Multiple showers with the Threshold Accepting method

- TA for multiple showers: algorithm and cost function
- 1st test: Cheat scoring, greedy algorithm with 100 GeV qq
- Algorithm refinements: Single d-quark events at 50 GeV
- Algorithm tests: qq 100 GeV
- Next steps

Simulated annealing

Physically motivated method to quickly approximate the global minimum of a function ("cost") of a discrete system.

Variant: Threshold Accepting (TA).

choose **initial configuration**;

for (**threshold T =** $T_0 \geq T_1 \geq T_2 \geq \ldots \geq T_n = 0$) {

choose **new configuration** which is small perturbation of the old configuration;

if (cost(new configuration) < cost(old configuration) + T)

old configuration = new configuration;

}

(journal of computational physics 90, 161 1990)

Threshold accepting: implementation in the hadronic clustering

initial configuration:

assign the seeds to every shower

NEW!

new configuration:

randomly associate some unassociated cluster to some track, or un-associate

thresholds T = chosen adapted to cost differences and running time resource

Cost function

1. Energy\momentum costs:

- After seed setting: Probability that track with momentum p is correctly assigned to cluster with energy E:

  $$P_{E\backslash p} = 1/\sqrt{(2\pi\sigma^2)}\ \mathrm{Exp}\{-\tfrac{1}{2}(E-p)^2/\sigma^2\}$$

- with estimated cluster energy uncertainty

  $\sigma = 0.7\ \sqrt{p}$ [GeV] (as in standard algorithm)

Cost function

2. Geometrical costs:

- Joining (breaking) one good link with score s, the probability of reconstruction correctness increases (decreases),

$$P \to P \bullet f(s)$$
$$\to P/f(s)$$

- Function f(s) to be determined ...

Cost function

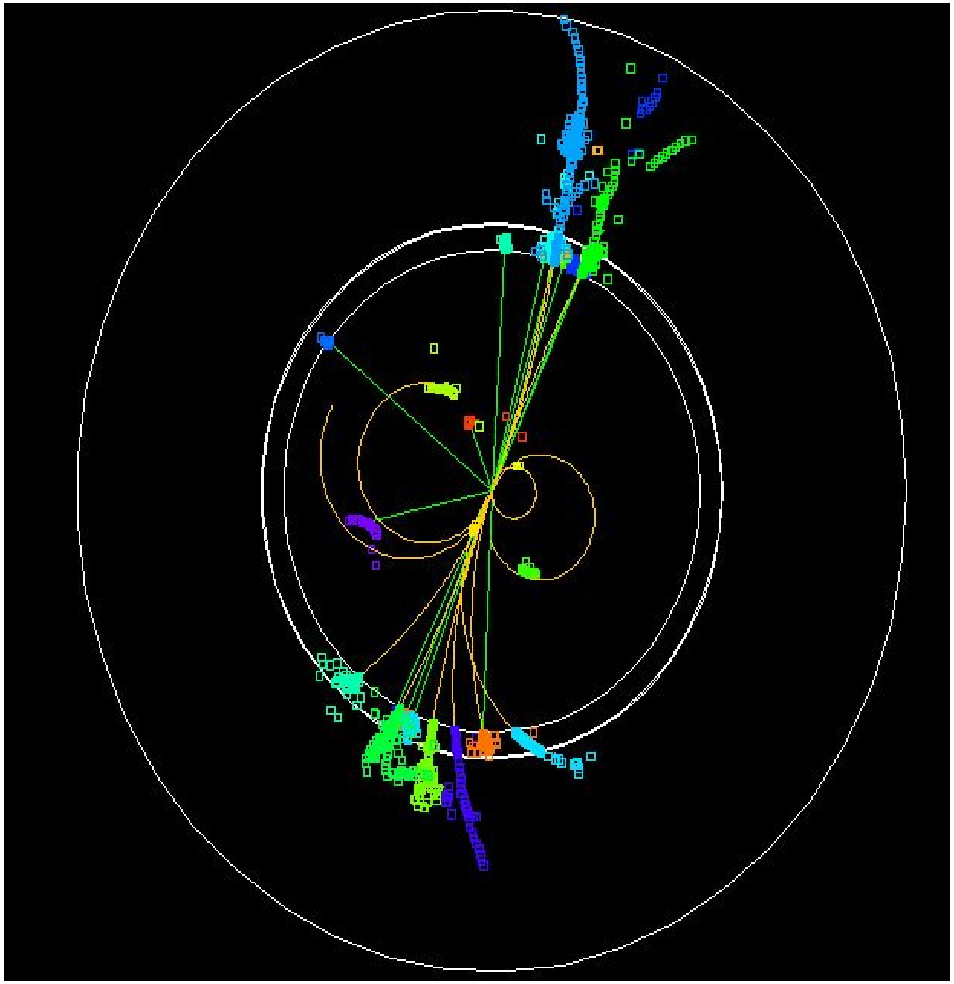
Put together:

for every joined (broken) link

- $P = (1/\sqrt{(2\pi)}\sigma \, \mathrm{Exp}\{-½(E-p)^2/\sigma^2\})^{\alpha} \cdot (/) \, f(s)$,
  α expresses relative reliability of E\p- vs. geo-metrical costs
- To be maximized! Or minimize the negative logarithm:
  cost function = $\alpha \cdot (c + \sigma + ½(E-p)^2/\sigma^2) - (+) \ln f(s)$

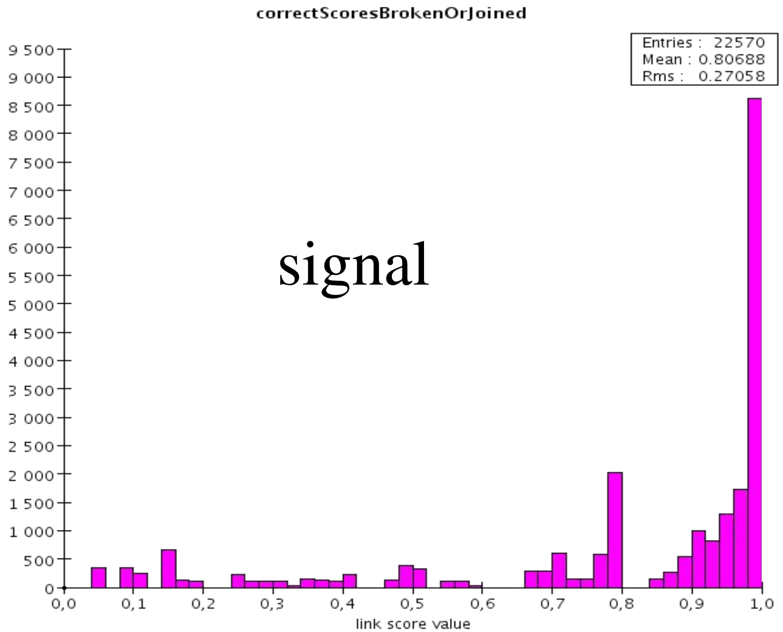
1st test: Cheat scoring, greedy algorithm with 100 GeV qq

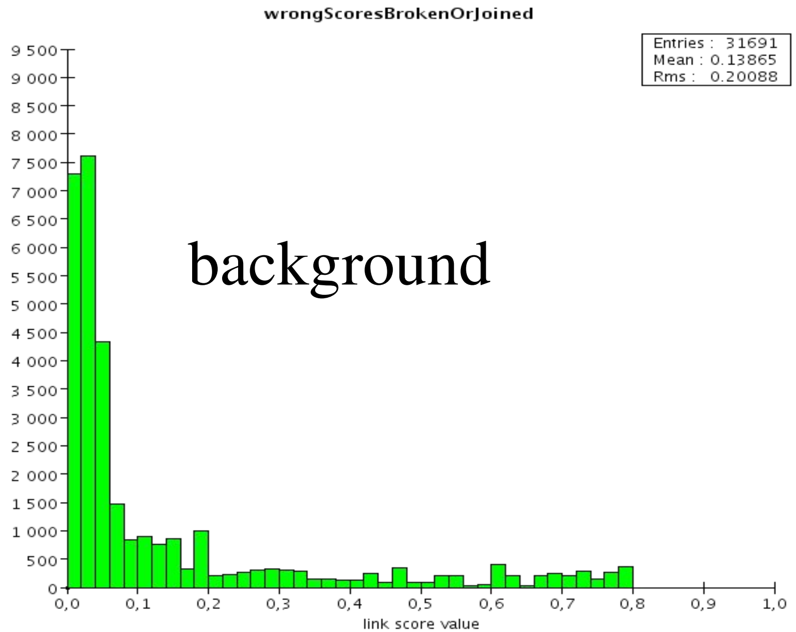
- Cheat scoring: link score = 1 for correct link (= 0 for incorrect link)
- Therefore: every cluster linked to at most one seed, no annealing necessary, T={0} (10000 times) with no E\p costs and simply subtract(add) 1 when joining correct (incorrect) link

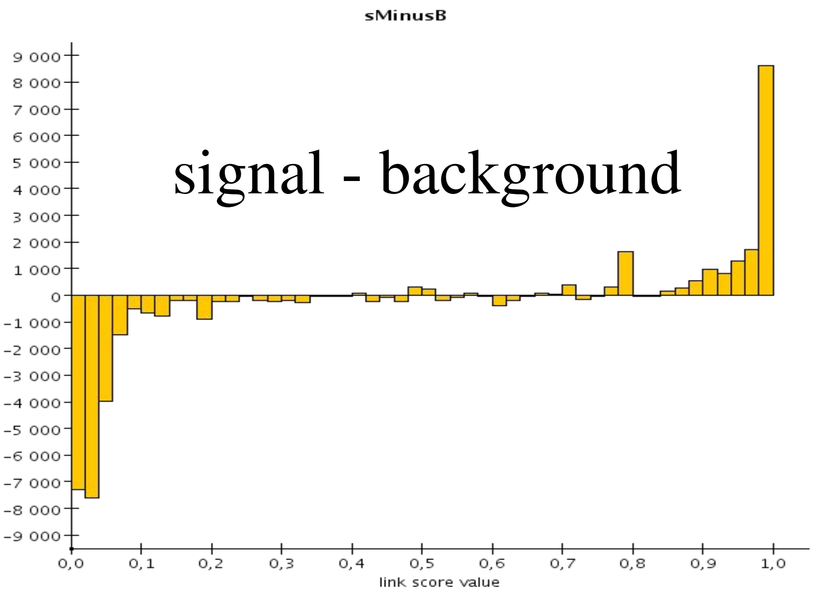


qq 100 GeV event 0

Real scoring: looking for f(s)



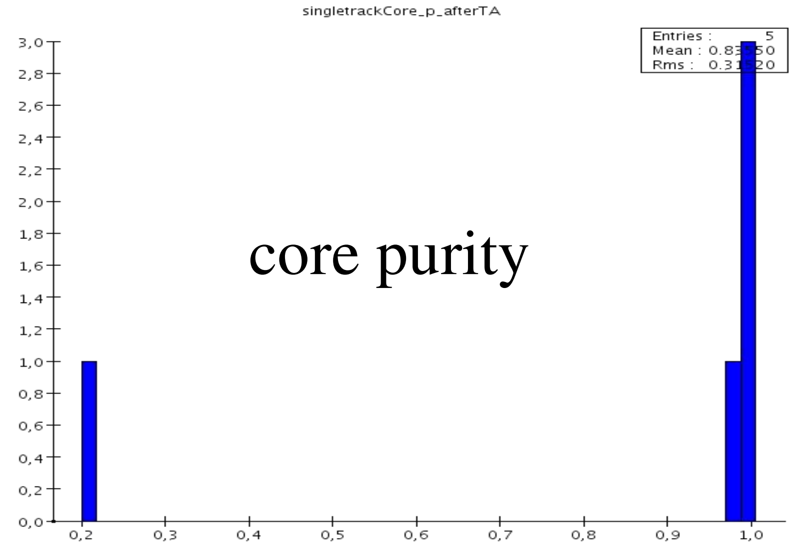


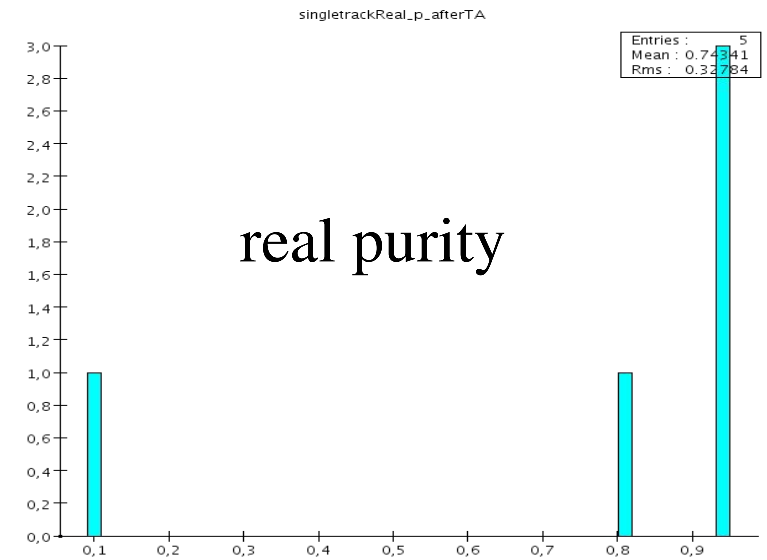


- tried to exploit scoring quantitatively by adding (signal-background) in given score bin when breaking link
- however: simple recipe gave better results: add 1 (subtract 2) when breaking link above 0.7 (below 0.4)

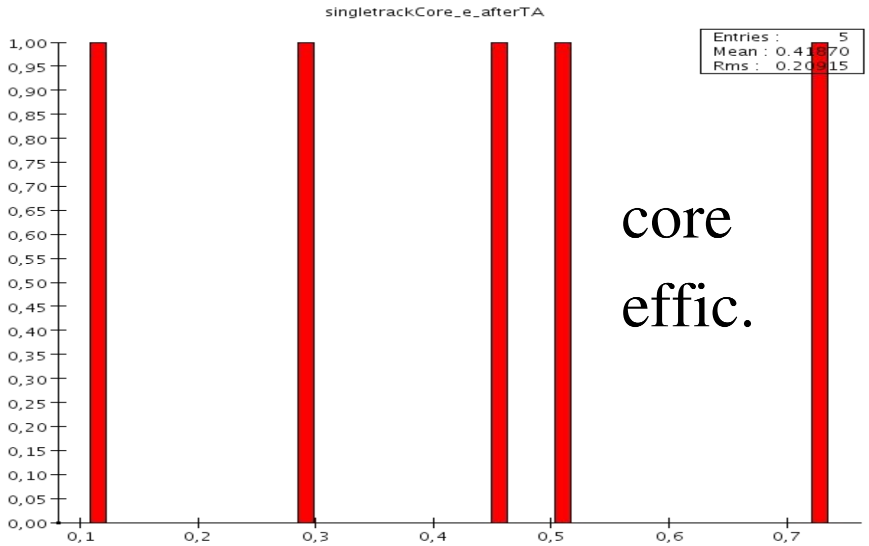
single d-quark at 50 GeV, single events

- Shown: file 0, event 0; best α=0 !
- differences to Mat's shower building:
  - follow not only the maximum link
  - no “full propagation”
  - work on all tracks simultaneously
- purities: seem ok

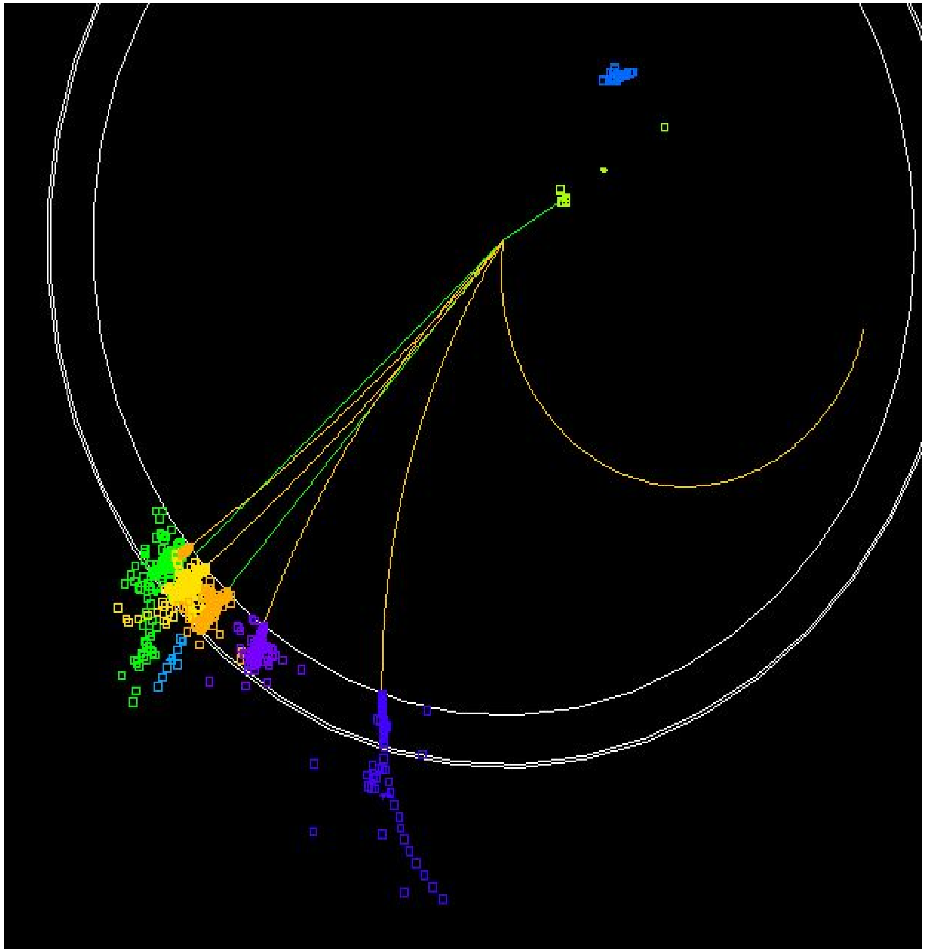




single d-quark at 50 GeV, single events

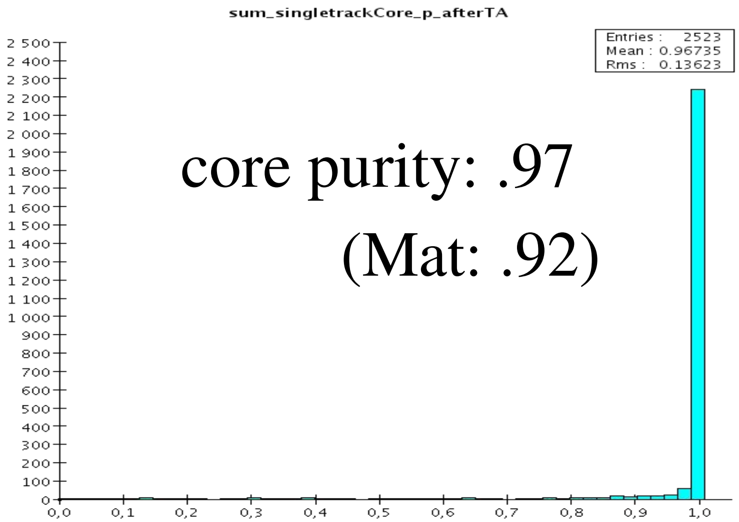


- efficiencies: to be improved

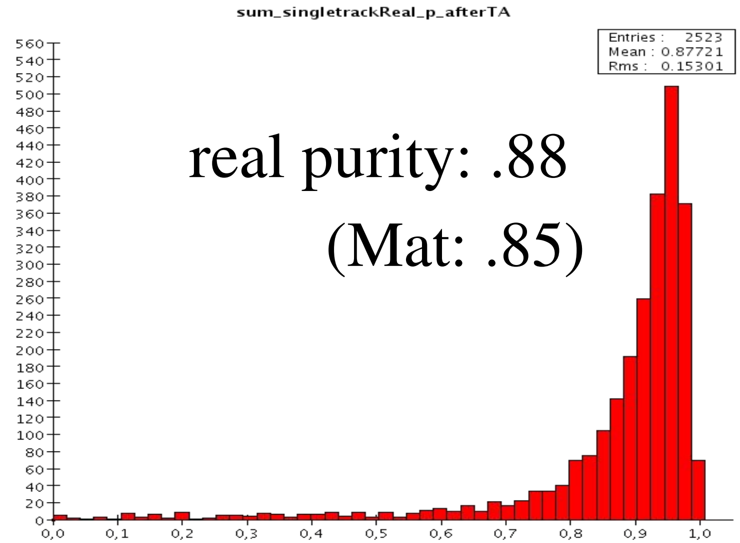


qq 100 GeV, α=0, 500 events

count non-existing link with score 0!

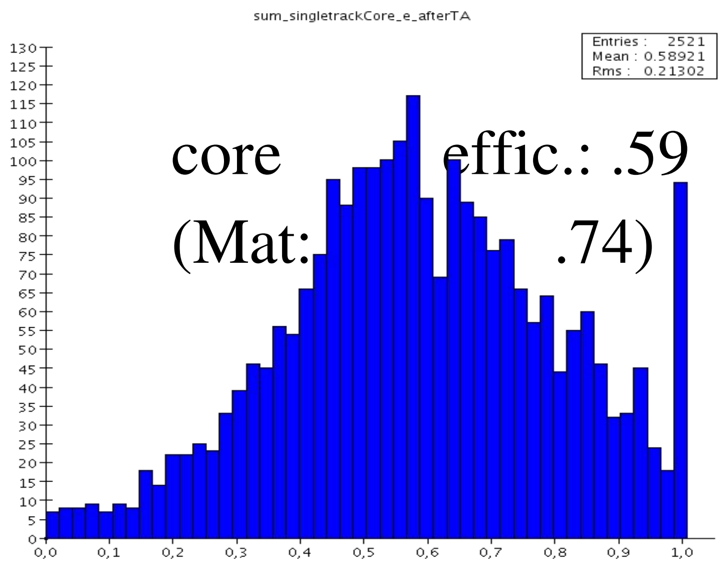


core purity: .97
(Mat: .92)

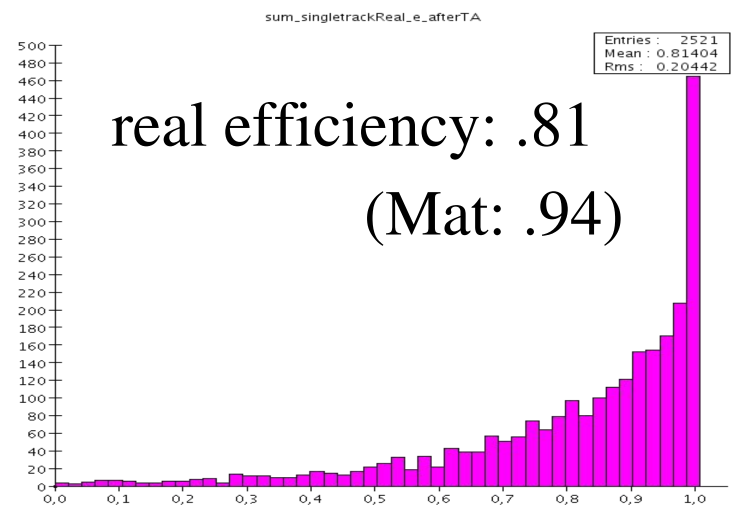


real purity: .88
(Mat: .85)

- good purity



core effic.: .59
(Mat: .74)

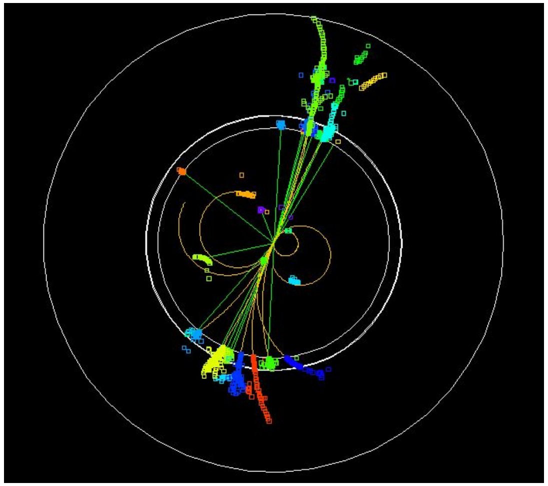


real efficiency: .81
(Mat: .94)

- efficiency to be improved

Next steps

- Quantitative studies
- Reduce running time
  - only useful "small variations"
  - cheaper data structures
  - local data structures
  - charged had. energy calculation: simpler ... incrementely
  - better random numbers ... deterministic version?
- Improve efficiency: more local calculation
- Resolution calculation qq 100 GeV



- qq 500 GeV