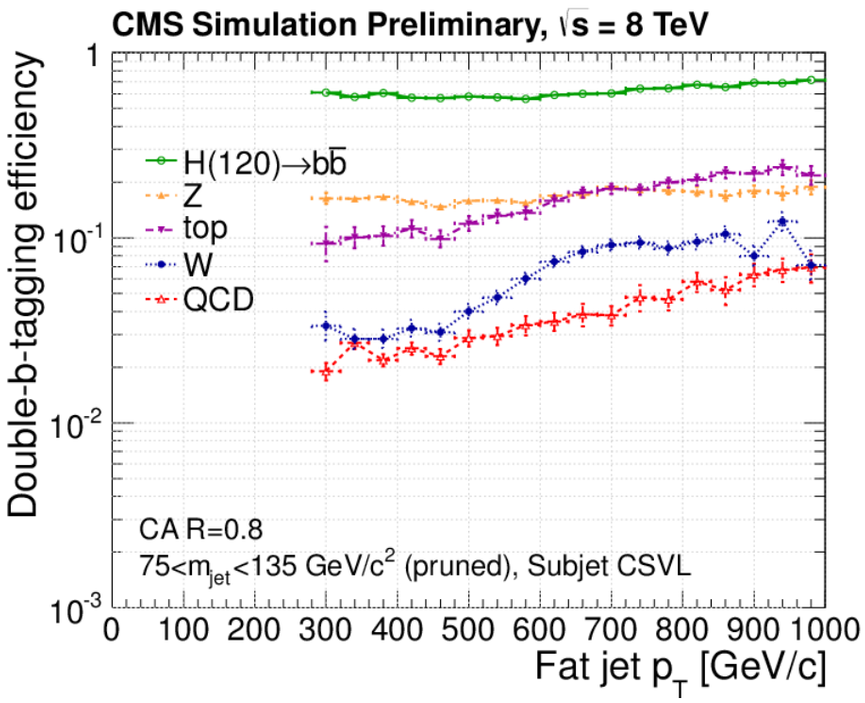
- Qjets and N-subjettiness are most powerful single variables

Try combining observables into an MVA discriminant; a small improvement



subjet b-tagging

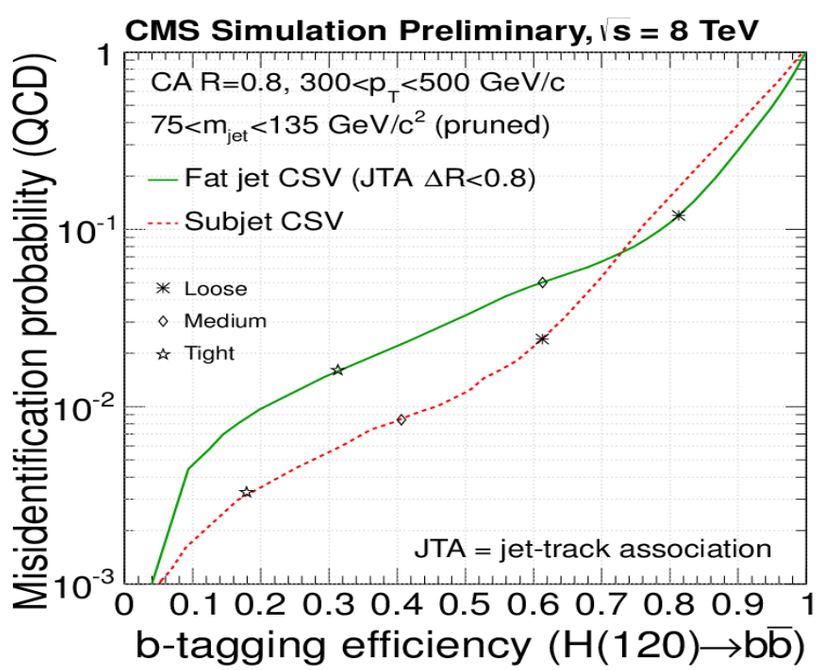
CMS PAS BTV-13-001

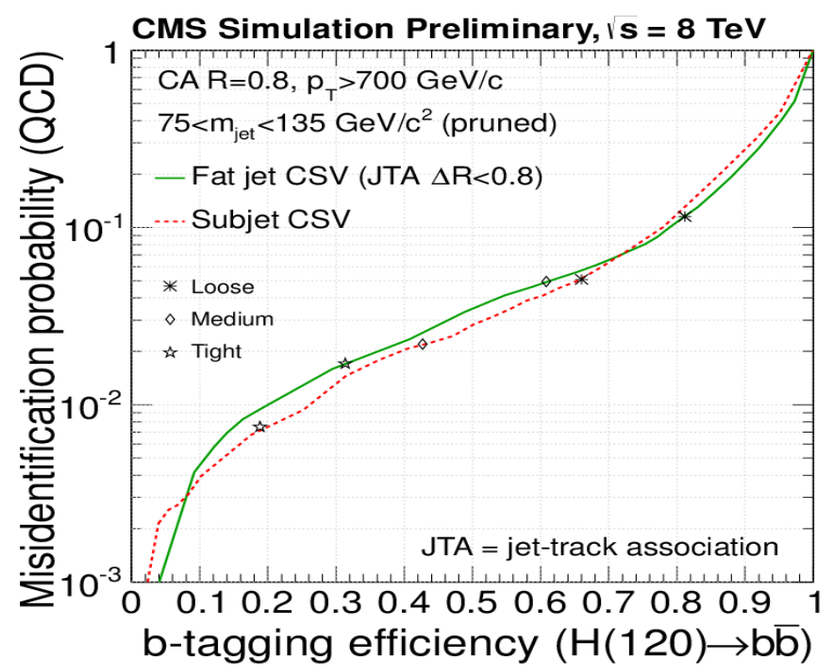


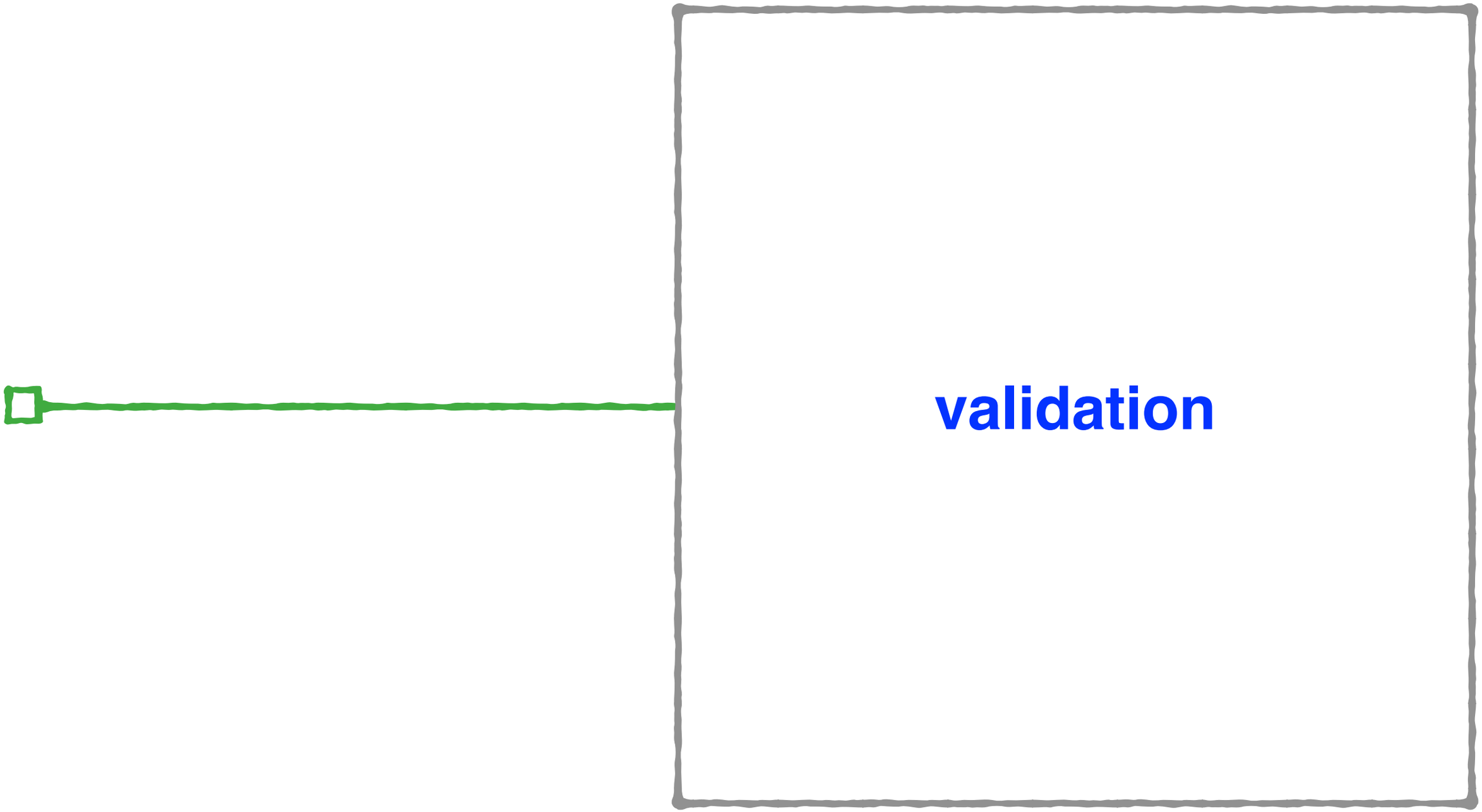
Extending standard b-tagging techniques, can **apply b-tagging to subjets**

Applicable both for boosted **tops** and boosted **Higgs** searches

B-tagging the subjets is generally better than tagging the whole fat jet, both similar at high p_T



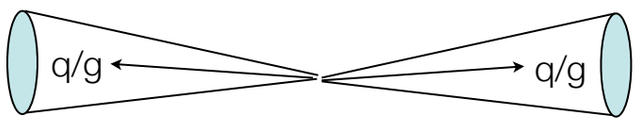



validation

validation

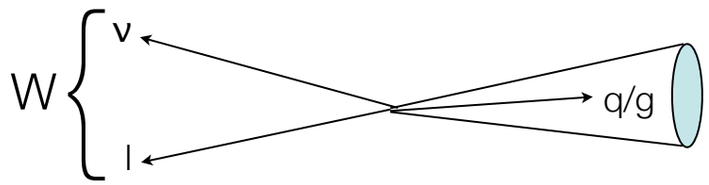
CMS PAS JME-13-006

dijets



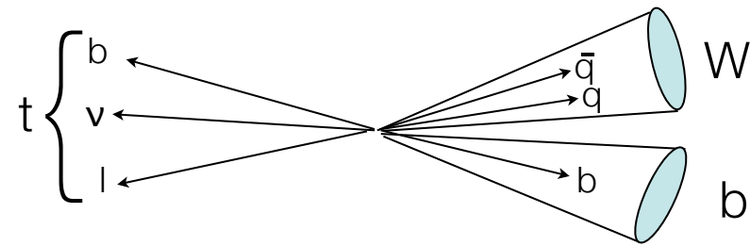
higher pT regime

W+jets

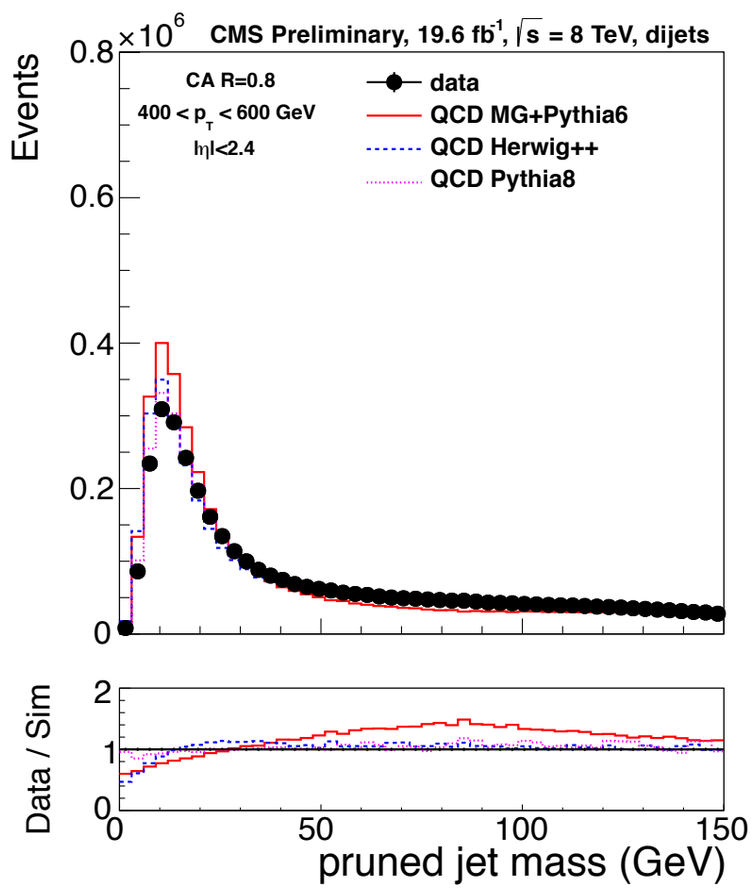


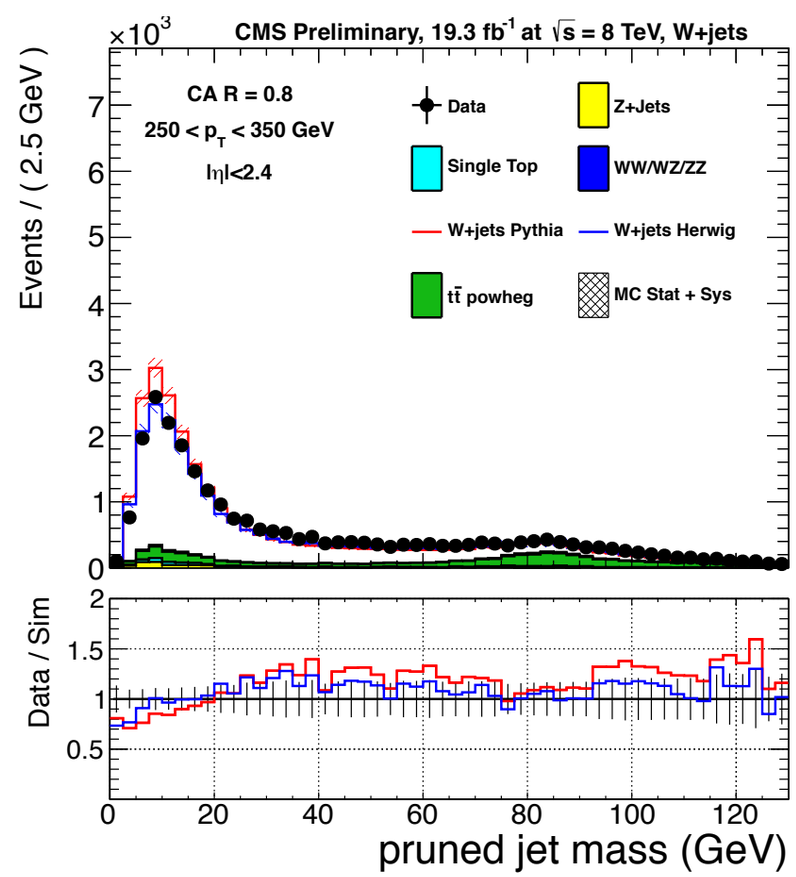
lower pT regime

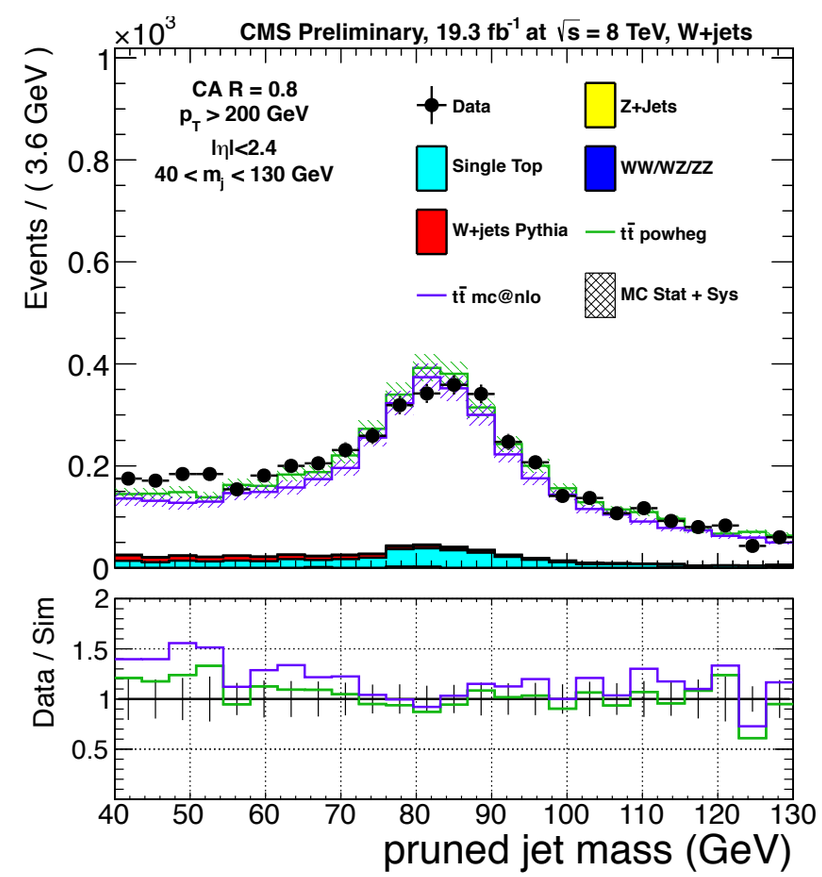
ttbar



contains real W's





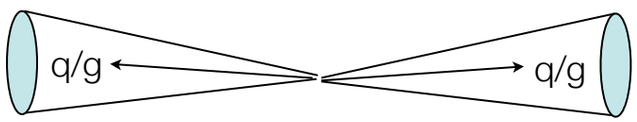


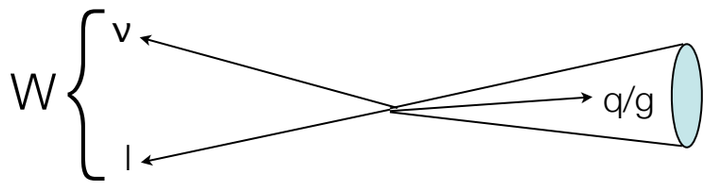
validation

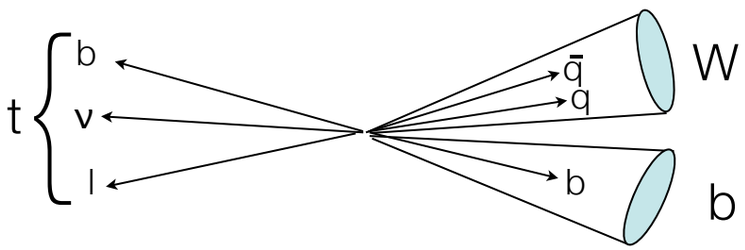
CMS PAS JME-13-006

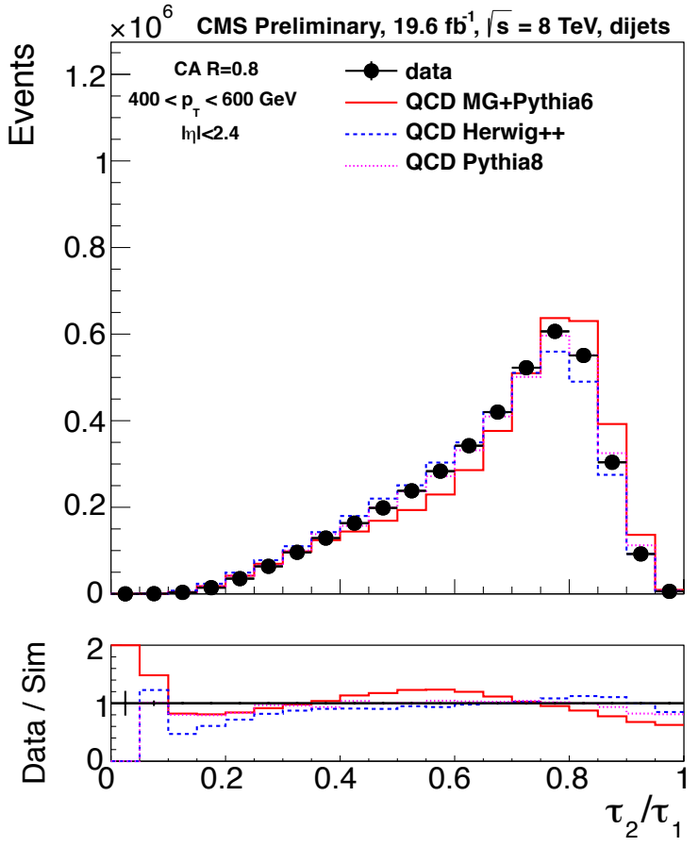
Good agreement between simulation and data

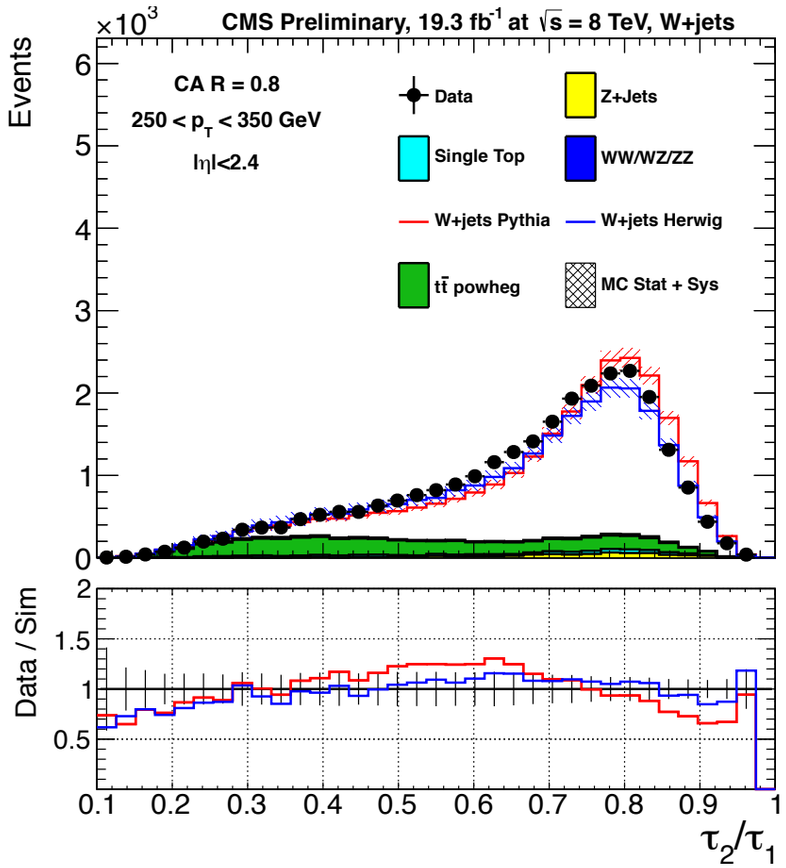
Pythia8 (4C) tends to do the best job describing data

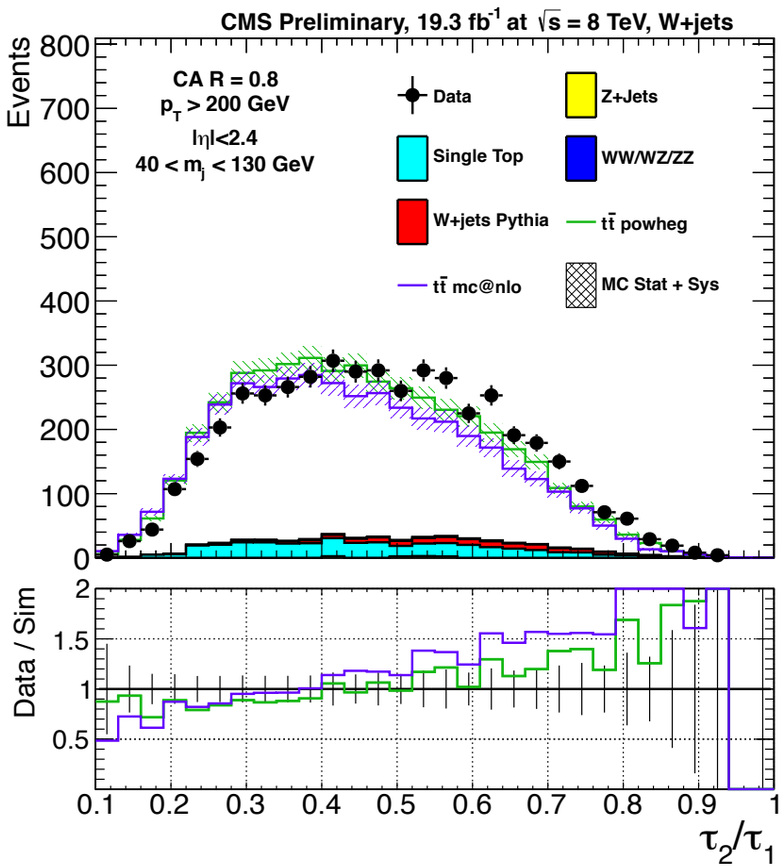




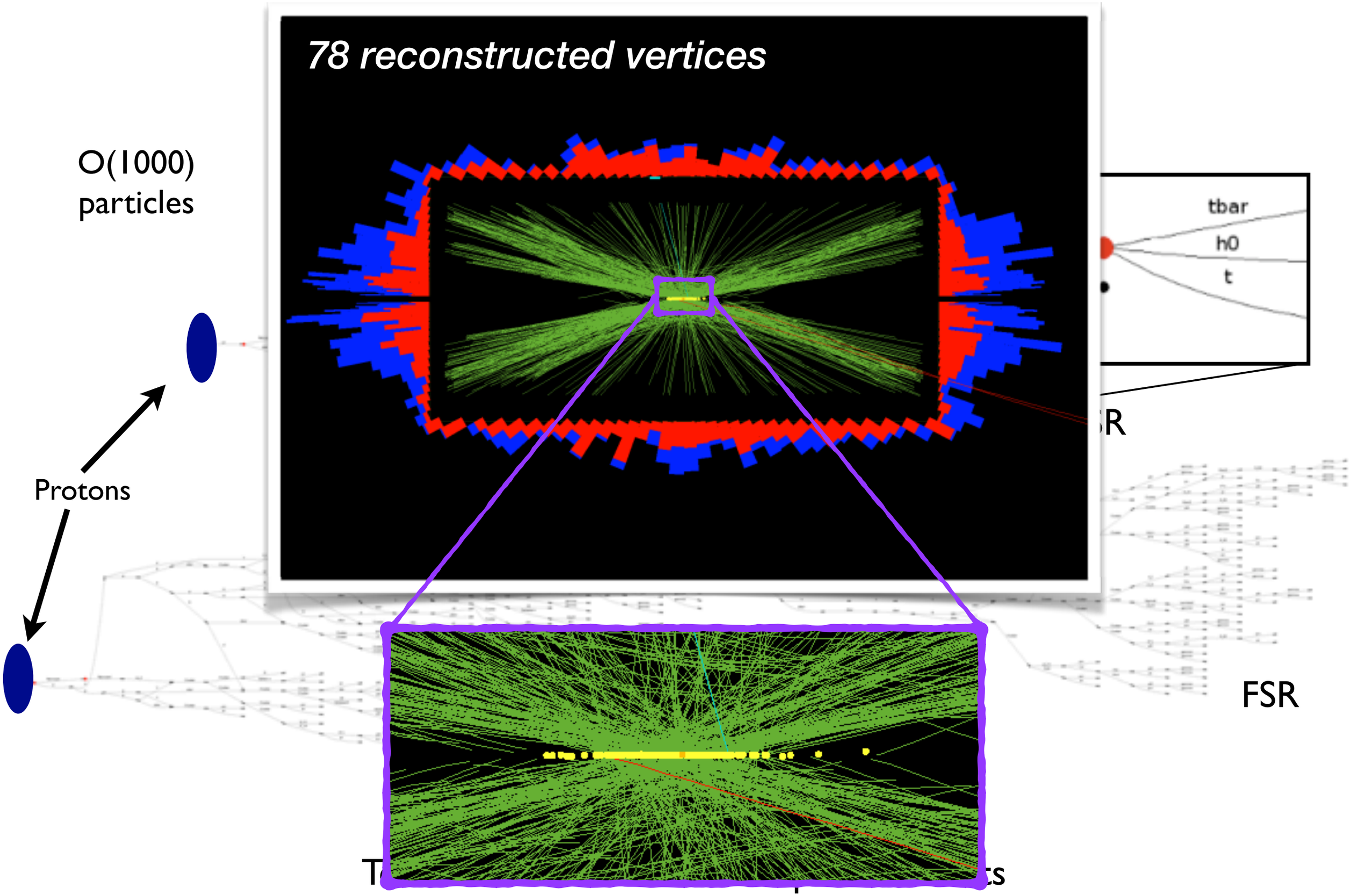








introduction



Searches for New Particles Using Cone and Cluster Jet Algorithms: A Comparative Study

Michael H. Seymour
Department of Theoretical Physics, University of Lund,
Sölvegatan 14A, S-22362 Lund, Sweden

Z.Phys. C62 (1994) 127-138

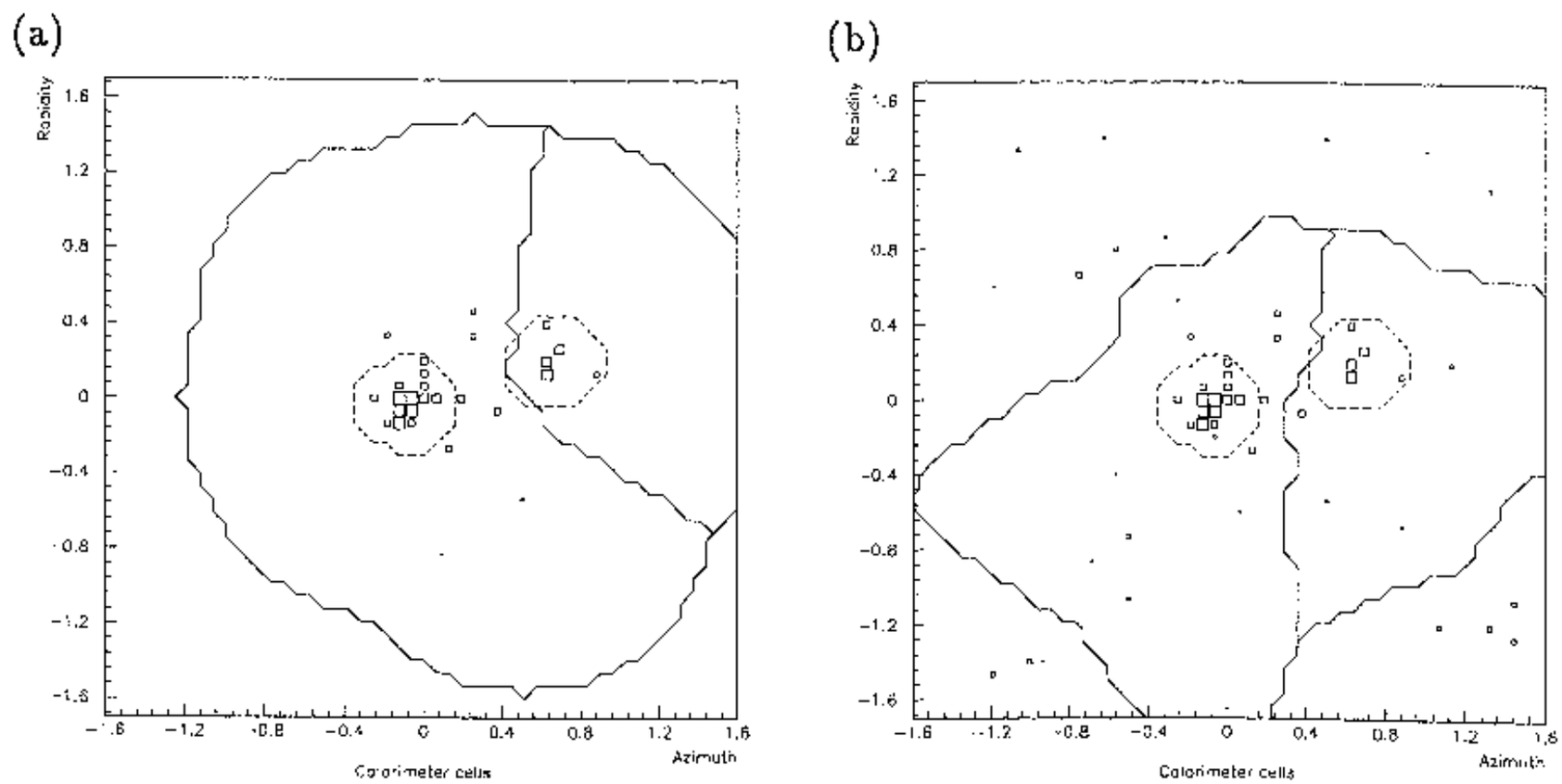


Figure 2: A hadronic W decay, as seen at calorimeter level, (a) without, and (b) with, particles from the underlying event. Box sizes are logarithmic in the cell energy, lines show the borders of the sub-jets for infinitely soft emission according to the cluster (solid) and cone (dashed) algorithms.

impossible becomes possible

Breakthrough paper on Hbb search at LHC

previously was thought to be impossible to do H→bb at the LHC

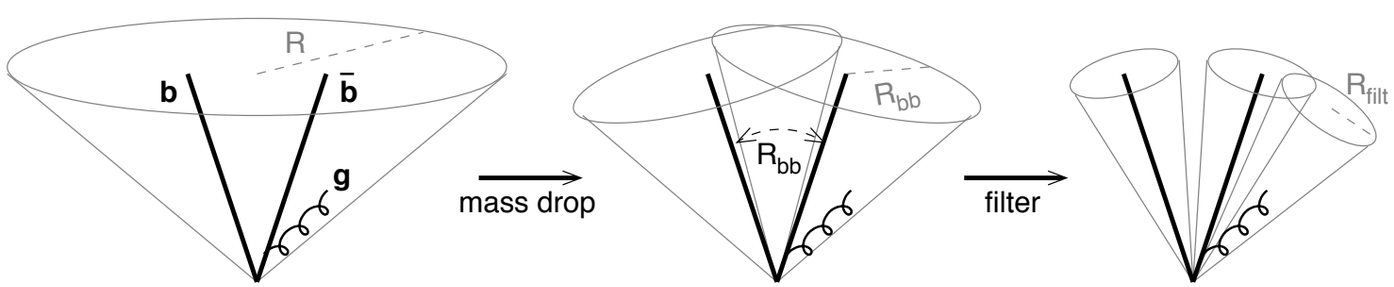
Go to boosted regime (high Higgs p_T), backgrounds are reduced and use substructure techniques, one fat R=1.2 jet, to identifying Higgs candidates

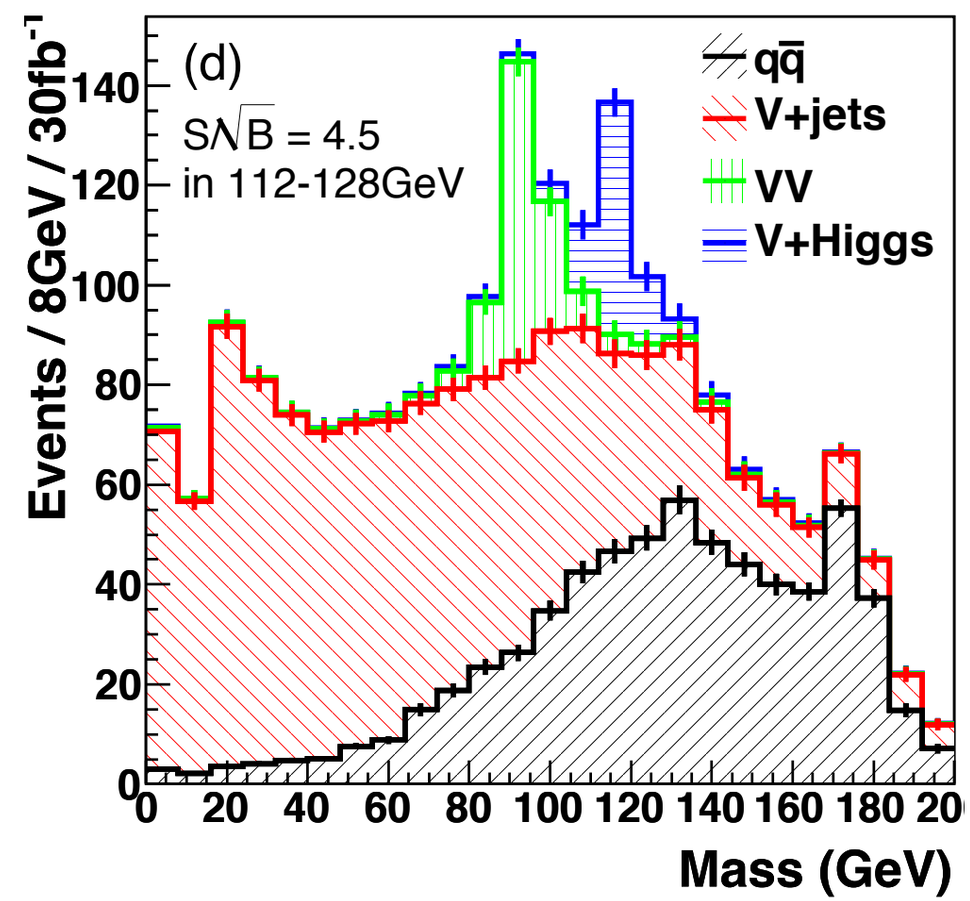
Phys.Rev.Lett. 100 (2008) 242001

Jet substructure as a new Higgs search channel at the LHC

Jonathan M. Butterworth, Adam R. Davison
Department of Physics & Astronomy, University College London.

Mathieu Rubin, Gavin P. Salam
LPTHE; UPMC Univ. Paris 6; Univ. Denis Diderot; CNRS UMR 7589; Paris, France.

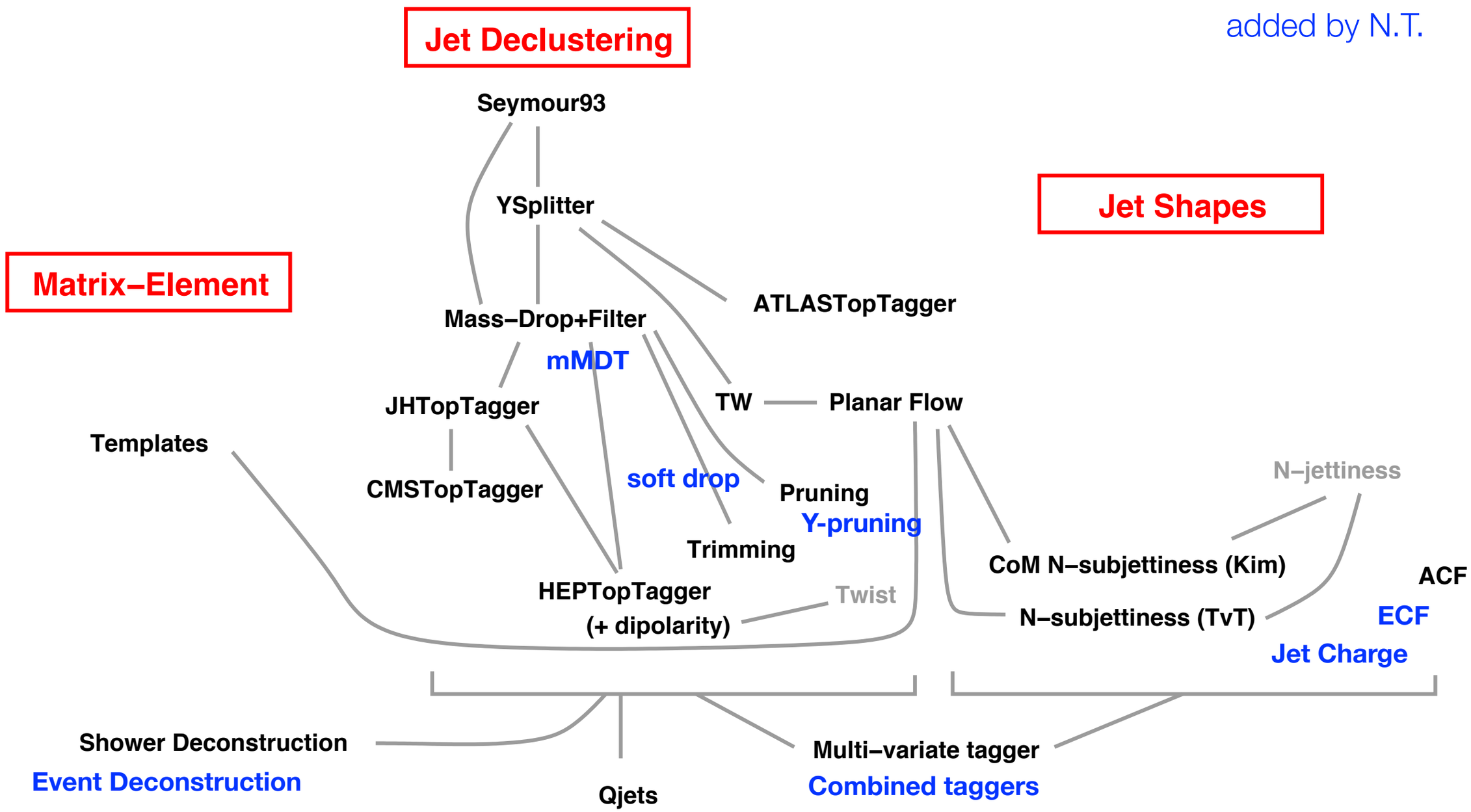




tools

Graphic from Gavin Salam circa 2012

added by N.T.

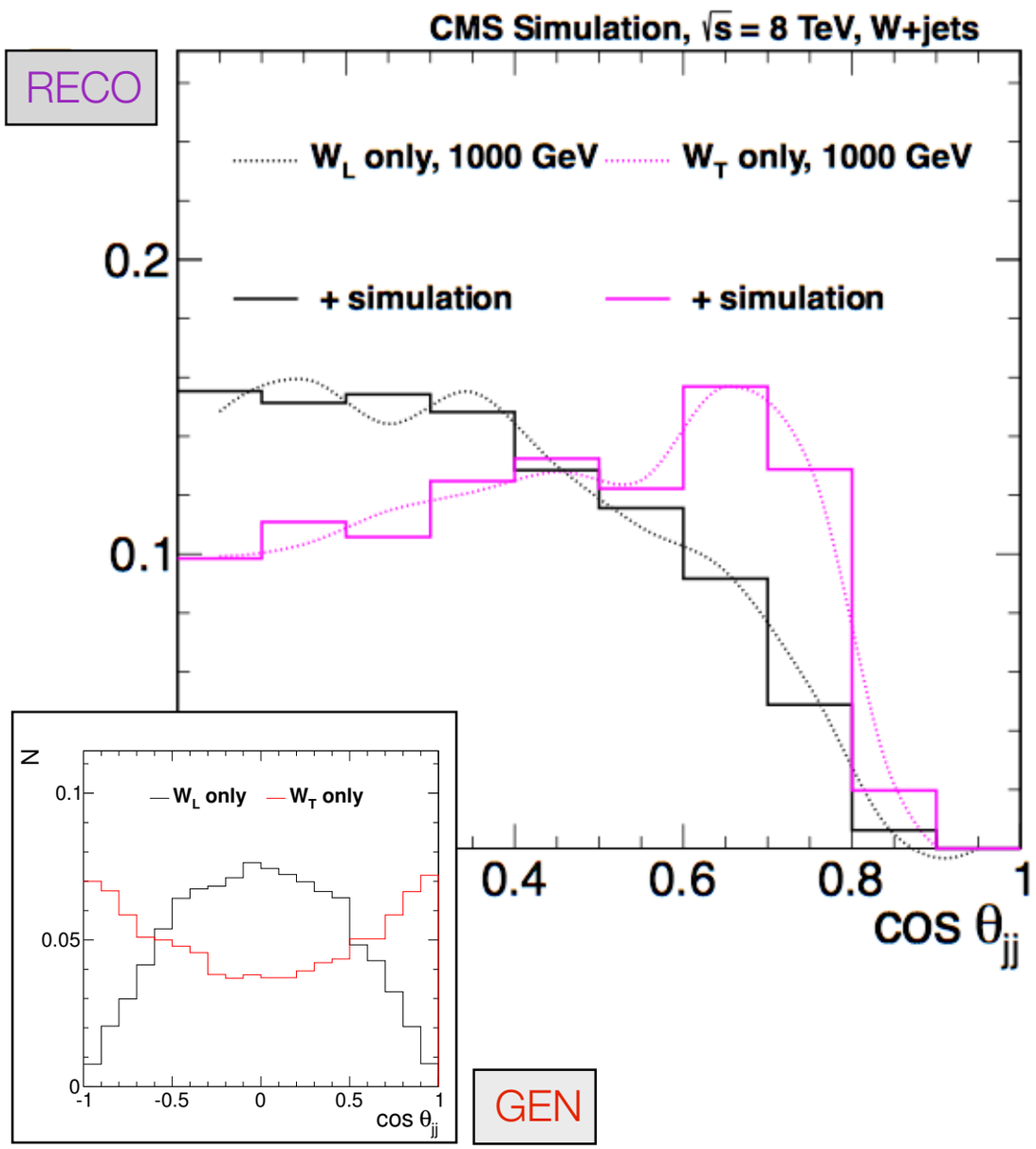


apologies for omitted taggers, arguable links, etc.

beyond bump hunting

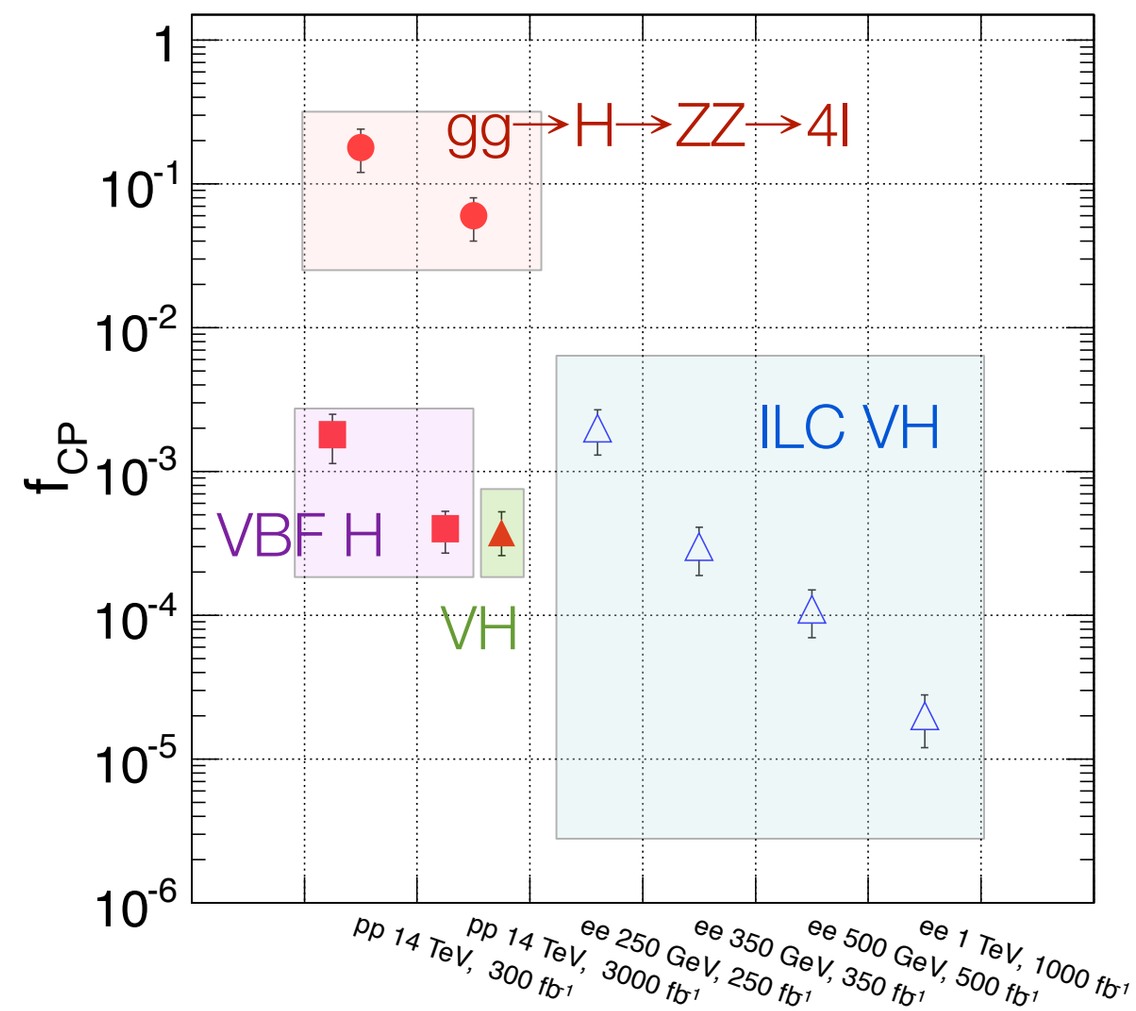
Angular analysis:

W polarization with subjet kinematics



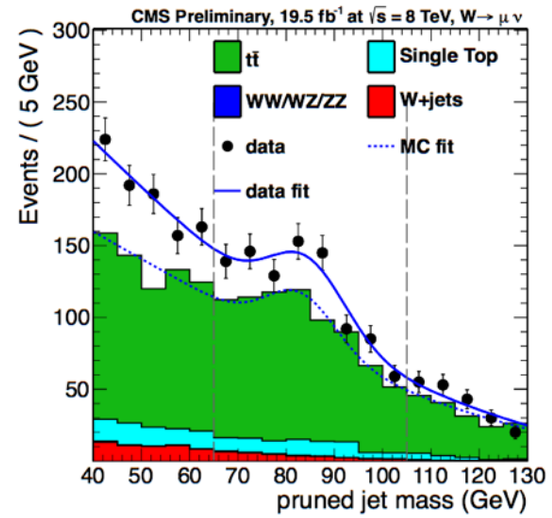
anomalous HVV couplings:

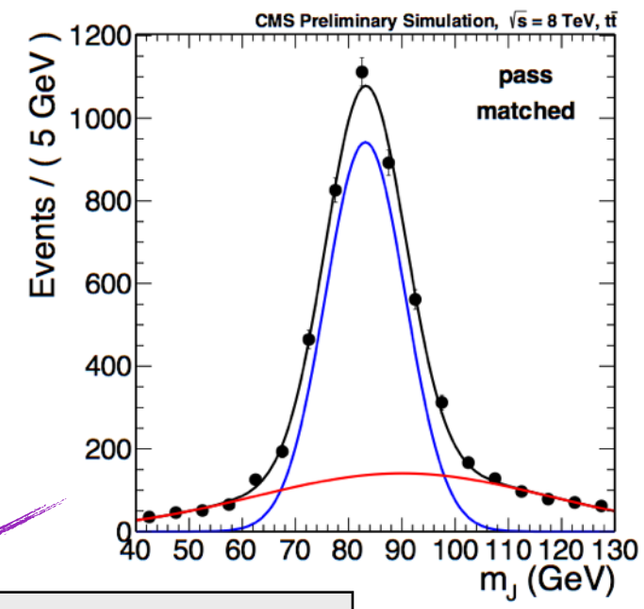
search for CP violation in Higgs sector potentially most sensitive in the boosted VH(bb) phase space

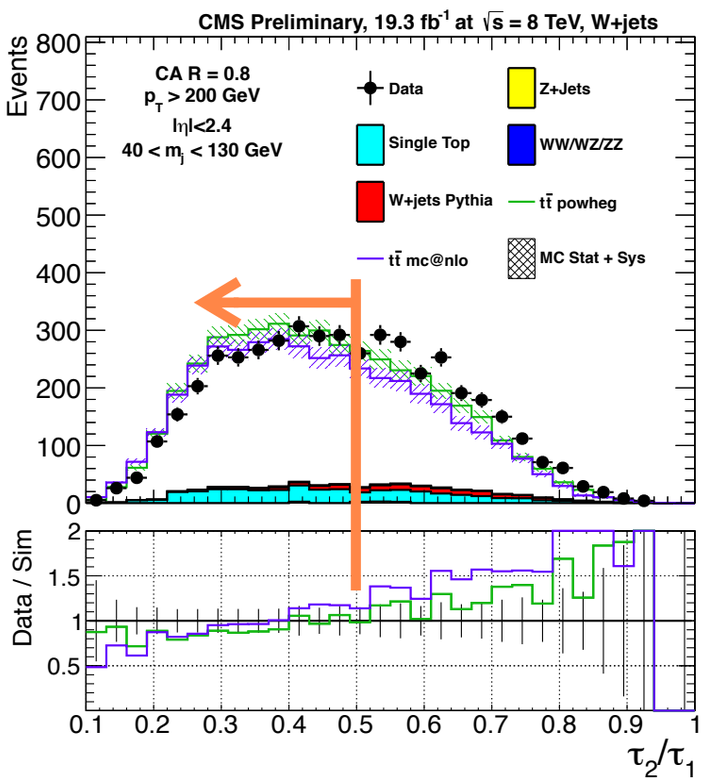


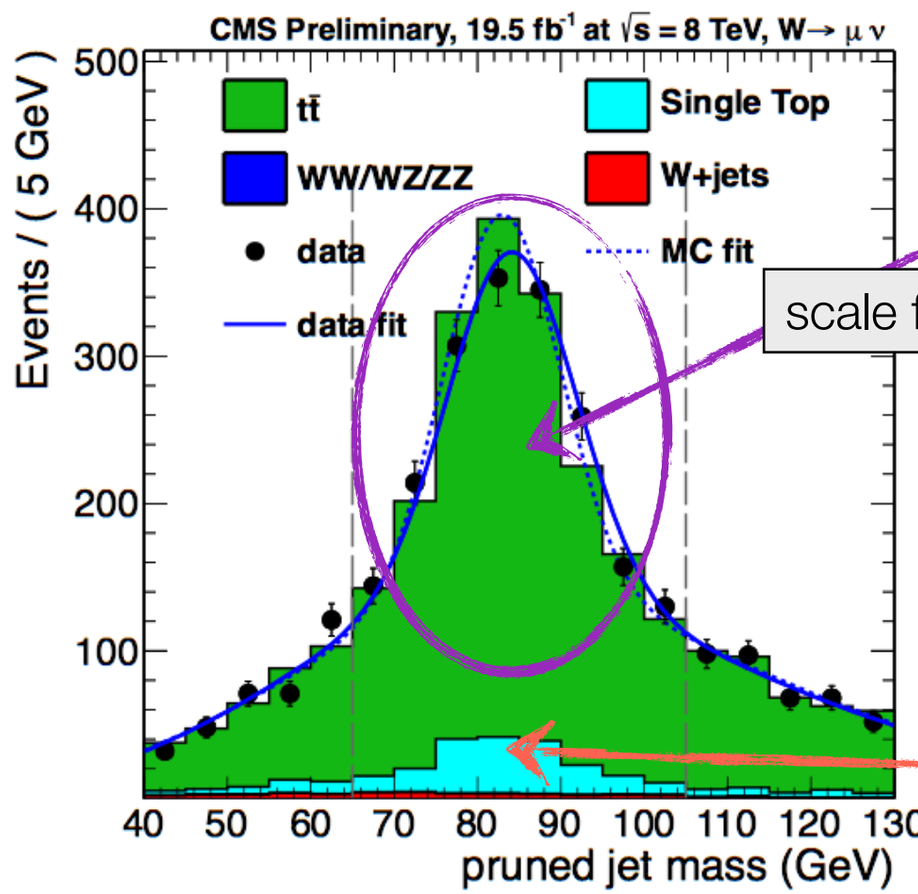
data-driven methods

Data-driven estimation of W-tagging scale factors via simultaneous fit to pass/fail samples

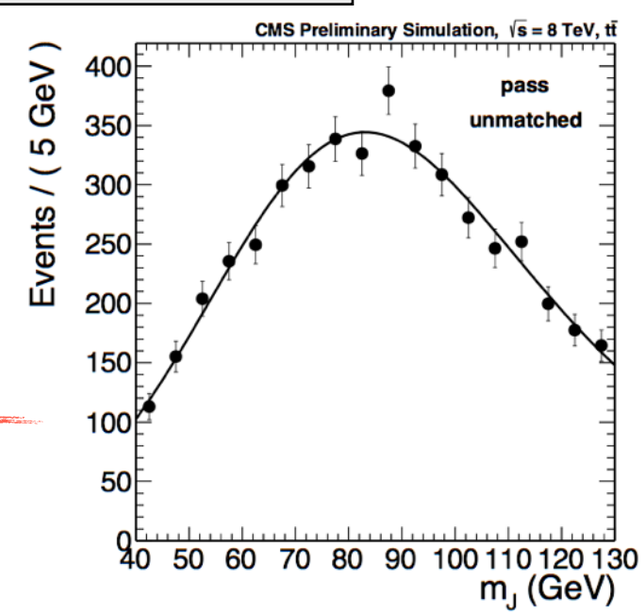








scale factor = 0.93 ± 0.08



mean (MC) = 83.2 ± 0.3 GeV
mean (data) = 84.4 ± 0.4 GeV
σ (MC) = 7.1 ± 0.4 GeV
σ (data) = 7.4 ± 0.6 GeV