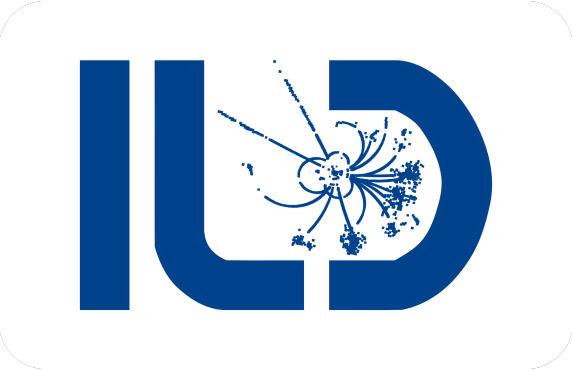
Status of the new ILD Software Simulation and Baseline Detector



the ILD software working group

Coordination	F. Gaede A. Miyamoto
Generator	M. Berggren J. Tian
Simulation	D. Jeans S. Lu
Reconstruction	L. Tran
MC production	A. Miyamoto H. Ono

- introduction
- generator
- detector simulation
- reconstruction
- grid production
- outlook

Large-scale production of ILD simulation samples last done for Detector Baseline Document studies ~ 5 years ago.

now planning to produce new large set of simulated events
full set of S.M. samples

- detector optimisation
  in particular a comparison of “large” and “small” detector models

- physics studies with latest version of reconstruction tools

There have been significant updates to software tools in the last five years, which we plan to make use of

- latest versions of WHIZARD event generator (WHIZARD 2.x)

- dd4hep for detector description

- new reconstruction tools

- ilcdirac system for accessing grid computing resources

generator

largely within LC generator group:

see Junping's presentation earlier in this workshop for full status

WHIZARD is the workhorse of our event generation

enjoy close contact with authors

regular face-to-face meetings: March'17, Dec'16, ...

plan to upgrade from WHIZARD 1.96 (DBD) → WHIZARD 2.5

significant updates to central algorithms

extra features implemented : e.g. lcio output

most aspects look fine, and ready for production

however a few outstanding issues remain, in particular

colour flow in 4f events: interface to Pythia

ISR treatment: getting both reasonable p_T and cross-sections

these **must** be addressed before proceeding with large production

alternative tools

lower-priority investigations starting into

e.g. BHWIDE, BDK/BDKRC, Pythia8, Madgraph

for comparisons and cross-checks

detector simulation

Sub-detector contacts	
Si tracking	M. Vos
TPC	D. Tsionou
Calorimeters	D. Jeans
Forward calos.	B. Pavlik
Services, MDI	K. Buesser
Muons	N. D'Acenzo

Detector simulation framework changed from
Mokka (used for DBD) to the relatively new
dd4hep framework

dd4hep developed within AIDA framework

our DBD baseline detector model (“ILD_o1_v05”) has been ported to this new framework

new / updated detector models being developed in dd4hep
- reflect changes in detector design
- different models for optimisation

large-scale production of full samples will be performed with 2 models
others can be used for smaller-scale comparisons

Large event samples will be simulated in two detector models

"Large" – rather close in size to the DBD detector "ILD_l4_v02"

"Small" – reduce outer radius of TPC by ~30 cm "ILD_s4_v02"
bring in exterior detectors (lower cost)
increase B-field from 3.5 → 4.0 T (recover lost performance)

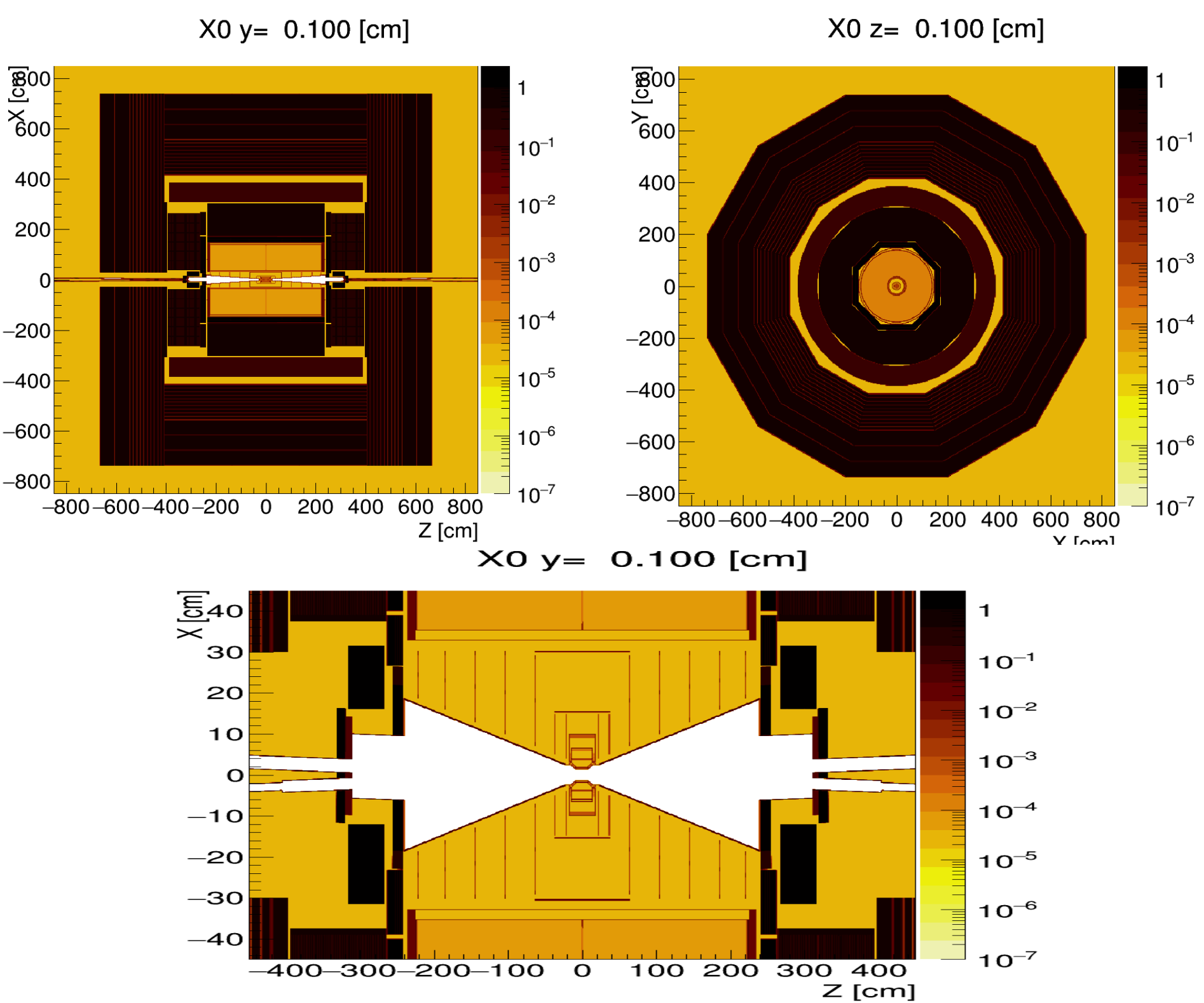
most sub-detectors are largely unchanged from the DBD, except:

– updated design of forward calorimeters (new L*)

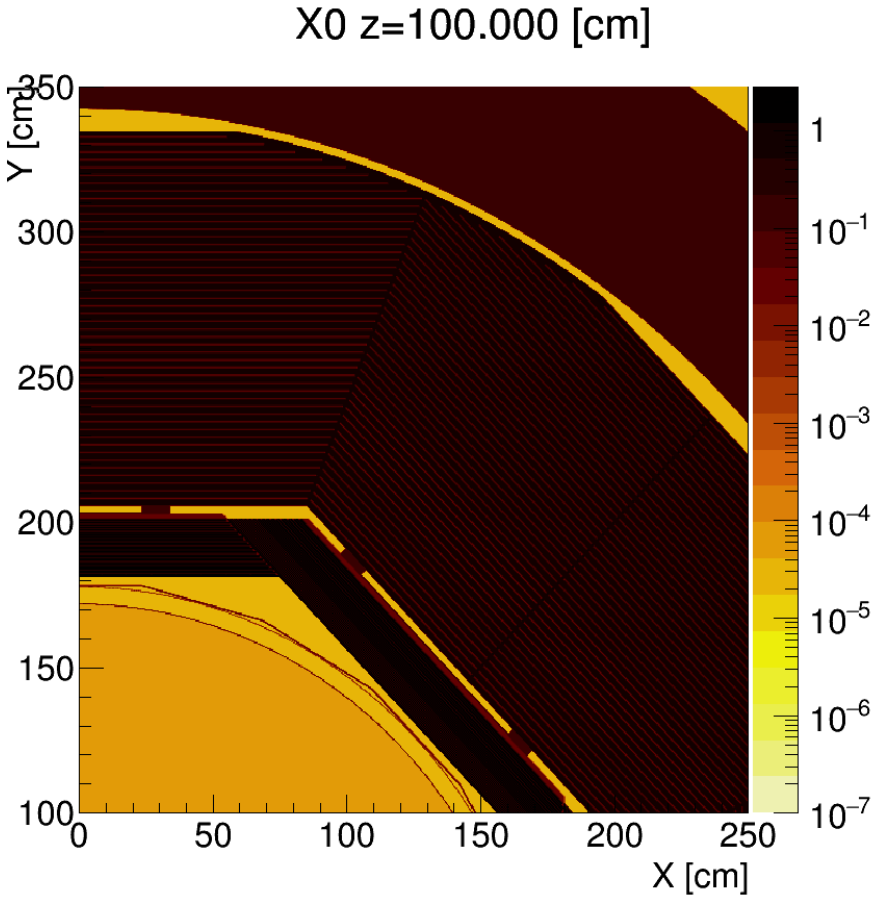
– less technically aggressive ECAL design
→ ~3cm extra thickness,
recovered by reducing uninstrumented TPC volumes

– "hybrid" HCAL simulation model
simultaneously simulate hits in both scintillator tile and RPC layers
implemented in "Tesla" geometry

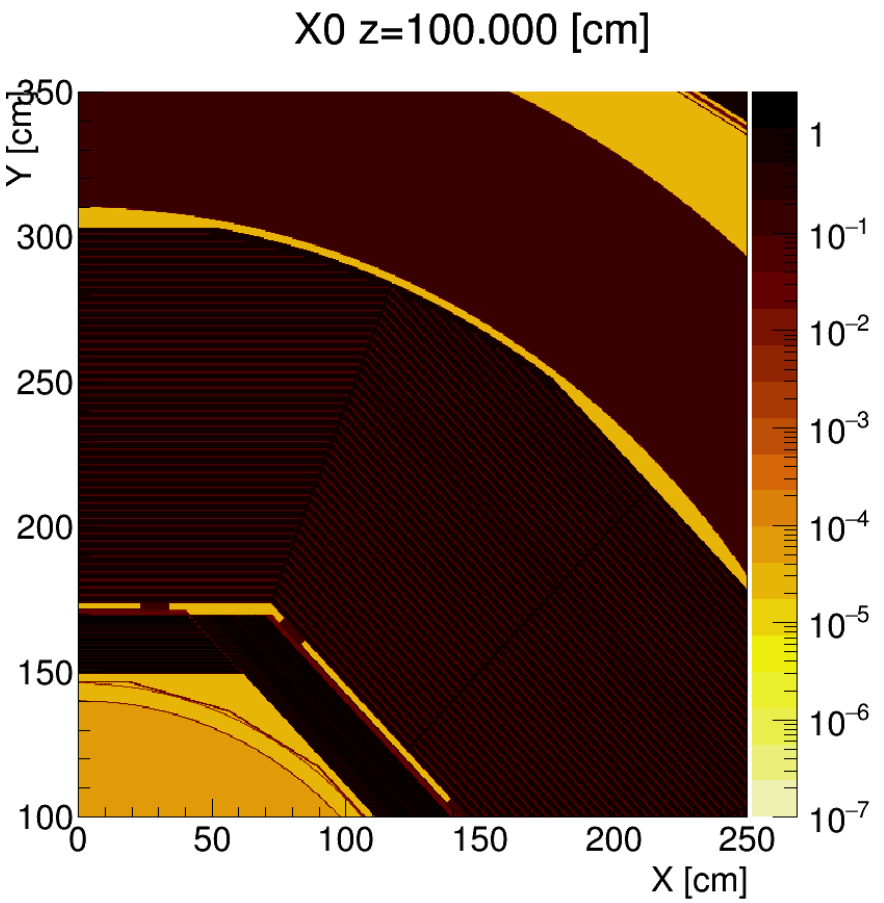
simulation descriptions of these two models have been prepared


X0 y= 0.100 [cm]
X [cm]
Z [cm]
X0 z= 0.100 [cm]
Y [cm]
X0 y= 0.100 [cm]
X [cm]
Z [cm]

large
ILD_l4_v02

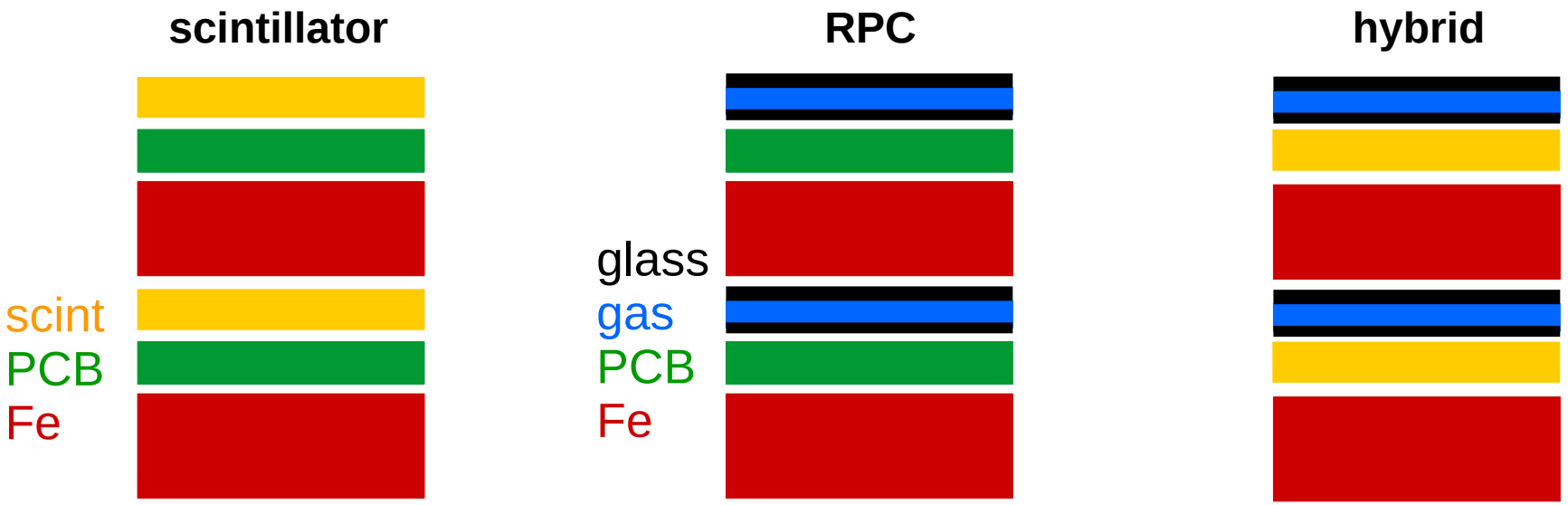


small
ILD_s4_v02



thicker ECAL, reduced TPC, at different radii

combined technology simulation (simplified)



"combined" or "hybrid" model: not a proposal for a real hybrid calorimeter,
rather a time- and effort-saving software trick

simulates both gas gap and scintillator: hits stored in separate collections

allows comparison of both technologies in single simulation
→ saves time and effort of simulation group
→ allows event-by-event, even shower-by-shower comparisons

if material within gap is not too different from the individual models,
should give good description

these simulation models are close to final,
but a few aspects need to be confirmed by subdetector groups

– inner tracking detectors
path of VTX cables
move to pixel detectors in SIT ? with single BX tagging ?

– TPC : from where to take the 3cm eaten by ECAL
regions between sensitive volume and inner & outer field cages ?

– ECAL : need final details of updated dimensions

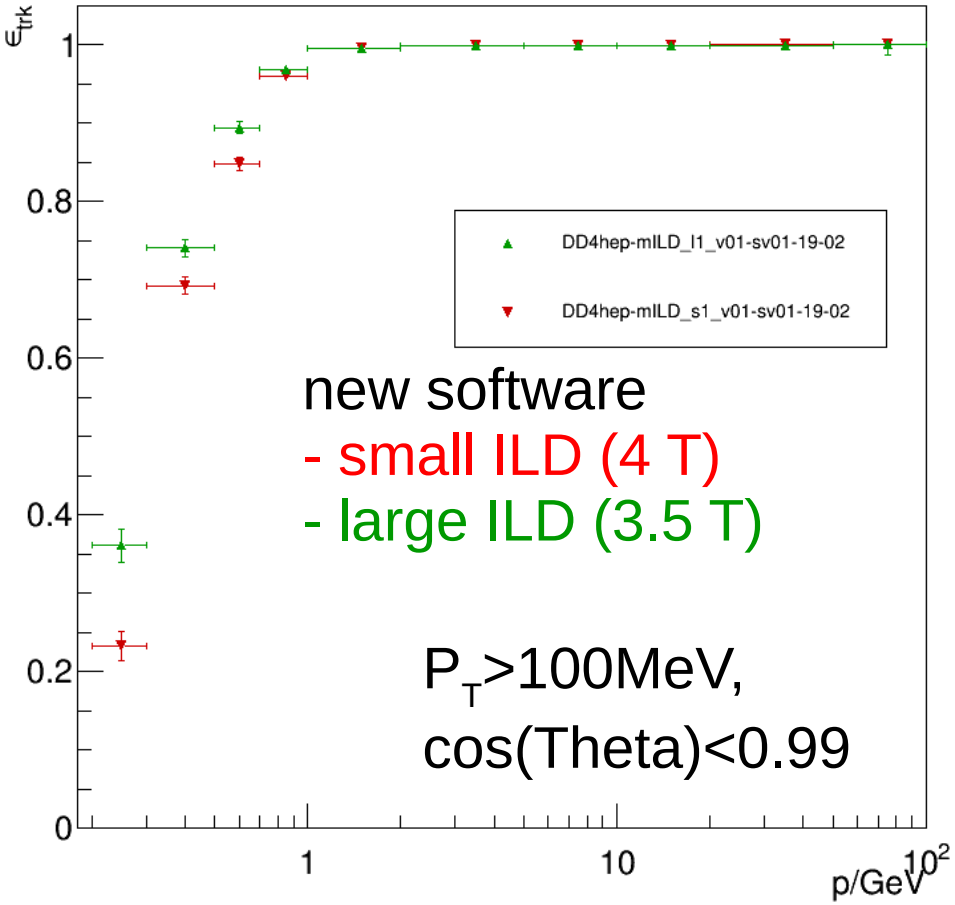
– HCAL : GEANT4 version and physics list to use
(should be >= 10.2.p02)

to progress efficiently, we would like to have answers a.s.a.p.
preferably today

reconstruction

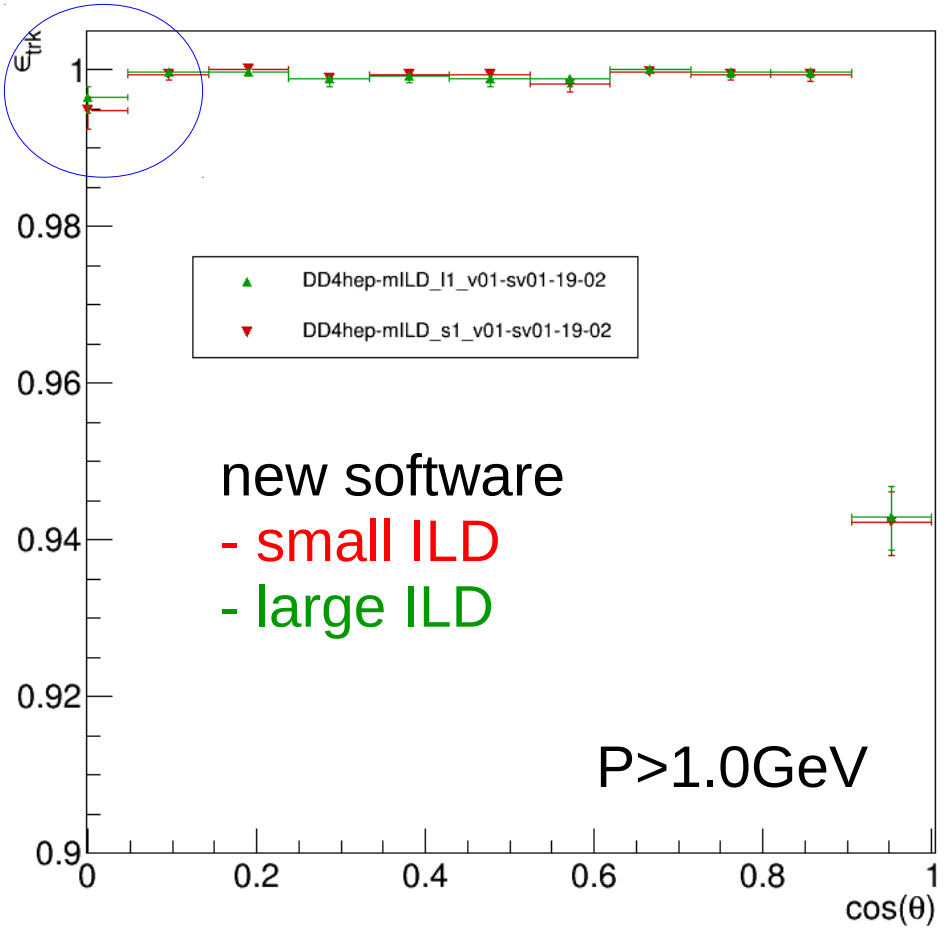
tracking efficiency in good shape

small loss at cos(theta)~0
to be understood



new software
- small ILD (4 T)
- large ILD (3.5 T)

P_T>100MeV,
cos(Theta)<0.99



new software
- small ILD
- large ILD

P>1.0GeV

- hit purity>90%, >= 4 SiHits, distance from IP < 10mm.

Jet energy resolution with PandoraPFA

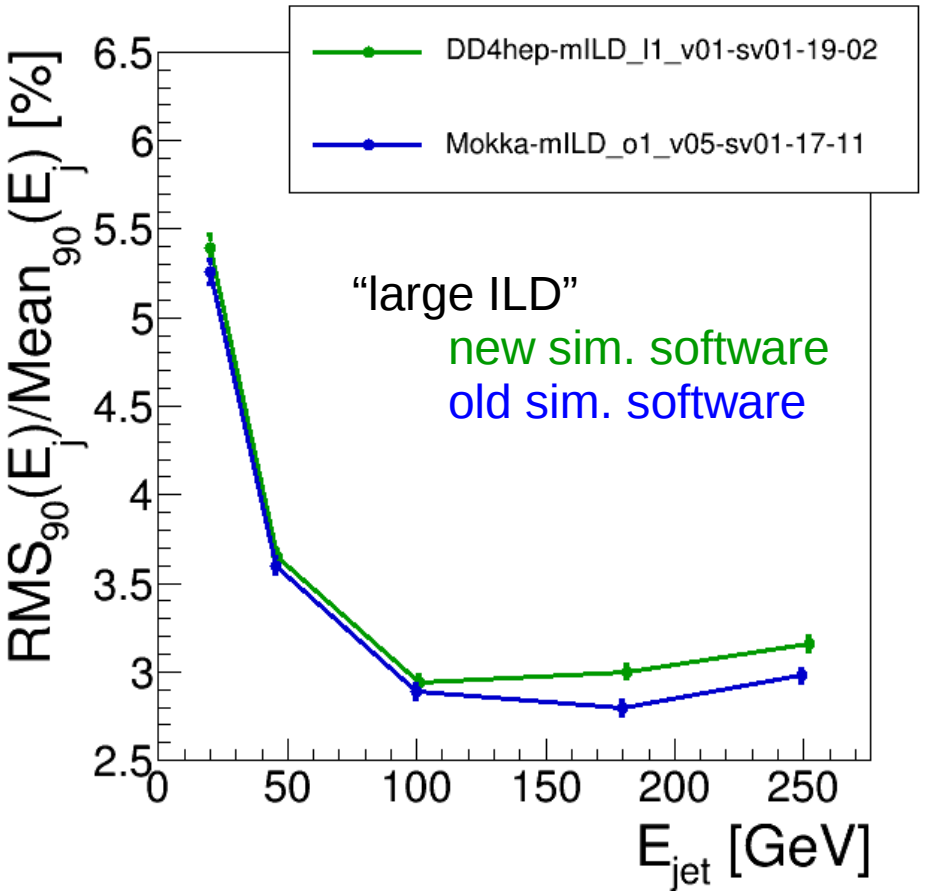
- a lot of work over last months to understand JER :

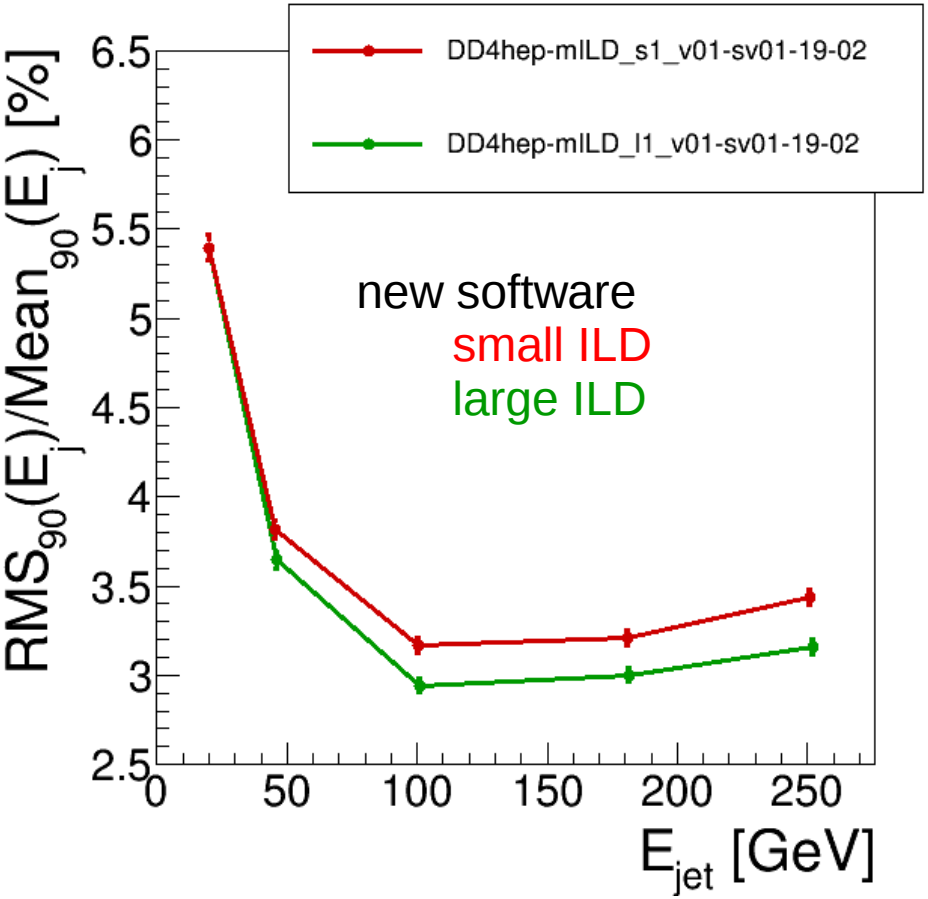
S. Green, S. Lu, L. Tran

- results now solid and understood

- new models still to be re-calibrated

di-quark (uds)
cos theta_thrust < 0.7
correct for neutrinos





MCProduction overview

- Role of the WG is
  - Take generator inputs from /ilc/prod/ilc/mc-dbd/generated
  - Produce DSTMerged files in /ilc/prod/ilc/mc-dbd/ild/dst-merged, keeping logs.

- Task of the WG are,

  1) Develop and maintain tools to perform the role.
     - ILCDirac production tool developed by CLICdp has been selected as a bases. We can share computing and human resources and avoid duplicate works.
     - Still the developments of ILD scripts are necessary to meet ILD needs

  2) Do actual production; Define input files, submit production, check results, collect logs, update web, etc.

- Production workflow.
  - 4 main steps : (1) Split stdhep files, (2) Simulation production, (3) Reconstruction, (4) DSTmerge and collect logs.
  - ILCDirac production are used for (2), (3) and partially (1). With this tool, the production task is automatic thanks to ILCDirac.
  - Other steps have been performed manually (interactively) and most real time consuming. Not yet integrated in ILCDirac workflow.

Recent activities

- 2 scripts to initiate production: DBD(old) and Opt2017(new) were developed.
  - DBD: 3 productions(StdhepSplit&Sim&Rec) to produce 1 process_ID
    - Convenient if many stdhep files for 1 process_ID
  - Opt2017: 2 productions(Sim&Rec) for many process_IDs in one directory.
    - Since April 2017.
    - Stdhepsplit does not work properly yet.

- MCProductions since March this year : Done with DBD Mokka/Marlin
  - See detail at our web log, https://ild.ngt.ndu.ac.jp/elog/dbd-prod/
  - 250GeV:
    - 2f_highM_z : 2 procs, 13M evts (contain <0.01% of erroneous events)
    - H mumu : 4 procs, 80k evts
  - 500GeV
    - e+e-Nγ, vvNγ: :6 procs: ~12M evt
    - 4f_ww_sl, 4f_ww_l, 4f_zz_sl : 6 procs, 427k evts
    - aa_4f, 5f (leptonic) : 50 procs, 446k evt
    - Hmumu : 4 procs, 16k
  - Lesson: Production of small number of events is not efficient.

- Plan of MCProduction with DBD soft
  - Produce unprocessed events in existing stdhep files (250 & 500 GeV), aiming minimize a request of small samples.
  - Produce new samples. New stdhep file required. A general purpose samples, not for a very specific physics analysis, is preferred.

Towards the optimization production

- Calibration samples have been produced to help code validation.
  - uds- and single-particle-events are produced when ILCSoft is released.
  - Data available now are at /ilc/prod/ilc/mc-opt/ild/[sim|rec]/calib/[uds|single]/<detector>/<ildconfig>
    - <detector> = ILD_o1_v05, ILD_l1_v01, ILD_s1_v01
      <ildconfig> = v01-19-02
  - Dirac UserJob tools are used, because file name convention does not follow the convention used for the production of physics events.

- Optimization production
  - Existing scripts will be modified to meet new needs. Items to be tested includes
    - generator files in slcio files.
    - DD4Sim and Marlin in ILD production framework.
    - Background overlay: not only aa_lowpt but also GuineaPig.
      - Make sure they work well in GRID environment.
  - Following detail information in advance are useful for us.
    - Gen samples : Only physics channel ? statistics ? Exact file location ?
    - Background processes: how much ?

    A test production to estimate CPU, storage, time needs will be necessary

outlook and summary

to get any meaningful results by the end of this year,
we need to set milestones

generator

2 show-stopping issues (ISR, colour flow)
to be fixed in WHIZARD2
→ in hands of WHIZARD authors

fall-back : generator samples from DBD era

simulation

– inner tracking detectors
  path of VTX cables
  move to pixel detectors in SIT ? with single BX tagging ?

– TPC : from where to take the 3cm eaten by ECAL
  regions between sensitive volume and inner & outer field cages ?

– ECAL : need final details of updated dimensions

– HCAL : GEANT4 version and physics list to use
  (should be >= 10.2.p02)

we propose to prepare a final model in the **next few days**,
  according to our best knowledge, and your feedback today

short period time for validation by sub-detector contacts
  if no complaints, proceed

freeze sub-detector models in **mid-July**
  then produce first test samples for physics groups,
    before summer break

reconstruction

central reconstruction tools basically OK
some re-tuning/calibration required

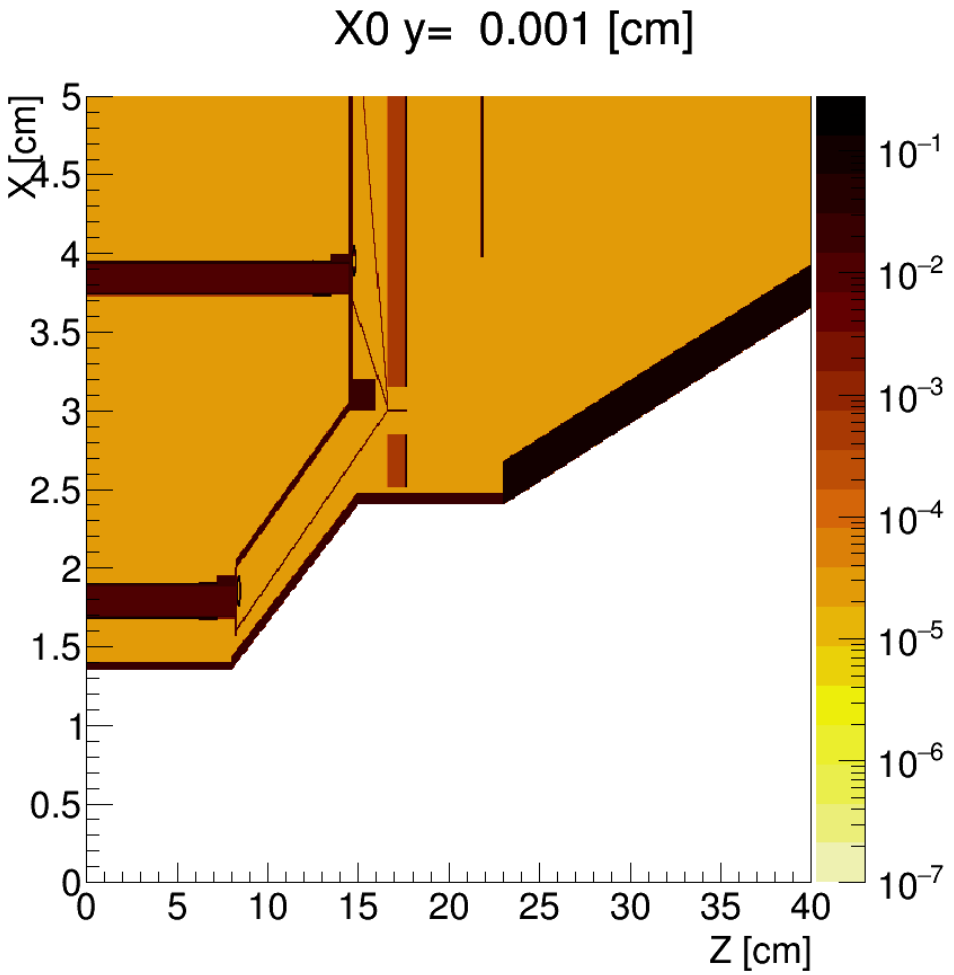
availability of SDHCAL-related tools unclear (at least to me)

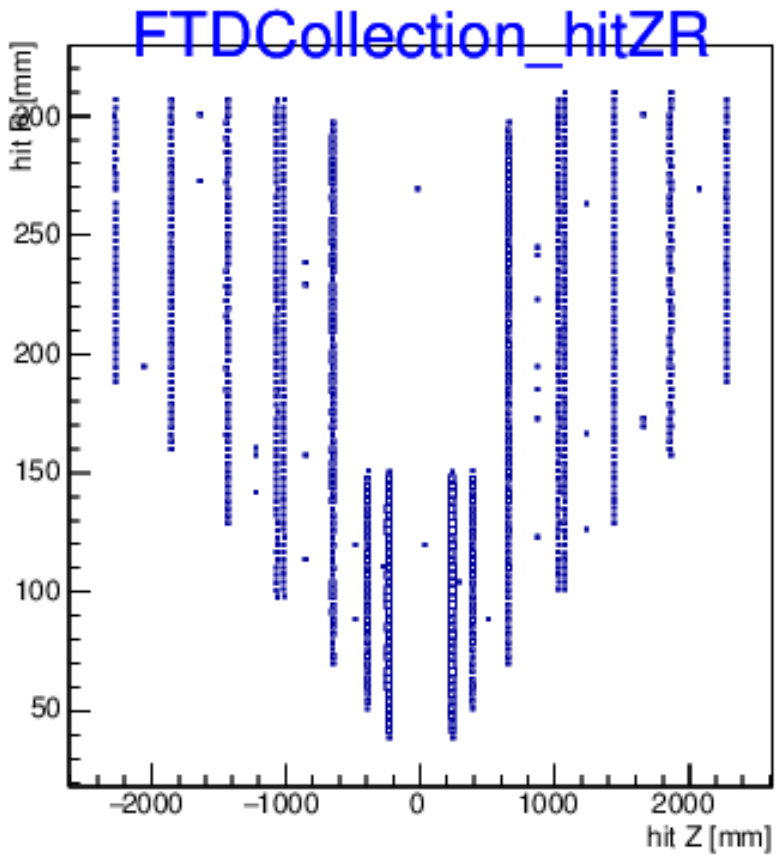
grid production

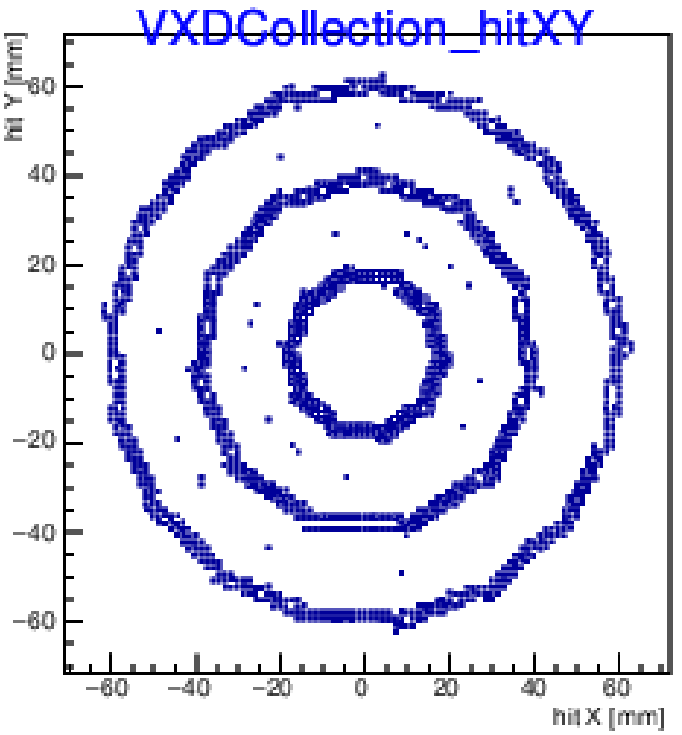
ready to stress-test in large production

end

path of vertex cables ?



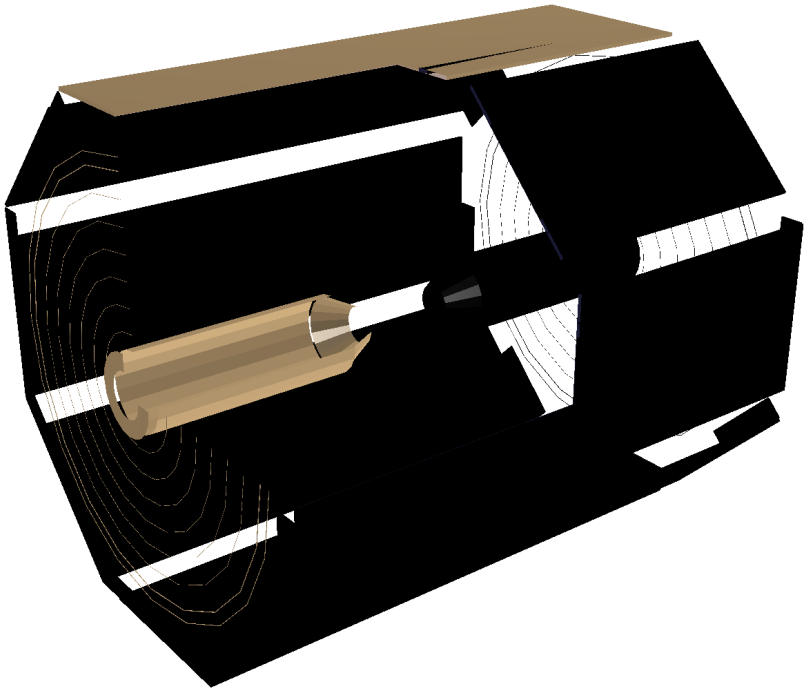




from lyon meeting:

simulated hit positions in VTX, SIT show some bug

rather rare, probably OK to remove at digitisation



services

HCAL & SIT services now added