Simple and Complex Objects: Strategies for Event Reconstruction at the LHC

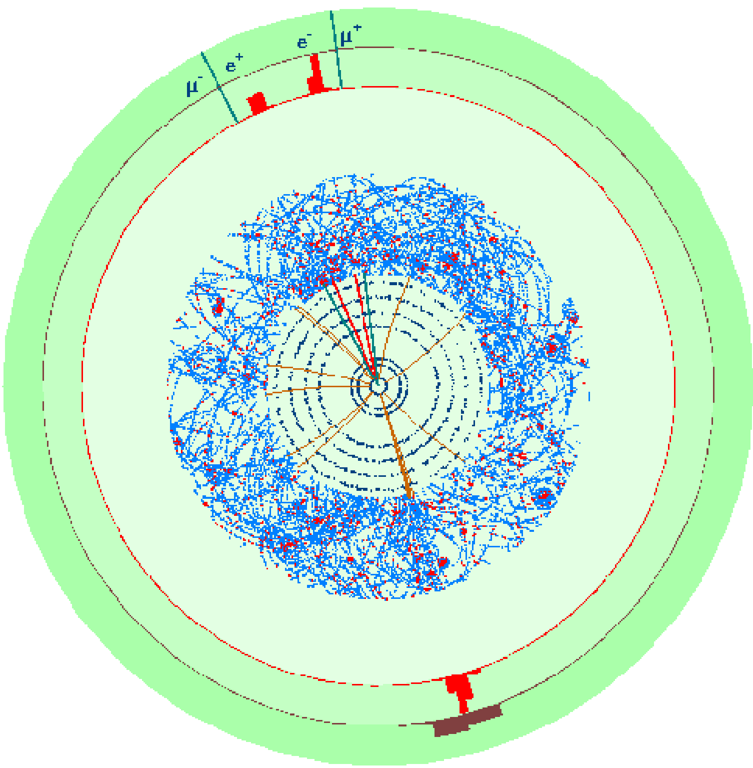
Lecture I: Reconstructing Simple Objects

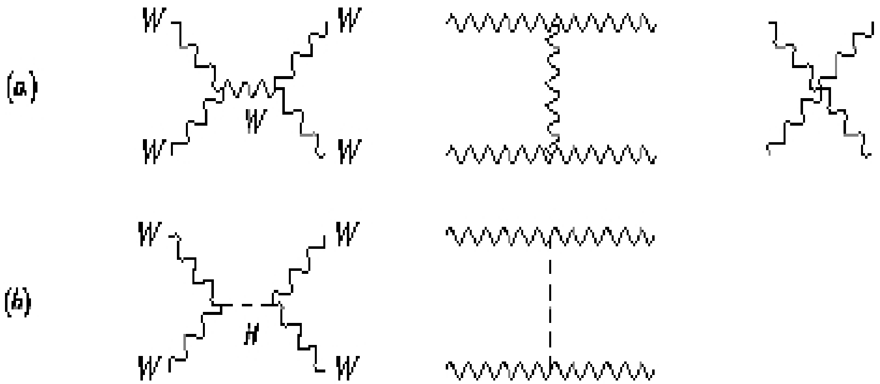
Marjorie Shapiro
UC Berkeley/LBNL
August 10

The Experimental Challenge:

Translate From:

To:





Pattern of digitized hits in a complex detectors

Interpretation in terms fundamental hard scattering process

Overview of Lectures:

Lecture 1: Reconstruction of Simple Objects

- Collider Basics
  - Rates and Cross Sections
  - Choice of Coordinates
  - Min Bias, Underlying Event
- Reconstruction Strategies
- Object Reconstruction Part I
  - Tracks
  - Jets
  - Charged Leptons: electrons

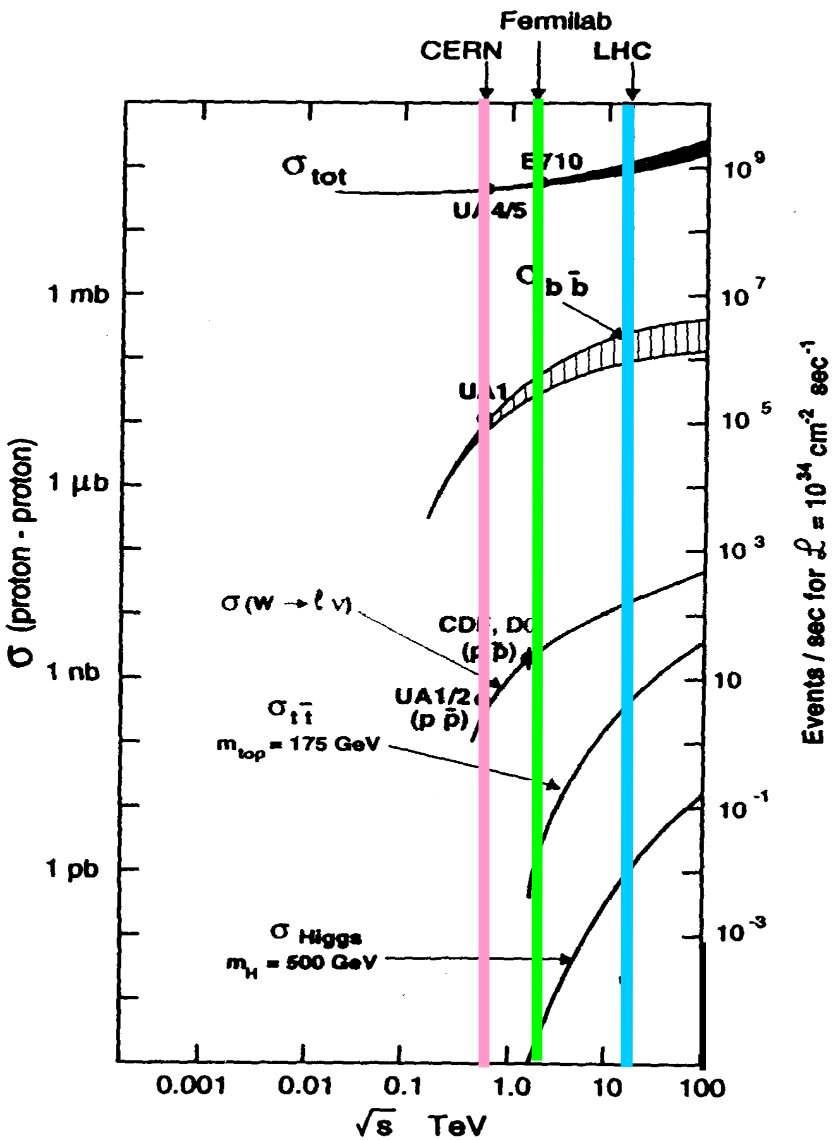
Lecture 2: Confronting the SM

- Object Reconstruction Part II
  - Charge Leptons: μ and τ
  - Neutrinos (and LSP)
- Finding W's and Z's
- Top

Lecture 3: TeV Scale Physics

- Object Reconstruction Part III
  - Photons
- Higgs
- SUSY

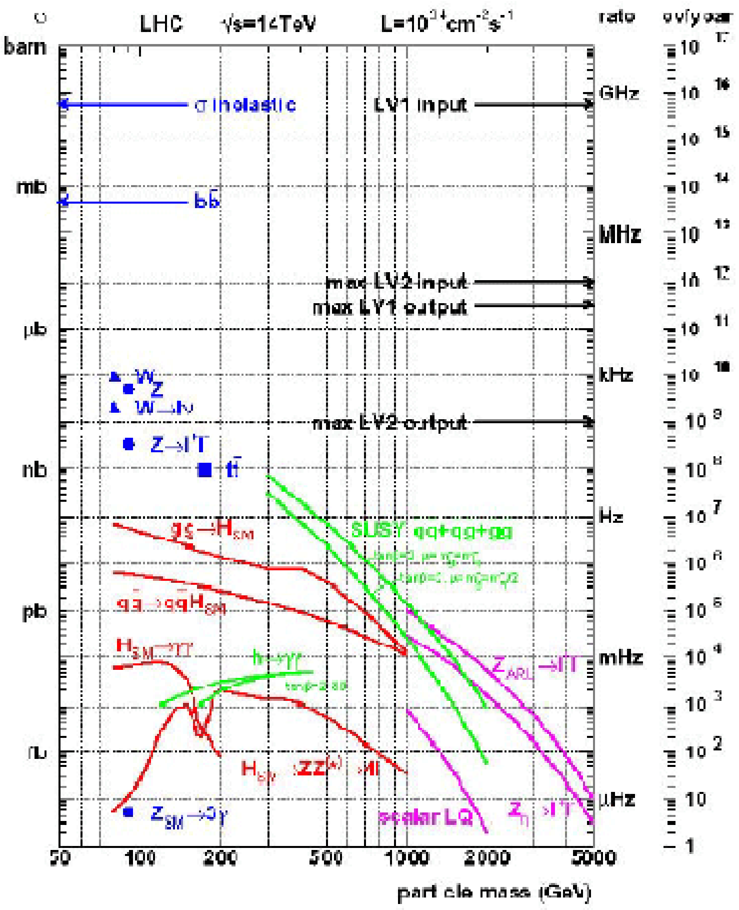
Collider Basics I: Cross Sections



- Rates Determined by:
  - Hard Scattering σ
  - Parton Luminosity
- QCD Processes Dominate
  - EW rates lower by α / α_{strong}
- Cross Sections Decrease Rapidly with s
  - Heavy particles difficult to produce

Implications for LHC

- Something happens every crossing
  - 25 inelastic evts/crossing at 10^{34} "Pile-up"
- Must Select Events of Interest: Trigger
  - Must know what you've thown out
  - Analysis must be trigger-aware
- Jets Dominate Hard Scattering Rate
  - Can isolate EW processes <u>only</u> they have something besides jets, eg leptons
  - Potential source of bckgnd "Fakes"
  - Detector mis-measurements can induce false signals
- W, Z: Bckgnd for Top, Higgs, SUSY

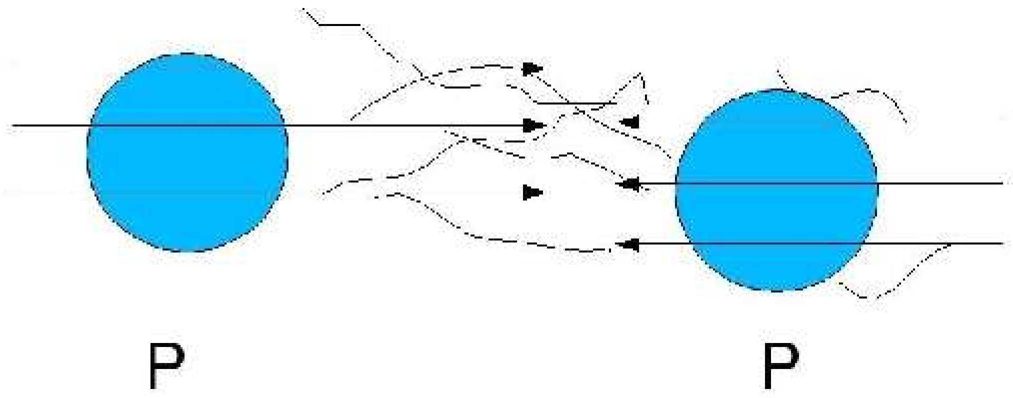


Analysis Strategy: Begin With Largest Cross Section and Work Down

- Characterize Bulk of Cross Section "Soft Physics"
  - Tracks
- Identify Dominant 2 $\rightarrow$ 2 QCD Processes
  - Jets
- Develop Strategies for Selecting EW Processes
  - e, μ, τ, ν. γ
- Reconstruct Heavy Objects Produced Strongly
  - Top, SUSY(?)
- Understand Discovery Potential for Low Rate EW Processes
  - Higgs

Soft Physics

- Bulk of Inelastic Cross Section: Large Impact Parameter, Soft Collisions
- Low Momentum Transfer $\rightarrow$ Cannot Use Perturbative QCD
  - Fireballs
  - Regge Theory
  - Multiple Parton Interactions
- Qualitative Features:
  - Limited P_T wrt Beamline
  - Longitudinal Distribution Dominated by Phase Space



Consider the Invariant Phase Space Factor:

$$\frac{d^3P}{E} = d\phi \frac{dP_T^2}{2} \frac{dP_{||}}{E}$$

$$E\frac{d\sigma}{d^3P} = \frac{1}{\pi}\frac{d\sigma}{dP_T^2\, dy}$$

Where: $$y = \frac{1}{2}\ln\left(\frac{E+P_{||}}{E-P_{||}}\right)$$ "rapidity"

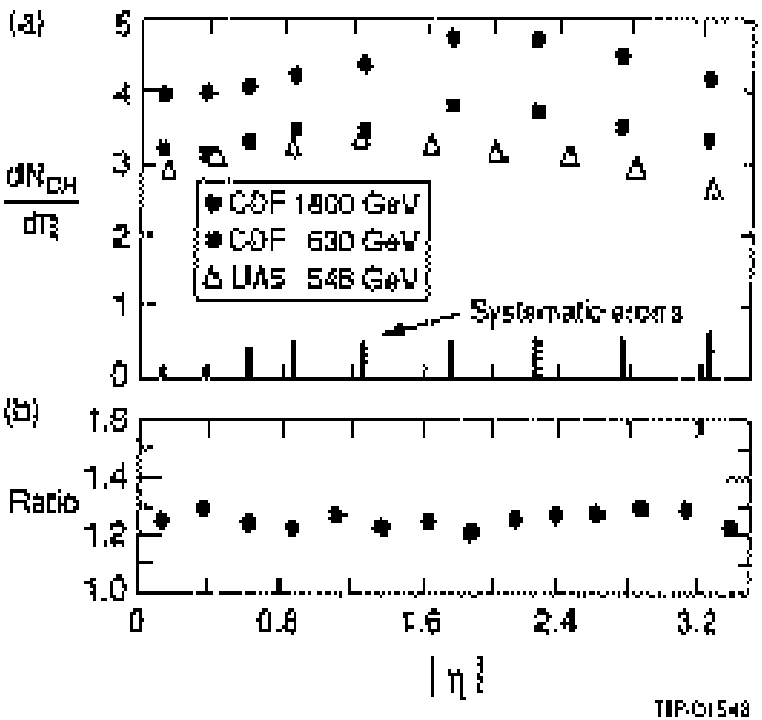
$$dy = \frac{dP_{||}}{E}$$

Note: $$y' = y - y_f \text{ where } y_f = \tanh^{-1}\left(\frac{v}{c}\right)$$ ← relativistic boost

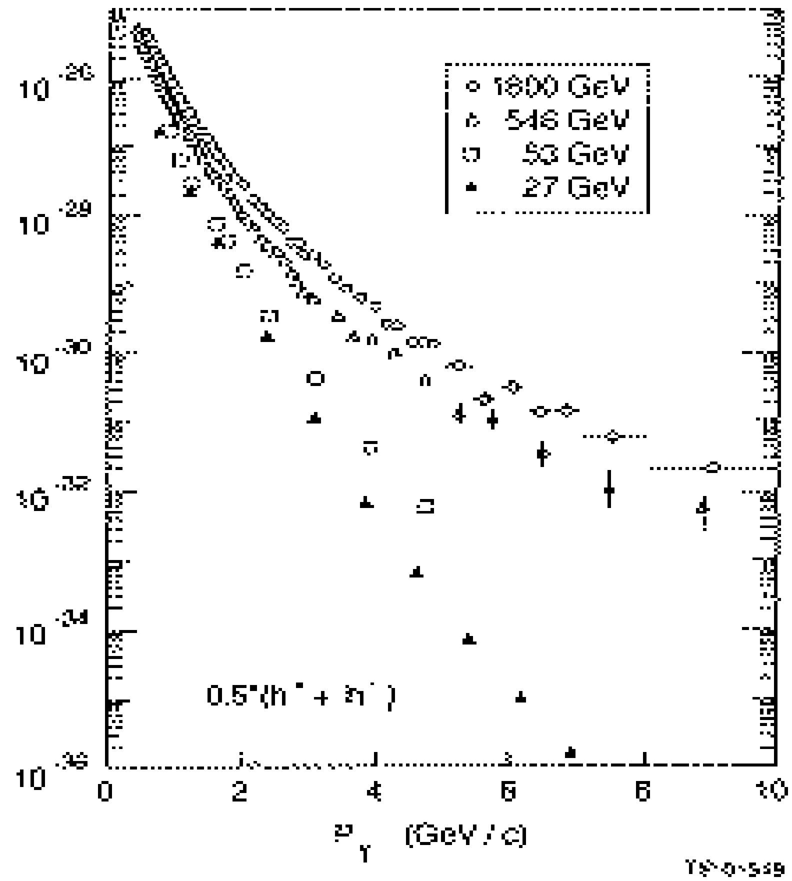
$$y \sim -\ln\left(\tan\frac{\theta}{2}\right) \equiv \eta$$ ← "pseudo-rapidity"

Natural Variables to Described Particle Production: P_T, η, ϕ

Particle Production in “Min Bias” Events



Particle Product Flat in η and increases ~ $\ln(E_{CM})$



At large E_{CM} high P_T tail

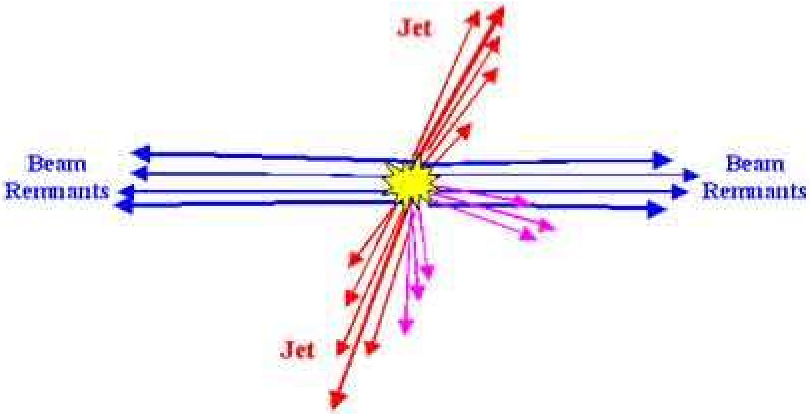
→ onset of Hard Scattering

Some Comments on Pile-up

- At LHC expect $dN_{Ch}/d\eta \sim 6.5$ for min bias
  - Haze of additional particles at low P_T
  - Makes pattern recognition difficult
  - Degrades calorimeter resolution
- Probability of a second hard scatter event very small, even at full luminosity
- Can significantly effect measurements where we sum over a large number of detector cells (eg Total Energy in Calorimeter)
- Reduce sensitivity by requiring a minimum energy per cell

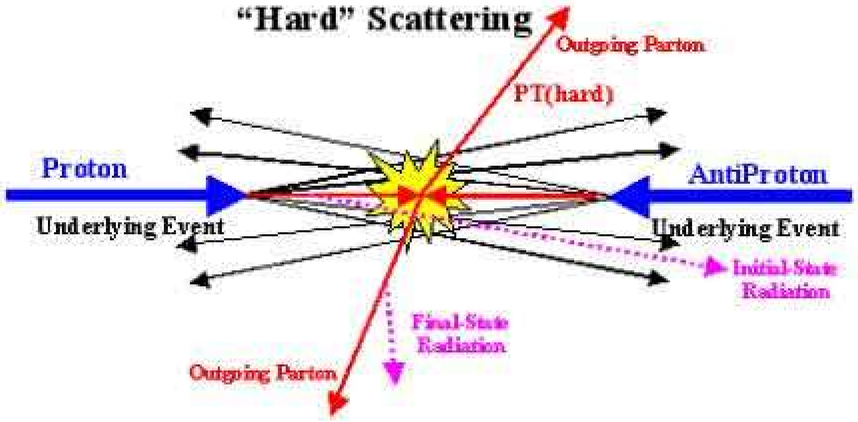
Underlying Event and Initial State Radiation

- Hard Collision leaves remnants of incoming p's moving in Beam Direction



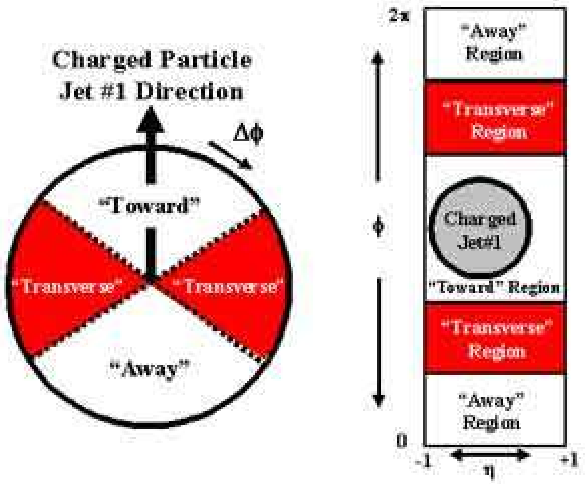
- "Initial State" gluon radiation largely co-linear with incoming partons: same basic structure

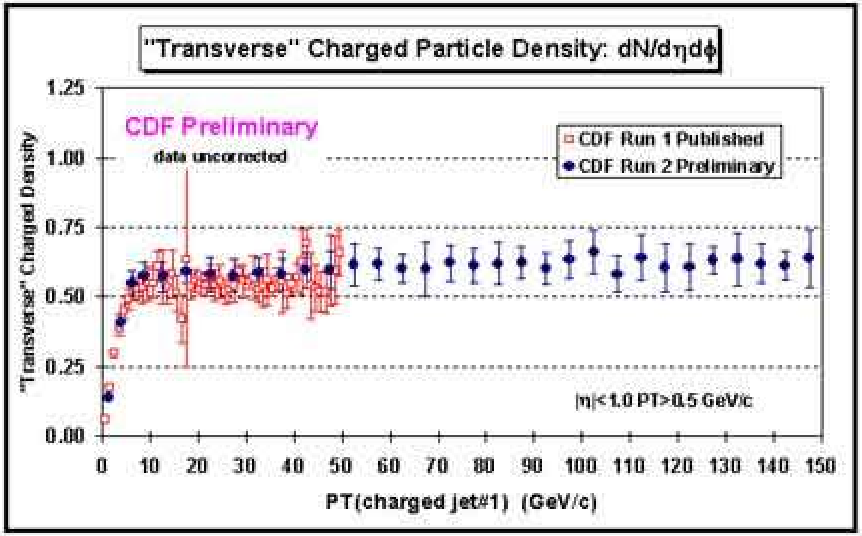
Soft particles distributed uniformly in η

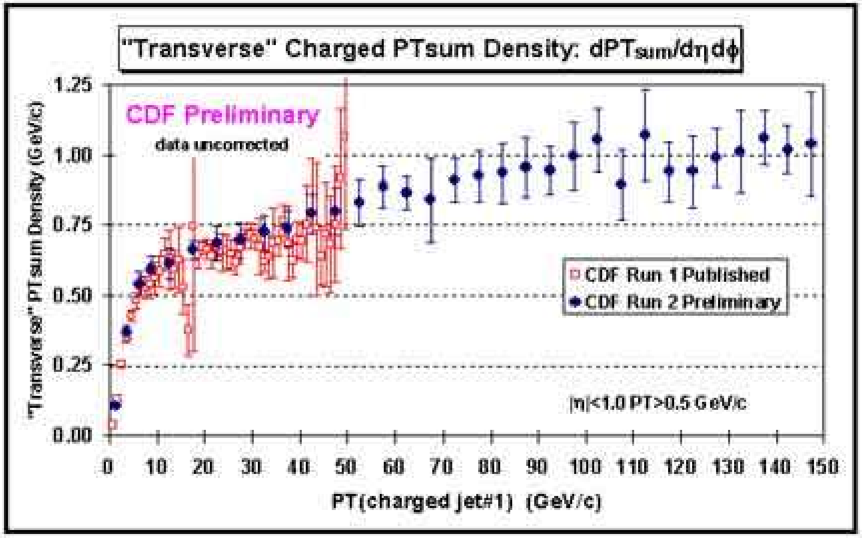


Track Distributions from Underlying Event

- Look at 90^o from jet direction







- Approx constant Particle multiplicity
- Energy density increases with hard scattering scale

Stages of Object Reconstruction

- Interpretation of Digitized Channel Info as Hits
  - Combine neighboring channels (clustering)
  - Transform coordinates (time-to-distance, local-to-global)
- Feature Extraction in Individual Sub-Systems
  - Pattern Recognition
- Fitting
  - Determine energy, momentum and/or position
- Correlating Among Sub-Systems
  - Track extrapolation to calorimeter, muon systems
- Refining Object
  - Best estimate of parameters using all information

Comments on Object Reconstruction

- Strategies may be detector and application specific
  - Trigger reconstruction: “Region of Interest” driven
  - “Seeded reconstruction:” Use info from one detector as starting point for another
  - Parameters of algorithm tunable (trade-off signal:background)
- Algorithms may be iterative to improve performance
  - Position and incident angle corrections depend on track parameters
  - Calibrations may depend on interpretation (eg electrons and photons can have different calibration constants)

Offline Production

- LHC reconstruction very complex
  - Many channels, many hits: Reconstruction is slow
- Large data collection rate and large event size
  - TB of storage required
- Cannot support bulk reconstruction by individual physicists
- Organized Production effort

  RAW $\rightarrow$ RECO $\rightarrow$ AOD $\rightarrow$ TAG $\rightarrow$ NTUPL
- Results available through Data Delivery System
  - "Analysis" is performed on output of Offline reconstruction

Output of Offline Production

- Collections of Candidate Objects
  - Tracks, Jets, Electrons, μ, τ, vertices, missing energy, heavy flavor
- Selections Performed Using Loose Criteria
  - Can tighten during analysis phase
- No attempt to uniquely identify objects
  - Same energy deposition may appear as jet, e and γ candidate
- Support for multiple algorithms
  - Best jet algorithm for Top analysis may not be the best algorithm for QCD studies

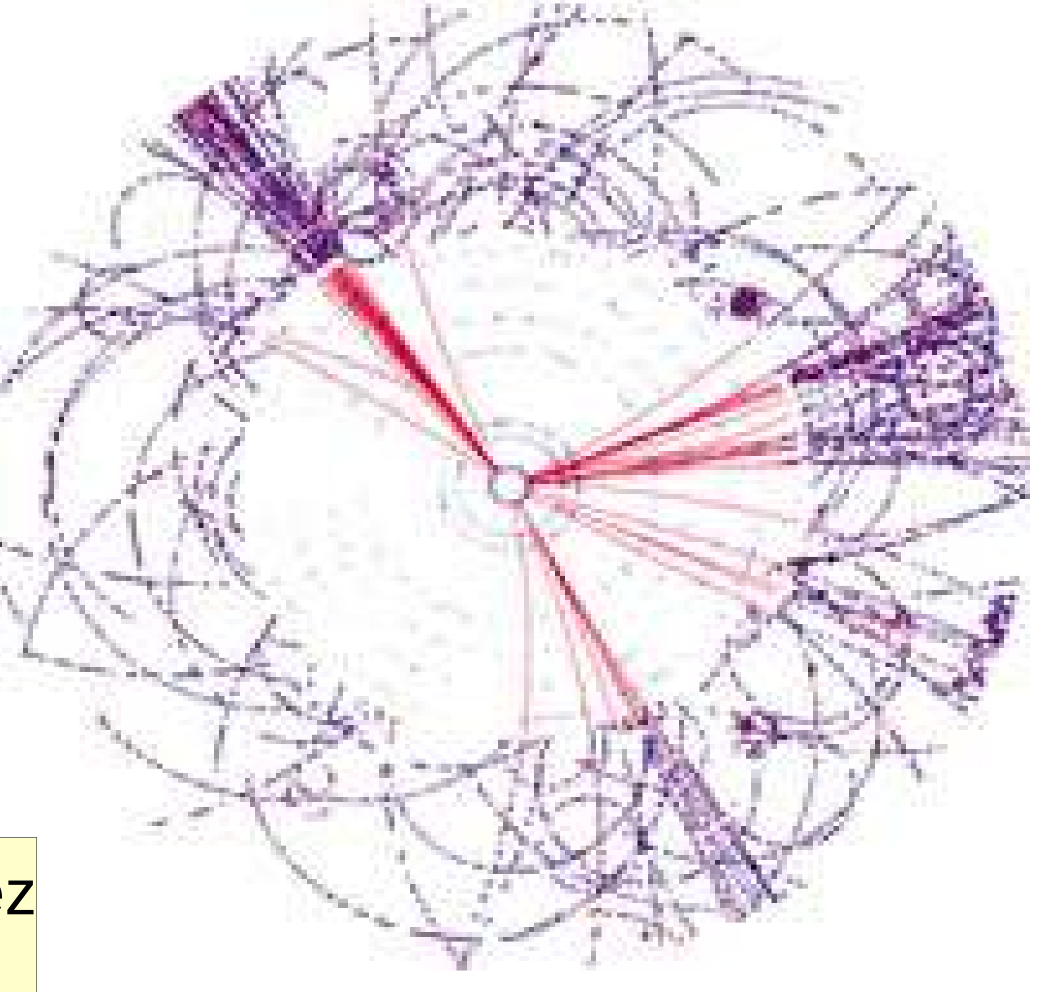
Physicists impost consistent interpretation during analysis phase

Object 1: Tracks

- Reconstruction of trajectory of charged particles

Measure:

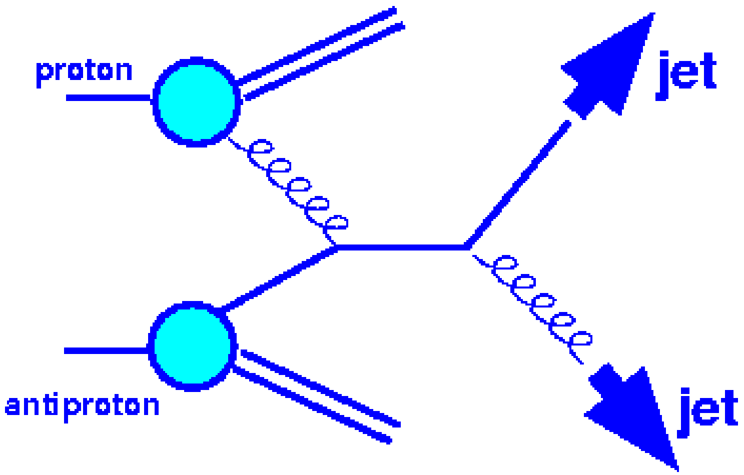
- Momentum
- Charge
- Vertex information

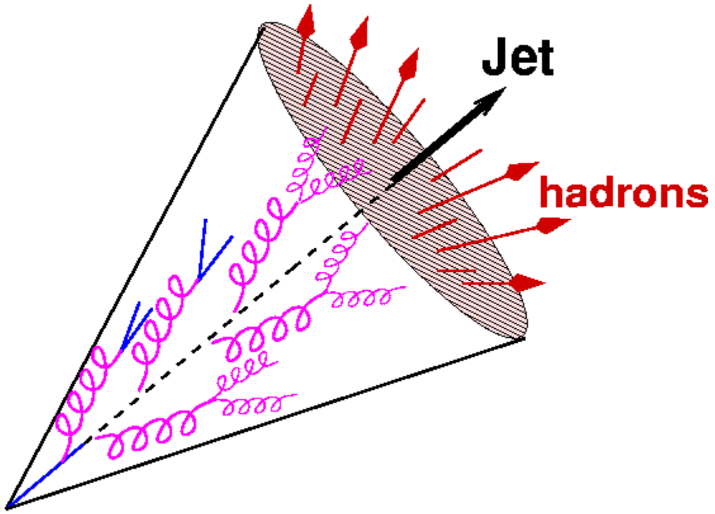


Details in A. Dominguez talk Monday

Object 2: QCD Jets

- 2 $\rightarrow$ 2 elastic scattering of quarks and gluons

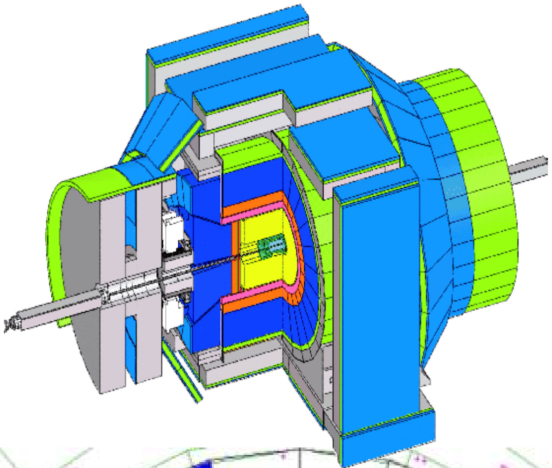


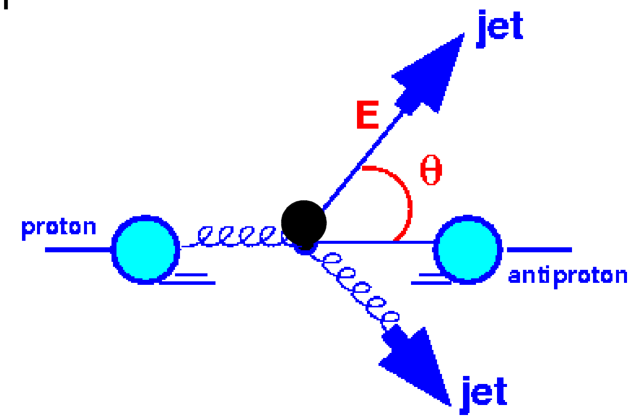


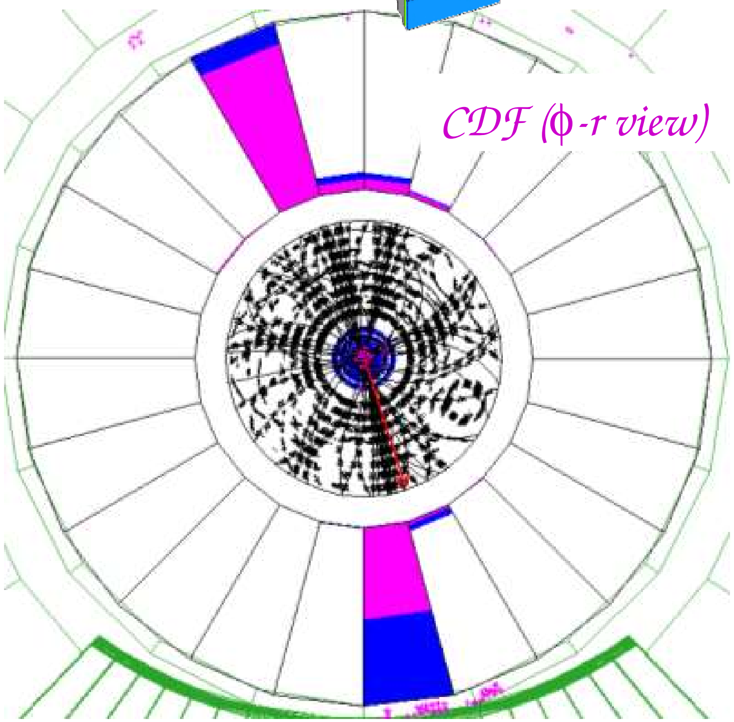
- Strategy
  - Calorimeter based pattern recognition
  - Associate tracks with the jets after calorimeter jet found
  - Primary vertex needed to calculate p_T

What do Jets Look Like ?

(the highest P_T jet event from CDF)

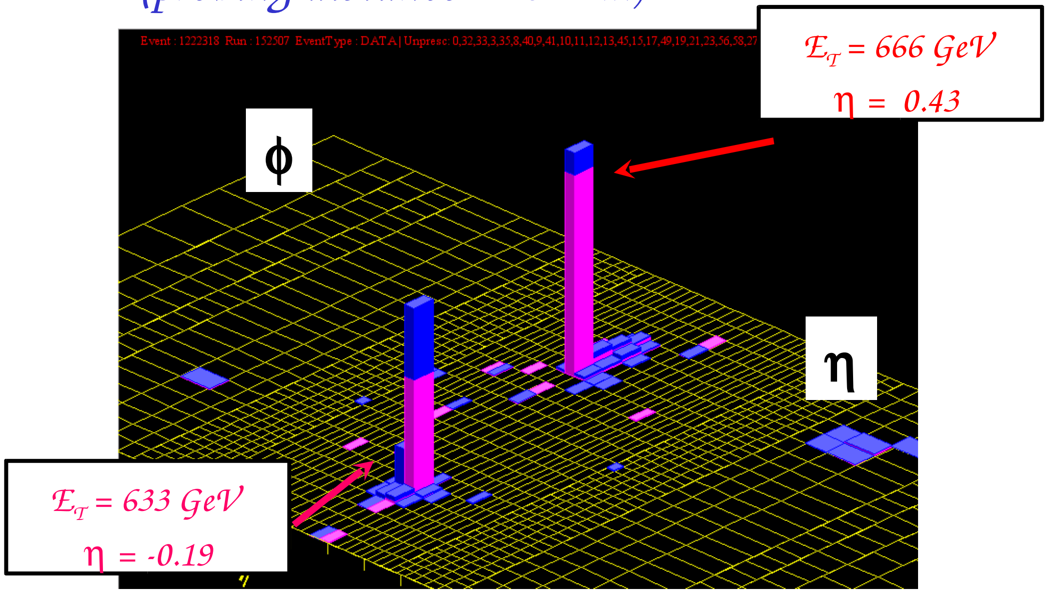






Dijet Mass = 1.36 TeV

(probing distance ~10^{-19} m)



Comments on Jet Reconstruction

- Quarks and Gluons are colored objects: Hadrons are not
  - Mapping of collections of particles to partons is tricky
  - Many experimental and theoretical details
    - See talks from J. Huston and B Heinemann next week
- But, these lectures need to use some basic facts about jets
  - Will discuss the baseline reconstruction algorithm and leave the hard stuff for next week

Jet Reconstruction:What Variables Do We Use?

- We have seen that natural variables are P_T ,η and ϕ
- But we don't measure P_T in calorimeter
  - Make pseudo-particles from calorimeter cells
  - Project calorimeter data onto a uniform $\eta - \phi$ grid
  - Treat each calorimeter cell as massless particle

Calorimeter “Tower”

Energy: From Calorimeter

Direction: Project to Origin

A Simple Cone Algorithm for Finding Jets (your mileage might vary)

- Jets are circles when projected in $\eta - \phi$ space
- To reject fluctuations in underlying event and pileup:
  - Start with a "seed" tower above fixed E_T (E_{ESeed})
  - Draw a circle in $\eta - \phi$ space (Cone Size: 0.4 to 1)
  - Include all towers with above a fixed E_T (E_{tmin})
  - Calculate E_T centroid
  - Iterate list of towers until stable
- This is the "pattern recognition" phase

Defining Jet Energy and Momentum

- For our purposes, the basic definition is enough:

$$E_{jet} = \sum_{Towers} E_i$$

$$E_T = \left(\sum_{Towers} E_i \right) \sin\theta centroid$$

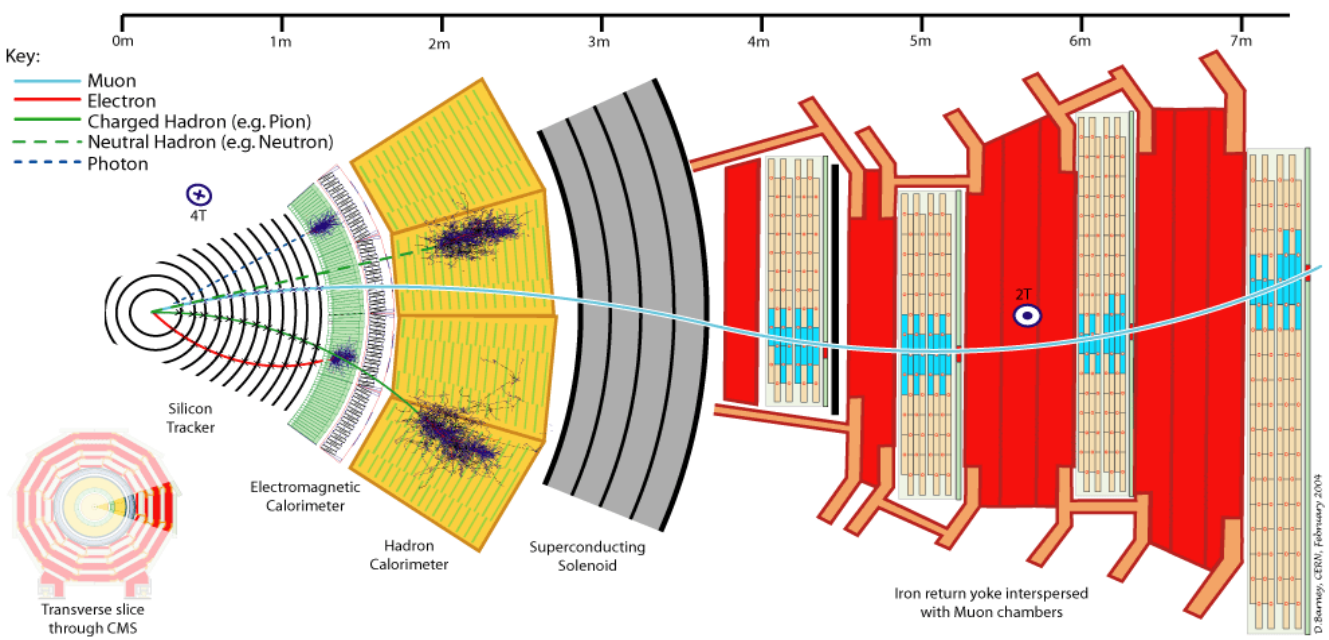
$$\vec{p}_{jet} = \sum_{Towers} E_i \hat{n}$$

$$p_T = \sqrt{p_x^2 + p_y^2}$$

Jets defined this way have “mass” ~ 10 GeV

Objects 3-5: Charged Leptons (e, μ, τ)

- Must extract lepton signal from much larger jet bckgnd
- Requires correlation of information among detectors
- Selected based on properties of each lepton species



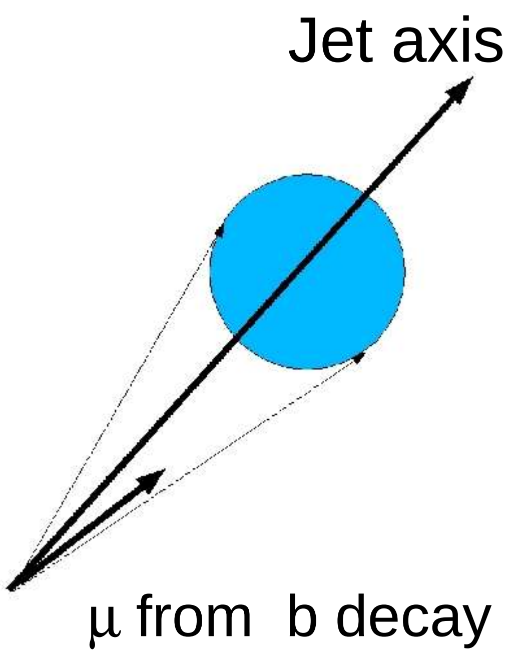
Comments on Lepton ID

- Lepton ID involves simultaneous selection on a number of measurements
- Trade-off between efficiency and rejection
  - Tunable handle to customize selection to specific physics analysis
- Two approaches:
  - Cut-based algorithms: selection on fixed values for each variable
  - Neural Net or Multivariant analysis: More complex treatment of correlations

In both cases, majority of work is determining correct choice of variables and parameterizing signal and bckgnd input distributions

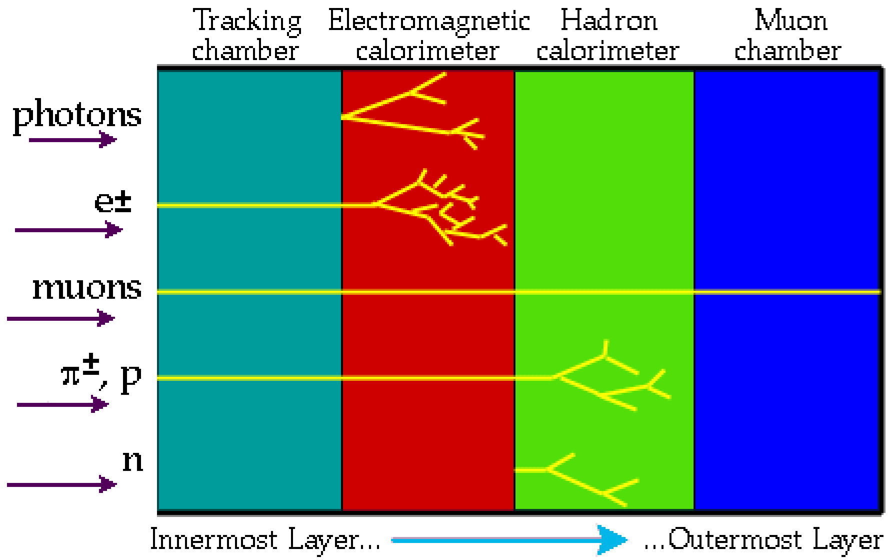
Physics Dependent Lepton ID Efficiencies

- Some selection variables used in lepton ID are environment dependent
  - Energy deposition in calorimeter affected by nearby tracks
- Classify leptons as isolated or non-isolated
  - Leptons produced in the decay of high mass objects are usually emitted far from other jets and leptons
  - Leptons produced in the decay of b and c quarks are usually buried in jets
- Warning: Efficiency will be P_T and process dependent: more when we get to SUSY



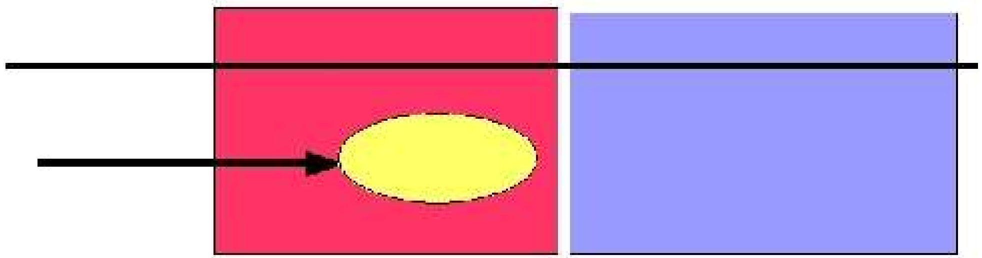
Object 3: Electron Reconstruction

- Electrons signature:
  - Energy Deposition in EM Calorimeter
  - Track pointing at the energy deposition and with momentum consistent with calorimeter energy
  - Little or no energy in hadron calorimeter

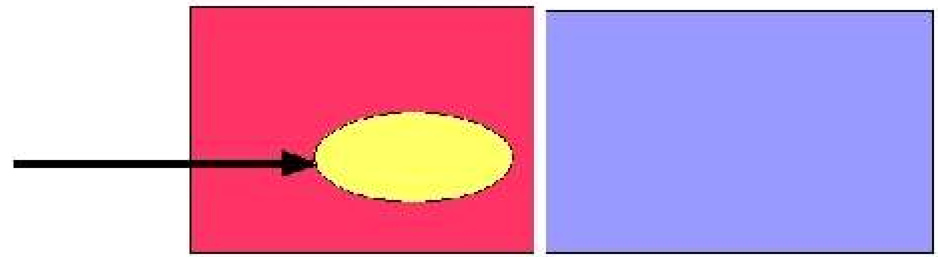


Backgrounds for Electron ID

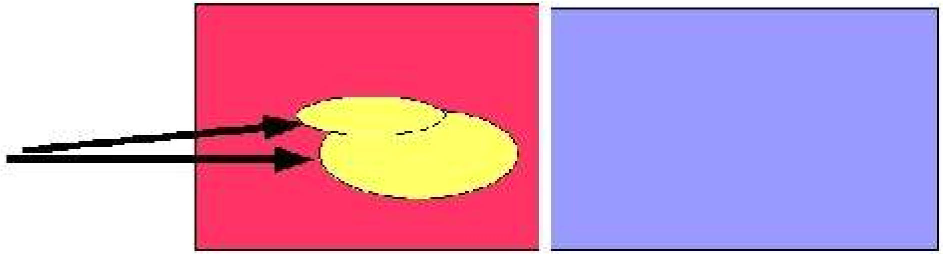
π^0 and non-interacting π^+



Early showering π^+



Photon Conversions

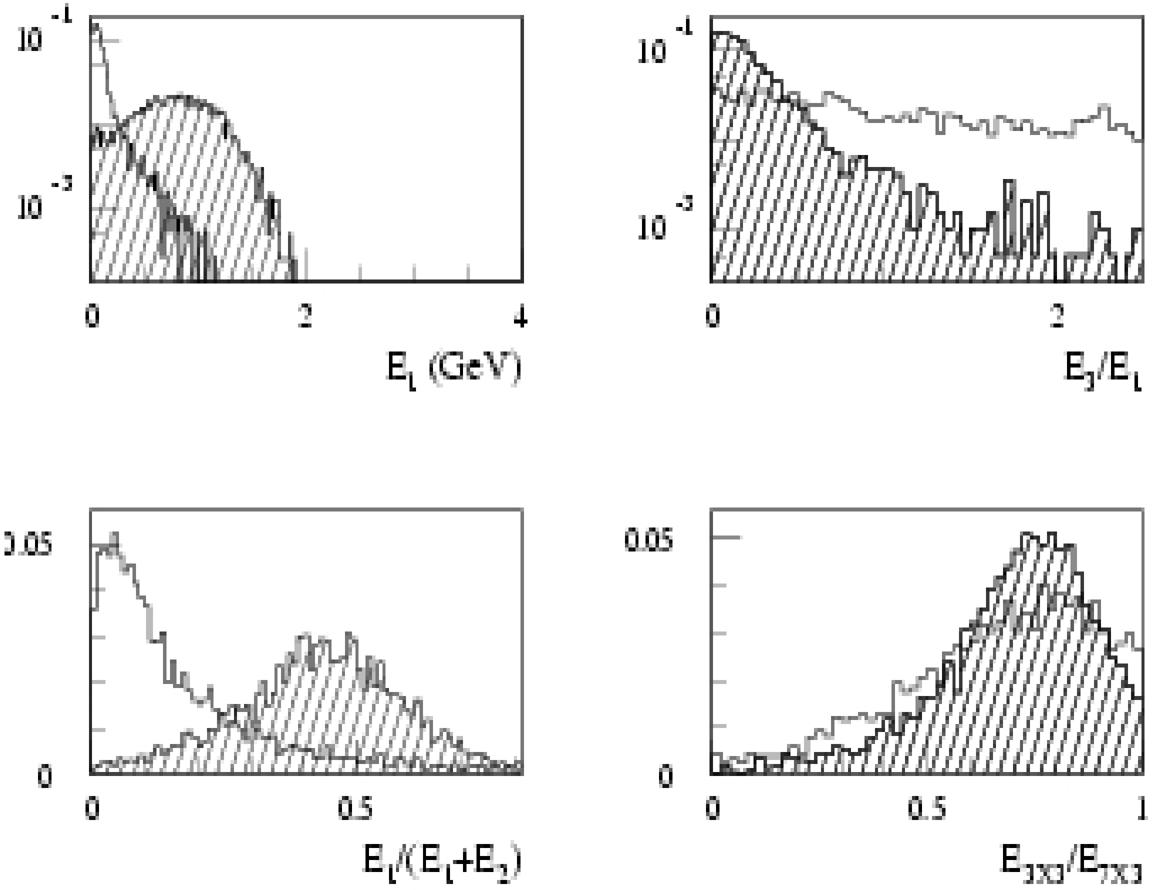


Electron ID: Rejection of Background (I)

Choice of variables depends on detector. Some possibilities:

- Shower Shape Variables:
  - Longitudinal shape: ratio of energy in depth segments of calorimeter
  - Transverse shape: Hadron showers typically wider than electrons (also rejects π^0 π^+ overlap)
  - Had/EM: Expect very little energy deposit in HAD calorimeter

Some ATLAS Examples: Shower Shape Variables



Comparison of Distributions for Electrons and Jets

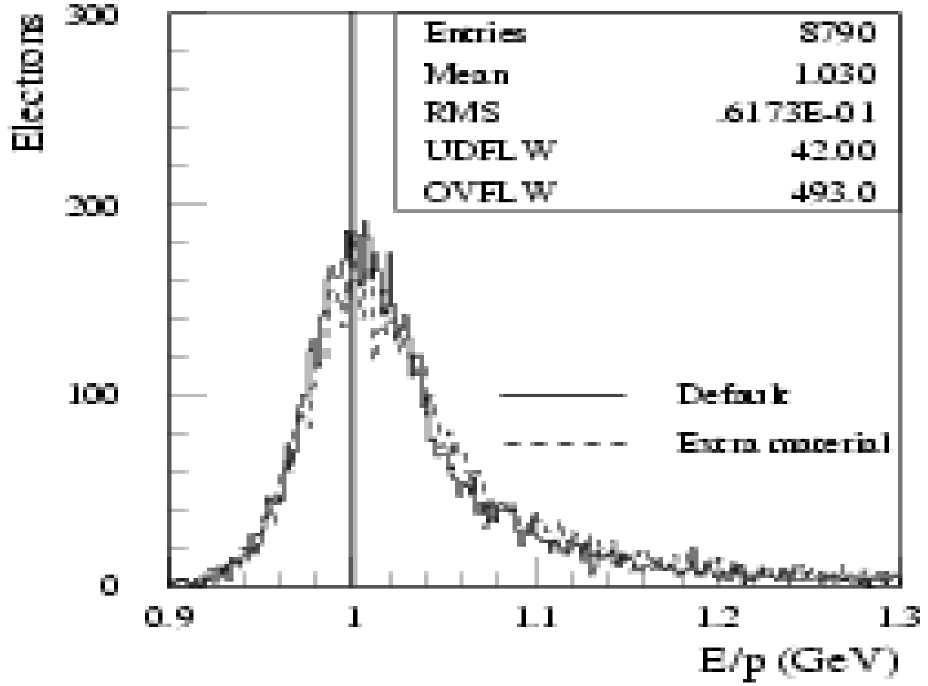
Electron ID: Rejection of Background (II)

- Track-Shower Matching:
  - E/P: Ratio of energy in calorimeter to momentum in tracker
  - Pointing: Compare extrapolated position of track to position of EM cluster

Caution:

- Significant material in LHC trackers means electron bremstahlung
- Correct modeling of material distribution necessary both for defining selection criteria and for estimating efficiency

E/P Distribution: ATLAS



20 GeV Electrons

Electron ID: Rejection of Background (III)

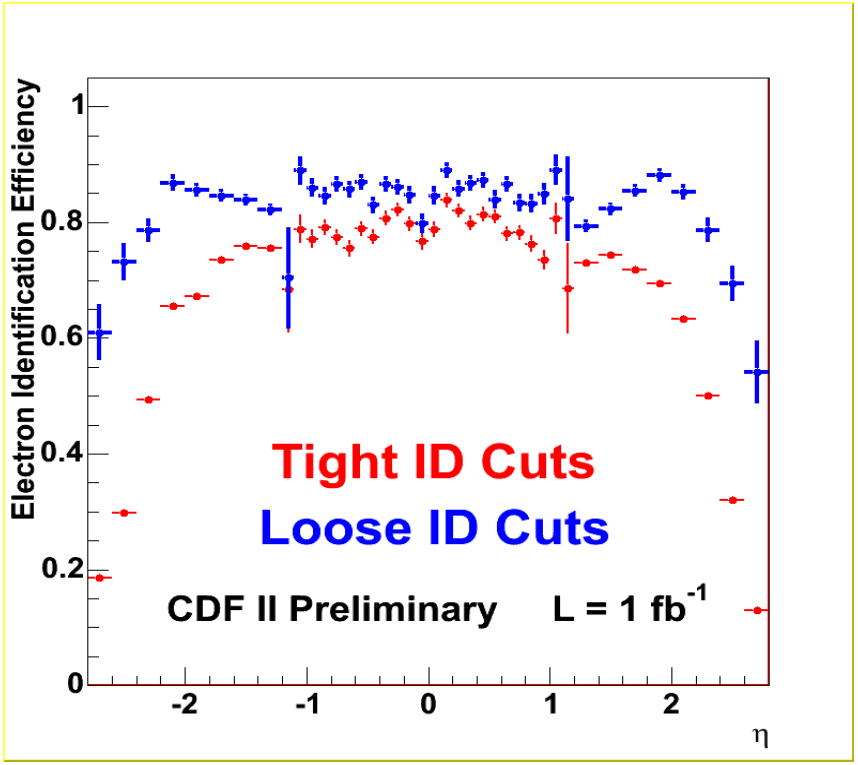
- Large amount of material also means photon conversions are an issue (photons from π^0)
  - Explicit removal of conversions:
    - Require hits in pixel layer (most of material outside this)
    - Look for second track from conversion: cut on reconstructed mass and angle

Electron ID: Rejection of Background (IV)

- Isolation:
  - Study ratio of energy in annulus round electron to enegy of electron
  - As noted above: Does not work for all physics processes
- Transitions Radiation and dE/dx:
  - CDF drift chamber measures dE/dx: sensitive to particle velocity: helps for low momentum e
  - Atlas tracker has TR function: Can require high energy deposition hit, at cost of efficiency

Efficiency of Electron Selection

- Measure when possible using real data:
  - W from no-track trigger to measure tracking efficiency
  - Z with one tight electron and with loose selection
- Use simulation to extrapolate kinematics and correct for environmental issues (eg isolation)



Conclusions for Today

- Successfully reconstruction of physics objects requires knowledge of QCD production
- Complexity of events, large rate and large event size means reconstruction must be done communally
- Analysis starts with collections of candidates, refines selections and imposes consistent interpretation
- Object reconstruction mirrors detector and physics signatures
  - Today: Jets and Electrons
  - Tomorrow: μ, τ, ν, W, Z, top