Computing Performance

Run II Computing Review

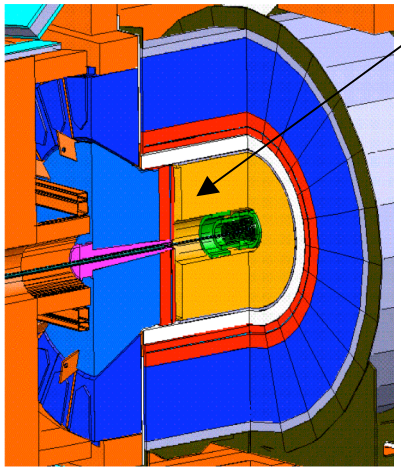
September 14th 2005

Dave Waters, University College London

- Status & physics performance of the reconstruction algorithms :
  - Tracking
  - Calorimetry
  - Particle ID
  - Jets & B-tagging

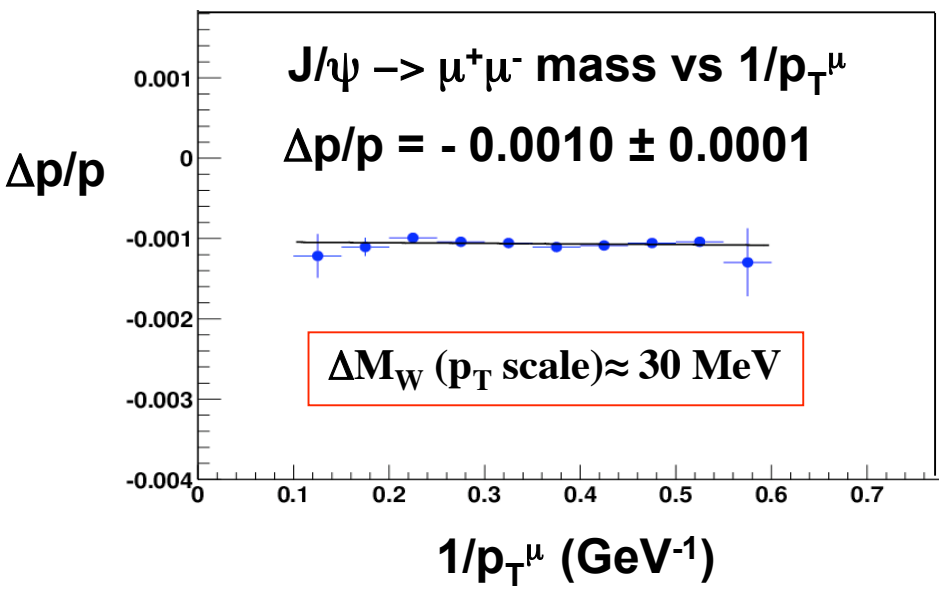
  } + lepton ID
- Luminosity performance
- CPU performance
- Online implications of offline code performance
- Simulation refinements
- Conclusions & outlook

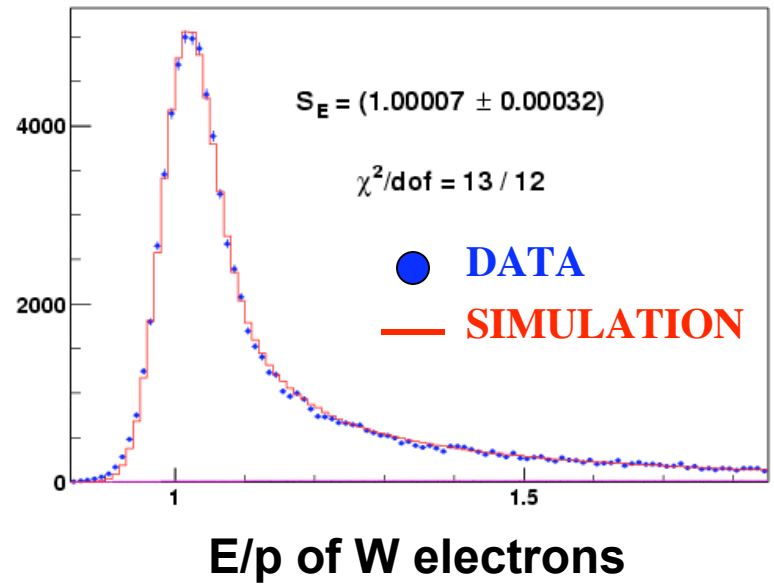
Tracking Performance



Central Outer Tracker

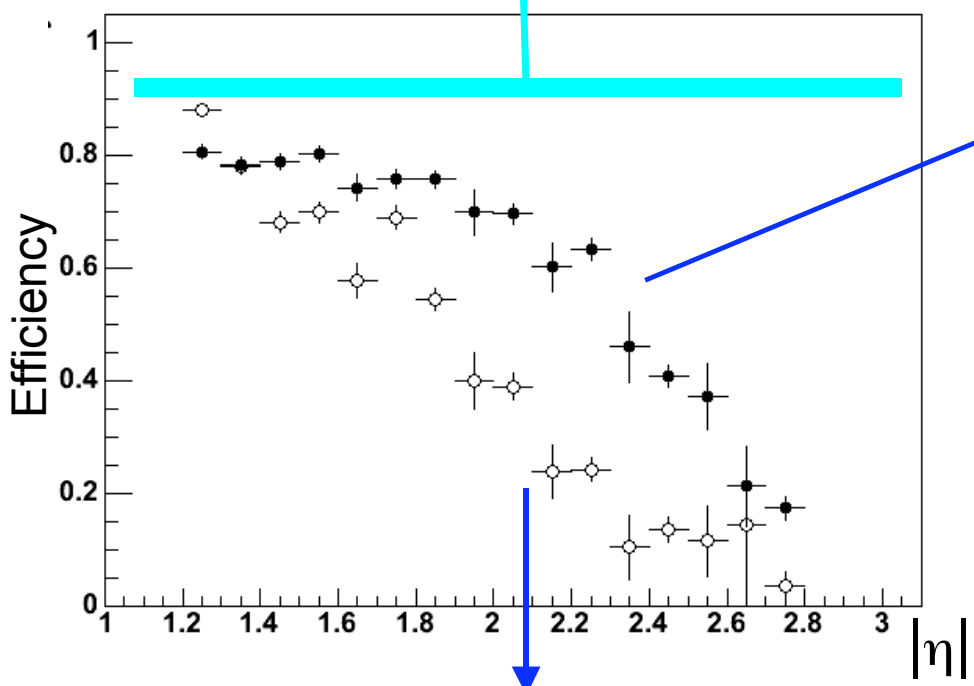
- The heart of CDF : the key to tracking and setting of energy/momentum scales.
- Extremely efficient : 99.6 % (isolated); > 96% (non-isolated) particles.
- With recent improvements (post-pattern recognition hit finding), algorithms stable and performing well at least to ~ 2×10^{32} cm^{-2} s^{-1}
- New alignment : wire positions to 10 μm.



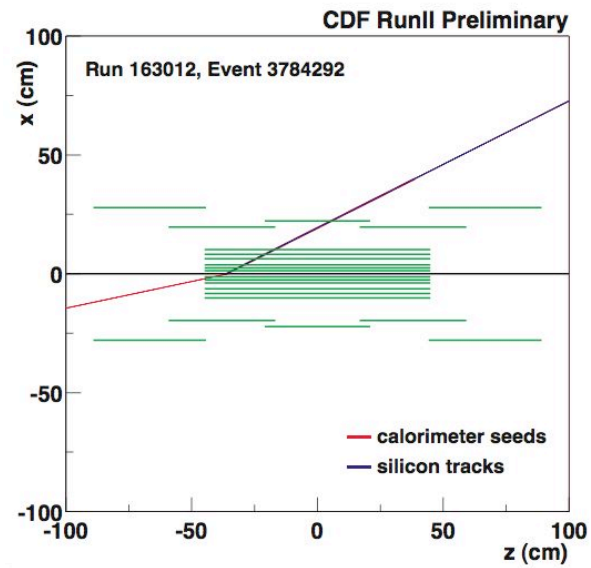


Forward Tracking & Lepton ID

electrons with no track requirement



forward electron tracks (seed track finding using calorimeter info)



all tracks :

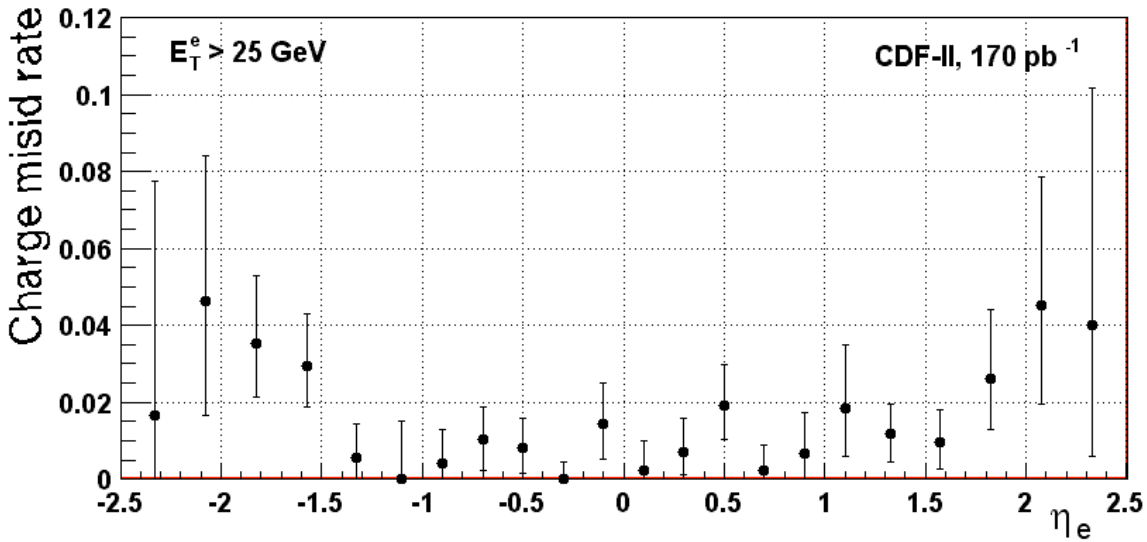
- inside-out
- silicon stand-alone

NIM A538 (2005) 249

Forward Tracking & Lepton ID

Can also optimize algorithms and/or cuts for resolutions rather than efficiency :

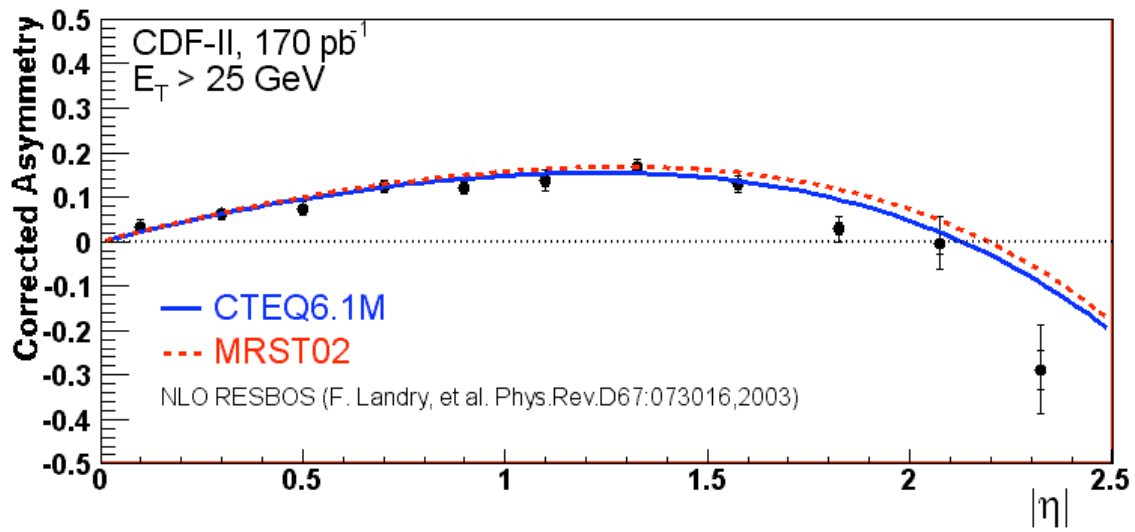
- p_T
- lifetime measurements
- charge identification



<u>W (Lepton) Charge Asymmetry</u>

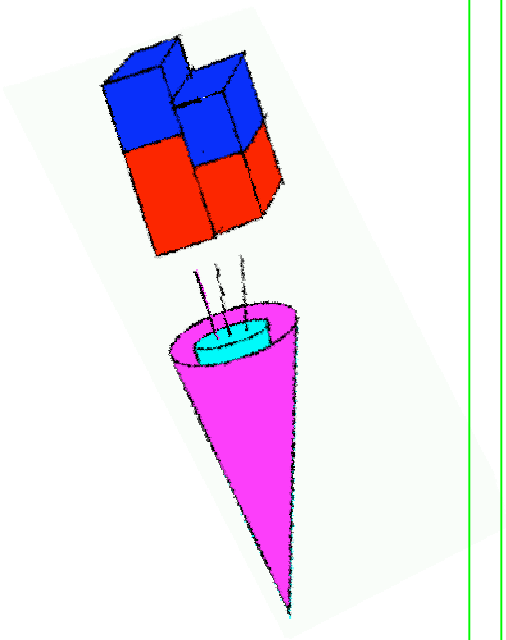
Input to PDF's

$$\frac{d\sigma(e^+)/d\eta_e - d\sigma(e^-)/d\eta_e}{d\sigma(e^+)/d\eta_e + d\sigma(e^-)/d\eta_e}$$

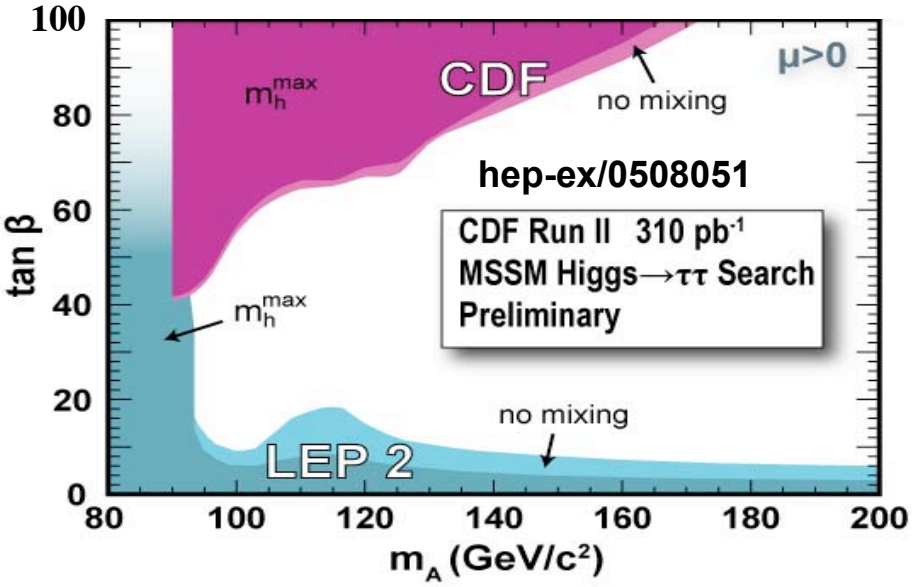


Taus

- Complex high-level object, incorporating :
  - tracking
  - calorimetry
  - shower-max & π^0 reconstruction

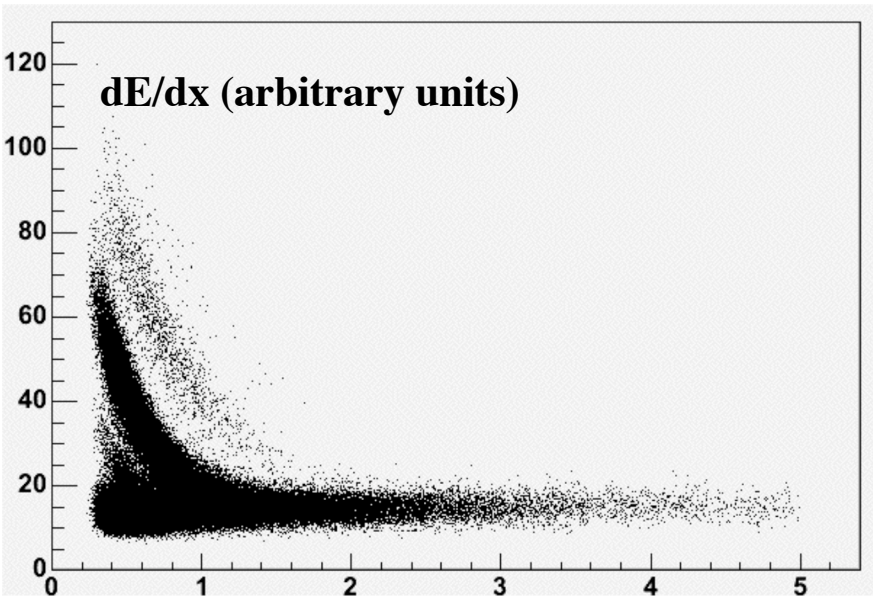


- Offline: ID algorithms largely converged. Some differences due to demands of different analyses.
- Online: suite of tau triggers (leptonic + hadronic decay modes).
- $Z \rightarrow \tau^+\tau^-$ calibration signal established.
- Very important search mode :

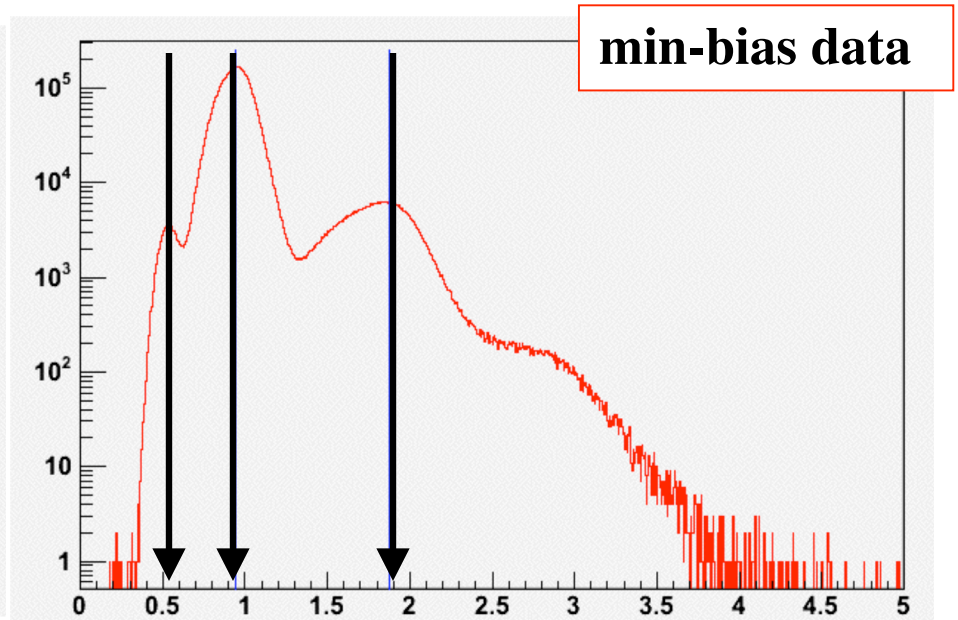


Particle ID : dE/dx

- 96 charge measurements allow excellent COT particle ID
- well tested & packaged CDF-standard particle ID code corrects for detector inhomogeneities and time variation : *tools developed for one analysis* (heavy-flavor physics) *readily adopted by other analyses* (baryon production in min-bias events).
- silicon dE/dx has been demonstrated (adds very little to COT for most analyses)

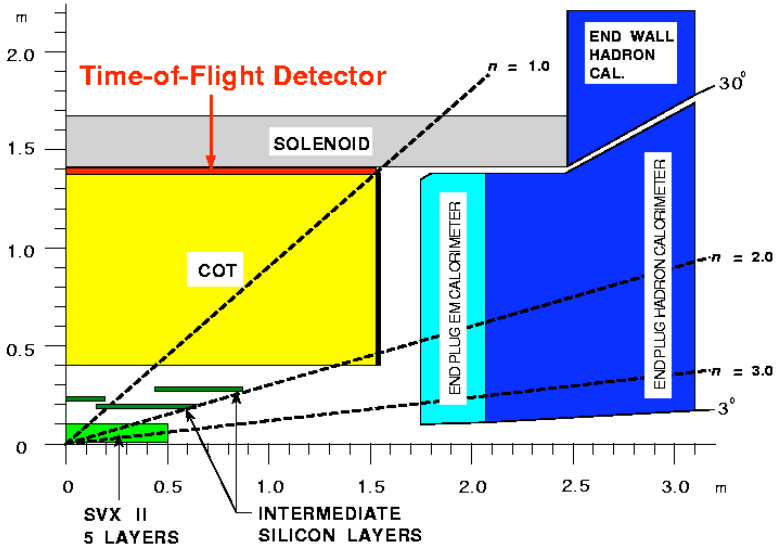


track momentum (GeV/c)

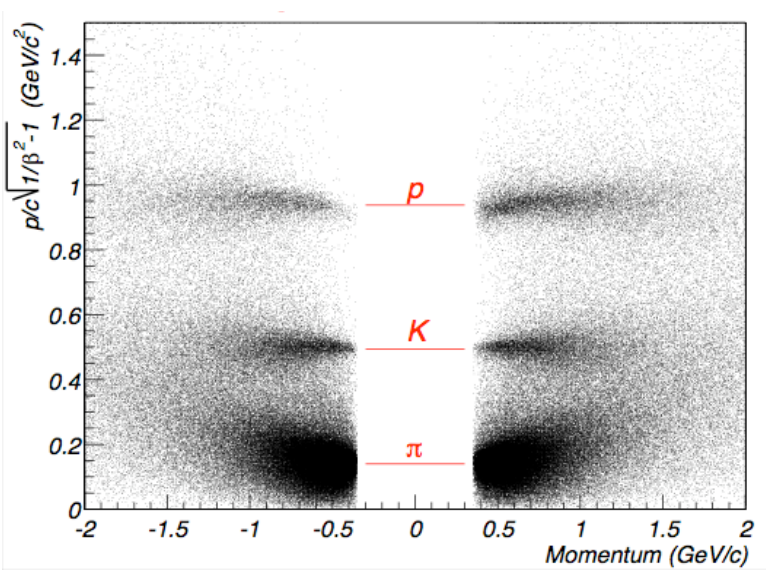


K *p* *d*

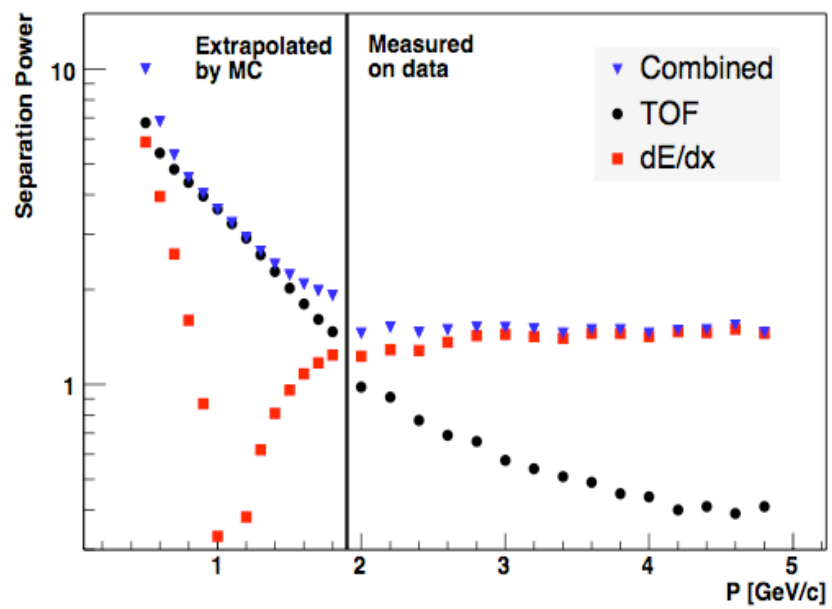
Particle ID : Time of Flight



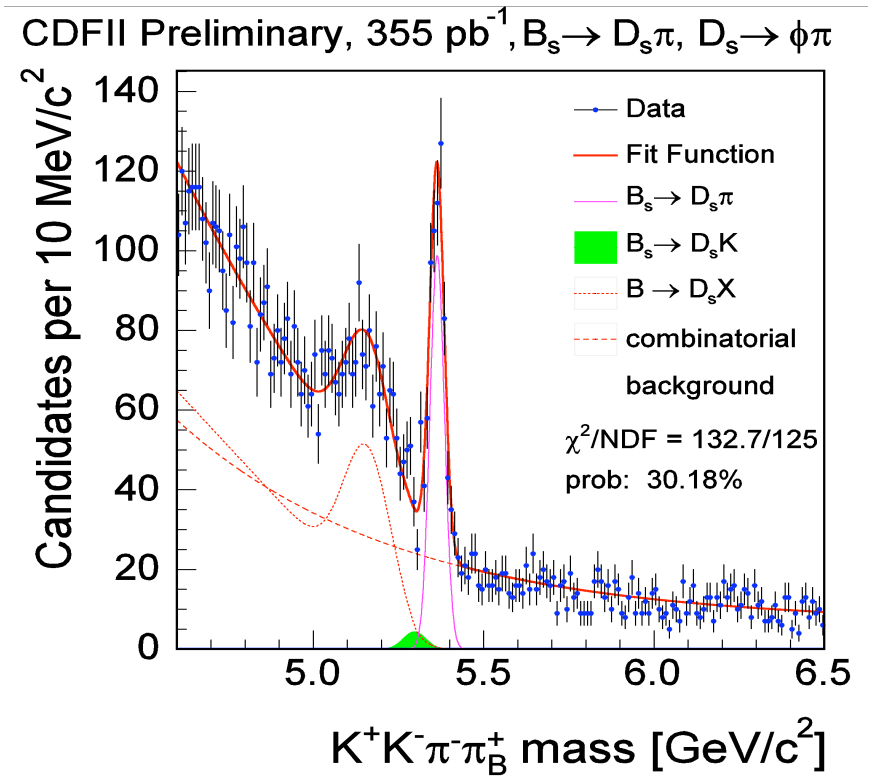
- Large effort to derive corrections (as functions of track position, angle, particle type, etc.), calibrations & develop accurate simulation.
- Close to design resolution of 100 ps.
- Combined (complementary) particle ID :



K-π separation ("sigmas")



Flavour Physics



Example : B_S

Demanding analysis bringing together all the aspects of tracking & particle ID :

- secondary vertex and lifetime determination
- dE/dx
- time-of-flight

Future :

A few low level improvements :

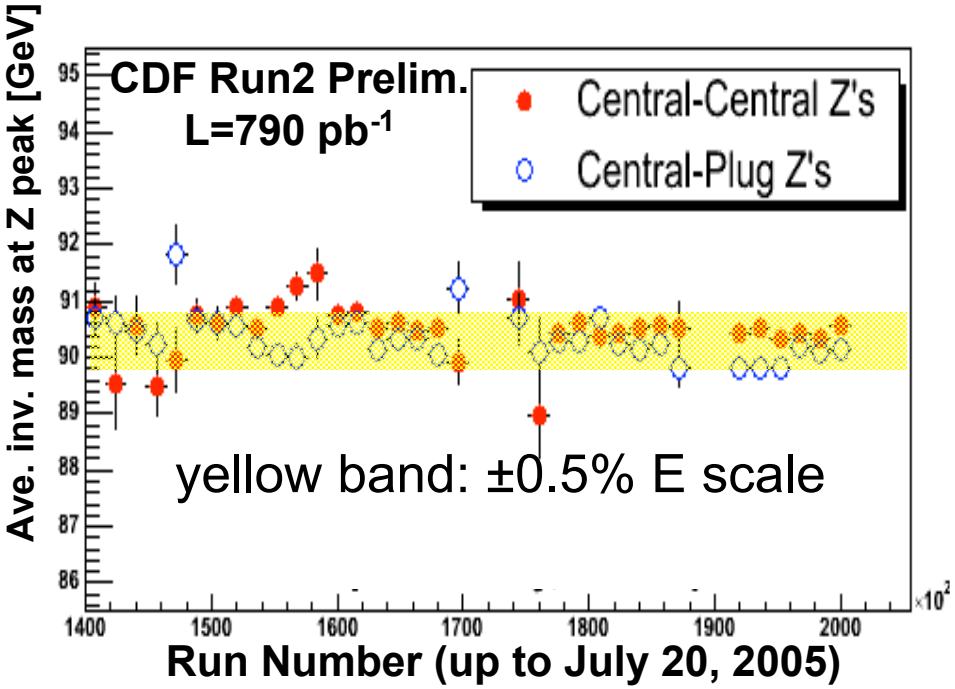
- one more SVX & relative alignment iteration after feeding in latest COT alignment

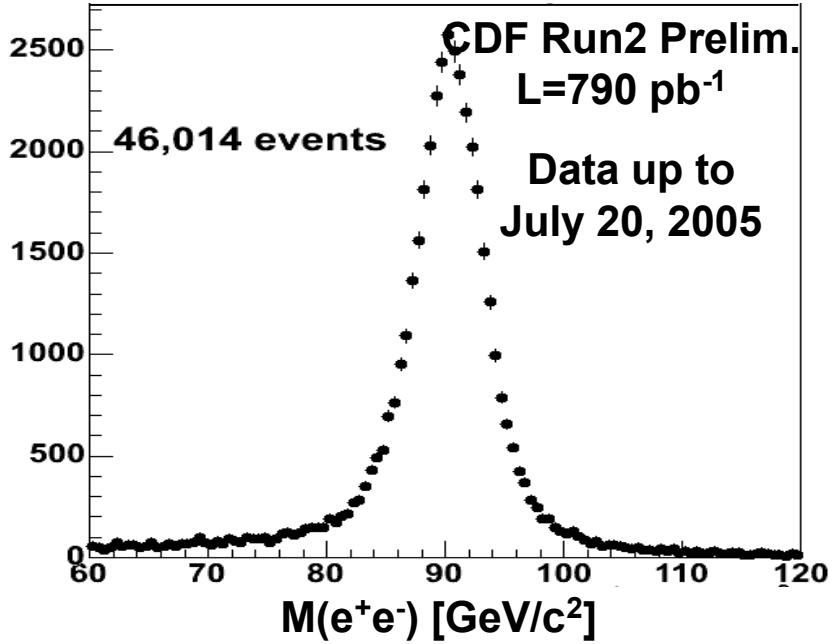
Mainly high level algorithm developments :

- flavor tagging
- understanding of backgrounds
- etc …

Calorimeter Performance

- Algorithms (jet, electron ID) now very stable.
- New Run IIb detectors already incorporated into production reconstruction :
  - new central pre-shower detector
  - EM timing
- Further development will be in higher level (user/ntupling) realm.
- Recent progress has been operational - fast calibrations for production :





Jets

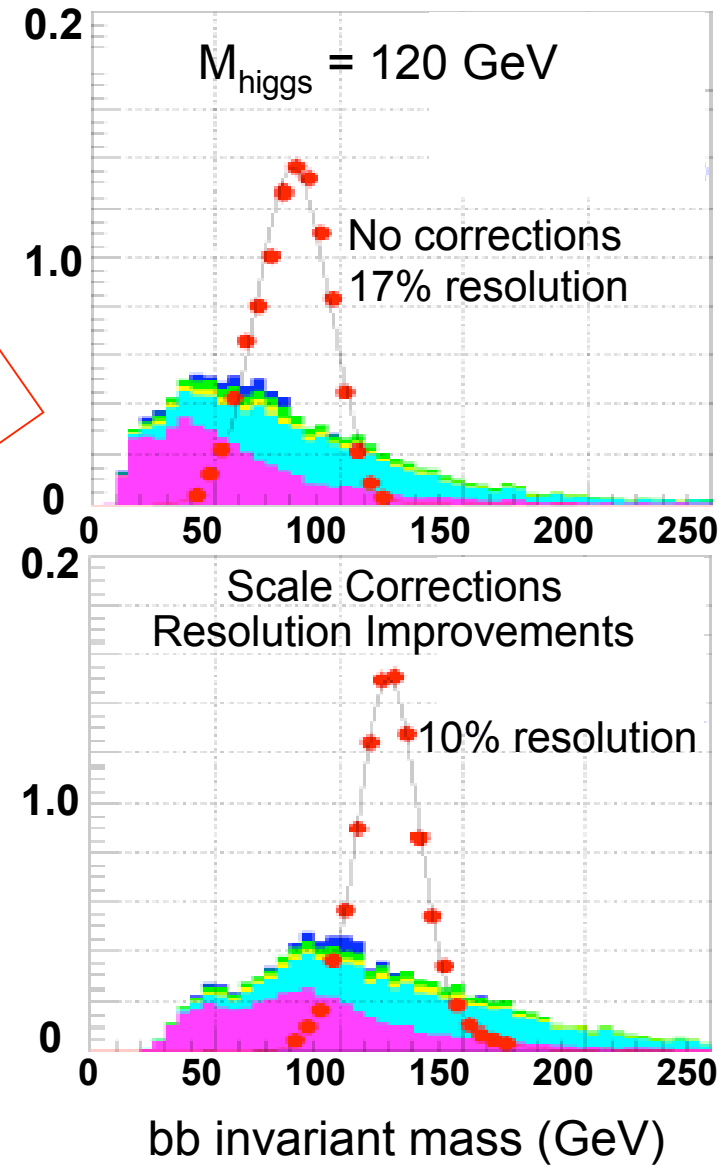
Jet energy scale uncertainty:

- precision measurements (M_{top}) & searches (Higgs)
- now ~2.5% uncertainty for jets in top decays
- further improvements:
  - generators, higher order QCD
  - better scale for $E_T > 100$ GeV region
  - complete by end of this year

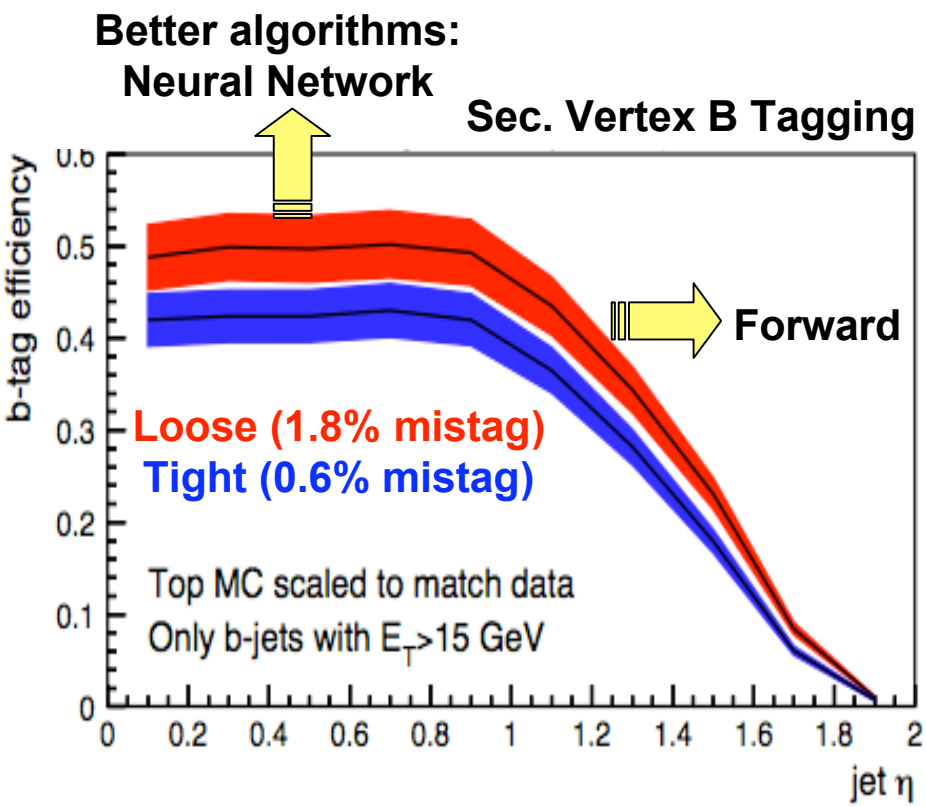
Jet energy resolution:

- searches (Higgs, …)
- currently 17%, goal 10-11% (@ M_H =120GeV)
- further improvements:
  - combine track, calorimeter Info: 2%
  - expand cone size: 2%
  - b-jet specific corrections:1-2%
  - better algorithms: 1-2%
  - complete by spring 2006

improvements all "post-production"



B-Tagging



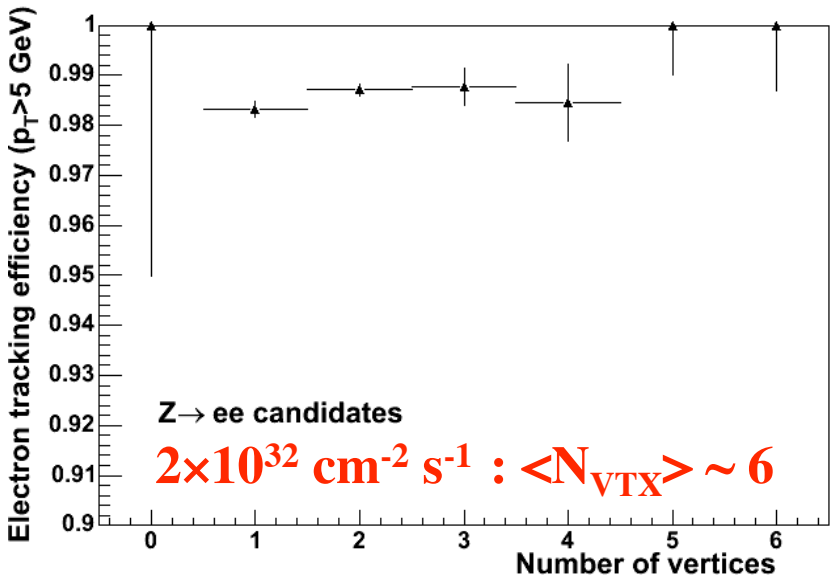
- Importance of 3rd generation for top physics, Higgs & searches.
- Different analyses can pick points in space of *efficiency vs. fake-rate*.
- Approaching design performance, but not there yet.
- Other taggers also becoming mature :
  - soft muons
  - jet probability
- Ongoing efforts :
  - increased forward acceptance
  - soft electron tagger
  - neural net based taggers

Performance at High L_{INST}

Low-level :

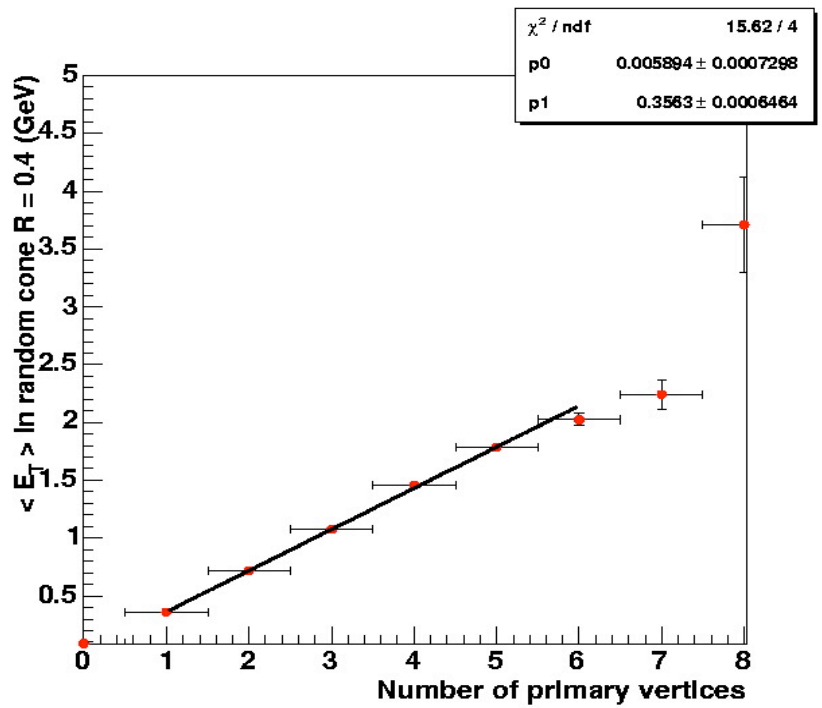
COT tracking most vulnerable to high occupancies :

- hit recovery maintains good efficiency & resolution up to ~ 2×10^{32} cm^{-2} s^{-1}
- no evidence for any loss in efficiency in current dataset
- will monitor performance closely as peak luminosities further increase

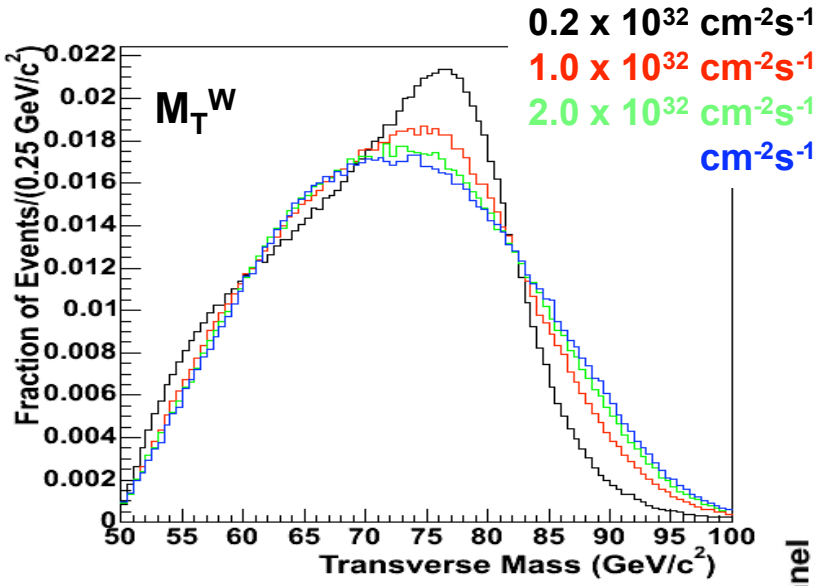


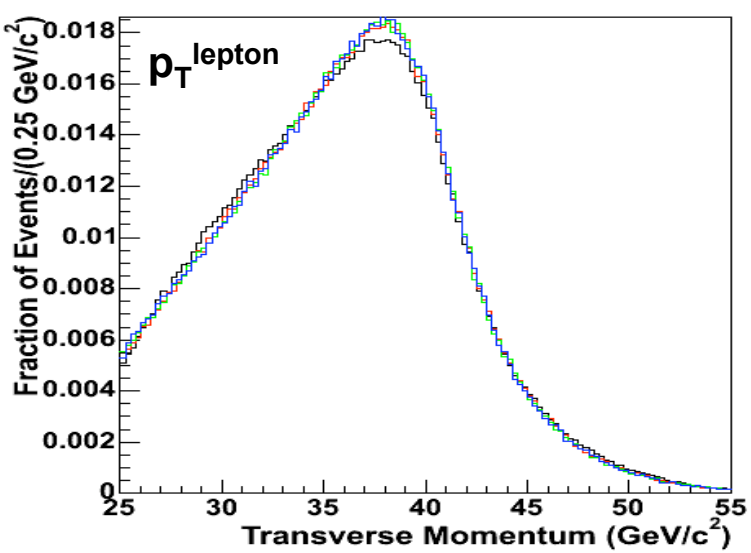
Higher-level :

- Important analysis variables are f(N_{VTX}) :
  - isolation energies
  - underlying event corrections
- Check for non-linearities at very high-L
- Need good simulation (see later)
- Use observables less dependent on L_{INST}



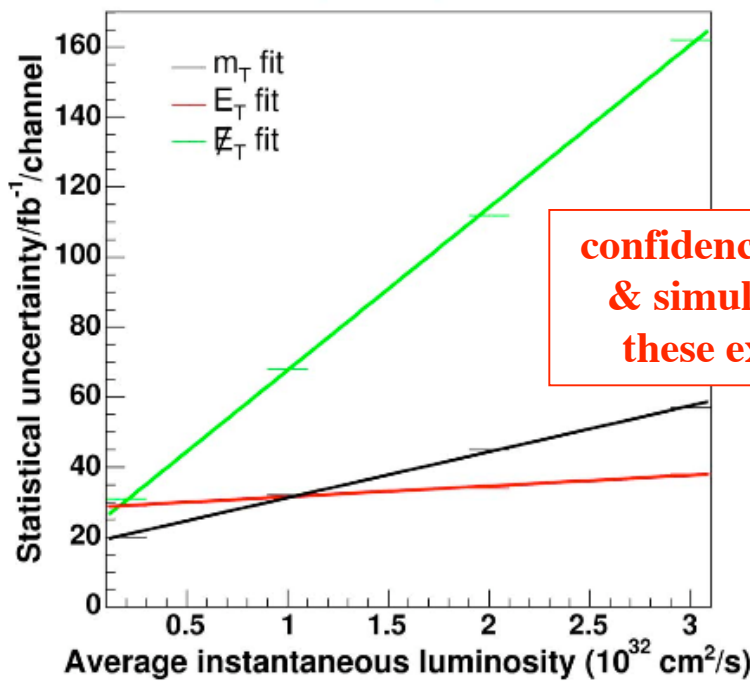
Design Analyses for High L_{INST}





e.g. W Mass Measurement

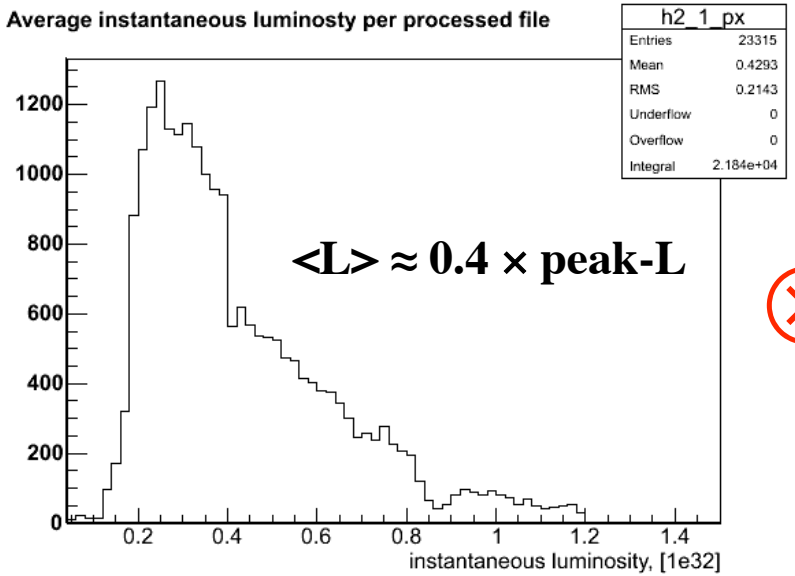
- Missing-E_T resolution becomes progressively worse at higher instantaneous luminosities.
- Lepton p_T resolution almost unaffected.
- Switch over point is ~ 1×10^{32} cm^{-2} s^{-1}
- Data taken at all luminosities will be useful.



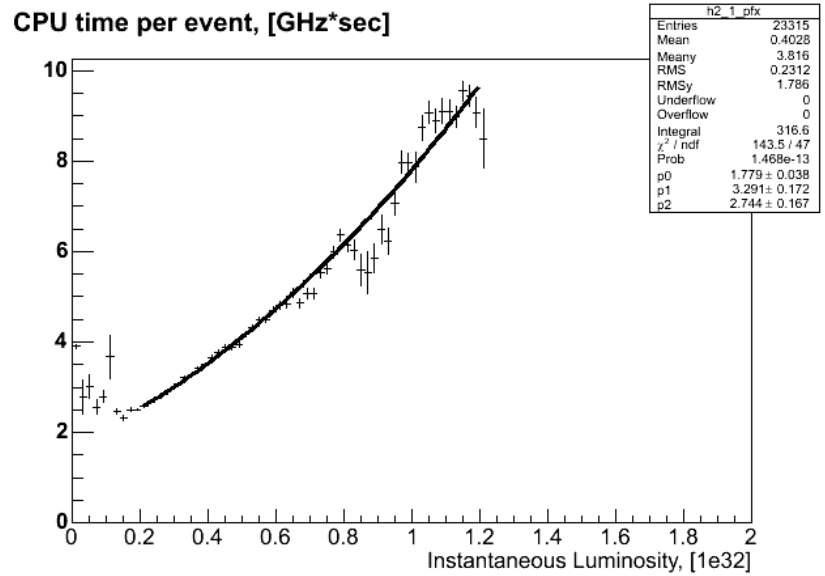
confidence in algorithms & simulation to make these extrapolations

CPU Performance

- Increased occupancy leads to higher processing times
- Expect linear ($\propto N_{TRK}$) + combinatoric terms ($\propto N_{TRK}^{k}$)



⊗



$$\Rightarrow \langle CPU \rangle_{peak\text{-}L = 3 \times E32} \approx 2.5 \times \langle CPU \rangle_{peak\text{-}L = 1 \times E32}$$

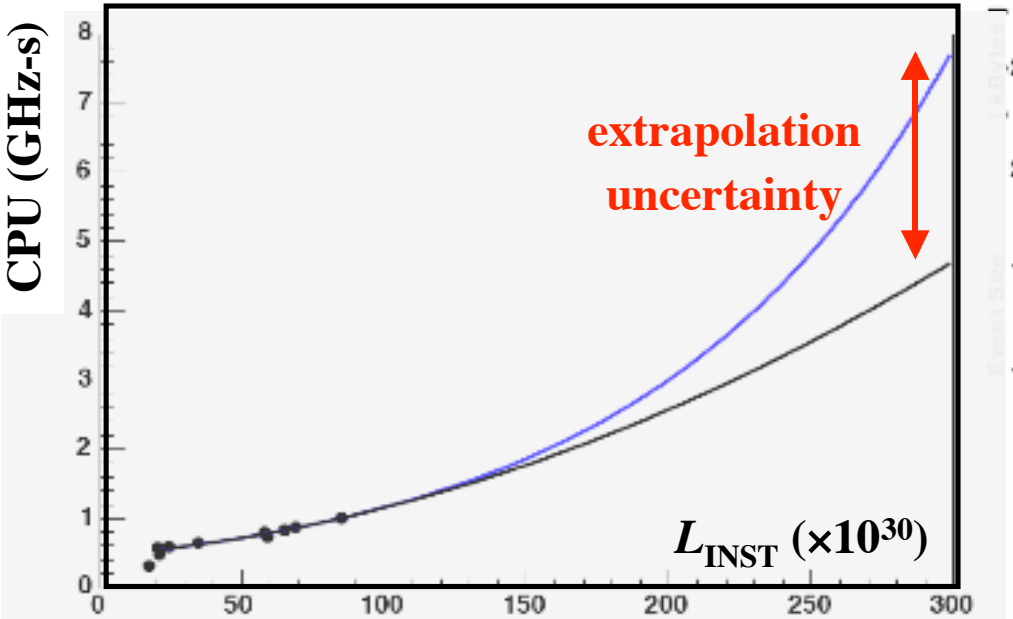
- ➢ Still dominated by linear growth term
- ➢ Manageable CPU resource requirements for lifetime of experiment

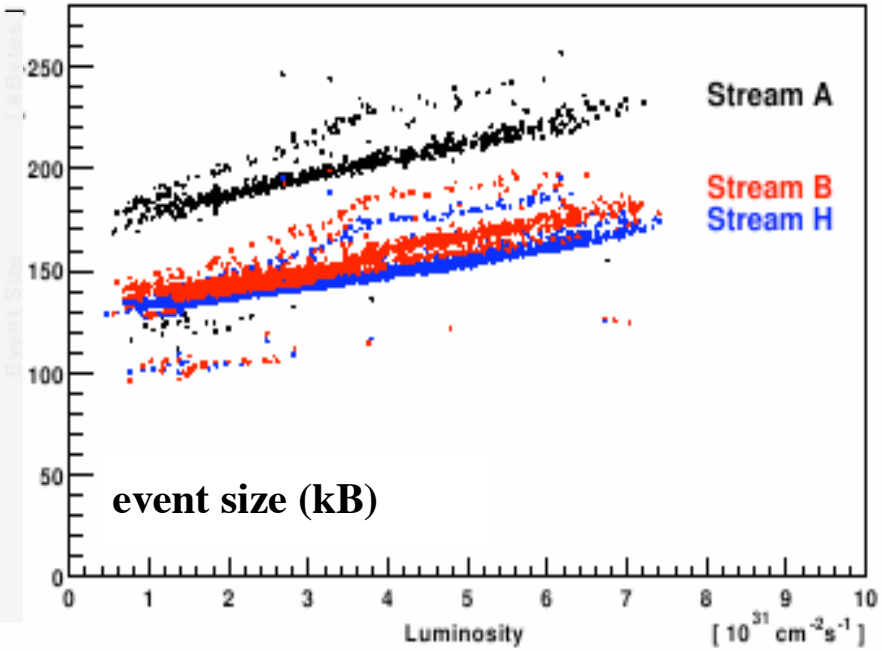
Online Implications

- Offline software & algorithms are used online :
  - CPU performance more critical
  - event size critical
  - dirtier environment

- Strategy :
  - Optimization
  - CPU & I/O trade-offs
  - Remove unnecessary reconstruction (e.g. z-side clustering)
  - Different thresholds online & offline
  - *Different algorithm online*

planning to avoid this !

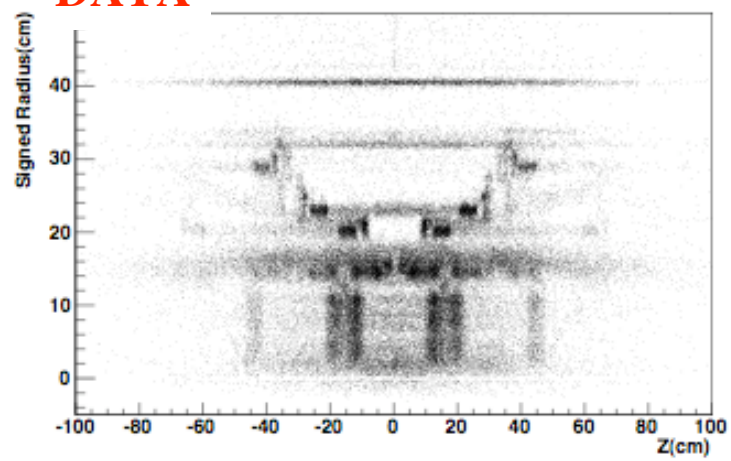




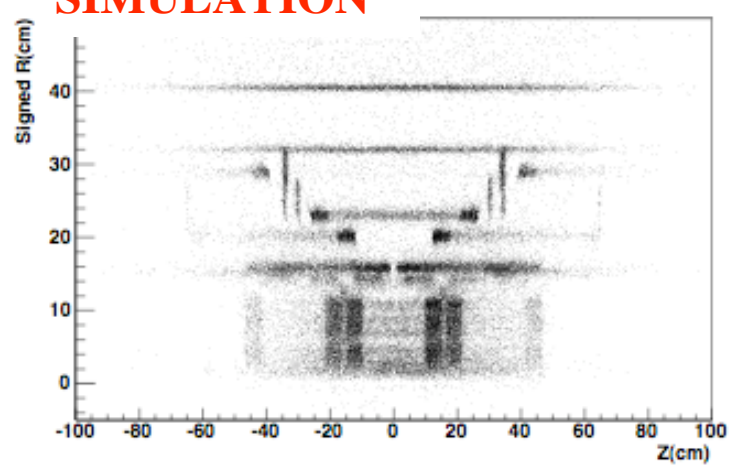
Simulation Refinements

- Simulation infrastructure is now very stable.
- Material tunings converging (good to ≤ 15%)
- Recent improvements to tracking simulation:
  - t_0 distribution
  - hit resolutions
- Some work needed to include and tune description of new detectors (CPR).
- Certain analyses use fast simulations; can be validated against full GEANT simulation.

DATA

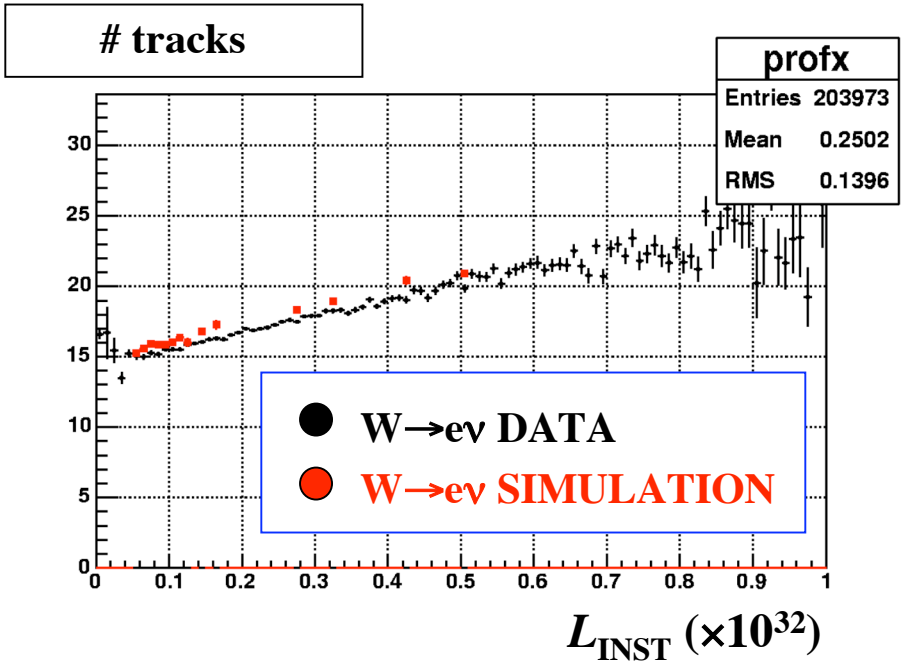


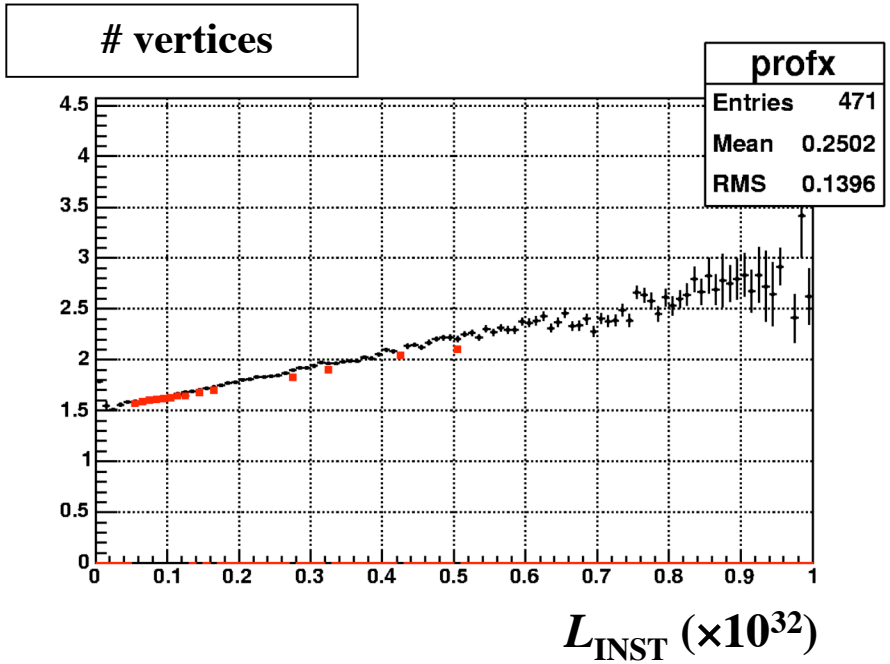
SIMULATION



Simulation Refinements

- **New** : run-dependent Monte Carlo
- Tracks changes detector configuration.
- Simulate the effects of higher instantaneous luminosities through the addition of overlapping min-bias events : Poisson distributed according to L_{INST} (run).
- MC run list chosen to reproduce L_{INST} distribution of data.
- MC datasets easily *extended* (not *replaced*) as further collider data is accumulated.





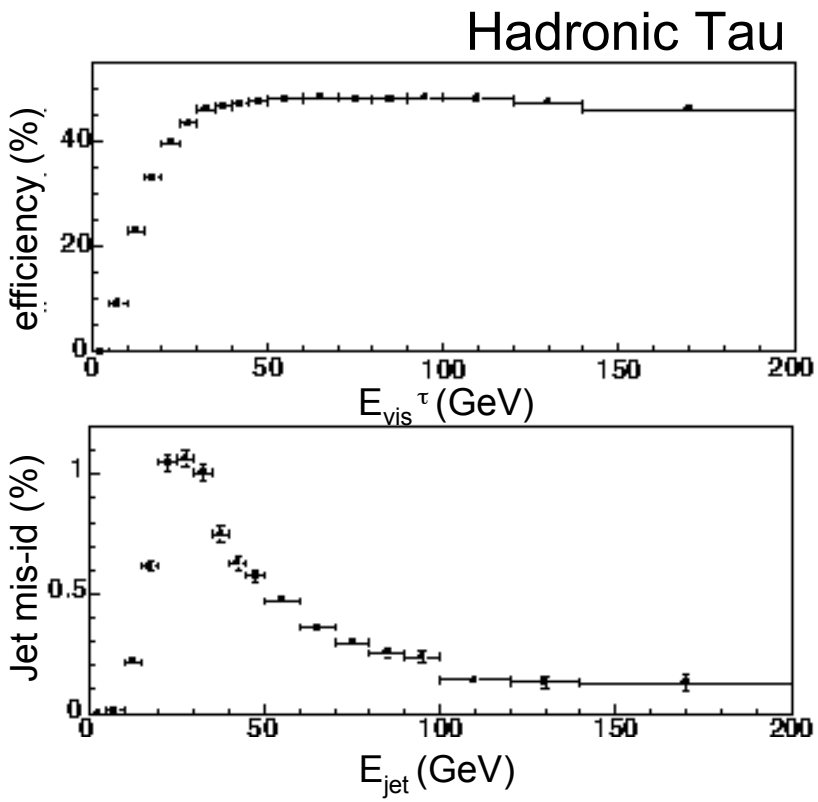
Conclusions & Outlook

- Low-level software & algorithms development is approaching asymtote.
  Only a few exceptions :
  - Forward tracking → enhanced lepton ID, b-tagging performance etc.
- Of course, there are cautions :
  - Effect of increased instantaneous luminosity : we have planned for this, but extrapolations are non-linear with sometimes large uncertainties
  - (Un)-anticipated changes in detector configuration & performance

- High-level developments : much more work to be done, but mainly in user-domain.
  For example :
  - Final sets of jet & other corrections.
  - Advanced analysis techniques - neural networks etc.

- CDF production reconstruction is ready for $1fb^{-1}$ analyses … and beyond.

Backup


Hadronic Tau
efficiency (%)
40
20
0
0
50
100
150
200
E_{vis}^{τ} (GeV)
Jet mis-id (%)
1
0.5
0
0
50
100
150
200
E_{jet} (GeV)