# Physics at the LHC
## Lecture 3: Jet Physics at the LHC

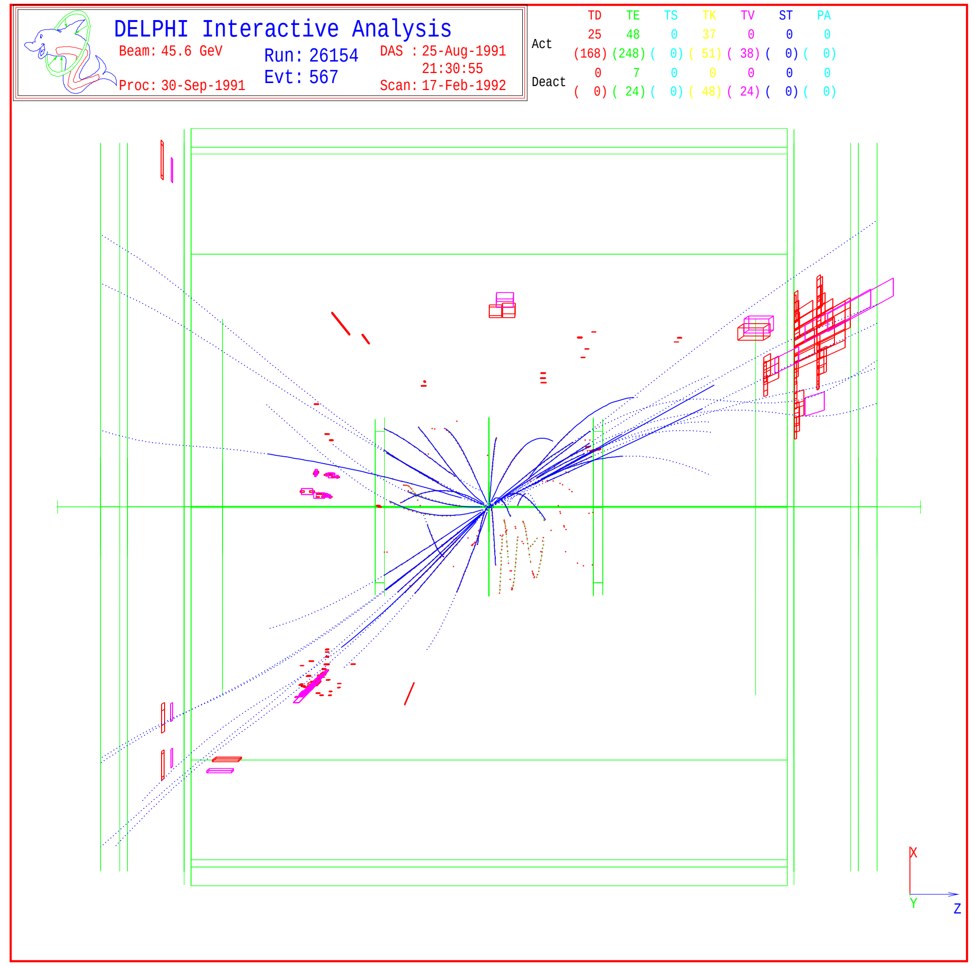
Klaus Mönig, Sven Moch

(klaus.moenig@desy.de, sven-olaf.moch@desy.de)

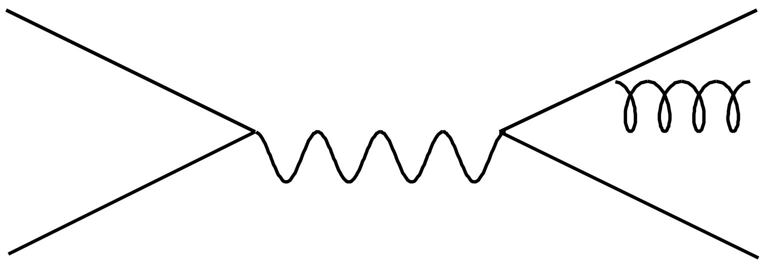
Wintersemester 2009/2010

# Sort of introduction: Jets in $e^+e^-$

Experiment: Measures hadrons




Theory: calculates final states with quarks and gluons



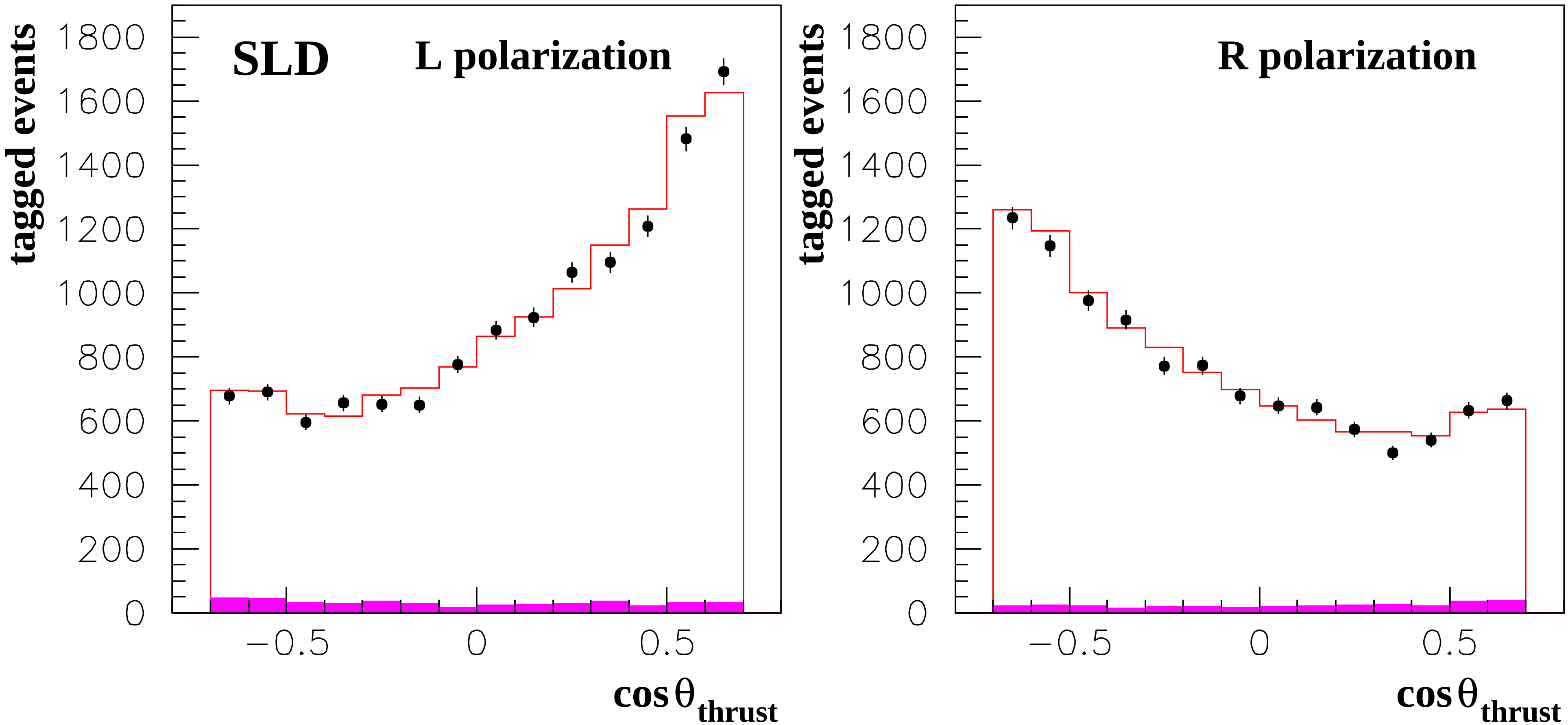
Must try to understand transition partons $\rightarrow$ hadrons

Must try to reconstruct partons from hadrons

# Experimental observation:

- Hadrons come in bundles (jets)
- Jets remember parton momentum

Signed $\cos\theta$ distribution of jets in polarised $e^+e^-$ scattering

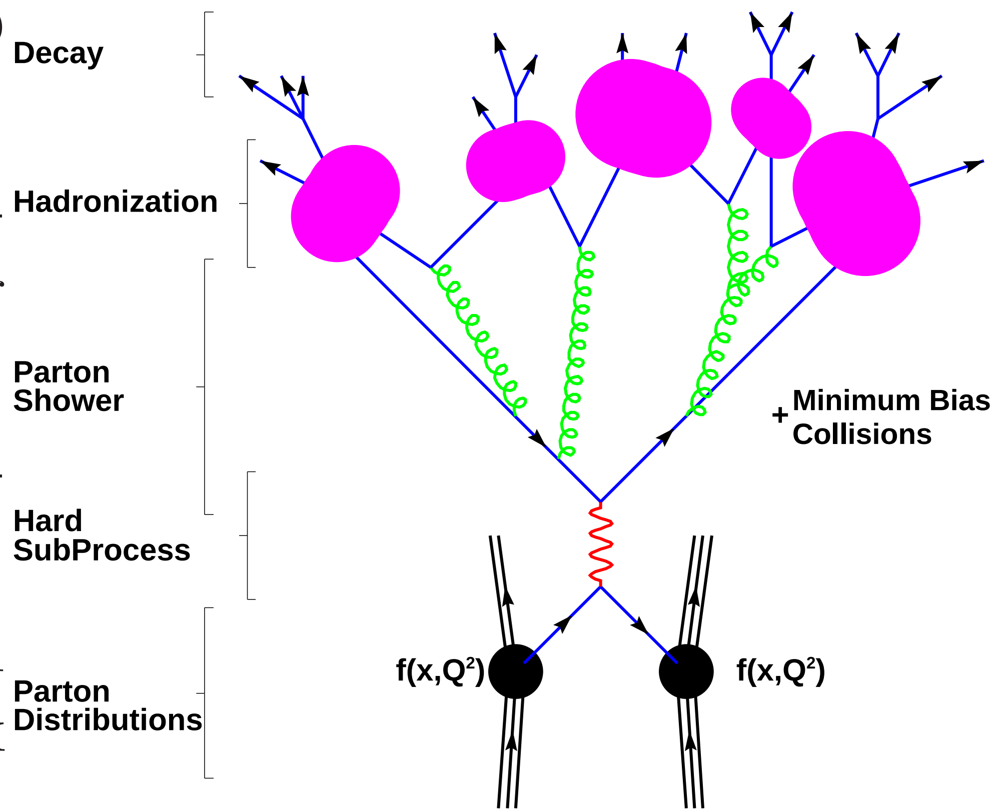


# Model of fragmentation

- Quarks and gluons radiate gluons
- Gluons split into $q\bar{q}$ pairs
- "final state" partons rearrange into hadrons

## MC description of gluon emission/splitting (parton shower)

- Radiation according to QCD splitting functions
- Interference dealt with by ordering principle (e.g. angular ordering)
- Shower stops at a scale of typically 1 GeV
- First one or two radiations can be according to exact matrix element



# Fragmentation models

## String fragmentation (Pythia, Lund model)

- Quarks span a string between them
- When the quarks move apart string tension increases
- When the tension reaches a critical value string breaks creating a new $q\bar{q}$ pair at the new ends
- When the energy is small enough hadrons are formed

## Cluster fragmentation (HERWIG)

- Remaining gluons split into $q\bar{q}$ pairs
- $q\bar{q}$ pairs rearrange into colour singlet clusters
- Clusters decay isotropically

# Jet algorithms

Try to "undo" fragmentation

Warning: Hadrons are colour singlets, quarks and gluons are colour triplets/octets $\Rightarrow$ quark/gluon "reconstruction" can never be exact

General jet algorithm:

- Define distance measure $d_{ij}$ for pair of particles
- Define combination algorithm

Jet algorithm

- Calculate $d_{ij}$ for all pairs and find $d_{ij,min}$
- If $d_{ij,min} > d_{cut}$ STOP
- Combine particles corresponding to $d_{ij,min}$
- Restart

## Distance measure:

- Most obvious choice: invariant mass (JADE algorithm)

  In practice massless approximation $d_{ij} = \frac{E_i E_j}{s}(1 - \cos\theta)$

  – was successfully used for QCD studies

  – algorithm tends to cluster all low energy particles first ➡ not so good for parton reconstruction

- To solve this problem replace mass by relative transverse momentum ($k_T$, Durham algorithm)

  $d_{ij} = \frac{\min(E_i^2 E_j^2)}{s}(1 - \cos\theta)$

  – equally well behaved for QCD studies

  – prefers to combine low angles ➡ closer to physics of parton showers

## Combination procedure

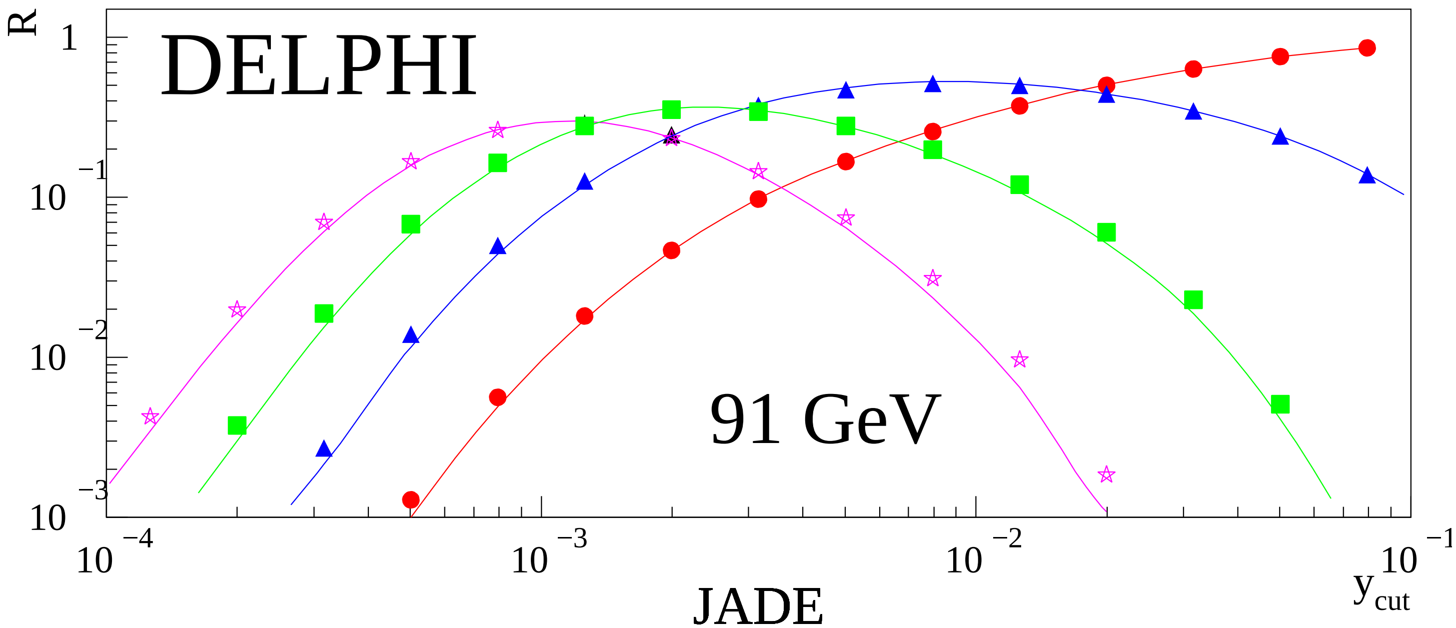
- Most obvious: add 4-momenta $p_n = p_i + p_j$
- Quarks and gluons are massless, two alternatives in use:
  - add 3-momenta and calculate energy assuming $m = 0$
  - add 4-momenta and rescale 3-momentum so that $m = 0$

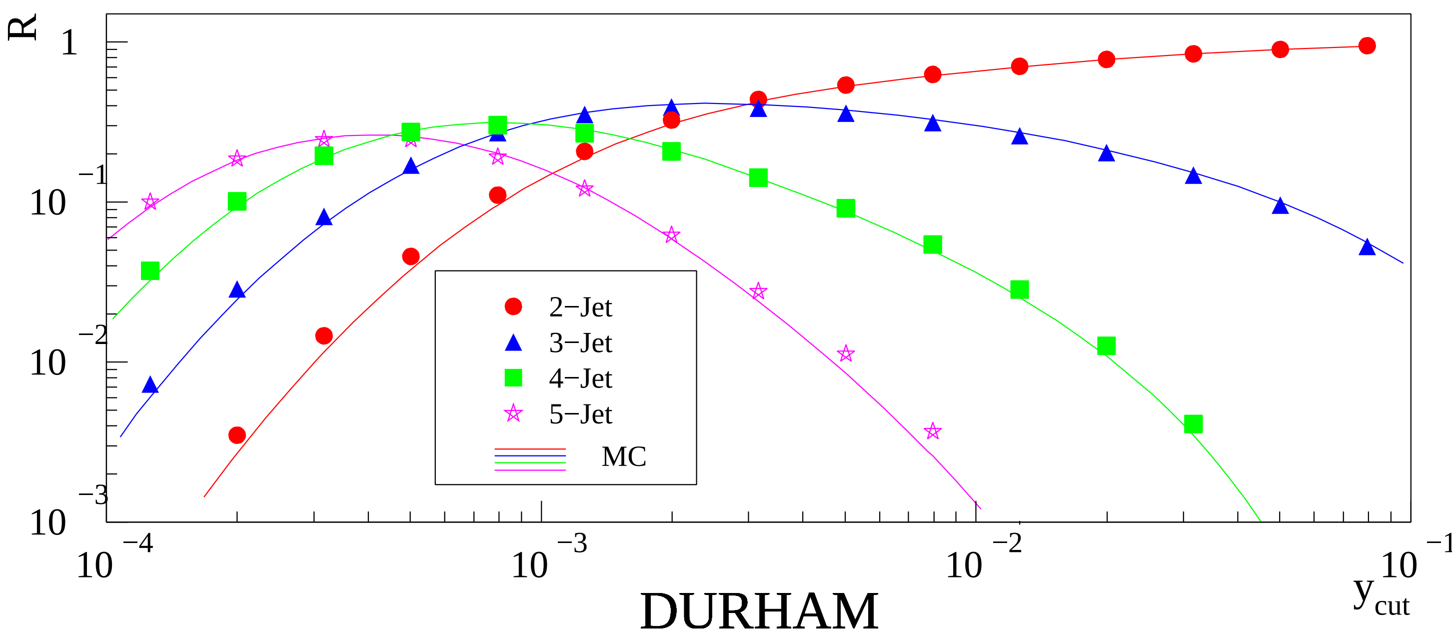
## Infrared and collinear safety:

- QCD Feynman graph diverges for $p_g \longrightarrow 0$
  ➡ algorithm must be stable when particle with $p \approx 0$ is added
- QCD Feynman graph diverges for splitting with $\theta \longrightarrow 0$ ➡ algorithm must be stable when particle is split into two with $\theta \approx 0$

ok for JADE and $k_T$

# Jet rates in $e^+e^-$



# Jets in pp

## Differences to $e^+e^-$

- Protons disappear as colour non-singlets in the beampipe
- Final state is boosted and algorithms not Lorenz invariant
- The underlying event adds activity in the detector
- At high luminosity there are additional minimum bias events that cannot be separated

Must adapt $k_T$ algorithm

New algorithms in pp ($p\bar{p}$): cone algorithms

## Adaptation of $k_T$ algorithm

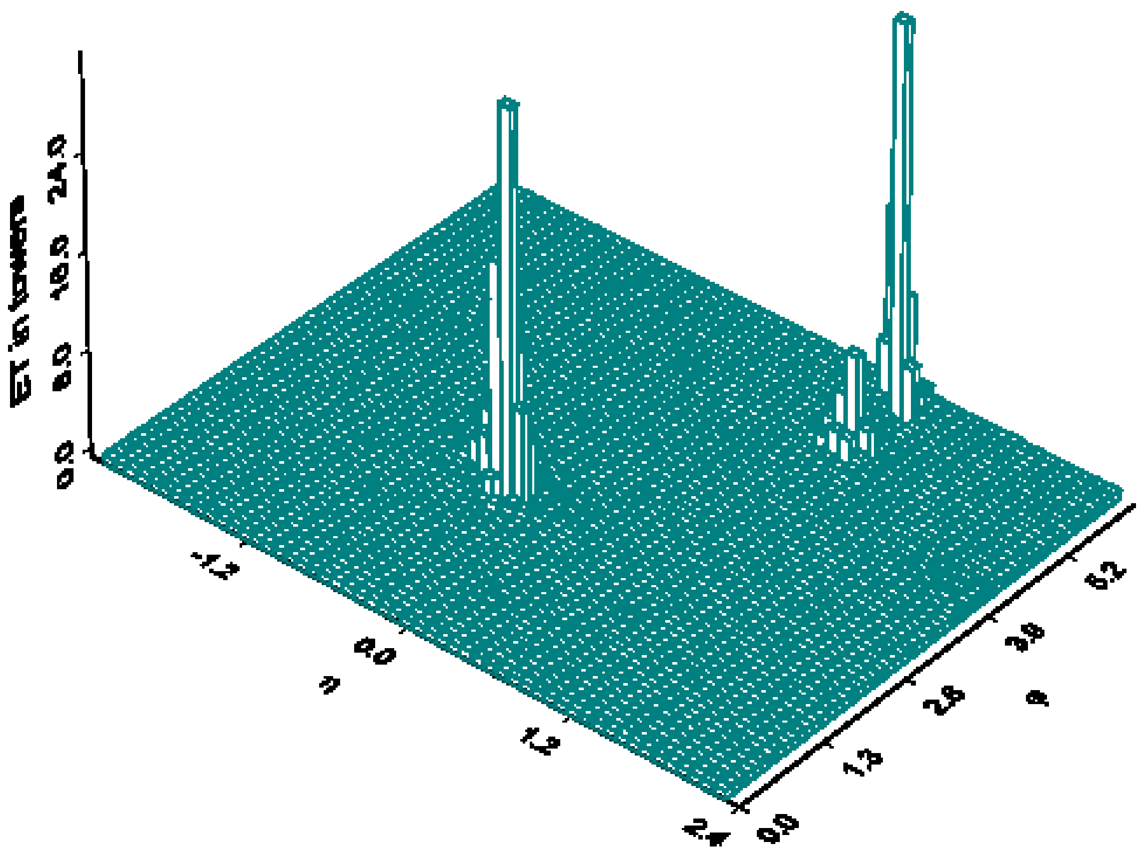
- Replace $1 - \cos\theta_{ij}$ by $\Delta R_{ij} = \sqrt{(y_i - y_j)^2 + (\phi_i - \phi_j)^2}$
- Distance $d_{ij} = \min(p_{t,i}^2, p_{t,j}^2)\frac{\Delta R_{ij}}{D^2}$ ($D$ adjustable parameter)
- Add to pairs also single particles $d_i = p_{t,i}^2$
- If minimum is a particle: Define as jet and remove from list
- If minimum is a pair combine and start again
- Stop if nothing left

### Features of the $k_T$ algorithm

- Every hadron is uniquely assigned to a jet
- Every hadron is assigned to a jet
  - few hadrons that belong to a given parton are missing
  - significant noise from underlying event and minimum bias
- Jets have complicated shapes

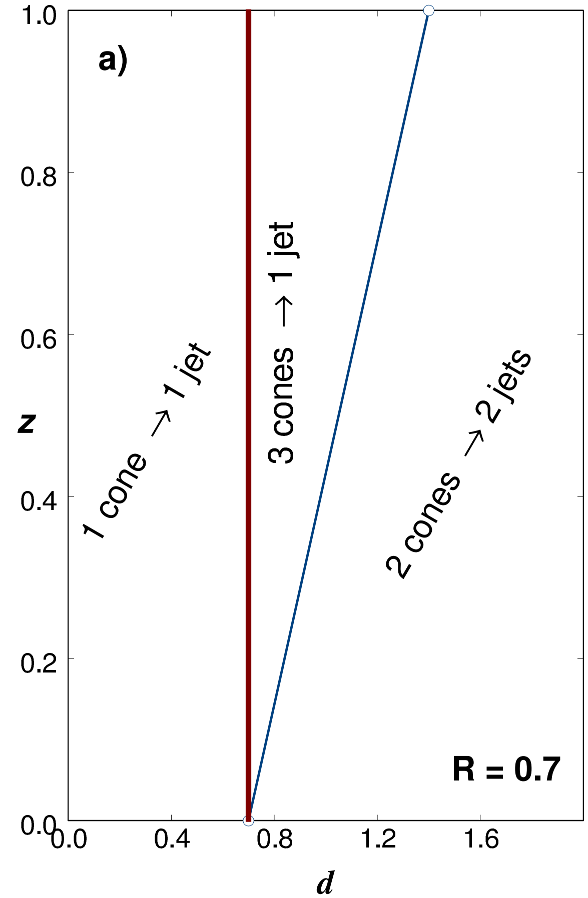
# Cone algorithms

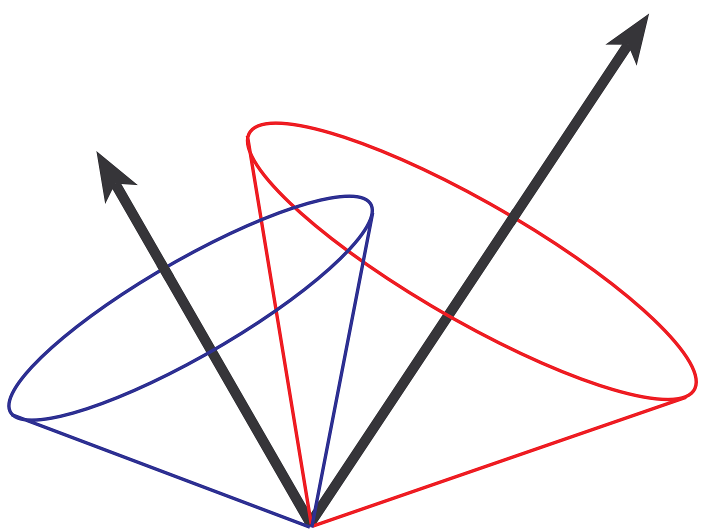
- Naively imagine a jet as a energy flow within a cone in $(y, \phi)$ space
- Consequently 1st $\mathrm{p\bar{p}}$ jet algorithms add energy within a cone
- Iterative procedure
  - Start with a cone containing some energy and opening angle $R$
  - Calculate the cone centre e.g. by adding 4-momenta
  - Recalculate energy in cone
  - Iterate until a stable cone is reached.

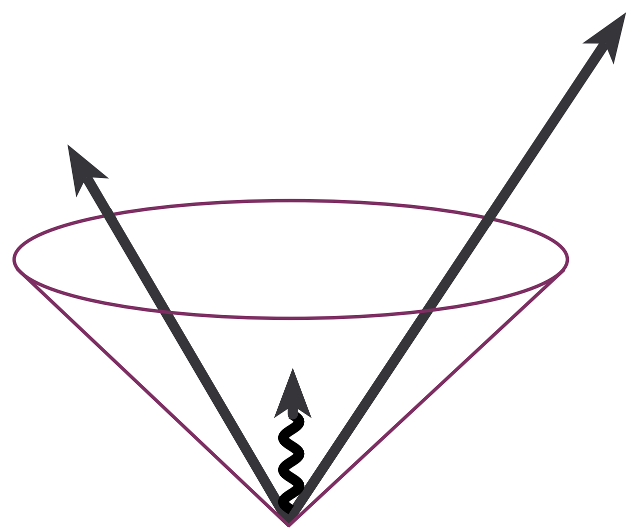


## Stability problems

- Solution is not unique
- Usually seeds are used in experiment ➡ infrared unsafe (partially solved by artificial seed between two real ones (midpoint algorithm))
- Large fragmentation corrections in cases where two jets are merged into one
- Jets may overlap and splitting procedure is deeded





- A new cone algorithm exists that is equivalent to a seedless one solving the theoretical problems (SISCone)
- Anyway it turns out that the theoretical uncertainties are only on the 10% level

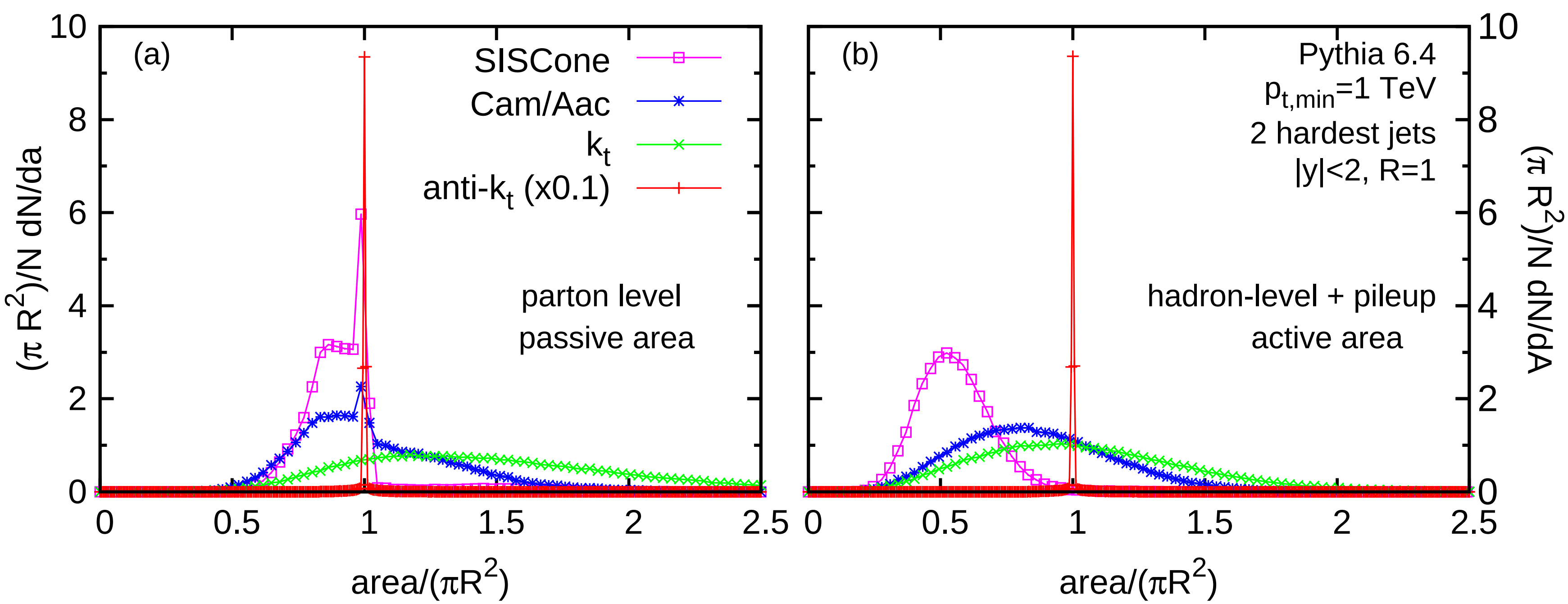
## Features of cone algorithms

- Low energy hadrons are not all included in jets
  - energy missing for event reconstruction
  - lot of underlying event/pileup rejected
- Jet shapes are usually round ➠ makes underlying event, pileup, noise corrections easier

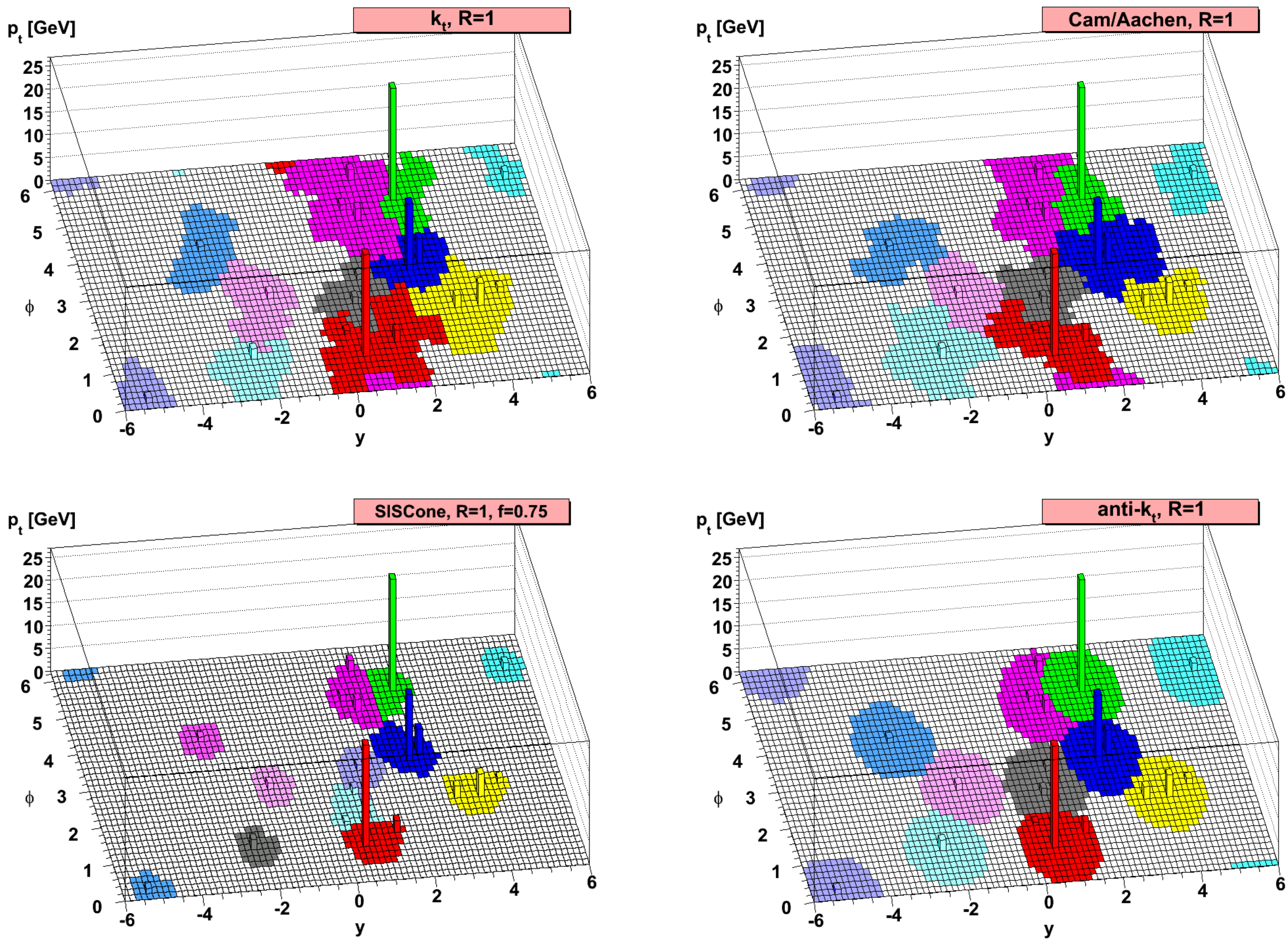
## New idea: anti $k_T$ algorithm

Define new distance measure: $d_{ij} = \min(p_{t,i}^{-2}, p_{t,j}^{-2})\frac{\Delta R_{ij}}{D^2}$

- First cluster high energy with high energy and high energy with low energy particles
- This keeps jets round, with well defines area
- Algorithm still infrared and collinear safe!

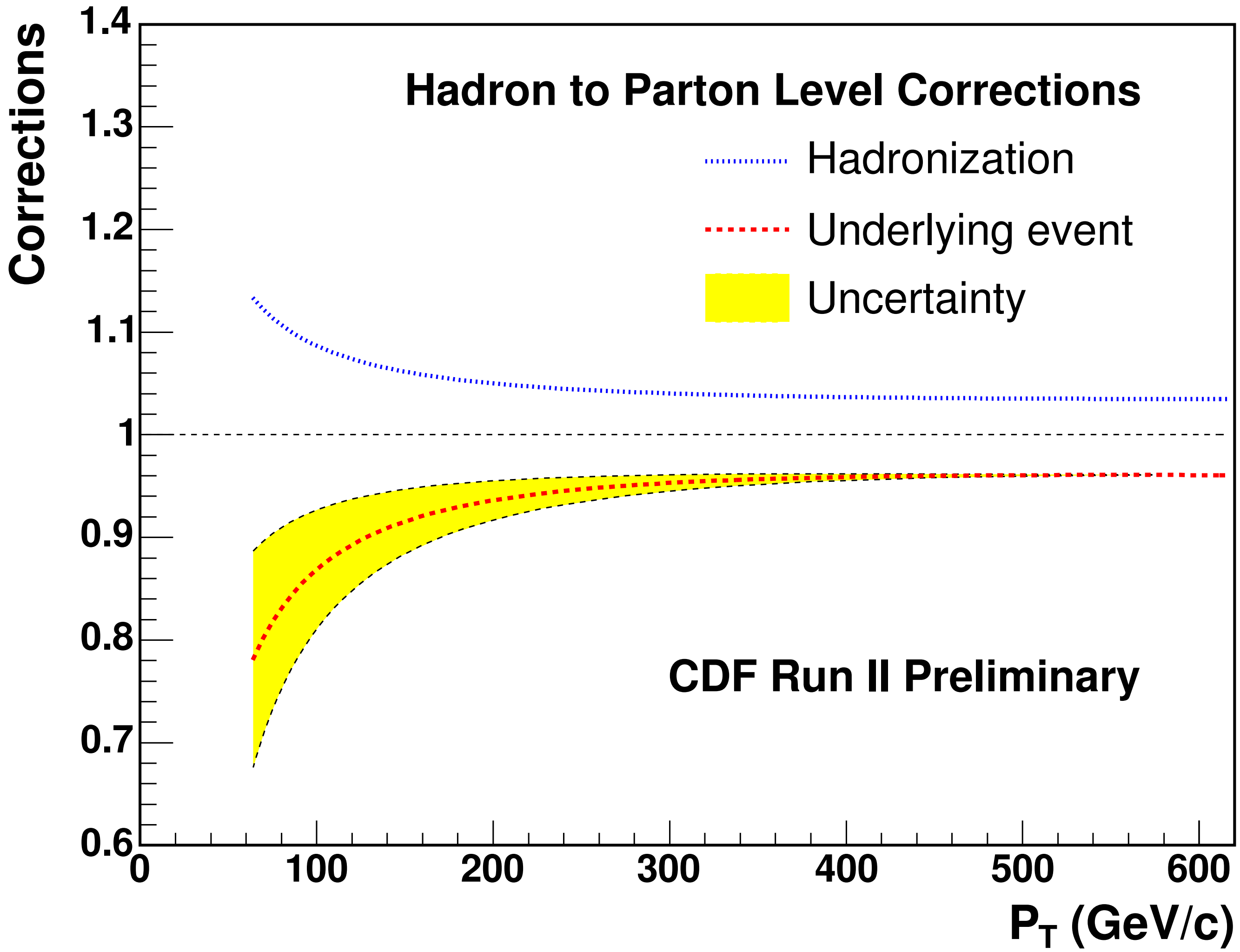


# Typical shapes for IR and collinear safe algorithms

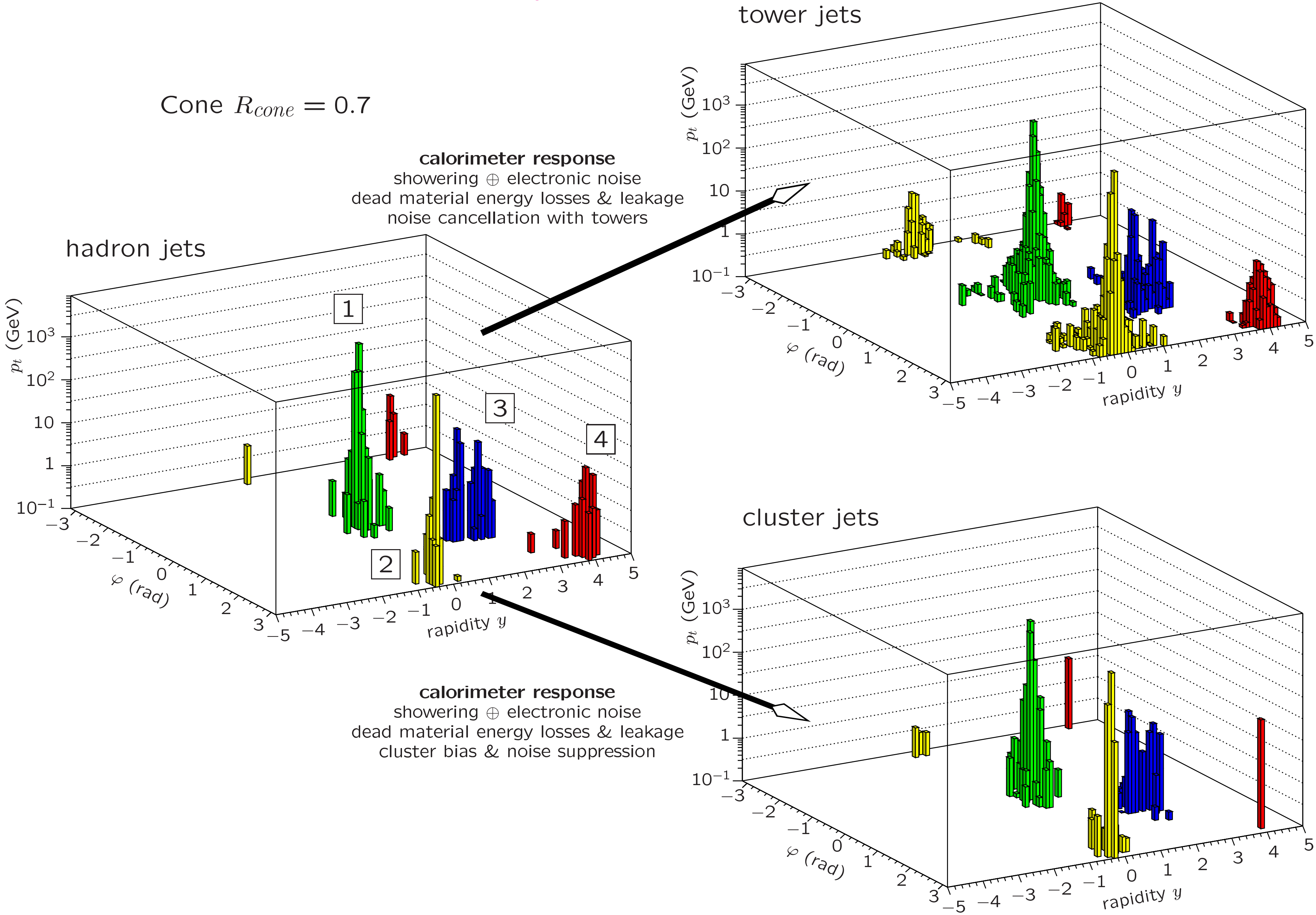


# Experimental issues

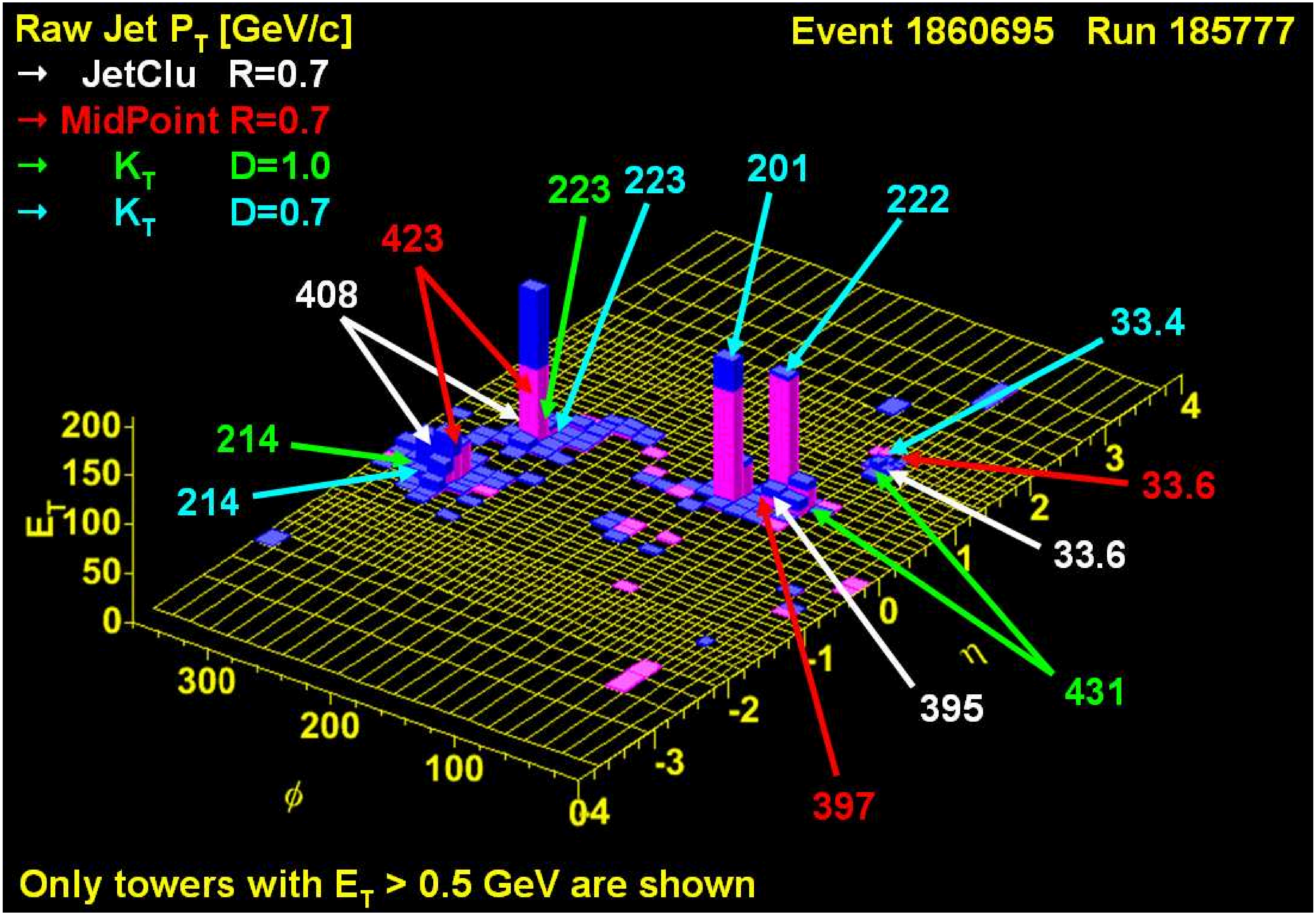
- Some part of the jet is outside the cone ➟ needs corrections
- Energy from the underlying event or pileup gets into the cone
- Treatment of noise in the calorimeter cells affects reconstructed jets



# Dependence of jets on calorimeter treatment



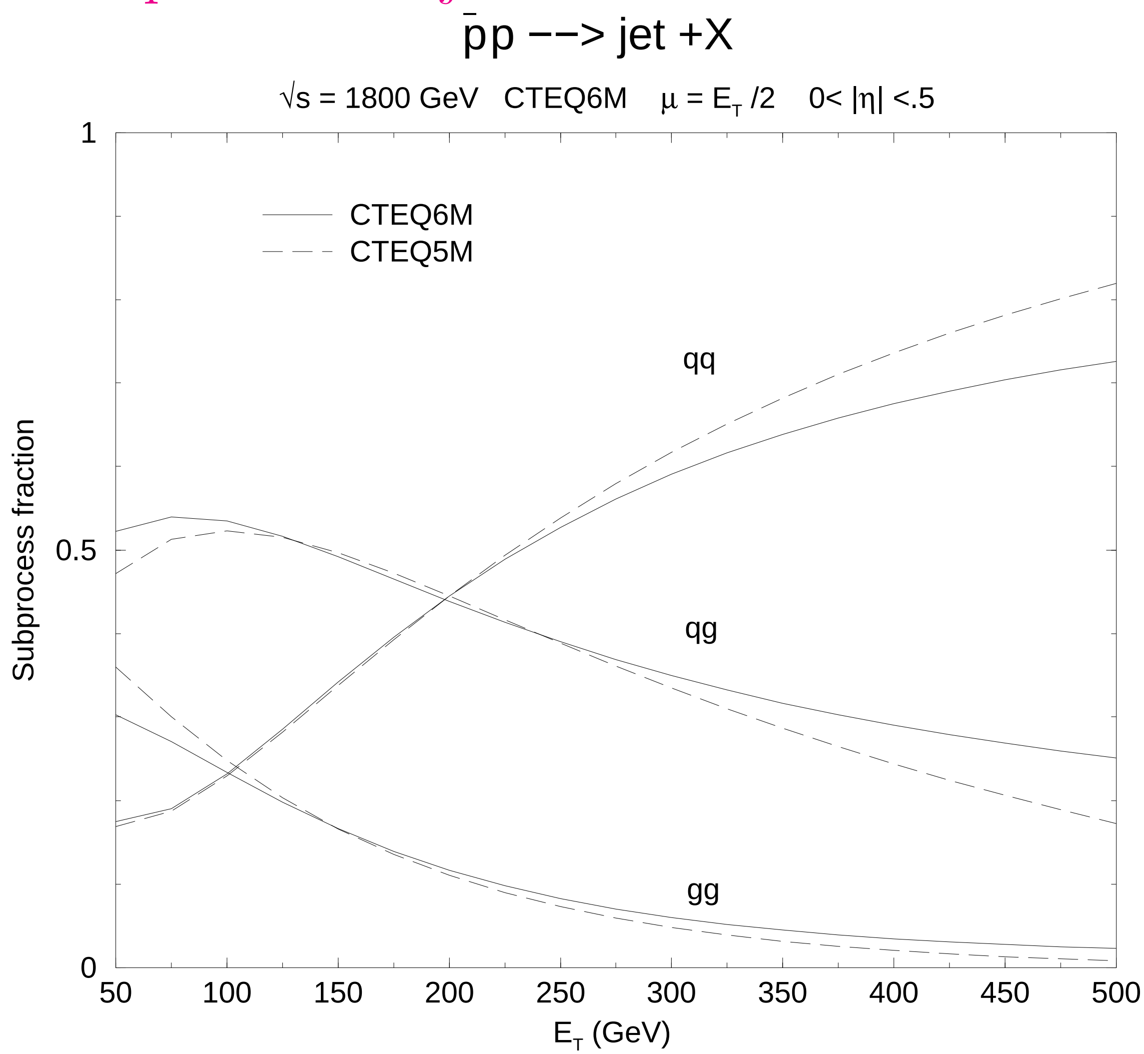
# Results of different jet algorithms for one CDF event



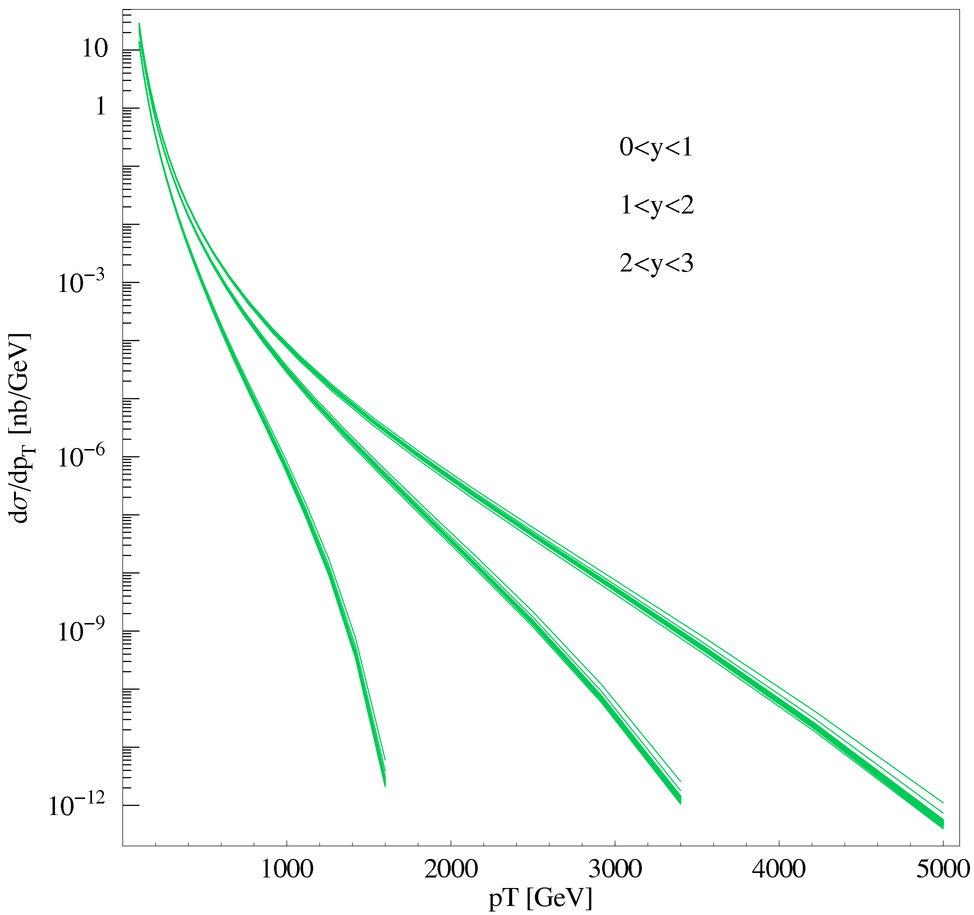
# QCD predictions for jet-rates

- Jet-events originate from gg, qg, qq scattering
- They can be calculated in QCD integrating over the PDFs
- At medium energies qg dominates, at high energies qq is dominant

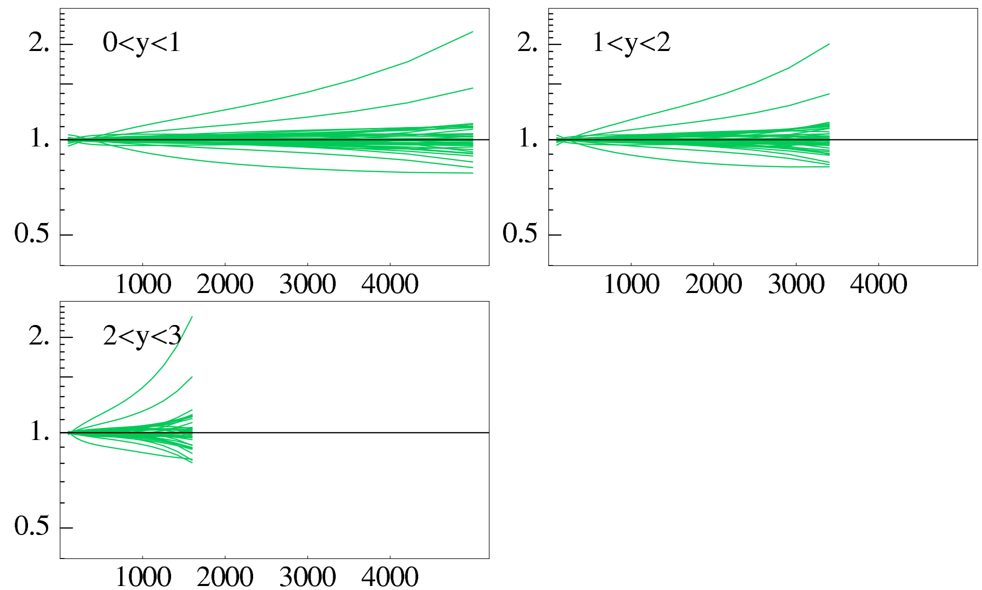
Composition of jet events at the Tevatron



## Jet cross sections at the LHC




## Uncertainties due to PDFs are of the order 20-30%

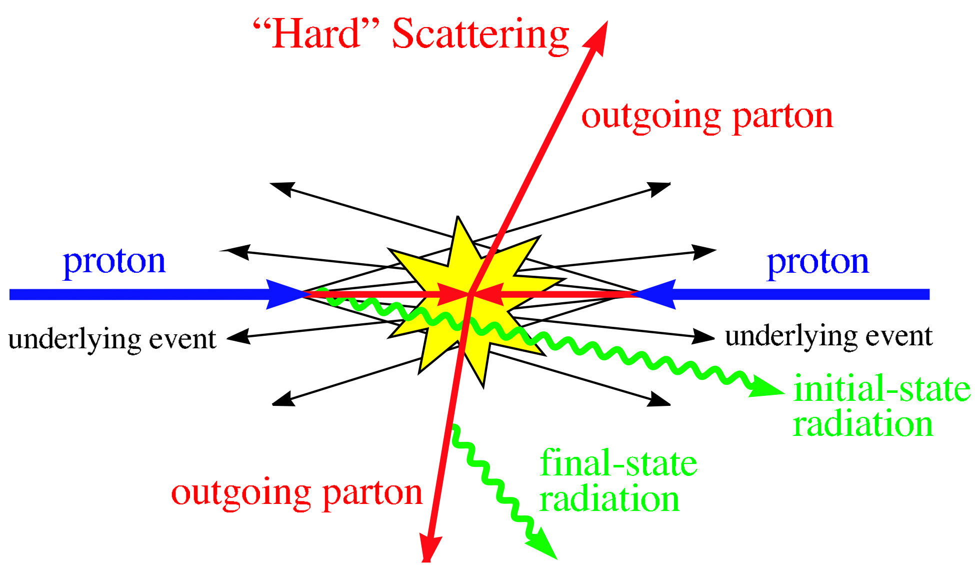



## Rates at $\mathcal{L} = 10^{34}\mathrm{cm}^{-2}\mathrm{s}^{-1}$

- 1000 ev./s for $p_T > 100$ GeV
- 1 ev./s for $p_T > 1$ TeV

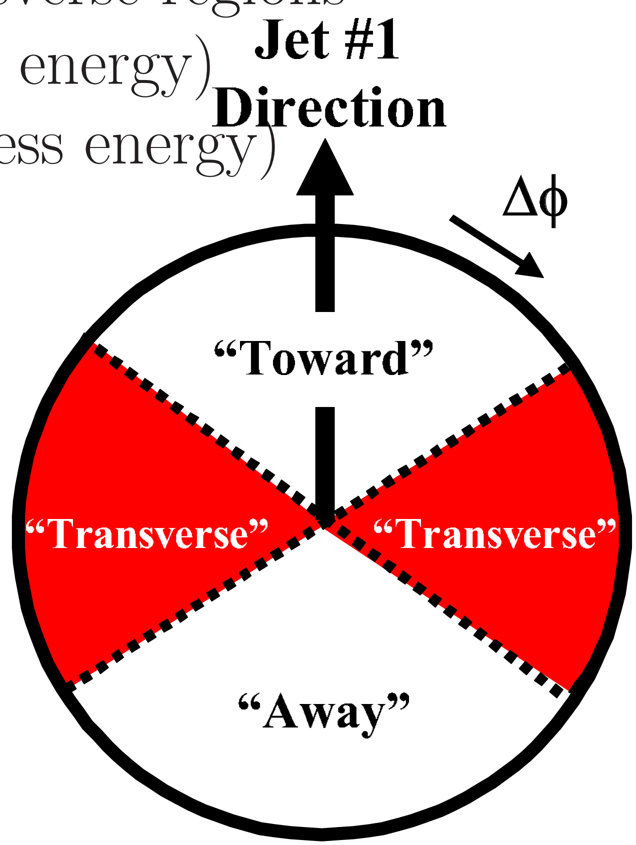
# The underlying event

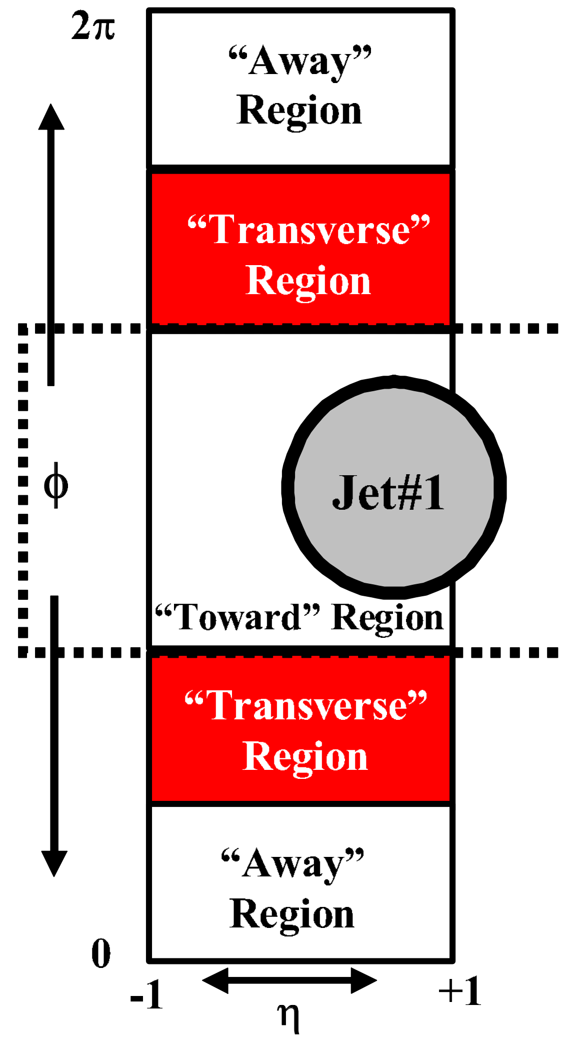
- At the LHC the parton-parton cross section integrated over the PDFs exceeds the proton-proton cross section
- This is interpreted as several parton-parton scatterings during one inelastic pp event
- There are indications that the hard partons concentrate in the core of the event
- For this reason the underlying event does not simply look like minimum bias



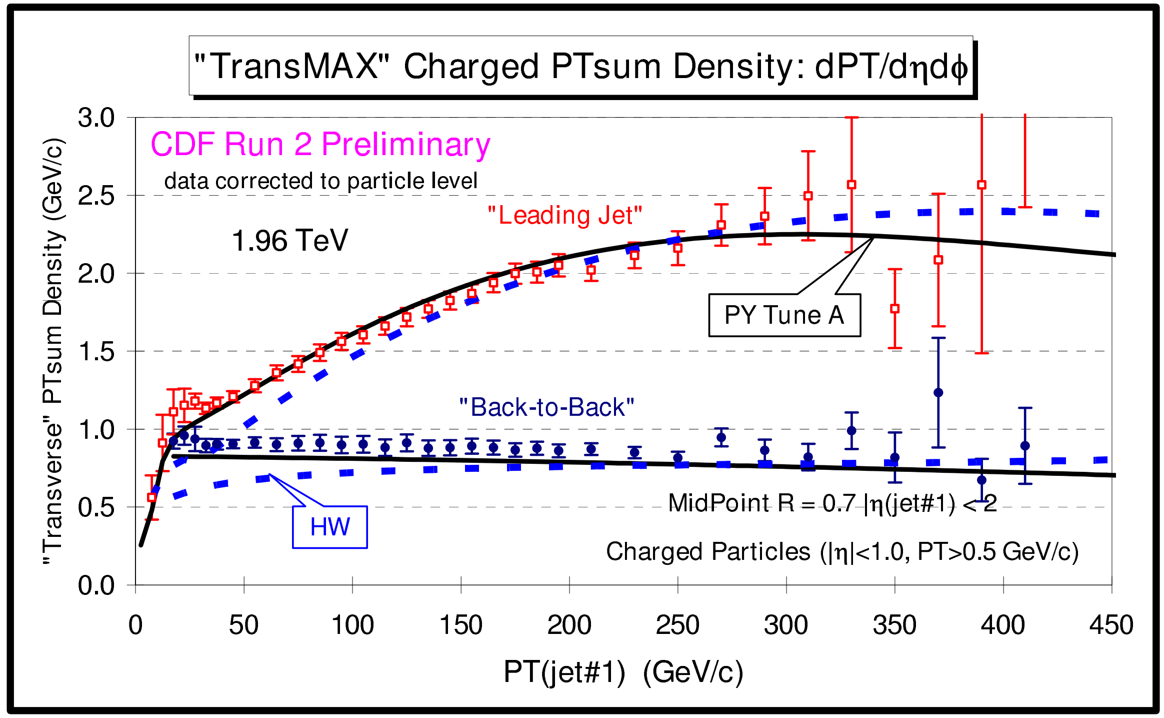
# Analysis of the underlying event:

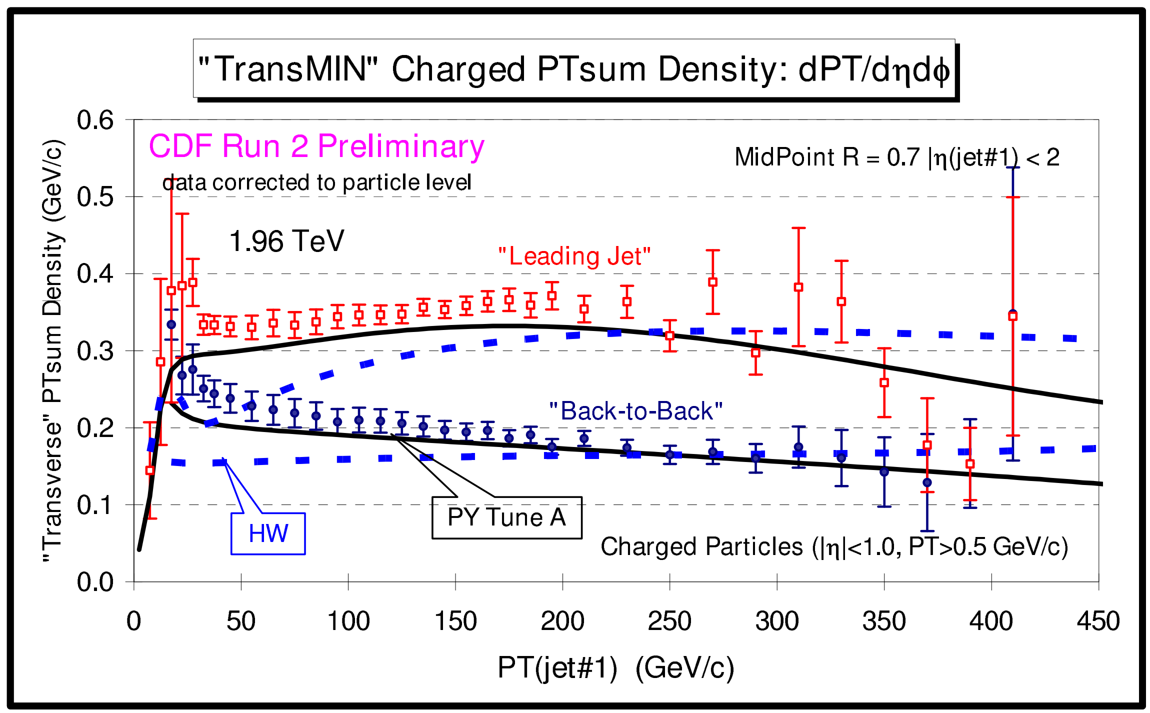
- 2-jet events are back-to-back
- Select events with one hard jet
- The opposite region should contain the 2nd jet
- Look into the transverse regions
  TRANSmax (more energy)
  and TRANSmin (less energy)



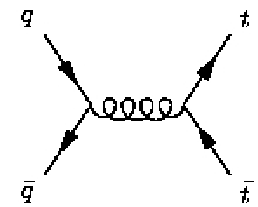



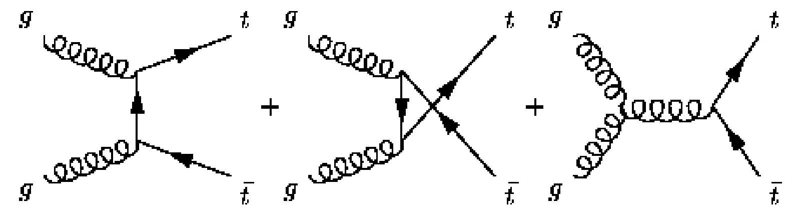
- Generators without multi-parton interactions cannot describe the data
- However generators with UE can be tuned to agree with the data
- Warning: it cannot be expected that the Tevatron tunes describe the LHC data



# Top as example for jets at the LHC




Top pair production:

- $\mathrm{q\bar{q}}$ and gg, gg largely dominates at LHC

Top cross section $\sim 800\,\mathrm{pb}$
$\Rightarrow$ 80 000 000 events in 100 $\mathrm{fb}^{-1}$

Top decays: $t \to bW$ (99.8%) with

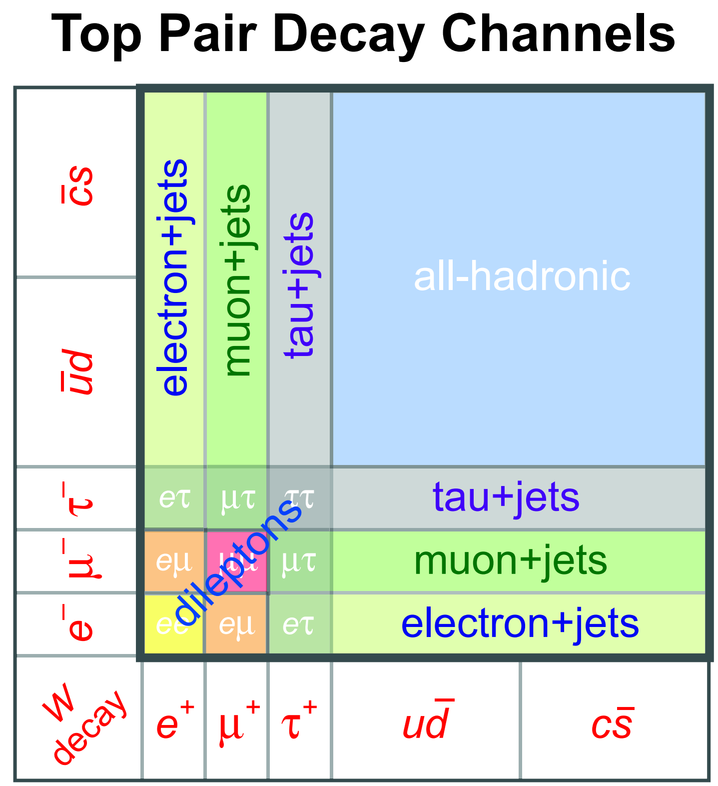
$$
\begin{array}{ll}
W \to \mathrm{q\bar{q}} & 2/3rd \\
W \to \ell\nu,\ \ell = e,\mu & 2/9th \\
W \to \tau\nu & 1/9th
\end{array}
$$

(Rest is $t \to s, d\,W$)

➠ Always have jets in top events

➠ Always have b-quarks

This means for $t\bar{t}$ events:




- 45% 2 b-jets + 4 light jets
  - everything can be reconstructed
  - however large pairing ambiguities
  - large QCD backgrounds
- 30% 2 b-jets, 2 light jets, $1\ell$ , $1\nu$
  - the neutrino can be reconstructed with a 2-fold ambiguity using the W-mass
  - pairing ambiguities are less
  - lepton strongly suppresses QCD background
- 5% 2 b-jets, $2\ell$ , $2\nu$
  - clean samples
  - however few constraints for reconstruction
- Rest contains $\tau$s ➟ difficult

## Intermezzo: b-tagging at colliders

- b-quarks decay semileptonically with $BR(b \to \ell X) = 2 \times 10\%$
  - can be used for b-tagging
  - however low efficiency from the beginning
  - leptons inside jets where fake rate is high
- b-quarks have significant lifetime ($c\tau \sim 0.5$mm)
  - e.g. flight distance of 50 GeV B-meson: $\sim 5$mm
  - impact parameter w.r.t. primary vertex: $\sim 500\mu$m
  - resolution of modern vertex detectors: $\sim 10\mu$m

➡ can use vertexing for b-tagging

# Impact parameter methods

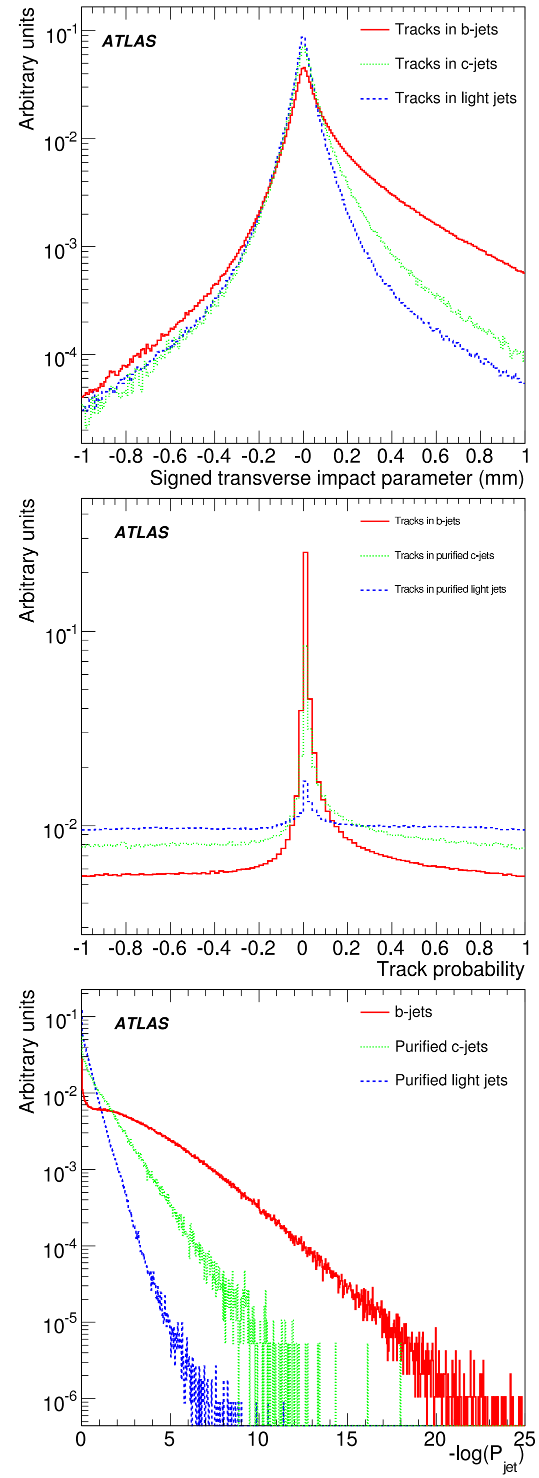
- Signed track impact parameter gives already good sensitivity for b-tagging
- Calculate probability for optimal use

$$\mathcal{P}_i = \int_{d_i/\sigma_{d,i}}^{\infty} \mathcal{R}(x)dx$$

- Tracks in a jet/event can be combined

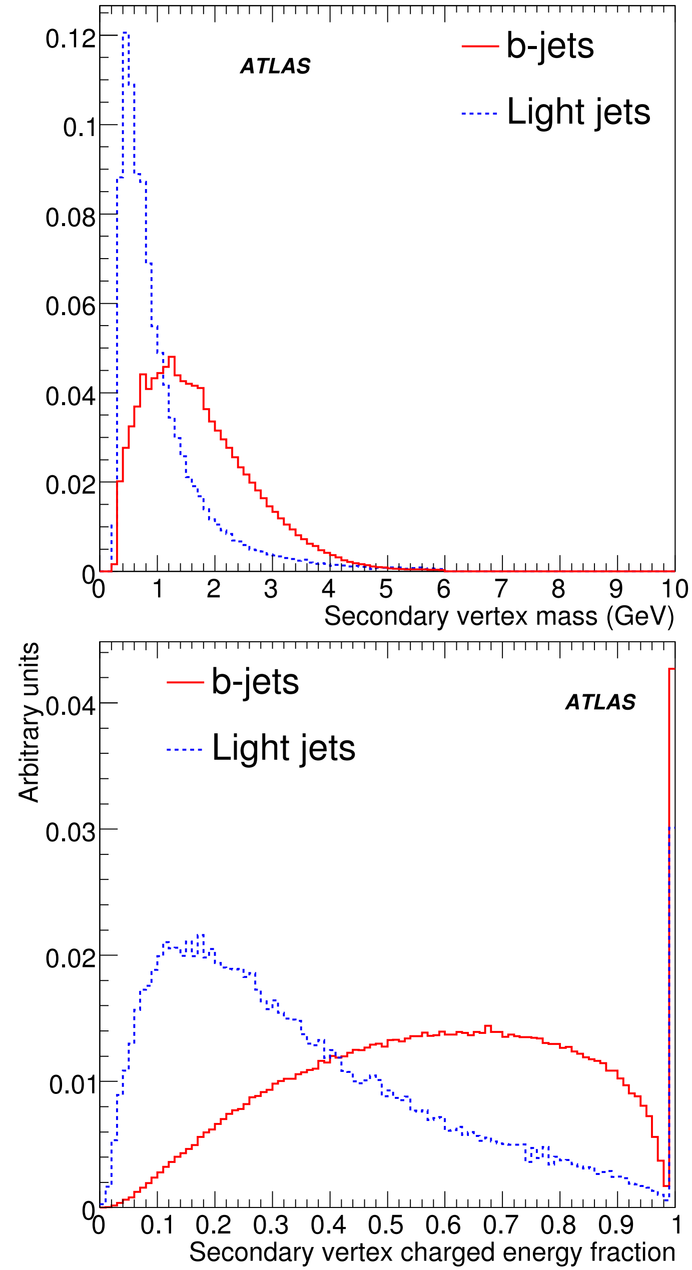
$$\mathcal{P}_0 = \prod_i^N \mathcal{P}_i \qquad \mathcal{P} = \sum_0^{N-1} \frac{(-\ln \mathcal{P}_0)^j}{j!}$$

- In principle this method gives an optimal separation of b- and light jets
- However very sensitive understanding of tracking

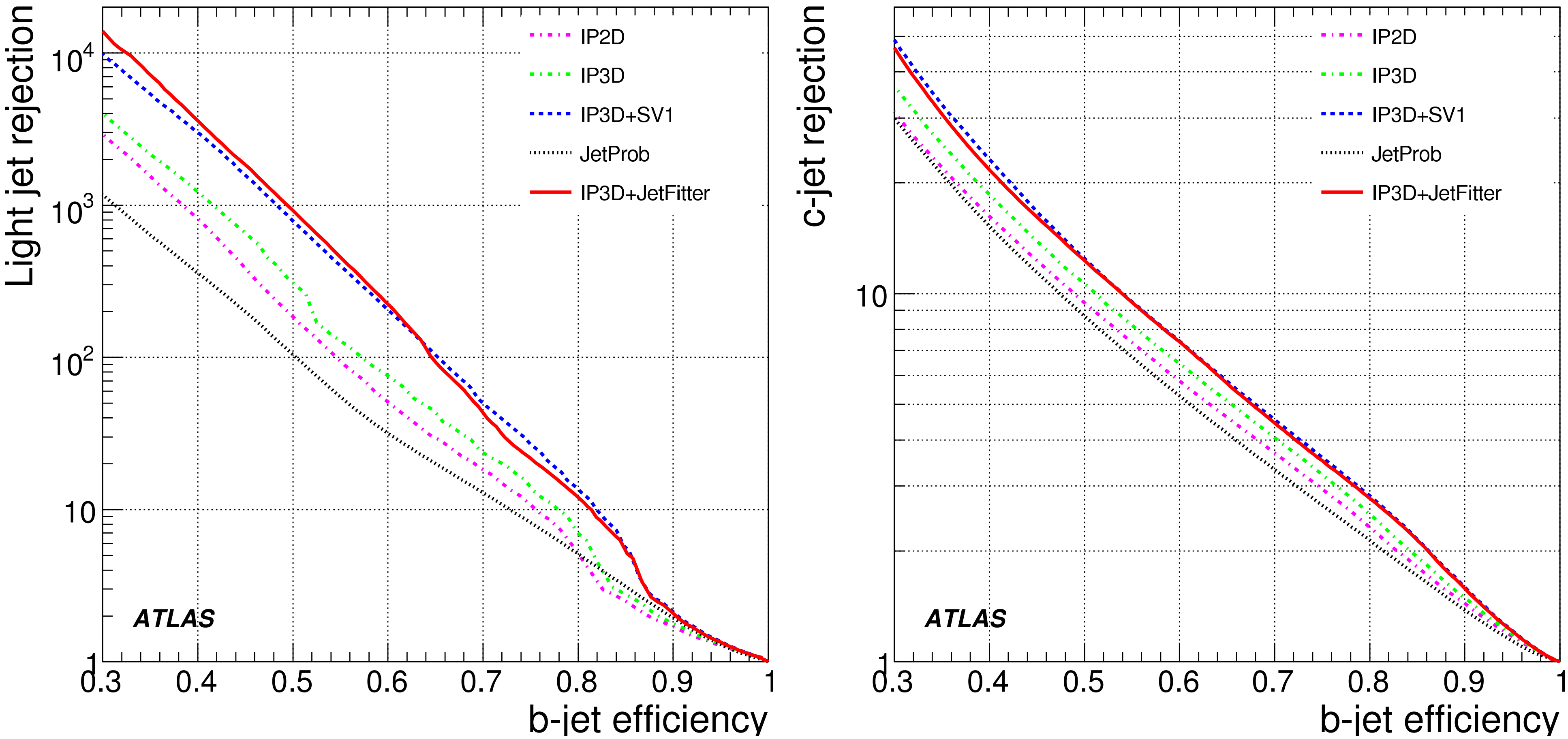


# Enhancement/alternative: secondary vertices

- Secondary vertices are not faked so easily by reconstruction problems
- The vertex mass gives a good separation especially to c-quarks
- The energy of the fitted particles normalised to the jet energy makes use of the hard b-fragmentation (average B energy is $\sim 80\%$ of jet energy)



With the available methods e.g. a light quark rejection of $10^3$ and a c-quark rejection of 10 can be achieved for 50% b-efficiency ($t\bar{t}$ events)



# Selection of $\mathrm{t\bar{t}}$ events

Concentrate on mixed decays:

- $E_{T,miss} > 20\,\mathrm{GeV}$ (neutrino!)
- 1 isolated lepton with $P_T > 20\,\mathrm{GeV}$
- 2 b-jets with $p_T > 40\,\mathrm{GeV}$ and $\geq 2$ light jets with $p_T > 40\,\mathrm{GeV}$

(At the beginning of data taking b-tagging can be dropped at the price of a larger background)

Hadronic W reconstruction

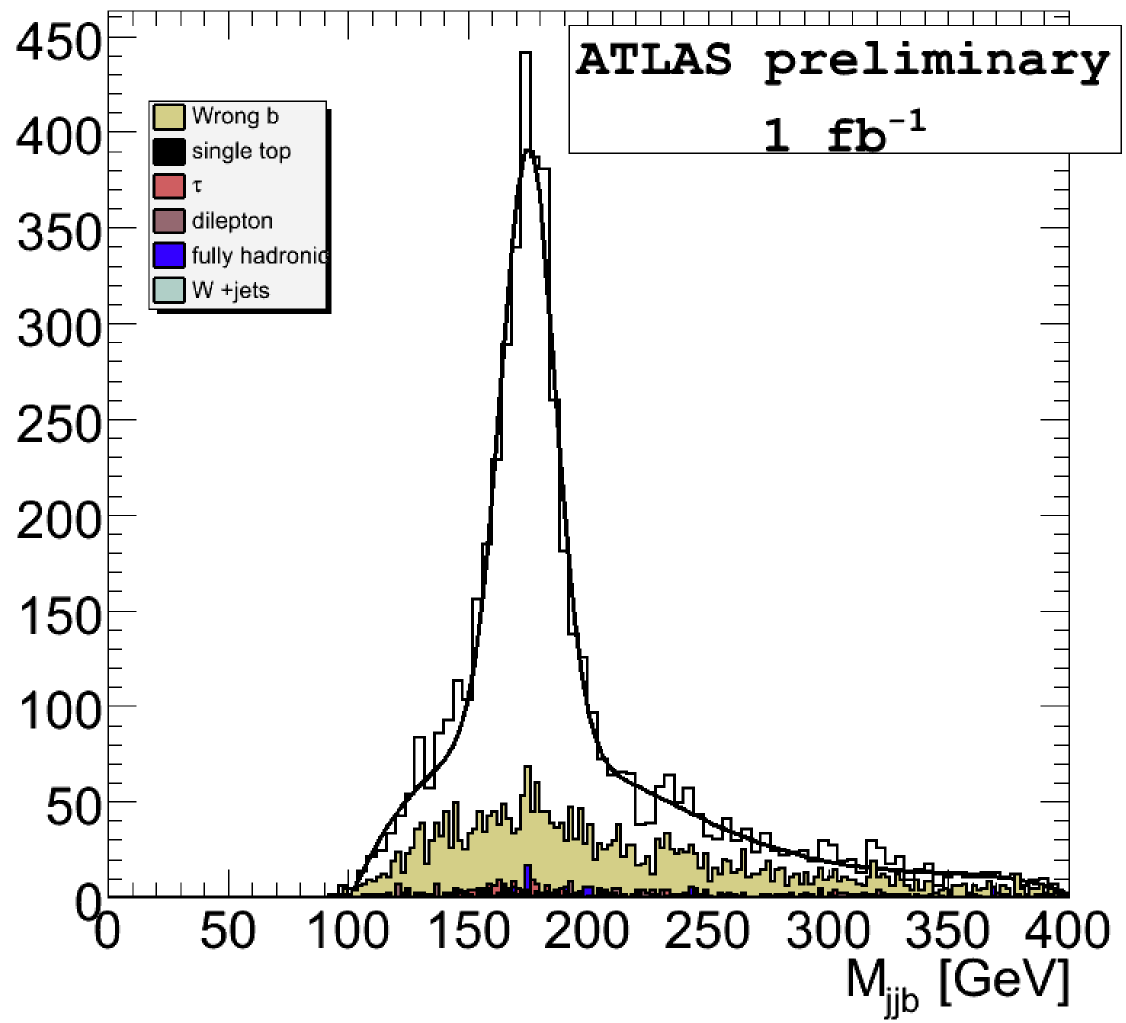
- Accept light jet pair if consistent with $m_\mathrm{W}$ at $3\sigma$
- Rescale jets to $m_\mathrm{W}$ using a $\chi^2$ technique
- Cut again on jet-jet mass around $m_\mathrm{W}$

W-b association

- Take the combination with minimal $\Delta R$

# Top mass distribution for the top-mass analysis

- Selected top-sample has very little non-$t\bar{t}$ background
- For mass determination most serious background is combinatorial background

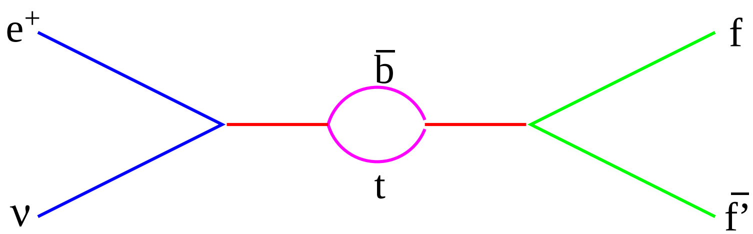


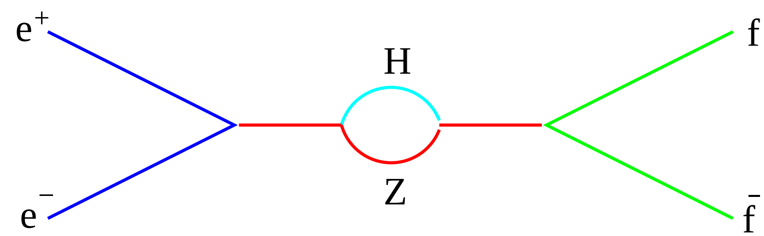
# Measurements of the top-quark mass

## Why is the top mass interesting?

## SM:

- Electroweak precision data are affected by loop corrections

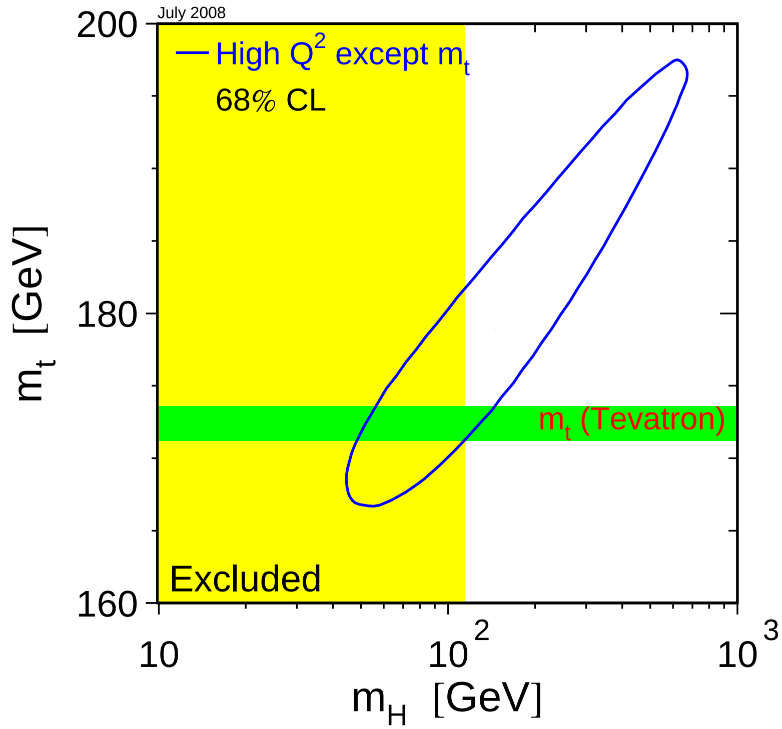


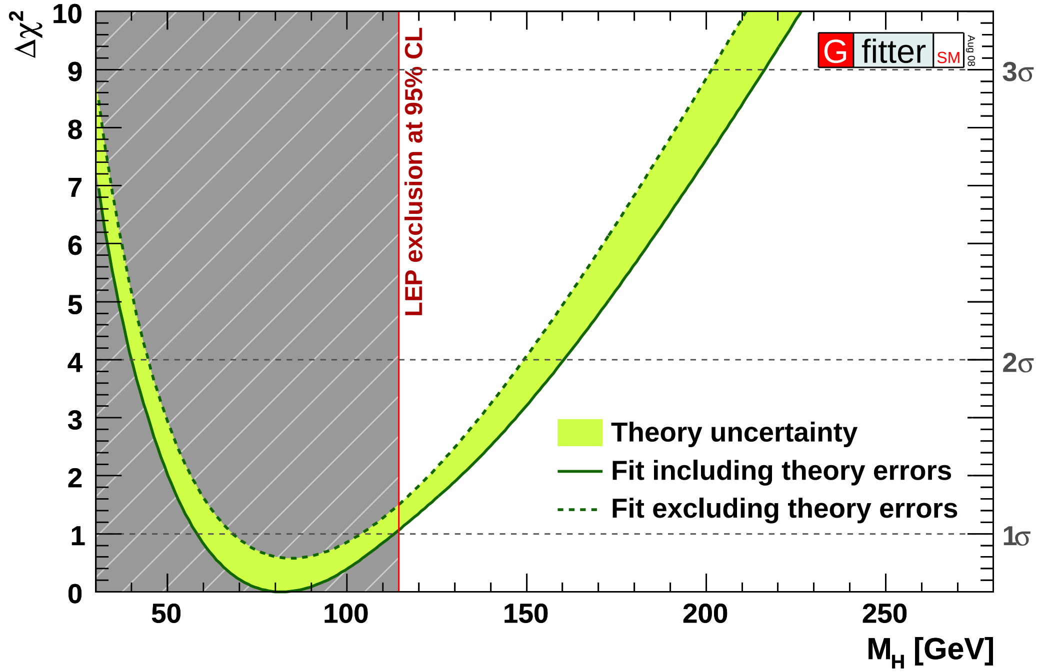




e.g. $m_{\mathrm{W}}^2 = \frac{1}{2} m_{\mathrm{Z}}^2 \left( 1 + \sqrt{1 - \frac{4\pi\alpha}{\sqrt{2} G_{\mathrm{F}} m_{\mathrm{Z}}^2} \frac{1}{1 - \Delta r}} \right)$

- Can be used e.g. to constrain $m_{\mathrm{H}}$
- Top-quarks corrections are quadratic ➠ need to be known to get useful results

# For this need $m_t$ measurement of $\mathcal{O}(1\,\mathrm{GeV})$




# Beyond SM

- Some models like SUSY predict the Higgs mass from the model parameters
- Here the $m_t$ corrections can be of order $\Delta m_H/\Delta m_t \sim 1$

➠ In principle a much better top mass is needed

Current uncertainty from the Tevatron: $\Delta m_{\mathrm{t}} = 1.2\,\mathrm{GeV}$

This does not yet include

- errors from colour reconnection effects
- uncertainties from the mass definition

which might add up again to $1\,\mathrm{GeV}$

Expectation at the LHC: $\Delta m_{\mathrm{t}} \lesssim 1\,\mathrm{GeV}$

- totally dominated by systematics
- largest experimental uncertainty: energy scale of b-jets
- errors from QCD might be of similar size

# Search for rare top decays

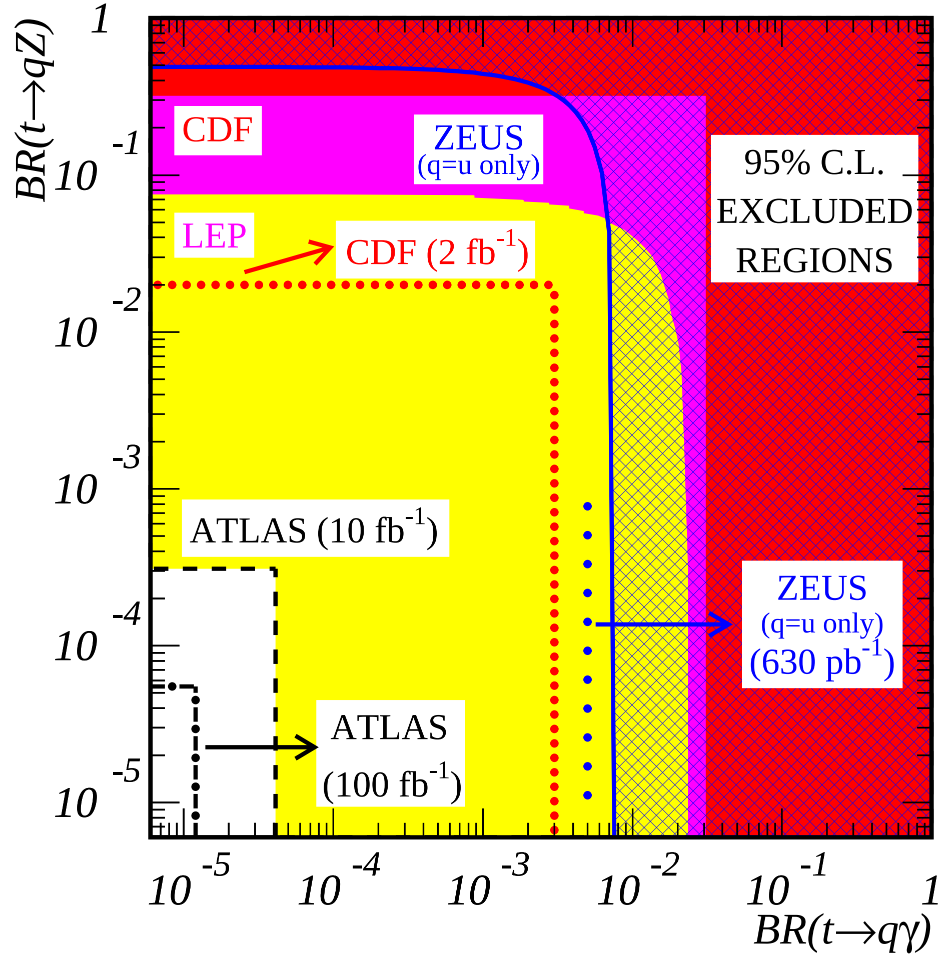
The SM predicts FCNC top decays ($t \to Zq$, $\gamma q$, $gq$) on the $10^{-14} - 10^{-12}$ level

In some new physics scenarios $10^{-4}$ can be reached

Example for a $t \to Zq$, $Z \to \ell\ell$ selection:

| Description of Cuts | Signal $t \to Zq$ $\varepsilon$ (%) | Background Processes $Z$+jets Nevt | $Z+W$ Nevt | $tt$ Nevt |
|---|---|---|---|---|
| Preselection 3 leptons, 2 jets | 80.2 | $3.7\times10^5$ | 2941 | $11.7\times10^5$ |
| 3 leptons, $p_T^\ell > 20$ GeV/$c$ | 43.3 | 945 | 1778 | 1858 |
| $\not{p}_T > 30$ GeV | 32.7 | 80 | 1252 | 1600 |
| 2 jets, $P_T^{jet} > 50$ GeV/$c$ | 19.8 | 31 | 225 | 596 |
| $m_Z \pm 6$ GeV | 16.8 | 24 | 180 | 29 |
| one $b$-tag | 8.2 | 10 | 28 | 10 |
| $m_t \pm 24$ GeV | 6.1 | 0 | 2 | 5 |

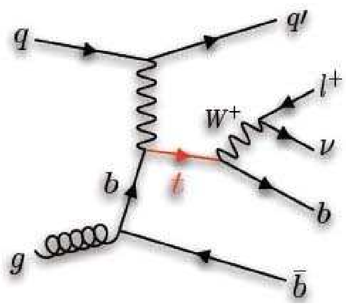
Even the LHC can only scratch the interesting region!



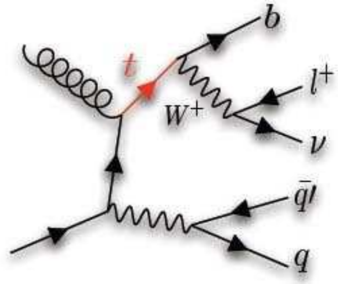
# Measurement of single top production

## Feynman graphs

t-channel

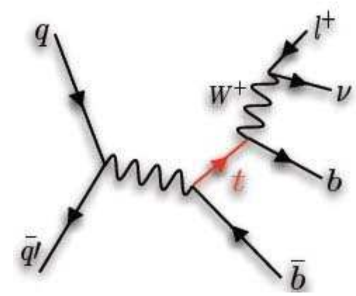



$\sigma = 260\,\mathrm{pb}$




$\sigma = 60\,\mathrm{pb}$

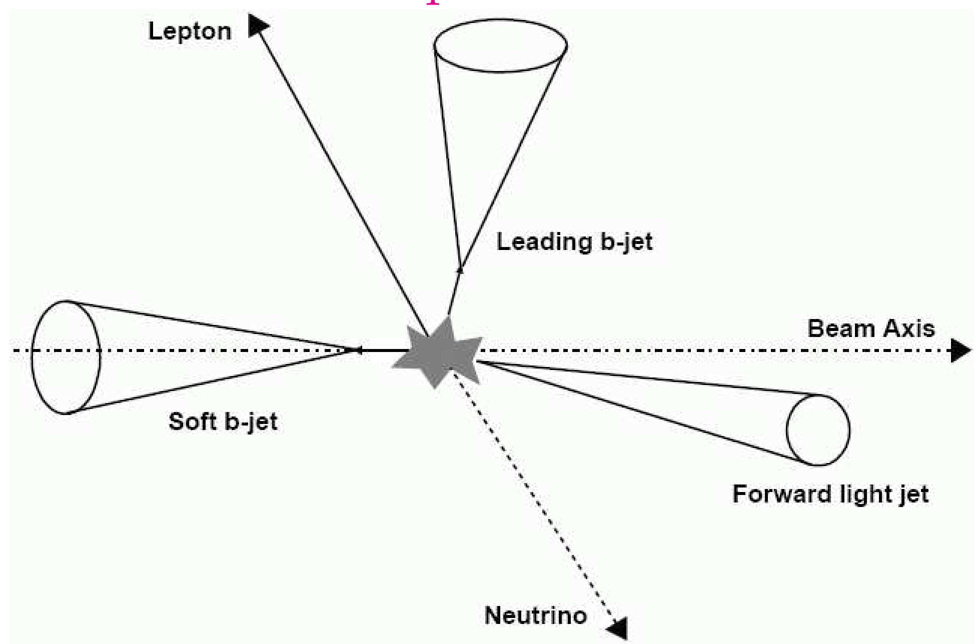
s-channel




$\sigma = 10\,\mathrm{pb}$

All channels are sensitive to $V_{tb}$

## t-channel analysis




- Special cuts for event topology
- Main background is $\mathrm{t\bar{t}}$ and some W+jets
- Sample analysis gave $\varepsilon = 1.4\%$ and S/B=0.8
- Statistical error no problem, however systematics can be in 20% region

# Conclusions of 3rd lecture

- Quarks and gluons always end up in jets
- Most interesting physics at the LHC involves final state quarks (and gluons) ➟ jets
- There is always an arbitrariness in the definition of jets
- Top quark production is an example of jet production at the LHC