SiD Simulation Status Report

Norman Graf (SLAC) SiD Workshop, RAL April 14,2008

Software Development Overview

- Physics analyses.
 - Refine and strengthen the arguments for the ILC.
- Detector design.
 - Integrated system design for an optimal detector.
- Event reconstruction.
 - Demonstrate that proposed detector systems can conduct the physics program, and allows cost-benefit optimization to be done.
- Testbeam and prototype infrastructure.
 - Support subsystem design, readout and analysis.

Simulation Mission Statement

- Provide full simulation capabilities for Linear Collider physics program:
 - Physics simulations
 - Detector designs
 - Reconstruction and analysis
- Need flexibility for:
 - New detector geometries/technologies
 - Different reconstruction algorithms
- Limited resources demand efficient solutions, focused effort.
 - Strong connections between university groups, national labs, international colleagues.

Goals

- Facilitate contribution from physicists in different locations with various amounts of available time.
- Use standard data formats, collaborate & cooperate.
- Provide a general-purpose framework for physics software development.
- Develop a suite of reconstruction and analysis algorithms and sample codes.
- Simulate benchmark physics processes on different full detector designs.
- Software should be easy to install, learn, use.
 - Support via web based forums, tutorials, meetings.

From Physics Studies to Benchmarking

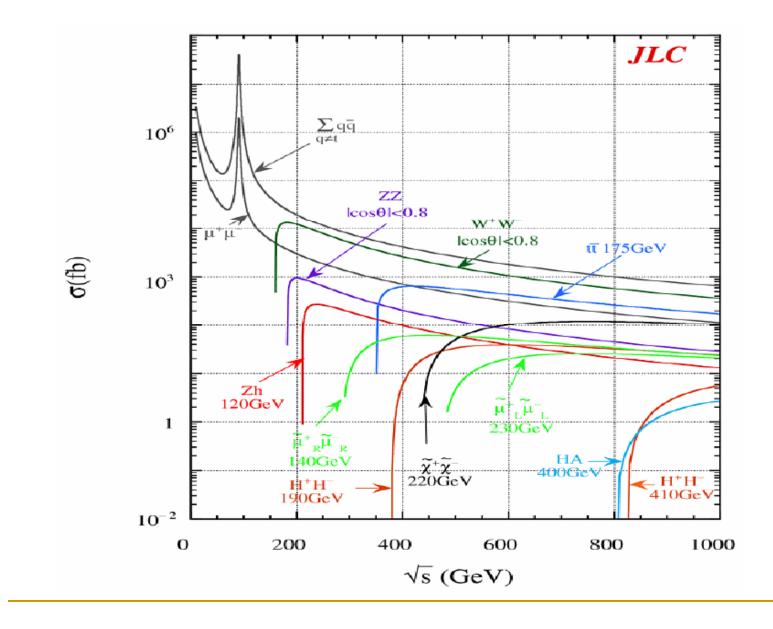
- We believe that the physics case for a TeVscale linear e⁺ e⁻ collider has been made.
- The emphasis of analyses now shifts towards
 - Optimization, evaluation and comparison of detector choices
 - Realities required by engineering: e.g. amount and distribution of readout and support material, ...
 - Realities required by realistic detector response simulations: e.g. electronics digitization, noise, ...
 - Realities required by reconstruction algorithms: e.g. track finding & fitting, PFA, jetfinding, ...

Motivation for Common Benchmarks

- Detector concepts will naturally seek to optimize their designs using physics processes.
- The wider community would like to see demonstrated physics capabilities from mature detectors, e.g. with a reasonable level of engineering, costing, etc.
- Natural, then, especially with limited resources, to agree upon a common set of analyses to be used by all the concepts in the LOI process.

Standard Model Backgrounds

- All analyses used in the context of the detector optimization and LOI process will need an inclusive sample of the Standard Model Background.
- Will provide the SM sample centrally in stdhep format, for all concepts to use.
- Provide all information and tools necessary to produce specific signal samples individually with exactly the same setup as this SM sample.



Whizard SM Sample

- Generate an inclusive set of MC events with all SM processes
- WHIZARD Monte Carlo used to generate all 0,2,4,6-fermion and t quark dominated 8fermion processes.
- 100% e⁻ and e⁺ polarization used in generation. Arbitrary electron, positron polarization simulated by properly combining data sets.
- Fully fragmented MC data sets are produced.
 PYTHIA is used for final state QED & QCD parton showering, fragmentation, particle decay.
- Events are weighted!

Standard Model Sample

- Full 2ab⁻¹ SM sample available via ftp from SLAC.
- Each file corresponds to a particular initial e-/e+ polarization and final state

ftp://ftp-lcd.slac.stanford.edu/ilc/whizdata/ILC500/

cumbersome to work with for end user

Have to mix polarizations by hand

Each file contains only processes of one type, so need to run over complete data set (thousands of files) to get faithful subset.

500 fb⁻¹ sample of these events generated with 80% e⁻, 30% e⁺ polarizations, randomly mixed events from all processes

- <u>ftp://ftp-lcd.slac.stanford.edu/ilc/ILC500/StandardModel/</u>
- Roughly 400 million events

Next Steps for SM Data Sample

- Remove 120 Higgs from n fermion final states at 500 GeV, and add explicit ffH, ffHH, etc. final states.
- Regenerate states with τ in final state using TAUOLA.
 - Coding done at DESY, generation in progress.
- Produce full SM data set at 250 GeV
 - Machine parameters have been obtained from the GDE, GuineaPig events being generated to provide pair background events and beam spectra to be used for event generation.

Additional Backgrounds

- GuineaPig pairs and photons (Cain too?)
 - Added crossing angle, converted to stdhep, available <u>here</u>.
- Muons and other backgrounds from upstream collimators & converted to stdhep.
 - Need to validate and understand normalizations.
- $\gamma\gamma \rightarrow$ hadrons generated as part of the "2ab⁻¹ SM sample."
- All events then capable of being processed through full detector simulation.
- Additive at the detector hit level, with time offsets, using LCIO utilities.
 - i.e. simulate response separately for signals and backgrounds, then add at digitization/reconstruction level.

Additional Event Generation Issues

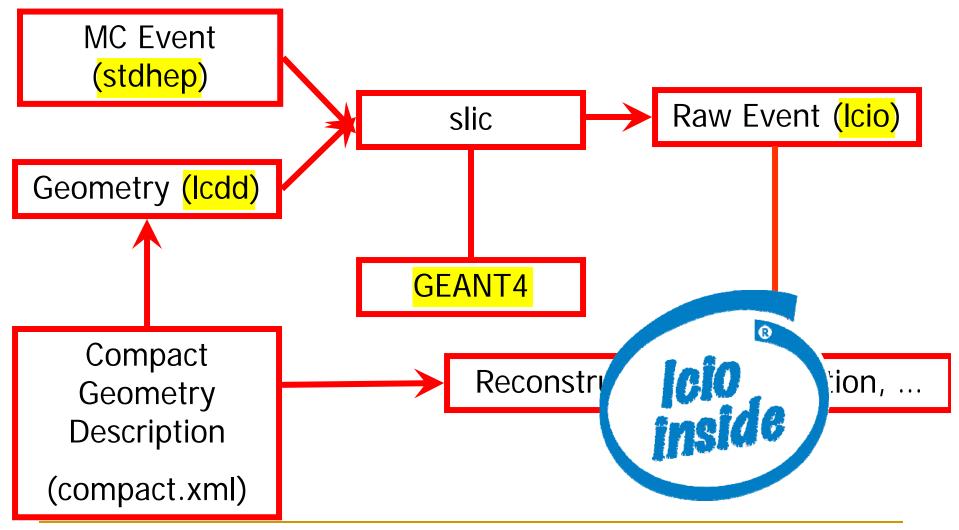
Crossing Angle

- Agree that events will be generated with 0° and the 14mr crossing angle will be accounted for at the time of simulation.
- Tuning of pythia parton fragmentation has come up.
 - Will proceed with the pythia defaults due to the large number of events already generated.
- Essential to fully document and maintain a provenance for all files (and analyses?)

Detector Modeling & Simulation

LC Detector Full Simulation

Items highlighted in yellow represent common standards



slic

- Number of internal optimizations and refactorings.
 - Should not be noticed by end users.
- Upgrades to recent version of Geant4 (9.1p1) has essentially eliminated problem of event aborts when particle tracking became stuck.
- Output file autonaming option provides provenance.
 panpyZmumuh120-0-500_SLIC-v2r3p10_geant4-v9r0p1_LCPhys_sid01.slcio
 input stdhep file name generator & version geant version physics list det name

slic from scratch:

cvs -d :pserver:anonymous@cvs.freehep.org:/cvs/lcd co SimDist cd SimDist

./configure

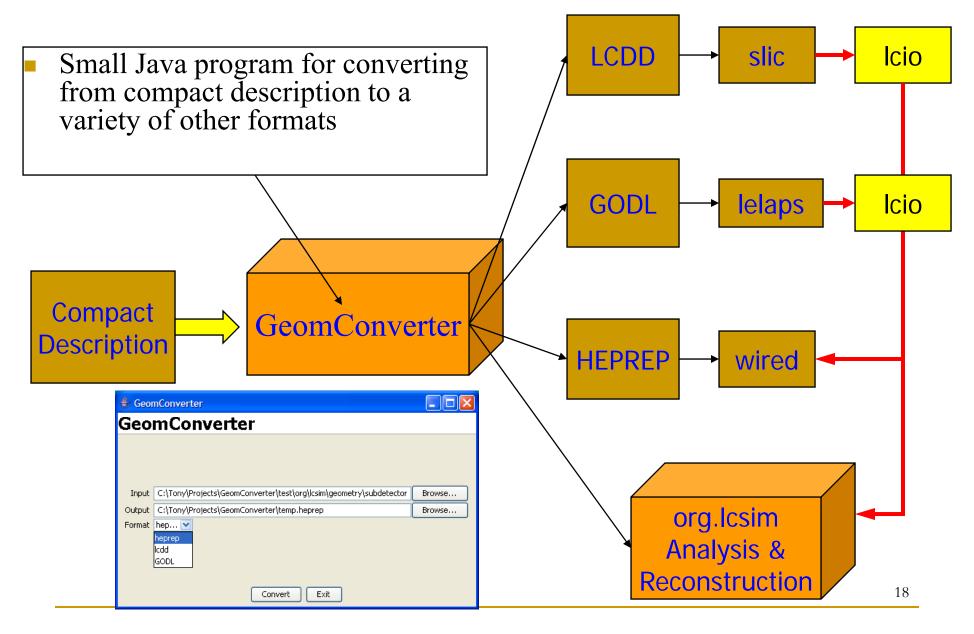
make

Binaries also available for Windows, Mac, Linuxes

Detector Variants

- Runtime XML format allows variations in detector geometries to be easily set up and studied:
 - Stainless Steel vs. Tungsten HCal sampling material
 - □ RPC vs. GEM vs. Scintillator vs. dual readout
 - Layering (radii, number, composition)
 - Readout segmentation (size, projective vs. nonprojective)
 - Tracking detector technologies & topologies
 - TPC, Pixels, Silicon microstrip, SIT, SET
 - "Wedding Cake" Nested Tracker vs. Barrel + Cap
 - Field strength
 - Far forward MDI variants

GeomConverter



LOI Detector Simulations

- Need to clarify exactly what is required for the LOI and what is deferred to the CDR/TDR.
- To-date the philosophy has been to provide a very flexible framework which would allow a number of detectors to be designed and studied very quickly.
- Answer the "big picture" questions first where the details do not matter, then only spend time and effort to study the engineering details on a roughly optimal design.

sid01

- The current model of the Silicon Detector is, therefore, quite simple, consisting of cylindrical barrel geometries and disks in the forward region.
- A lot of effort has gone into ensuring that the correct amount and distribution of material is included, but there are no cracks, gaps or overlaps.
- The detector description can be found at:

http://www.lcsim.org/detectors/sid01

EM Calorimeter Analysis

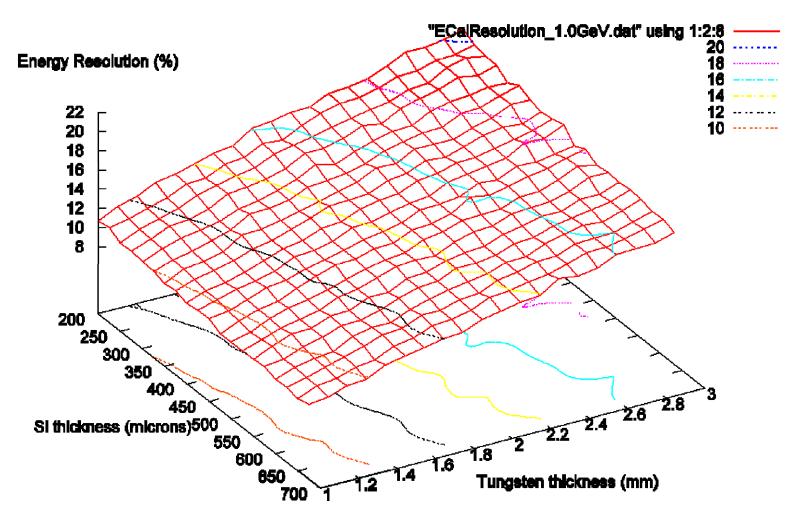
- Can we use simulations to design a detector with good response to photons?
- Start by investigating the intrinsic detector characteristics:
 - Energy linearity
 - Energy resolution
- Analyze the response to single photons.

Simple Geant4 study

- Generate simple sampling calorimeters composed of tungsten-silicon sandwiches.
- Create stacks sufficiently large to contain the full particle showers.
- Vary thicknesses of tungsten and silicon to study the impact on the energy resolution.
- Simulate the response to single photons of varying energy.
- Plot resolution as a function of tungsten and silicon thickness.

Resolution as $fn(d_W, d_{Si})$

Energy Resolution for 1GeV photons



EM Calorimeter Resolution & Linearity

- Repeat study in full detectors, study resolution and linearity for mixed ECal & HCal designs.
- The tools are available to design a system of calorimeters with good energy resolution and linearity of response to photons.
- The baseline silicon detector calorimeters provide an energy resolution of:

 $\sigma/E \sim 18\%/\sqrt{E}$ with ~ 0 constant term.

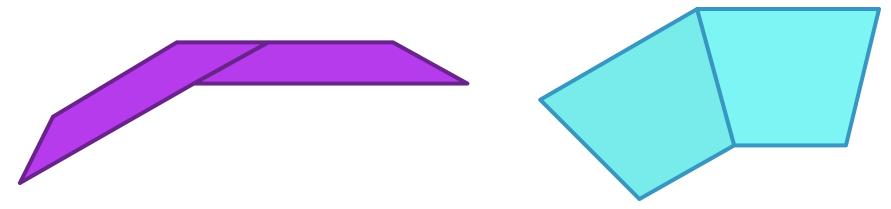
 However, if physics or benchmarking groups insist on better resolution, can straightforwardly come up with designs which provide that performance.

Calorimeter Improved Simulations

- Having settled on a concept with the requisite performance, will have to design a detector which can be built.
- Engineering will have to be done to come up with the plans, but the existing simulation package can already handle arbitrarily complex shapes.
- Can then study effects of support material, dead regions due to stay-clears, readout, power supplies, etc.
- However, hard work is in reconstructing and analyzing this, not simulating it.

Improved Calorimeter Simulations II

- Have two types of polygonal barrel geometries currently defined in the compact description:
- Overlapping staves: Wedge staves:



- Can define ~arbitrary layerings (and boundary stay-clears) within these envelopes to simulate sampling calorimeters.
- Additional geometries (e.g. tilted polyhedra) will be developed on an as-needed basis.
- Support for dual readout, e.g. optical processes, included.

Improved Silicon Tracking Detectors

- For the purposes of quickly scanning the parameter space of number of tracking layers and their radial and z positioning, etc. have been simulating the trackers as cylindrical shells or planar disks.
- Are now moving beyond this to be able to realistically simulate buildable subdetectors.
- Have always been able to simulate arbitrarily complex shapes in slic using lcdd, but this is a very verbose format.
- Introduced Geometry and Detector Element trees to handle arbitrary hierarchies of detector elements.
- Have now introduced tilings of planar detectors (simulating silicon wafers) into the compact xml description.

xml: Defining a Module

```
<module name="VtxBarrelModuleInner">
  <module_envelope width="9.8" length="63.0 * 2" thickness="0.6"/>
  <module component width="7.6" length="125.0" thickness="0.26"
                    material="CarbonFiber" sensitive="false">
                  <position z="-0.08"/>
  </module component>
  <module component width="7.6" length="125.0" thickness="0.05"
                   material="Epoxy" sensitive="false">
                 <position z="0.075"/>
  </module_component>
  <module component width="9.6" length="125.0" thickness="0.1"
                    material="Silicon" sensitive="true">
                  <position z="0.150"/>
  </module_component>
</module>
```

xml: Placing the modules

<layer module="VtxBarrelModuleInner" id="1">

<barrel_envelope inner_r="13.0" outer_r="17.0" z_length="63 * 2"/><rphi_layout phi_tilt="0.0" nphi="12" phi0="0.2618" rc="15.05" dr="-1.15"/><z_layout dr="0.0" z0="0.0" nz="1"/>

</layer>

```
<layer module="VtxBarrelModuleOuter" id="2">
```

<barrel_envelope inner_r="21.0" outer_r="25.0" z_length="63 * 2"/><rphi_layout phi_tilt="0.0" nphi="12" phi0="0.2618" rc="23.03" dr="-1.13"/><z_layout dr="0.0" z0="0.0" nz="1"/>

</layer>

```
layer module="VtxBarrelModuleOuter" id="3">
```

```
<barrel_envelope inner_r="34.0" outer_r="38.0" z_length="63 * 2"/><rphi_layout phi_tilt="0.0" nphi="18" phi0="0.0" rc="35.79" dr="-0.89"/><z_layout dr="0.0" z0="0.0" nz="1"/>
```

</layer>

```
layer module="VtxBarrelModuleOuter" id="4">
```


<barrel_envelope inner_r="46.6" outer_r="50.6" z_length="63 * 2"/>

<rphi_layout phi_tilt="0.0" nphi="24" phi0="0.1309" rc="47.5" dr="0.81"/><z_layout dr="0.0" z0="0.0" nz="1"/>

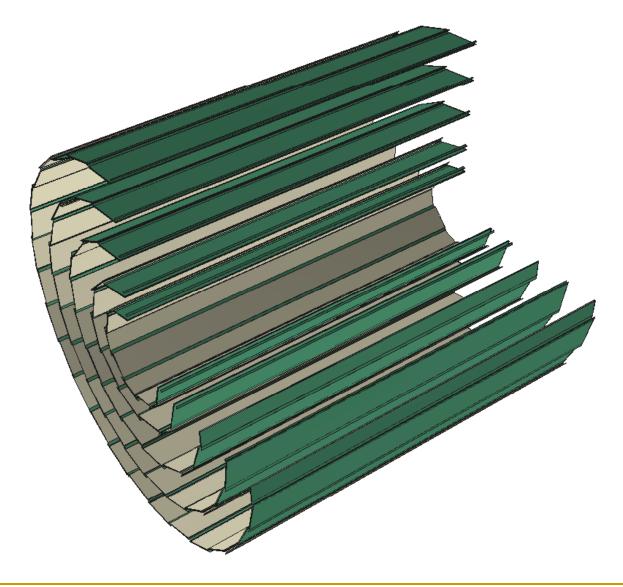
</layer>

```
layer module="VtxBarrelModuleOuter" id="5">
```

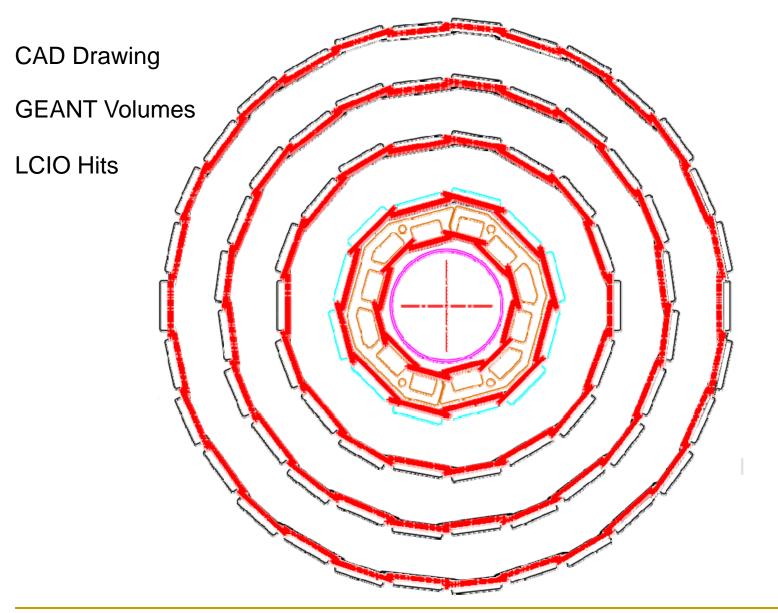
<barrel_envelope inner_r="59.0" outer_r="63.0" z_length="63 * 2"/><rphi_layout phi_tilt="0.0" nphi="30" phi0="0.0" rc="59.9" dr="0.77"/><z_layout dr="0.0" z0="0.0" nz="1"/>

</layer>

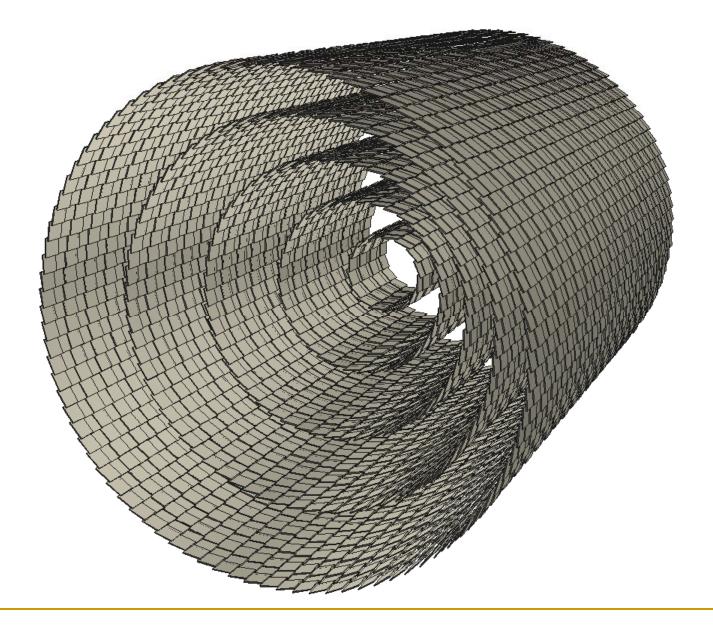
The Barrel Vertex Detector



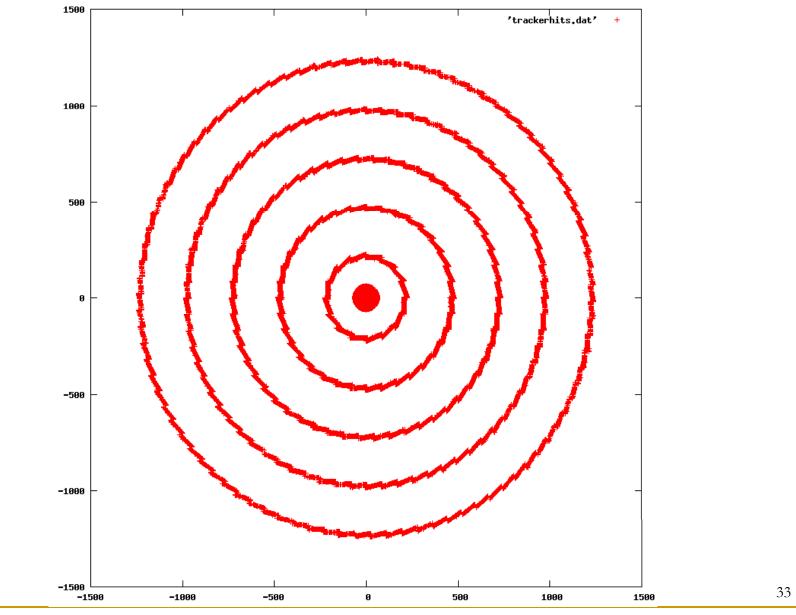
LCIO SimTracker Hits from Vertex



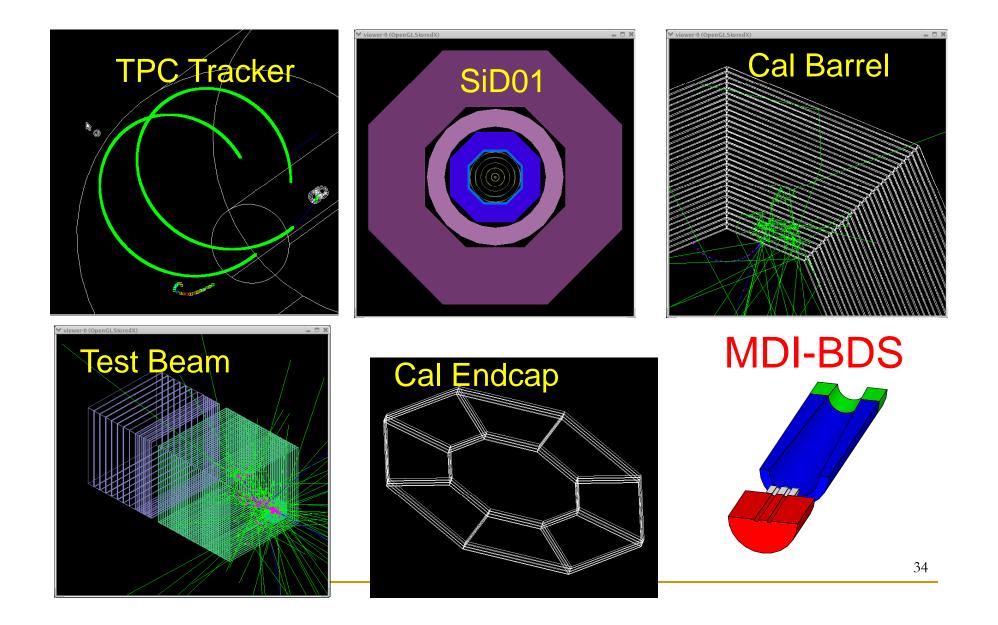
The Barrel Outer Tracker



LCIO SimTracker Hits from Tracker



Example Geometries (same exe)



More complex geometries

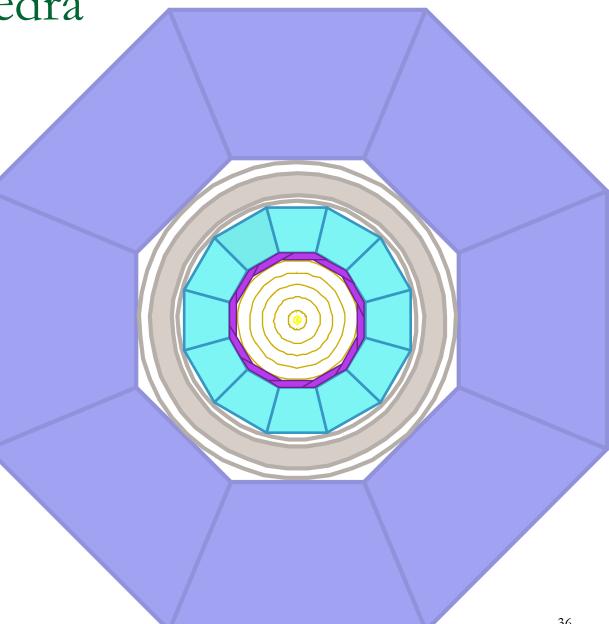
- Arbitrarily complex geometries can be accommodated in the lcdd detector description, but these will not be propagated to the reconstruction system.
 - May be appropriate for supports, readouts, and far-forward BDS/MDI elements.
- CAD to GEANT functionality implemented, and tested for simple elements, so engineering drawings for some elements can be adopted ~as-is.
- Will, of course, have an impact on performance due to use of BREPS in Geant.

sid01_polyhedra

Dodecagonal, overlapping stave EMCal

Dodecagonal, wedge HCal

Cylindrical Solenoid with substructure Octagonal, wedge Muon

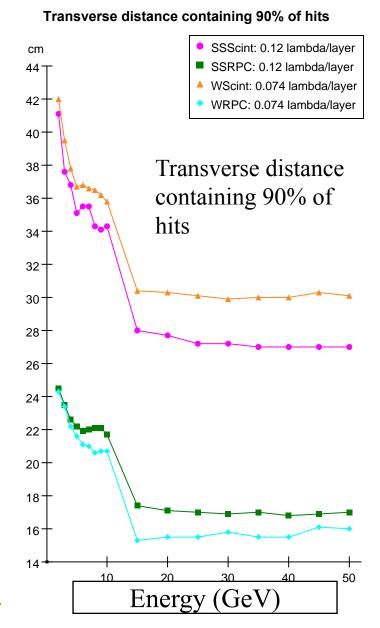


Defining the Detector(s) for the LOI

- As the detectors become more realistic, and therefore more complex, more interaction is needed between subdetector experts and the simulation group.
- The sim/reco group provides the tools to define detectors, but it is the responsibility of the detector subgroups to ensure that their design is being simulated.
- Can we find someone to take on the responsibility to lead the optimization process?
- Need to establish a Change Control Board for THE Silicon Detector.

Physics Lists & Hadron Showers

- Number of issues still unresolved with choice (or not) of showering models.
- For many neutral hadrons LHEP (Gheisha-inspired) is only choice.
- Transition regions still very worrisome.
- LCPhys is still default, but all lists available.
- FTFP_BERT looks good.



Diagnostic Histograms

- Diagnostic plots of event data
 - MCParticles, hits, clusters
- Runs on different detectors
 - must have compact description
 - also need sampling fractions
- Easy to use and setup
- on the web at

http://lcsim.org/SlicDiagWeb/

- SLAC CVS project
 - cvs –d :pserver:anonymous @cvs.freehep.org:/cvs/lcd co

2,000

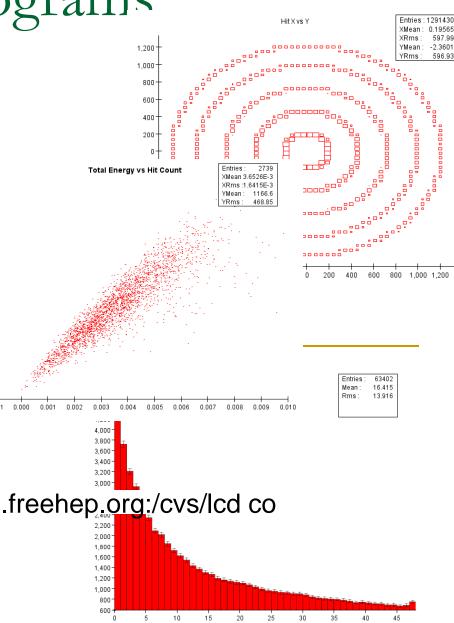
1,800

1,400

1,200⁻ 1,000⁻ 800⁻ 600⁻

> 400 200

- ./build.sh
- ./bin/SlicDiagnostics [...]



Event Samples

- Have generated canonical data samples and have processed them through full detector simulations.
- simple single particles: γ , μ , e, $\pi^{+/-}$, n, ...
- composite single particles: π^0 , ρ , K^0_S , τ , ψ , Z, ...
- Z Pole events: comparison to SLD/LEP
- WW, ZZ, tt, qq, tau pairs, mu pairs, Zγ, Zh
- beam pairs, muons, $\gamma\gamma \rightarrow$ hadrons, etc. backgrounds
- inclusive 1 ab⁻¹ Standard Model sample
- Web accessible <u>http://www.lcsim.org/datasets/ftp.html</u>
- Additive at the detector hit level, with time offsets.
 - □ Investigate effects of full backgrounds.

Beam Background Overlays

- Take output from full beam simulation (from IR/backgrounds group)
- Feed into full detector simulation
- Build library of simulated background bunches
- Overlay backgrounds on signal events at start of reconstruction
 - Adjust timing of hits (for TPC e.g.)
 - Add energy in calorimeter cells
 - Allows to change #bunches/train, bunch timing

Additional LOI Simulation Issues

Detector Geometries

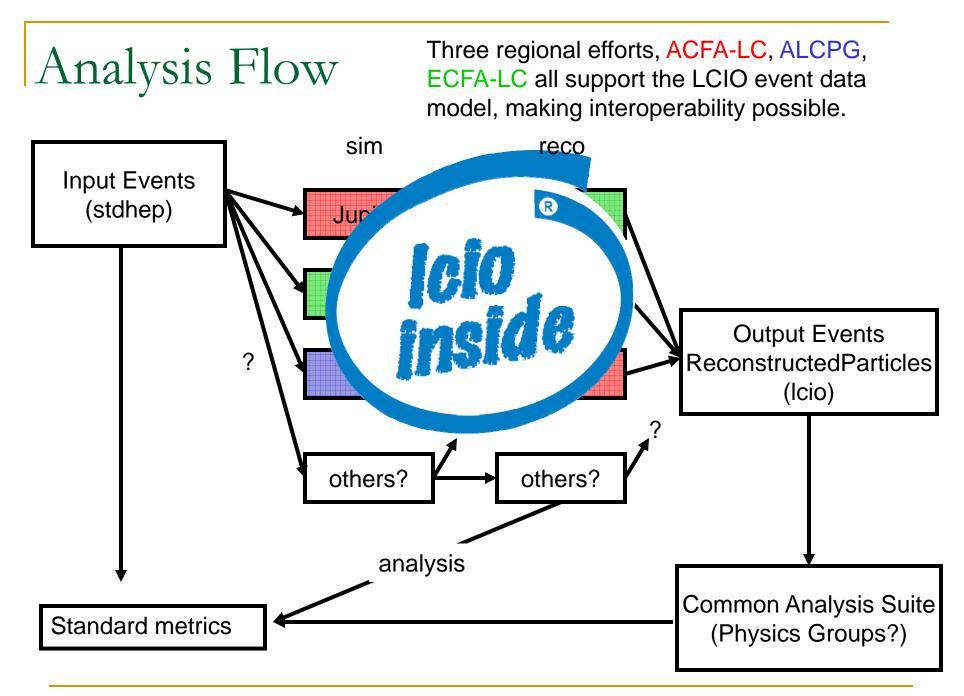
- Toolkit is capable of modeling very detailed detector geometries, including cracks, support structures and other "dead" material.
- In fullness of time, mature reconstruction will account for such issues.
- Question is, how much effort should be devoted to putting in effects in the simulation only to take them out again (or not) in the reconstruction?

Detector Magnetic Fields

 Implementing full field maps in Geant is very CPU consuming. Accounting for non-homogeneous fields in the reconstruction is also extremely CPU intensive. Propose to generate signals using simplified fields, full fields only for far-forward backgrounds.

Reconstruction

- This topic will be addressed by other speakers later today and tomorrow.
- Many of the reconstruction algorithms (track finding, fitting, calorimeter clustering, etc.) are in place.
- A number of weak links are identified and need work.
- Please do not reinvent the parts that have already been done!
- Use what we have!
- Adopt, adapt, improve.



Detector Description Improvement

- Each detector subgroup needs to decide very soon what level of detail it believes it needs, and what can be achieved, given the limited resources.
- All detector design files are in cvs, so concurrent development of the full detector can proceed in parallel.
 - Check out LCDDetectors, modify the compact.xml file to update your detector.
 - Check for geometrical overlaps or other conflicts
 - commit.
 - □ iterate.

Detector Description Process

- Depending on the magnitude of the changes, the reconstruction (as it is) may or may not have to change.
- Calibration samples will need to be run to determine things like the sampling fraction corrections for analog sampling calorimeters.
- Physics and some number of background samples will have to be processed.
- Reconstruction will need to be run.
- Analyses will have to be run.
- Iterate if time remains.

Detector Response Simulation

- Will need to process a large number of events through the full simulation, reconstruction and analysis chain for this exercise.
 - As always, very critically resource limited, not just CPU and disk space, but also personnel.

Summary

- The ALCPG sim/reco group provides SiD a mature, full-featured, detector simulation package, capable of modeling very detailed geometries.
 - How detailed should the simulations be for the LOI?
- LCIO event data model and file format allows exchange of simulated (& reconstructed) data between regional software packages.
 - Number of analyses presented at TILC08 exploited this functionality, e.g.
 - Jupiter GLD simulation input to MarlinReco PandoraPFA
 - slic SiD simulation fed to LCFI flavor-tagging package

Advice for the SiD LOI

"Don't get it right, get it written"

James Thurber

- Waiting for all of the elements to be designed in excruciating detail is a recipe for failure.
- It would be much better to proceed with what we have all the way through to the end, even if this means simplified detector designs and partial reconstruction, than to wait until the very end.
- Get a draft out soon, then refine it.

Additional Information

- SiD <u>http://silicondetector.org</u>
- Icsim.org <u>http://www.lcsim.org</u>
- ILC Forum <u>http://forum.linearcollider.org</u>
- Wiki <u>http://confluence.slac.stanford.edu/display/ilc/Home</u>
- org.lcsim <u>http://www.lcsim.org/software/lcsim</u>
- Detectors <u>http://www.lcsim.org/detectors</u>
- LCIO <u>http://lcio.desy.de</u>
- SLIC <u>http://www.lcsim.org/software/slic</u>
- JAS3 <u>http://jas.freehep.org/jas3</u>
- AIDA <u>http://aida.freehep.org</u>
- WIRED <u>http://wired.freehep.org</u>

Backup Slides

Benchmark Selection Process

- WWS Software panel (Akiya Miyamoto, KEK, Ties Behnke, DESY, Norman Graf, SLAC,) in consultation with the detector concepts and the WWS Roadmap Panel and starting from the Benchmark Panel Report has drafted a short list of processes which are:
 - consistent with the ILC baseline
 - sensitive to detector performance
 - not overly dependent on sophisticated analysis techniques.
 - i.e. emphasis is on demonstrating detector performance
 - Expect, and welcome, input from IDAG

Benchmark Processes

$e^+e^- \rightarrow ZH, H \rightarrow e^+e^-X, \mu^+\mu^-X (M_H = 120 \text{ GeV}, E_{cms} = 250 \text{ GeV})$

- tracking efficiency and momentum resolution
- material distribution in the tracking detectors
- EM shower ID, kink reconstruction (bremsstrahlung)
- Higgs Mass and cross section

$$e^+e^- \rightarrow ZH, H \rightarrow cc, \mu^+\mu^- Z \rightarrow v i (\bar{M}_H = 120 \text{ GeV}, E_{cms} = 250 \text{ GeV})$$

- heavy flavour tagging, secondary vertex reconstruction
- multi jet final state, c-tagging in jets, uds anti-tagging
- test anti-tagging by studying the H→gg
- BR(H $\rightarrow c\bar{c}$), BR (H $\rightarrow \mu^{+}\mu^{-}$)

Benchmark Processes

 $e^+e^- \rightarrow ZH, H \rightarrow cc, \mu^+\mu^{-}, Z \rightarrow qq$ (M_H=120GeV, E_{cms}=250GeV)

- in addition to the charm tagging, this final state tests the confusion resolution capability
- BR(H \rightarrow c \overline{c}), BR (H \rightarrow $\mu^{+}\mu^{-}$)

$$e^+e^- \rightarrow Z \rightarrow \tau^+\tau^- (E_{cms} = 500 \text{ GeV})$$

- tau reconstruction, aspects of particle flow
- π⁰ reconstruction
- tracking of very close-by tracks
- σ , A_{FB} , and τ decay mode efficiency and purity.

Benchmark Processes

 $e^+e^- \rightarrow tt, t \rightarrow bW, W \rightarrow qq' (M_{top}=175 GeV, E_{cms}=500 GeV)$

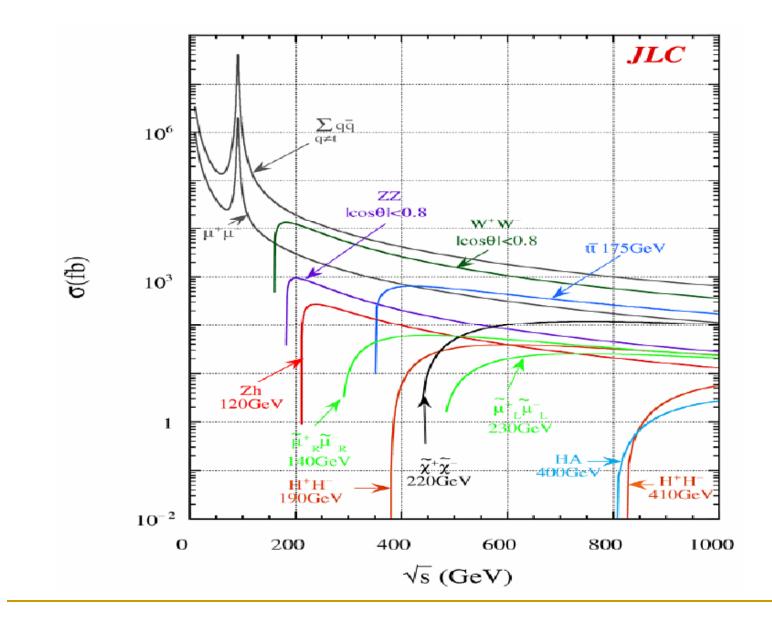
- multi jet final states, dense jet environment
- particle flow
- b-tagging inside a jet
- maybe lepton tagging in hadronic events (b-ID)
- tracking in a high multiplicity environment
- σ , A_{FB} , and m_{top}

$$e^+e^- \rightarrow \chi^+ \chi^- / \chi_2^0 \chi_2^0$$
 (E_{cms}=500 GeV)

- particle flow (WW, ZZ separation)
- multi-jet final states
- SUSY parameter is point 5 of Table 1 of hep-ex/0603010
- σ, and masses

Standard Model Backgrounds

- All analyses used in the context of the detector optimization and LOI process will need an inclusive sample of the Standard Model Background.
- Will provide the SM sample centrally in stdhep format, for all concepts to use.
- Provide all information and tools necessary to produce specific signal samples individually with exactly the same setup as this SM sample.



Event Generators

- No single MC generator is optimal for everything.
- However, Whizard is a multi-purpose Matrix Element generator.
 - Signals and backgrounds of all types (SM + MSSM) can be produced with the same settings
 - It contains all interferences, hence it is more accurate than generators like Pythia, especially for complex final states (6f and more)
 - Some inaccuracies remain, but benefits outweigh these minor issues.

Whizard SM Sample

- Generate an inclusive set of MC events with all SM processes
- WHIZARD Monte Carlo used to generate all 0,2,4,6-fermion and t quark dominated 8fermion processes.
- 100% e⁻ and e⁺ polarization used in generation. Arbitrary electron, positron polarization simulated by properly combining data sets.
- Fully fragmented MC data sets are produced.
 PYTHIA is used for final state QED & QCD parton showering, fragmentation, particle decay.
- Events are weighted!

Standard Model Sample

- Full 2ab⁻¹ SM sample available via ftp from SLAC.
- Each file corresponds to a particular initial e-/e+ polarization and final state

ftp://ftp-lcd.slac.stanford.edu/ilc/whizdata/ILC500/

cumbersome to work with for end user

Have to mix polarizations by hand

Each file contains only processes of one type, so need to run over complete data set (thousands of files) to get faithful subset.

500 fb⁻¹ sample of these events generated with 80% e⁻, 30% e⁺ polarizations, randomly mixed events from all processes

<u>ftp://ftp-lcd.slac.stanford.edu/ilc/ILC500/StandardModel/</u>

Next Steps for SM Data Sample

- Remove 120 Higgs from n fermion final states at 500 GeV, and add explicit ffH, ffHH, etc. final states.
- Regenerate states with τ in final state using TAUOLA.
 - Coding done at DESY, generation in progress.
- Produce full SM data set at 250 GeV
 - Machine parameters have been obtained from the GDE, GuineaPig events being generated to provide pair background events and beam spectra to be used for event generation.

Additional Signal Processes

- Detector concepts are free (and encouraged) to add additional processes to this list in order to optimize their designs or demonstrate the capabilities of the detectors.
- These should, however, be generated using conditions as close as possible to those used for the canonical samples.
- It is more important for this process that we use a common, well-understood set of events than it is to pick the "best" generator for each final state.
- We are comparing detector response, not making physics case for the machine.

Additional Backgrounds

- GuineaPig pairs and photons (Cain too?)
 - Added crossing angle, converted to stdhep, available <u>here</u>.
- Muons and other backgrounds from upstream collimators & converted to stdhep.
 - Need to validate and understand normalizations.
- $\gamma\gamma \rightarrow$ hadrons generated as part of the "2ab⁻¹ SM sample."
- All events then capable of being processed through full detector simulation.
- Additive at the detector hit level, with time offsets, using LCIO utilities.
 - i.e. simulate response separately for signals and backgrounds, then add at digitization/reconstruction level.

Additional Event Generation Issues

Crossing Angle

- Agree that events will be generated with 0° and the 14mr crossing angle will be accounted for at the time of simulation.
- Tuning of pythia parton fragmentation has come up.
 - Will proceed with the pythia defaults due to the large number of events already generated.
- Essential to fully document and maintain a provenance for all files (and analyses?)