2007 International Linear Collider Workshop

DESY, Hamburg, 30th May – 3rd June 2007

The LCFIVertex package

- ***Scope of the package***
- ***Validation results from fast MC SGV***
- ***Performance obtained with MOKKA / MarlinReco***
- ***Results for different tracking input options***
- ***Influence of hadronic interactions***

Sonja Hillert (Oxford)

on behalf of the LCFI collaboration

The development team

The LCFIVertex package was developed by

Sonja Hillert – *coordination and system test*

Ben Jeffery – *ZVTOP, LCIO/Marlin interface, working classes design and testing*

Erik Devetak – *flavour tag inputs calculation and testing, MC jet flavour*

Mark Grimes – *flavour tag procedure, vertex fitter*

Dave Bailey – *neural network code*

The authors thank the LCFI physics group for help and advice, in particular

Dave Jackson (advice on ZVTOP) and

Victoria Martin (test of vertex charge procedure for c-jets)

We would also like to thank

Frank Gaede, Ties Behnke and Norman Graf for fruitful dicussions

Dennis Martsch for producing a MOKKA test sample on the GRID

Alexei Raspereza for advice and for extending the tracking to provide covariance matrices

Introduction

- **The LCFIVertex package provides:**
  - **vertex finder ZVTOP with branches ZVRES and ZVKIN (new in ILC environment)**
  - **flavour tagging based on neural net approach**
    - **includes full neural net package**
    - **default: Richard Hawkings' algorithm, cf. LC-PHSM-2000-021 ,**

      but flexible to allow change of inputs, network architecture etc
  - **quark charge determination, initially limited to jets containing a charged 'heavy flavour hadron'**

- **software uses LCIO for input and output and is interfaced to MarlinReco;**

  tests for running the code in the JAS environment planned in the US (Norman Graf)

- **code available from the ILC software portal http://www-flc.desy.de/ilcsoft/ilcsoftware/LCFIVertex**

- **a new Zeuthen CVS repository, 'tagnet', has been created for neural net descriptions –**

  users who train networks for a specific physics process should make them available there!

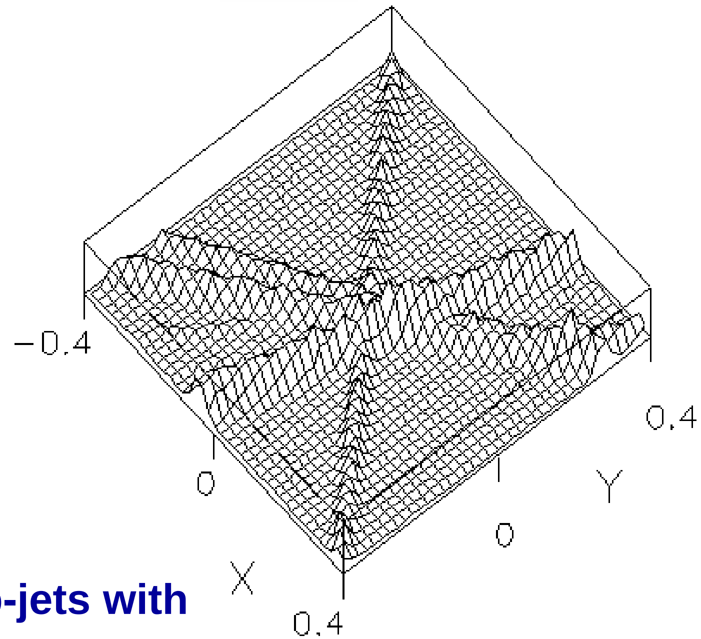
The ZVTOP vertex finder

D. Jackson,

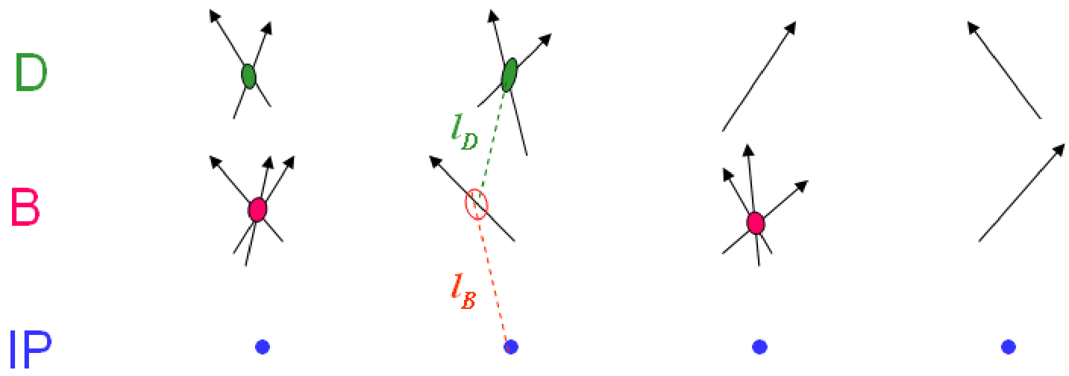
NIM A 388 (1997) 247

- **two branches: ZVRES and ZVKIN (also known as ghost track algorithm)**

- **The ZVRES algorithm: very general algorithm that can cope with arbitrary multi-prong decay topologies**
  - **'vertex function' calculated from Gaussian ´probability tubes´ representing tracks**
  - **iteratively search 3D-space for maxima of this function and minimise χ^2 of vertex fit**



- **ZVKIN: more specialised algorithm to extend coverage to b-jets with 1-pronged vertices and / or a short-lived B-hadron not resolved from the IP**

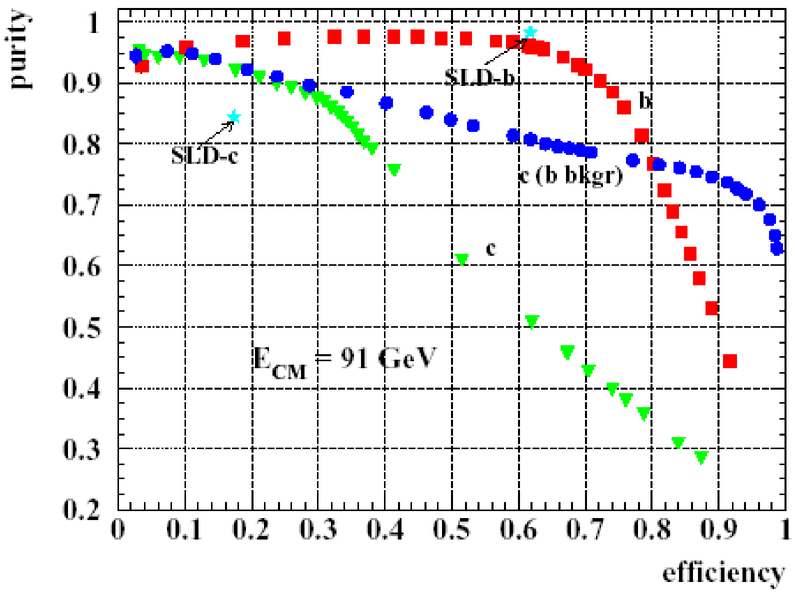


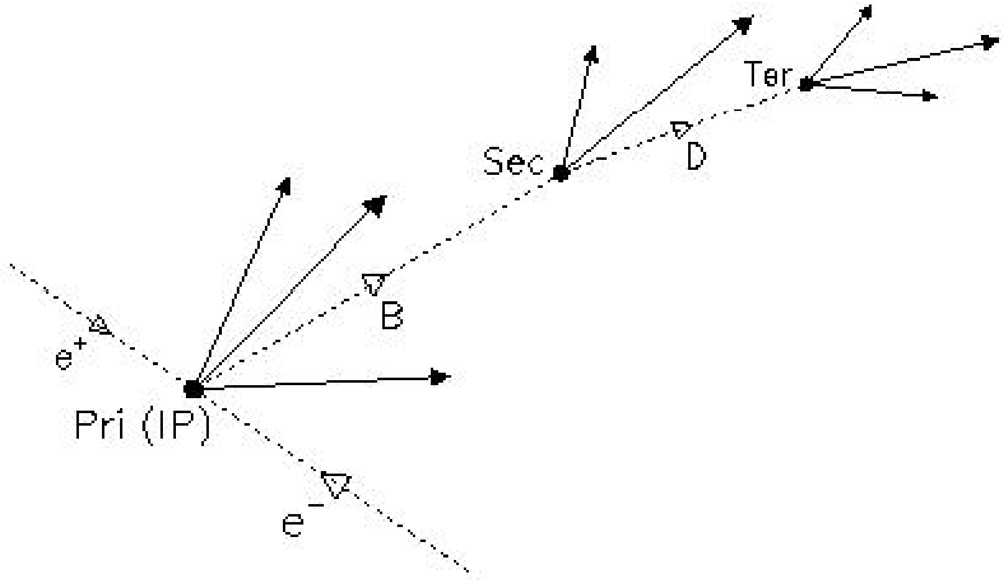
- **additional kinematic information (IP-, B-, D-decay vertex approximately lie on a straight line) used to find vertices**
- **should improve flavour tag efficiency and determination of vertex charge**

Flavour tag and quark charge sign selection

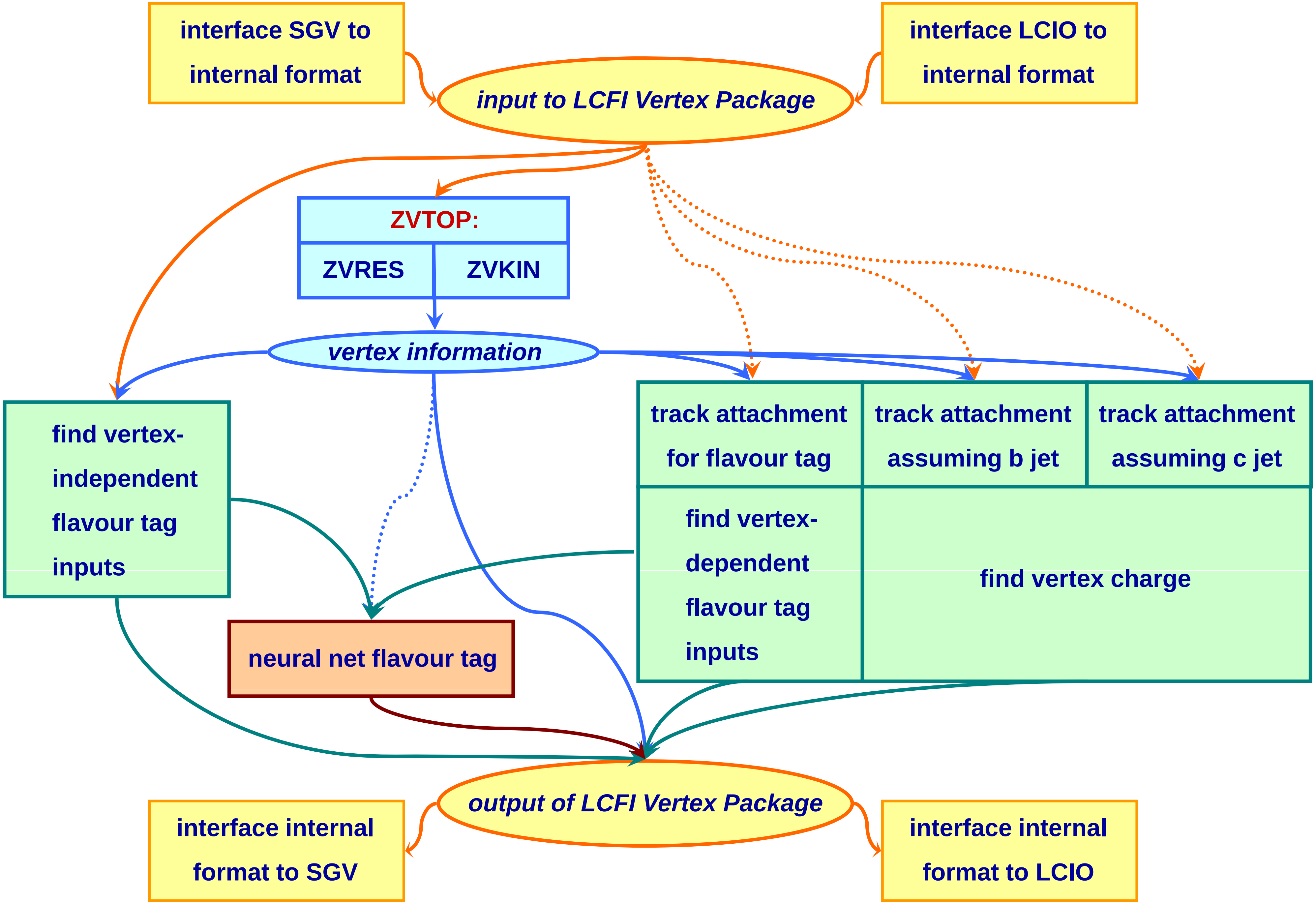
- aim of flavour tag: distinguish between b-jets, c- jets and light-quark / gluon jets
- heavy flavour jets contain secondary decays, generally observed as secondary vertices
- NN-approach to combine inputs; most sensitive: secondary vtx Pt-corrected mass & momentum

Klaus Desch/ Thorsten Kuhl





- for charged B-hadrons (40% of b-jets): quark sign can be determined from vertex charge: need to find all stable tracks from B-decay chain
- probability of mis-reconstructing vertex charge small for both charged and neutral cases
- neutral B-hadrons require ‘charge dipole’ procedure from SLD still to be developed for ILC



interface SGV to internal format
interface LCIO to internal format
input to LCFI Vertex Package
ZVTOP:
ZVRES
ZVKIN
vertex information
find vertex-independent flavour tag inputs
track attachment for flavour tag
track attachment assuming b jet
track attachment assuming c jet
find vertex-dependent flavour tag inputs
find vertex charge
neural net flavour tag
output of LCFI Vertex Package
interface internal format to SGV
interface internal format to LCIO

Approach to validation of the package

- **Tests using SGV event reconstruction**

  permitted direct comparisons with results from FORTRAN version using identical input events

  - **standalone test of ZVRES, input / output directly from / to SGV common blocks**

  - **separate tests of Marlin processors for ZVRES, ZVKIN, flavour tag input calculation**

    FORTRAN-LCIO interface used to write out lcio file from SGV, read in by Marlin processor

    and used to feed values into internal working classes of our package

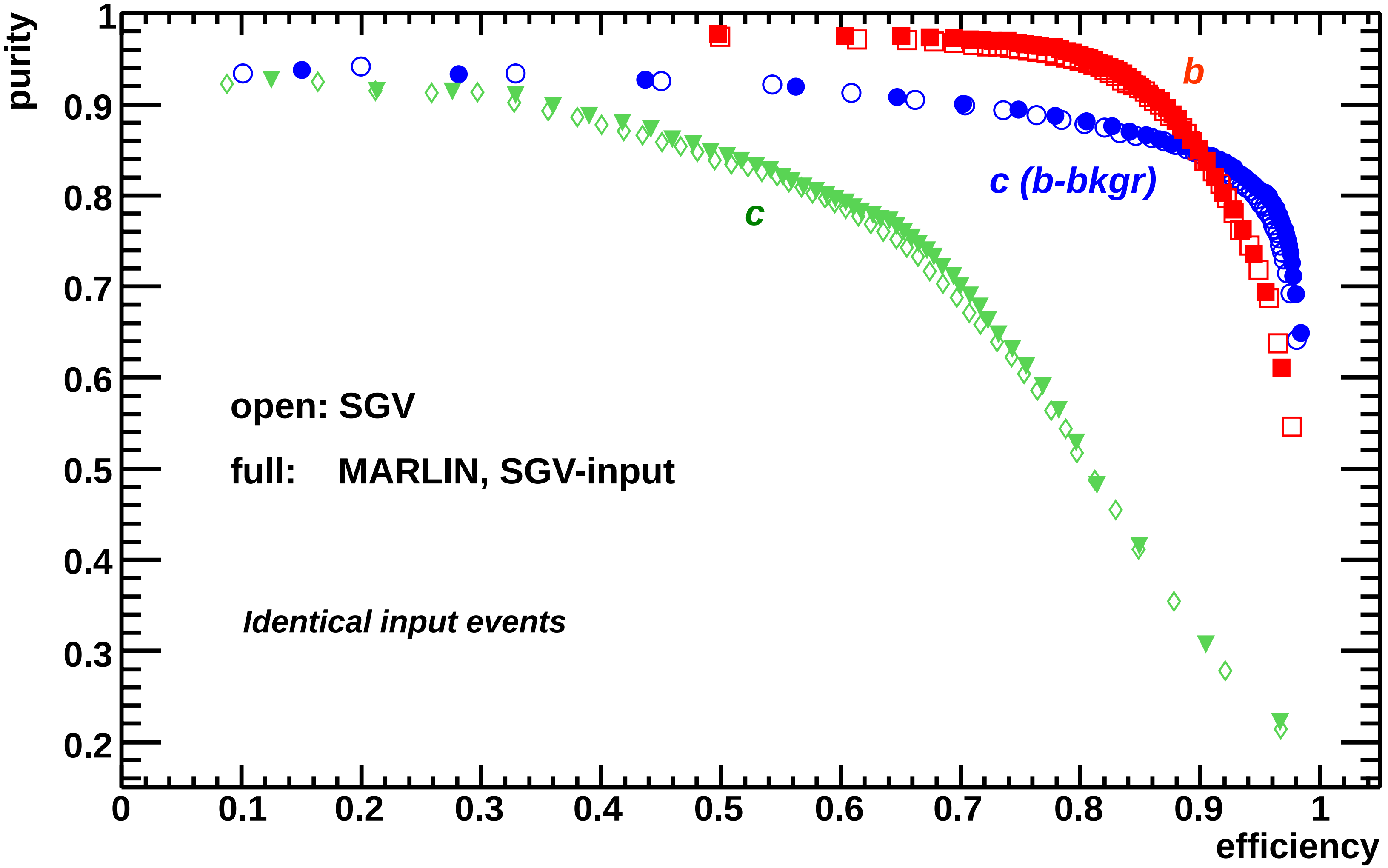
    results from those tests: Ben Jeffery's talk at ECFA workshop, Valencia, 2006

  - **full-chain test of ZVRES + flavour tag + vertex charge using same setup**

- **Tests using MarlinReco event reconstruction**

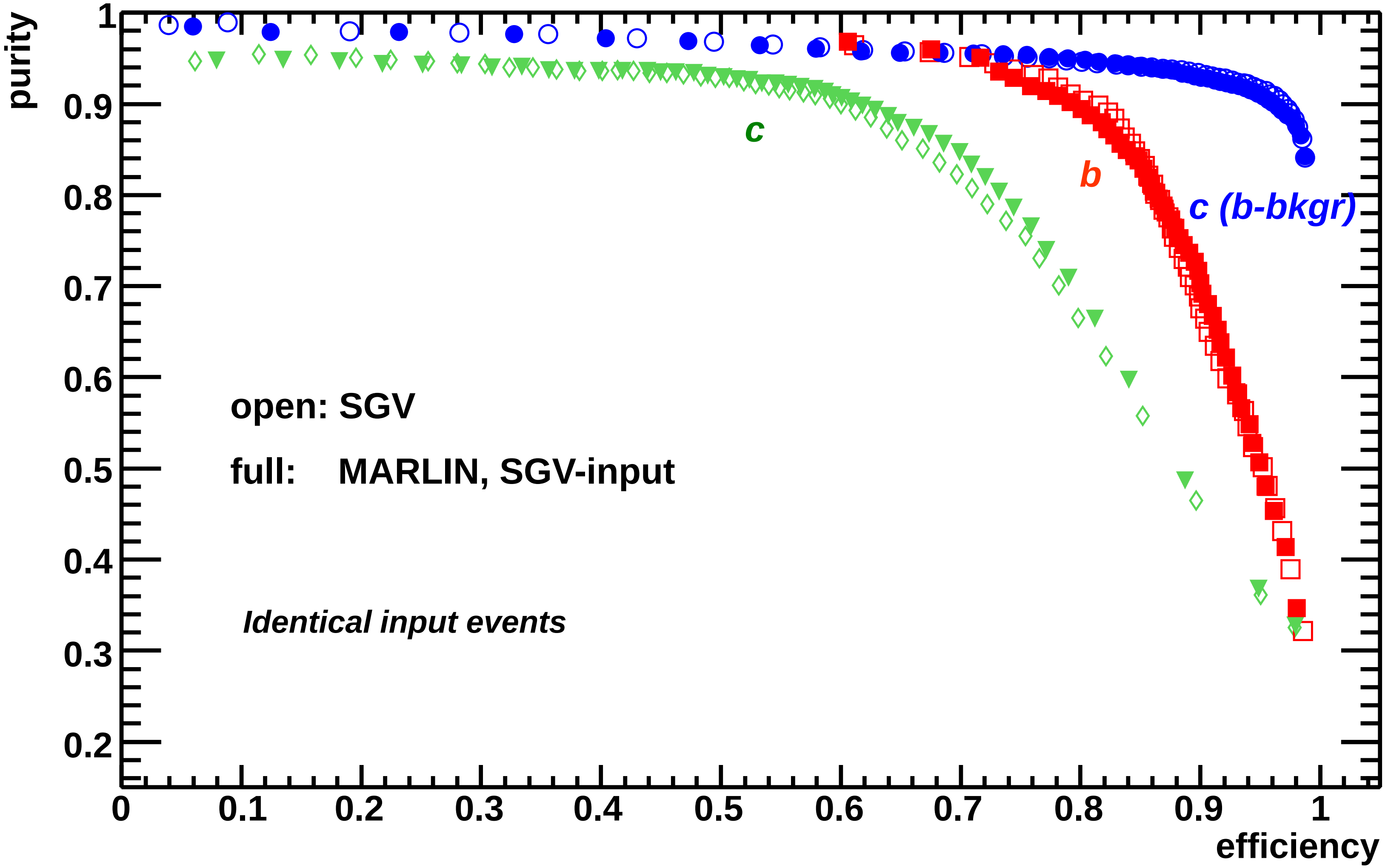
  - **full chain test repeated with PYTHIA events, passed through MOKKA**

Flavour tag performance at the Z-peak



- excellent agreement between the LCFIVertex Marlin code fed with SGV input and SGV

Flavour tag performance at sqrt(s) = 500 GeV

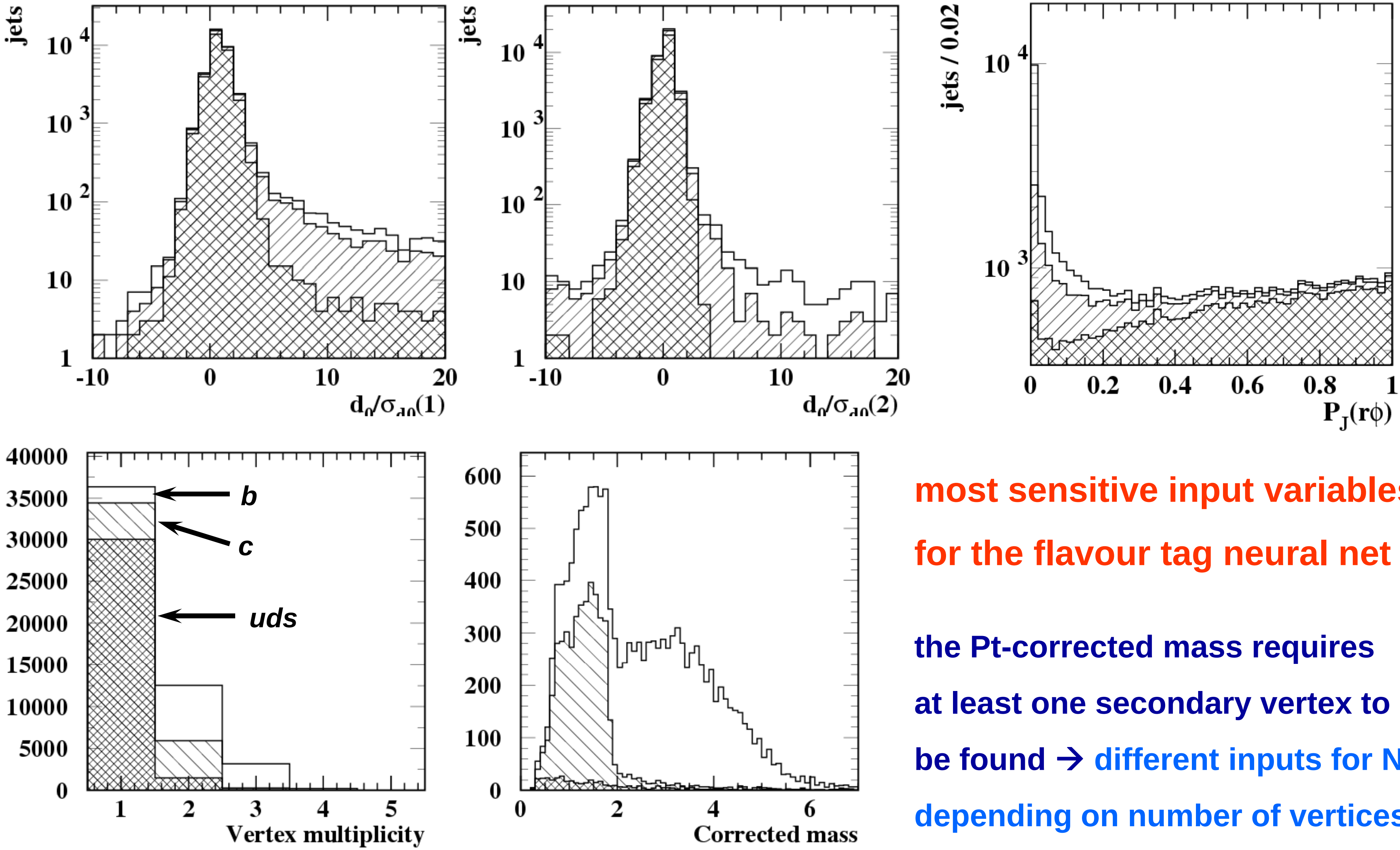


➢ excellent performance holds up over entire energy range relevant at the ILC

Performance obtained from MOKKA / MarlinReco

- **Tests using MarlinReco event reconstruction**
  - **full chain test repeated with PYTHIA events, passed through MOKKA v06-03 using the LDC01Sc detector model: simplified vertex detector geometry (cylinders)**
  - **thanks to Dennis Martsch for processing a test sample on the GRID**
  - **note: photon conversions switched off in GEANT, as these can easily be suppressed later**
  - **also, hadronic interactions in the beam pipe and in the vertex detector layers suppressed by a radius cut at the track selection level using MC information (optional, small effect, geometry currently hard-coded)**
  - **use simple digitizer and track cheater from Alexei Raspereza's tracking package thanks to Alexei for adding output of track covariance matrices in LCIO to his code**
  - **for jet finding using Durham algorithm from the Satoru jet finder package, with y-cut 0.04**
  - **default track selection of the LCFIVertex package was derived for previous BRAHMS study**

Some input distributions from LC-PHSM-2000-021

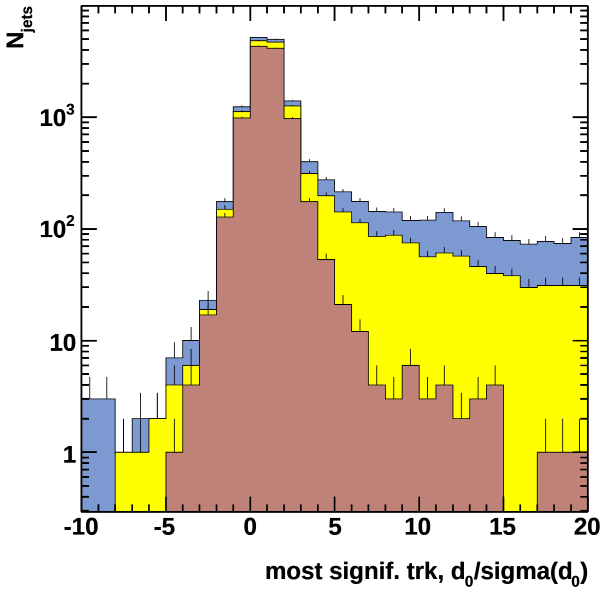


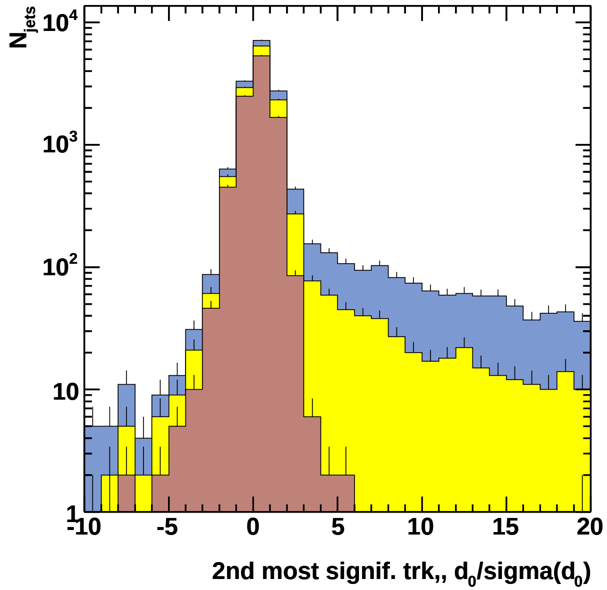
most sensitive input variables

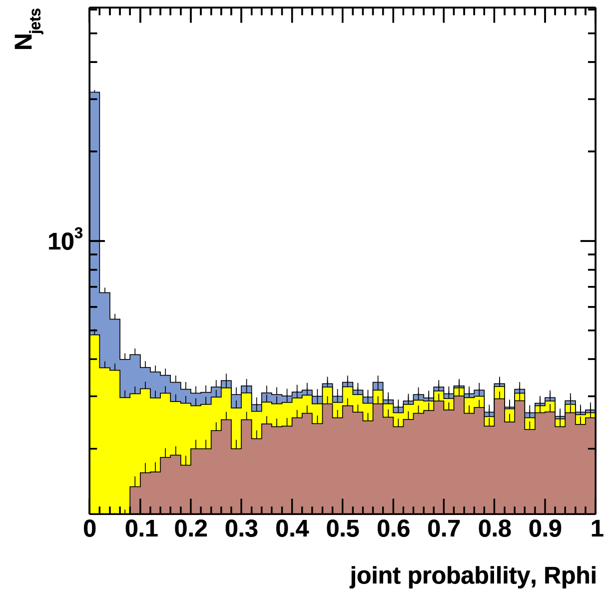
for the flavour tag neural net

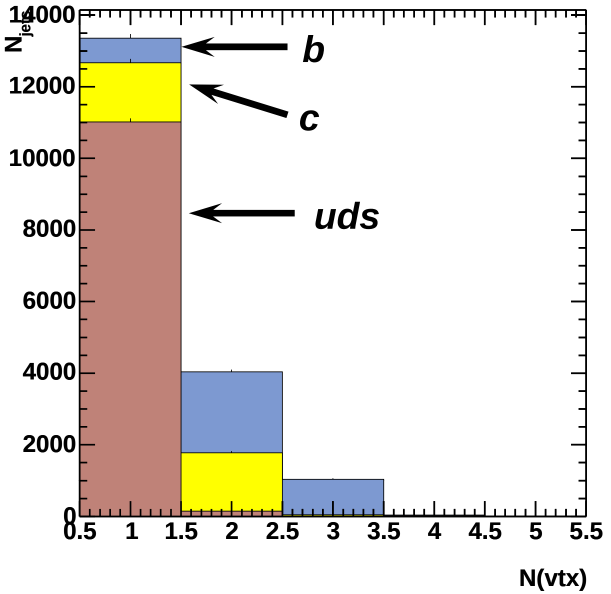
the Pt-corrected mass requires
at least one secondary vertex to
be found → different inputs for NN
depending on number of vertices

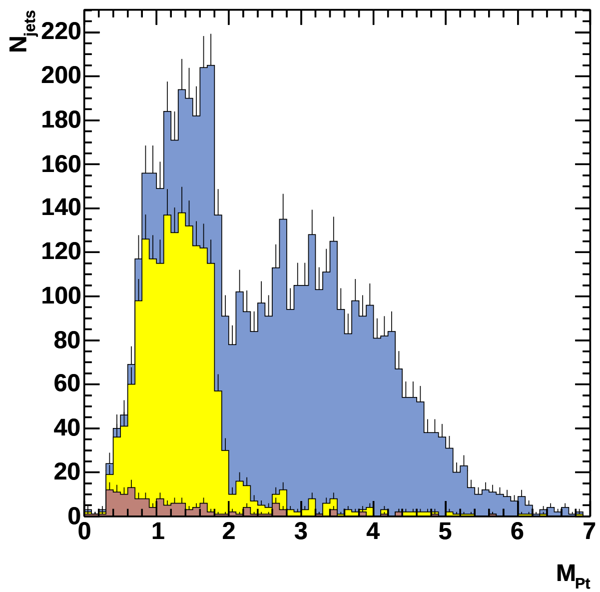
Input distributions from track cheater without TPC







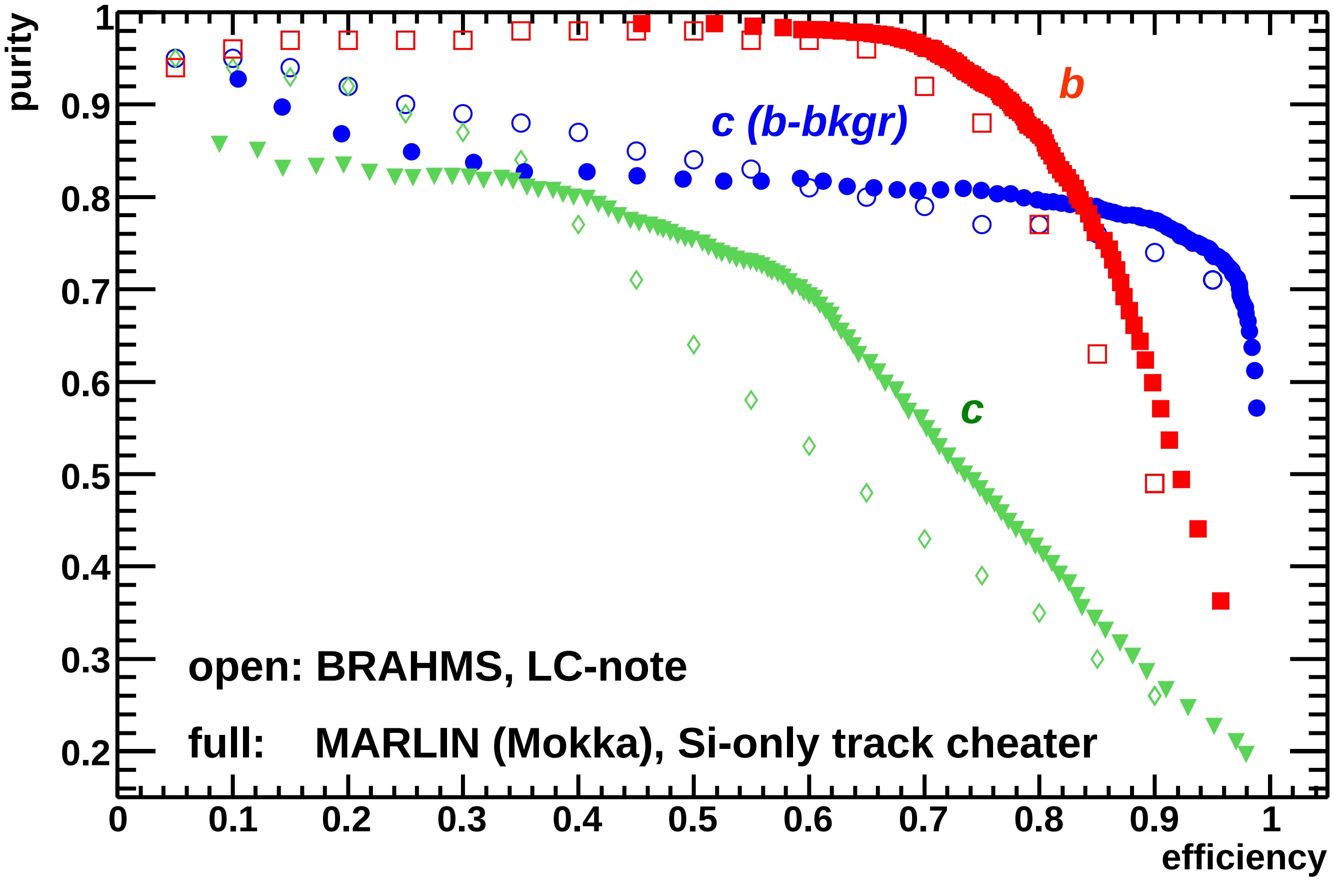




good agreement of MOKKA inputs with BRAHMS ones

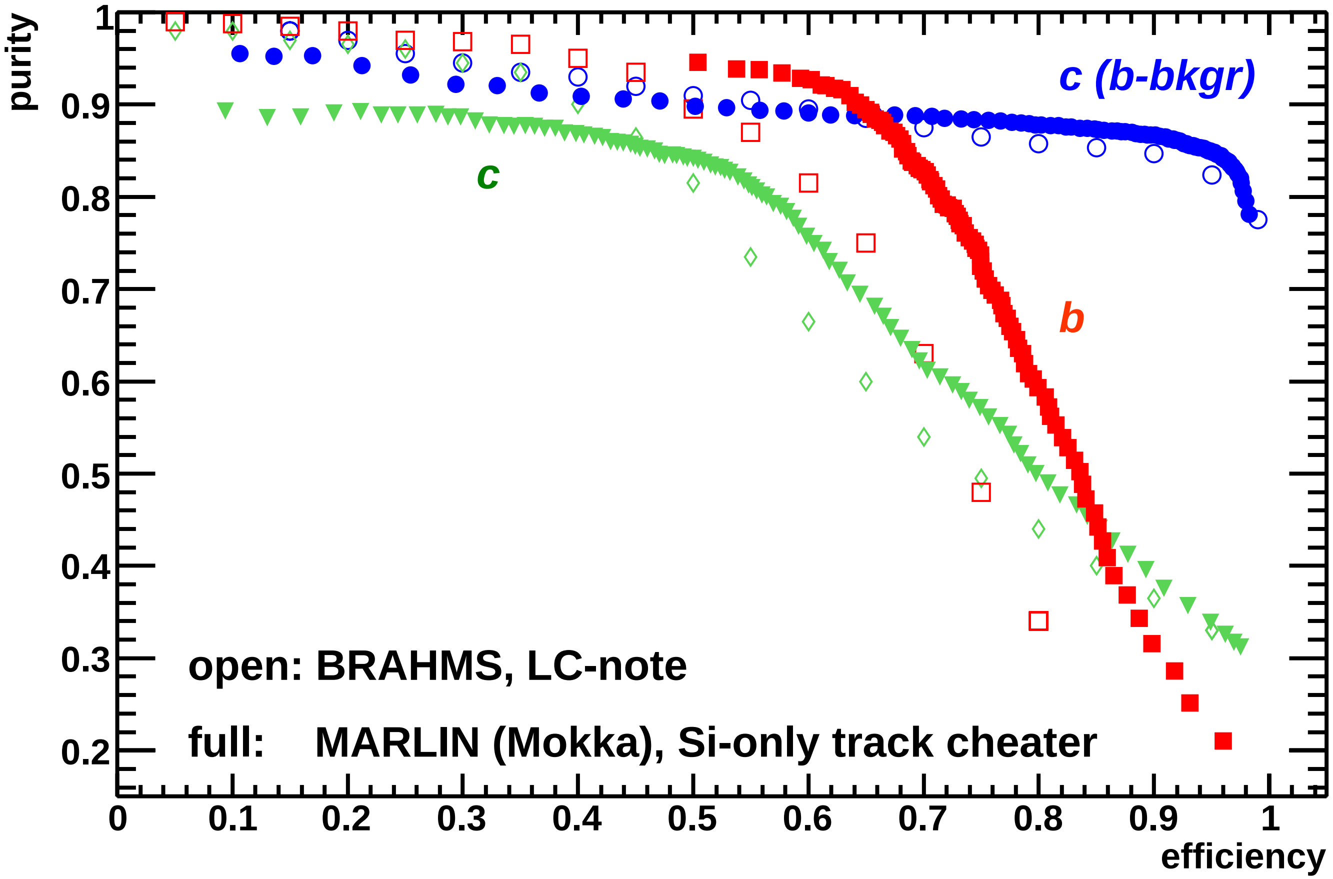
slightly better discrimination since MOKKA assumes point resolution 2 μm, BRAHMS assumed 3.5 μm

Resulting purity vs efficiency at the Z-peak



➢ at high efficiency MARLIN(MOKKA) with "Silicon-only" track cheater (VXD, SIT, FTD) gives better performance compared to LC-note result using tracking with pattern recognition

Purity vs efficiency at sqrt(s) = 500 GeV, Si-only



- excellent performance also found at higher energy
- note that at this energy a track momentum cut has to be applied for the jet finder to run (feature of Satoru jet finder)

Results for different tracking options

- performance of the code depends on quality of track information provided to it:
  initially a bug in the track cheater caused severe degradation in performance

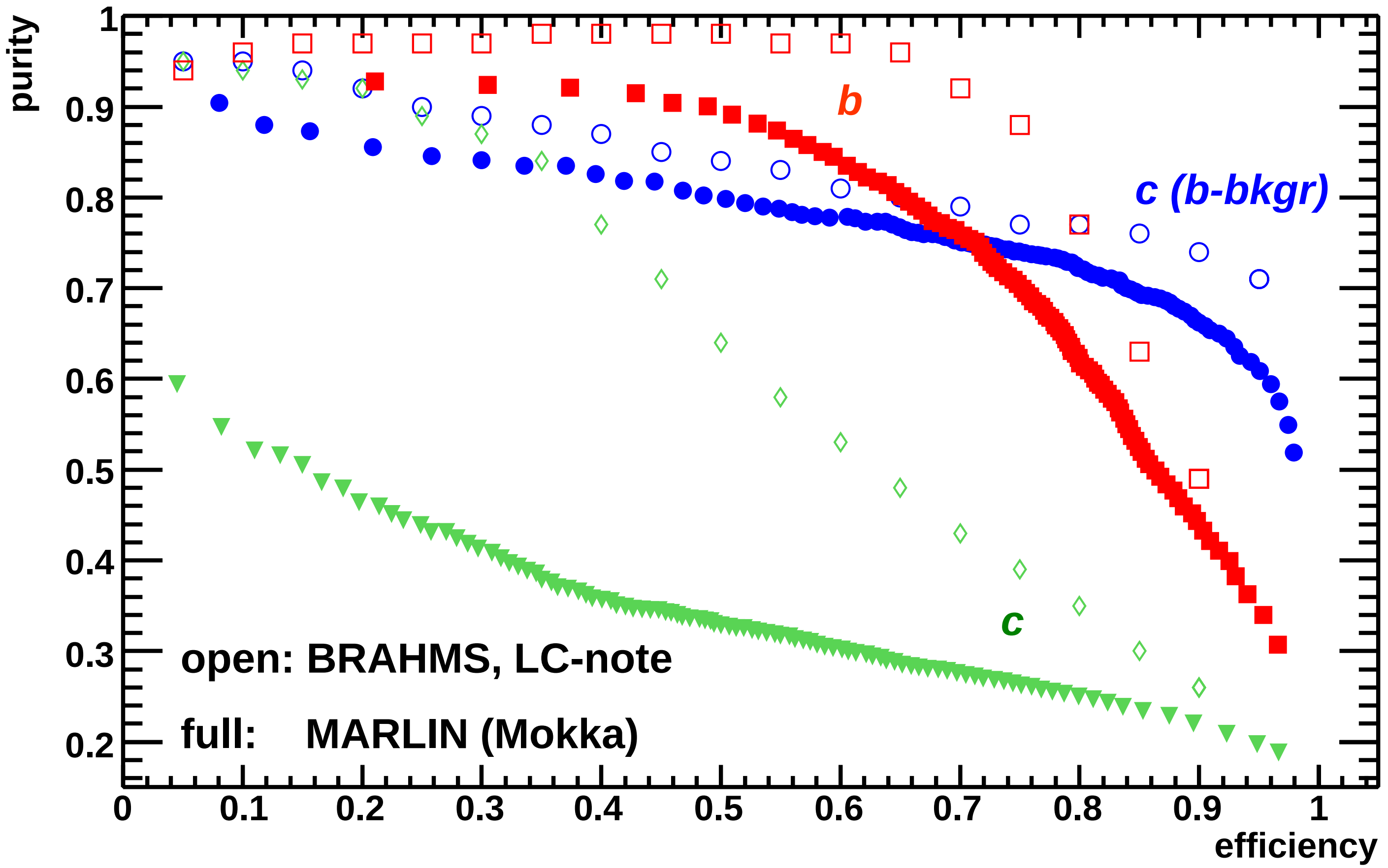
- for validation used track cheater by Alexei Raspereza;
  this provides two options: one using all trackers (VXD, TPC, SIT, FTD) and “Silicon only”

- performance study with the various LDC tracking options ongoing in close collaboration with Alexei, who plans to release a new version of the tracking code in mid June
  a lot of progress since Orsay simulation workshop and during LCWS,
  but work is still ongoing, following pages just show the current status

- for the April version of LDC tracking (shown at Orsay) for tagging recommend using the Si-only track cheater for the time being, as performance of the other options is degraded

- for the version that is current head of the CVS of LDC tracking find:
  good performance at the Z-peak for all options, at 500 GeV, only track cheater with Si-only

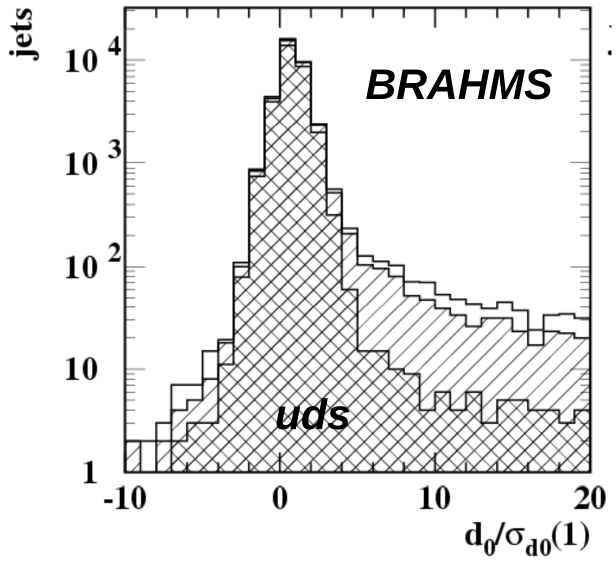
Performance with initial bug in track cheater

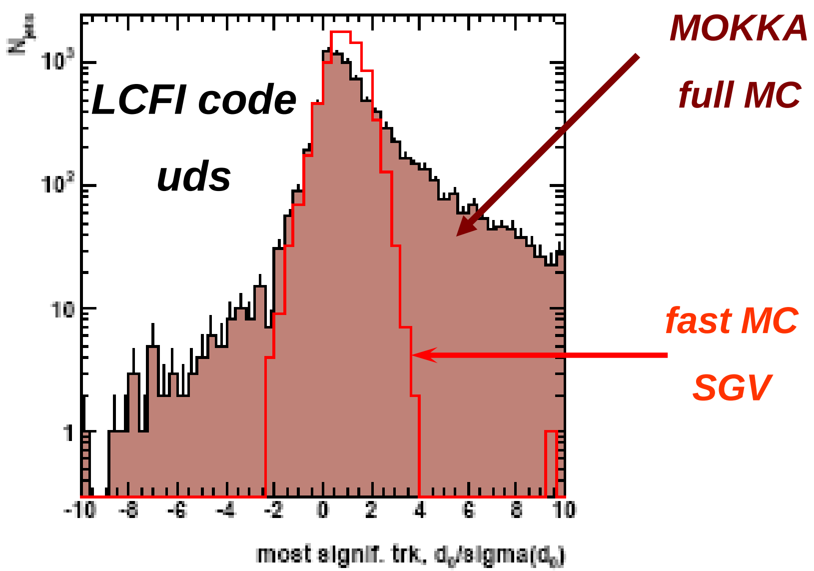


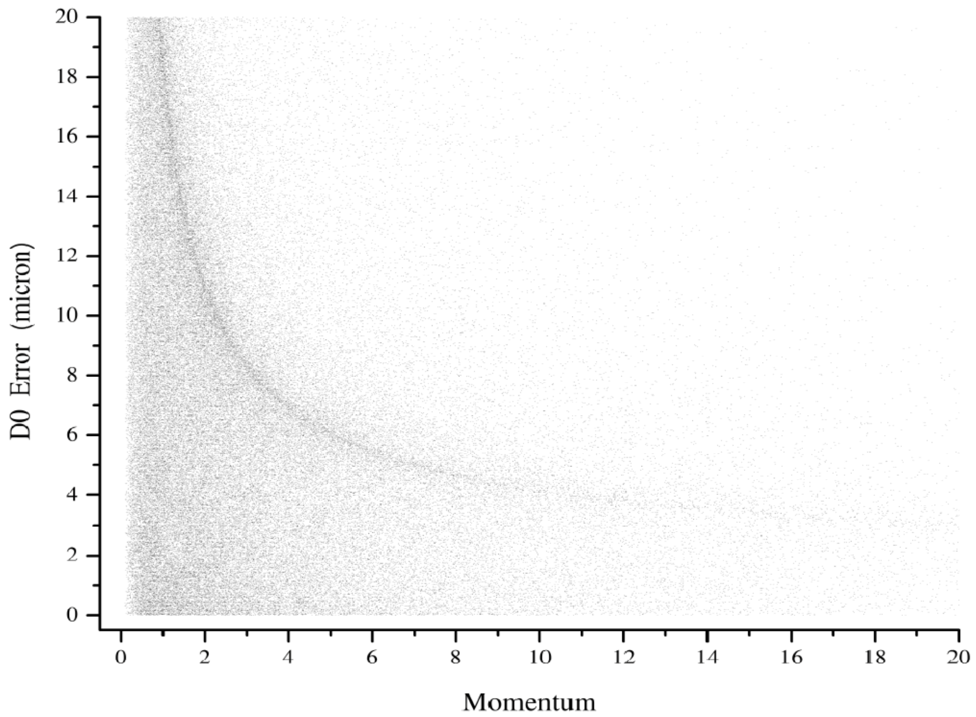
- initially obtained poor performance since there was a bug in the newly added track cheater covariance matrices → be careful to provide correct input to the package

Effect of track cheater bug on tagging inputs I

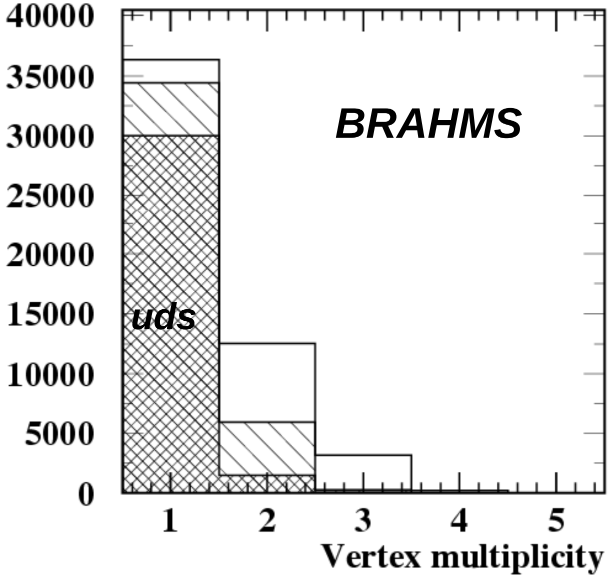
- looked at impact parameter significance of most significant track for uds jets
- MOKKA result (lower left) initially had exceedingly large tails compared to BRAHMS result (upper left)
- tracked down to a large number of tracks having unrealistically small impact parameter errors in R-phi

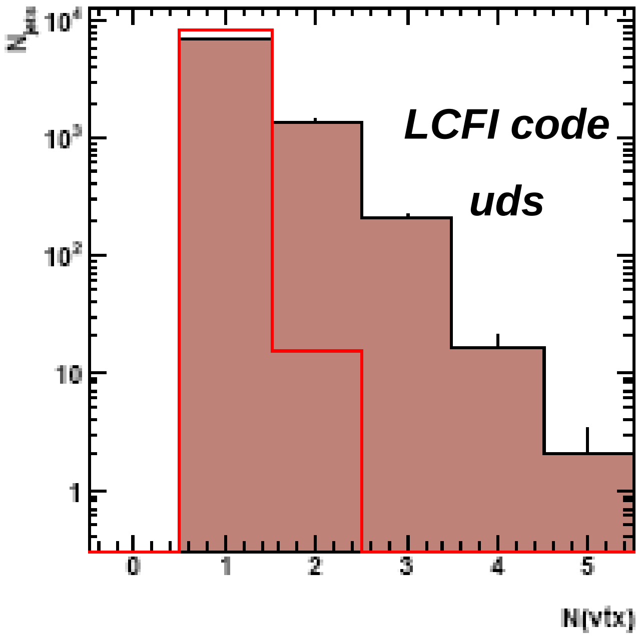




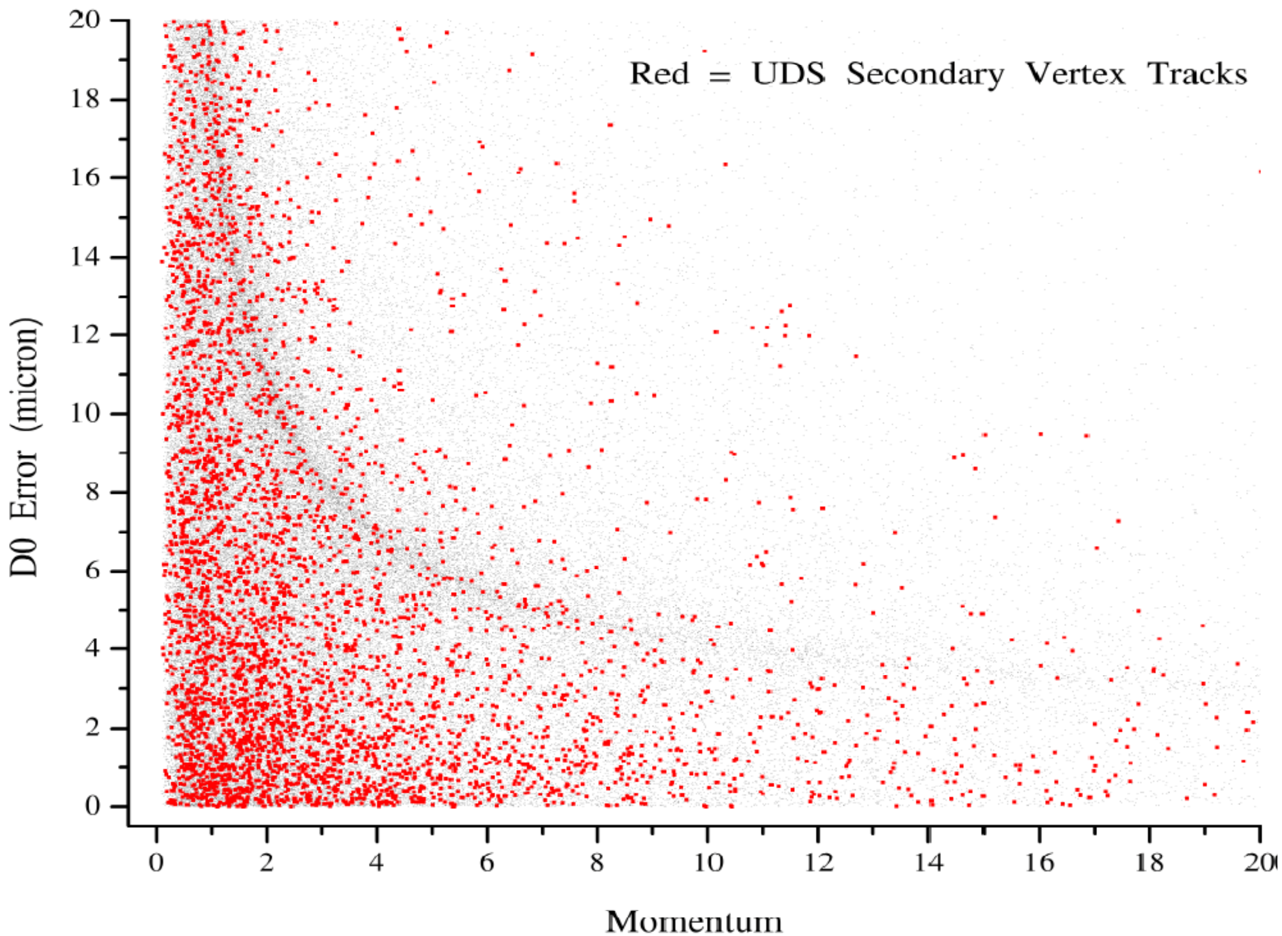


Effect of track cheater bug on tagging inputs II

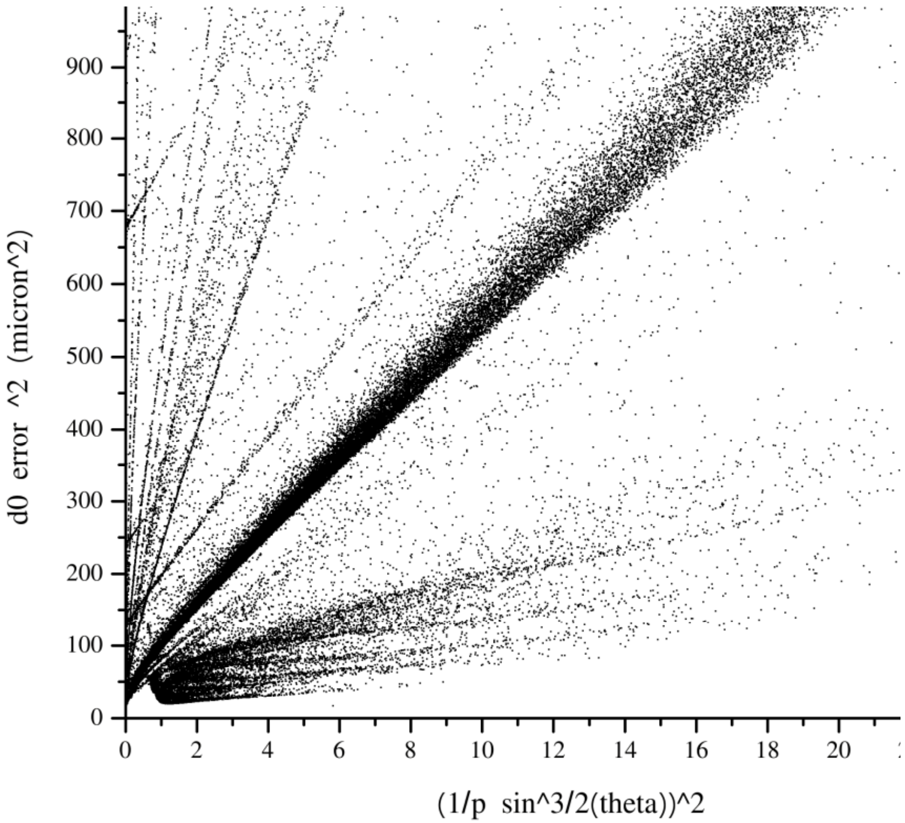


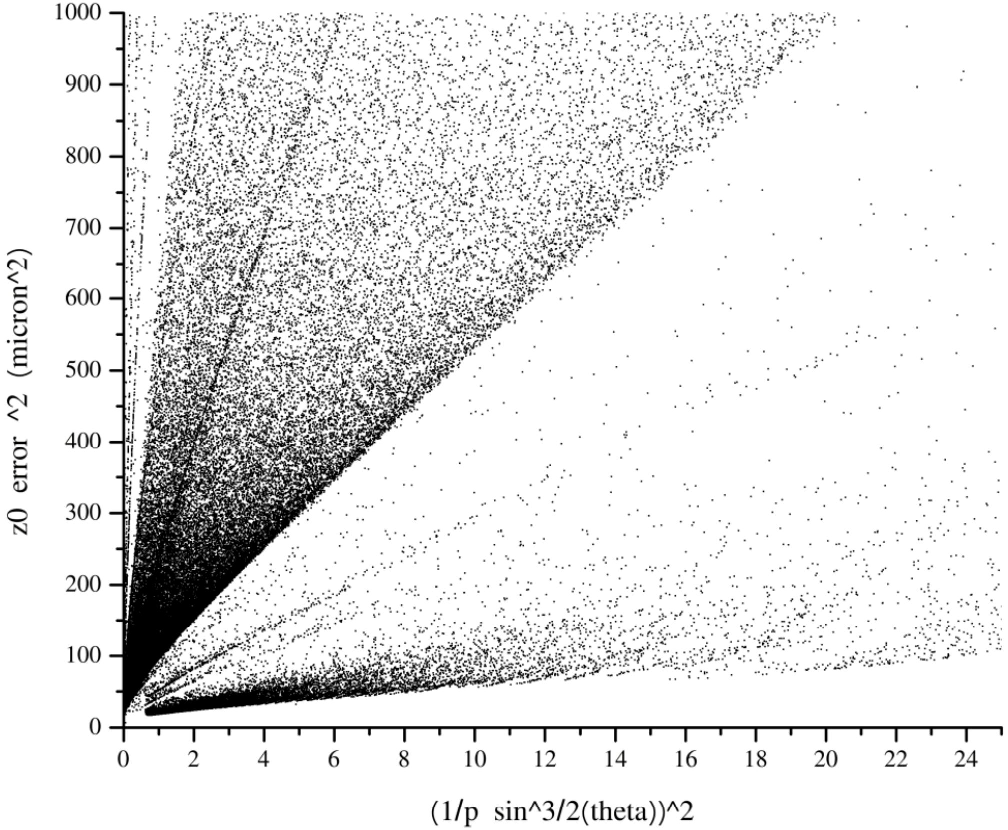


➢ this also showed up in an unusually large number of secondary vertices found in uds jets:



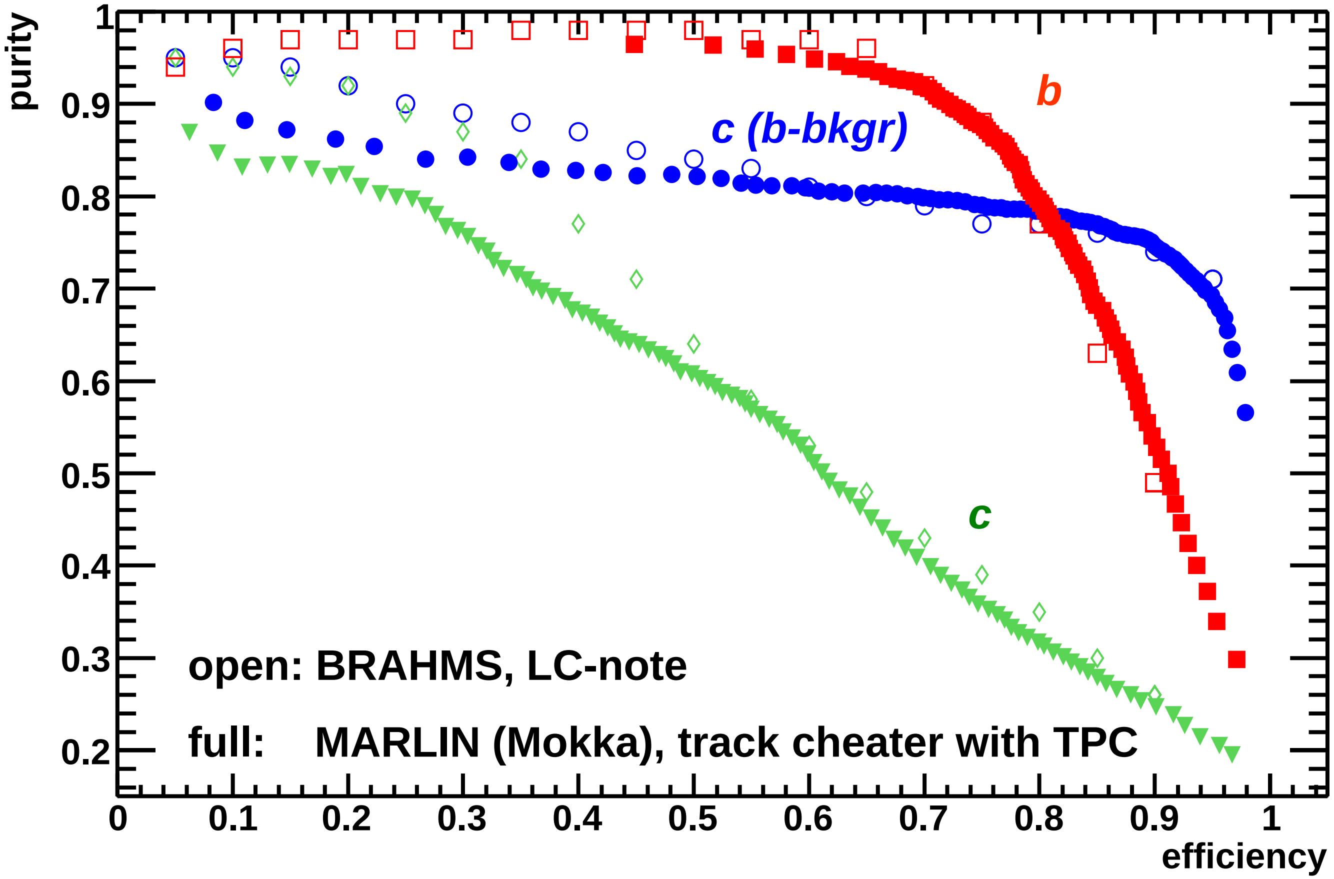
April version of track cheater, TPC included





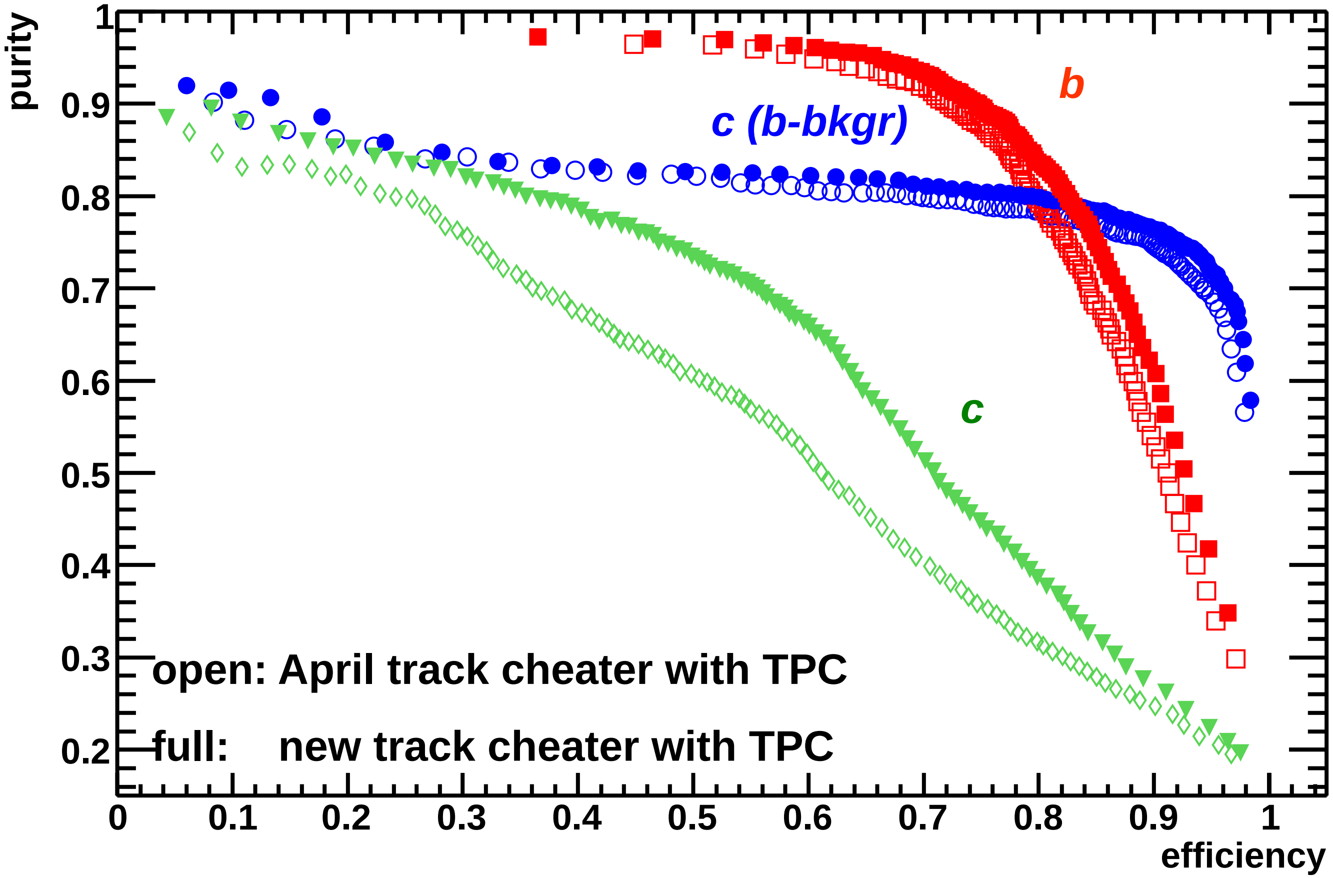
- **at the time of the Orsay workshop, these problems were much improved, but for the track cheater with TPC, there were still some impact parameter problems**

Z-peak, April version of track cheater with TPC included



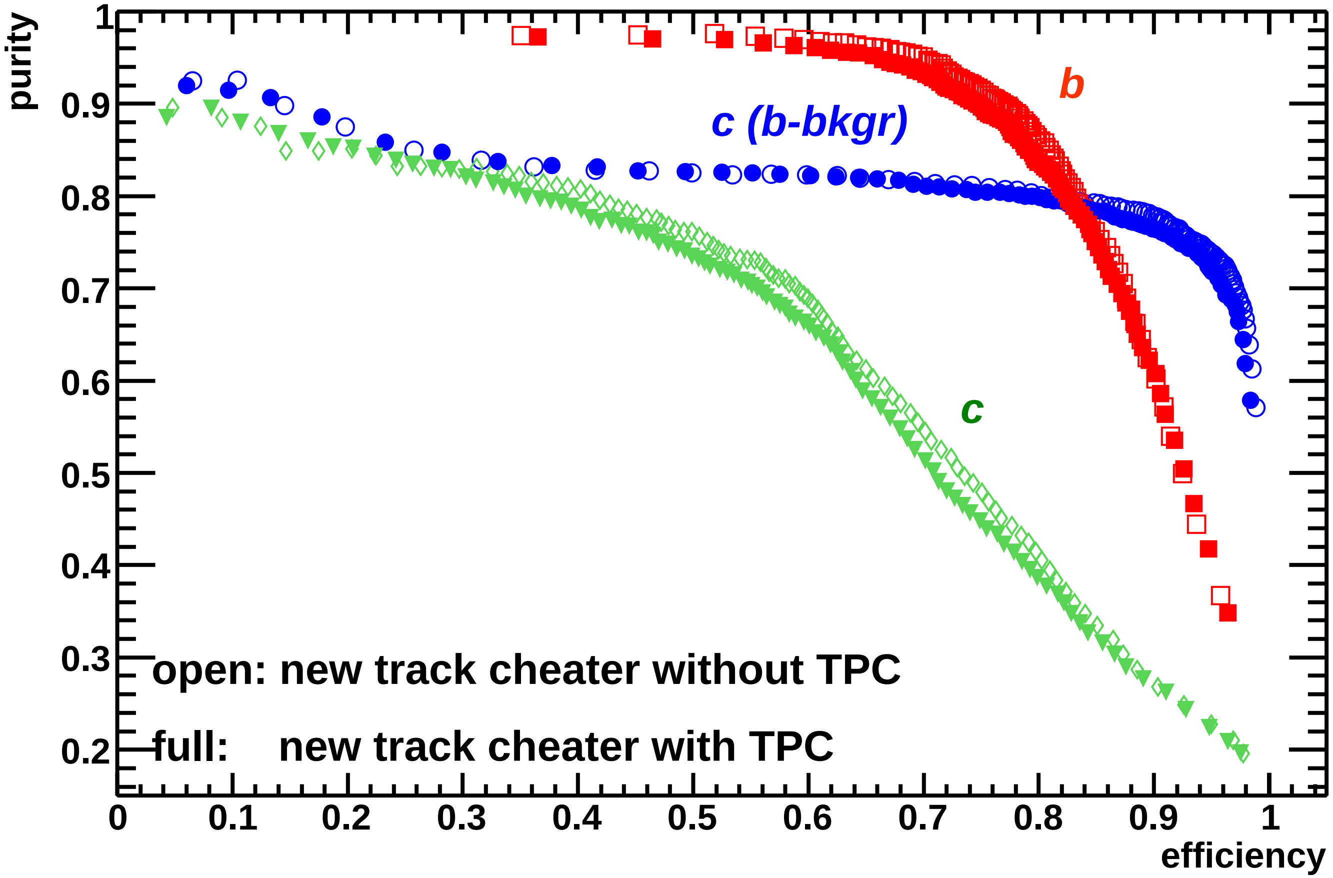
➢ **performance of April version of track cheater with TPC (used for Orsay simulation workshop) rather similar to BRAHMS result – could indicate these problems were present in BRAHMS**

Z-peak, track cheater with TPC, version comparison



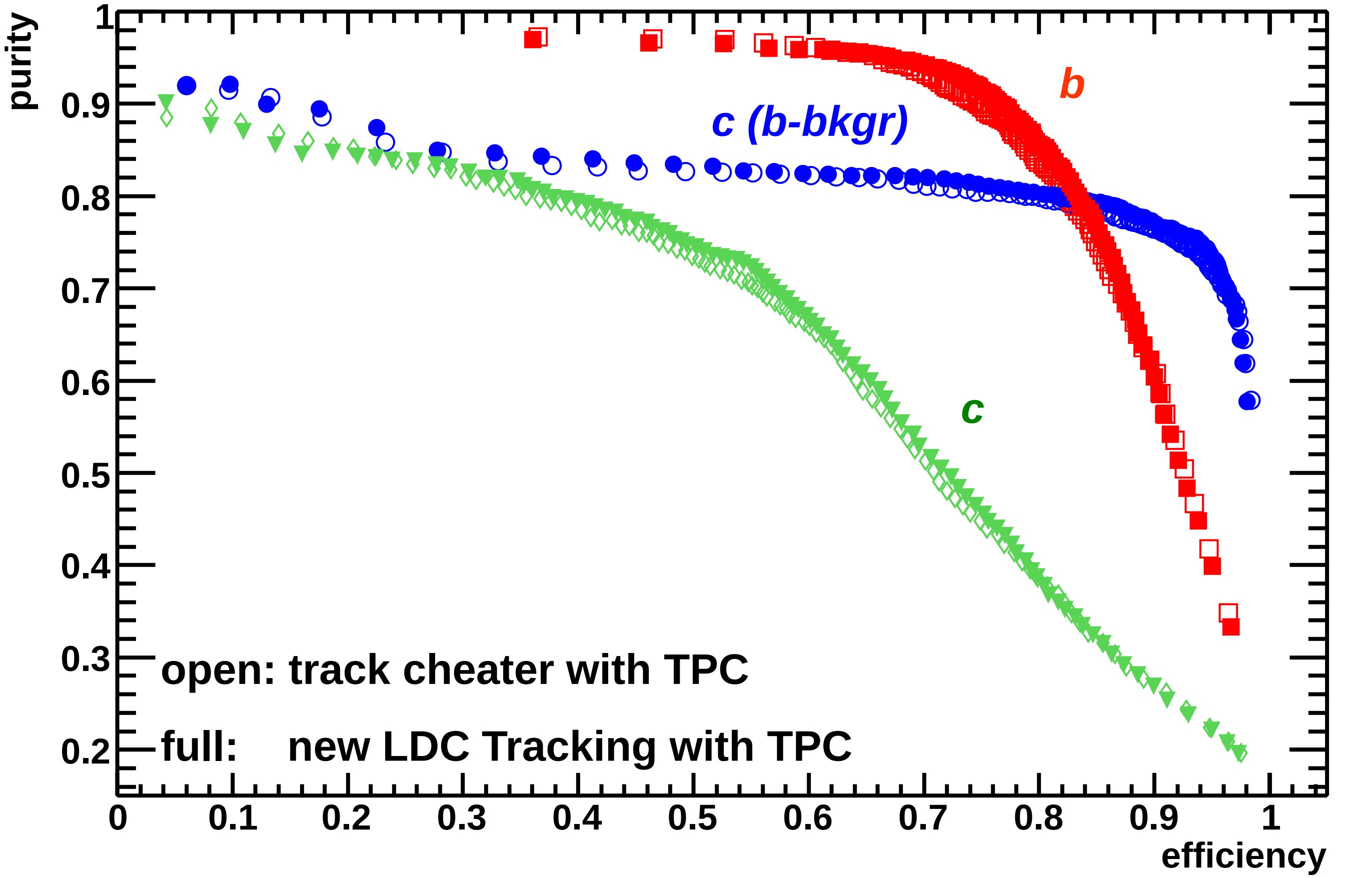
- ➢ in current code presented by Alexei at this conference, problem is corrected; performance of track cheater with TPC at the Z-peak considerably improved

Z-peak, new track cheater with and without TPC



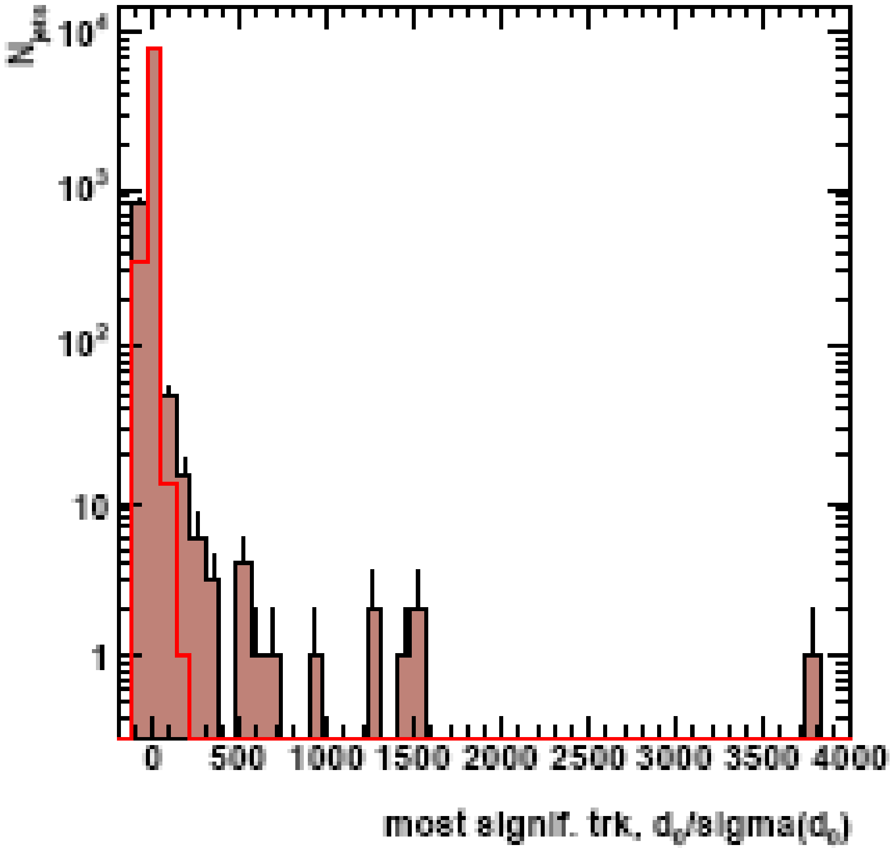
- ➢ **for this improved version, track cheater Z-peak performance with and without TPC very similar; option without TPC slightly better, under study**

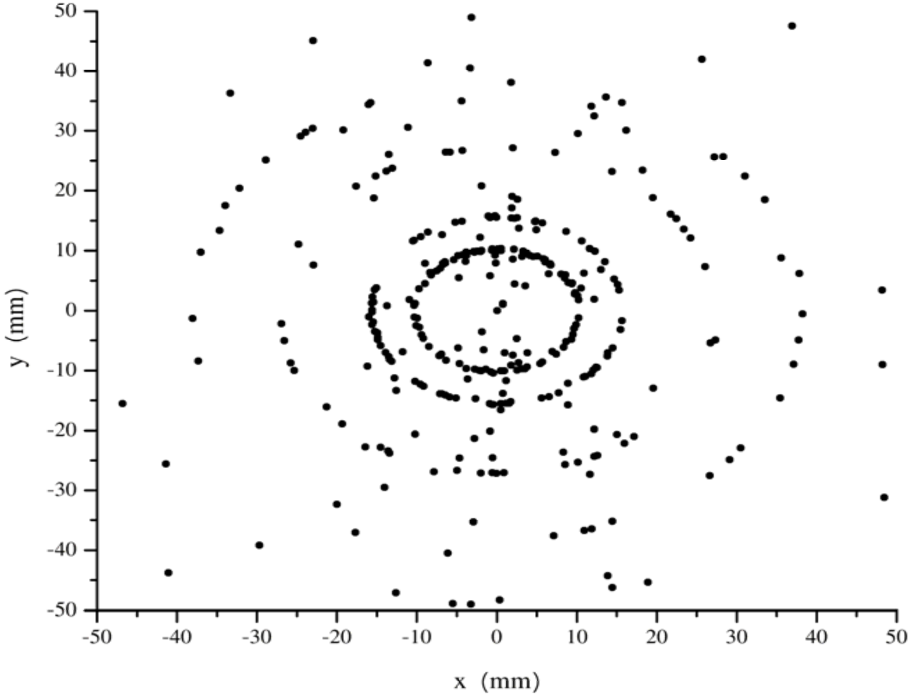
Z-peak, track cheater compared to full LDC tracking



- for improved version, track cheater and full LDC tracking perform very similarly at Z-peak; unexpectedly, full tracking is slightly better, being investigated

Hadronic interactions



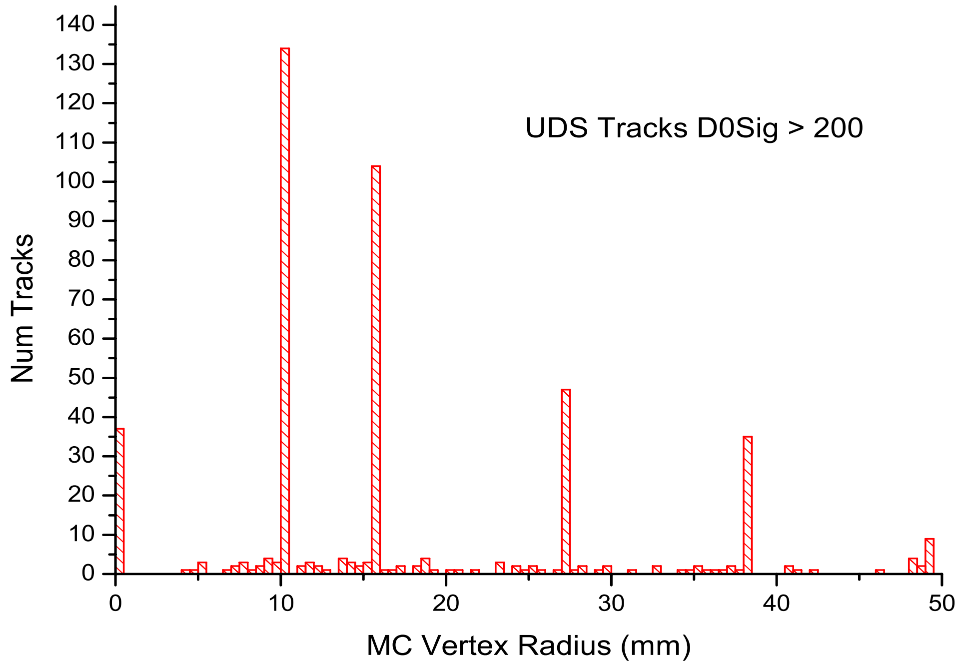


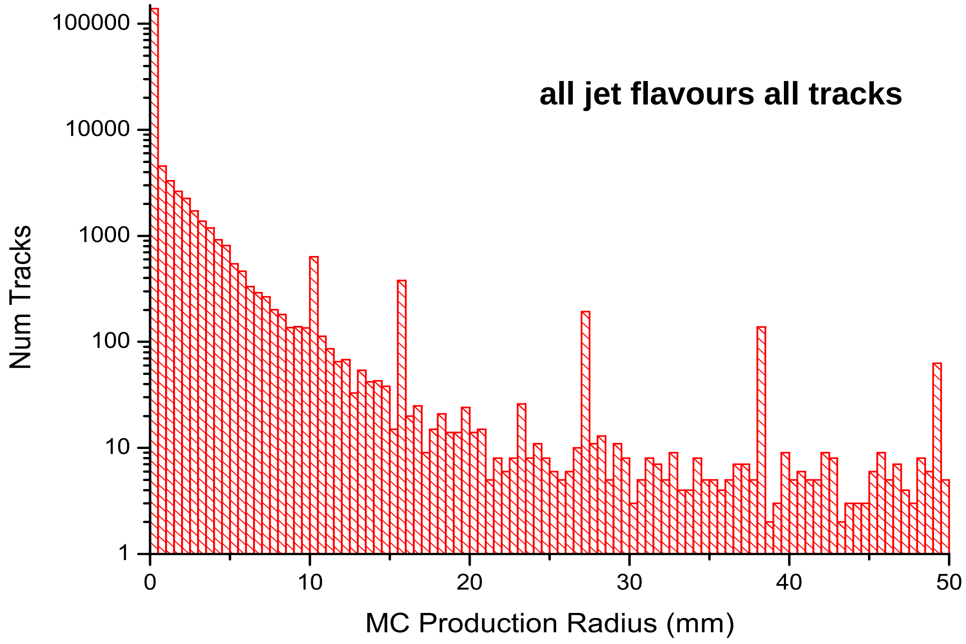
➢ **even after correcting track cheater bug, some tracks in uds jets have extremely high impact parameter significances (left) – these mostly have their MC origin in the material (see right)**

(Si-only track cheater, Z-peak sample)

Hadronic interactions

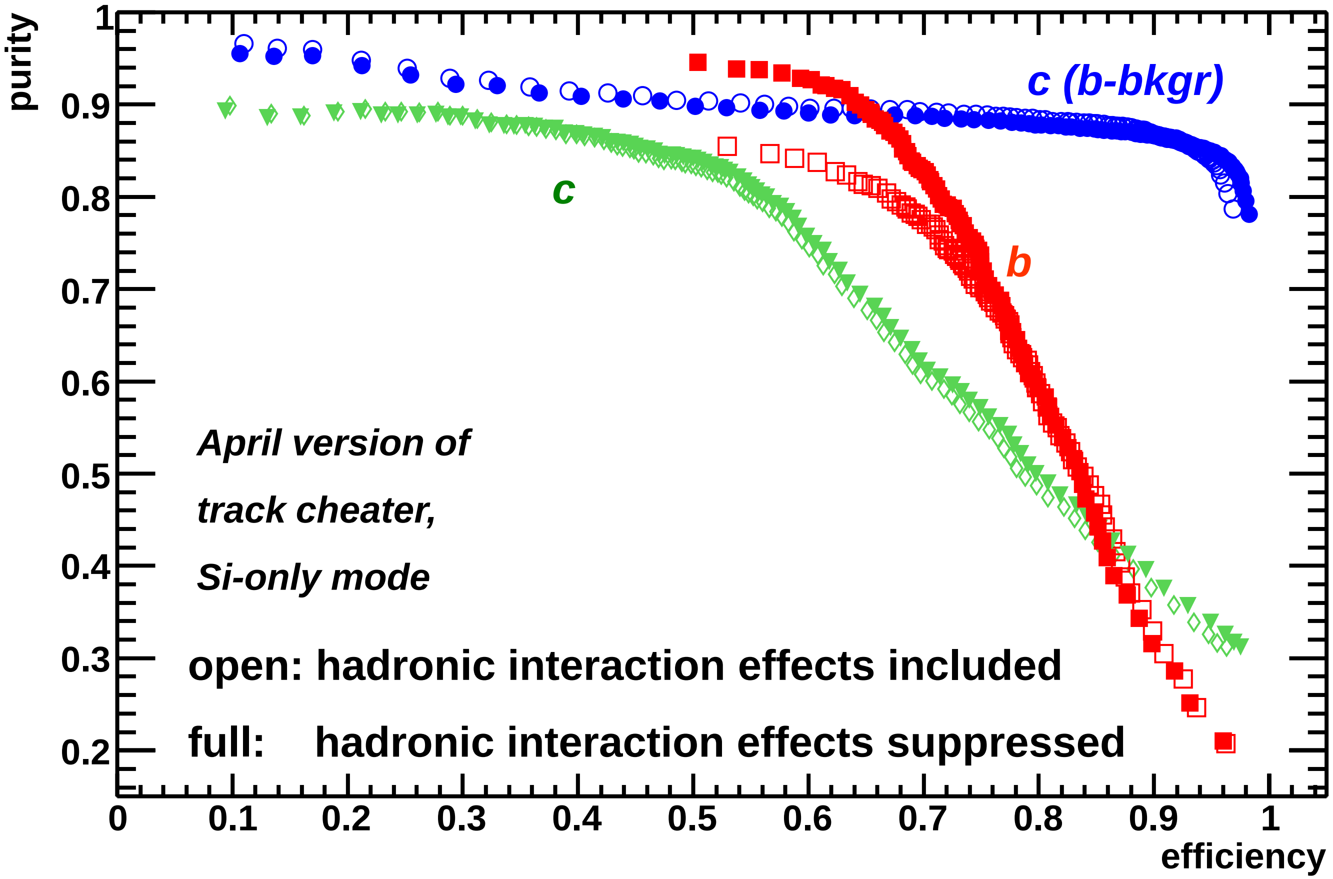
Ben Jeffery





- ➢ **left plot shows MC radius distribution for the same tracks as in 2D plot on previous page**
- ➢ **right plot: MC radius distribution for all tracks in the sample – note the logarithmic scale!**

Performance at 500 GeV w/wo hadronic interactions



- **effect on performance: negligible at the Z-peak, substantial at higher energy**
- **for the time being suppress tracks using MC information, in future proper treatment needed (known to be possible from earlier experiments, e.g. NA32)**

Summary

- The LCFIVertex package is available at http://www-flc.desy.de/ilcsoft/LCFIVertex ;
  provides the vertex finder ZVTOP, flavour tagging and vertex charge calculation for b- and c-jets
- core functionality has been extensively tested using the fast MC SGV, yielding excellent agreement and slightly better performance than FORTRAN code
- some aspects of the package will need further exploration, e.g. best use of the ZVKIN vertexing algorithm in the ILC environment
- first results using MOKKA input show good agreement with previous BRAHMS results
- performance of tracking code has significant effect on performance of vertexing/ flavour tagging
  results from track cheater with Si + TPC need further study,
  for current tracking version, recommend using Si-only track cheater when running our code
- effects of hadronic interactions significant for b-tag at high energy;
  currently suppressed using MC information, proper treatment needed and possible in future

Outlook: the next steps

- will add diagnostic features (plots of flavour tag inputs, ‘raw’ numbers for tags etc) and further documentation in the near term future

- performance study with full LDC tracking beginning

- improvements to the code, to provide more realistic description of some effects, e.g. treatment of hadronic interactions, IP-treatment, refitting for joint probability parameters

- parameter tuning and training neural nets with full MC and full LDC tracking; systematic performance study varying input parameters, followed by retraining the networks

- exploration of ZVKIN branch of ZVTOP for flavour tag; study/tune at Z-peak, extension to 500 GeV, combine with ZVRES results for flavour tag

Please provide feedback and publish your flavour tag networks in the Zeuthen tagnet repository!

Additional Material

Flavour tag

- **Vertex package provides flavour tag procedure developed by R. Hawkings et al (LC-PHSM-2000-021) and recently used by K. Desch / Th. Kuhl as default**

- **NN-input variables used:**

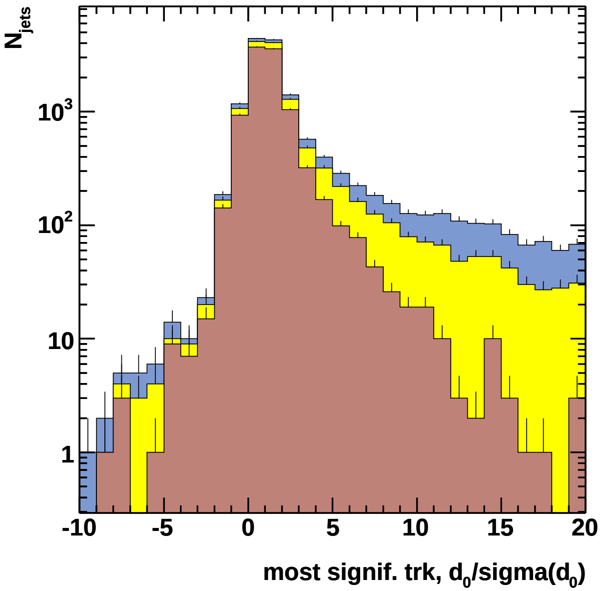
  - **if secondary vertex found: M_{Pt}, momentum of secondary vertex, and its decay length and decay length significance**

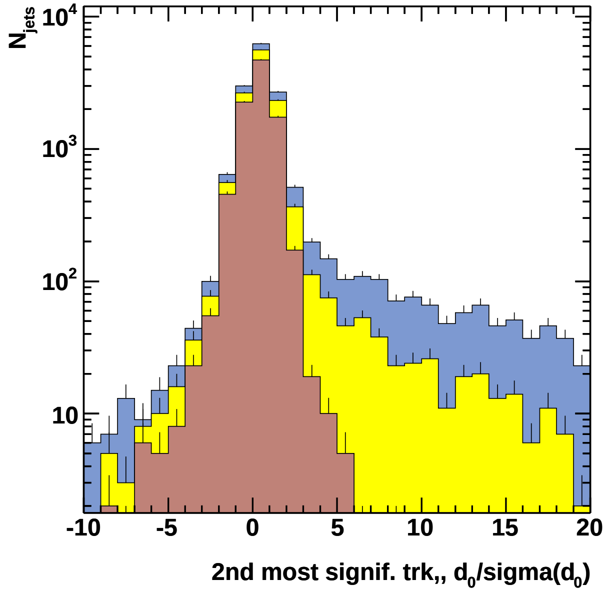
  - **if only primary vertex found: momentum and impact parameter significance in R-ϕ and z for the two most-significant tracks in the jet**

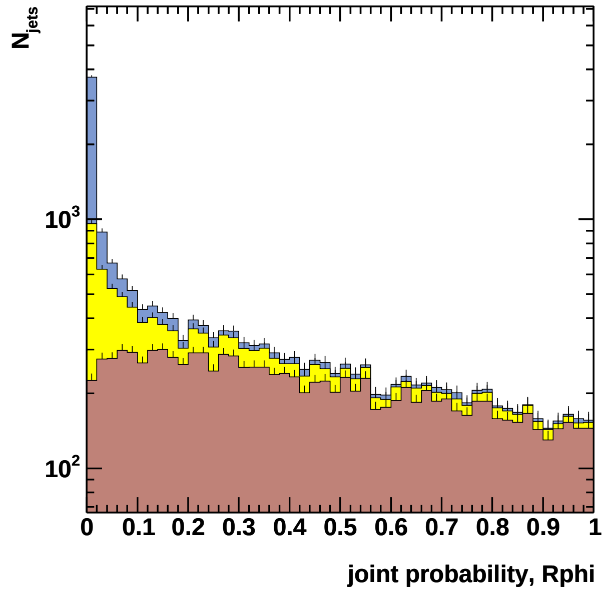
  - **in both cases: joint probability in R-ϕ and z (estimator of probability for all tracks to originate from primary vertex)**

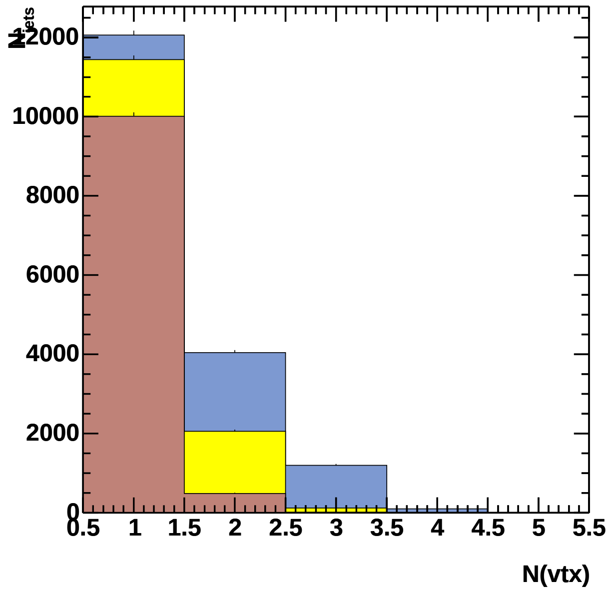
- **flexible enough to permit user further tuning of the input variables for the neural net, and of the NN-architecture (number and type of nodes) and training algorithm**

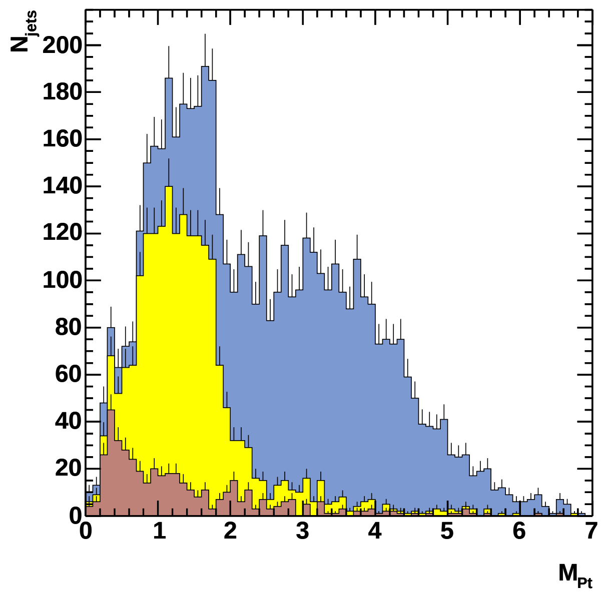
Input distributions from April track cheater with TPC



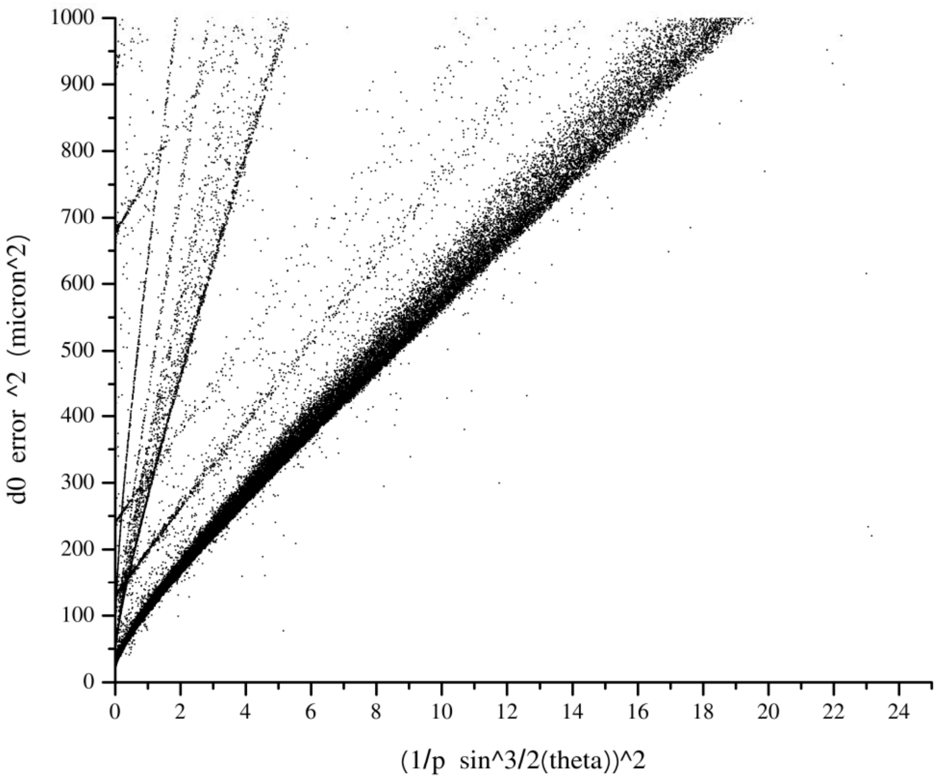


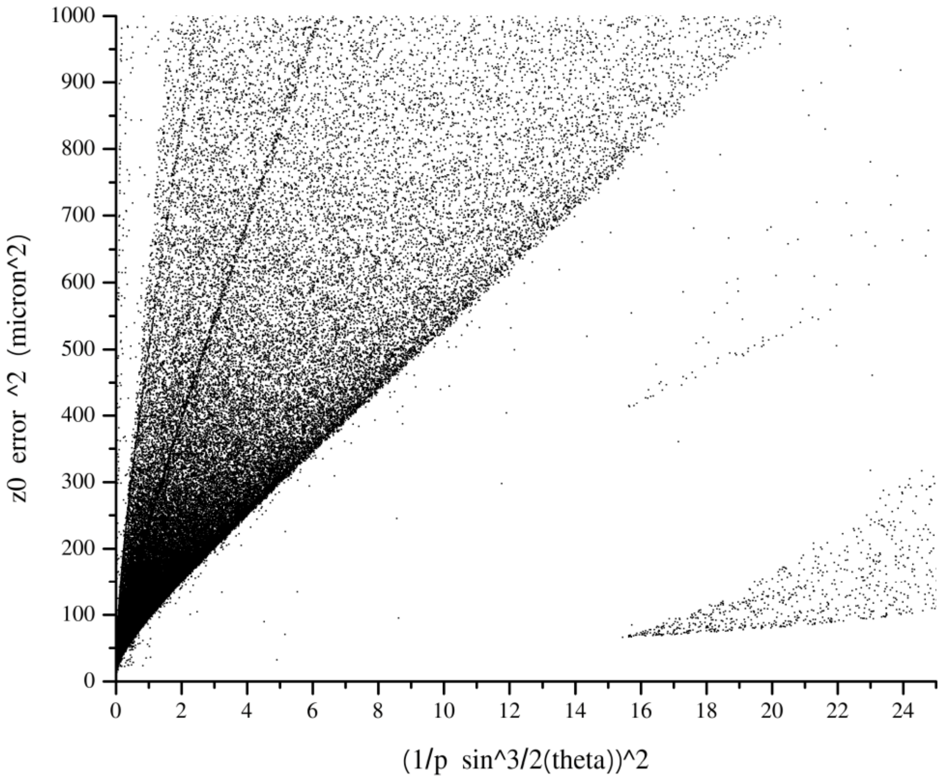






Intermediate track cheater, Silicon-only (Orsay)

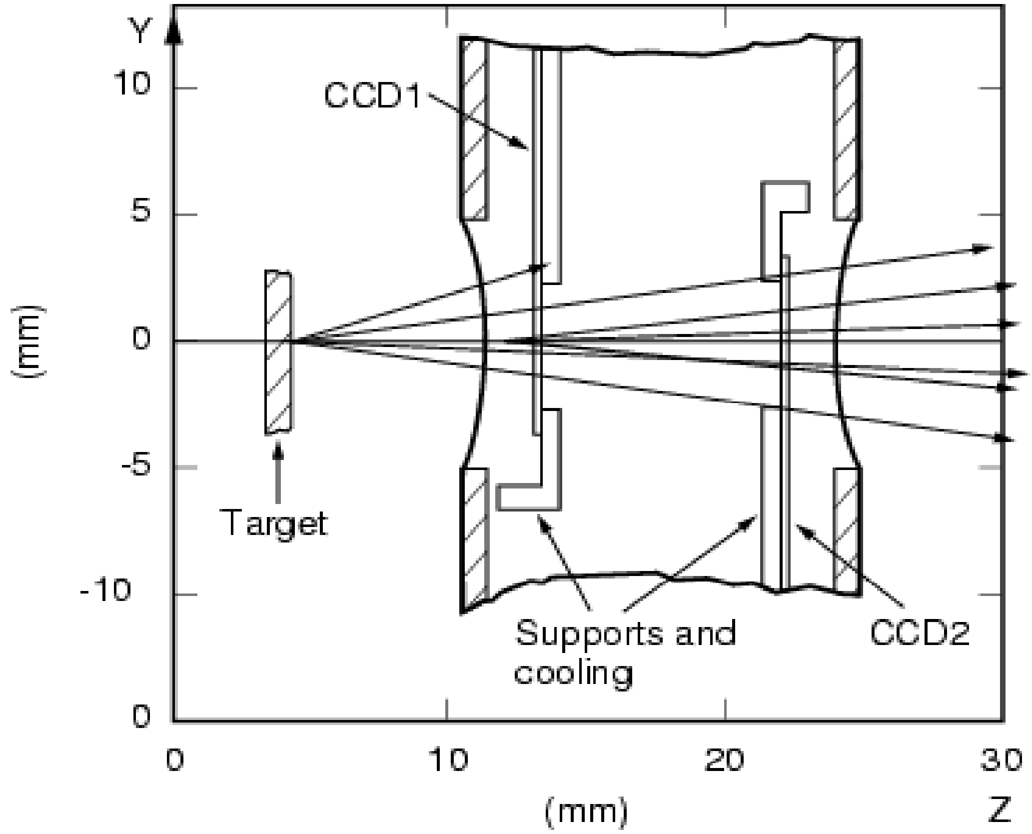




- after correction of the initial bug by Alexei Raspereza, dependence of track errors in Rϕ on momentum and angle looks OK, in z, almost OK, for "Silicon only" tracking

Hadronic interactions – previous experience

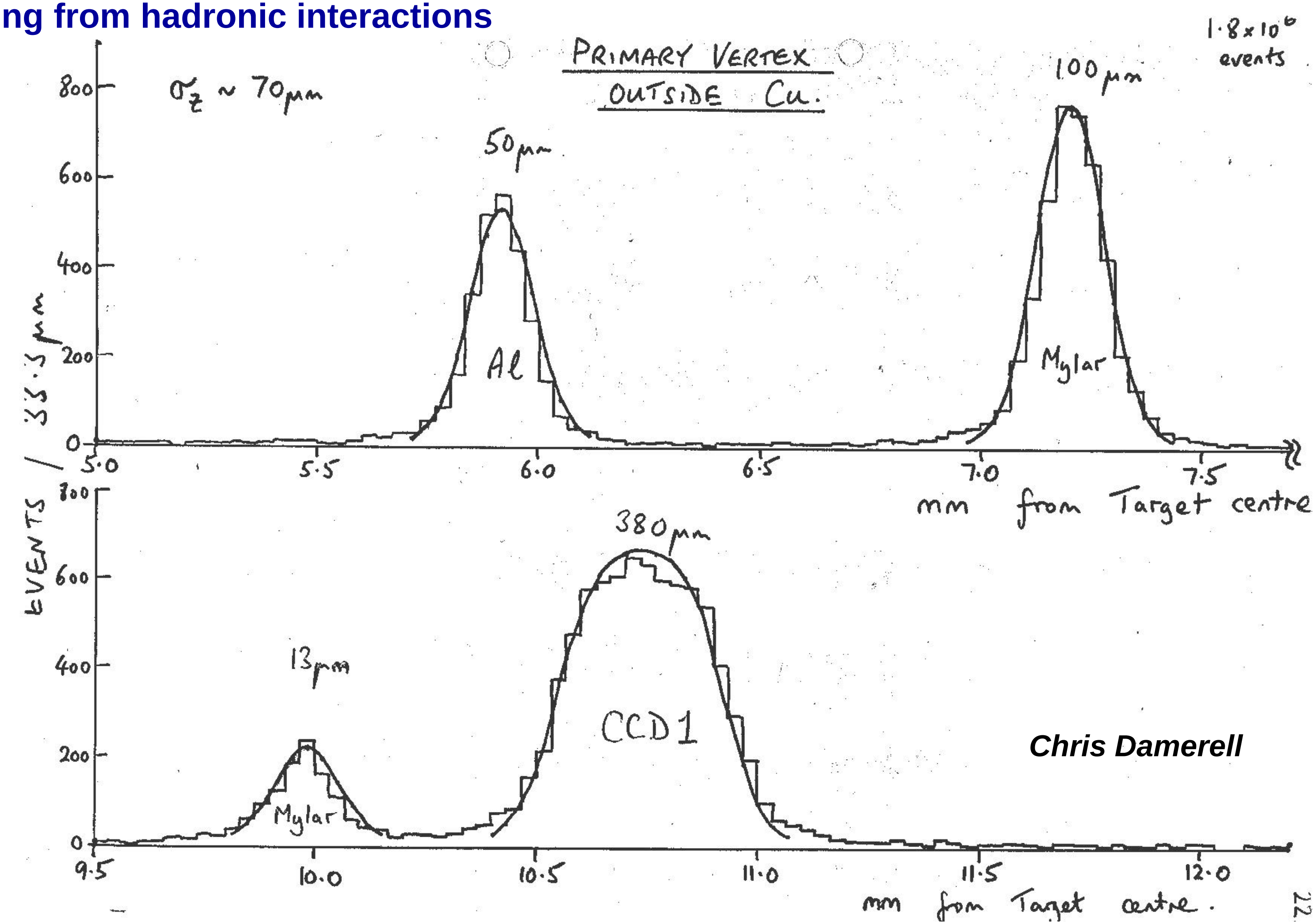
- NA32 (circa 1982) **First measurement of charm particle lifetimes with Si detectors**
- **Decays accepted from downstream edge of target to front face of CCD2**
- **But z regions of mylar window and CCD1 were excluded due to large backgrounds from secondary interactions**
- **Small efficiency loss, but events were then seen with ZERO background**



Chris Damerell

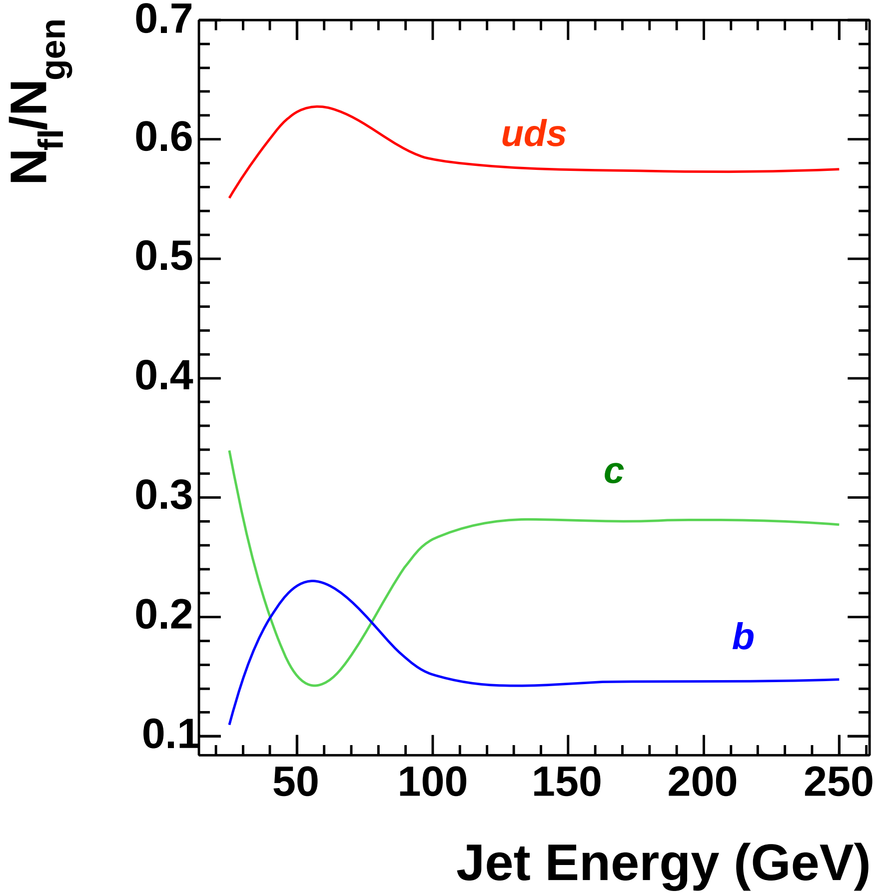
Hadronic interactions – previous experience

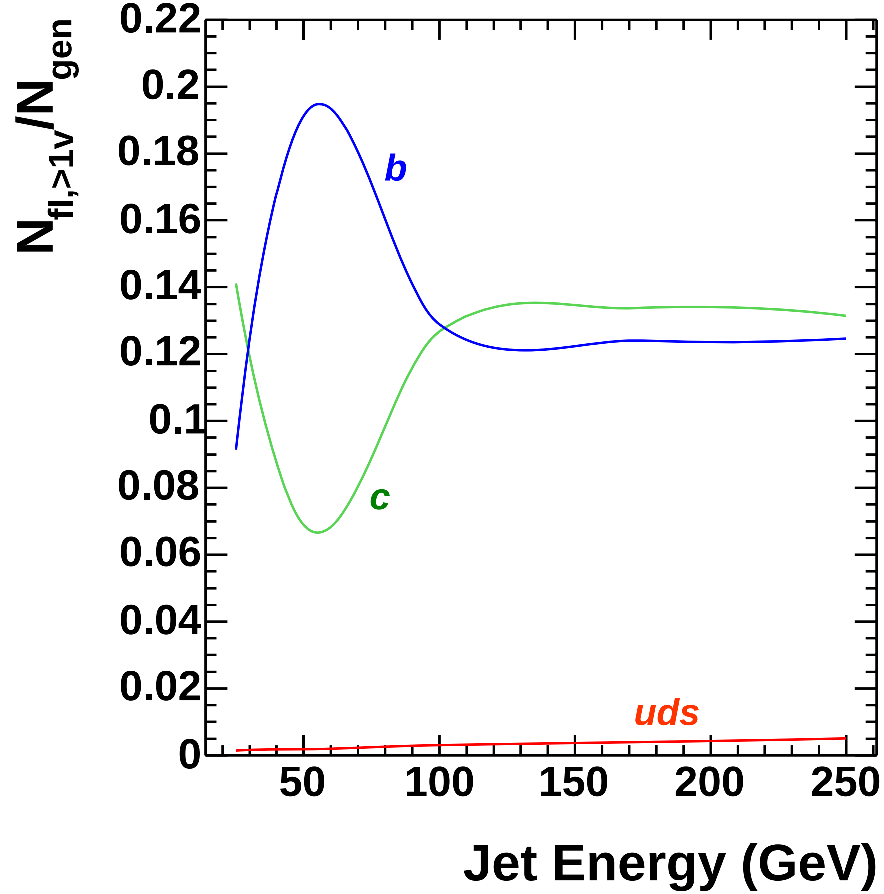
- procedure similar to ZVTOP has previously been very efficient to identify and suppress tracks stemming from hadronic interactions



Flavour composition of sample at different energies

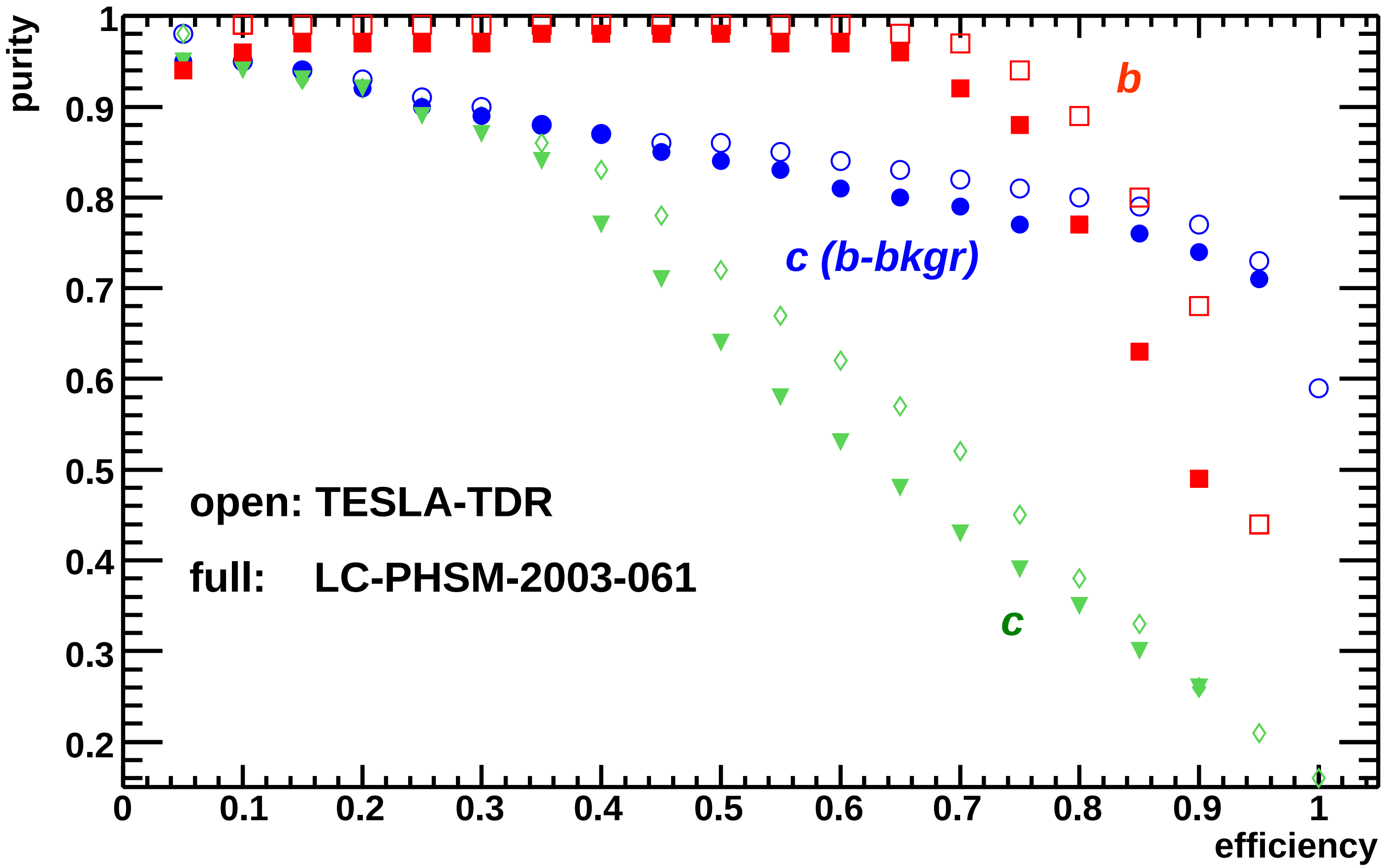
left: contributions of b, c and uds jets in generated sample, right: require > 1 ZVTOP vertex





fractions of b- and c-jets become more similar at higher energies → in that respect, b-tag becomes more challenging; increase in average decay length makes vertex finding easier

The two BRAHMS results in comparison



- LC-note result uses more realistic tracking and track selection derived from the sample used; performance slightly worse than previous TESLA-TDR result
- results shown for nominal layer thickness at time of TESLA-TDR of 0.064% X0