ILC Software and Tools workshop

LAL-Orsay, 2nd – 4th May 2007

The LCFIVertex package

- *Scope of the package*
- *Validation results from fast MC SGV*
- *Performance obtained with MOKKA / MarlinReco*

Sonja Hillert (Oxford)

on behalf of the LCFI collaboration

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, initally 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**

- **tutorials on how to use the code will be given in the afternoon session today**

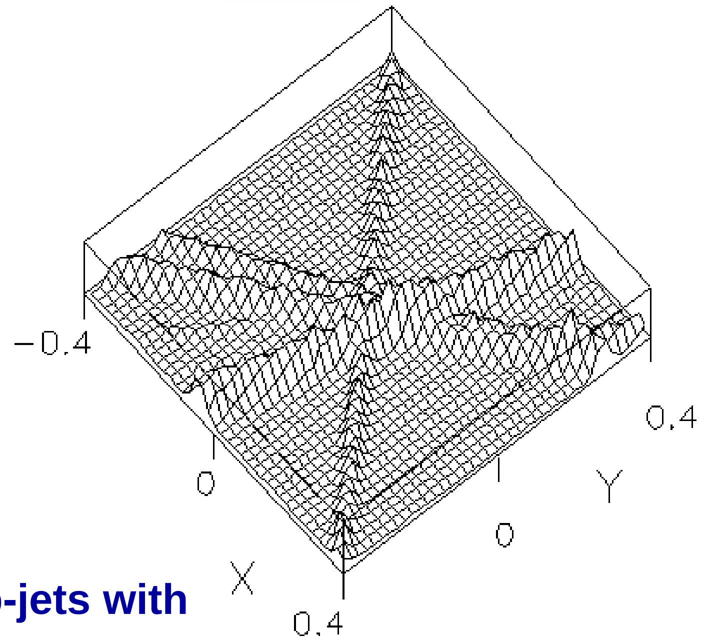
The ZVTOP vertex finder

D. Jackson,

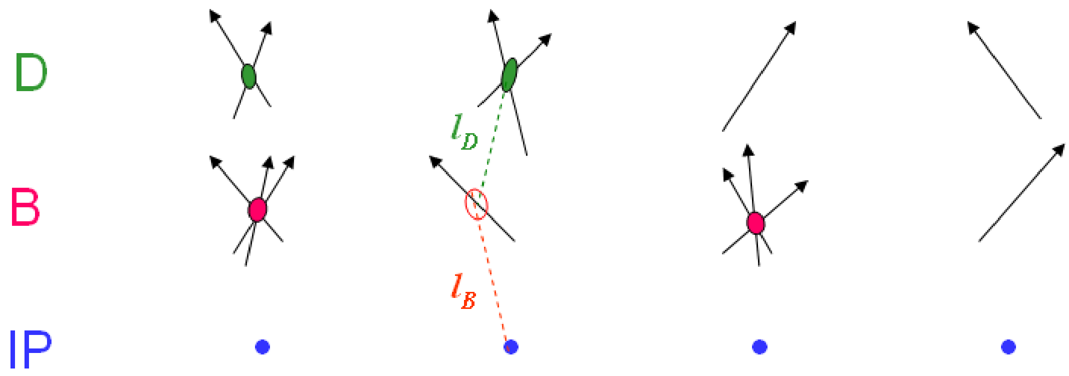
NIM A 388 (1997) 247

- **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**

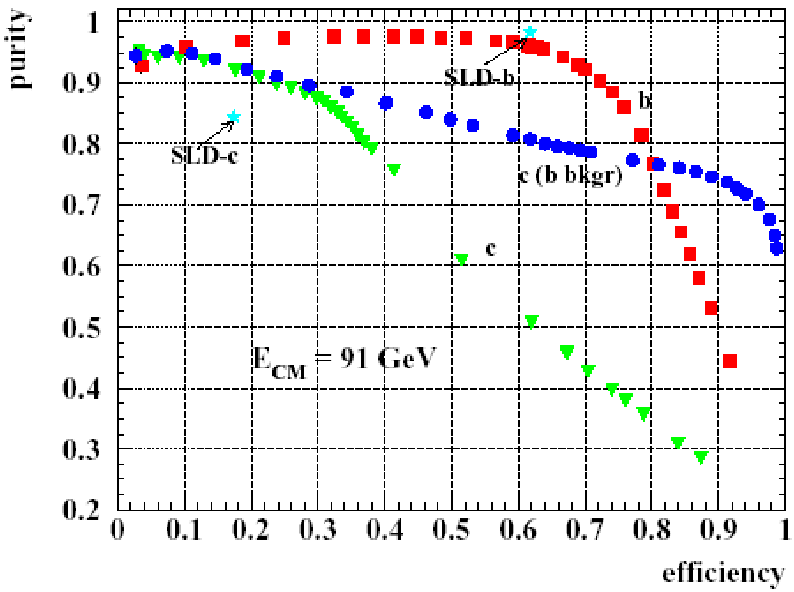


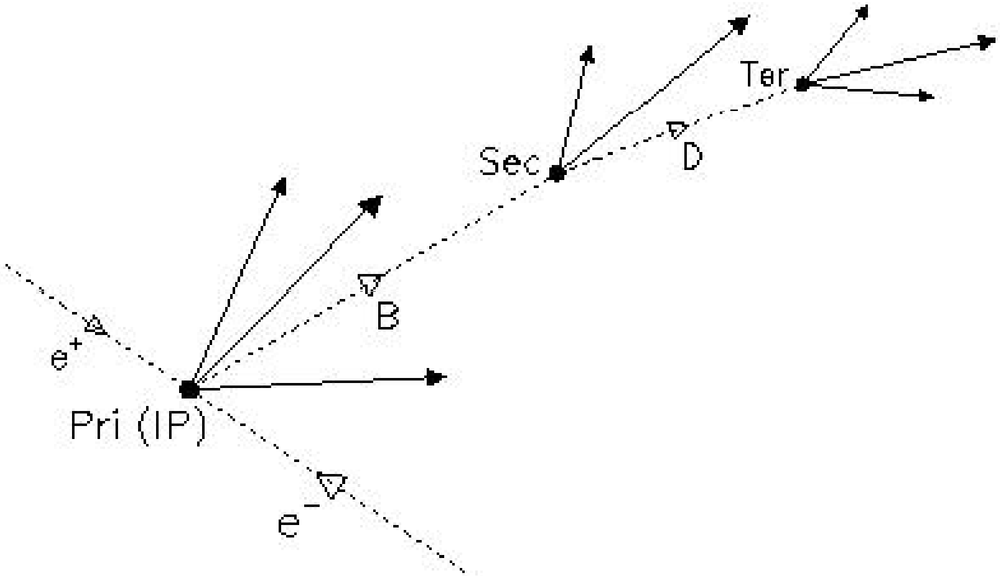
- **additional kinematic information (IP-, B-, D-decay vertex approximately lie on a straight line) used to find vertices**
- **should improve flavour tag efficiency and determination of vertex charge**

Flavour tag and quark charge sign selection

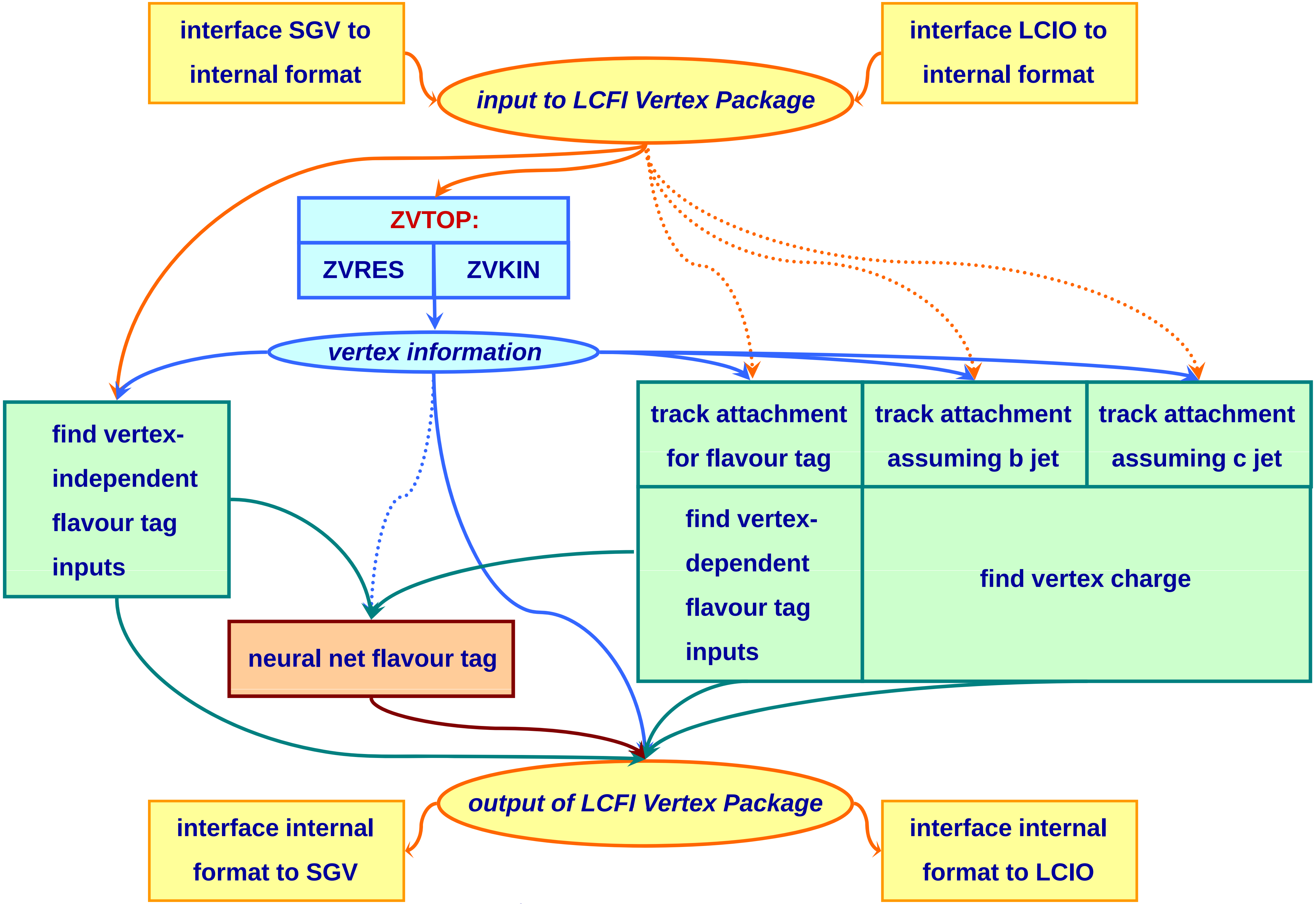
- aim of flavour tag: distinguish between b-jets, c- jets and light-quark / gluon jets
- heavy flavour jets contain secondary decays, generally observed as secondary vertices
- NN-approach to combine inputs; most sensitive: secondary vtx Pt-corrected mass & momentum

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

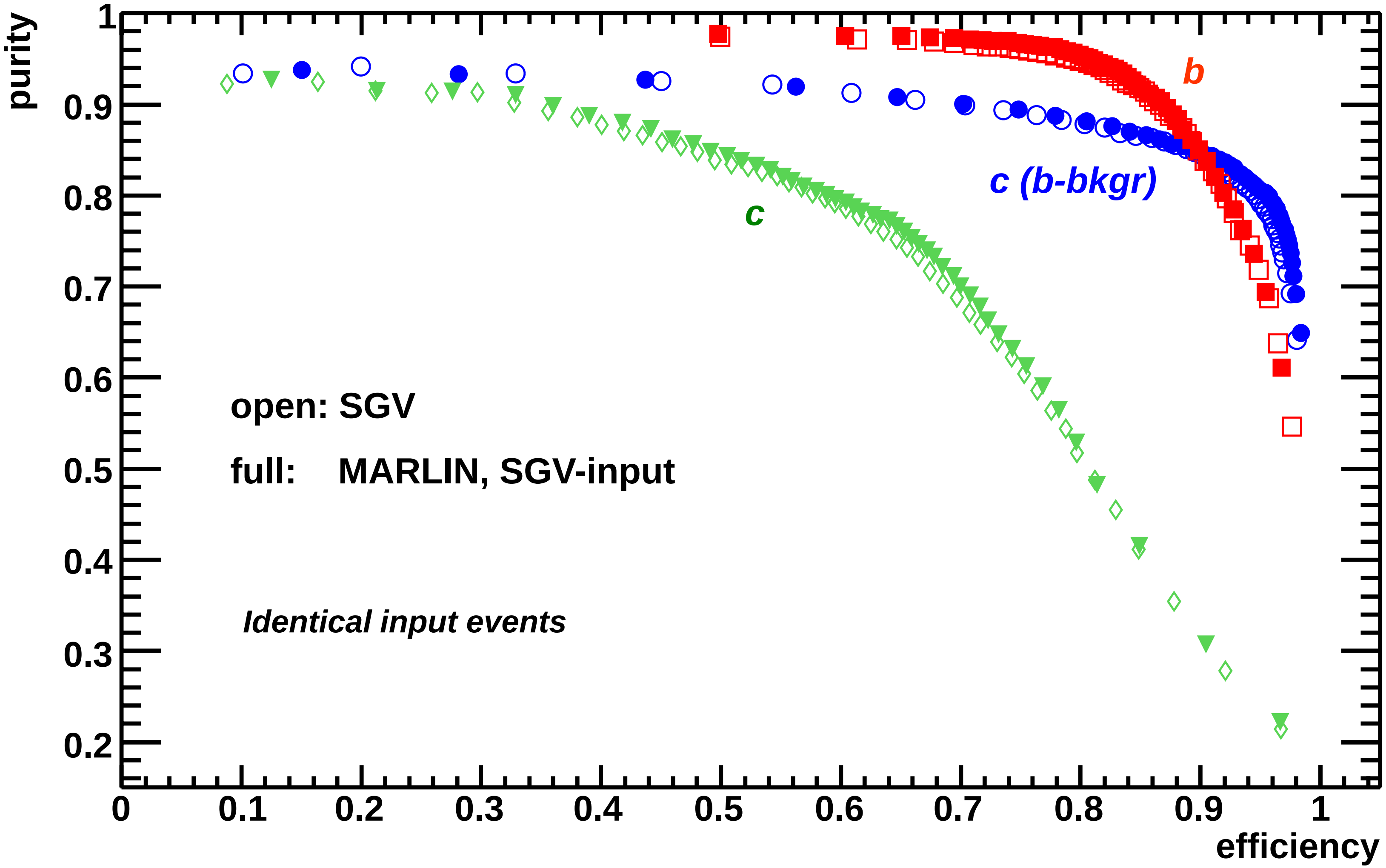
  results from those tests: Ben Jeffery's talk at ECFA workshop, Valencia, 2006

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

➢ **Tests using MarlinReco event reconstruction**

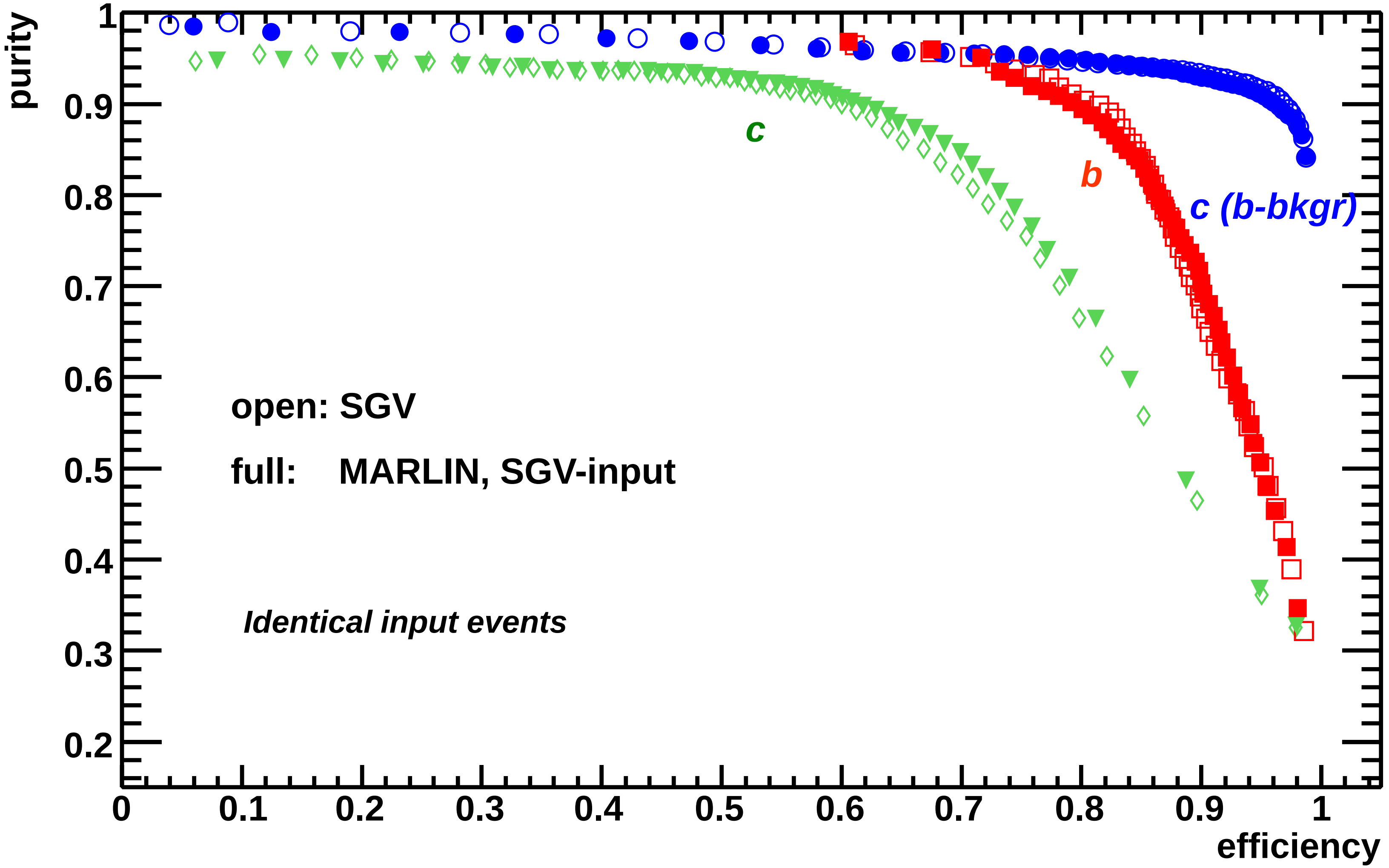
- **full chain test repeated with PYTHIA events, passed through MOKKA**

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

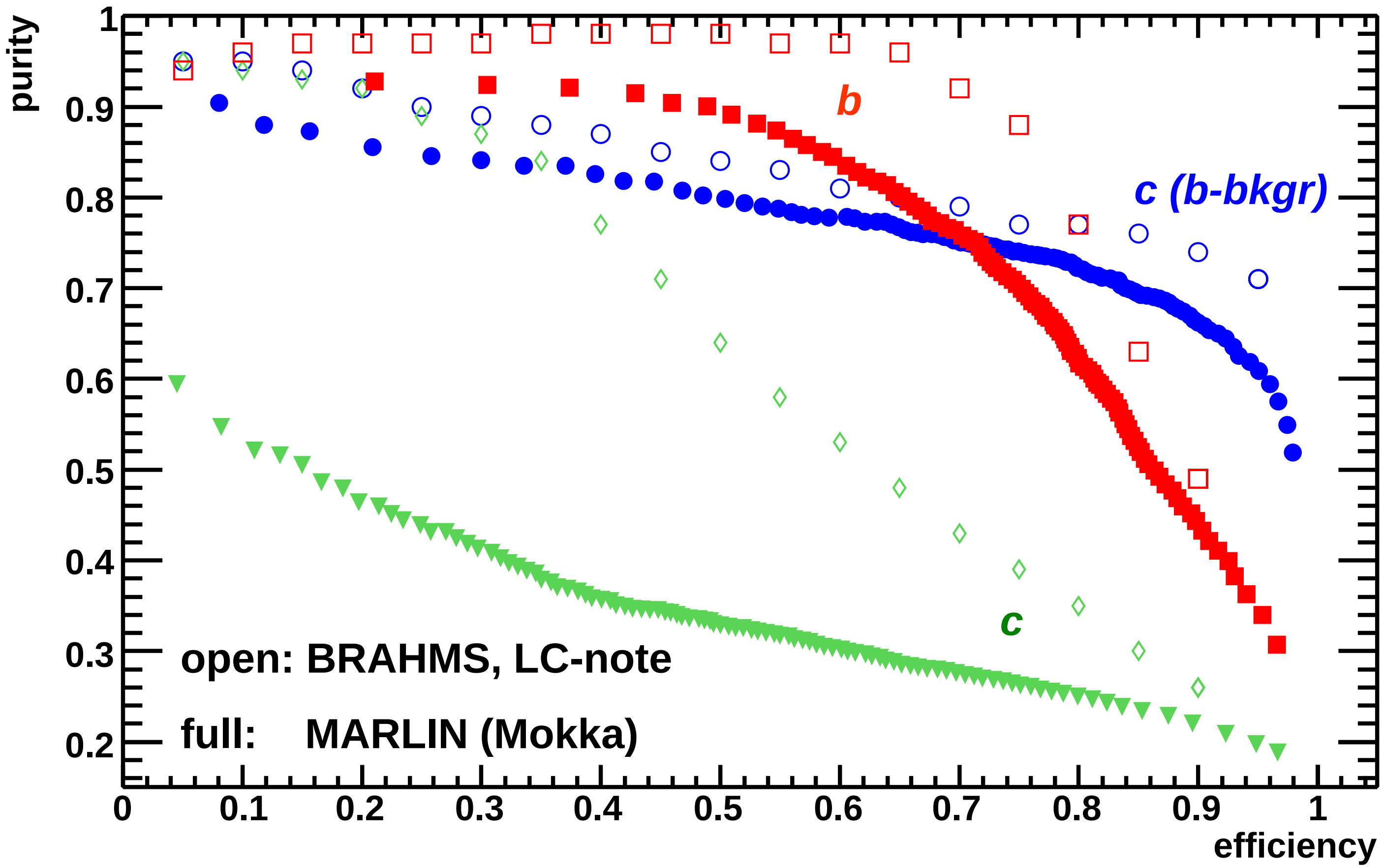
  - note: photon conversions switched off in GEANT, as these can easily be suppressed later
  - also, hadronic interactions in the beam pipe and in the vertex detector layers suppressed by a radius cut at the track selection level using MC information (optional, small effect, geometry currently hard-coded)

  - for tracking use Alexei Raspereza's track cheater (omit pattern recognition as in SGV) 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

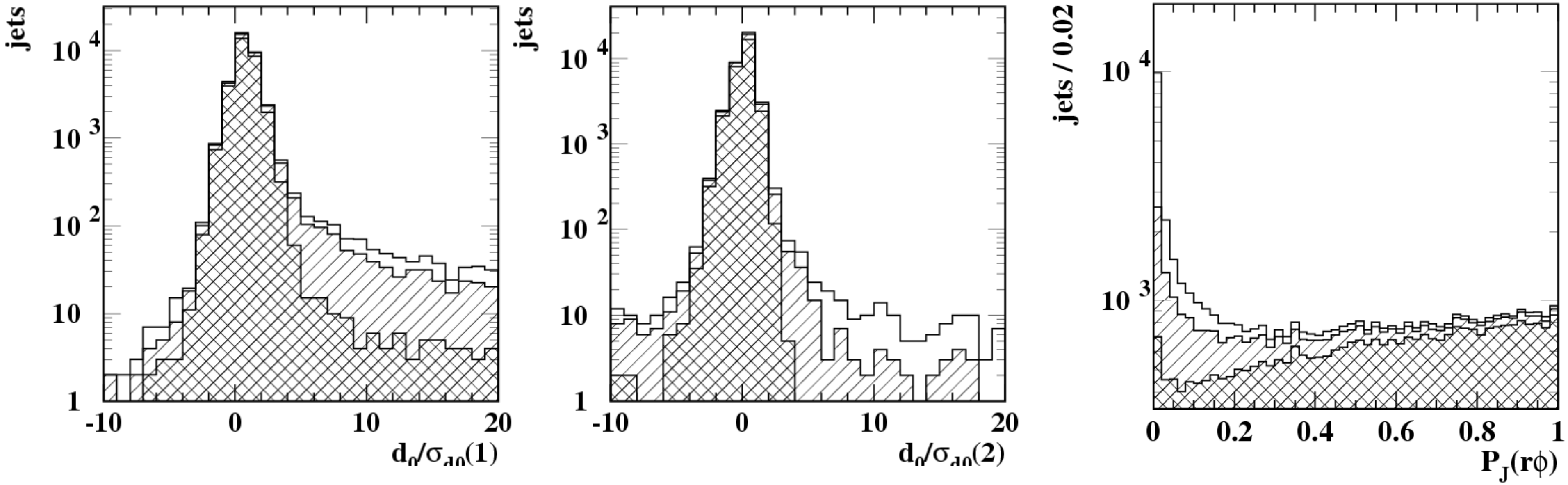
  - default track selection of the LCFIVertex package was derived for previous BRAHMS study

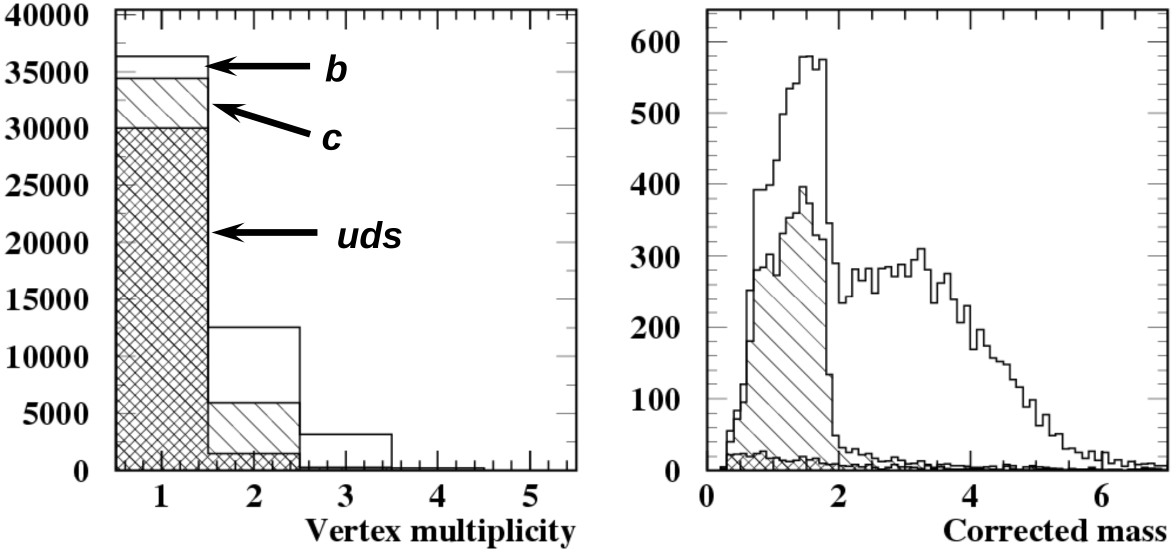
Initial MARLIN result with MOKKA input



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

Some input distributions from LC-PHSM-2000-021

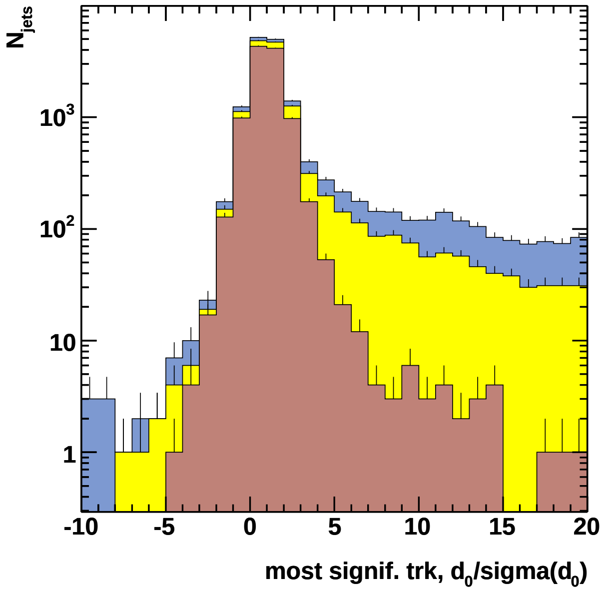


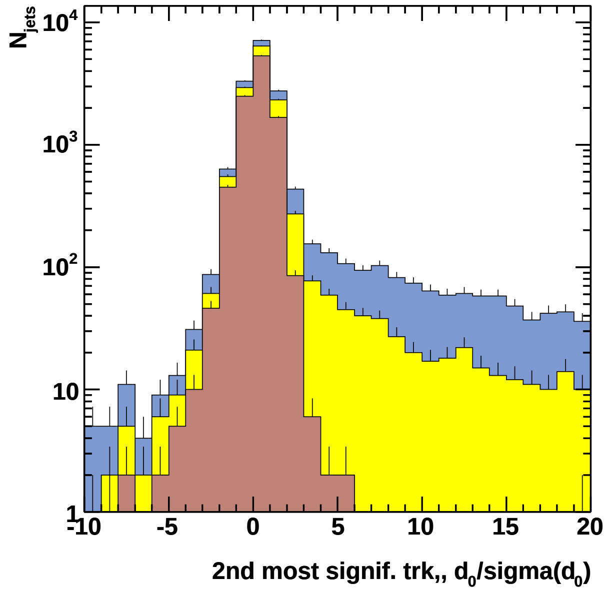


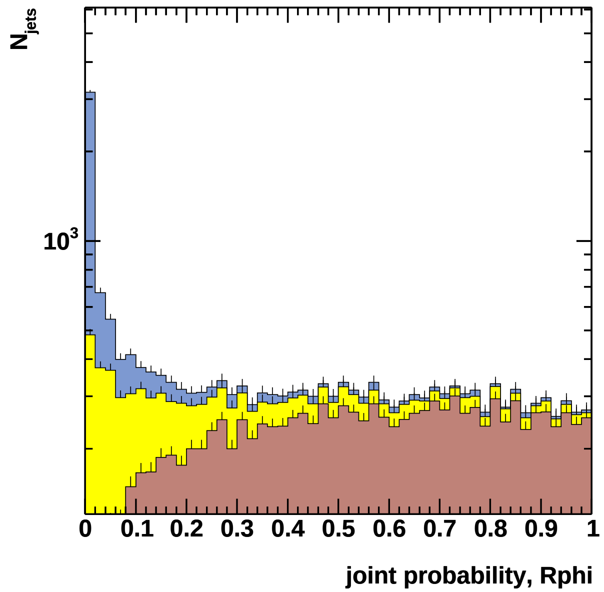
most sensitive input variables for the flavour tag neural net

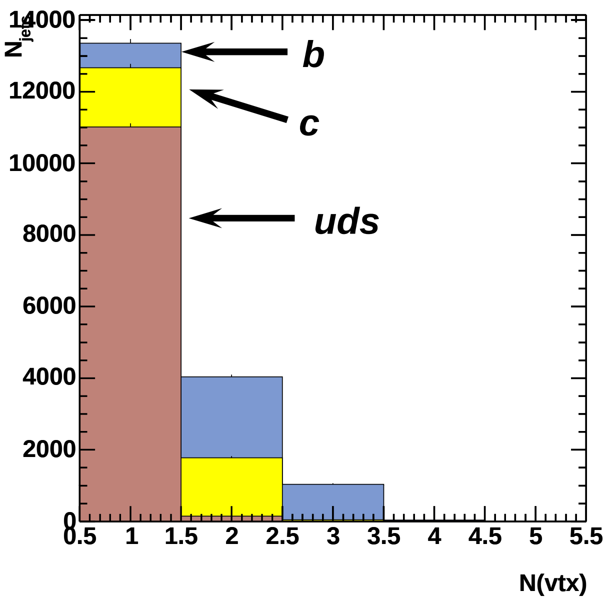
the Pt-corrected mass requires at least one secondary vertex to be found → different inputs for NN depending on number of vertices

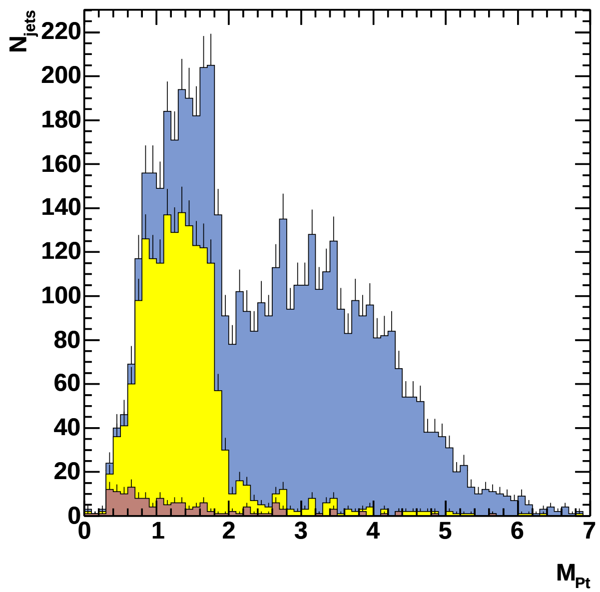
Input distributions from track cheater without TPC







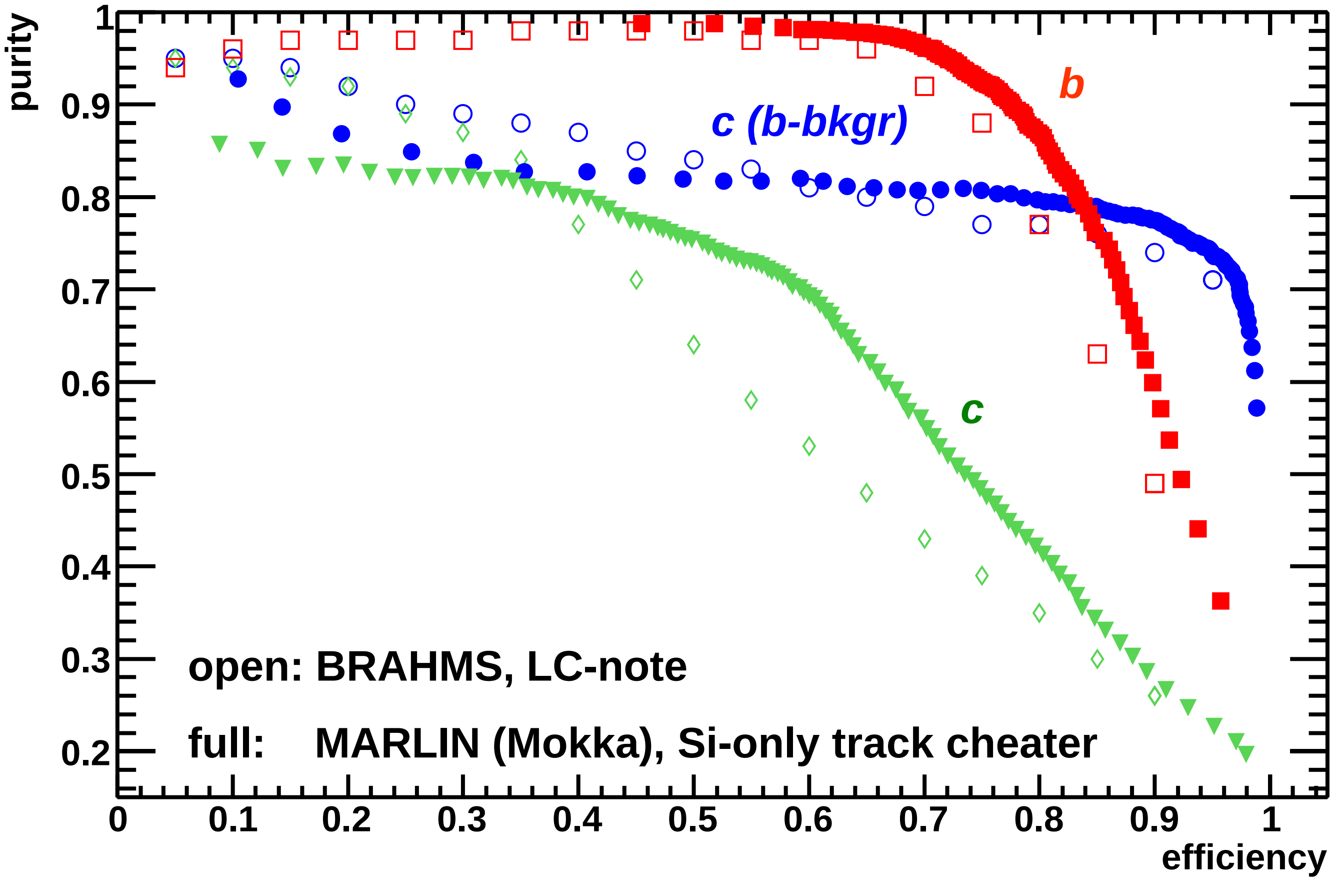




good agreement of MOKKA inputs with BRAHMS ones

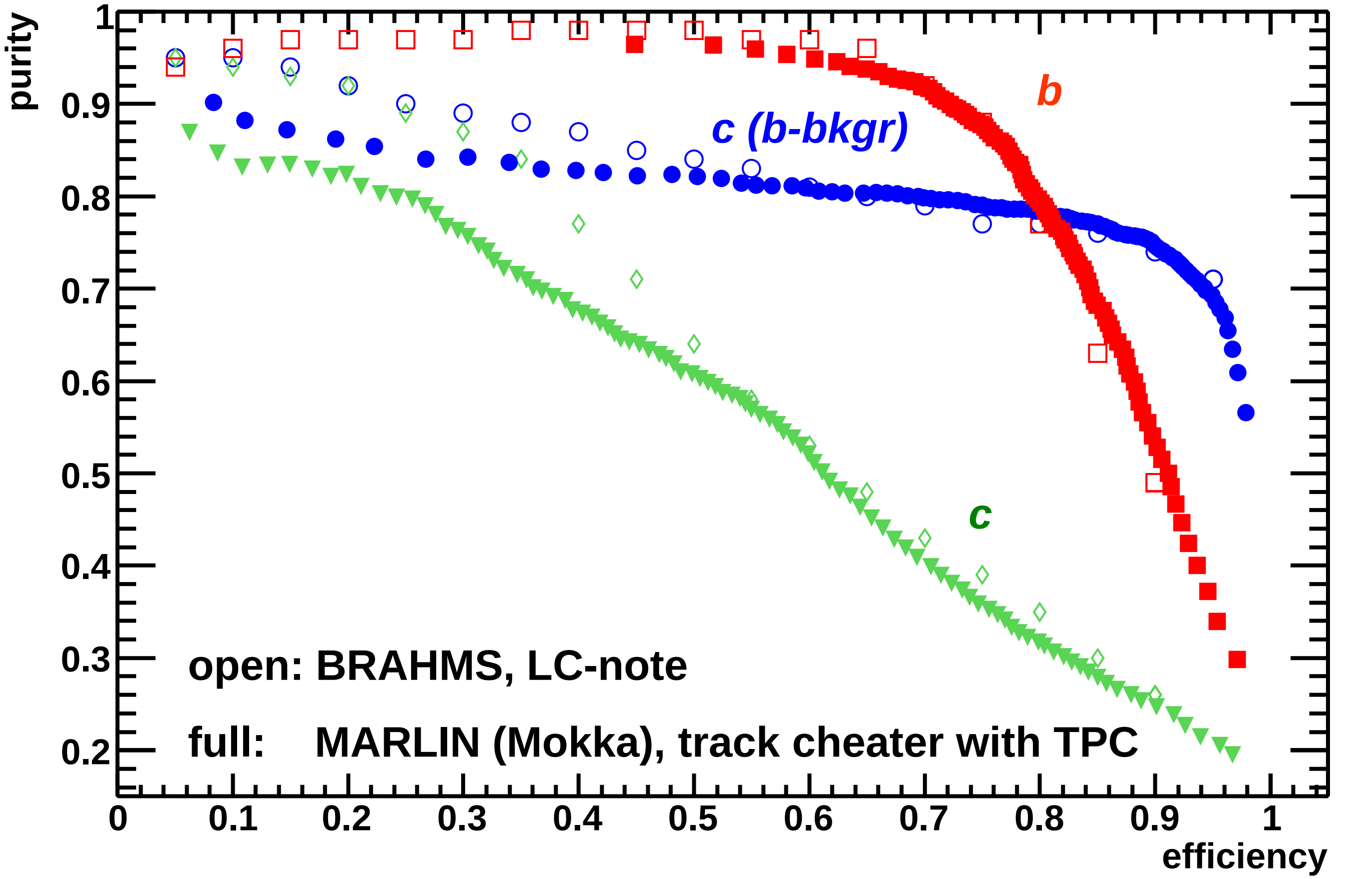
slightly better discrimination since MOKKA assumes point resolution 2 μm, BRAHMS assumed 3.5 μm

Resulting purity vs efficiency at the Z-peak



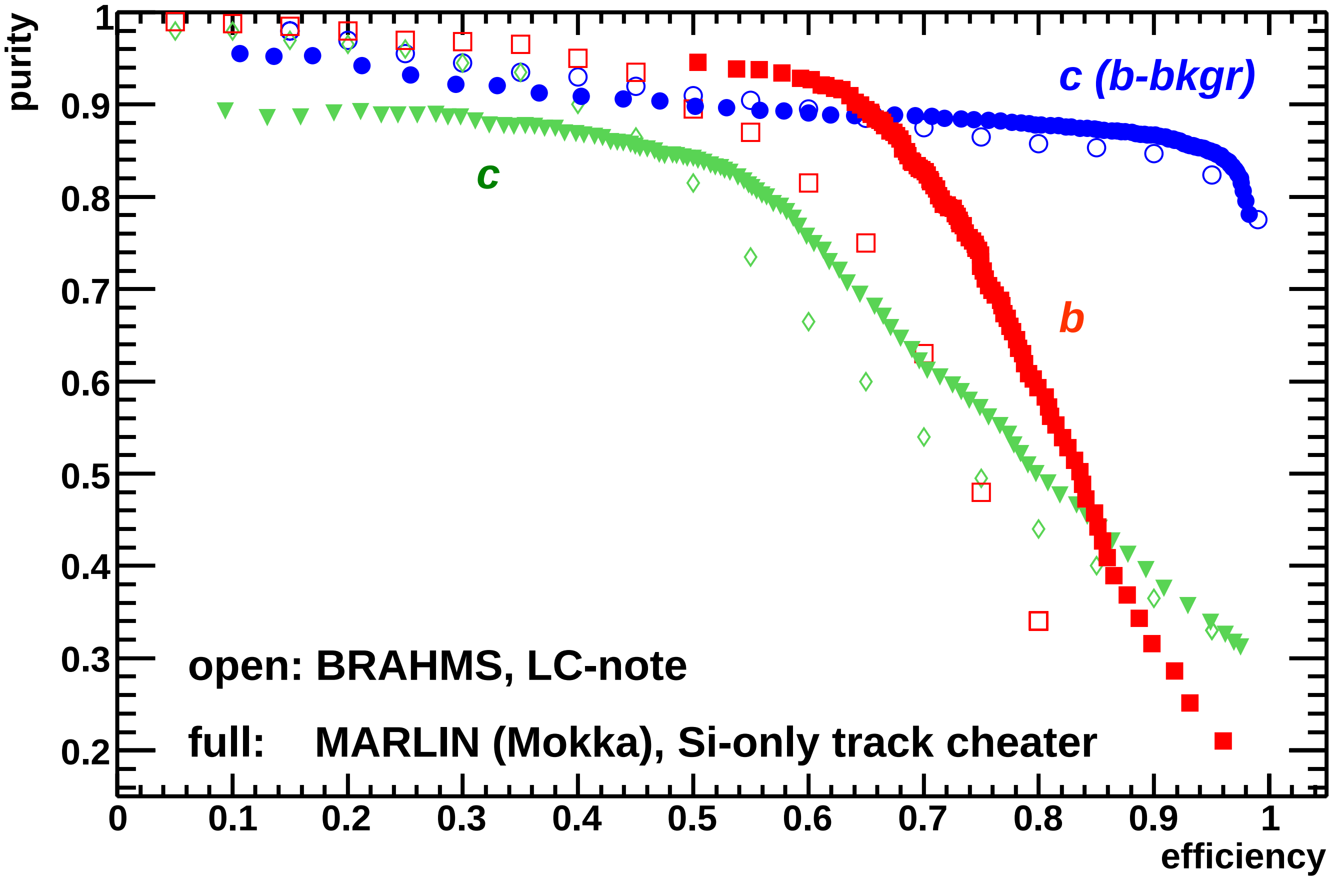
- at high efficiency MARLIN(MOKKA) with "Silicon-only" track cheater gives better performance compared to LC-note result using tracking with pattern recognition

Purity vs efficiency at the Z-peak, TPC included



- track cheater with TPC still has some problems, under investigation by Alexei
- this results in slight decrease of performance of tagging code, if run with this input

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 (this result: 100 MeV)

Summary and outlook

- **The LCFIVertex package is now 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**

- **work on the package will continue, e.g.**
  - **extend performance study to full LDC tracking (useful diagnostic)**
  - **add diagnostic features to increase user-friendliness, and further documentation**
  - **reassess flavour tagging procedure and vertex charge determination in ILC environment (tuning cuts, parameters & algorithms)**

Additional Material

Flavour tag

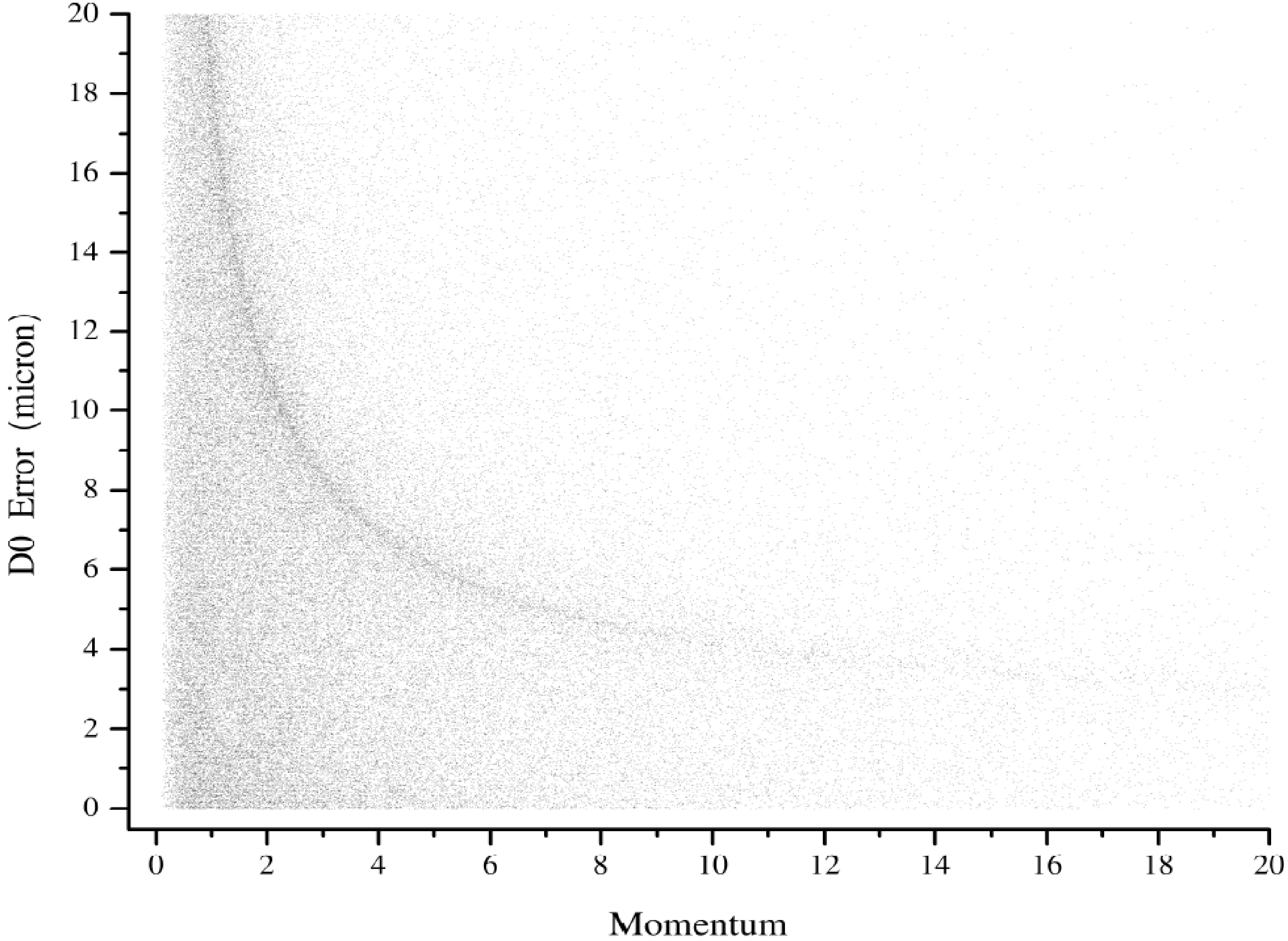
- Vertex package provides flavour tag procedure developed by R. Hawkings et al (LC-PHSM-2000-021) and recently used by K. Desch / Th. Kuhl as default

- 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

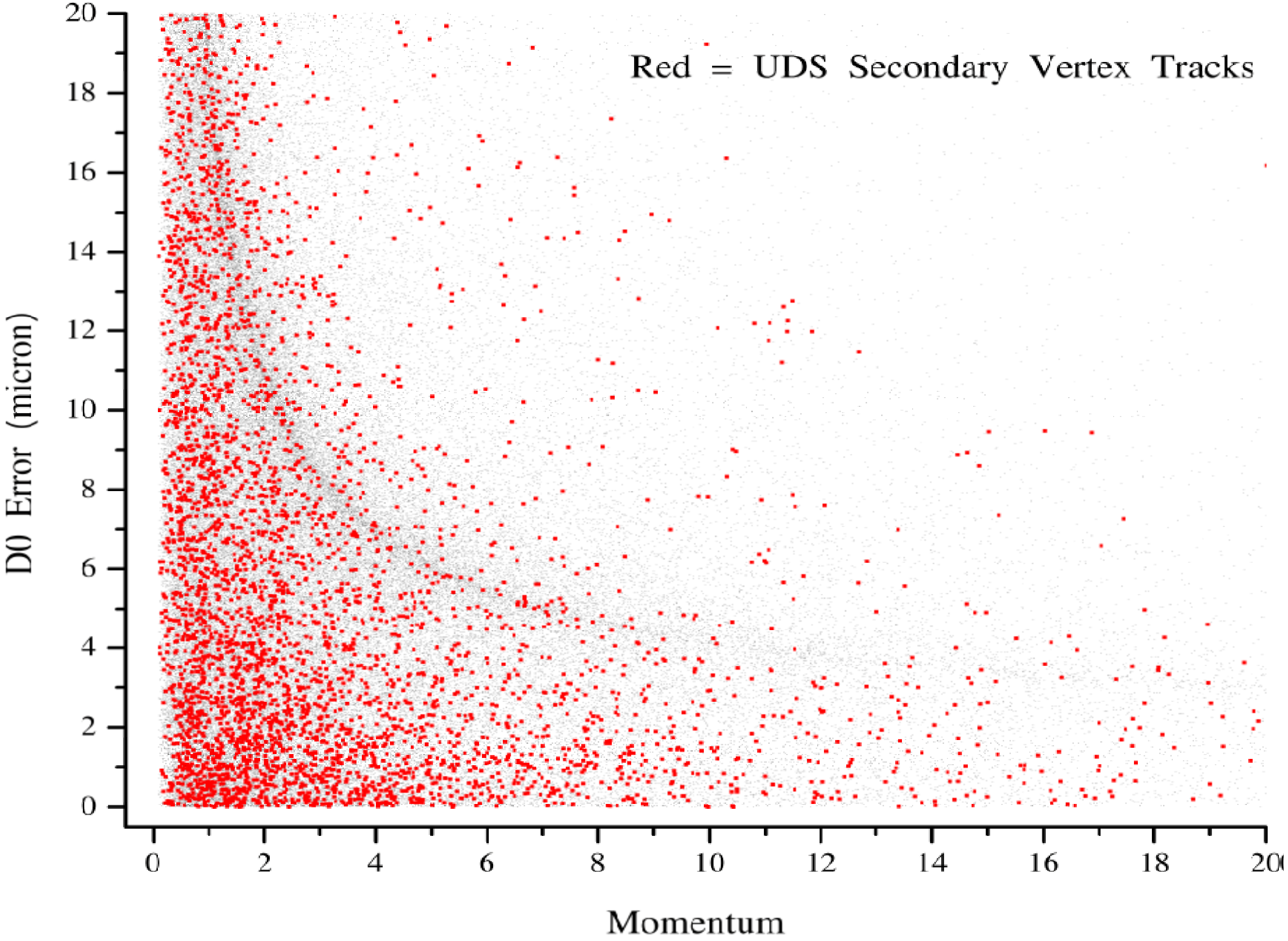
System Test result related to track cheater problem

- exceedingly large impact parameter significances for tracks in uds jets were traced down to a large number of tracks having unrealistically impact parameter errors in R-phi



Secondary vertices in uds-jets

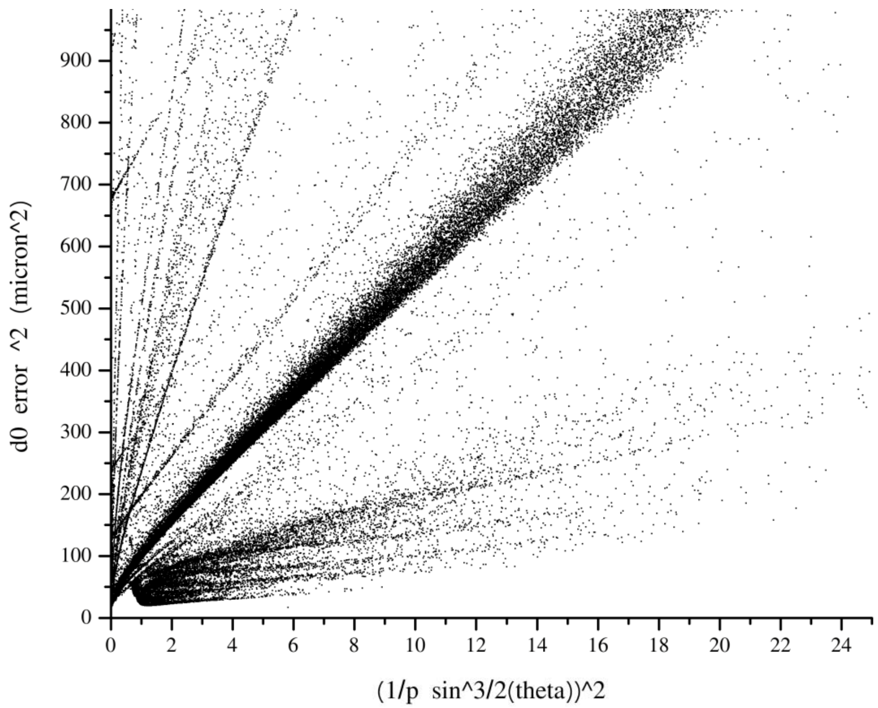
➢ **this also explains the unusually large number of secondary vertices found in uds jets:**

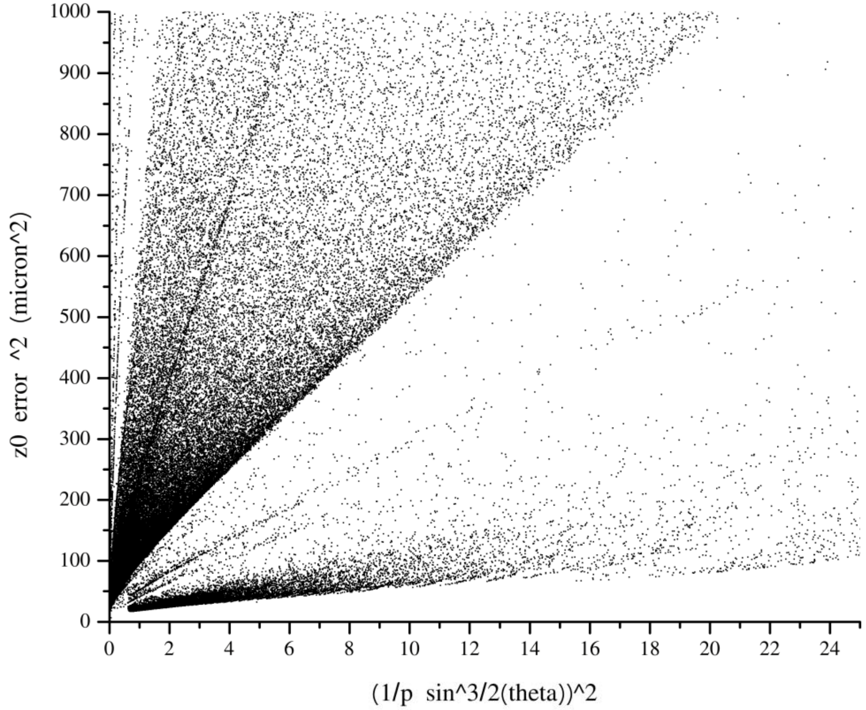


➢ **Alexei Raspereza contacted about this last Thursday**

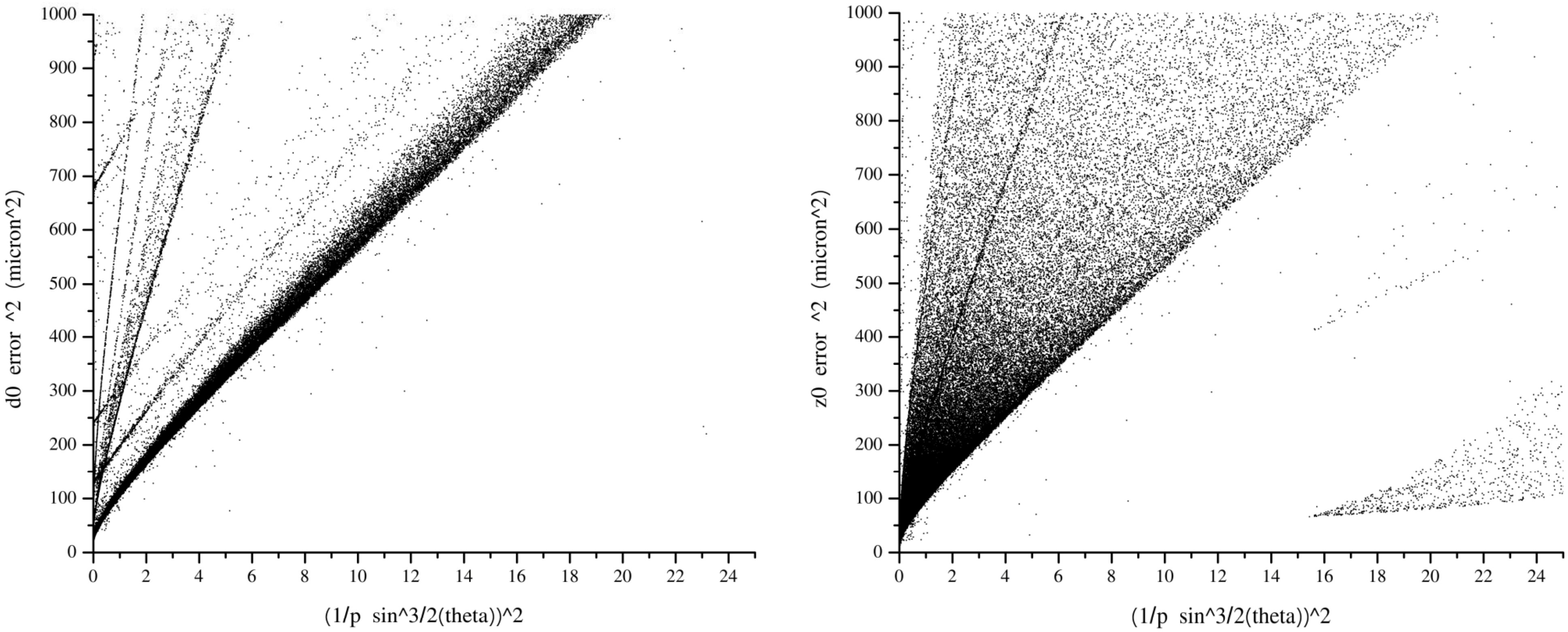
Intermediate track cheater, TPC included

- momentum and angular dependence of impact parameter error σ(d0) confirms that there are still problems when including the TPC in the track cheater run

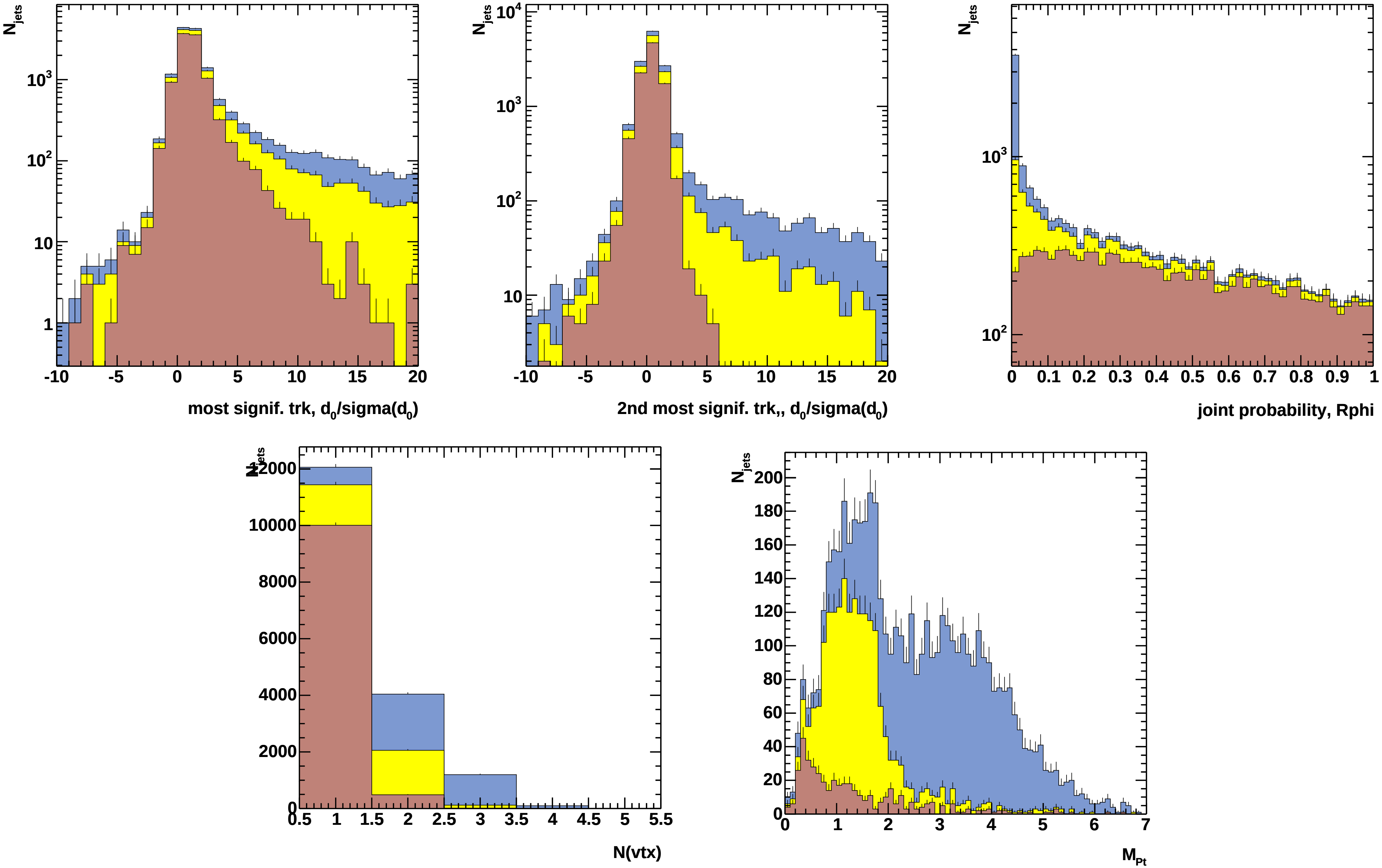




Intermediate track cheater, Silicon-only (no TPC)

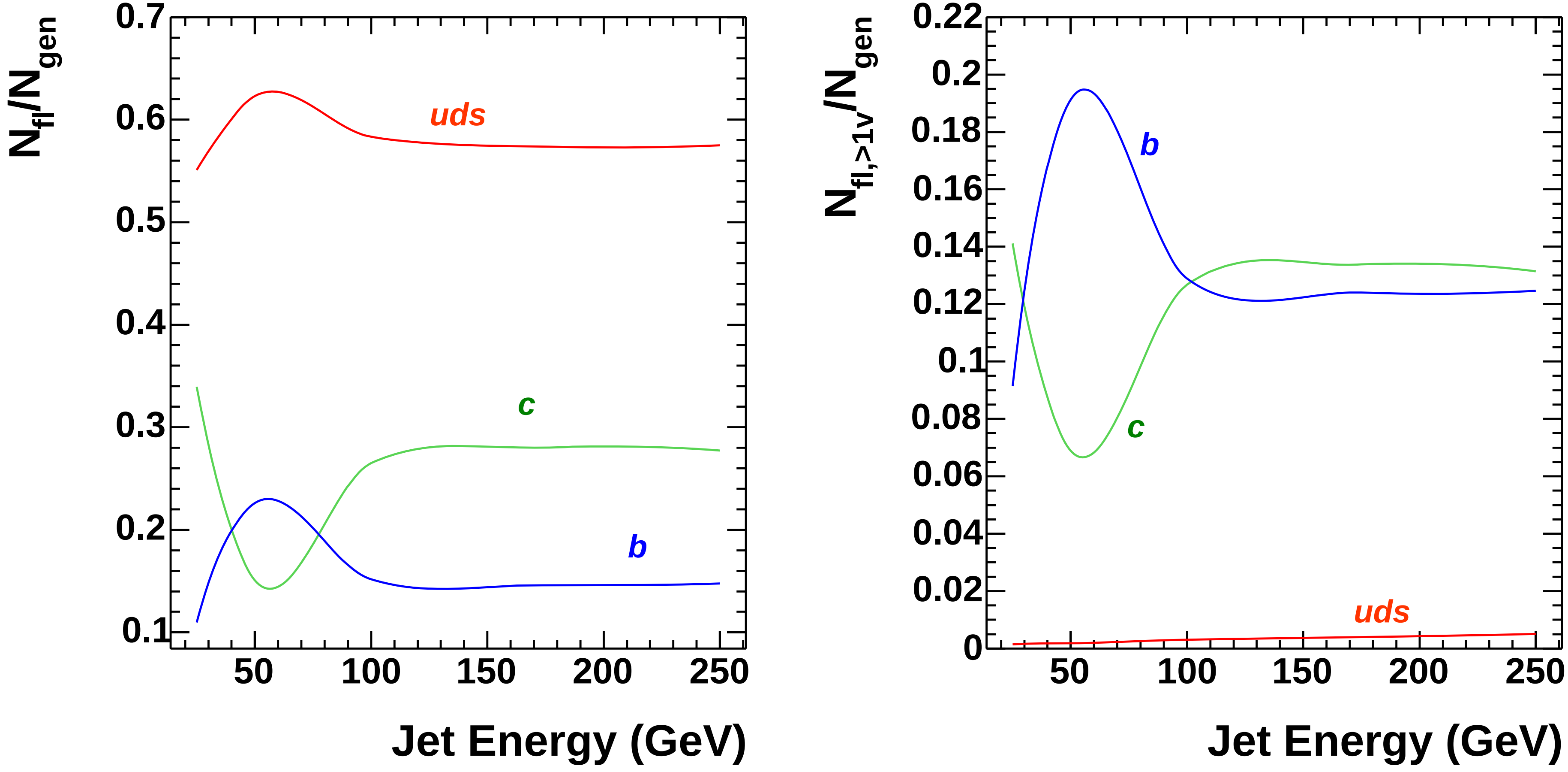


Input distributions from track cheater with TPC



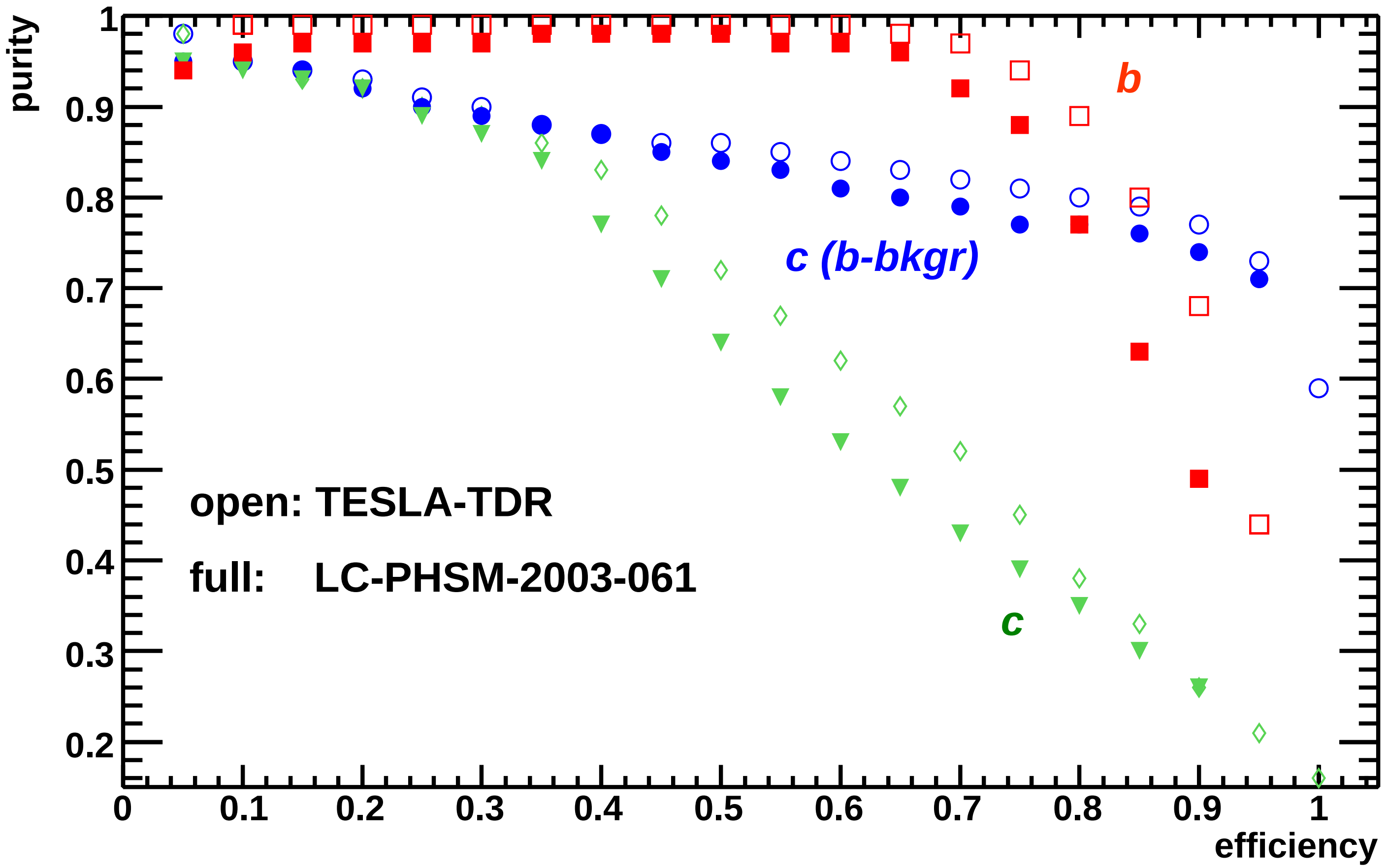
Flavour composition of sample at different energies

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**