Benchmarking SiD

Andrei Nomerotski, Univ.of Oxford
SiD Workshop, Abingdon, 14 Apr 2008

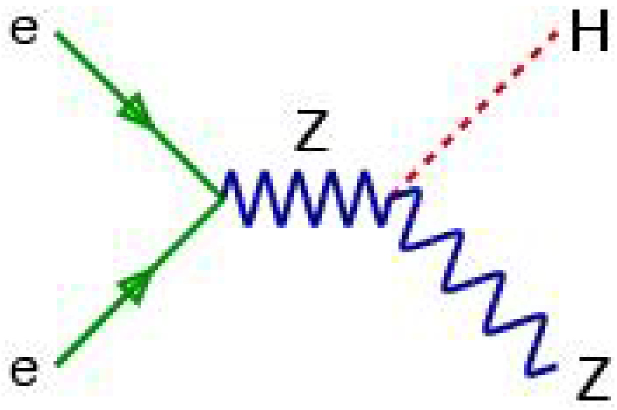
Physics at LC

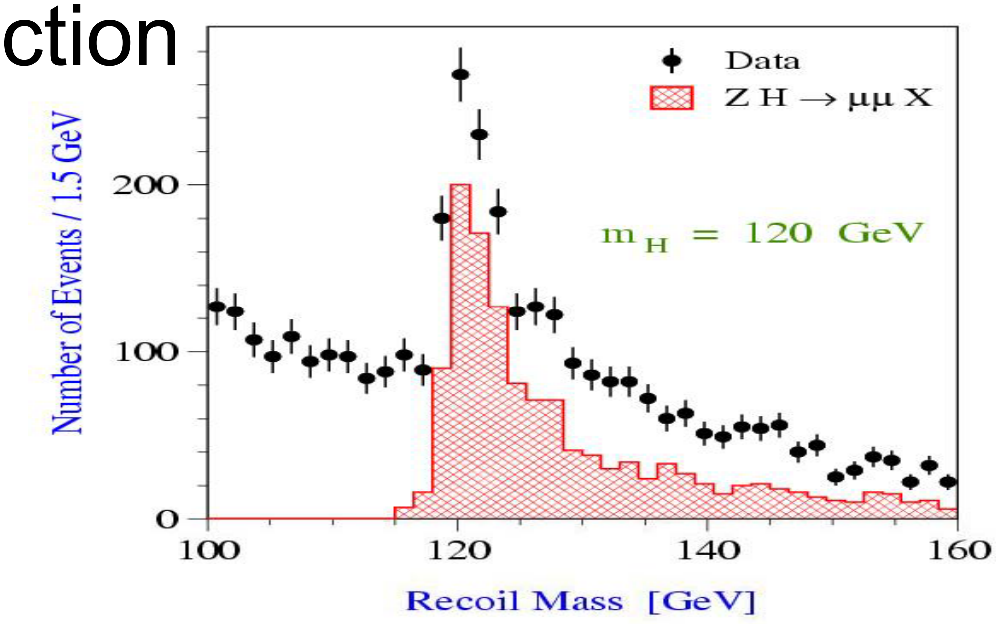
- Emphasis on
  - Higgs properties
  - Precision of measurements
  - SUSY parameters

But many other things are interesting and important

ZH

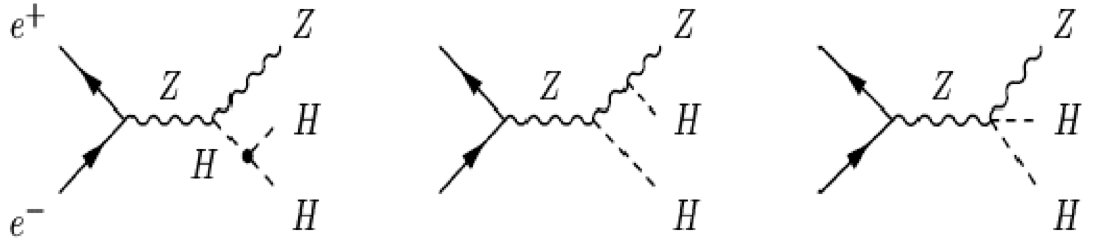
- Higgs:
  - Higgs can be reconstructed through recoil mass independently of its decay channel
  - Even for invisible decays
- Rely on Z reconstruction only



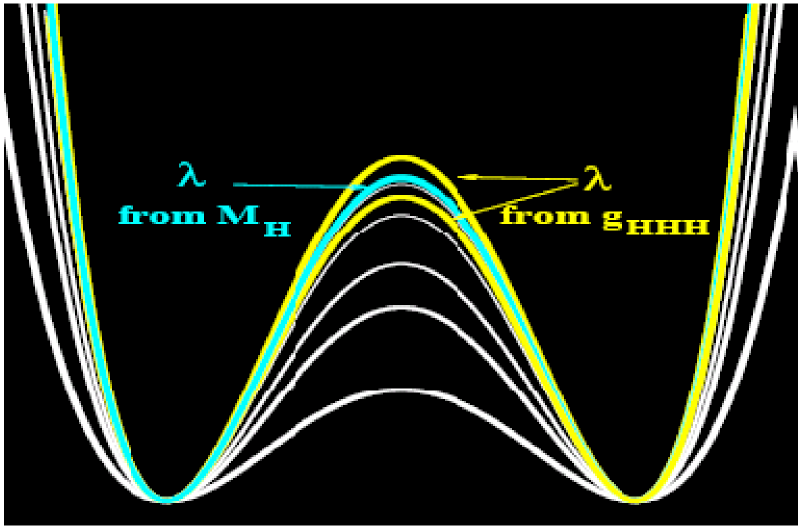


ZHH

Double Higgstrahlung: $e^+e^- \rightarrow H^0H^0Z^0$

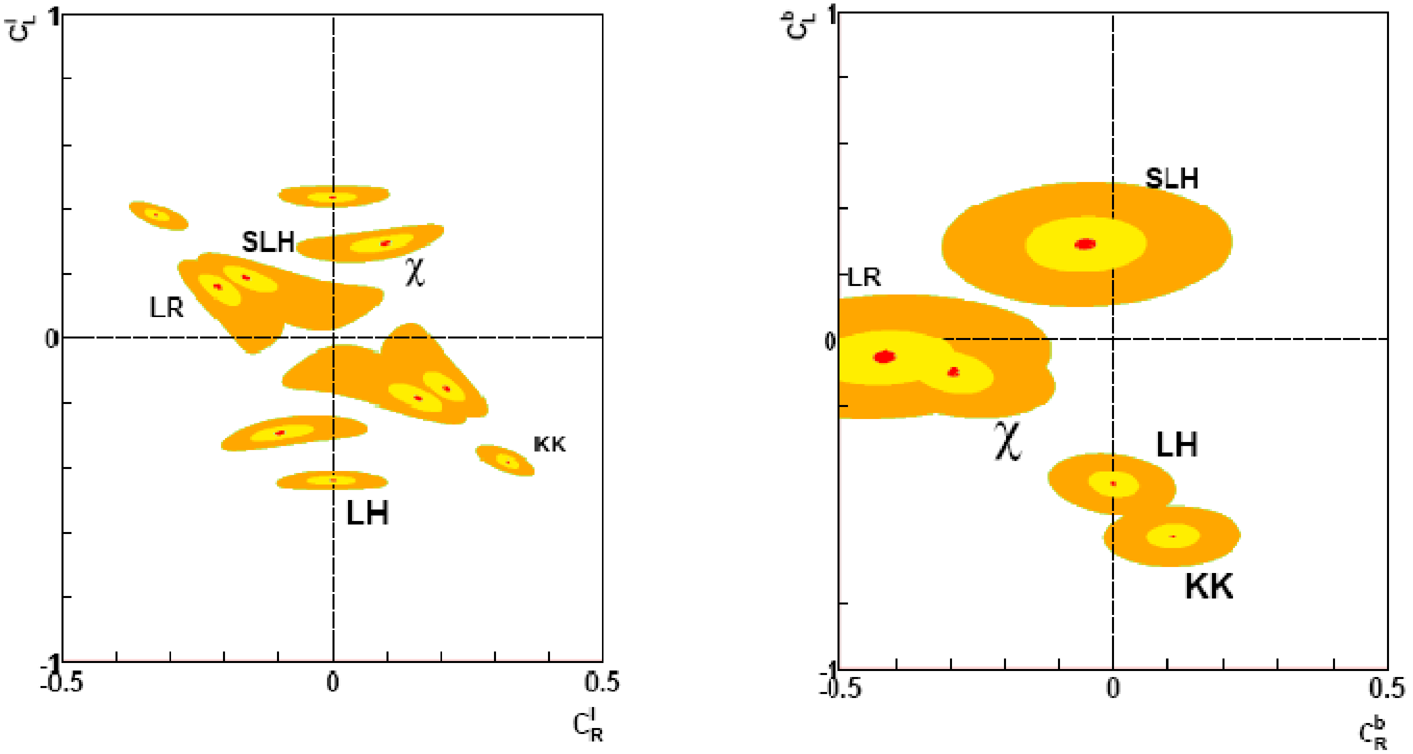


- Key to understanding of Higgs potential – mechanism of symmetry breaking
- Higgs potential can be derived independently from M_H and from g_{HHH} and compared
- Talk by Yiming Li tomorrow



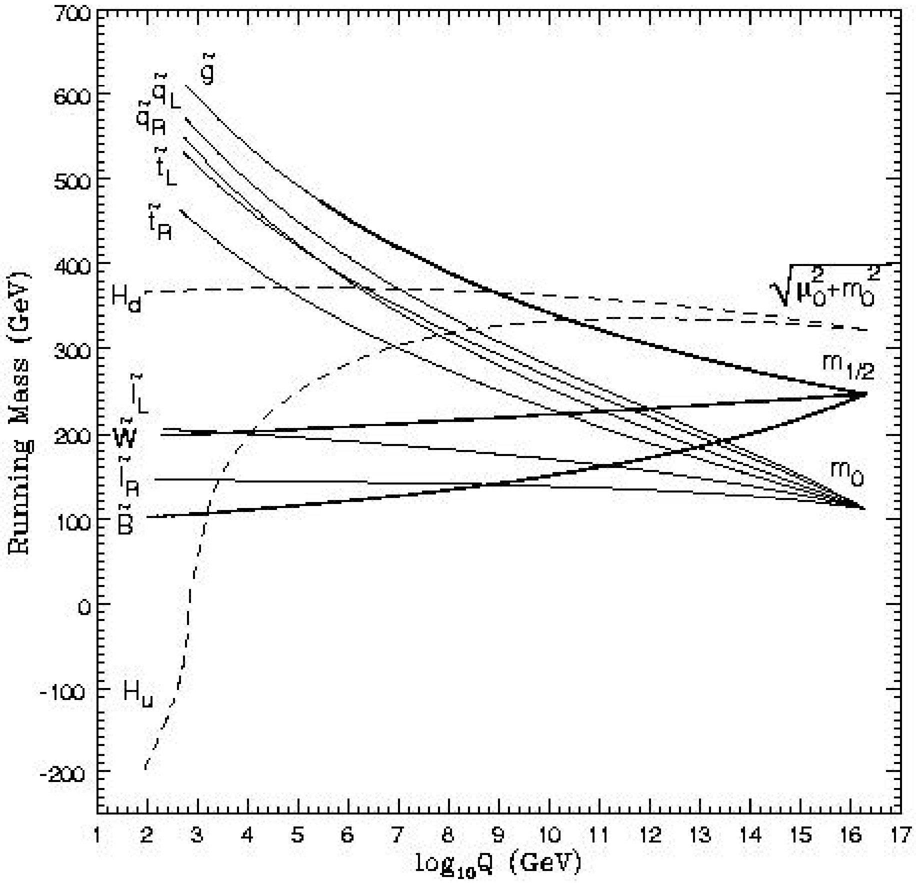
Precision

- Precision measurements of σ, A_{LR} and A_{FB} allows to distinguish between different Z' models: Little Higgs, Simplest Little Higgs, E_6, KK excitations, LR-symmetric
- Different channels (ex. Leptons and b-quark) allow to resolve ambiguitities
- In SM : $C_L^e = -\frac{1}{2} + \sin^2\theta_w \qquad C_R^e = \sin^2\theta_w$
- CM 500 GeV, m(Z')= 1,2,3 TeV; $1ab^{-1}$

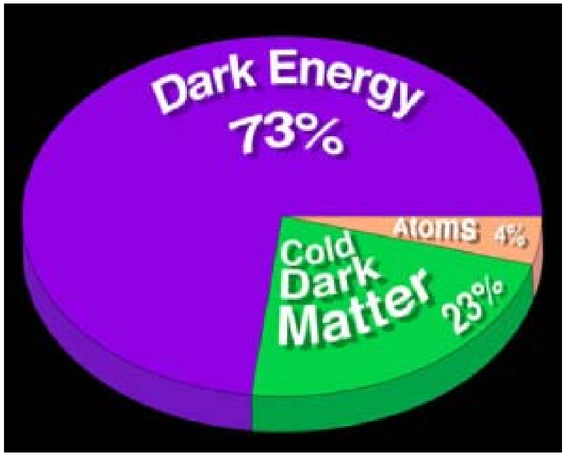


Supersymmetry

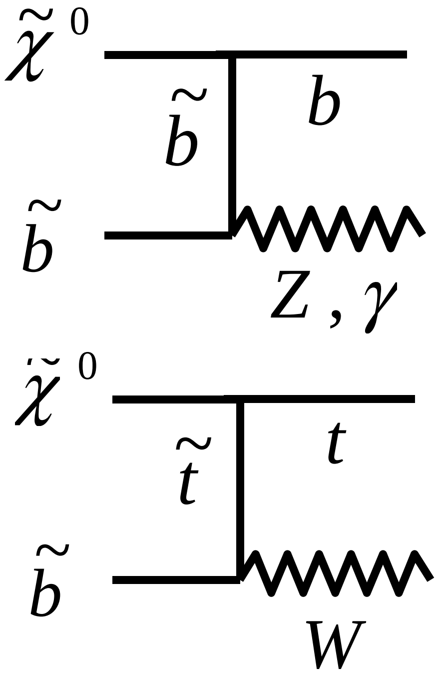
- Rich phenomenology
  - Mass spectrum and couplings are determined by several high mass scale parameters
  - They also define decay modes of SUSY particles → signatures
- Huge variety of signatures!



Cosmology Motivated SUSY Scenarios

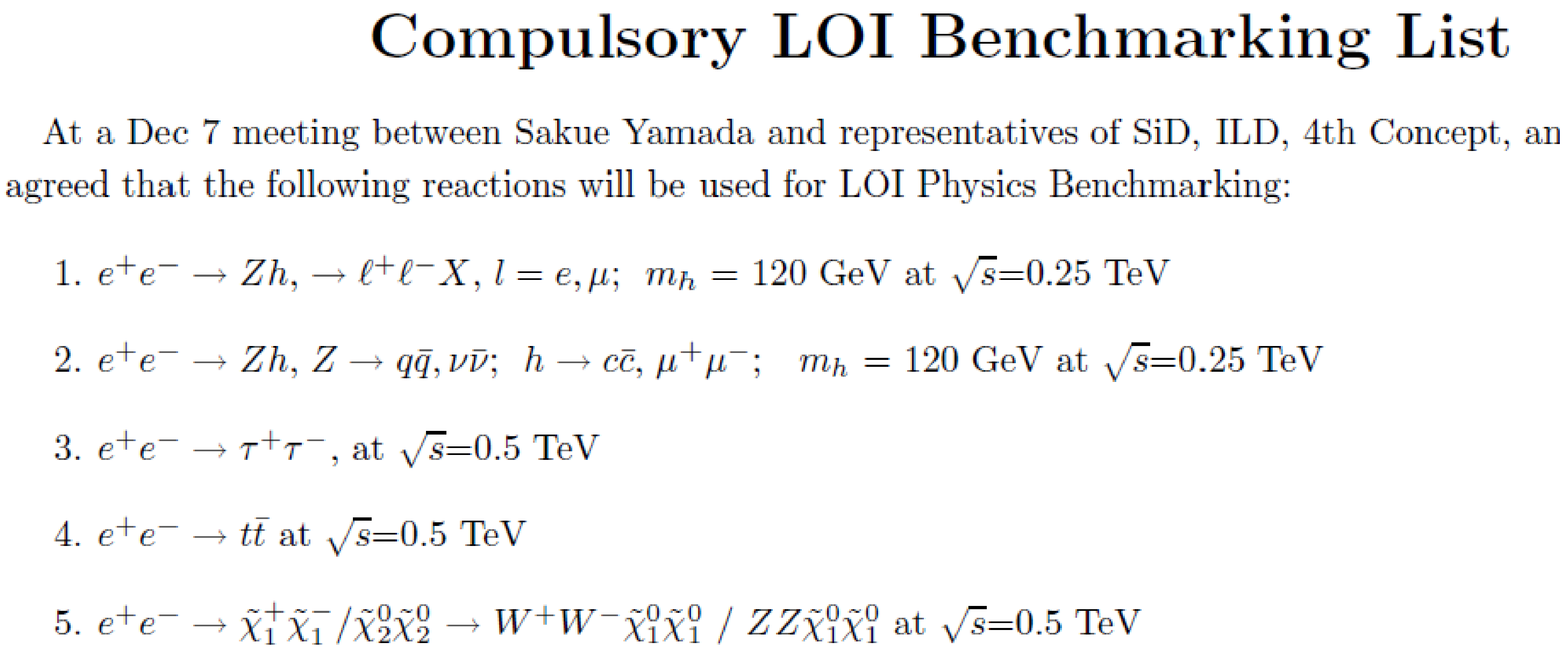


- Dark Matter is 23% of Universe – how to explain?
- In SUSY : DM is neutralino, co-annihilation with other SUSY particles regulate how much DM is left after Big Bang
  - Need small mass splitting
- Small mass split between LSP and NLSP = small visible energy in the detector
  - ee → stops, sbottoms, staus
  - Important case to motivate the massless Tracker with zero P_T cutoff
- Large two –photon backgrounds
  - Need to veto electron/positron in forward systems
- Talk by Tomas Lastovicka tomorrow



From Physics Studies to Benchmarking

- Emphasis of physics studies will shift towards
  - Evaluation and comparison of detector choices
  - Realities required by engineering: material (amount and distribution)
  - Realities required by reconstruction algorithms: tracking & PFA



Compulsory LOI Benchmarking List

At a Dec 7 meeting between Sakue Yamada and representatives of SiD, ILD, 4th Concept, an agreed that the following reactions will be used for LOI Physics Benchmarking:

1. $e^+e^- \to Zh, \to \ell^+\ell^- X, l = e, \mu;\ m_h = 120$ GeV at $\sqrt{s}$=0.25 TeV
2. $e^+e^- \to Zh,\ Z \to q\bar{q}, \nu\bar{\nu};\ h \to c\bar{c}, \mu^+\mu^-;\ \ m_h = 120$ GeV at $\sqrt{s}$=0.25 TeV
3. $e^+e^- \to \tau^+\tau^-$, at $\sqrt{s}$=0.5 TeV
4. $e^+e^- \to t\bar{t}$ at $\sqrt{s}$=0.5 TeV
5. $e^+e^- \to \tilde{\chi}_1^+\tilde{\chi}_1^- / \tilde{\chi}_2^0\tilde{\chi}_2^0 \to W^+W^-\tilde{\chi}_1^0\tilde{\chi}_1^0$ / $ZZ\tilde{\chi}_1^0\tilde{\chi}_1^0$ at $\sqrt{s}$=0.5 TeV

- 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 (Snowmass 2005)
- Had iterations in February and March to define observables more precisely

Benchmark Processes

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

- tracking efficiency and momentum resolution
- material distribution in the tracking detectors
- EM shower ID

- Observable: Higgs Mass and cross section

Benchmark Processes

$e^+e^-\rightarrow ZH$, $H\rightarrow cc$, $Z\rightarrow\nu\{$ (M_H=120 GeV, E_{cms}=250 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
- Observable: BR(H → cc)

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

- in addition to the charm tagging, this final state tests the confusion resolution capability
- Observable: BR(H → cc)

Benchmark Processes

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

- tau reconstruction, aspects of particle flow
- π^0 reconstruction
- tracking of very close-by tracks

- Observables: efficiency and purity for tau main decay modes with and without π^0

Benchmark Processes

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

- multi jet final states, dense jet environment
- particle flow
- b-tagging inside a jet
- tracking in a high multiplicity environment

- Observables: σ, A_{FB}, and m_{top}

- Talk tomorrow by Erik Devetak

Benchmark Processes

$e^+e^- \longrightarrow \chi^+\chi^- / \chi_2^0\chi_2^0$ (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

- Observables: σ and χ^+,χ_2^0 masses

Additional SiD Benchmarking Studies for LOI

6. $e^+e^- \to c\bar{c}, b\bar{b}$, at $\sqrt{s}$=0.5 TeV;

   $A_{FB}^{LR}(c)$ & $A_{FB}^{LR}(b)$

Talk today by Ben Jeffery

7. $e^+e^- \to Zhh$, m_h = 120 GeV at $\sqrt{s}$=0.5 TeV;

   g_{hhh}

8. $e^+e^- \to \tilde{\tau}_1\tilde{\tau}_1$, at Point 3 at $\sqrt{s}$=0.5 TeV;

   $M_{\tilde{\tau}_1}$ $\sigma(e^+e^- \to \tilde{\tau}_1\tilde{\tau}_1)$

9. $e^+e^- \to \tilde{t}_1\tilde{t}_1^* \to c\bar{c}\tilde{\chi}_1^0\tilde{\chi}_1^0$, $m_{\tilde{t}_1}$ = 120 GeV, $m_{\tilde{\chi}_1^0}$ = 100 GeV, at $\sqrt{s}$=0.5 TeV

   $M_{\tilde{t}_1}$, $\sigma(e^+e^- \to \tilde{t}_1\tilde{t}_1^*)$, $\cos\theta_{\tilde{t}}$

10. $e^+e^- \to \tilde{b}_1\tilde{b}_1^* \to bb\tilde{\chi}_1^0\tilde{\chi}_1^0$, at $\sqrt{s}$=0.5 TeV

    $M_{\tilde{b}_1}$, $\sigma(e^+e^- \to \tilde{b}_1\tilde{b}_1^*)$

11. $e^+e^- \to \mu^+\mu^-$, at $\sqrt{s}$=0.5 TeV

    Luminosity Weighted $\sqrt{s}$

12. H→ γγ

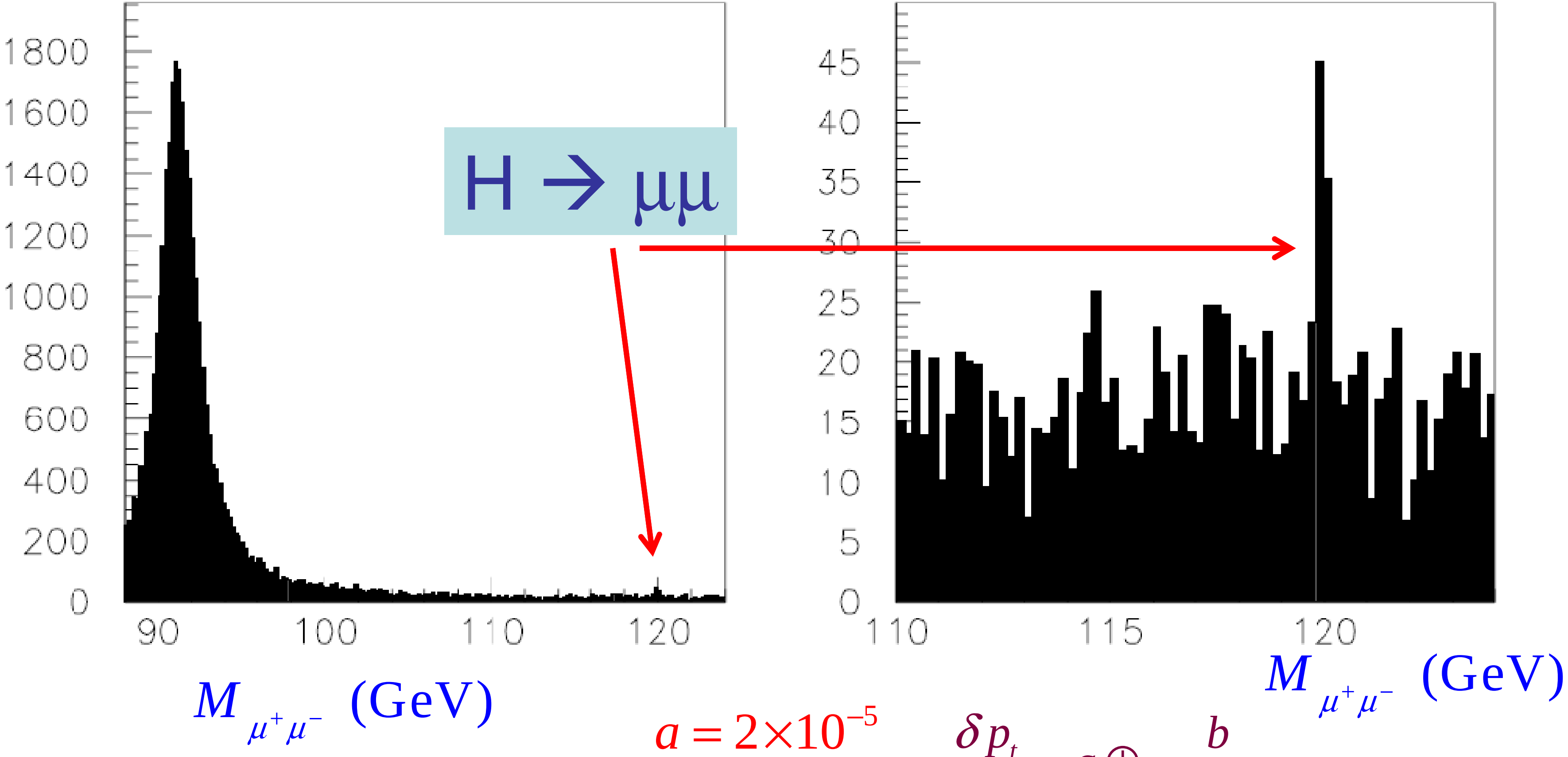
Detector concepts are free (and encouraged) to add additional processes to this list in order to optimize their designs or demonstrate the capabilities of the detectors

Benchmarking Processes

- Compulsory and additional processes will allow to benchmark subsystems
  - Vertexing
  - Tracking
  - EM and HAD Calorimetry
  - Muon system
  - Forward system
- and to compare SiD to other concepts

H→μμ

- One of important Higgs Br
- $M_{\mu\mu}$ distributions for NN>0.95 for signal and background summed
  - Need excellent mass resolution

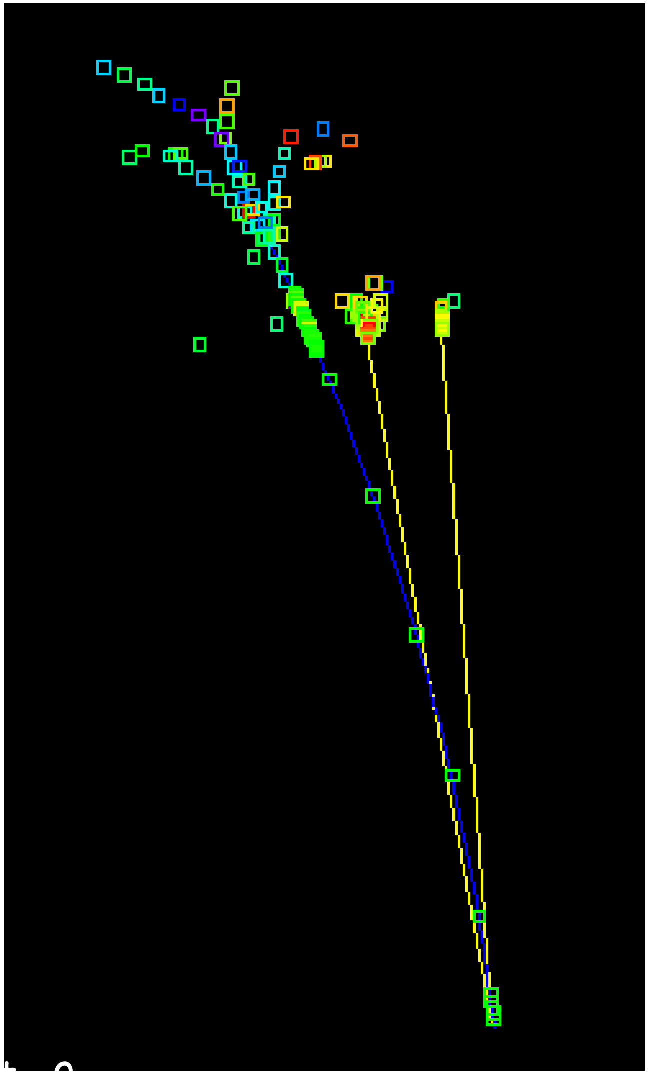


$a = 2\times10^{-5}$

$b = 1\times10^{-3}$

$$\frac{\delta p_t}{p_t^2} = a \oplus \frac{b}{p_t \sin\theta}$$

Importance of π^0



- $H \to \tau\tau$ process
- Tau polarization (from $\tau \to \rho\nu \to \pi^+\pi^o\nu$) allows to determine CP properties of Higgs
- Separation of clusters and reconstruction of π^o requires excellent segmentation of EMCAL (MAPS EMCAL)
- Also : using π^0 to constrain the vertex mass → improvements in b-tagging

Tools for Benchmarking

Java based lcsim.org framework

- lcsim.org FastMC
  - Smeared MC informaiton
- lcsim.org full MC: SLIC
  - GEANT based
- Perfect PFA
- Vertexing / Flavour tagging : LCFI package
- Track reconstruction and Full PFA

Analysis Model

- Use FastMC to develop analysis algorithms
- Use full MC and Perfect PFA as intermediate step to develop a realistic analysis

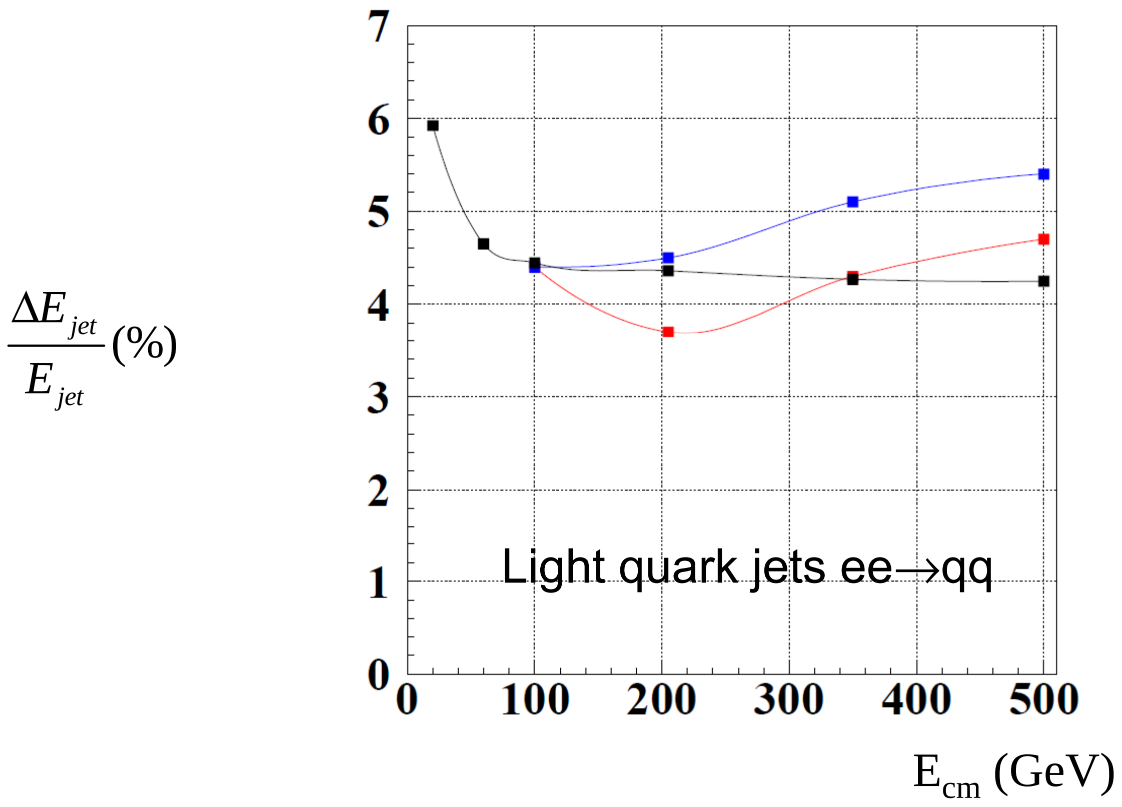
- Use realistic tracking and PFA for the analysis when ready
  - A drop-in replacement of algorithms

org.lcsim FastMC simulation of Calorimeter/PFA output

Use tracker momentum for all charged tracks within acceptance;
account for confusion term by blowing up single particle resolution for neutral hadrons

🟦 GLD PFA 🟥 LDC PFA ■ org.lcsim FASTMC with

$$\frac{\Delta E_{\gamma}}{E_{\gamma}} = \frac{0.18}{\sqrt{E_{\gamma}}} \qquad \frac{\Delta E_{n,K_L^0}}{E_{n,K_L^0}} = 0.28$$



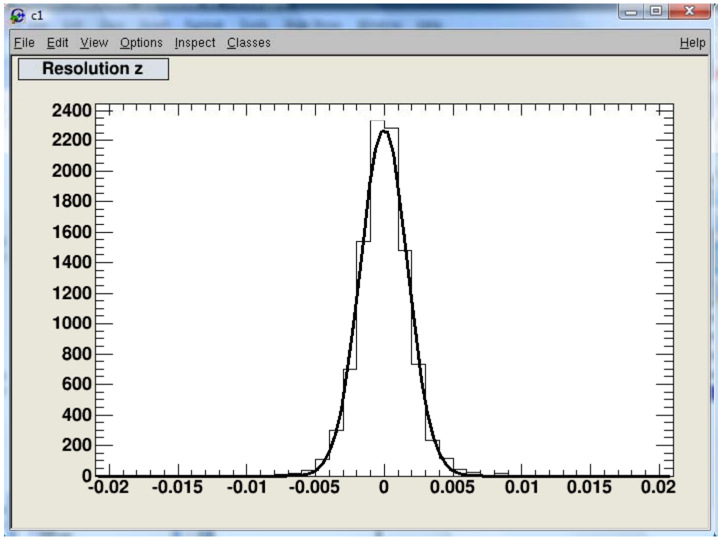
Perfect PFA

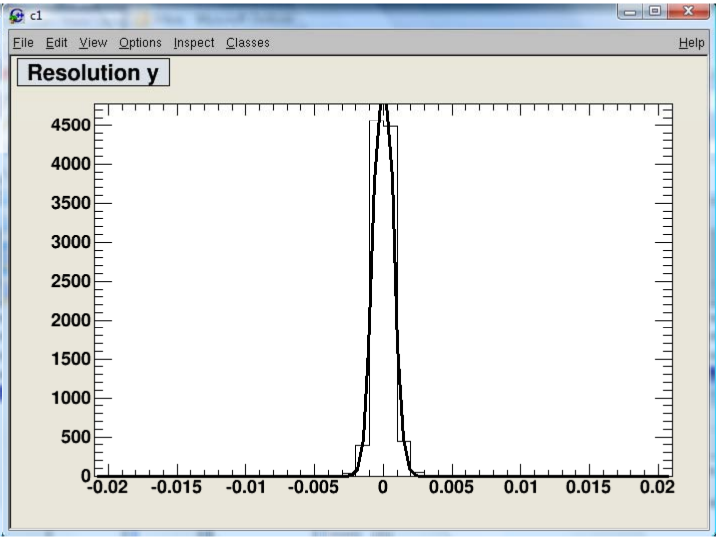
- Tracking
  - Define “trackable” charged particles
  - Smear as in FastMC
  - Full material effects (interactions and decays) before the calorimeter are taken into account in deciding which particles are actually tracked
- Neutrals
  - For all “non-trackable” particles, assign energy deposits in the calorimeters
  - Do neutral particle reconstruction using those deposits using perfect pattern recognition (no confusion term)
  - Use actual detector responses for energy and direction - so most of the nasty nonlinear, nongaussian effects are included

Primary Vertex Reconstruction with org.lcsim FastMC

LCFI vertexing package can be used with lcsim.org

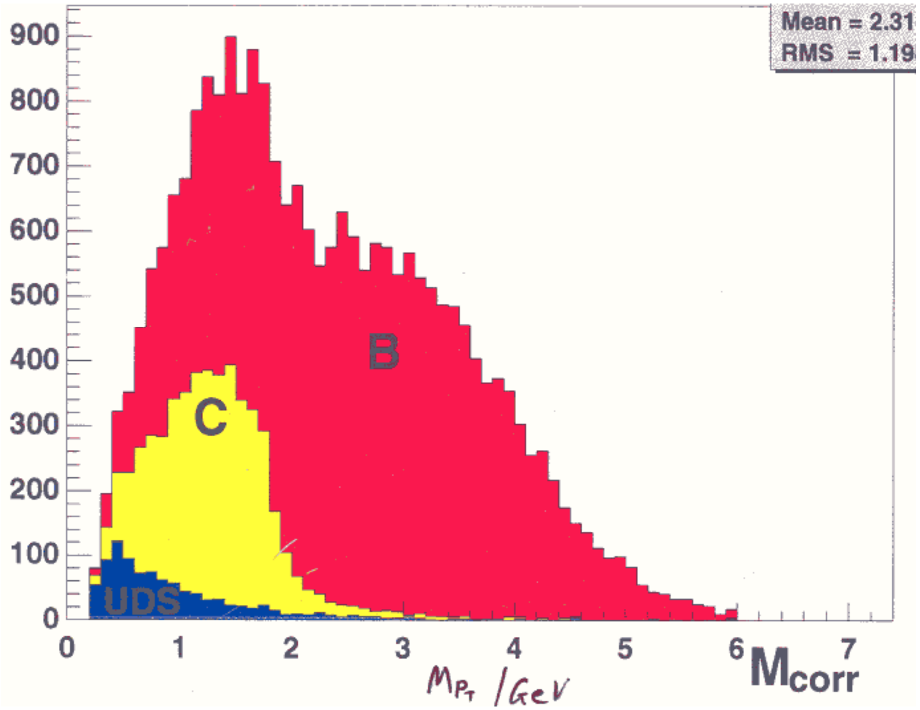
- Resolution for primary vertex in ZHH events:
  - 2 µm in z-direction
  - <0.8 µm in x-y plane (beamspot constr.) 2 µm (no beamspot constr.)
  - Resolution pulls are nice Gaussians with σ ~ 1.1 for all x,y and z





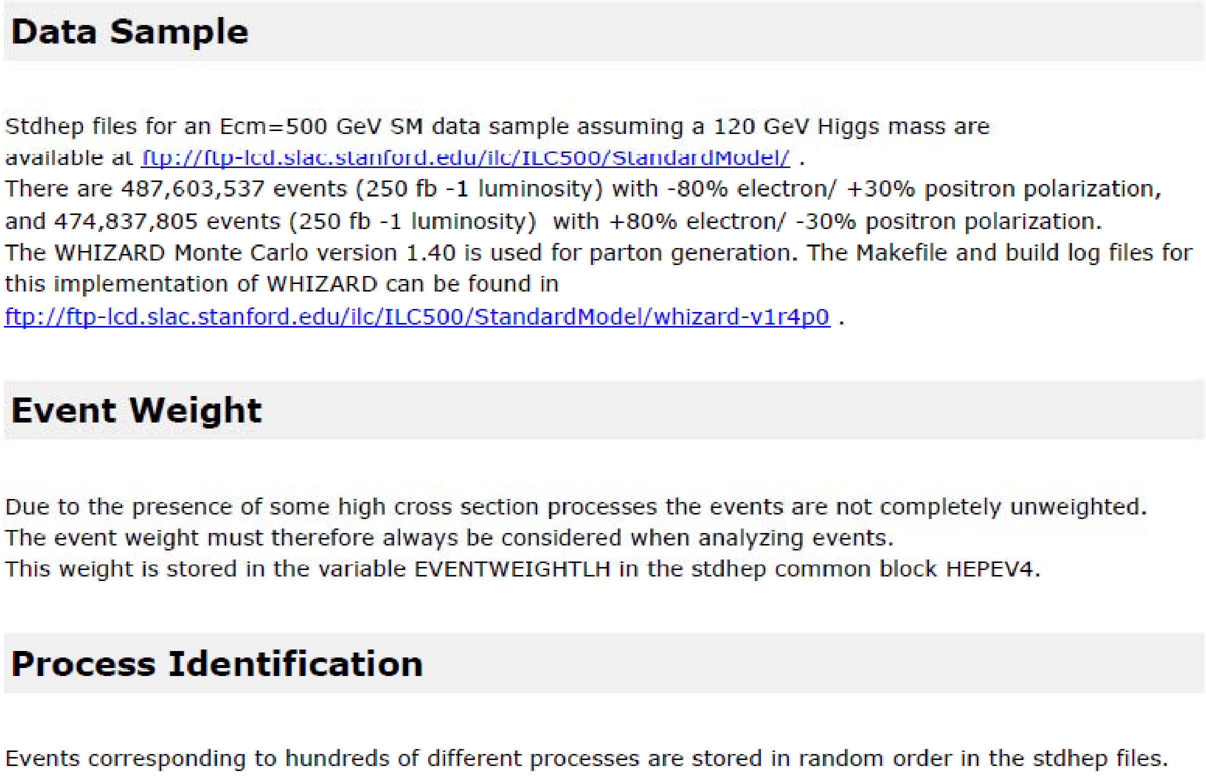
LCFI Flavour Identification

- Combine several variables into Neural Net
  - **Vertex mass**
  - Vertex momentum
  - Decay length
  - Decay length significance
  - Jet Probability
- Main contributors are Vertex Mass and Jet Probability



Standard Model Samples

- Full 2 ab^{-1} inclusive SM Data Sample available at SLAC via ftp
- Each file corresponds to a particular initial e-/e+ pol. and final state
- Web documentation
- .stdhep format, will be used by all concepts
- Need to add 250 GeV samples



Data Sample

Stdhep files for an Ecm=500 GeV SM data sample assuming a 120 GeV Higgs mass are available at ftp://ftp-lcd.slac.stanford.edu/ilc/ILC500/StandardModel/ .
There are 487,603,537 events (250 fb -1 luminosity) with -80% electron/ +30% positron polarization, and 474,837,805 events (250 fb -1 luminosity) with +80% electron/ -30% positron polarization.
The WHIZARD Monte Carlo version 1.40 is used for parton generation. The Makefile and build log files for this implementation of WHIZARD can be found in ftp://ftp-lcd.slac.stanford.edu/ilc/ILC500/StandardModel/whizard-v1r4p0 .

Event Weight

Due to the presence of some high cross section processes the events are not completely unweighted.
The event weight must therefore always be considered when analyzing events.
This weight is stored in the variable EVENTWEIGHTLH in the stdhep common block HEPEV4.

Process Identification

Events corresponding to hundreds of different processes are stored in random order in the stdhep files.

Status

- Several analyses done previously used both Fast and Full MC
  - ZH, ZHH, H→ μμ, ttbar …
- Recent developments with tools
  - Used LCFI package with lcsim.org samples
  - Used Perfect PFA for ZHH analysis

- ZHH, ttbar and sbottom channels will be covered in Benchmarking session tomorrow

Countdown for LoI

- April 2009 : submit LoI
- February 2009 : results available
- October 2008 : all MC samples available
  - SM sample is most CPU intensive
- August 2008 : start production
  - Reconstruction ready
- July 2008 : all analyses developed on fastMC and PPFA, generator level samples ready

How to contribute

- We look for people to lead several analyses – both compulsory and additional

- Physics studies are ideal for newcomers, fast track way to contribute to LoI

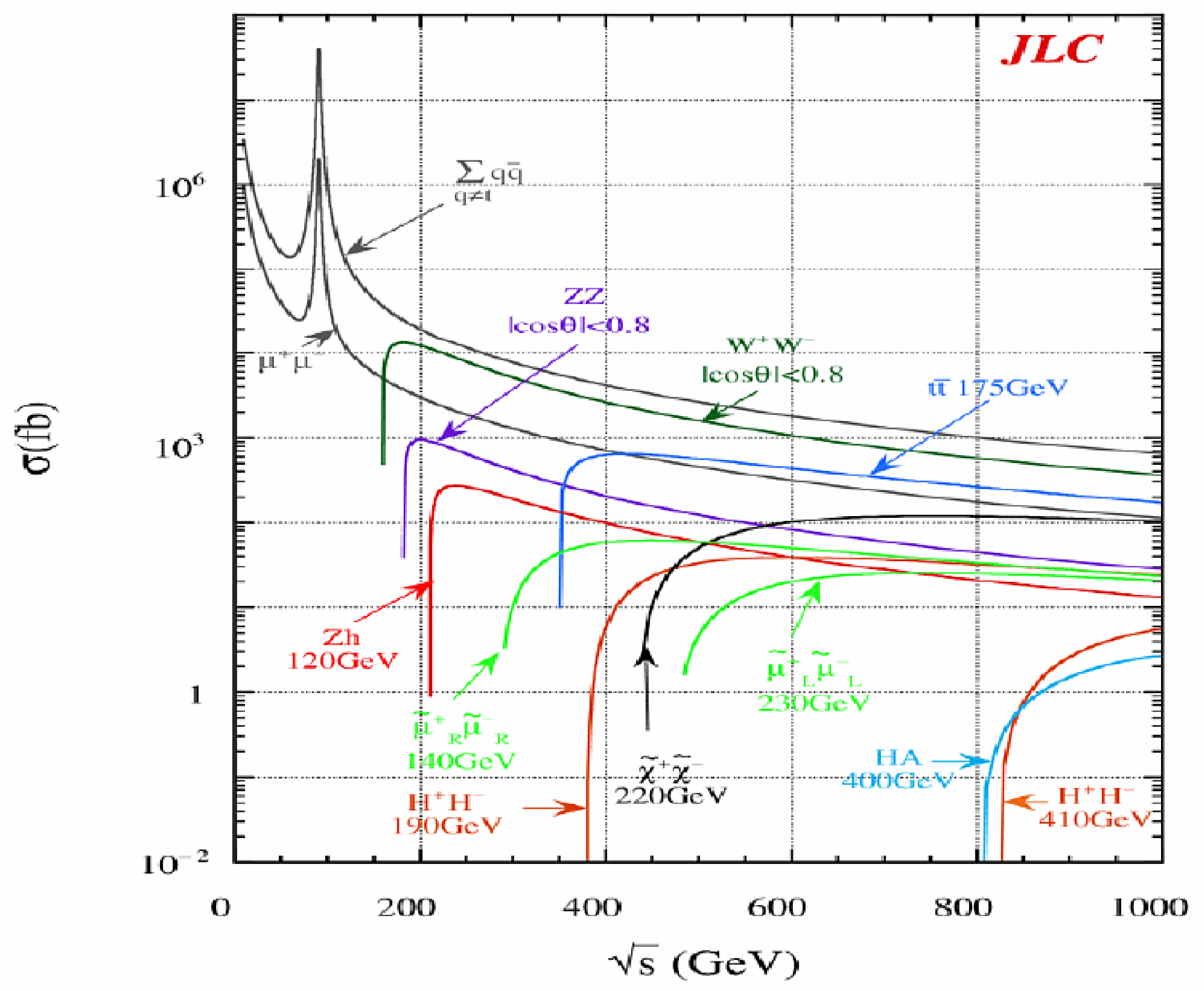
Benchmarking Manpower

1. $e^+e^- \to Zh, \to \ell^+\ell^- X$, $l = e, \mu$; $m_h = 120$ GeV at $\sqrt{s}$=0.25 TeV	SLAC
2. $e^+e^- \to Zh$, $Z \to q\bar{q}, \nu\bar{\nu}$; $h \to c\bar{c}, \mu^+\mu^-$; $m_h = 120$ GeV at $\sqrt{s}$=0.25 TeV	Michigan/Oxford
3. $e^+e^- \to \tau^+\tau^-$, at $\sqrt{s}$=0.5 TeV	?
4. $e^+e^- \to t\bar{t}$ at $\sqrt{s}$=0.5 TeV	RAL/Oxford
5. $e^+e^- \to \tilde{\chi}_1^+\tilde{\chi}_1^- / \tilde{\chi}_2^0\tilde{\chi}_2^0 \to W^+W^-\tilde{\chi}_1^0\tilde{\chi}_1^0$ / $ZZ\tilde{\chi}_1^0\tilde{\chi}_1^0$ at $\sqrt{s}$=0.5 TeV	SLAC
6. $e^+e^- \to c\bar{c}, b\bar{b}$, at $\sqrt{s}$=0.5 TeV;	Oxford
7. $e^+e^- \to Zhh$, $m_h = 120$ GeV at $\sqrt{s}$=0.5 TeV;	Oxford
8. $e^+e^- \to \tilde{\tau}_1\tilde{\tau}_1$, at Point 3 at $\sqrt{s}$=0.5 TeV;	Texas A&M ?
9. $e^+e^- \to \tilde{t}_1\tilde{t}_1^* \to c\bar{c}\tilde{\chi}_1^0\tilde{\chi}_1^0$, $m_{\tilde{t}_1} = 120$ GeV, $m_{\tilde{\chi}_1^0} = 100$ GeV, at $\sqrt{s}$=0.5 TeV	Lancaster
10. $e^+e^- \to \tilde{b}_1\tilde{b}_1^* \to b\bar{b}\tilde{\chi}_1^0\tilde{\chi}_1^0$, at $\sqrt{s}$=0.5 TeV	Oxford/Montenegro
11. $e^+e^- \to \mu^+\mu^-$, at $\sqrt{s}$=0.5 TeV	SLAC
12. H→ γγ	RAL

Backups

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.


JLC
σ(fb)
10^6
10^3
1
10^-2
0
200
400
600
800
1000
√s (GeV)
Σ q≠t qq̄
μ+μ−
ZZ |cosθ|<0.8
W+W− |cosθ|<0.8
tt̄ 175GeV
Zh 120GeV
μ̃+R μ̃−R 140GeV
H+H− 190GeV
χ̃+χ̃− 220GeV
μ̃+L μ̃−L 230GeV
HA 400GeV
H+H− 410GeV

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 8-fermion 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
  - ftp://ftp-lcd.slac.stanford.edu/ilc/ILC500/StandardModel/

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, to be incorporated soon.
- Produce full SM data set at 250 GeV
  - Need agreed-upon machine parameters.

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.

Producing Signal Samples

- All necessary files to set up Whizard in exactly the same way as done for the SM sample:
  - http://confluence.slac.stanford.edu/display/ilc/Standard+Model+Data+Samples
- Beam Parameters can be set up using information at:
  - ftp://ftp-lcd.slac.stanford.edu/ilc/ILC500/StandardModel/whizard-src/user.f90
  - ftp://ftp-lcd.slac.stanford.edu/ilc/ILC500/StandardModel/guinea-pig/ilc_0500_may05_run05_seed06/
- Fragmentation can be set up using information at:
  - ftp://ftp-lcd.slac.stanford.edu/ilc/ILC500/StandardModel/a6f/include/ilc_fragment_call.f90
  - ftp://ftp-lcd.slac.stanford.edu/ilc/ILC500/StandardModel/a6f/include/calc_a1sq_a2sq.f90

Additional Backgrounds

- GuineaPig pairs and photons (Cain too?)
  - Added crossing angle, converted to stdhep, available here.
- 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 Simulation 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.
- Detector Magnetic Fields
  - Implementing full field maps in Geant is very CPU consuming. Propose to generate signals using simplified fields, full fields for far-forward backgrounds.
- Essential to fully document and maintain a provenance for all files (and analyses?)