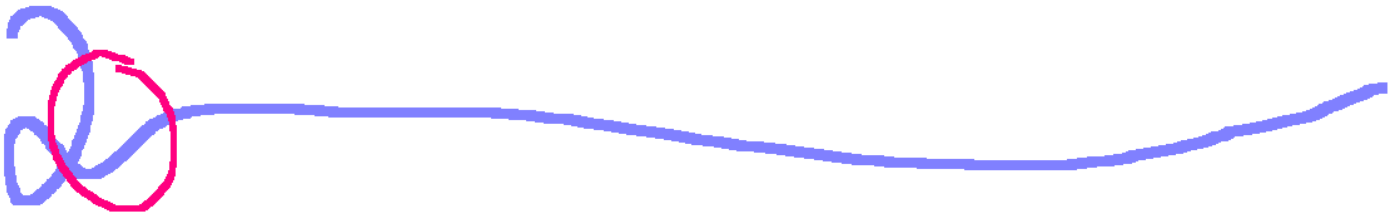
Progress on tau reconstruction software



Yu. Gershtein
Brown U
April 3, 2001

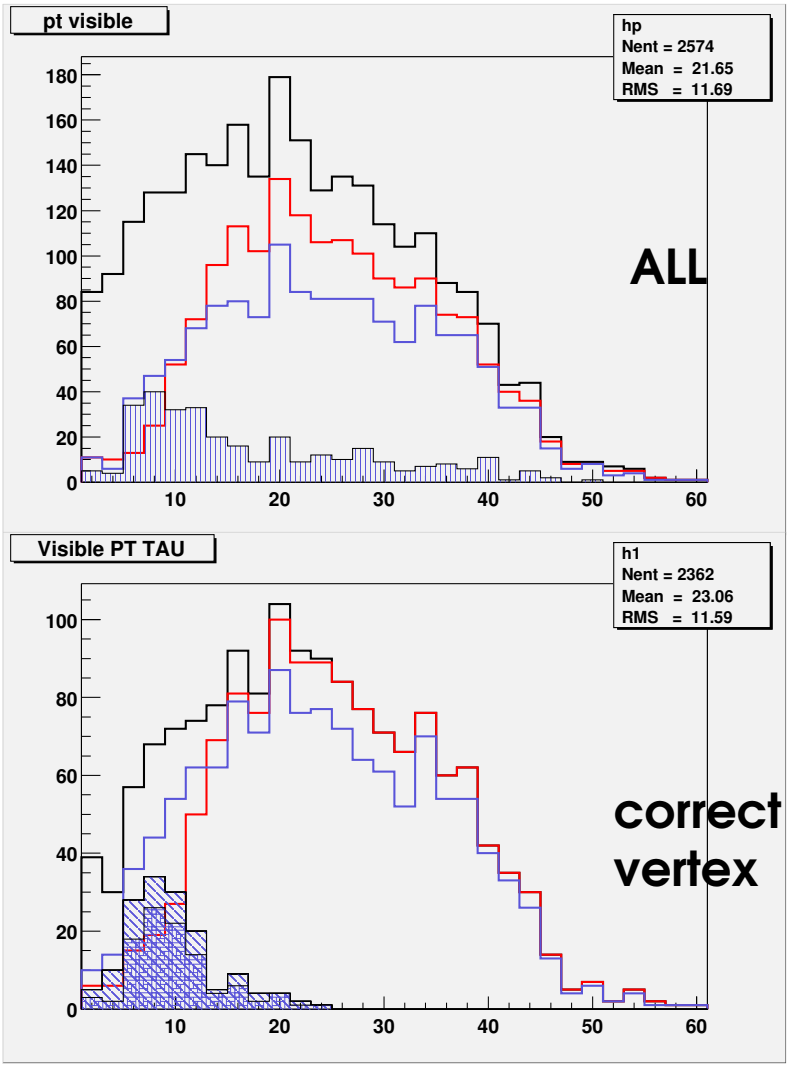
Outline

- P08
- What have we learned during last few months
- What absolutely has to be present in P09
- What is the most efficient way of making it happen
- Details of "P09" code
- Changes beyond P09

What is tau reconstruction

- P08: *(and P05)*
  - Cone clusters in calorimeter are the seeds
    - Et>8 GeV
  - Tracks are matched to the cluster
    - misvertexing reduces efficiency
  - PS clusters in a wide window are matched to the cluster
    - ghosts, misvertexing
  - Global jet variables are calculated
    - rms, profile, cal. layer fractions
- Structure is not perfect, I.e. there is no logical separation between reconstruction and calculation of variables relevant for TauID; many calculations are duplicated

Efficiency and Misvertexing



Black – all MC taus

Red - with Cal clus

Blue – with 4 GeV track

Blue hatched – with 4 GeV track but without a cluster

- Eventually when tracking and veretxing is perfect should not be a problem. But need cure now!

Tracking in Taus

taureco

- tau tracks are all tracks with
  - $|\Delta\varphi| < 10°$
  - $|\Delta\theta| < 10°$
- cuts are made on:
  - track RMS
  - number of tracks

Tracks from tau have right mass and width by const-ruction.
It doesn't matter much how the rest of tracks distributed:
What matters is their total transverse energy

proposed

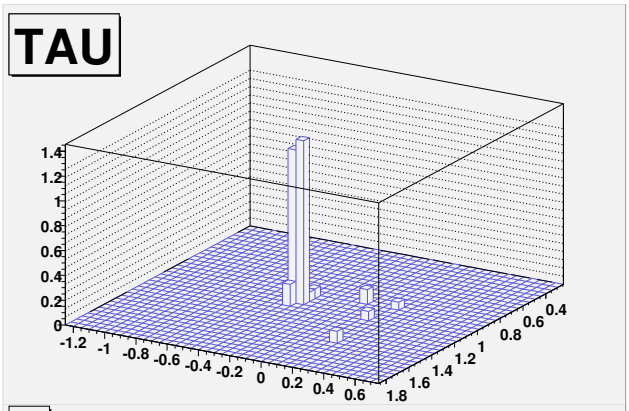
- order tracks in AOI by p_t
- start with highest p_t track
- go down the list and add tracks only if:
  - Ntrk ≤ 3
  - mass < 2 GeV
  - track p_t > 1.5 GeV
  - charge is O.K
- Make cut on transverse energy of all tracks which are left: $\sum_i E_i \sin\Theta_i$

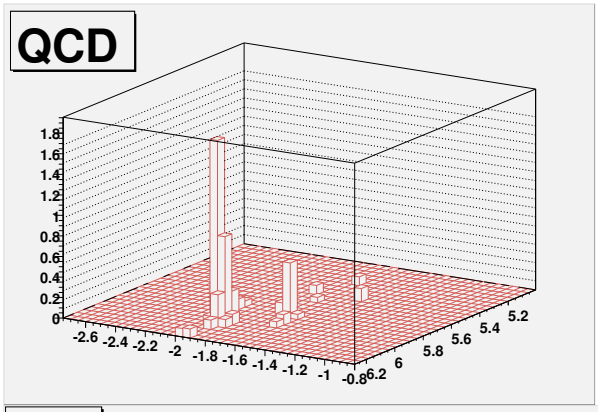
Energy in EM3

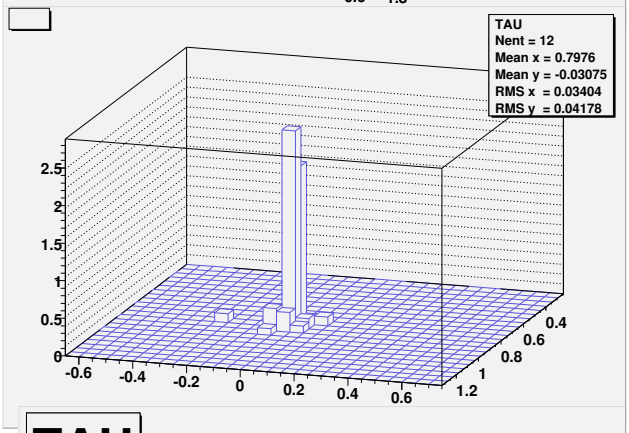
- EM3 is the most finely segmented portion of the calorimeter and more than half of all EM energy is deposited there
- Can try to distinguish
  - few high-pt photons (tau)
  - many low-pt photons (QCD)
- Implemented Nearest Neighbor clustering of EM3 cells of the tau jet

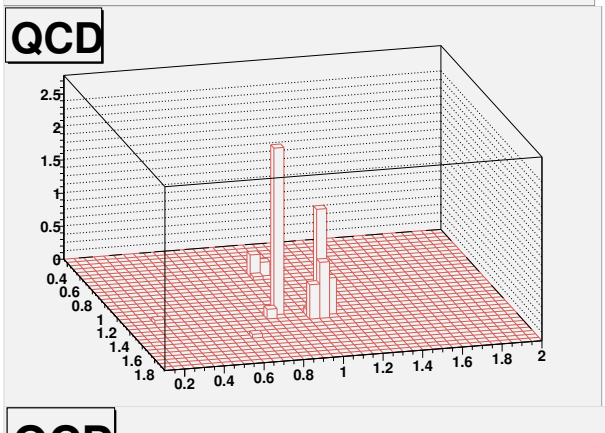
Displays of "Typical" Events

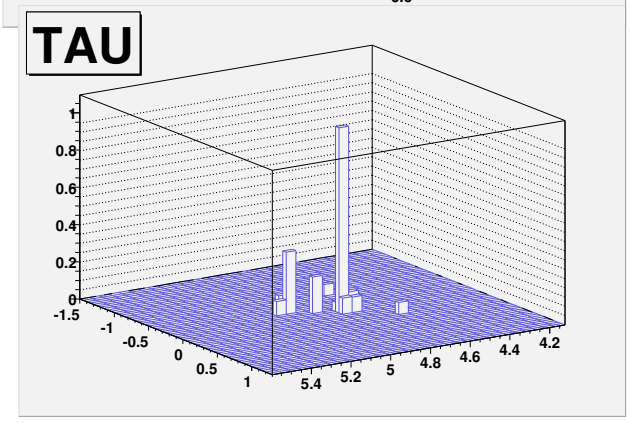
- Cell threshold: 50 MeV -> high enough to kill all noise

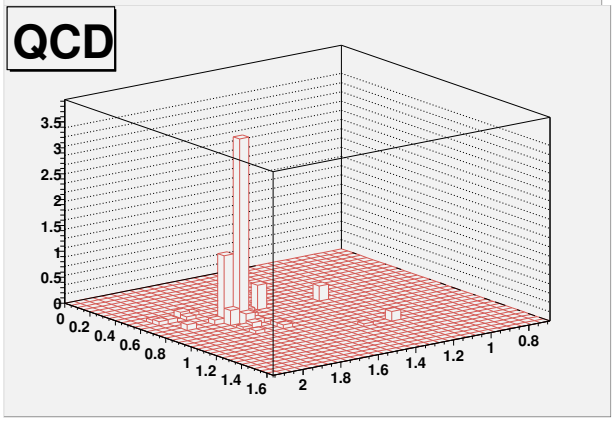






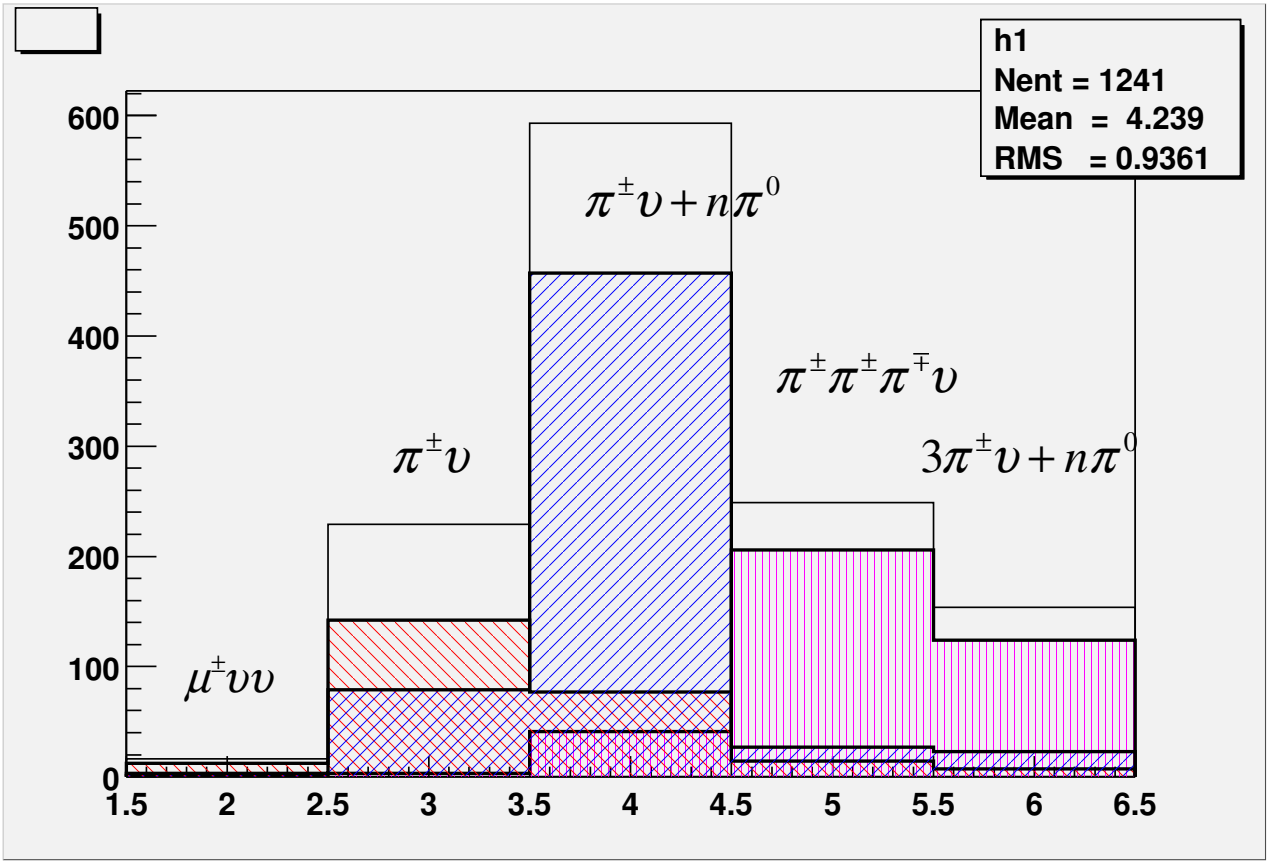






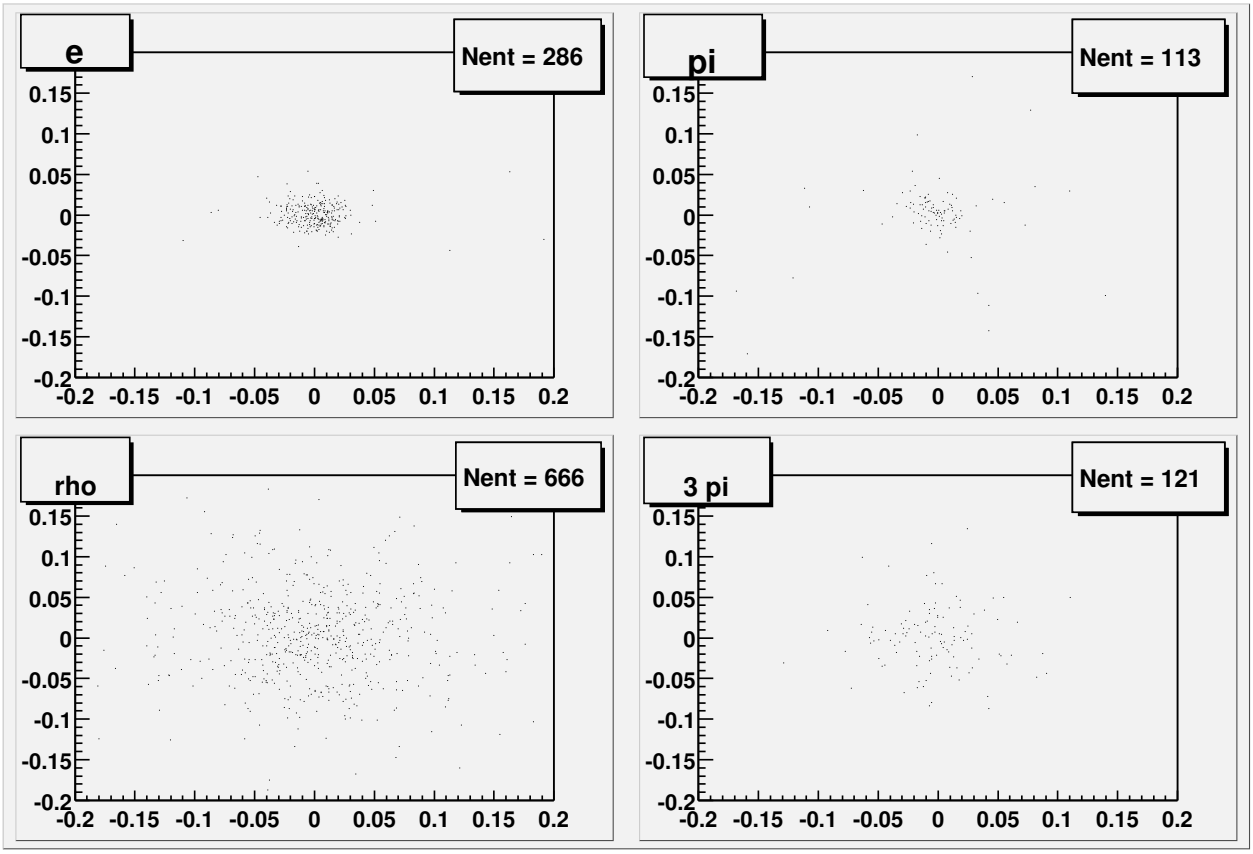
Three Tau Types

1. $N_{trk}=1$, $Et_{em3clus}<0.8$ GeV
2. $N_{trk}=1$, $Et_{em3clus}>0.8$ GeV
3. $N_{trk}>1$



EM clusters

- Match between extrapolation of a track into PS and EM sub-cluster (for $3\pi\pi^0$ track direction is taken at the vertex)



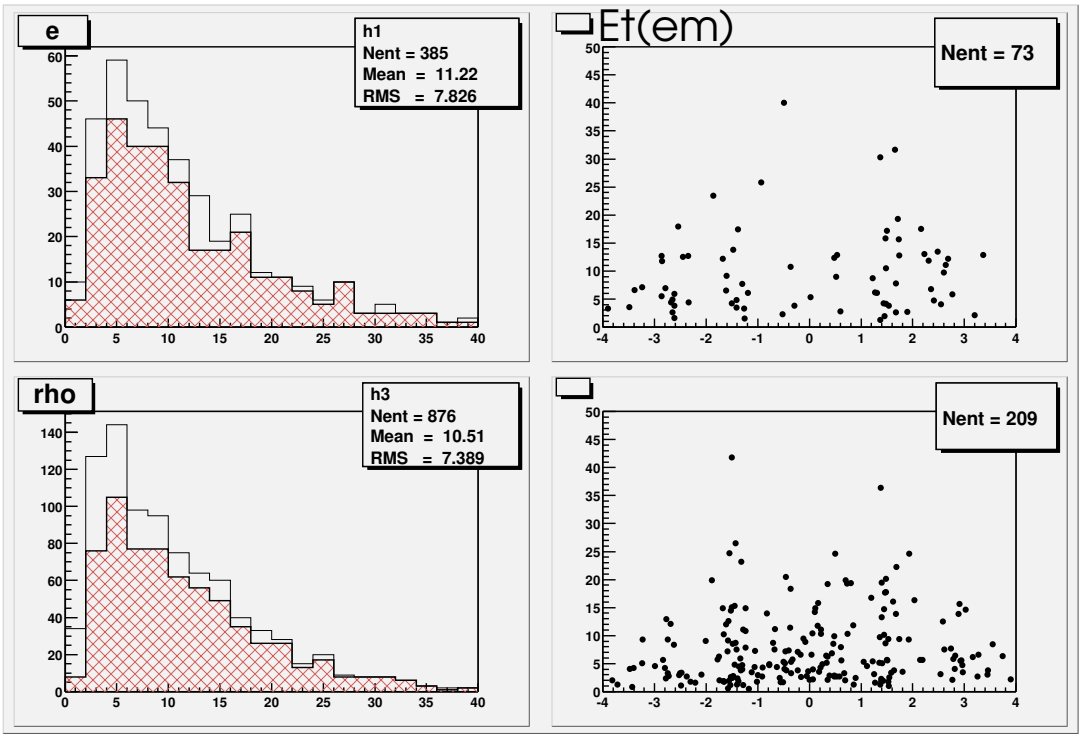
Matching Preshower

- PS match same way as in EM ID:
  - 0.05 eta-phi for CPS
  - 0.10 theta-phi for FPS

Open hist – all EM clusters

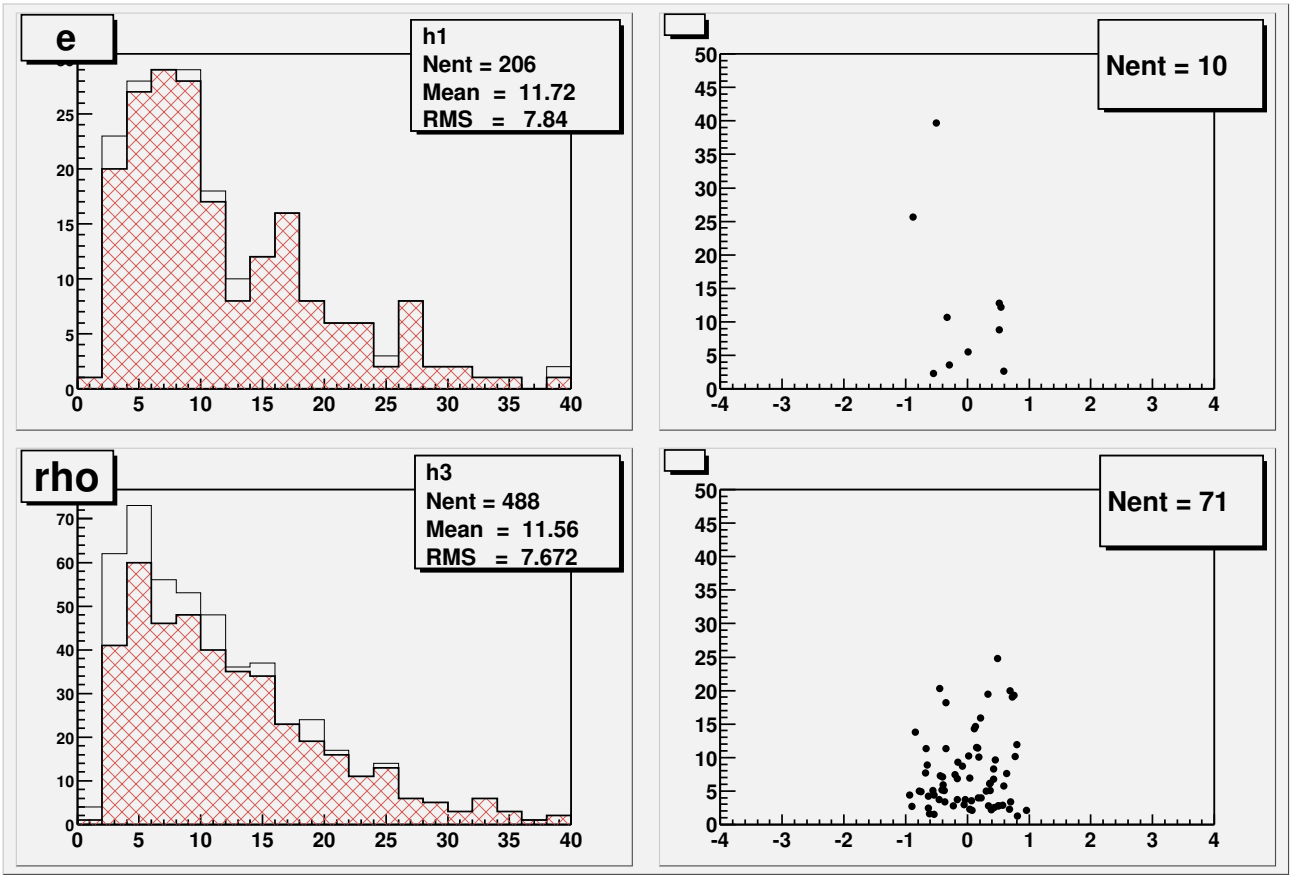
Red-hatched – with PS hit

EM clusters without PS hits



Et(em) eta (em)

CPS only



Eff for electrons – 196/206=95% -> too low! Matching?

For photons from rho - 85%

Proposed strategy

- Calorimeter clusters seeded from high towers and isolated tracks
- Tracks are matched without PV bias and divided into "tau"-tracks and "underlying" tracks
- EM clusters are found inside tau jets
- PS hits are looked for in a narrow window w/respect to the EM cluster

Why is it so important?

- I believe it is essential to have this structure in P09
- If we do not implement it in P09 DØ will not have certified and farm-produced tau Ntuples until p10 -> FALL
- If we succeed and put all PHYSICS in the root-tuples we can start developing ID TOOLS based on them

P09 Schedule

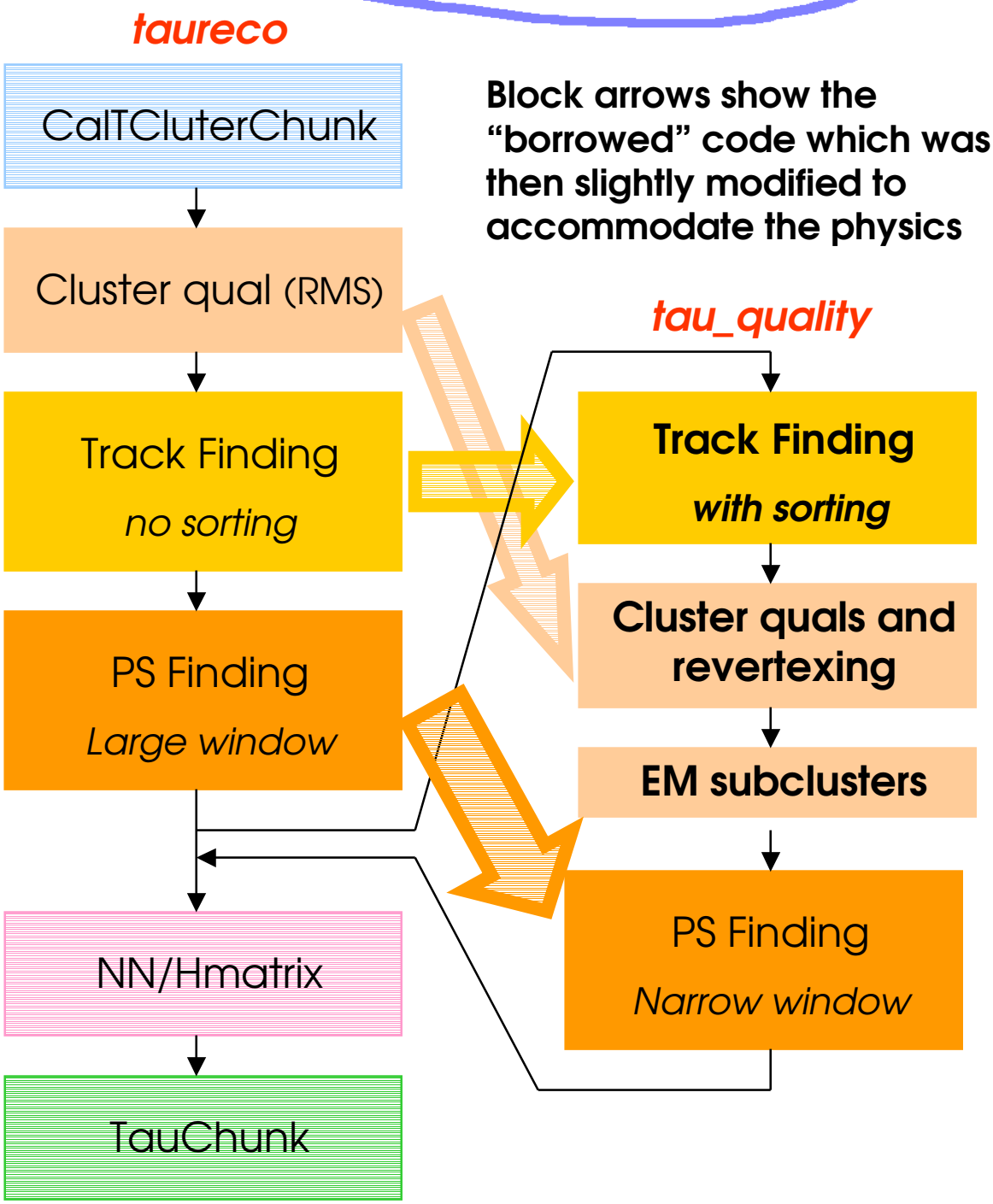
- **April 16, 2001:** All MAJOR improvements to DØRECO MUST be released in this week's test build. Any major modifications that miss this date must be targeted for next production release.
- **May 7, 2001:** No new functionality will be allowed in this week's test build. Only bug / mistake fixes will be allowed in this build. (and there better not be many...)
- **May 21, 2001:** Production version is created, and formal certification begins.
- **June 1, 2001:** Reports on certification samples at a SPECIAL Algorithms meeting (during an off-week). Declaration of the production worthiness of executable will then be made.

TIGHT, BUT POSSIBLE

The Design Problem

- The code that does these things was developed to calculate discriminating variables, not to reconstruct tau's
- The urgency to implement these methods in the reconstruction package was not recognized

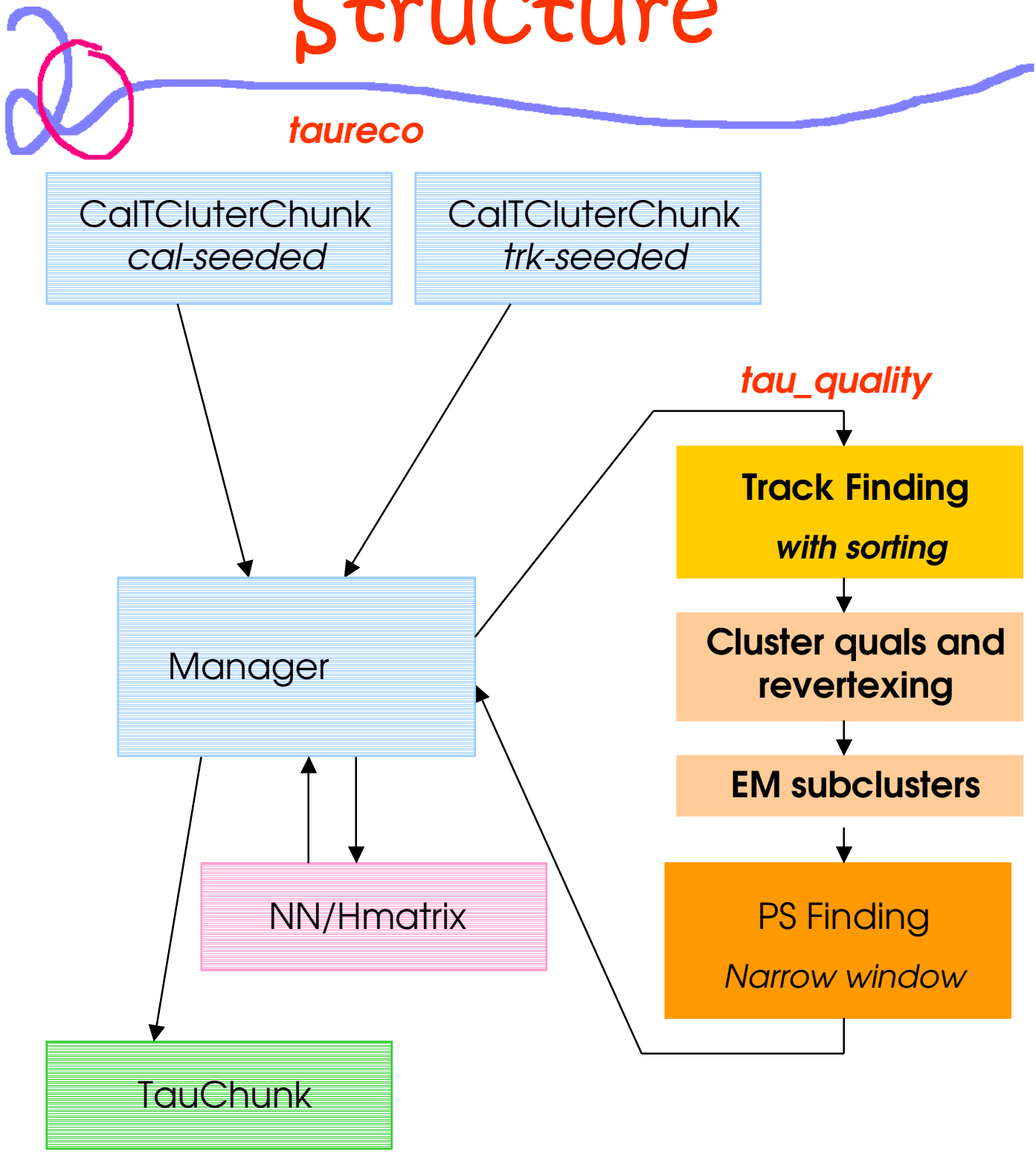
Current Structure



Is there a way to make it into P09?

- The code which has all the functionality exists (thanks to enormous help from Serban)
- In terms of program structure it inherited most of the shortcomings of P08 and added some new ones
  - the separation between "reconstruction" and "ID" is blurred even more
  - there are too many packages with names that do not often reflect their function
- But see bullet #1

New (not improved) Structure



Reconstruction

- two clustering packages
  1. **tau_clusterreco** - simple cone
  2. **tau_trkcluster**
     - seed is an intersection point of isolated track and preshower
     - output mimics simple cone
     - by design is not sensitive to misvertexing
- both packages have common RCP parameters for cone sizes, etc.
- Foreseen settings
  - Rcone = 0.4-0.5
  - Risolation cone = 0.7
  - Et min = (1) 6-8 GeV; (2) 1-2 GeV
  - Pt trk > 5 GeV $\sum$Et(0.7)<5 GeV

(may go to fraction…)

Reconstruction

- **tau_quality**
  - does most of the reconstruction
    -> bad name
    - track matching
    - cluster re-vertexing
    - EM sub-clustering
    - PS matching to EM clusters
  - also calculates quantities for TauID -> bad design
  - puts calculated quantities into container class **TauQualityInfo**

Reconstruction

TauQualityInfo : kinem::QualityInfo

```
enum {UNDEFINED, PI_NU, RHO_NU, THREE_PRONG} type();
Et(), EM12_Et(), EM3_Et(), EM4_Et()   "tauet_7", etc
Et_core(), EM12_Et_core(), EM3_Et_core(),
   EM4_Et_core()                      "tauet_3", etc
quality();
vector<int> Iquant()
    [0] - N(tau tracks)          "tau_nt"
    [1] - N(EM3 cells)
    [2] - N(EM3 clusters)        "tau_nclus"
vector<float> Fquant()
    [0] - Et(track 1)            "tau_ett1"
    [1] - Et(track 2)            "tau_ett2"
    [2] - Et(track 3)            "tau_ett3"
    [3] - M(tracks)              "tau_mt"
    [4] - Et(non-tau tracks)     "tau_ettr"

    [5] - Et(tau em part)        "tau_empt"
    [6] - Mass(tau em part)      "tau_emm"
    [7] - eta (tau em part)      "tau_emeta"
    [8] - phi (tau em part)      "tau_emphi"
```

Reconstruction

TauQualityInfo : kinem::QualityInfo

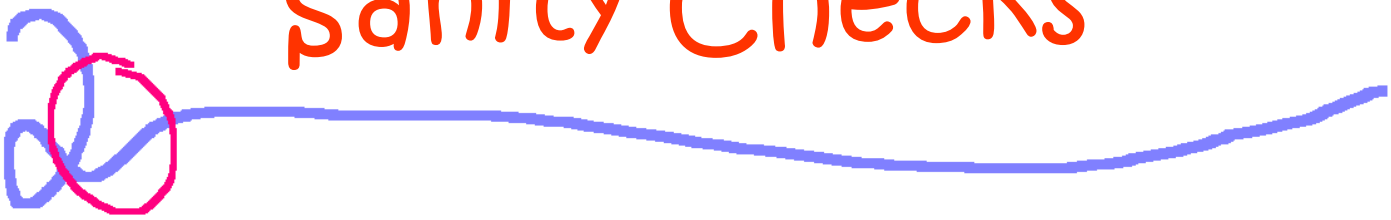
`vector<float> Fquant()` - *continued…*

```
[9] - Z of the tau vertex          "tau_tz"
[10]- eta of tau tracks            "tau_teta"
[11]- phi of tau tracks            "tau_tphi"
[12]- dphi at PS                   "tau_tphi1"
[13]- dphi at EM3                  "tau_tphi2"

[14]- eta of first EM3 cluster          "tau_3eta1"
[15]- eta of second EM3 cluster         "tau_3eta2"
[16]- phi of first EM3 cluster          "tau_3phi1"
[17]- phi of second EM3 cluster         "tau_3phi2"
[18]- Et of first EM3 cluster           "tau_3et1"
[19]- Et of second EM3 cluster          "tau_3et2"

[20]- eta of cluster core (re-vtx'ed) "tau_ceta"
[21]- phi of cluster core (re-vtx'ed) "tau_cphi"
[22]- et's of two hottest towers      "tau_q2t"
[23]- calorimeter RMS                 "tau_qrms"

[24]- eta of PS cluster (re-vtx'ed)   "taupseta"
[25]- phi of PS cluster (re-vtx'ed)   "taupsphi"
```

Reconstruction

- **taureco**
  - locates both **CalTClusterChunks**
  - locates PS and ChParticle chunks
  - creates **TauChunk**
  - loops over cal seeded clusters
    - *calls* **tau_quality**
    - *calls Hmatrix/NN*
    - *if RCP switch is set discards taus with* **quality()==POOR**
    - *inserts Tau in the chunk*
  - loops over track seeded clusters
    - *Checks if within 0.2 of cal seeded tau*
    - *calls* **tau_quality**
    - *calls Hmatrix/NN*
    - *if RCP switch is set discards taus with* **quality()==POOR**
    - *inserts Tau in the chunk*
  - Inserts **TauChunk** into event

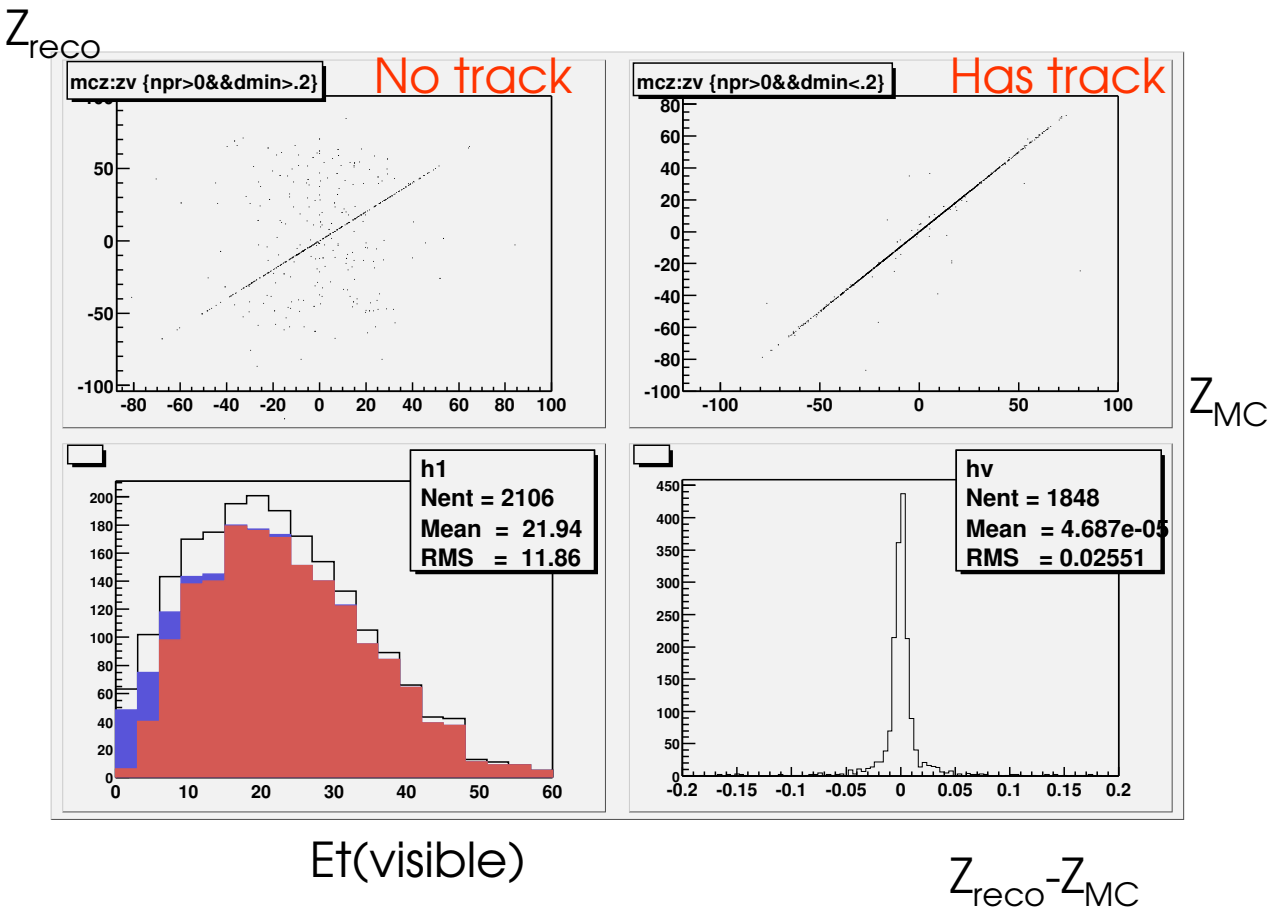
Sanity Checks



- A lot of code has been moved around
- Carefully checked if these reshuffling broke it

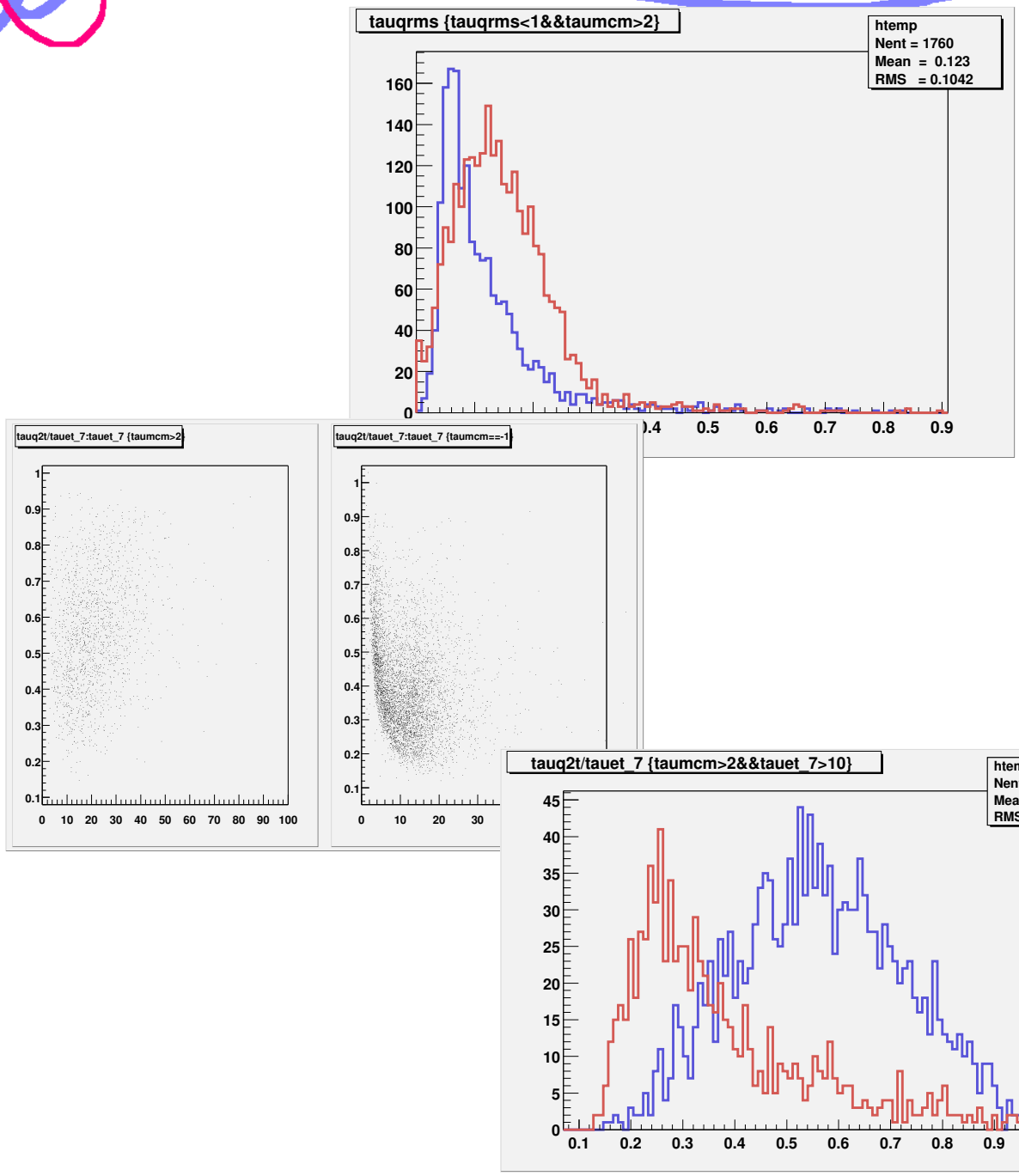
Finding Clusters

- Good!



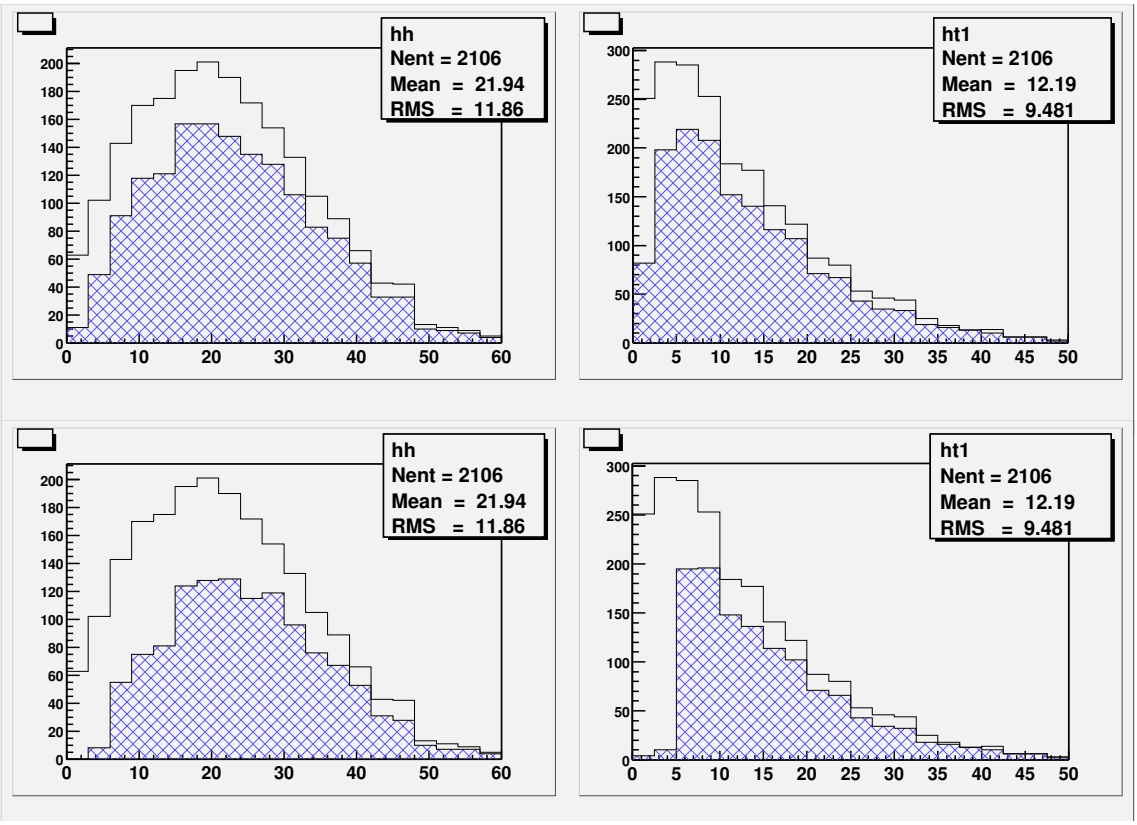
- Open histogram – all taus
- Blue – vertex is right (88% -> was 58%)
- Red – cluster is reconstructed

RMS and Profile



Matching tracks

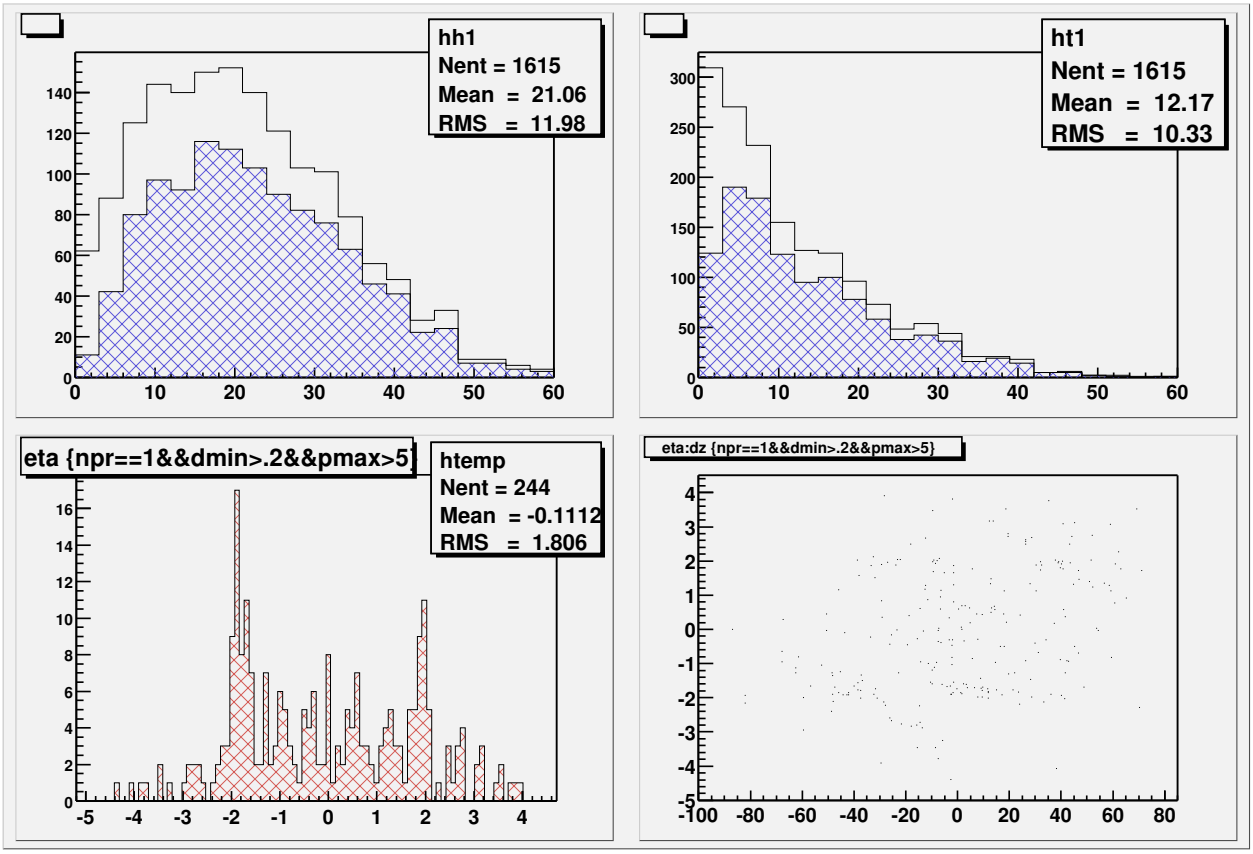
- As good as it was
  - pt>1.5 GeV = 75%
  - pt>5.0 GeV = 60%
- Some high pt tracks are lost



Et(visible) Et(MC track)

One Prongs

- Looks "familiar"

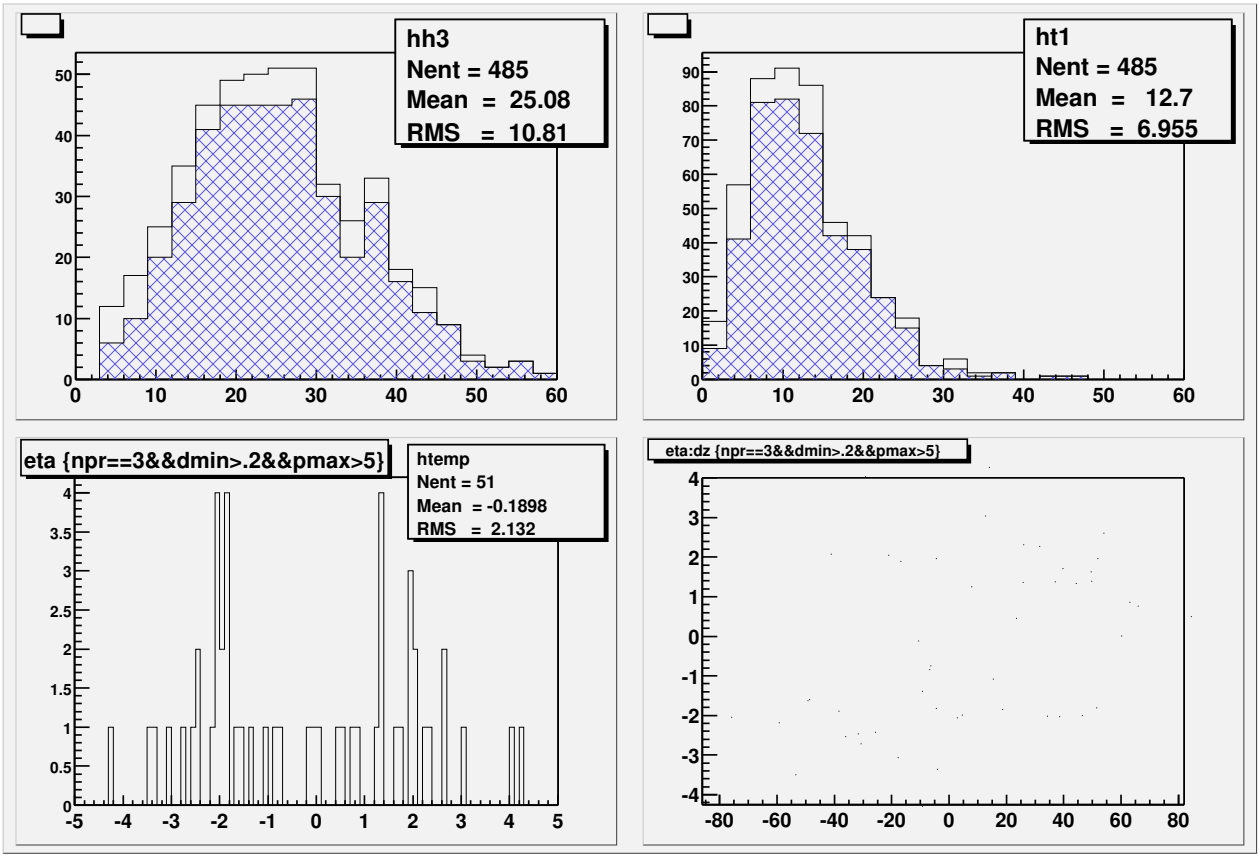


eta MC

Zvertex MC

Three Prongs

- Looks "familiar"



eta MC

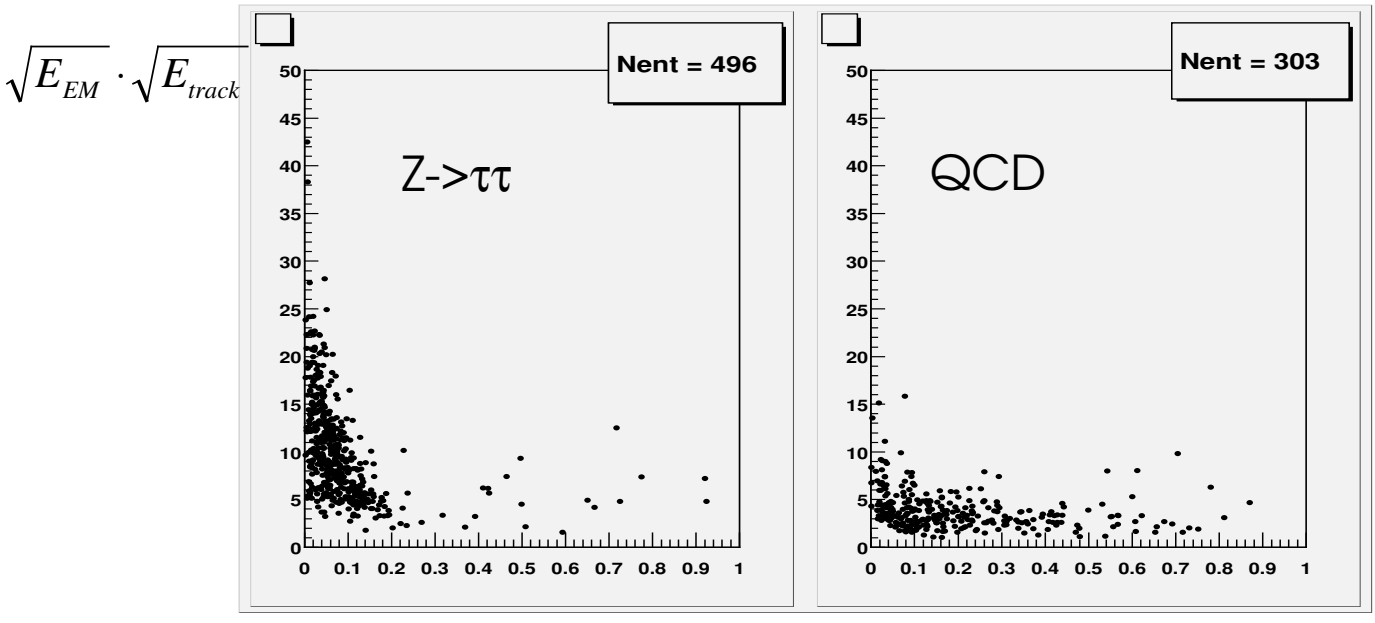
Zvertex MC

Mass (CC)

- For "RHO_NU" type

$$M = \sqrt{E_{EM}} \cdot \sqrt{E_{track}} \cdot \Delta\varphi$$

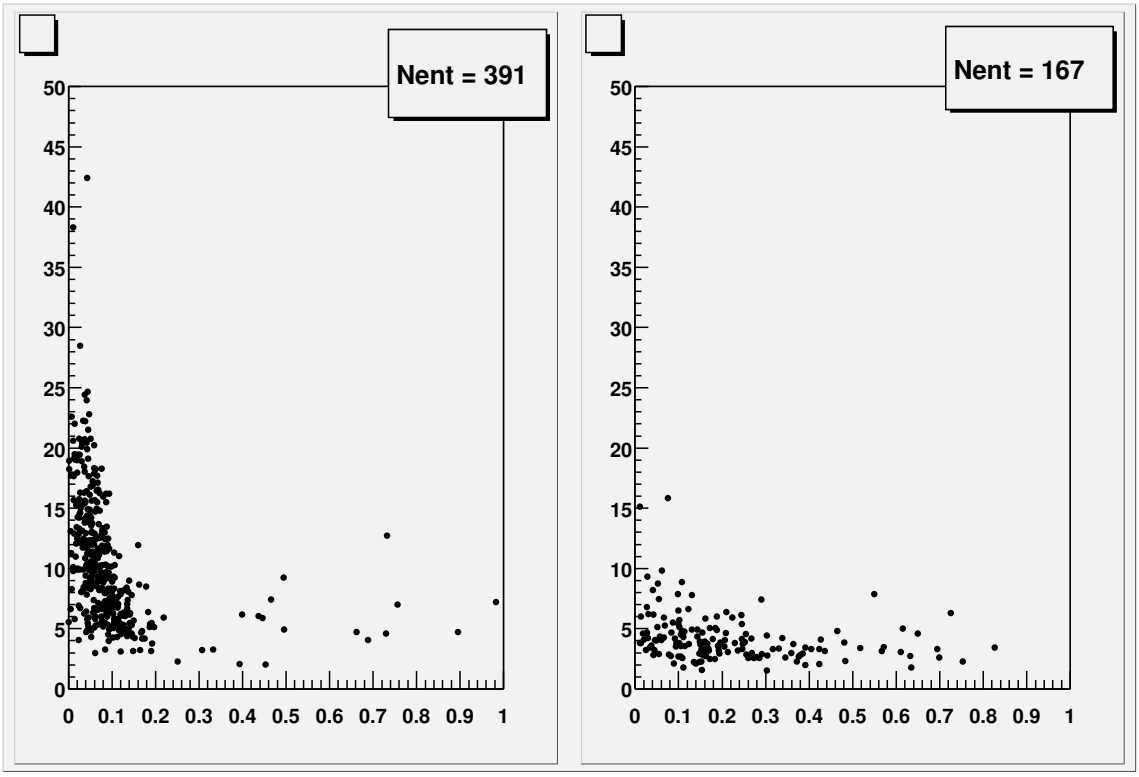
- Know EM sub-cluster energy and direction
- Know track energy and direction



Mass (CC)

- For events with matched CPS clusters can take direction from CPS

$$\sqrt{E_{EM}} \cdot \sqrt{E_{track}}$$



Plan for P09

- **April 9:** the new version of code is released. The released code is examined for crashes and used for studies to determine
  - cone size
  - feasibility of isolation cut in SimpleCone
  - isolation in ChargedParticle
- **April 16:** freeze TauQualityInfo contents and RCP parameters and produce large Ntuples of $Z \rightarrow \tau\tau$, $\chi^0\chi^\pm \rightarrow \tau\tau\tau$ and QCD jets
  - need disk space!!! 10 G
- **May 7:** all minor changes are implemented (I.e. particular choice of variables for NN from TauQualityInfo – no new calculations)
- **May 21:** we are on time and happy with the debugged code ☺

Beyond P09

(must start before June 1st)

- Move around the code yet again ☹
  - separate reconstruction and identification, reduce number of packages and rename some of them
  - …
- On the bright side: no changes in physics algorithms -> Ntuple content and all distributions (except NN) should remain the same -> very easy to debug
- Proceed full speed on development of a "blessed" ID tool which would start working on the Ntuple and eventually will make it into a framework package (but should retain possibility of running in both modes)