SiD Simulation and Reconstruction for the DBD

Norman Graf (SLAC)

ALCPG11, Eugene

March 22, 2011

slicPandora

- Frontend to the PandoraPFANew project.
  - LCIO binding, Geometry format, Sampling Fractions
- Uses geometry XML file generated by GeomConverter from compact detector.
- Reads input LCIO file with simulated events, tracks, and track states.
- Outputs LCIO file with Reconstructed Particles.
- Uses standard XML configuration file for PandoraPFA algorithm settings.
  - Pandora PFA (M. Thomson, J. Marshall)
  - slicPandora (N. Graf, J. McCormick, P. Speckmayer, C. Grefe, M. Stanitzki, J. Strube)

Workflow

- Generate LCIO events using slic and detector of your choice.
- Run SLIC output through LCSim tracking to generate Track and Track State collections.
- Take LCSim LCIO output and run in slicPandora to generate PFOs.
- Take LCSim LCIO output and run in uipfa to generate PFOs.

} Same Events

- Use LCIO analysis tool of your choice to work with PFOs (JAS3, ROOT, etc.).

slicPandora as analysis tool

- Same simulated Events can be analyzed using uipfa or slicPandora
  - study effect of algorithms/code on identical input

- Same stdhep input Events can be analyzed in different detectors
  - same input events can be processed through different detectors
    - compare digital to analog HCal with identical code

Calorimeter Calibration

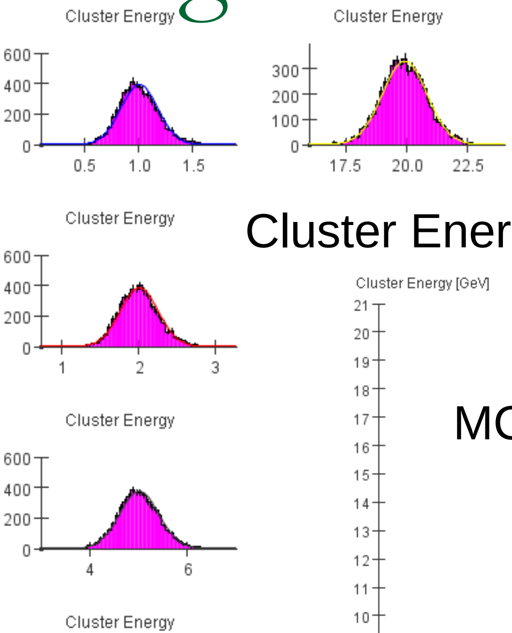
- The lcsim-cal-calib package is used to generate sampling fractions from single particle data.
  - Range of single particle types and energies simulated in slic.
  - Events are analyzed in lcsim using simple fixed-cone clustering algorithm to derive sampling fractions.
- Same events are run in slicPandora and fit again.
  - PandoraPFA uses different clustering algorithm, so sampling fractions must be refitted to match final cluster energies.
  - Final sampling fractions put into CalorimeterCalibration.properties file.

Optimization Data Sets

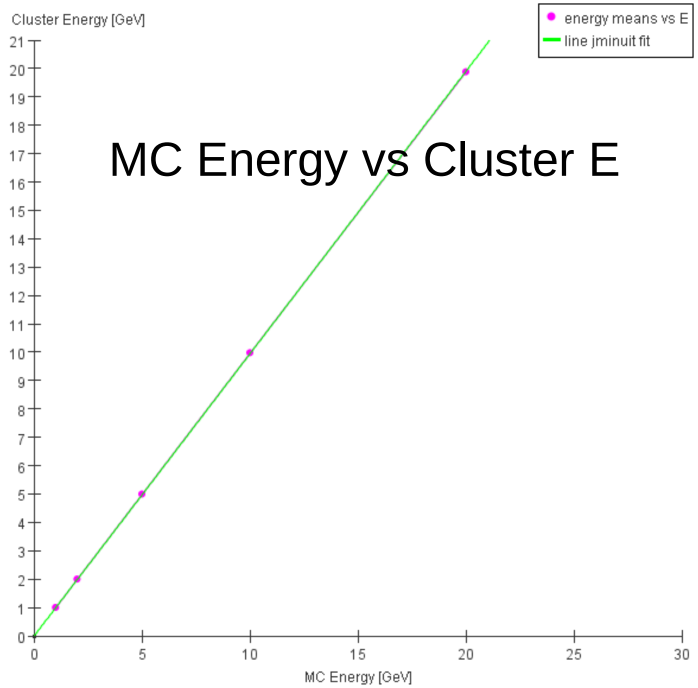
- Single $\mu^{\pm}$, γ, K^0_L, at fixed angles and energies for sampling fraction determination.
- Single particles (as above, plus $e^{\pm}$, $\pi^{\pm}$,$K^{\pm}$,$p^{\pm}$,…) at variable angles and energies to study clustering and tracking efficiency and resolution.
- Simple resonances (π^0,η,$\rho^{\pm}$) to study efficiency and resolution of two-particle states.
- Single quarks at fixed energies to study jet energy resolution (u,d,s).
- Single Z^0 at fixed energies to study dijet mass resolution.

Diagnostic Plots

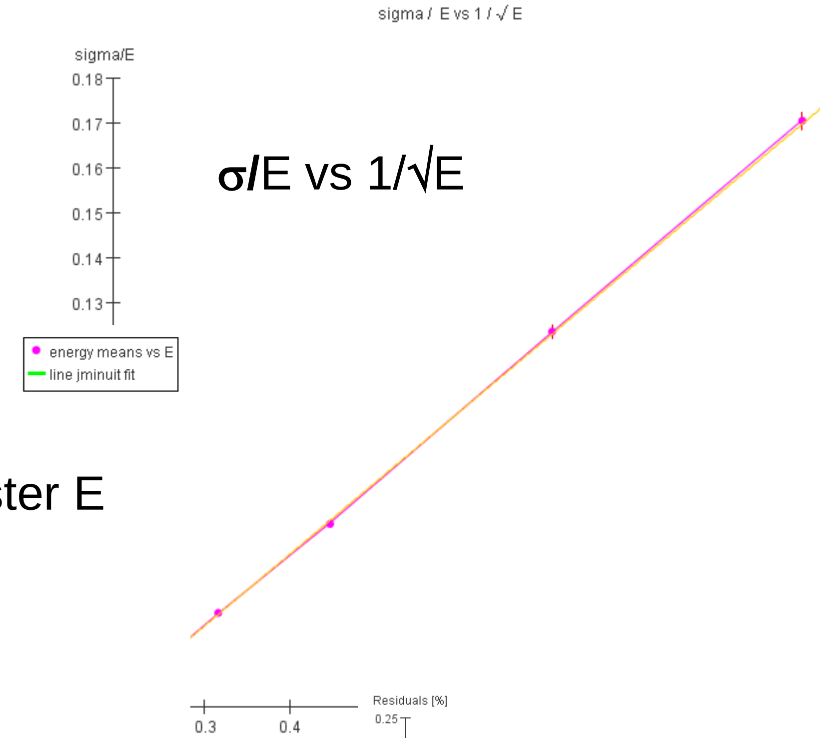
Cluster Energy



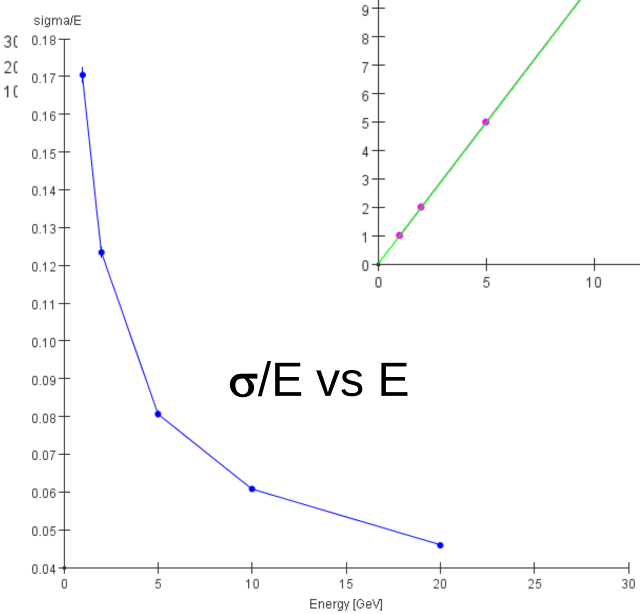
MC Energy vs Cluster E



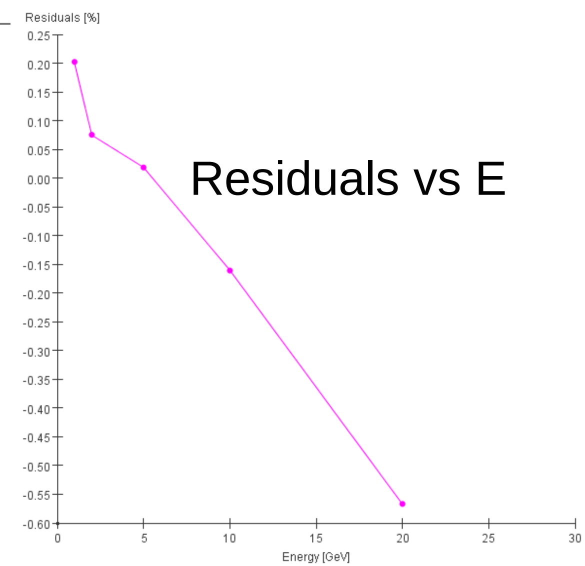
σ/E vs 1/√E



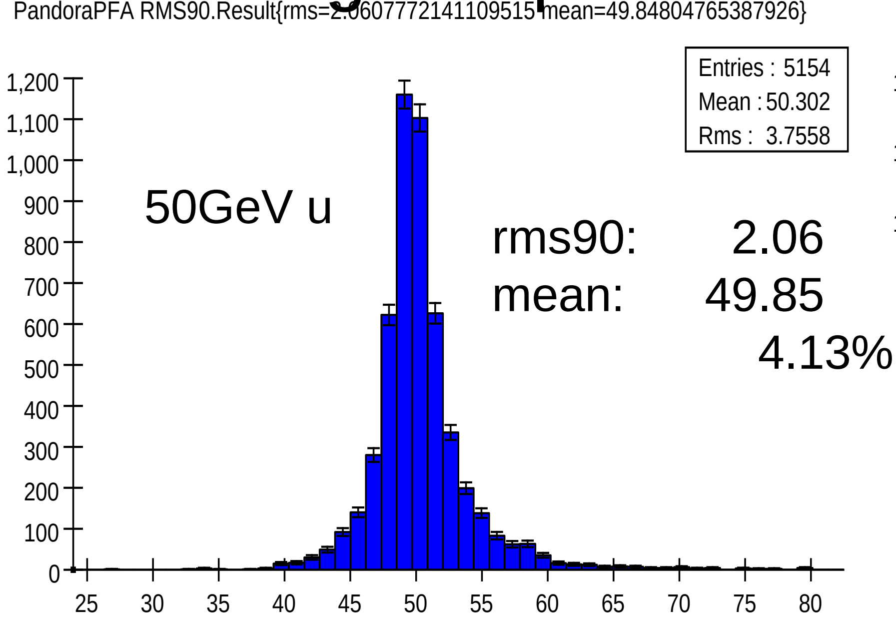
σ/E vs E

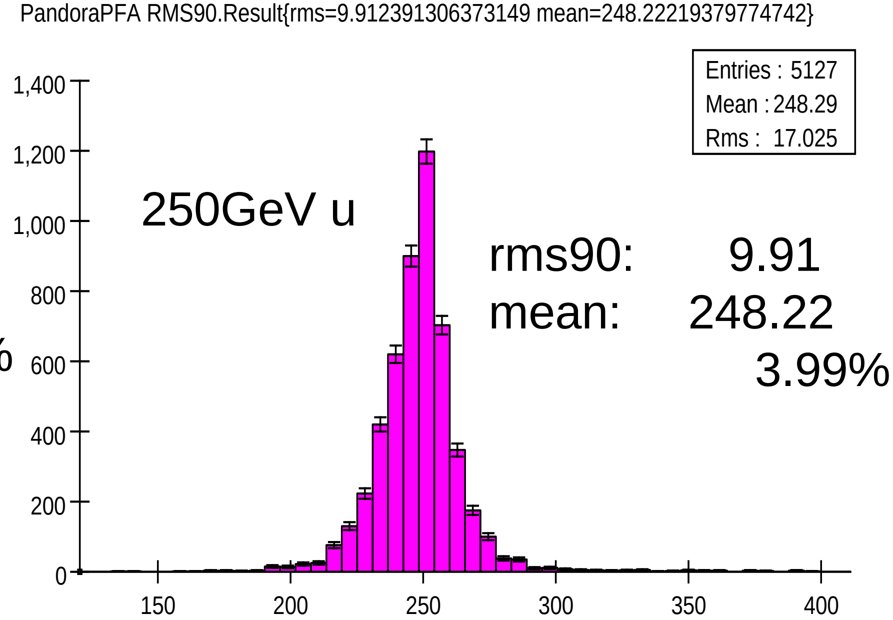


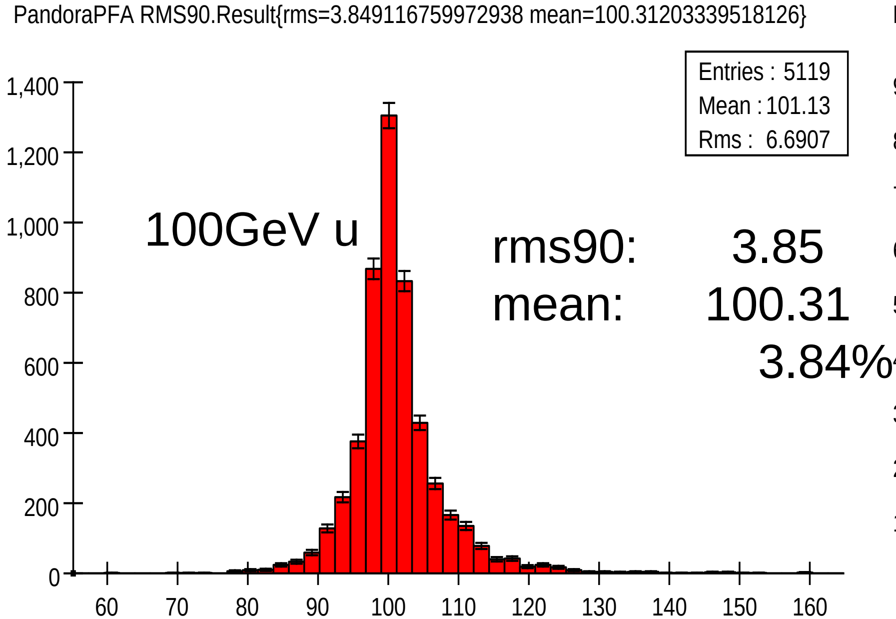
Residuals vs E

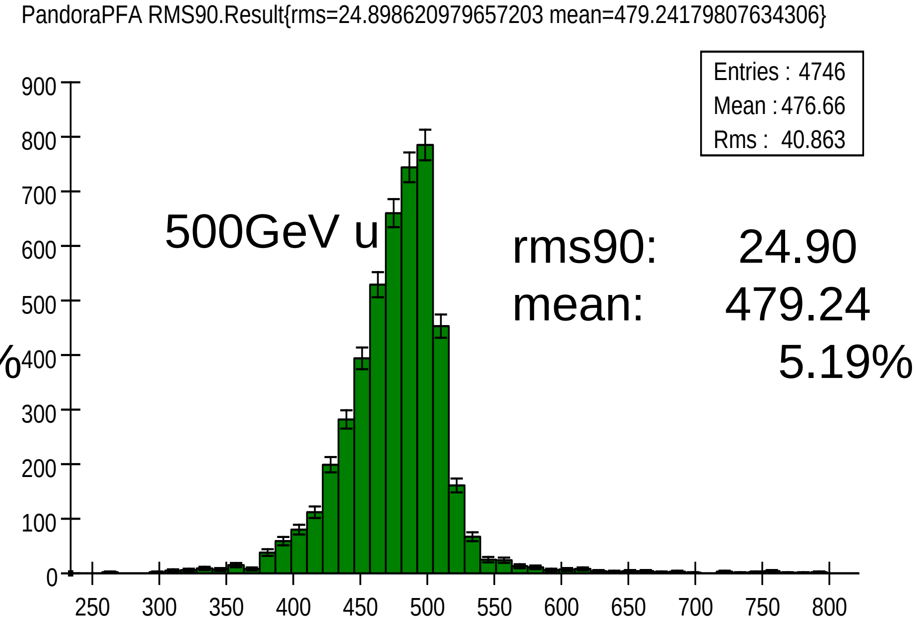


single u quark events sidloi3 slicPandora



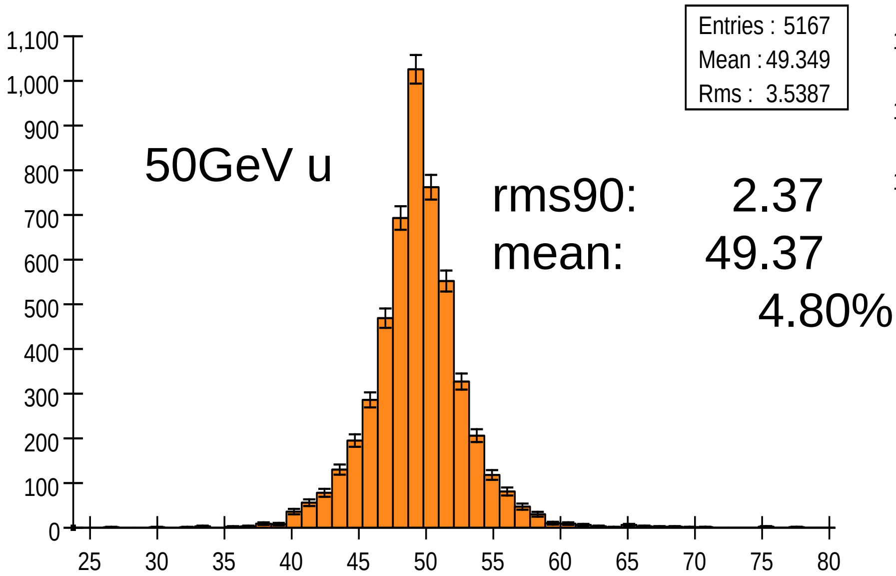




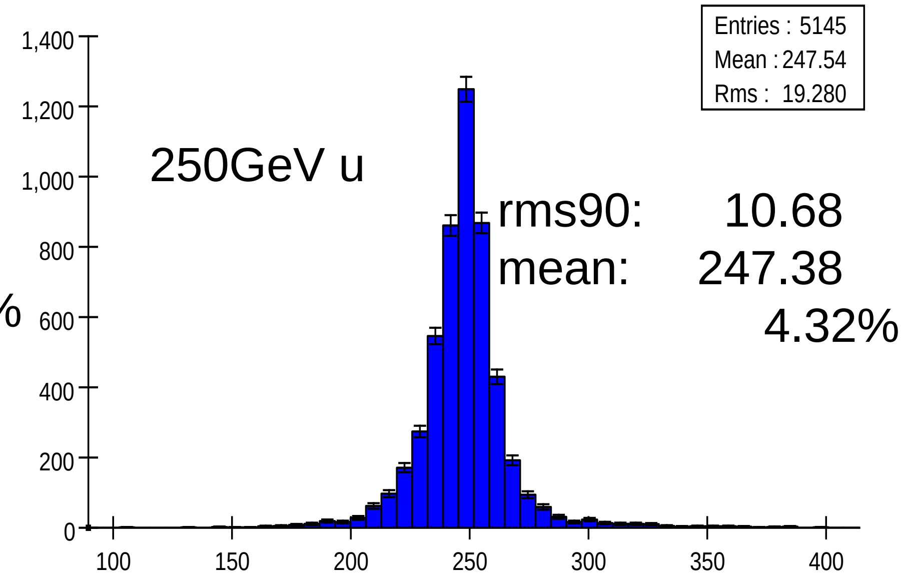


sidloi3 uipfa

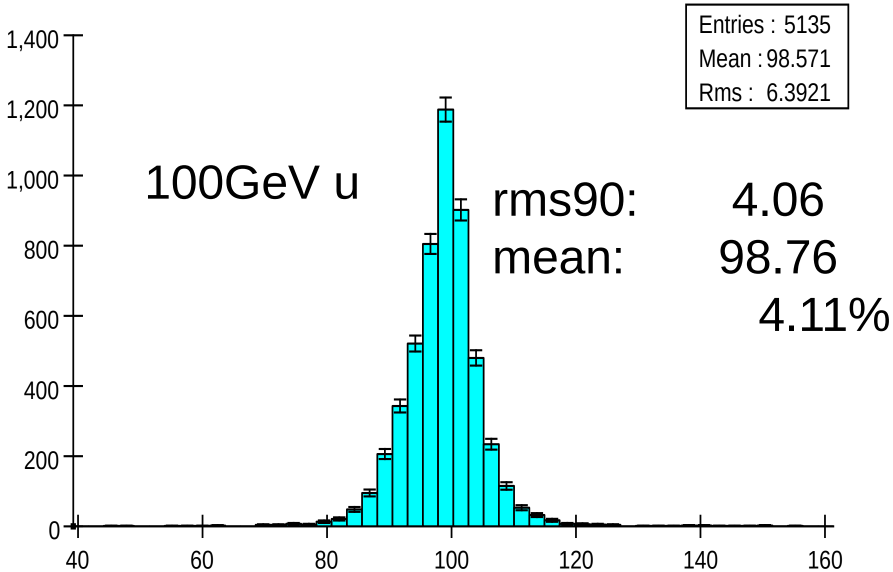
UIPFA RMS90.Result{rms=2.368183716488864 mean=49.36766548369459}



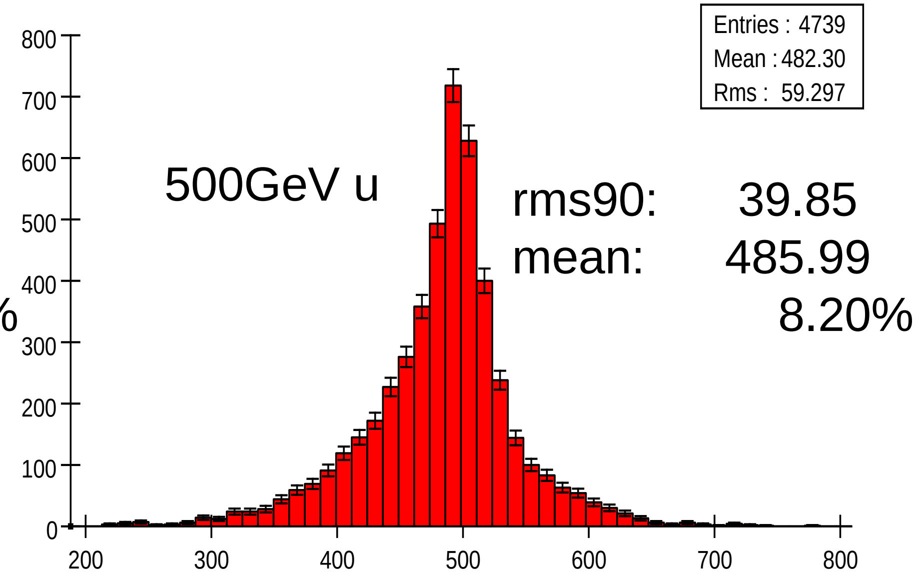
UIPFA RMS90.Result{rms=10.682481998693046 mean=247.3794150439412}



UIPFA RMS90.Result{rms=4.059309752969282 mean=98.76107533223772}

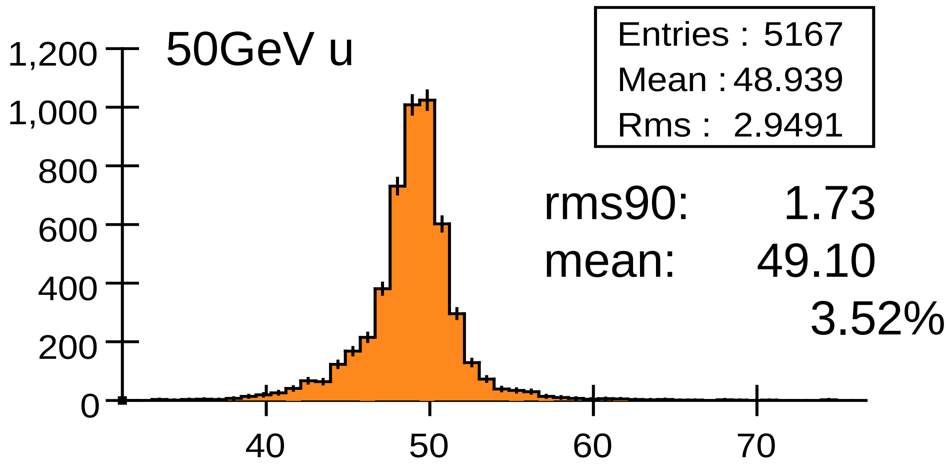


UIPFA RMS90.Result{rms=39.85497817941252 mean=485.98749022510214}

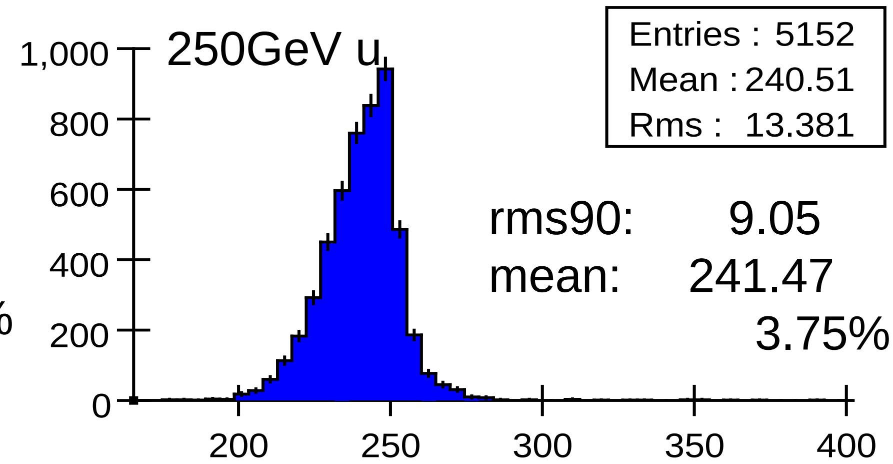


clic_sid_cdr pandora

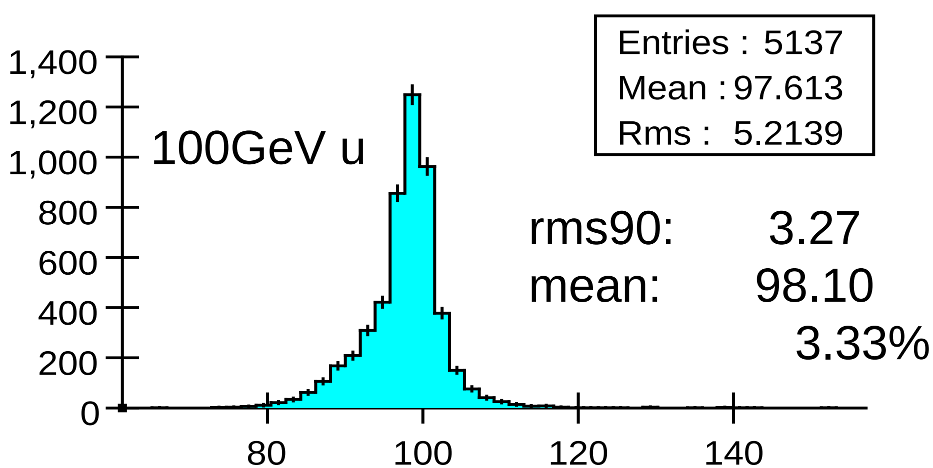
CLiC PandoraPFA RMS90.Result{rms=1.73406239...



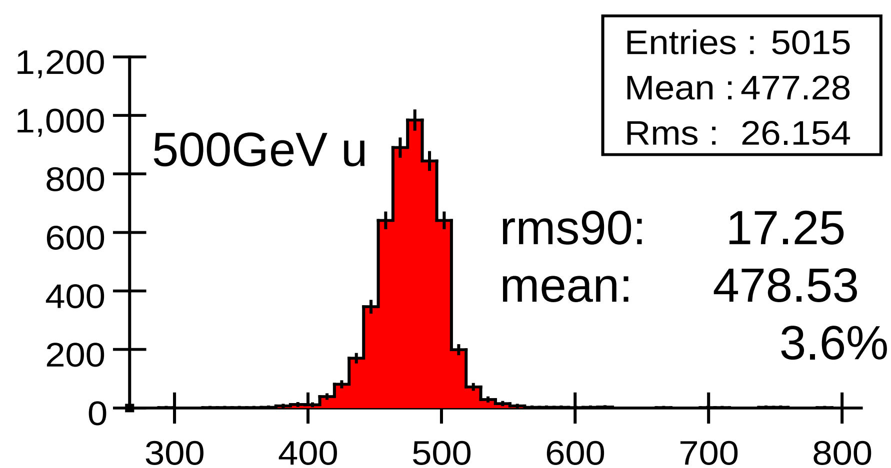
CLiC PandoraPFA RMS90.Result{rms=9.04880042...



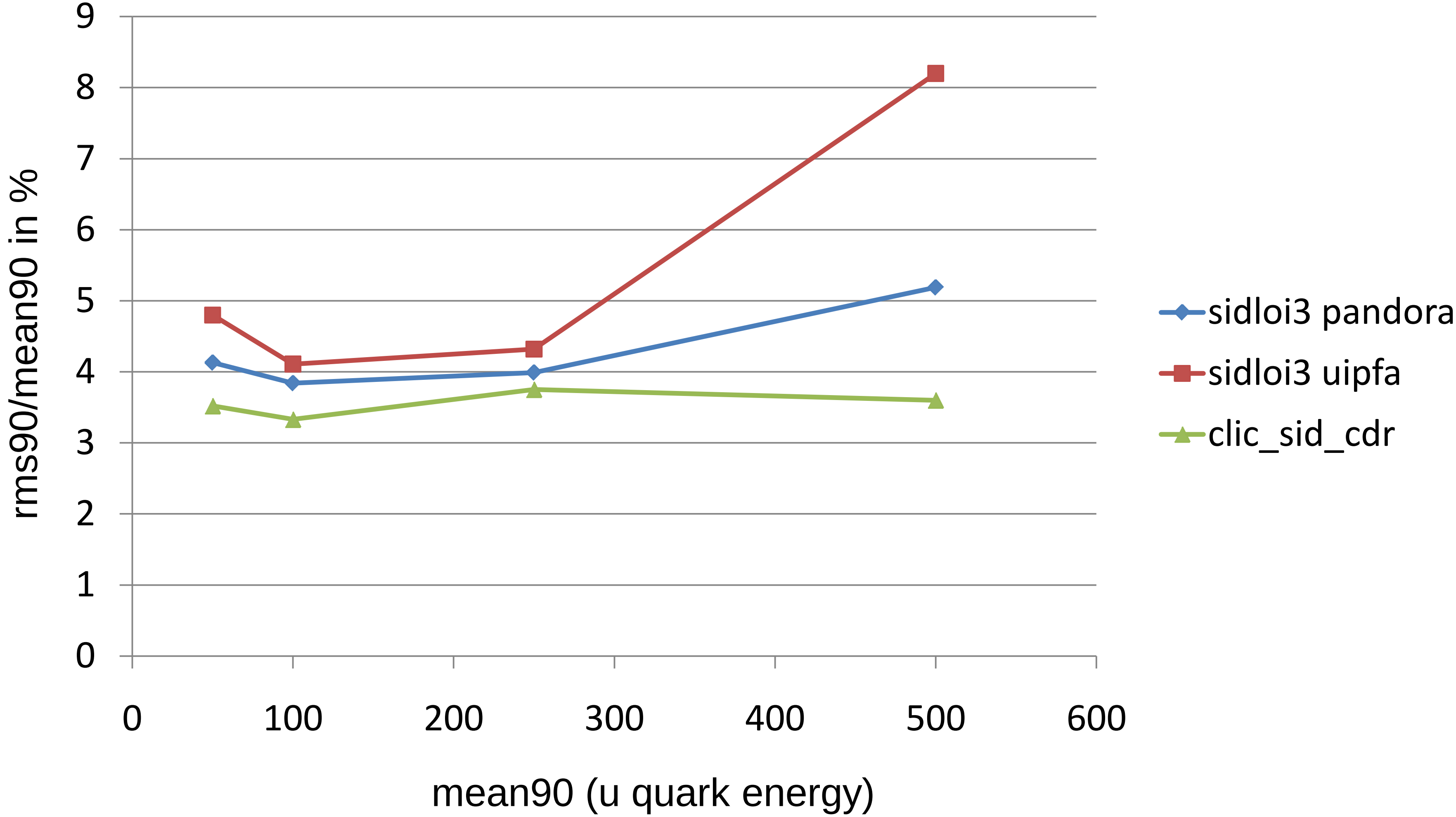
CLiC PandoraPFA RMS90.Result{rms=3.26997730...



CLiC PandoraPFA RMS90.Result{rms=17.2546253...



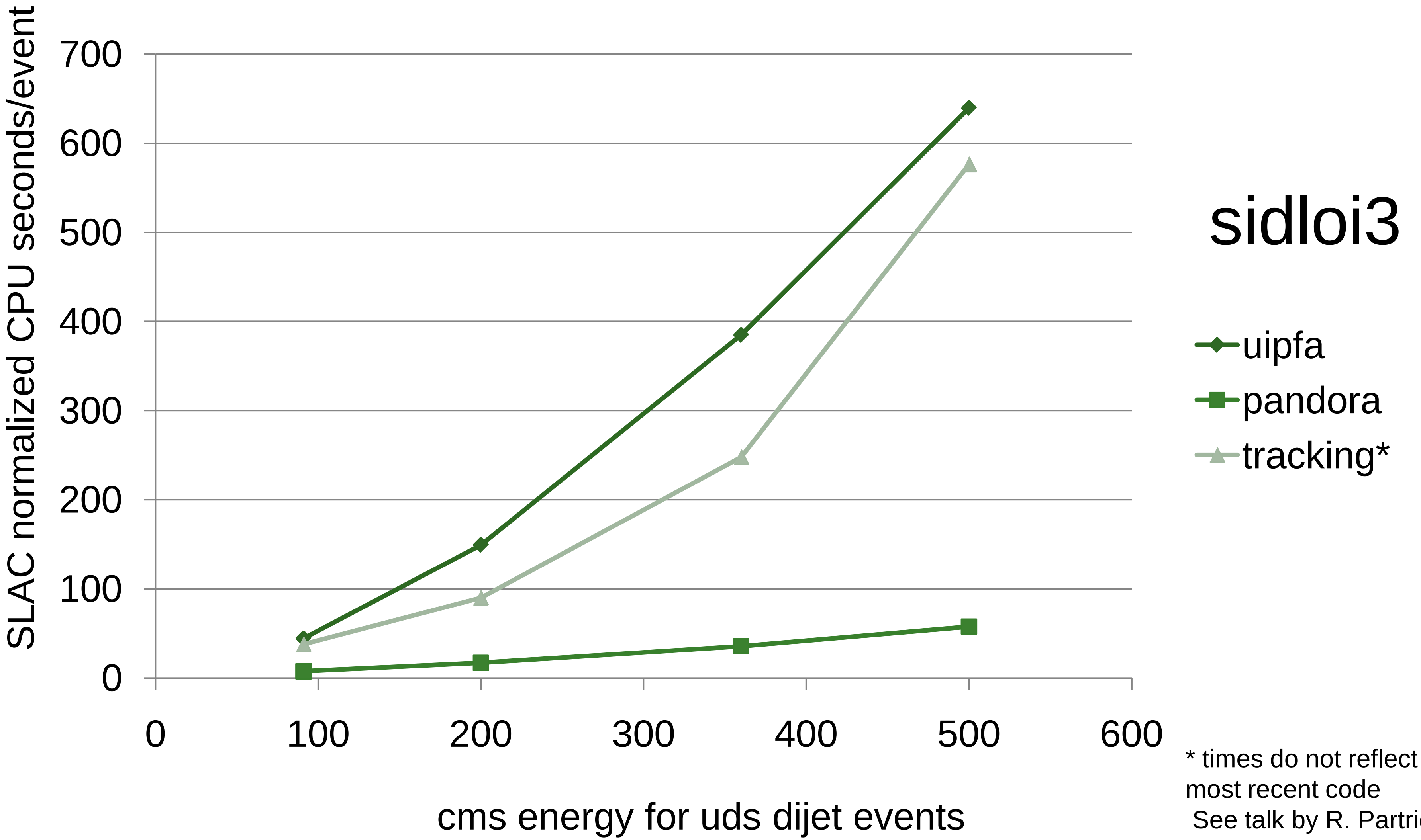
single u quark rms90/mean90 in %



Dijet Performance

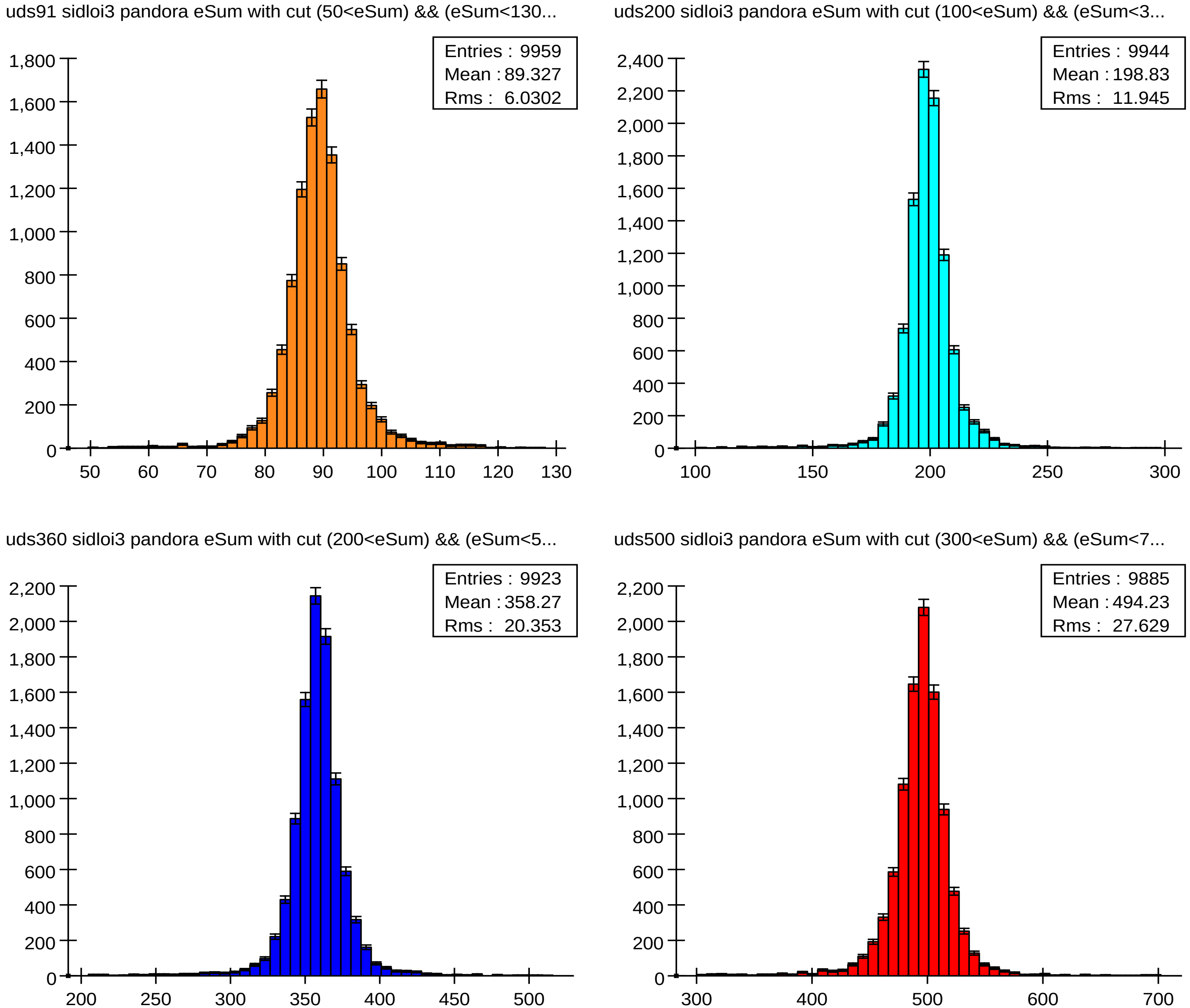
- light quark dijet events at fixed center-of-mass energies
- uds91, uds200, uds360, uds500
- Plot event energy, simple sum of energies of ReconstructedParticles

uds dijet Reconstruction Times

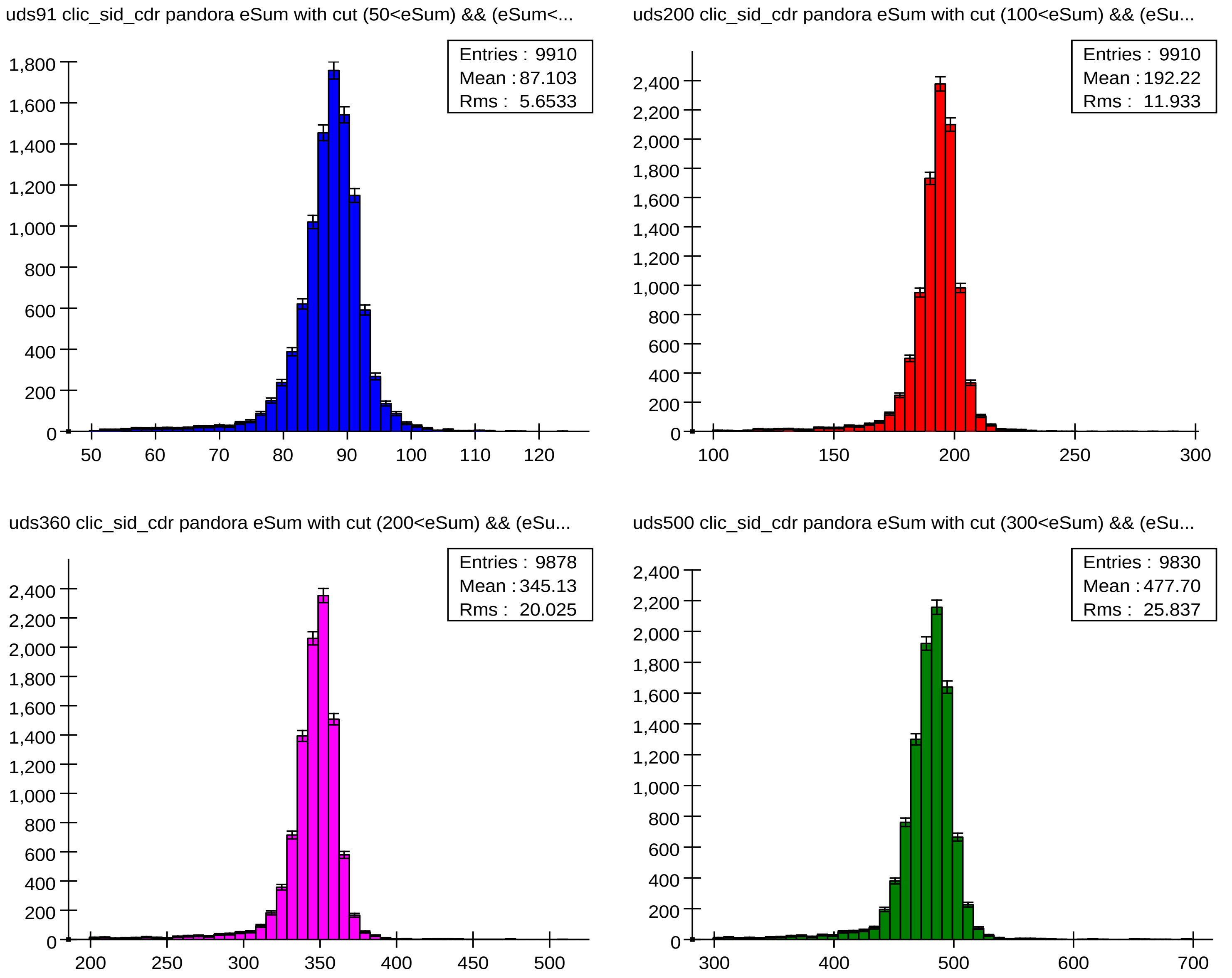


* times do not reflect most recent code
See talk by R. Partridge

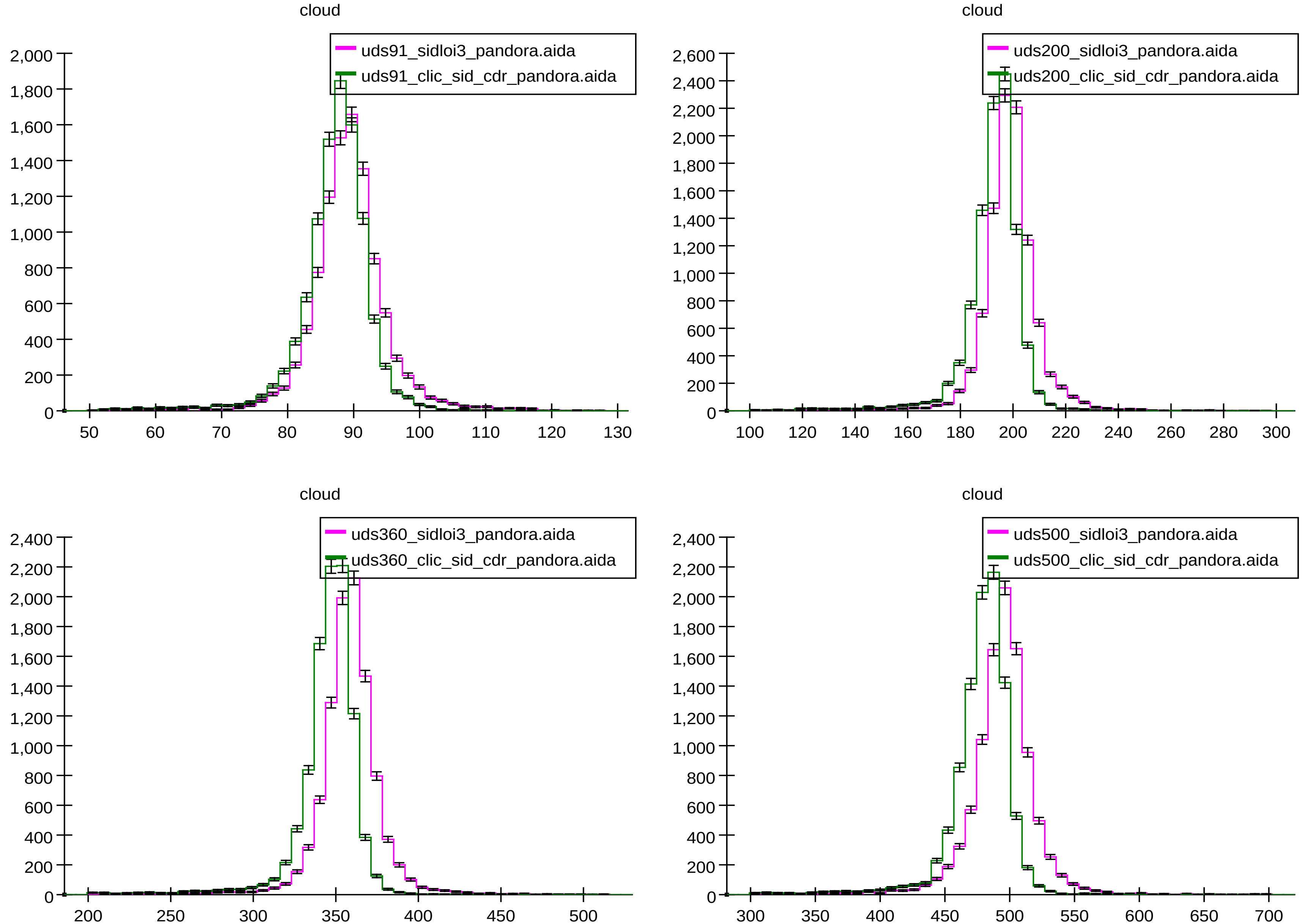
sidloi3 slicPandora event energy sum



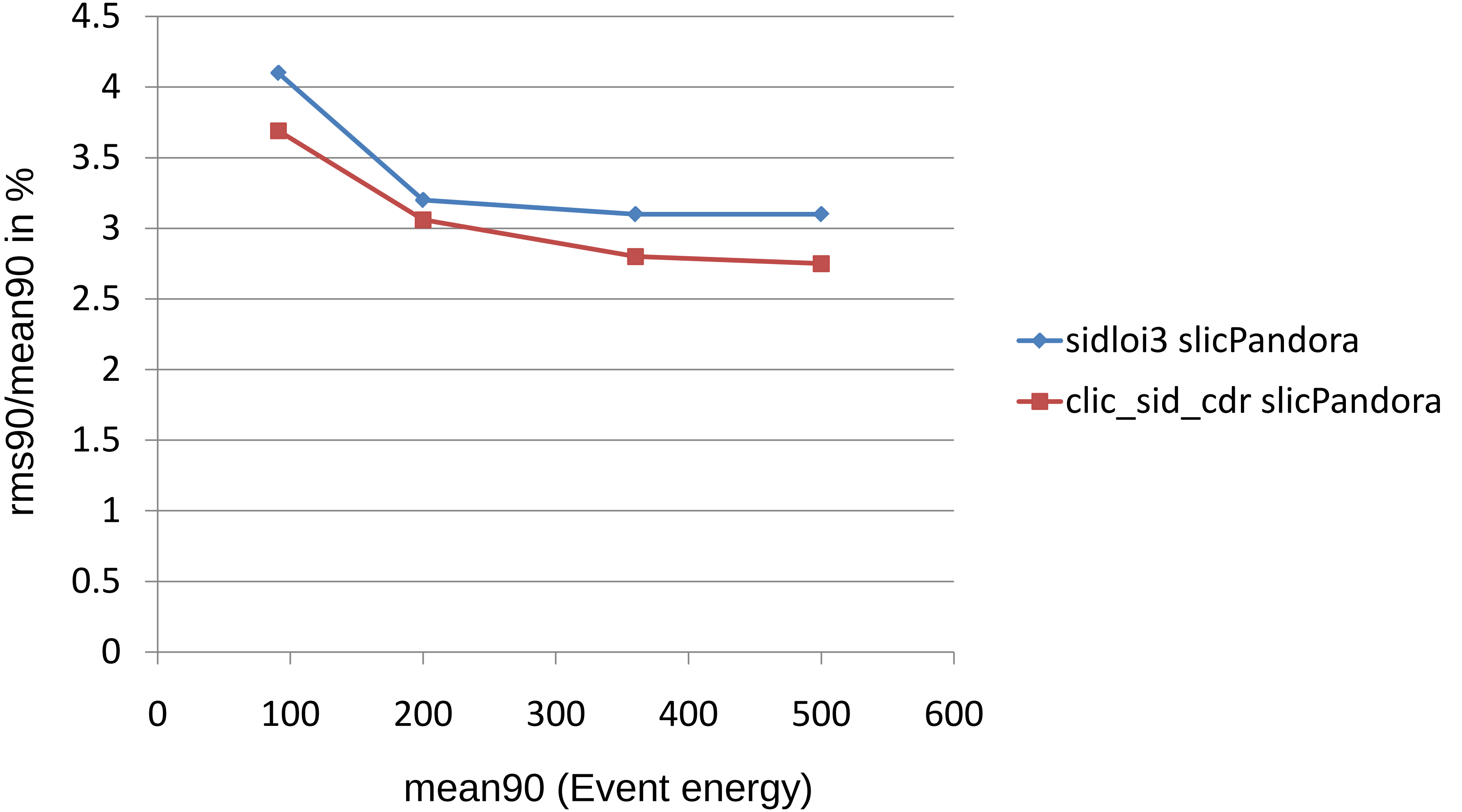
clic_sid_cdr slicPandora event energy sum



sidloi3 & clic_sid_cdr event energy sum



uds Event Energy rms90/mean90 in %



DBD Physics Benchmarking Preparations

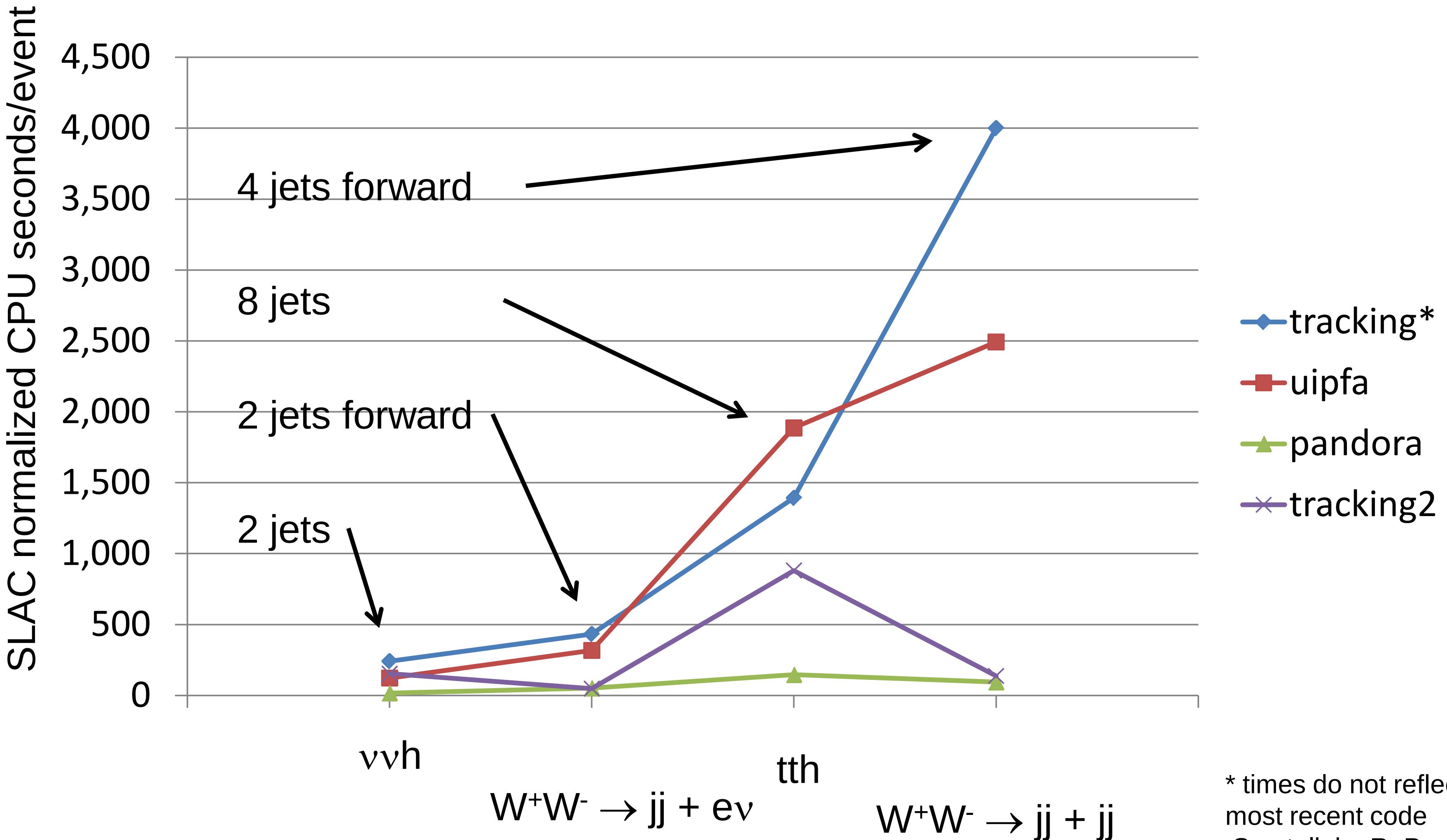
- Final details of the common event sample for the DBD Physics Benchmarking exercise are still being resolved.
- Interim solution was to generate a small sample (~10k events) of each of the signal processes
  - obtain data about processing times to inform planning for production
  - find bottlenecks in the sim→reco→analysis chain.
  - develop analysis strategies
  - optimize the detector?

pre-DBD Events @ 1 TeV

- Following Event Samples available
  - w33001 n1 n1 H $\nu\nu h$
  - w33002 n1 n1 H $\nu\nu h$
  - w33005 t T H tth
  - w33006 t T H tth
  - w33129 u D e1 N1 $W^+W^- \rightarrow jj + e\nu$
  - w33130 u D e1 N1 $W^+W^- \rightarrow jj + e\nu$
  - w33133 u D d U $W^+W^- \rightarrow jj + jj$
  - w33134 u D d U $W^+W^- \rightarrow jj + jj$

stdhep files generated by T. Barklow

1TeV DBD Reconstruction Times



* times do not reflect most recent code
See talk by R. Partridge

pre-DBD Events @ 1 TeV

- Events processed through sidloi3 detector model.
- lcsim track reconstruction (including latest fix to endcap hits) has been run.
- SlicPandora has been run.
- Collections of ReconstructedParticles are available for analysis.
- Events have also been processed through clic_sid_cdr as an *ad hoc* detector optimization exercise (and perhaps inclusion in CDR/DBD?).
  - Using slac batch resources (N.Graf)
  - Will use Grid (Dirac?) for full production.

Common Data Samples: Changes Since the LOI

- Distribute Event Generation between KEK, DESY and SLAC
- Include initial state particles and final state polarization and color flow in event record
- Improved data base for event generation information
- Include amplitudes with CKM-suppressed vertices in event generation
- Use particle aliasing to reduce the number of distinct WHIZARD processes (let the WHIZARD program do the flavor sums)

4-Fermion Production

- With aliasing, number of processes changed from 45 without CKM-suppressed final states to about 18 including CKM-suppressed final states
- Mikael Beggren has created a script that can do the MC integration, generate all 4-fermion events with a specified lumi, fill a status file, copy stdhep files to grid and output other info (.log, .in, .out, .prc.,… files) to a web directory.
- Mikael recently tested the script by generating all 4-fermion processes excluding those with final state electrons and produced 1 ab-1 equiv overnight.

6–Fermion Production

- Aliasing has allowed us to consolidate the processes
- Compilation and MC integration take much longer than before because of the CKM–suppressed vertices. However, compilation and integration only has to be done once, so this ultimately should not hold us up.

8-Fermion production

- Production of tth suspended for the moment at KEK
- Problems with ordering of final state fermions in the interface between WHIZARD and PYTHIA (not an uncommon problem)
- Still looking for a solution to the problem of generating 8-fermion and 10-fermion backgrounds to ttH

25

Event Generation

- Hope to get full audit of total number of events needed for the physics analyses:
  - Signal events for 1 ab^{-1} sample (all four polarizations)
  - Additional signal events for algorithm development, statistics smoothing, etc.
  - Physics backgrounds
  - Machine backgrounds
    - GuineaPig pairs,
    - $\gamma\gamma\rightarrow$hadrons,
    - upstream muons, etc.

Next Steps

- Settle on detector design to be used in the DBD simulation.
  - at minimum, update sidloi3 to “latest” detector design
  - Optimize detector? If so, what parameters? What metrics for improved performance? Who?
- Simulate response of sid_dbd to DBD input events.
  - How many, and of what type?
- Overlay backgrounds.
- Run reconstruction and prepare for analysis.

Summary I

- A number of updates to the track reconstruction have occurred recently. Track finding time has improved, in some cases significantly, and fake rates, especially in the forward region, have been reduced.
  - See talks by R. Partridge
- Updates to uipfa are ongoing. Expect to see impact of changes on physics events
  - See talks by R. Zaidan, R. Cassell, & U. Mallik

Summary II

- The technical interface between the lcsim simulation environment and the PandoraPFA reconstruction is essentially complete.
  - Geometry definition is automated
    - compact.xml → GeomConverter → detector_pandora.xml
  - Sampling fraction derivation is automated
    - lcsim-cal-calib + single particle generation
- sidloi3 and clic_sid_cdr being studied, (others?).
- Have first estimates for timing for simulation and reconstruction.
  - Once total event count is known, can predict resources needed to complete the task.

Drop-Dead Dates?

- Why wait? Why not start now?
  - Stress-tests all of the steps in the sim-reco-analysis chain.
  - Allows analyses to be refined, and perhaps iterated.
  - If we finish early, and we are “good enough”, then we are done.
  - If we finish early, and are willing and able to iterate, we have the opportunity to do so, using a “turn-the-crank” apparatus, knowing what metrics to use to optimize.