Jets of a new generation

Grégory Soyez

IPhT, CEA Saclay

In collaboration with Gavin Salam and Matteo Cacciari

Mini workshop on ATLAS physics — CPPM, Marseille — Nov 29-30 2010

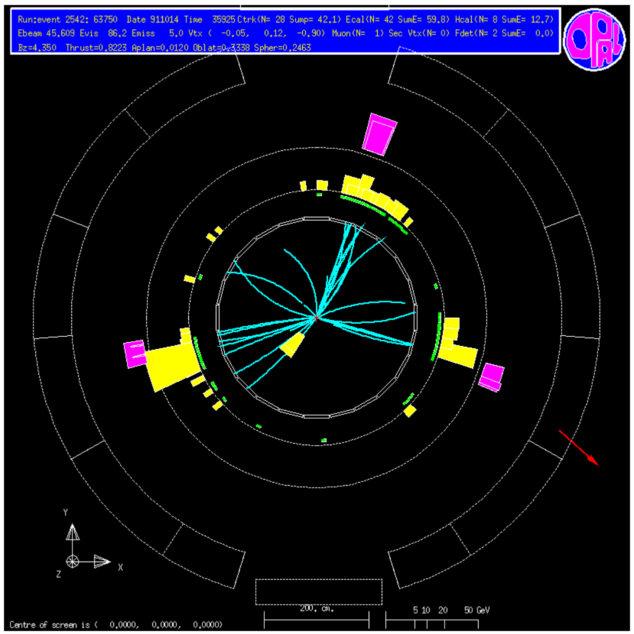
Plan

- Basic jet definitions
  - Motivation: the need for a jet definition
  - Situation today: meeting the 1990 requirements

- New directions
  - Subtracting pileup using jet areas
  - Jet substructure: UE filtering, boosted object tagging
  - Optimisation: kinematic dijet reconstruction

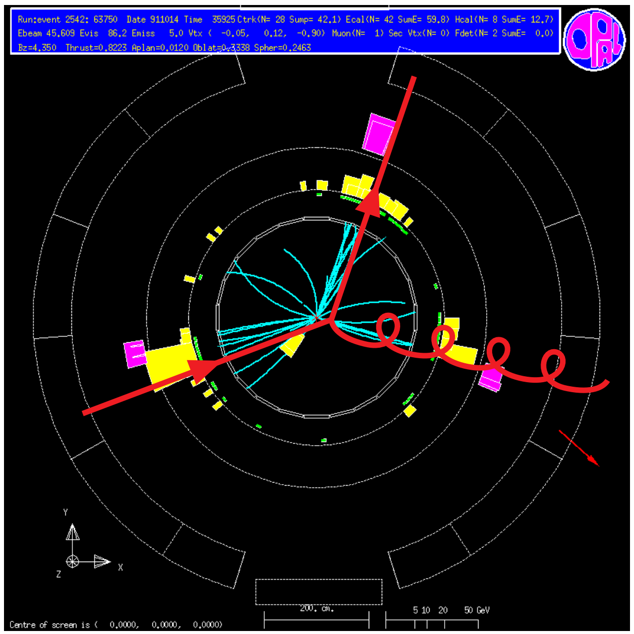
Jet definitions

“Jets” $\equiv$ bunch of collimated particles $\cong$ hard partons



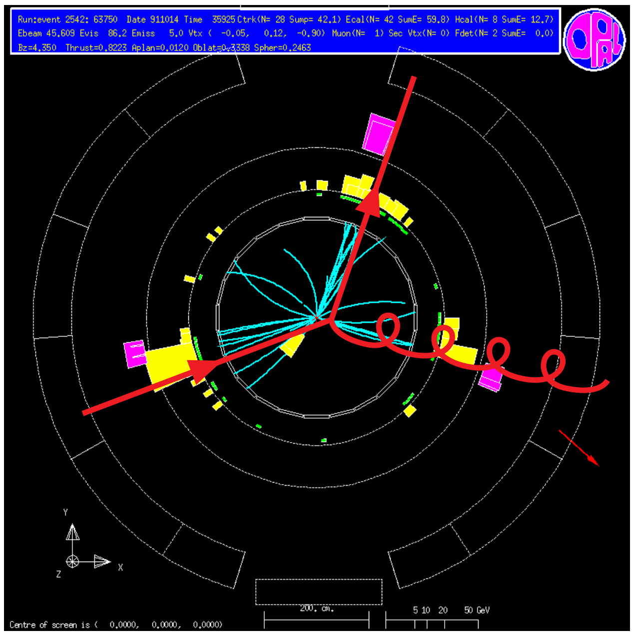
Jet definitions

“Jets” $\equiv$ bunch of collimated particles $\cong$ hard partons



Jet definitions

“Jets” $\equiv$ bunch of collimated particles $\cong$ hard partons



In practice: use a <u>jet definition</u>

particles $\{p_i\}$ $\xrightarrow[\textbf{definition}]{\textbf{jet}}$ jets $\{j_k\}$

algorithm: the recipe (insufficient!)

definition: algorithm + params

“Jet=hadron” too simplistic: What opening for “collimated”? NLO?

Jet algorithms: a big family

Recombination:

- k_t algorithm
- Cambridge/Aachen alg.
- anti-k_t algorithm

Cone:

- SISCone
- CDF JetClu
- CDF MidPoint
- D0 (run II) Cone
- PxCone
- ATLAS Cone
- CMS Iterative Cone
- PyCell/CellJet
- GetJet

Jet definitions: constraints

SNOWMASS accords (FermiLab, 1990)

Several important properties that should be met by a jet definition are [3]:

1. Simple to implement in an experimental analysis;
2. Simple to implement in the theoretical calculation;
3. Defined at any order of perturbation theory;
4. Yields finite cross section at any order of perturbation theory;
5. Yields a cross section that is relatively insensitive to hadronization.

Jet definitions: constraints

SNOWMASS accords (FermiLab, 1990)

Several important properties that should be met by a jet definition are [3]:

1. Simple to implement in an experimental analysis;
2. Simple to implement in the theoretical calculation;
3. Defined at any order of perturbation theory;
4. Yields finite cross section at any order of perturbation theory;
5. Yields a cross section that is relatively insensitive to hadronization.

20 years later, these are only recently satisfied!!!

Jet algorithms: the situation now

Recombination:

- k_t algorithm
- Cambridge/Aachen alg.
- anti-k_t algorithm

Cone:

- SISCone
- CDF JetClu
- CDF MidPoint
- D0 (run II) Cone
- PxCone
- ATLAS Cone
- CMS Iterative Cone
- PyCell/CellJet
- GetJet

Jet algorithms: the situation now

Recombination:

- k_t algorithm
- Cambridge/Aachen alg.
- anti-k_t algorithm

Implementation:
FastJet
www.fastjet.fr

Cone:

- SISCone
- CDF JetClu
- CDF MidPoint
- D0 (run II) Cone
- PxCone
- ATLAS Cone
- CMS Iterative Cone
- PyCell/CellJet
- GetJet

Inside jet definitions

- Recombination: successively recombine the closest pair

$$d_{ij} = \min(k_{t,i}^{2p}, k_{t,j}^{2p})(\Delta y_{ij}^2 + \Delta\phi_{ij}^2)$$

  Stop at distance R

  - $p = 1$: k_t algorithm (very close to QCD)

    [Catani, Dokshitzer, Seymour, Webber, 93]

  - $p = 0$: Cambridge/Aachen (C/A) algorithm (substructure studies)

    [Dokshitzer, Leder, Moretti, Webber, 93]

  - $p = -1$: anti-k_t algorithm (circular/rigid jets)

    [Cacciari, Salam, GS, 08]

- Cone: $\approx$ flow of energy in a cone (of fixed R) centred on the cone centre: SISCone

  [Salam, GS, 07]

Finite perturbative cross-section: only consider infrared-and-collinear-safe algorithms

Inside jet definitions

- Recombination: successively recombine the closest pair

$$d_{ij} = \min(k_{t,i}^{2p}, k_{t,j}^{2p})(\Delta y_{ij}^2 + \Delta\phi_{ij}^2)$$

Stop at distance R

- $p = 1$: k_t algorithm (very close to QCD)

  [Catani, Dokshitzer, Seymour, Webber, 93]

- $p = 0$: Cambridge/Aachen (C/A) algorithm (substructure studies)

  [Dokshitzer, Leder, Moretti, Webber, 93]

- $p = -1$: anti-k_t algorithm (circular/rigid jets)

  the default at the LHC [Cacciari, Salam, GS, 08]

- Cone: $\approx$ flow of energy in a cone (of fixed R) centred on the cone centre: SISCone [Salam, GS, 07]

Finite perturbative cross-section: only consider infrared-and-collinear-safe algorithms

New generations

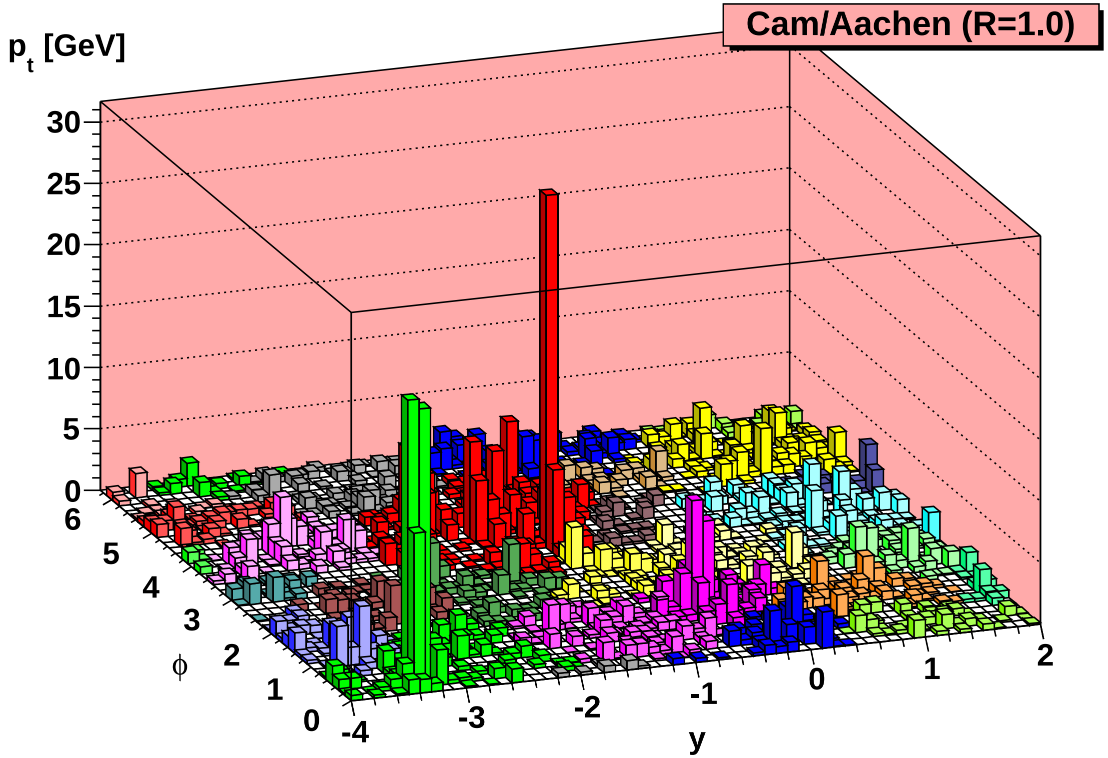
Challenges

Idea: set of solid algorithms: optimise their usage

- Handle pileup
- Tag boosted objects/Study substructure
- clean UE contamination to jets
- optimize reconstruction

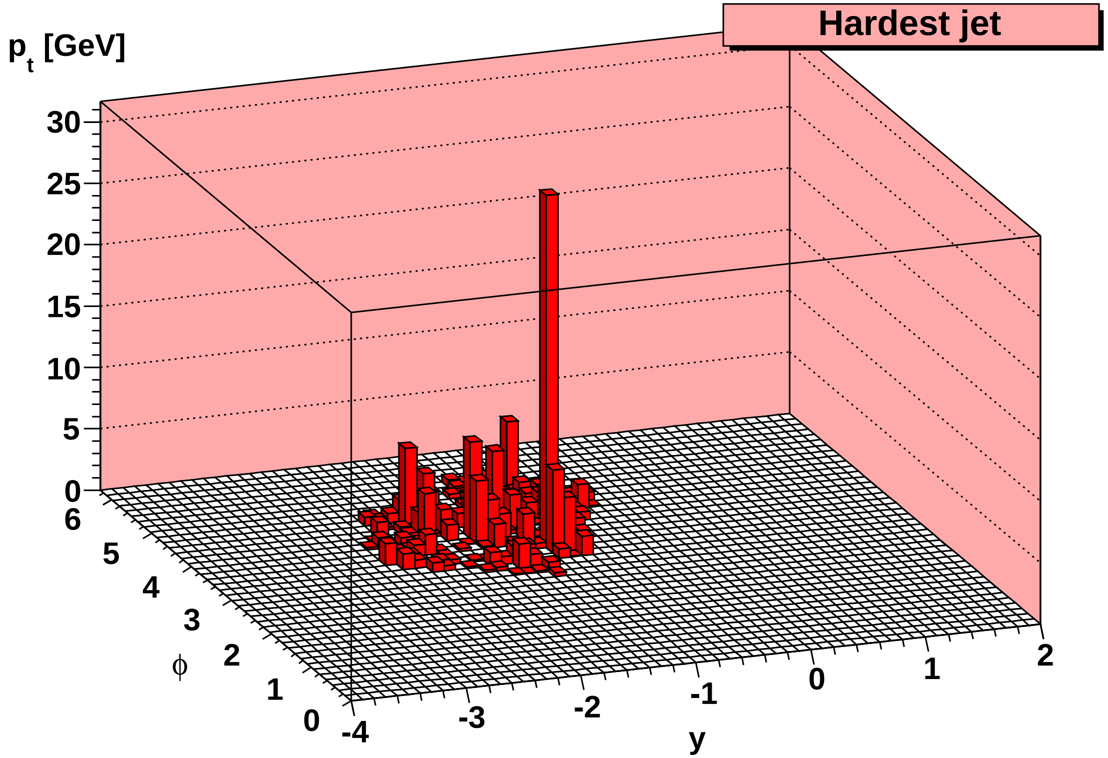
New generations
1. Jet substructure

Filtering



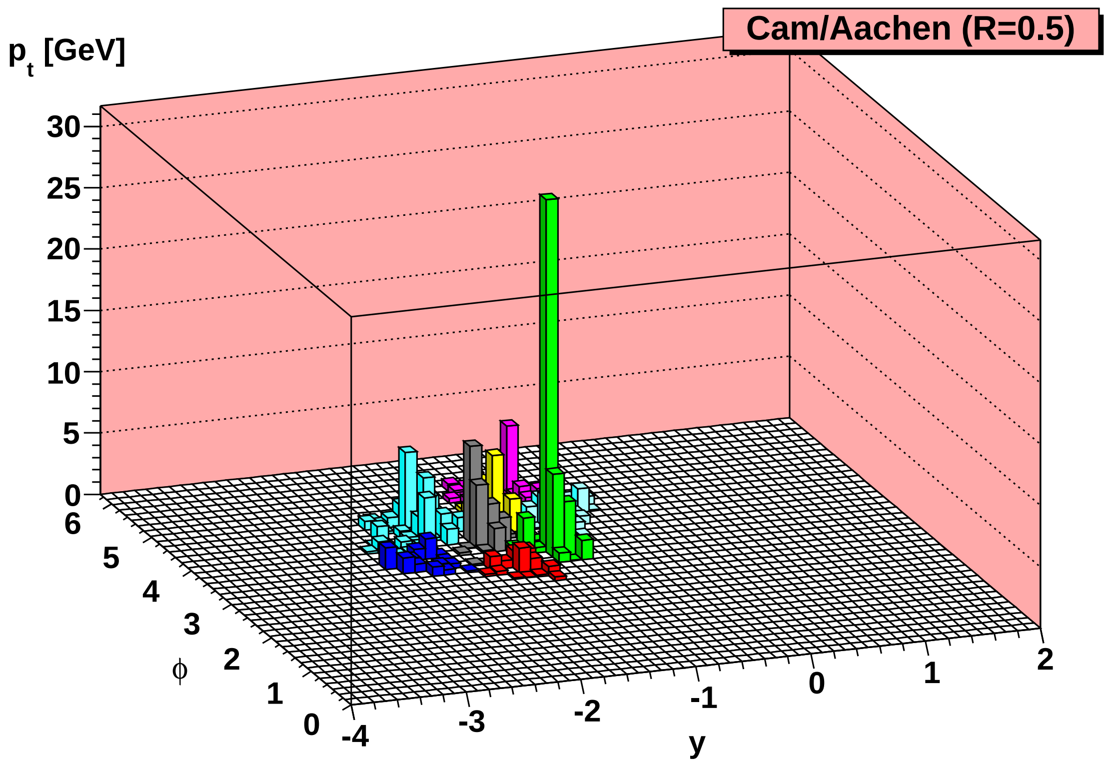
- cluster with Cambridge/Aachen(R)

Filtering



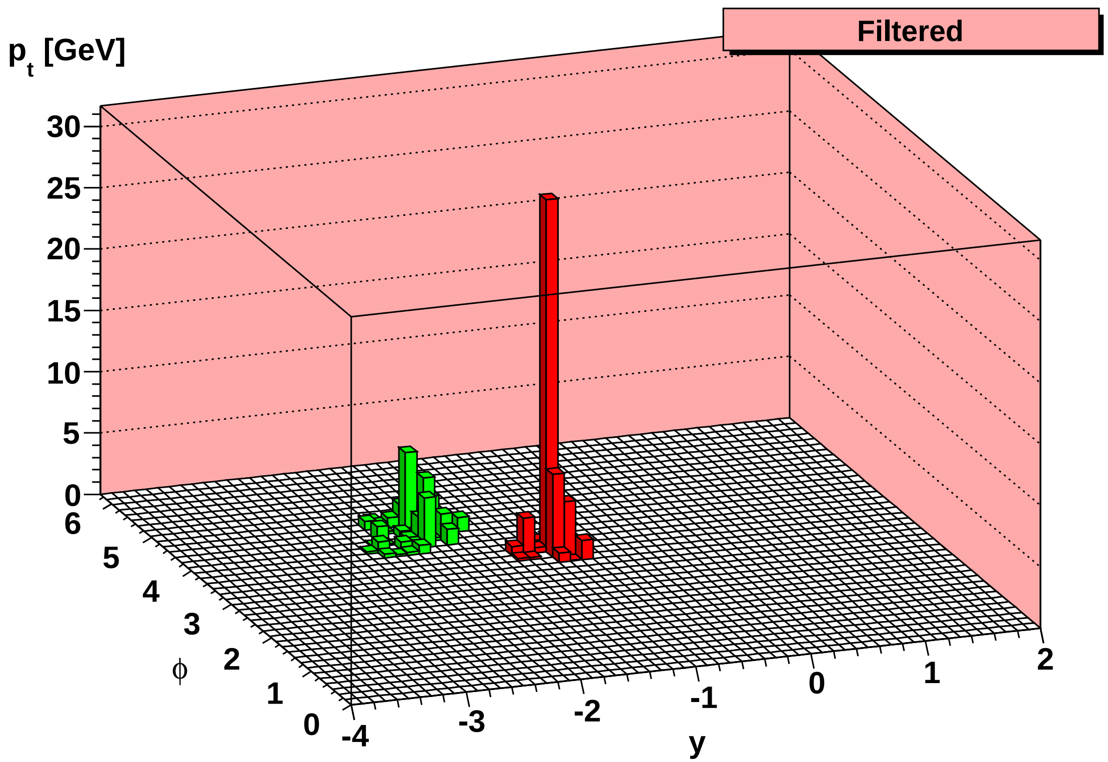
- cluster with Cambridge/Aachen(R)
- for each jet

Filtering



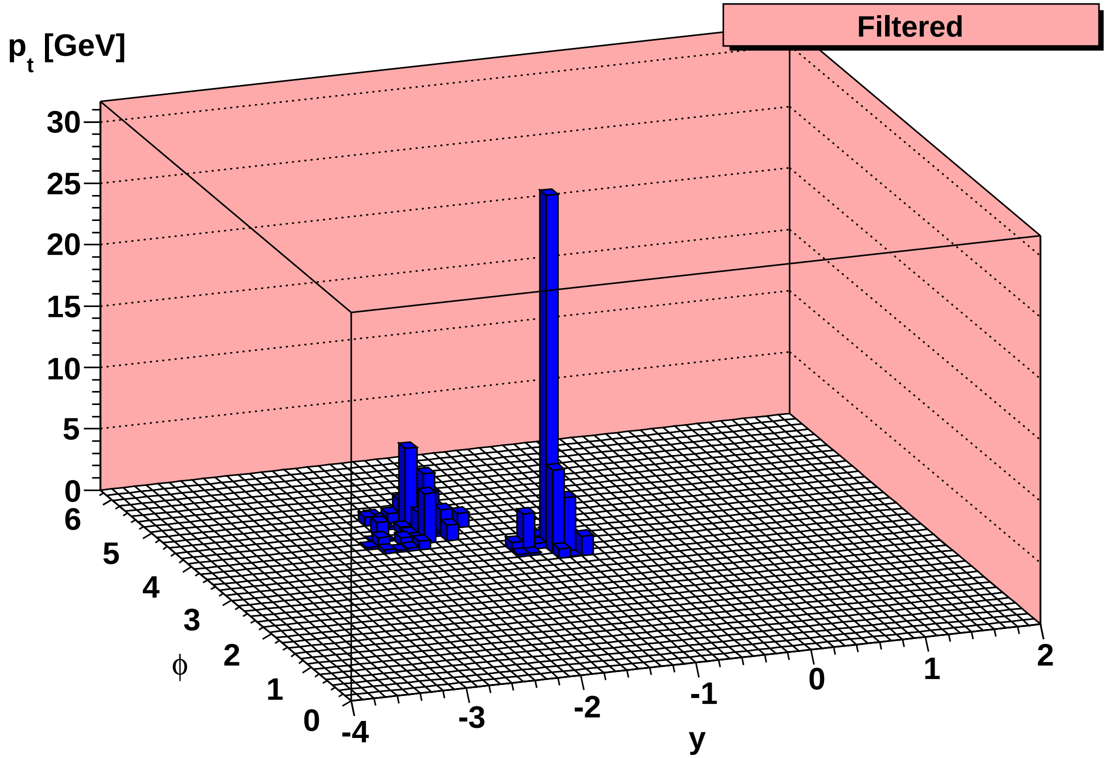
- cluster with Cambridge/Aachen(R)
- for each jet
  - recluster with Cambridge/Aachen(R/2)

Filtering



- cluster with Cambridge/Aachen(R)
- for each jet
  - recluster with Cambridge/Aachen(R/2)
  - keep the 2 hardest subjets

Filtering



- cluster with Cambridge/Aachen(R)
- for each jet
  - recluster with Cambridge/Aachen(R/2)
  - keep the 2 hardest subjets

Idea:

✓ keep perturb. radiation

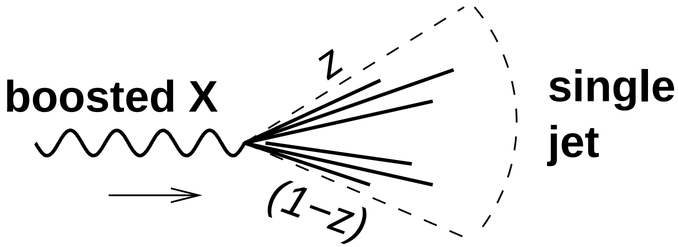
✓ remove UE

- Proven useful for boosted jet $H \to b\bar{b}$ tagging
  [J.Butterworth, A.Davison, M.Rubin, G.Salam, 08]
- Proven useful for kinematic reconstructions
  [M.Cacciari, J.Rojo, G.Salam, GS, 08]
- Similar: trimming [D.Krohn,J.Thaler,L-T.Wang,10]

Boosted object tagging

Problem:

boosted heavy object $\Rightarrow$ decays reconstructed in a single jet



$$R \gtrsim \frac{m}{p_t}\frac{1}{\sqrt{z(1-z)}}$$

How to disantangle that from a QCD jet?

Idea: substructure *e.g.* look inside the jet

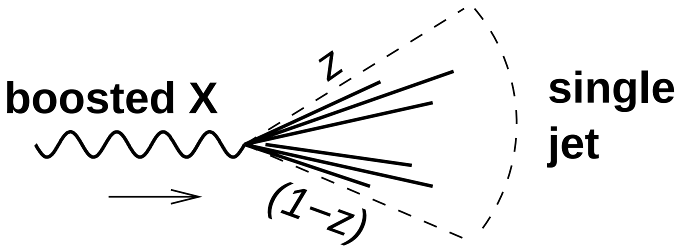
Various methods: mass drop, pruning, use Jade distances, asymmetry cuts,...

Applications: (examples)

- 2 decay products: $W \to q\bar{q}$, $H \to b\bar{b}$
- 3 decay products: $t \to qqb$, $\tilde{\chi} \to qqq$
- busier: $t\bar{t}H$

Example: boosted Higgs

[J.Butterworth, A.Davison, M.Rubin, G.Salam,08]



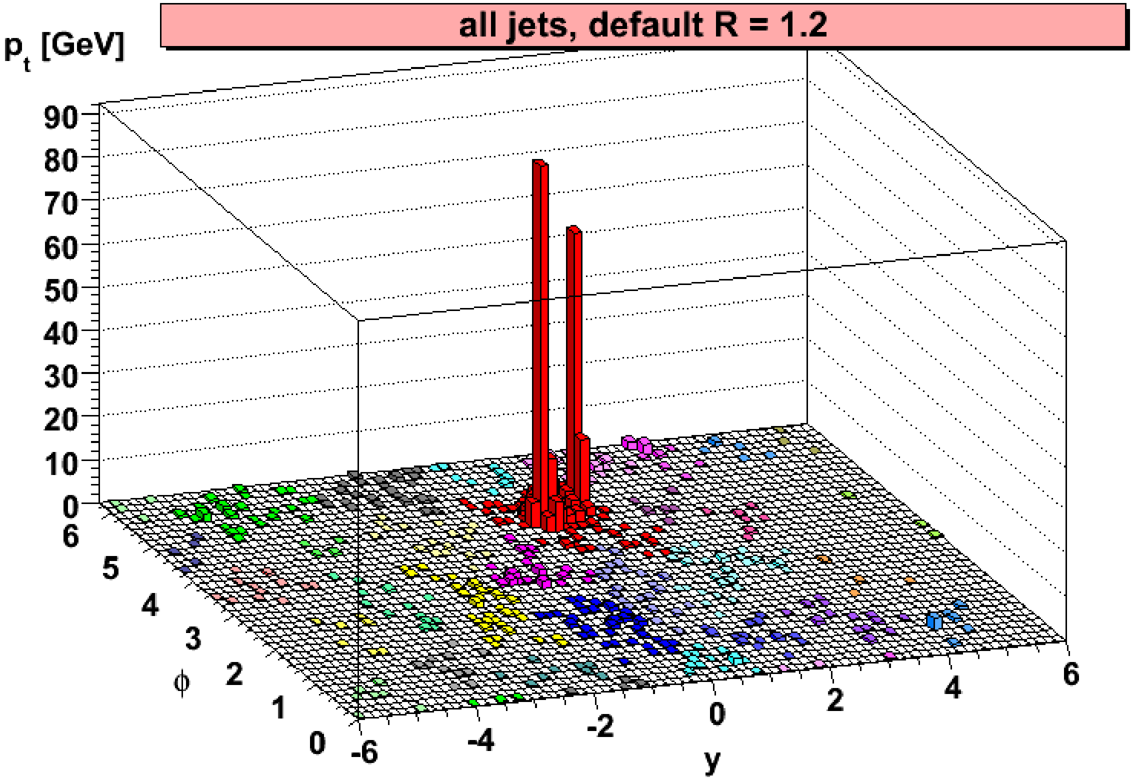
$$R \gtrsim \frac{m}{p_t} \frac{1}{\sqrt{z(1-z)}}$$

Method: start with a hard (C/A, radius R) jet j

1. Undo the last clustering $\rightarrow j_1, j_2$
2. If $\max(m_1, m_2) < 0.67m$, we have a mass drop, else back to 1
   idea: find the 2 b-jets, dynamically find R_{bb}
3. Require symmetric splitting $y_{12} \approx \frac{\min(z_1,z_2)}{\max(z_1,z_2)} > 0.09$, else go to 1
   idea: remove QDC asymmetric splittings
4. Require 2 b taggings
5. Filter *i.e.* uncluster down to R_{filt}, keep the 3 hardest subjets
   idea: keep “hard” QCD radiations, reduce UE

Boosted Higgs: one event, effects on S/B

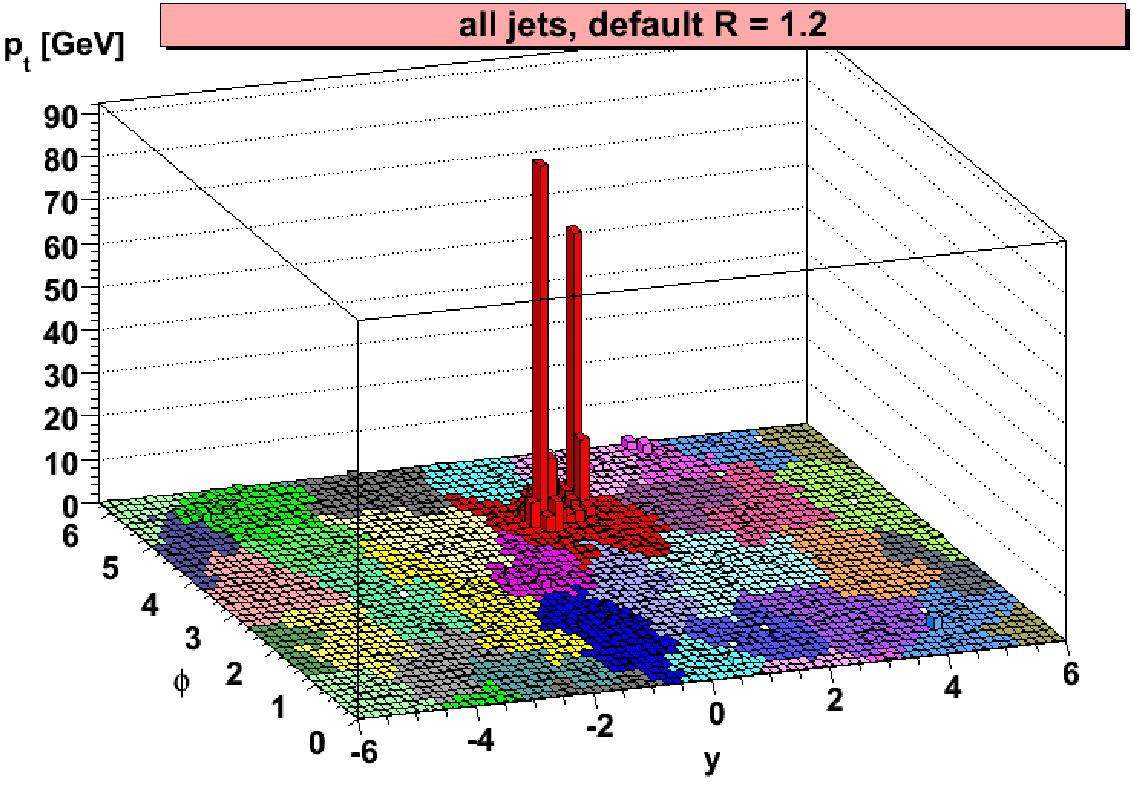
Herwig 6.510 + Jimmy 4.31 + FastJet 2.3



Cluster C/A, R=1.2

Boosted Higgs: one event, effects on S/B

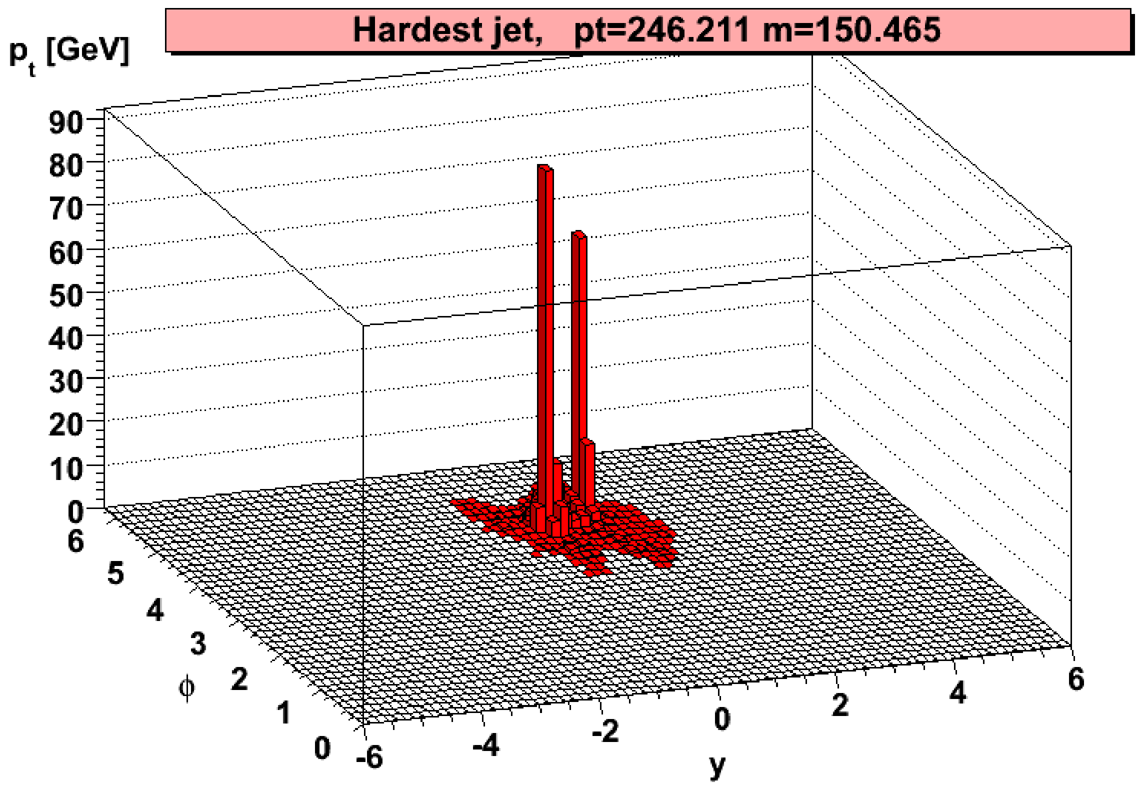
Herwig 6.510 + Jimmy 4.31 + FastJet 2.3



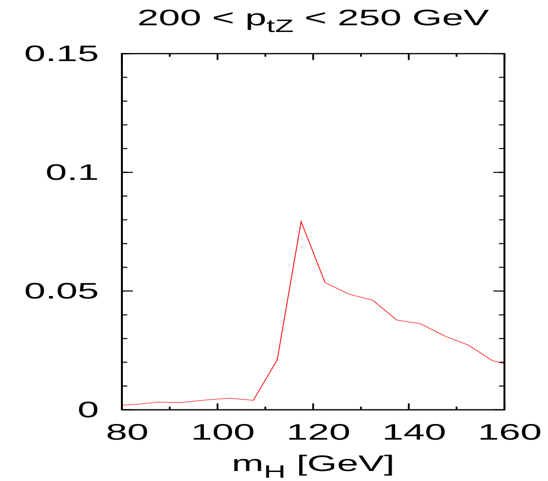
Show jets more clearly

Boosted Higgs: one event, effects on S/B

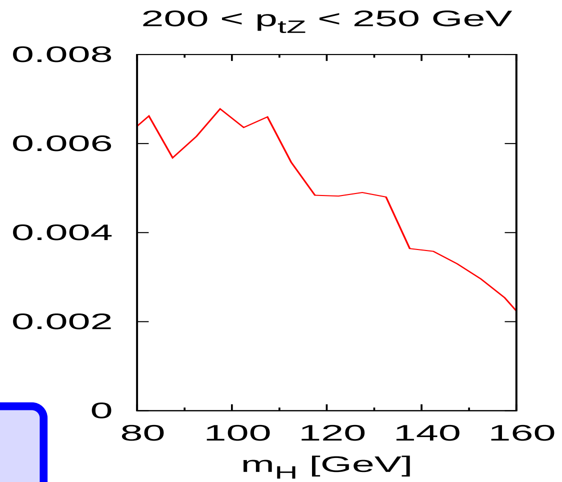
Herwig 6.510 + Jimmy 4.31 + FastJet 2.3



HZ Signal



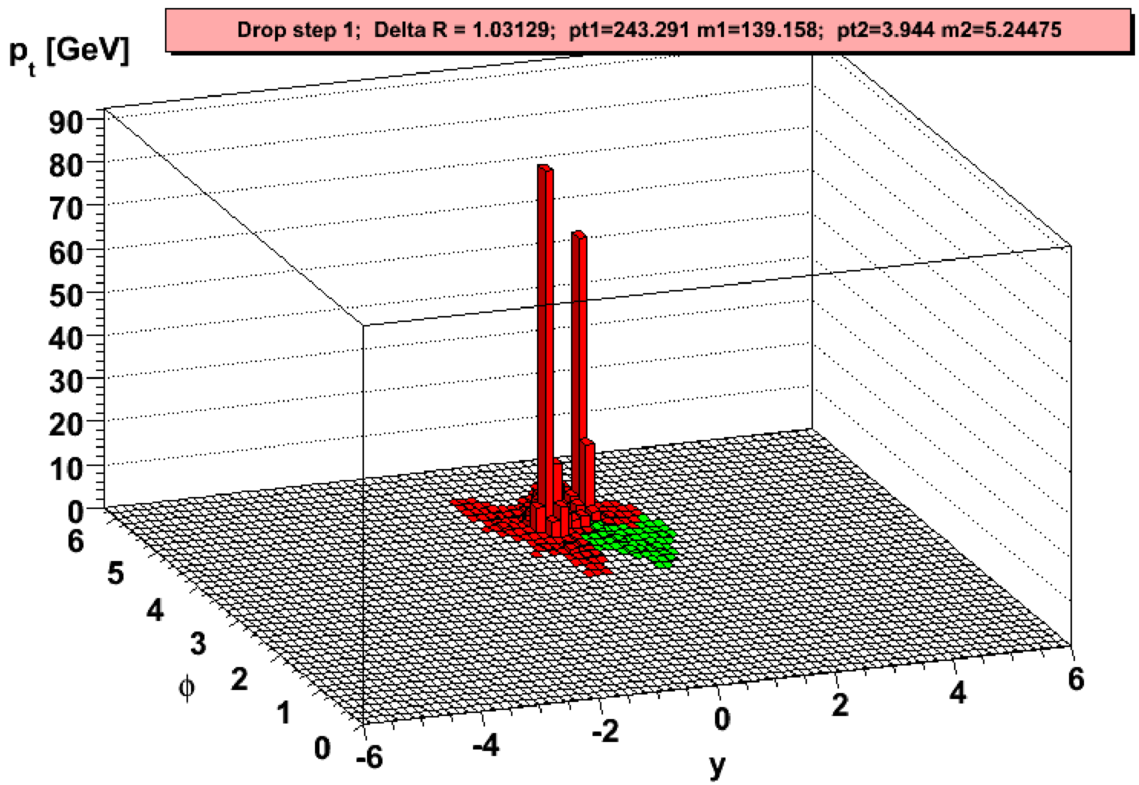
Zbb Background



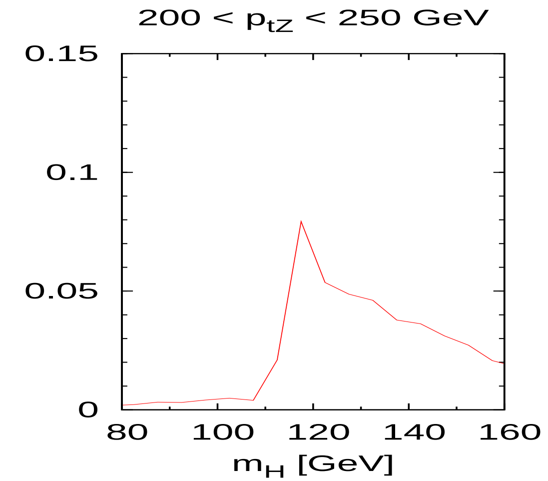
Hardest jet ($m = 150$ GeV)

Boosted Higgs: one event, effects on S/B

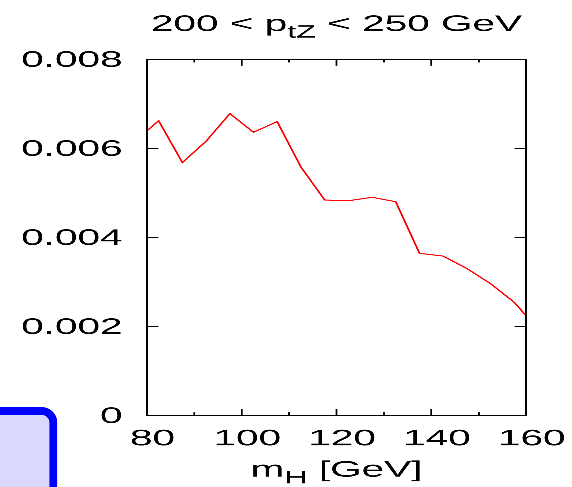
Herwig 6.510 + Jimmy 4.31 + FastJet 2.3



HZ Signal



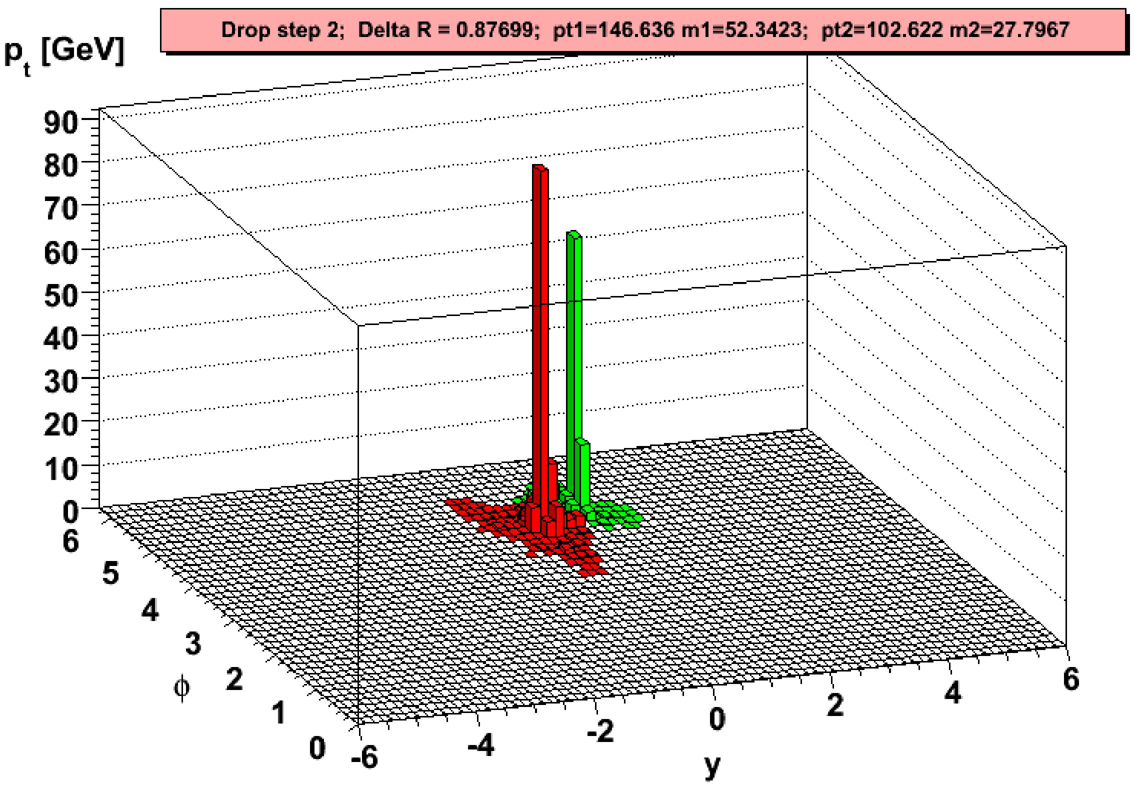
Zbb Background



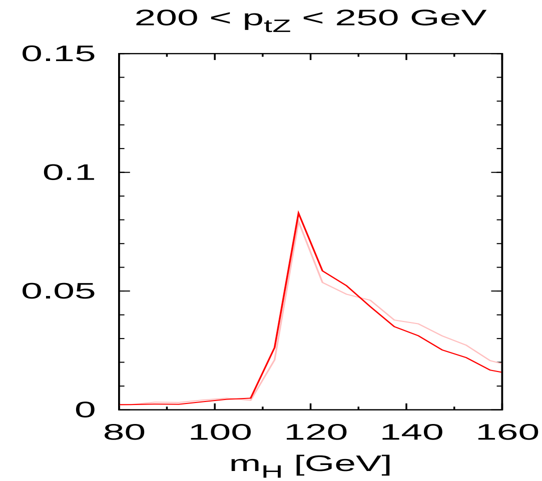
Split: $\frac{\max(m_1, m_2)}{m} = 0.92$, repeat ($m = 150$ GeV)

Boosted Higgs: one event, effects on S/B

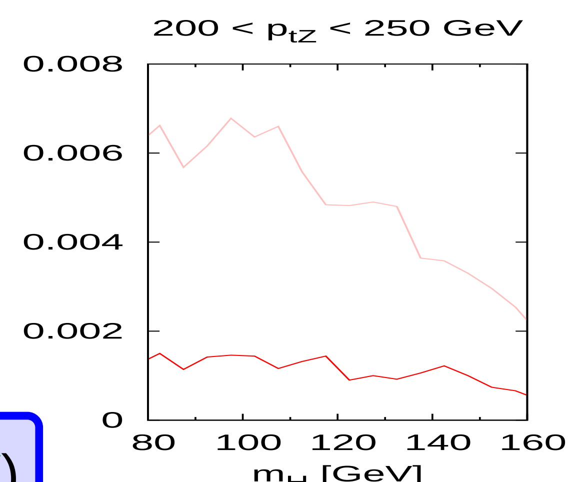
Herwig 6.510 + Jimmy 4.31 + FastJet 2.3



HZ Signal



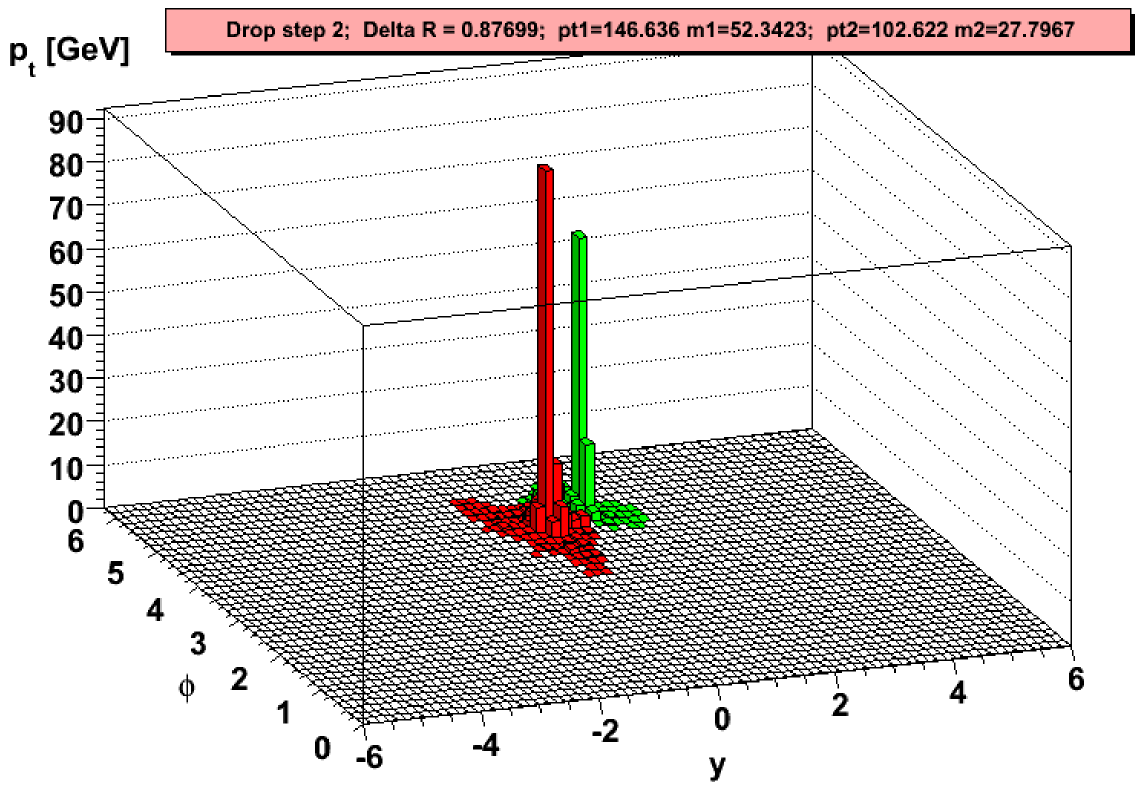
Zbb Background



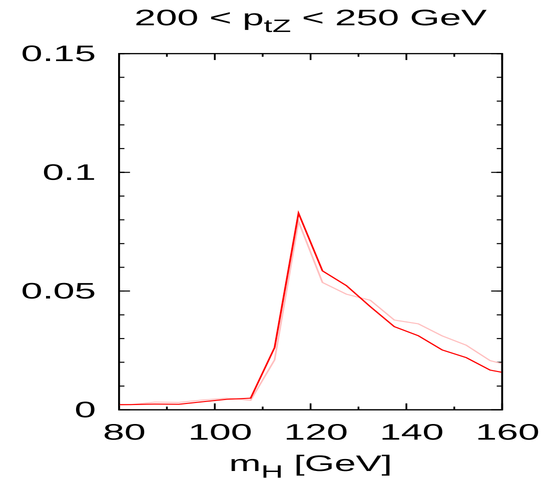
Split: $\frac{\max(m_1, m_2)}{m} = 0.37$, mass drop ($m = 139$ GeV)

Boosted Higgs: one event, effects on S/B

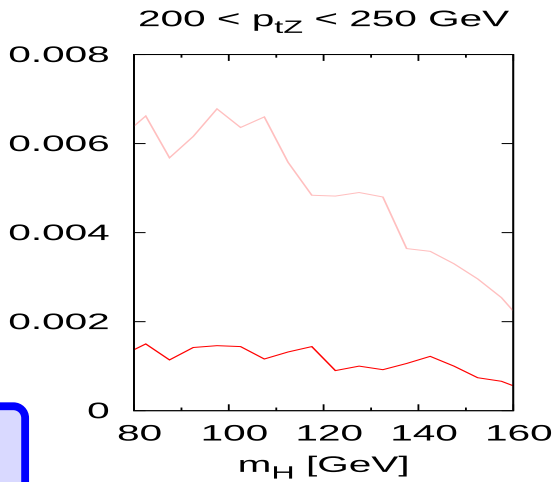
Herwig 6.510 + Jimmy 4.31 + FastJet 2.3



HZ Signal



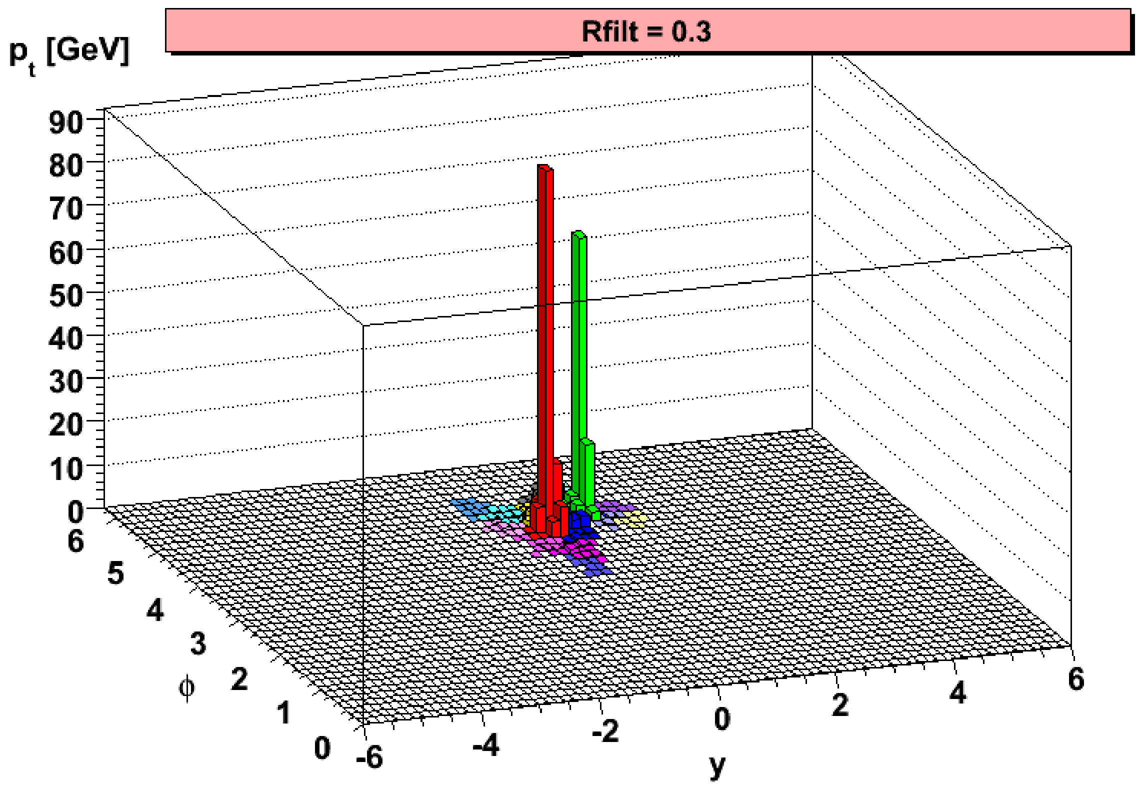
Zbb Background



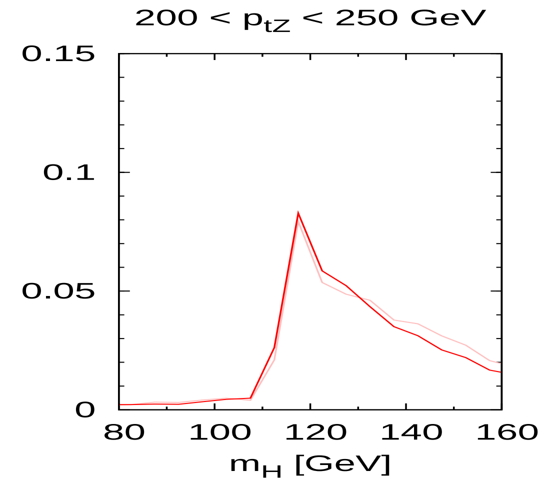
Split: $y_{12} = 0.7$, 2 b tags $\Rightarrow$ OK ($m = 139$ GeV)

Boosted Higgs: one event, effects on S/B

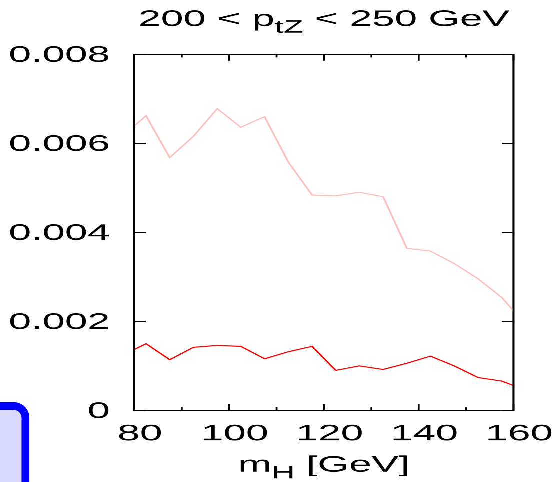
Herwig 6.510 + Jimmy 4.31 + FastJet 2.3



HZ Signal



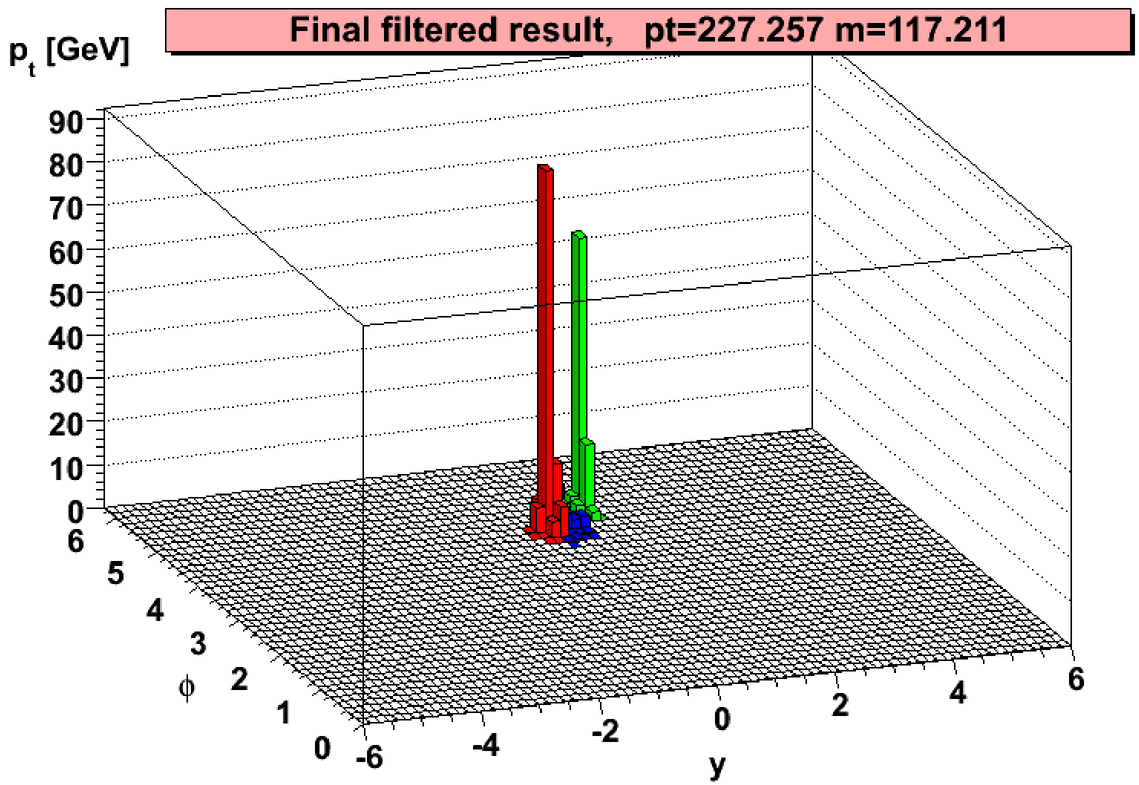
Zbb Background



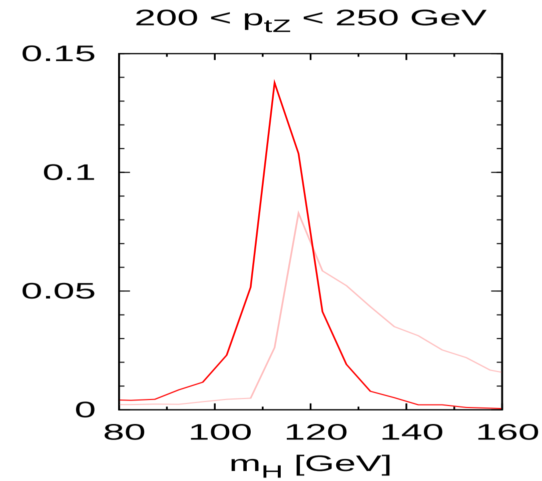
Re-cluster: $R_{\text{filt}} = 0.3$

Boosted Higgs: one event, effects on S/B

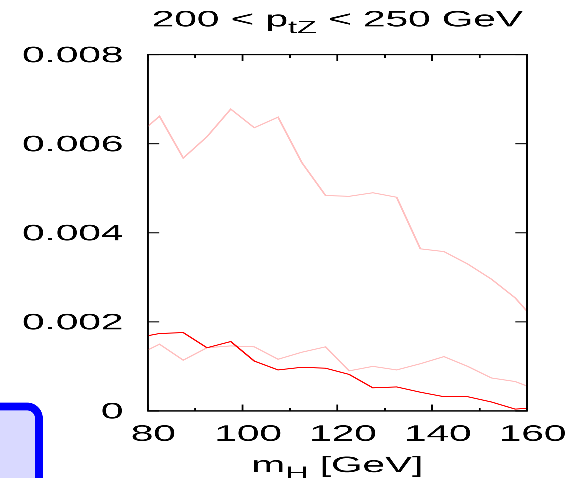
Herwig 6.510 + Jimmy 4.31 + FastJet 2.3



HZ Signal

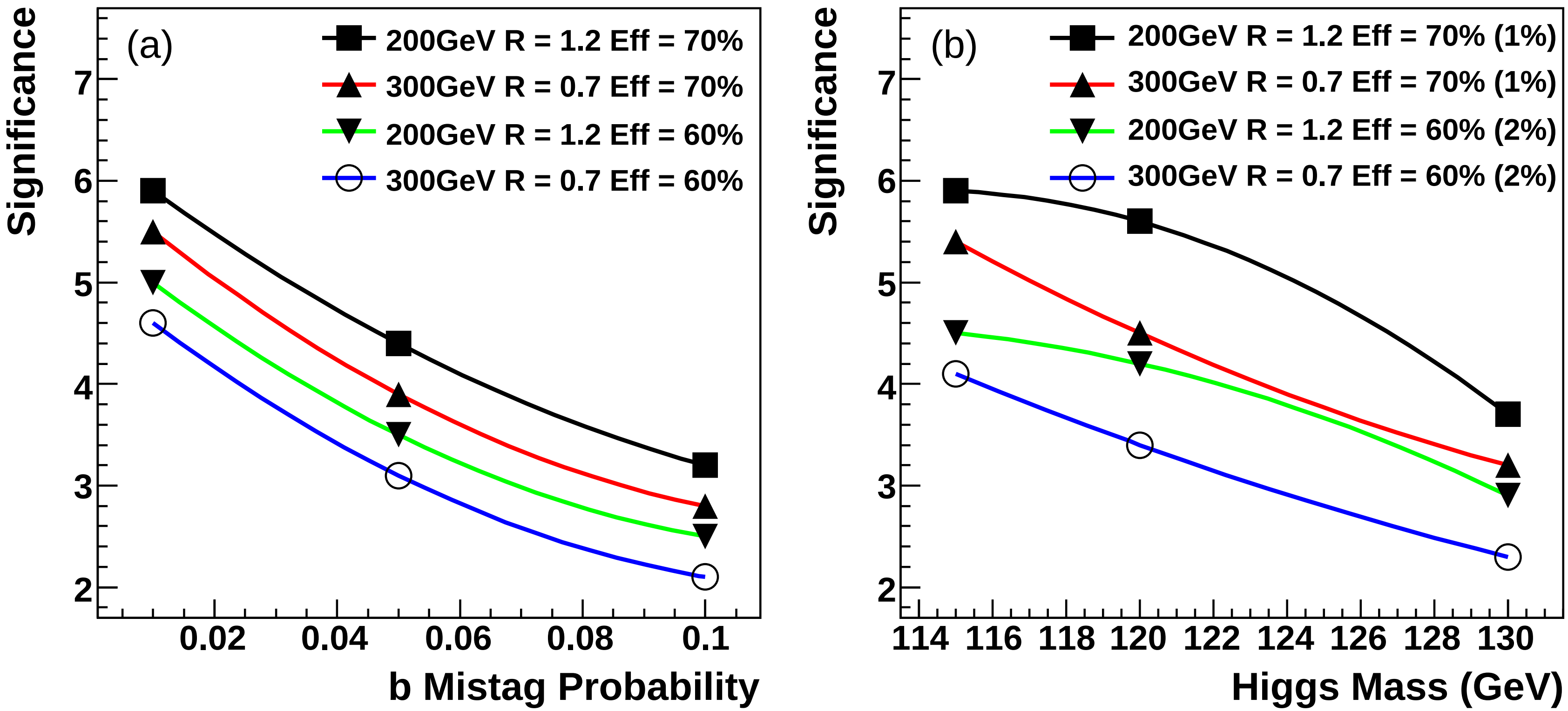


Zbb Background



Filter: keep 3 hardets ($m = 117$ GeV)

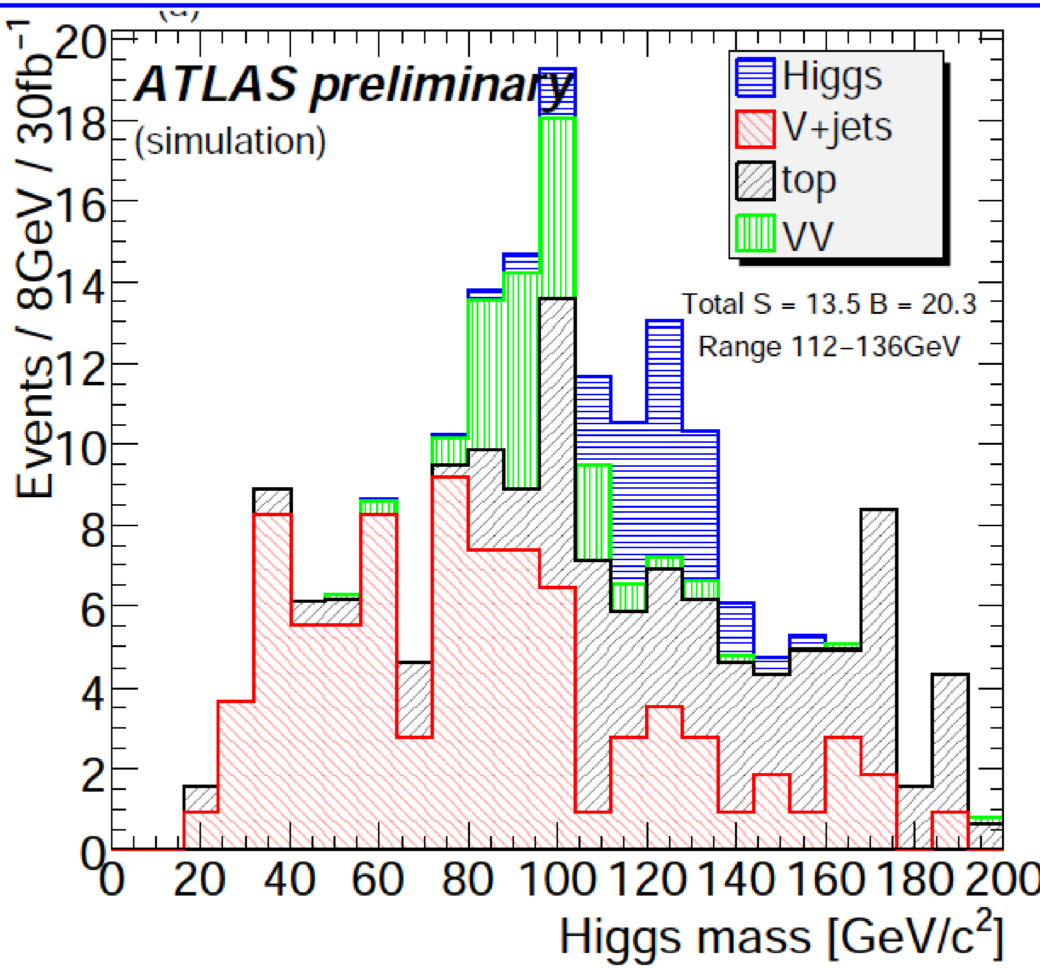
Boosted Higgs: one event, effects on S/B



More than 3σ for most scenarios (30 fb^{-1})

Filter: keep 3 hardjets ($m = 117$ GeV)

Boosted Higgs: one event, effects on S/B



Under consideration in ATLAS (here $WZ \to l\nu b\bar{b}$, 30 fb^{-1})

Filter: keep 3 hardjets ($m = 117$ GeV)

New generations
2. pileup subtraction using jet areas

[M.Cacciari,G.Salam, 07]

[M.Cacciari,G.Salam,GS, 08]

[M.Cacciari,J.Rojo,G.Salam,GS, 10]

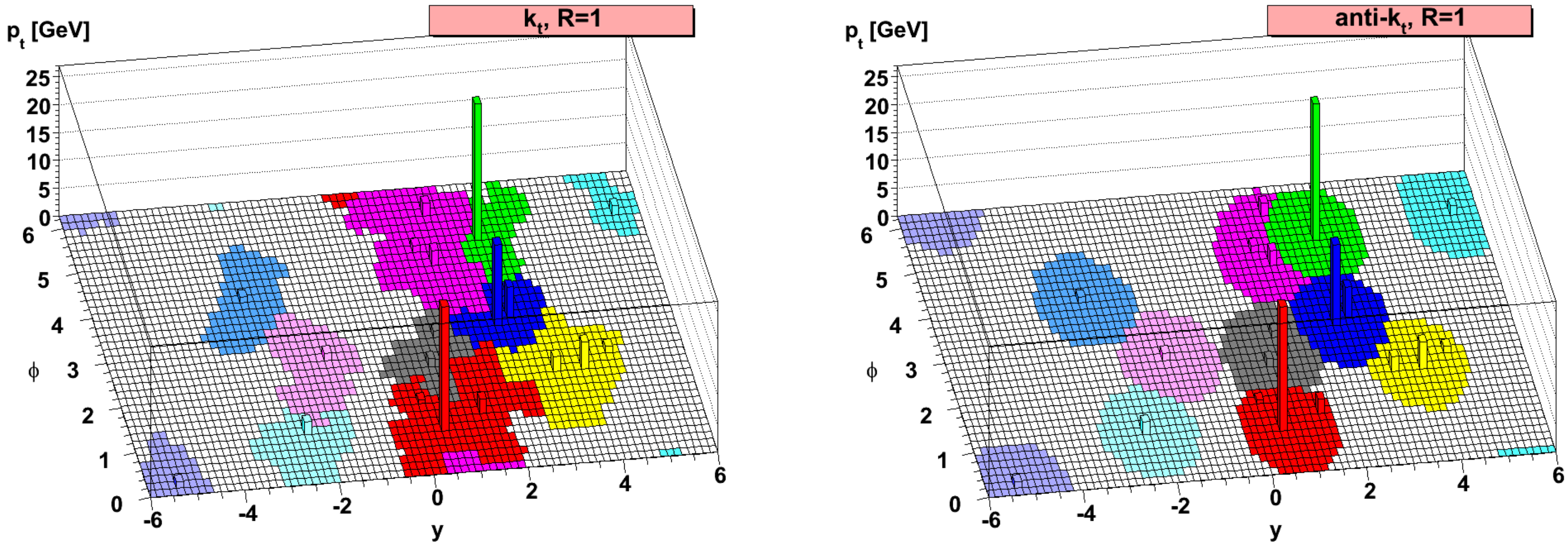
Central formula

$$p_{t,\mathrm{jet}}^{(\mathrm{sub})} = p_{t,\mathrm{jet}} - \rho_{\mathrm{bkg}} A_{\mathrm{jet}}$$

Central formula

$$p_{t,\mathrm{jet}}^{(\mathrm{sub})} = p_{t,\mathrm{jet}} - \rho_{\mathrm{bkg}} A_{\mathrm{jet}}$$

- jet area:
  - defined to mimic the reaction to the background
  - implemented in FastJet
  - analytic handle

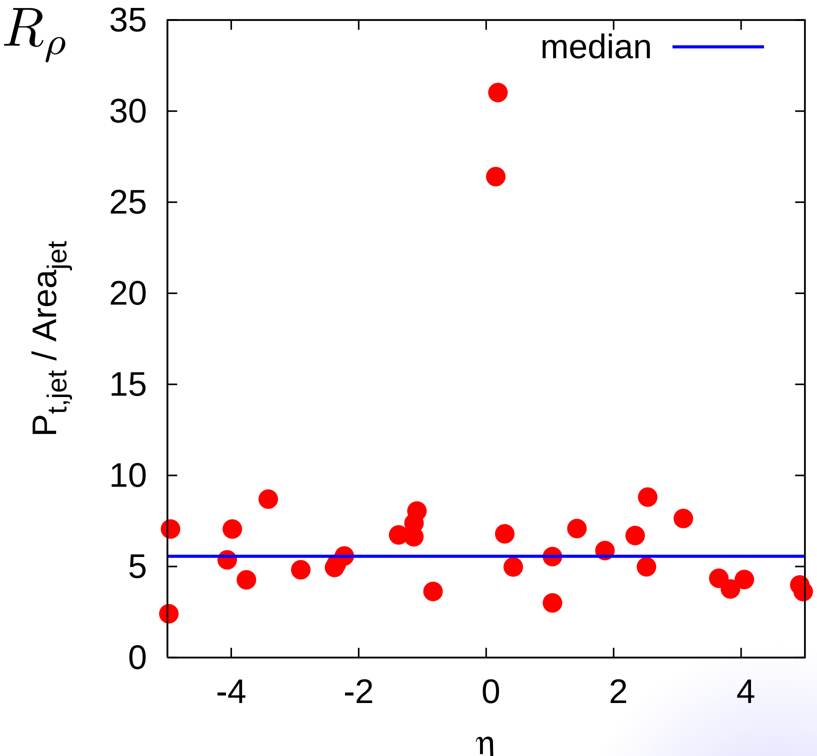


Central formula

$$p_{t,\text{jet}}^{(\text{sub})} = p_{t,\text{jet}} - \rho_{\text{bkg}} A_{\text{jet}}$$

- jet area:
- ρ_{bkg}, the background p_t density per unit area
  - Cluster with k_t of C/A with "radius" R_ρ
  - Estimate ρ_{bkg} using

$$\rho_{\text{bkg}} = \underset{j \in \text{jets}}{\text{median}} \left\{ \frac{p_{t,j}}{A_j} \right\}$$



Subtraction uncertainties

- Background fluctuations: (inside an event!)

$$p_{t,\mathrm{jet}} = p_{t,\mathrm{jet}}^{\mathrm{hard}} + \rho_{\mathrm{bkg}} A_{\mathrm{jet}} \pm \sigma_{\mathrm{bkg}} \sqrt{A_{\mathrm{jet}}}$$

Hint: reduce A_{jet} *e.g.* using filtering

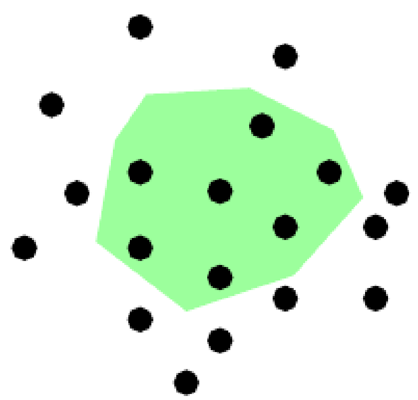
Subtraction uncertainties

- Background fluctuations: (inside an event!)

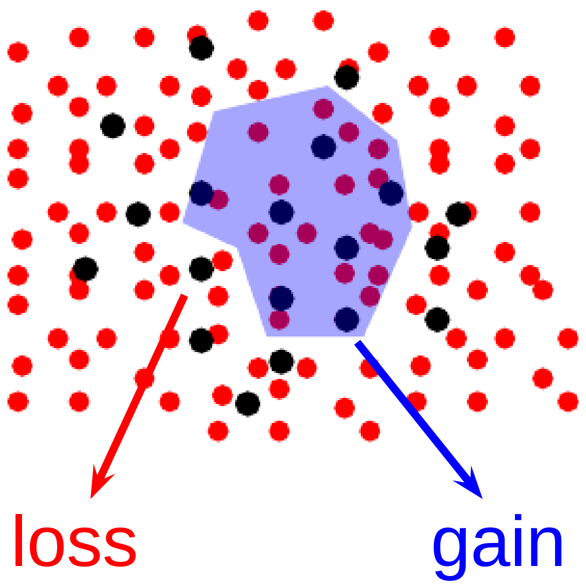
$$p_{t,\text{jet}} = p_{t,\text{jet}}^{\text{hard}} + \rho_{\text{bkg}} A_{\text{jet}} \pm \sigma_{\text{bkg}} \sqrt{A_{\text{jet}}}$$

- Back-reaction:

No background



With background



Hint: use anti-k_t (rigidity!)

Subtraction uncertainties

- Background fluctuations: (inside an event!)

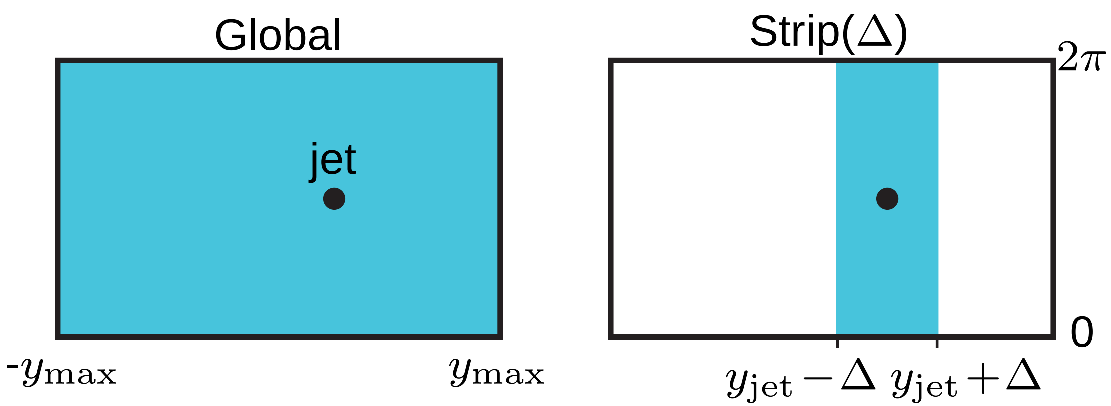
$$p_{t,\mathrm{jet}} = p_{t,\mathrm{jet}}^{\mathrm{hard}} + \rho_{\mathrm{bkg}} A_{\mathrm{jet}} \pm \sigma_{\mathrm{bkg}} \sqrt{A_{\mathrm{jet}}}$$

- Back-reaction:

- Background non-uniform (*e.g.* rap dependence)

  Use jets in a *local range* to estimate ρ_{bkg}

$$\rho_{\mathrm{bkg}} = \underset{j \in \mathcal{R}}{\mathrm{median}} \left\{ \frac{p_{t,j}}{A_j} \right\}$$



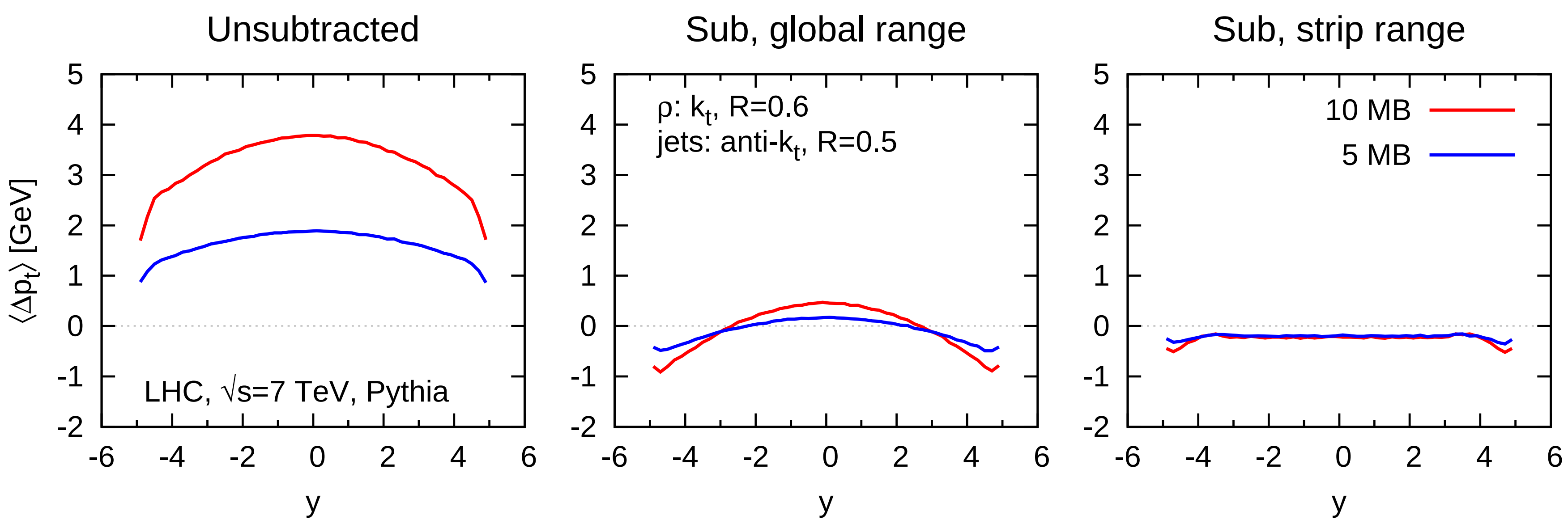
+ exclude the (*e.g.*) 2 hardest jets

Subtraction uncertainties

- Background fluctuations: (inside an event!)

$$p_{t,\text{jet}} = p_{t,\text{jet}}^{\text{hard}} + \rho_{\text{bkg}} A_{\text{jet}} \pm \sigma_{\text{bkg}} \sqrt{A_{\text{jet}}}$$

- Back-reaction:
- Background non-uniform (*e.g.* rap dependence)



New generations
3. Optimisation

[M.Cacciari,J.Rojo,G.Salam,GS, 08]

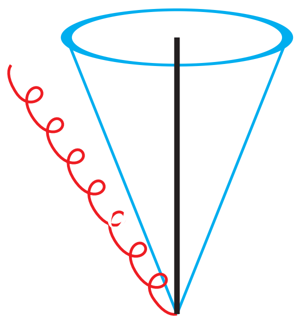
[M.Dasgupta,L.Magnea,G.Salam, 08]

[GS, 10]

Optimisation: underlying idea

Competition between

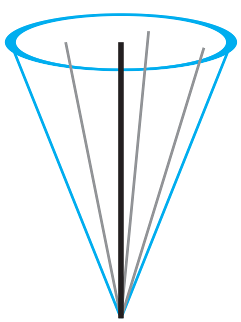
- catching perturbative radiation



Out-of-cone radiation:

$$\langle \delta p_t \rangle \propto -\int_R \frac{d\theta}{\theta} \sim -\log(1/R)$$

- not catching soft background radiation (underlying event)



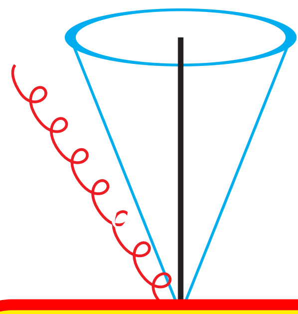
$\langle \delta p_t \rangle \sim$ Soft contents $\propto$ jet area $\sim R^2$

the coefficients depend on the algorithm

Optimisation: underlying idea

Competition between

- catching perturbative radiation

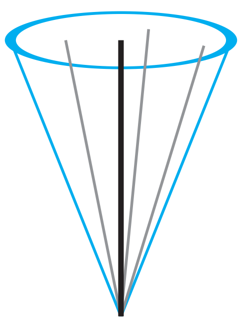


Out-of-cone radiation:

$$\langle \delta p_t \rangle \propto -\int_R \frac{d\theta}{\theta} \sim -\log(1/R)$$

What is the optimal jet definition (algo+R!)?

- not catching soft background radiation (underlying event)



$\langle \delta p_t \rangle \sim$ Soft contents $\propto$ jet area $\sim R^2$

the coefficients depend on the algorithm

Framework

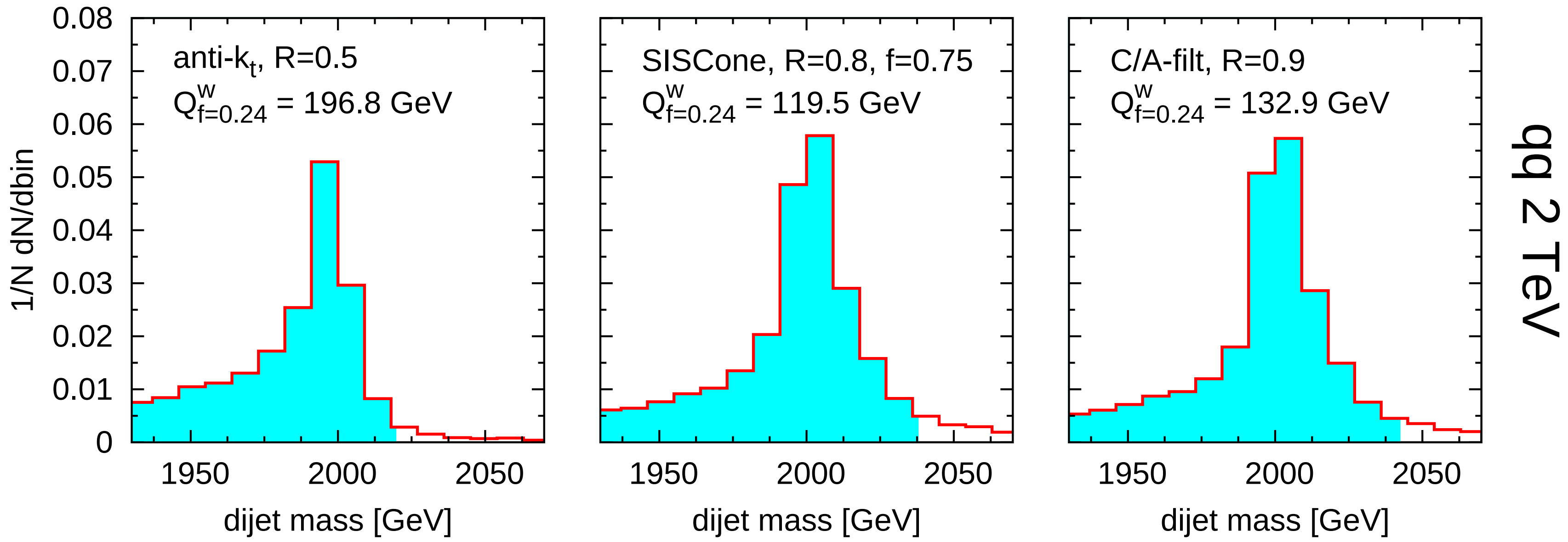
Example process to illustrate various effects:

$$Z' \to q\bar{q} \to 2 \text{ jets}$$

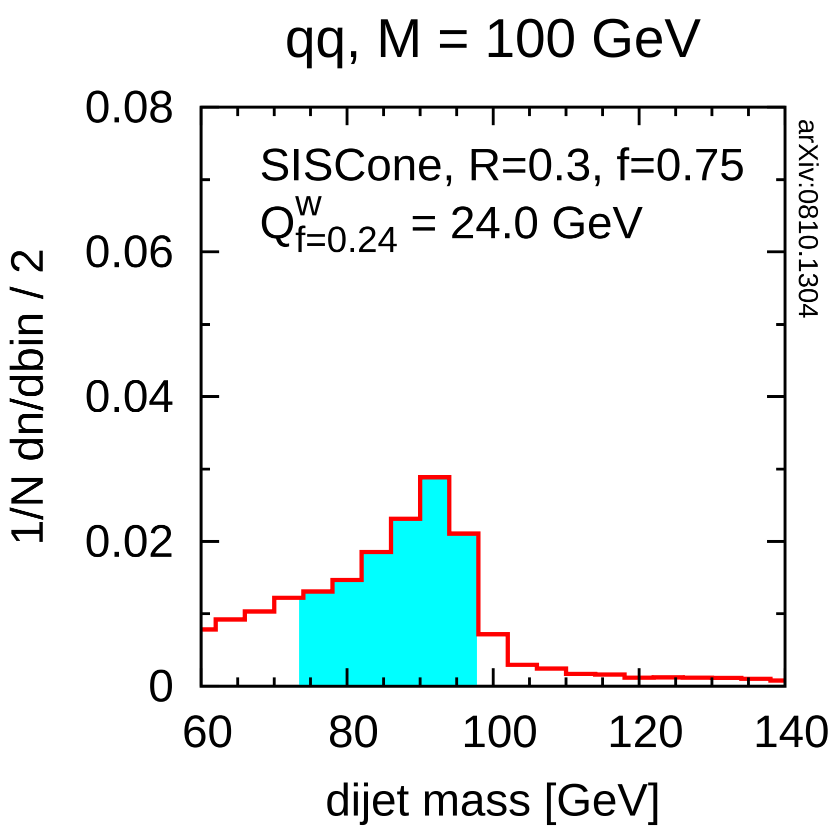
- $M_{Z'}$ can be varied (between 100 GeV and 4 TeV)
- Also valid for $H \to gg$ to study gluon jets
- Reconstruction method:
  - get the 2 hardest jets: j_1 and j_2
  - reconstruct the Z': $m_{Z'} = (j_1 + j_2)^2$

  Look how the mass peak is reconstructed
- Also $t\bar{t}$ with full hadronic decay for multijet tests

Observations



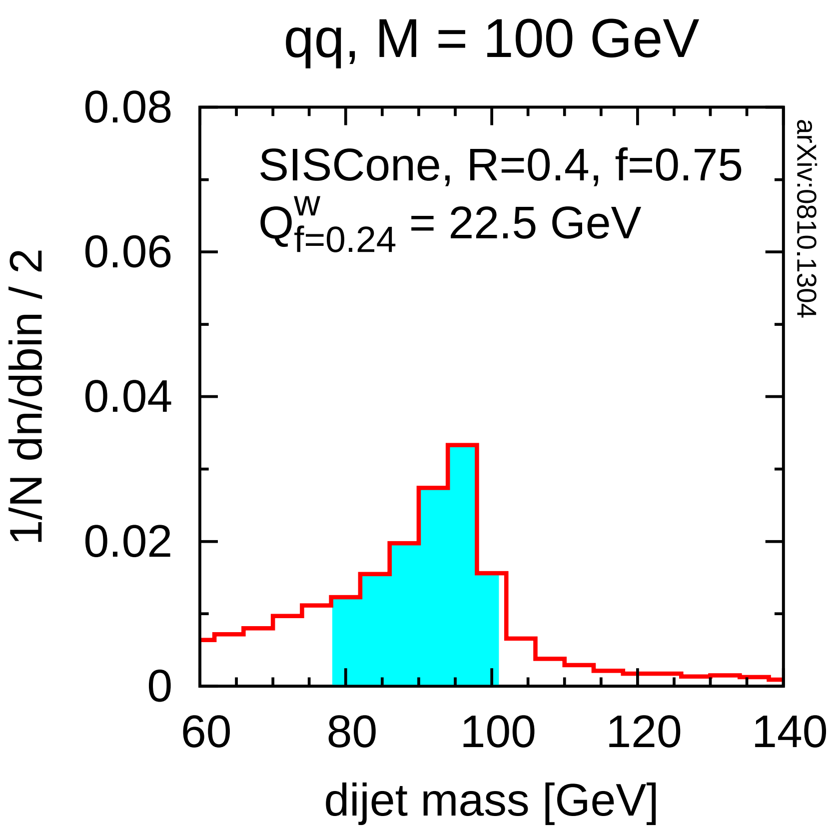
Observations



Histogram:

fixed mass, algorithm

Observations

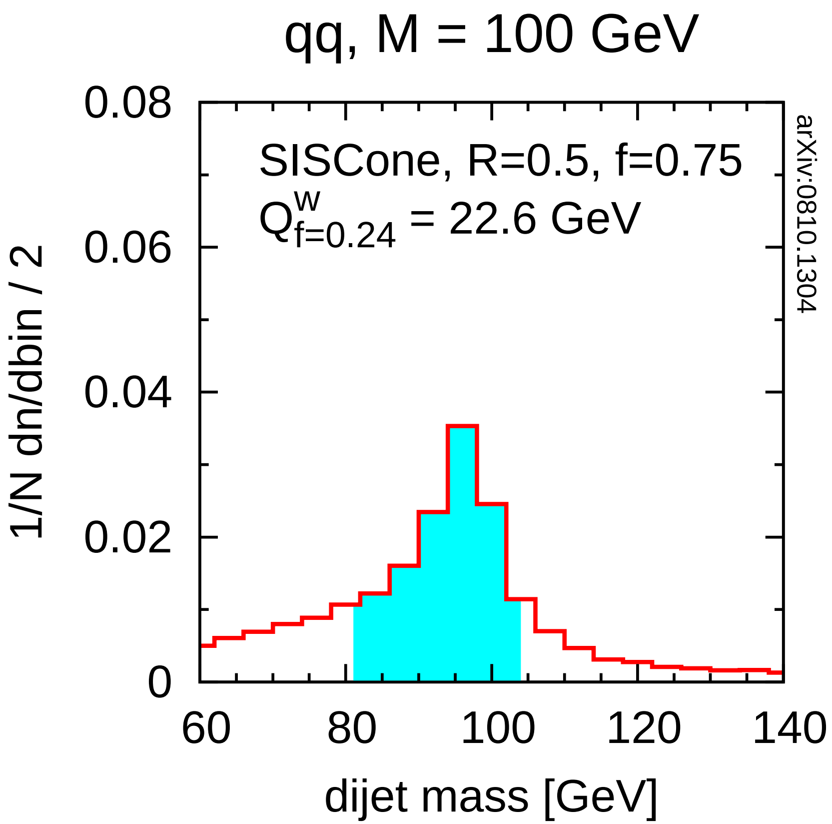


Histogram:

fixed mass, algorithm

vary R

Observations

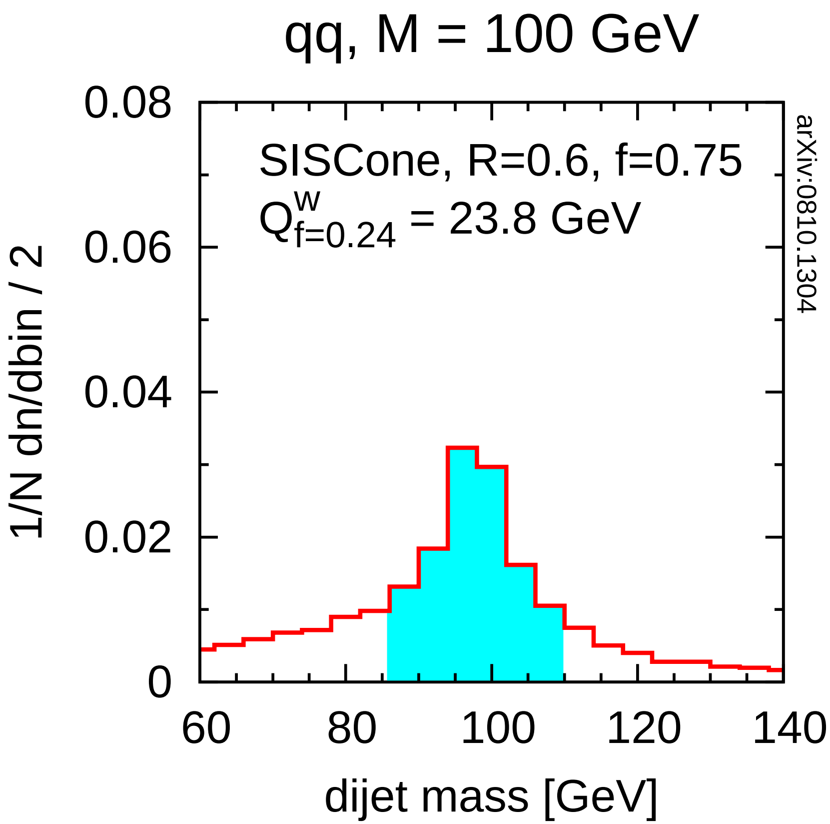


Histogram:

fixed mass, algorithm

vary R

Observations

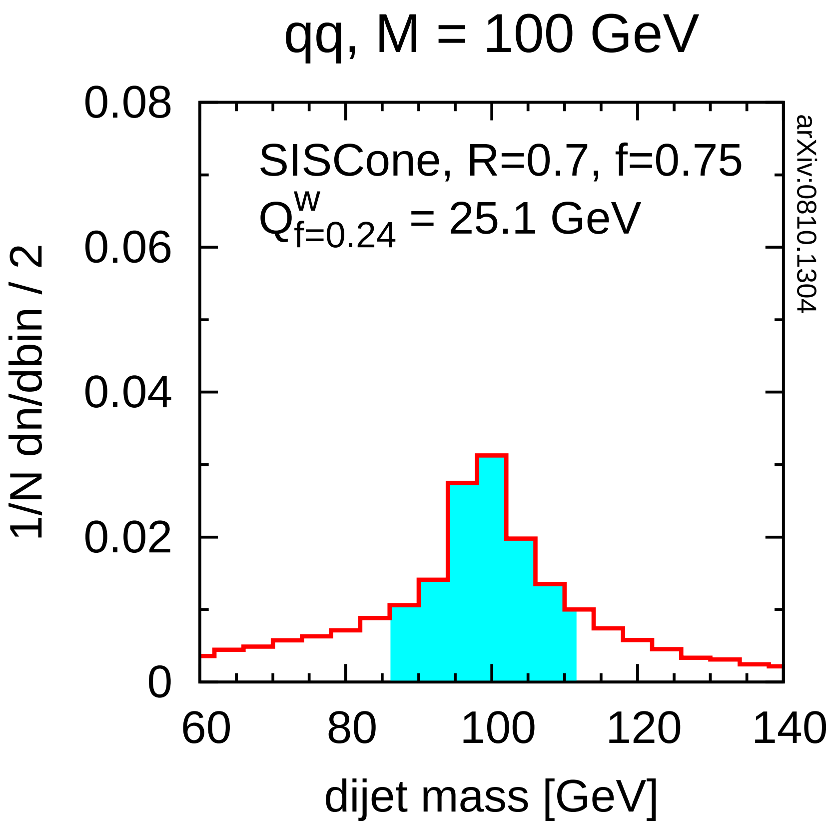


Histogram:

fixed mass, algorithm

vary R

Observations

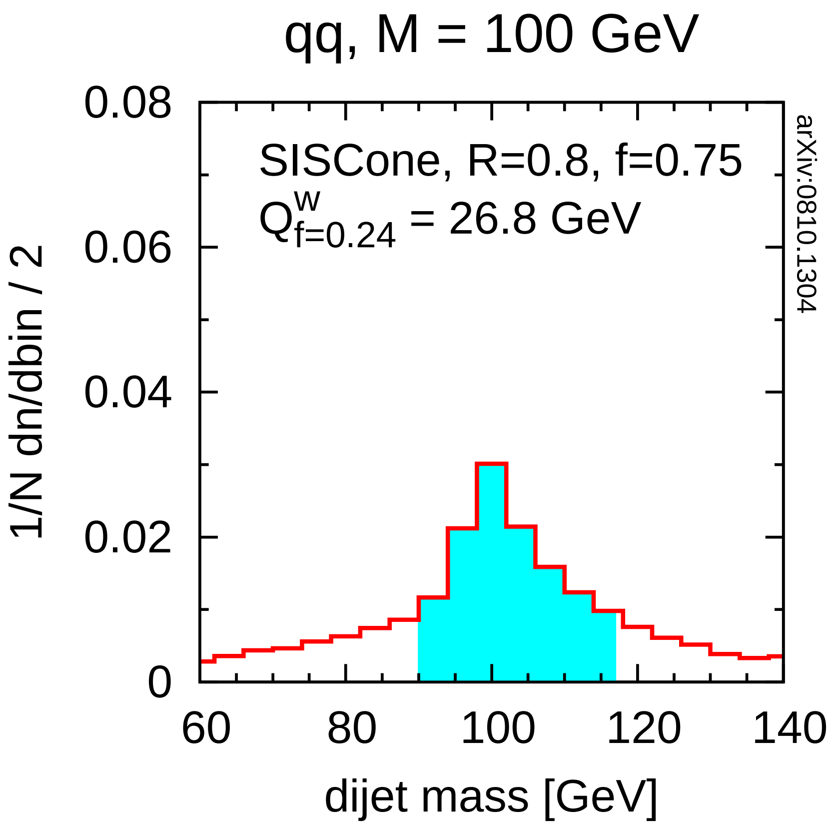


Histogram:

fixed mass, algorithm

vary R

Observations

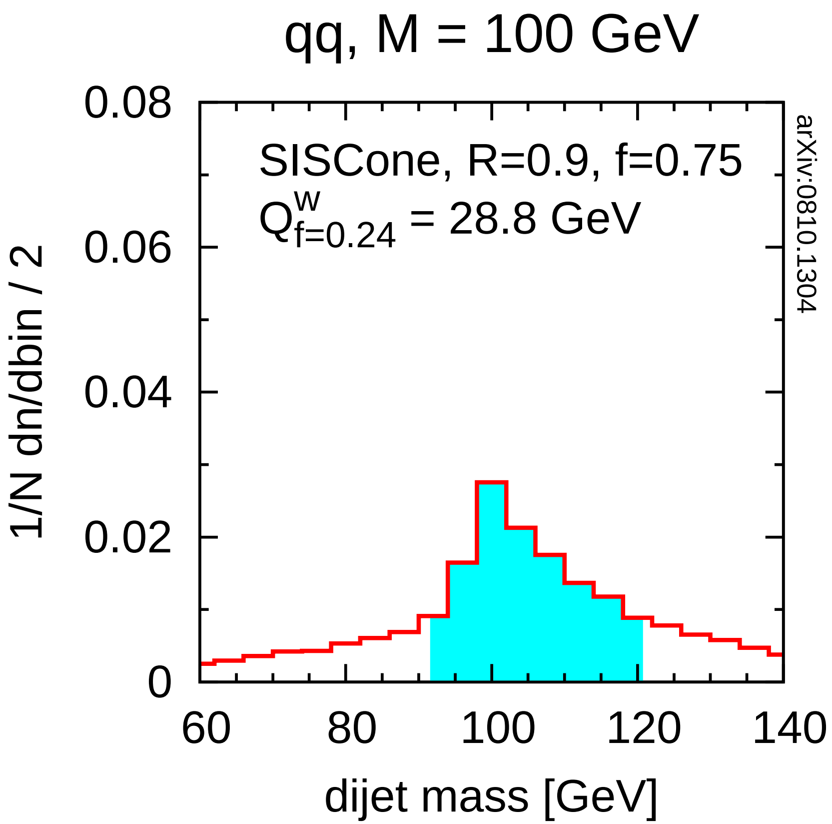


Histogram:

fixed mass, algorithm

vary R

Observations

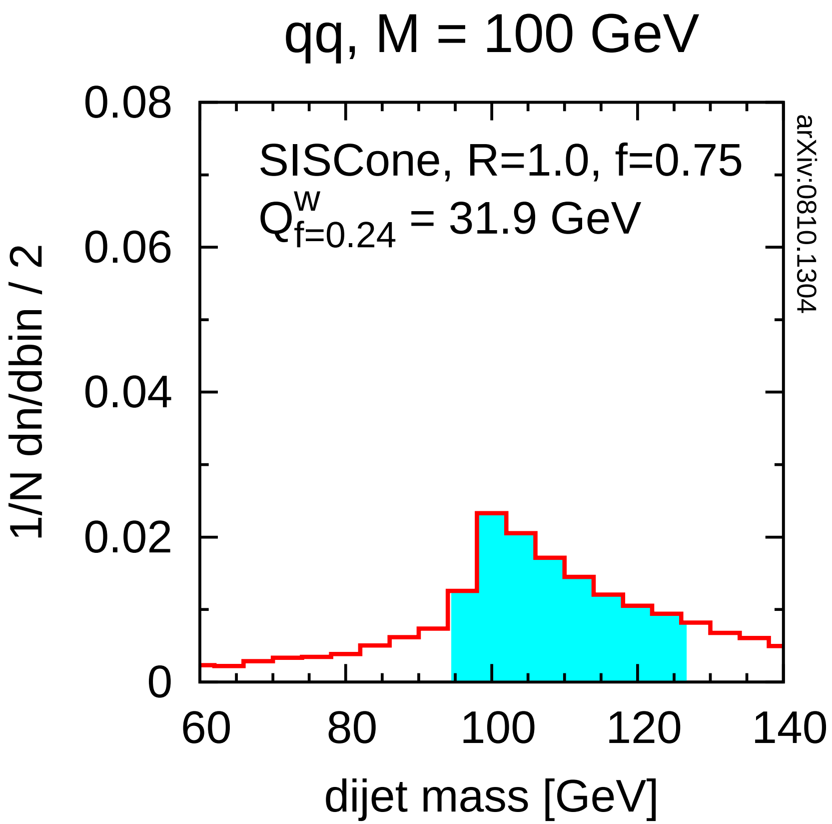


Histogram:

fixed mass, algorithm

vary R

Observations

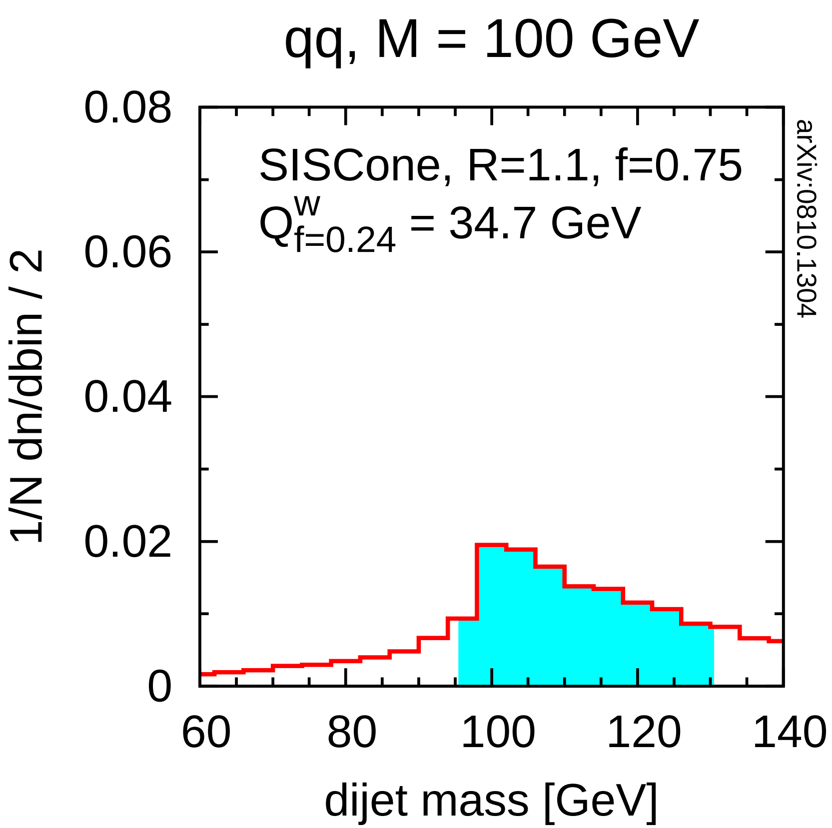


Histogram:

fixed mass, algorithm

vary R

Observations

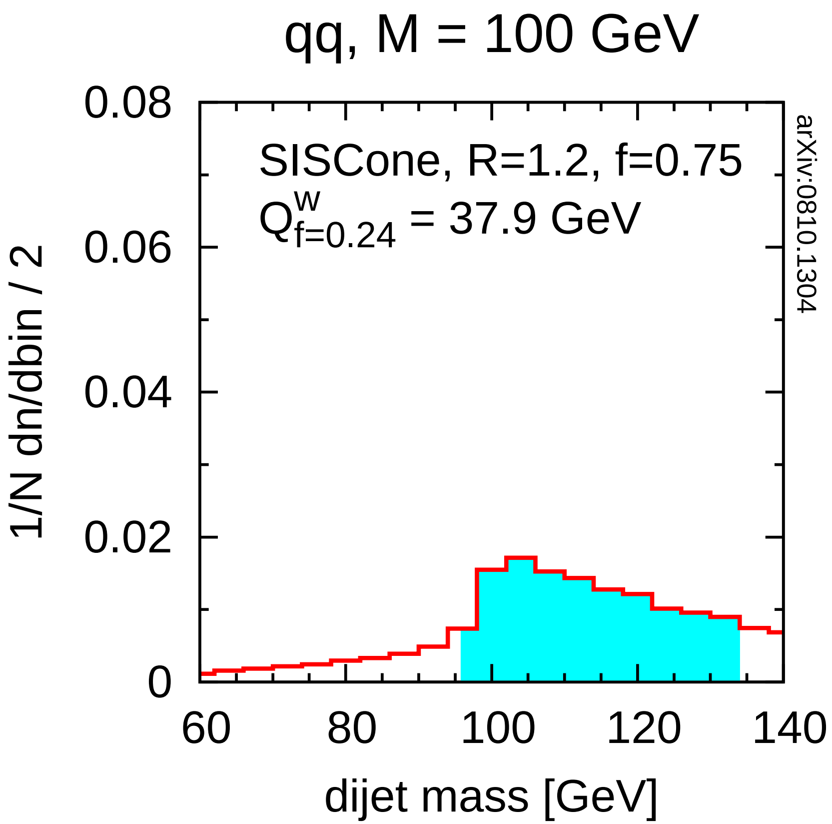


Histogram:

fixed mass, algorithm

vary R

Observations

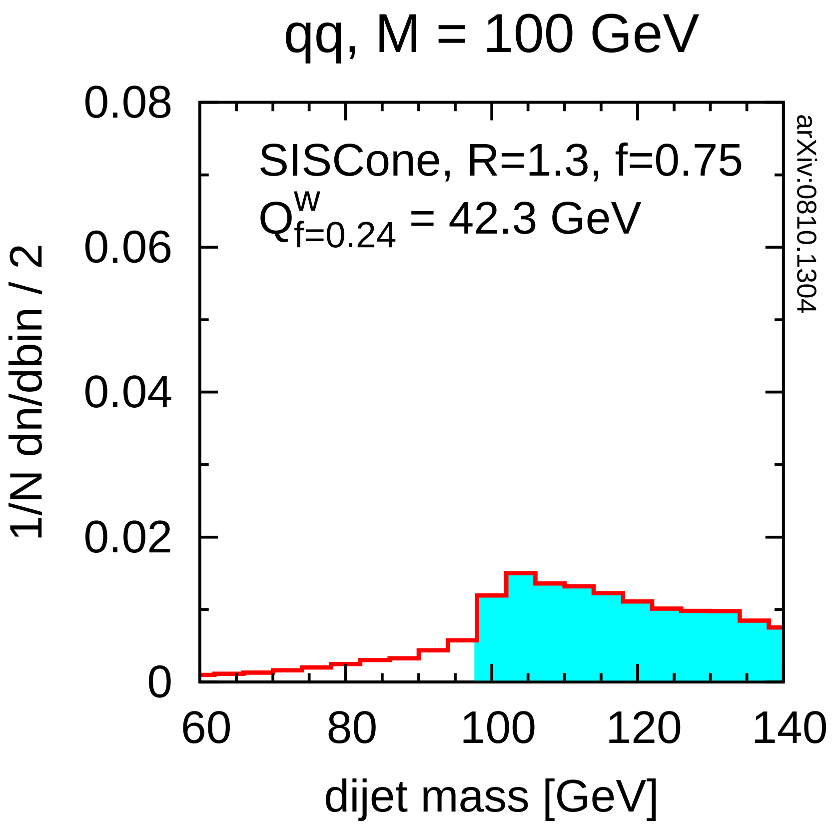


Histogram:

fixed mass, algorithm

vary R

Observations

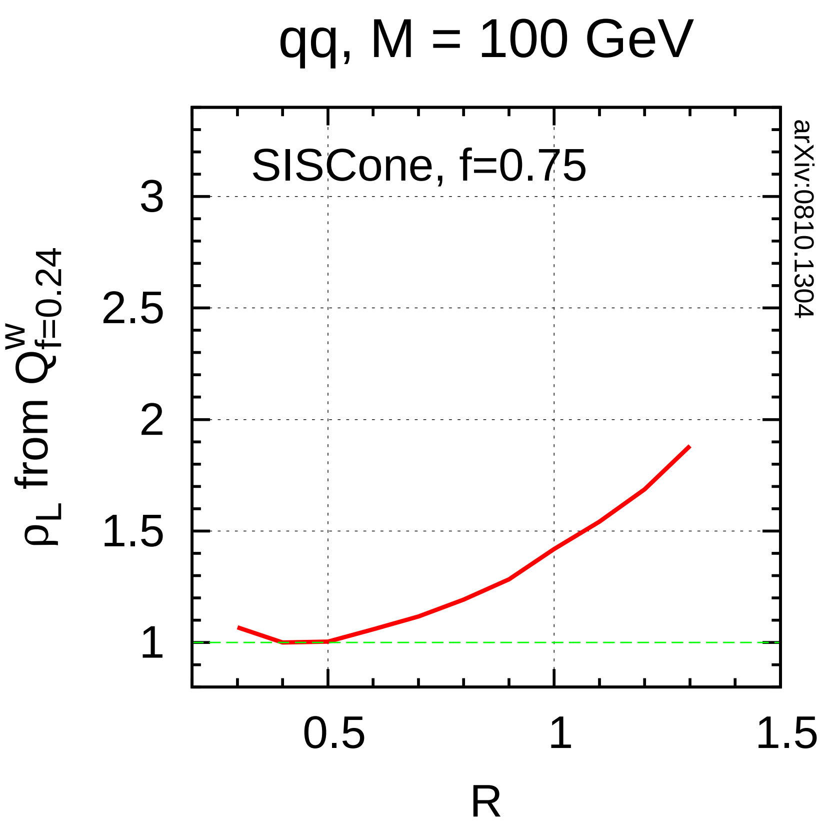


Histogram:

fixed mass, algorithm

vary R

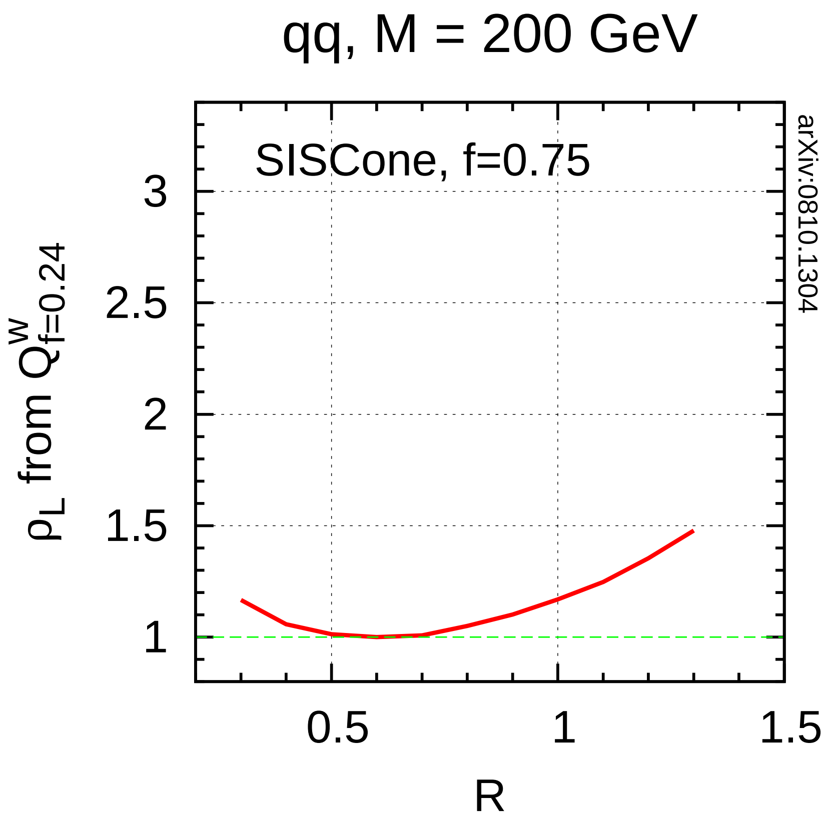
Observations



Width vs. R:

fixed algorithm

Observations

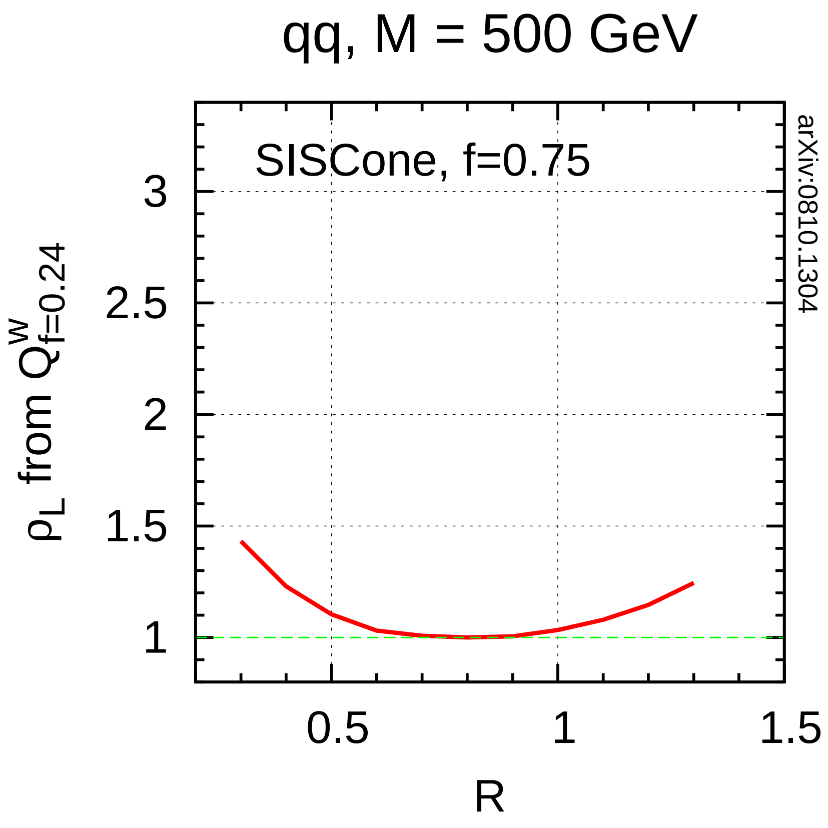


Width vs. R:

fixed algorithm

vary M

Observations

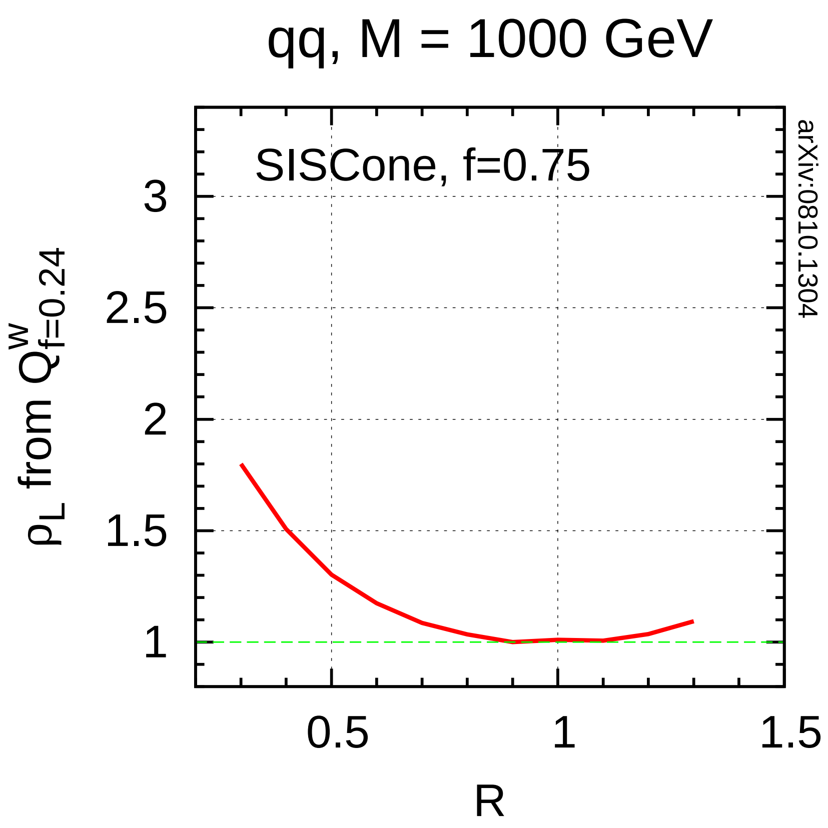


Width vs. R:

fixed algorithm

vary M

Observations

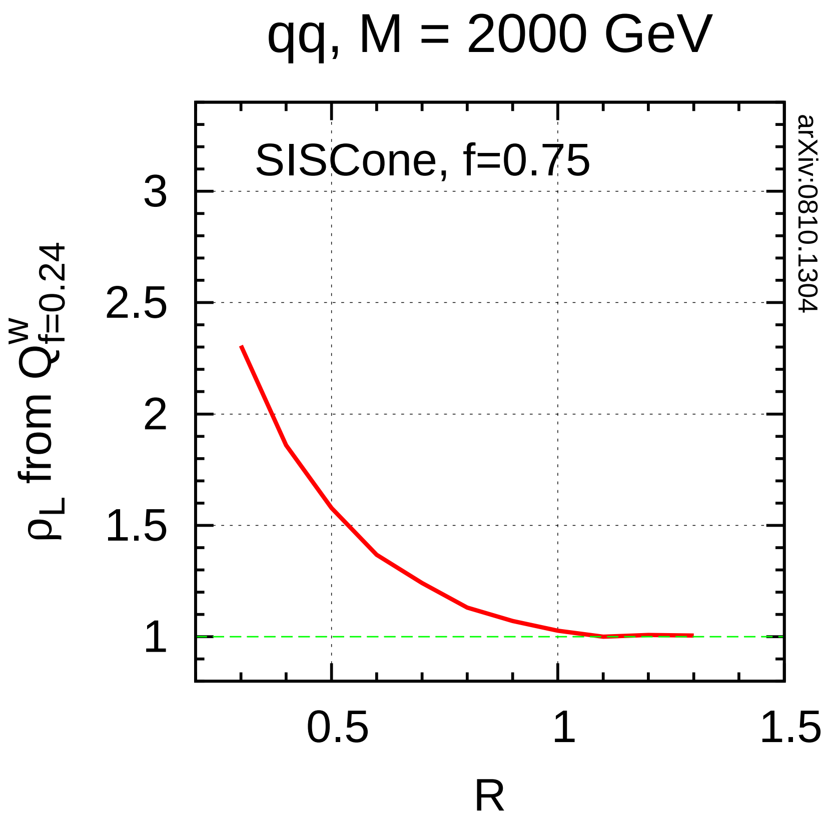


Width vs. R:

fixed algorithm

vary M

Observations



Width vs. R:

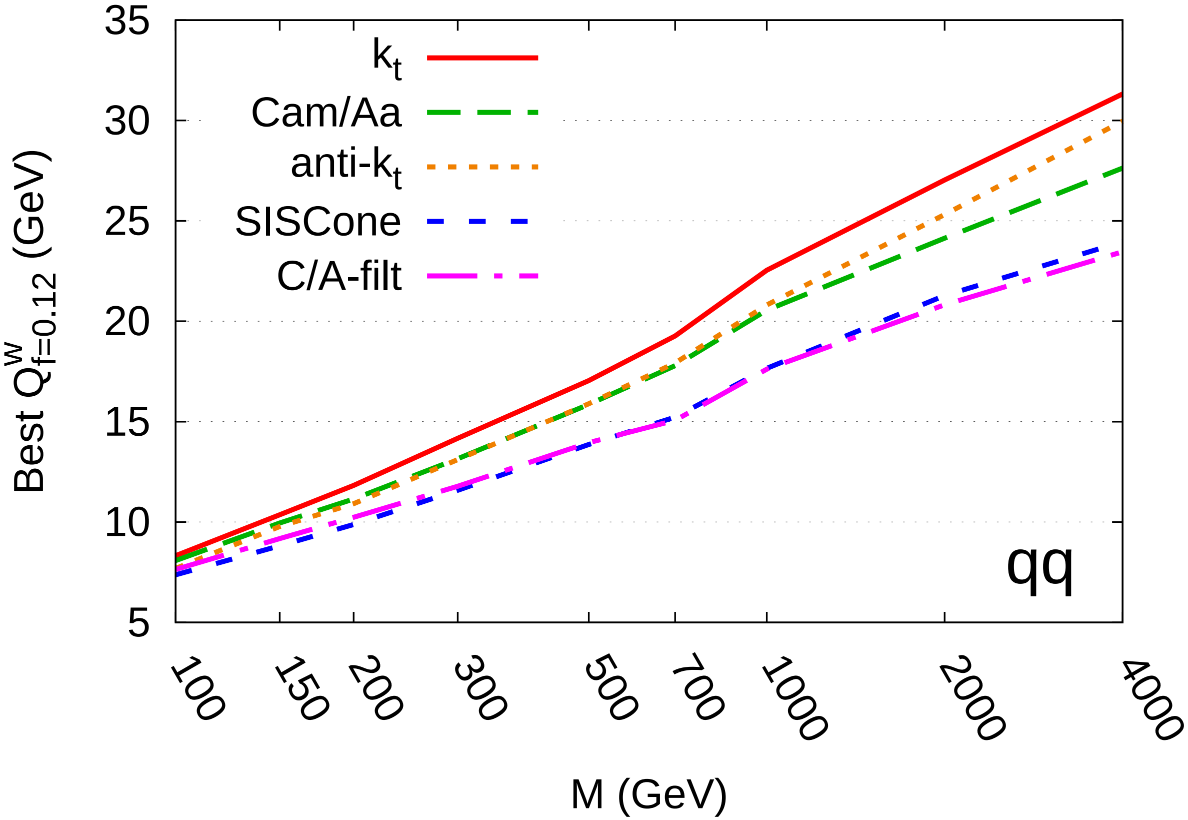
fixed algorithm

vary M

In summary:

- width vs. R: strong R dependence

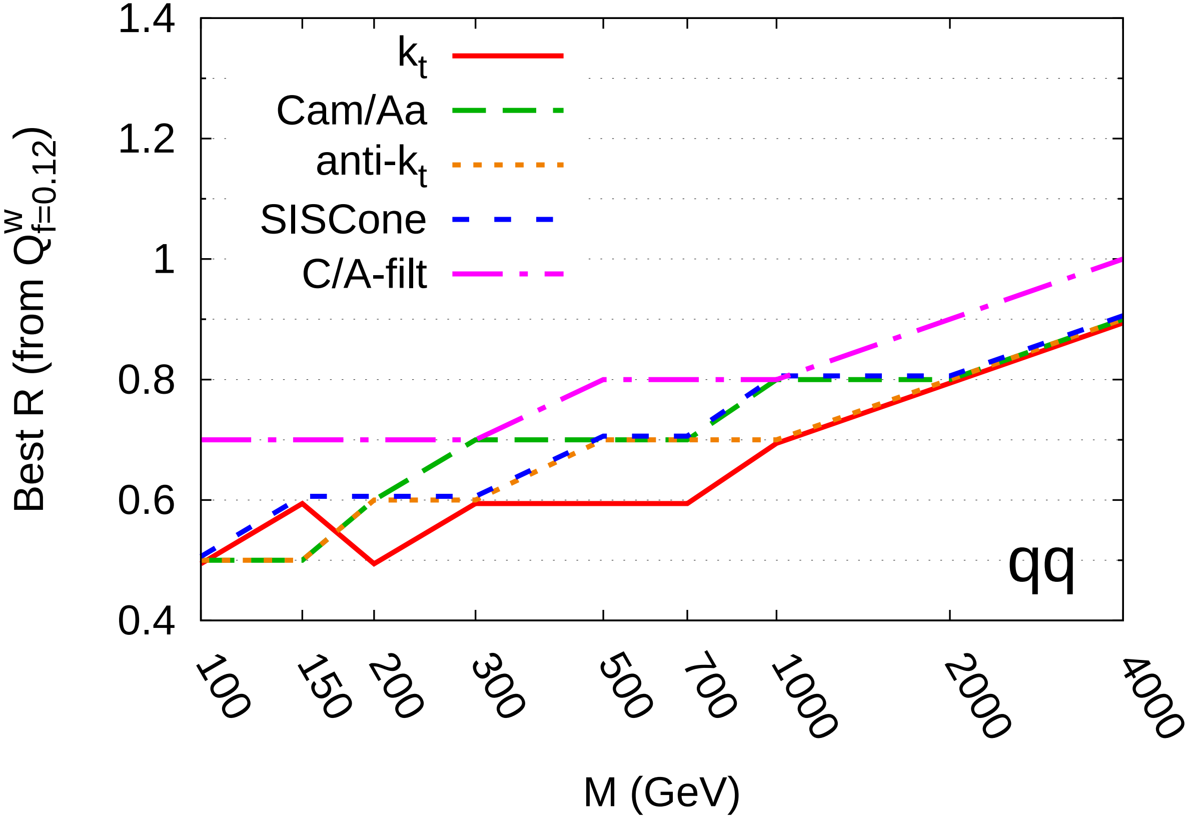
Observations



In summary:

- width vs. R: strong R dependence
- optimal width vs. M: SISCone, C/A(filt) preferred

Observations



In summary:

- width vs. R: strong R dependence
- optimal width vs. M: SISCone, C/A(filt) preferred
- optimal R vs. M: R_{best} increases with M

Towards analytics

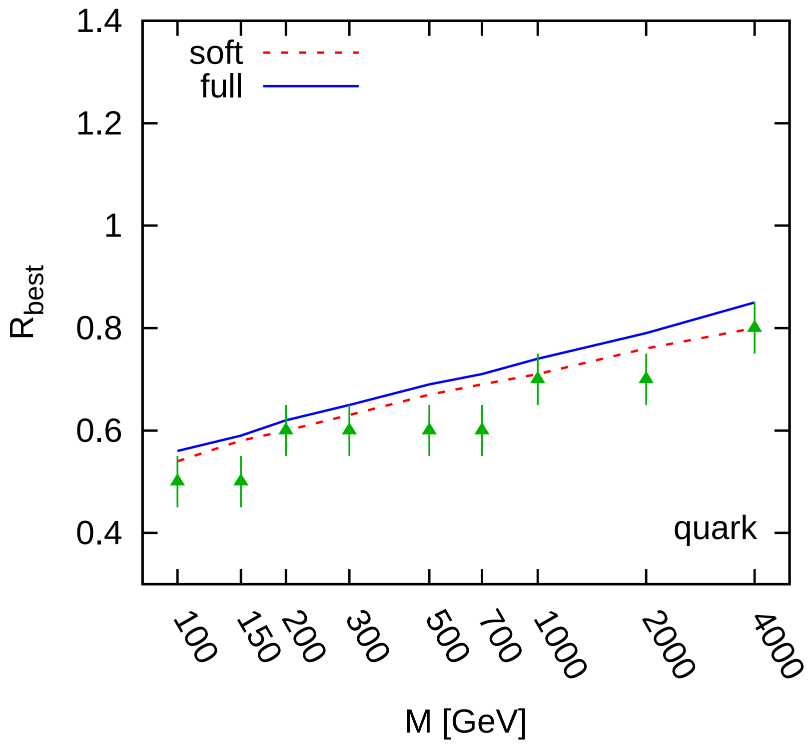
Analytic computation of the histogram including:

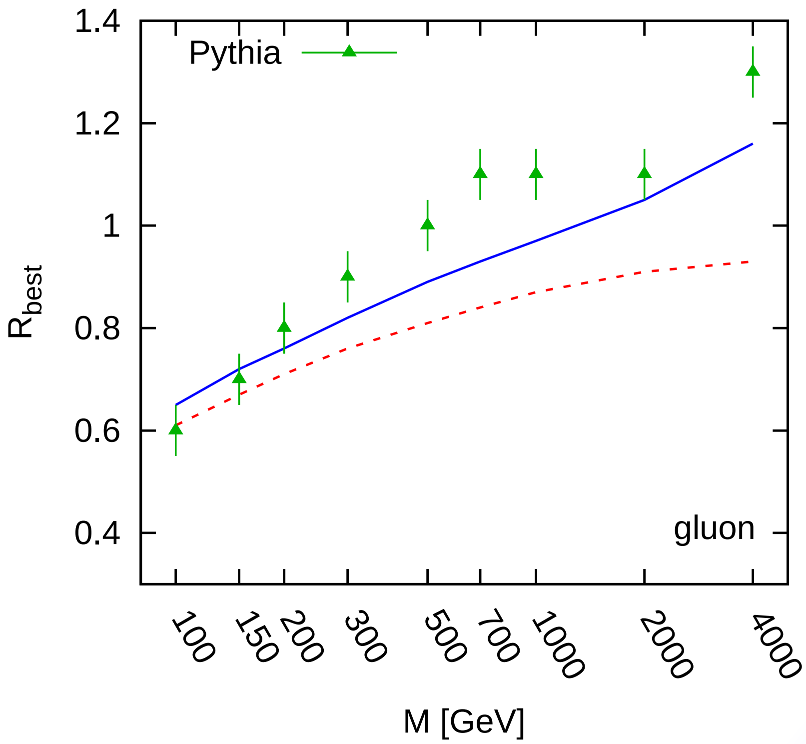
- pert. final-state radiation: loss $\propto \alpha_s M \log(1/R)$
- pert. initial-state radiation: gain $\propto \alpha_s M R^2$ (+PDFs)
- UE contamination: gain $\propto \rho_{\mathrm{UE}} R^2$

Towards analytics

Analytic computation of the histogram including:

- pert. final-state radiation: loss $\propto \alpha_s M \log(1/R)$
- pert. initial-state radiation: gain $\propto \alpha_s M R^2$ (+PDFs)
- UE contamination: gain $\propto \rho_{\mathrm{UE}} R^2$





Conclusions

- Finally a set of jet algs meeting the fundamental requirements
  i.e. Infrared-and-collinear-safe and fast
- Allows better/advanced usage of jets at the LHC
  - jet areas for background subtraction
  - jet substructure for boosted-objects tagging
  - UE-sensitivity using filtering/trimming
  - towards analytic understanding/optimisation
- Future: improve in those directions

Future of FastJet

FastJet 2.5/3.0 on its way:

- Interface: a jet knows about its clustering, *e.g.*

```
clust_seq.constituents(jet);
  → jet.constituents();
```

- Generic additional info in PseudoJet: `jet.extra_info()`

- `Selector` for selecting objects in a list, *e.g.*

```
Selector jet_sel =
  SelectorMaxAbsRap(2.5) && SelectorPtMin(20);
jets = jet_sel(clust_seq.inclusive_jets());
```

- Improved bkgd subtraction: `BackgroundEstimator`

- FastJet tools *e.g.* `Filter`