Particle Identification Issues
for Linear Collider Detectors

Classification:

Linear Collider Detector – Particle Identification

Institution and Personnel requesting Funding

Colorado State University
Robert J. Wilson, Professor
David Warner, Engineer

Wilson has designed and built particle identification systems in experiments such as the SLD experiment Cerenkov Ring Imaging Device (CRID); BaBar Detector of Internally Reflected Cherenkov light (DIRC). He has been the coordinator of the Particle ID subgroup of the American Linear Collider Physics Group, contributed to the Snowmass Resource Book, and has given numerous presentations on hadron ID issues for Linear Collider experiments.

Collaborators

Blair Ratcliff (SLAC) – subsystem hadron ID performance and hardware
Anthony Johnson and Norman Graf (SLAC) – software infrastructure

Project Leader

Robert J. Wilson
wilson@lamar.colostate.edu
(970) 491-5033

Project Overview

A primary goal of the next linear collider is to provide detailed investigations of fundamental physics in the 500-1000 GeV energy regime that are not possible with a hadron collider. While Particle Identification (PID) in the broad sense will certainly play a central role [1], the extent to which identification of stable hadrons (pi, K, p) is required remains an open question. The issue has particular relevance for detectors without gas-based tracking systems, such as the SiD detector concept, which lack even rudimentary hadron ID [2]. The primary purpose of this proposed research is to support the core of activity of a Linear Collider Detector Particle ID group.

A similar research plan proposed several years ago received strong support from the review panel at the time but, as with many other priority requests, the available level of support allowed for only slight progress. However, as a more concrete timeline for detector Conceptual Design

Reports is now being developed, the issue has reach a level of urgency that warrants a renewed attempt to provide more concrete information on which to base detector design decisions.

We propose to build on previous work in three areas: (1) Investigation of the need for particle identification in linear collider physics analyses, with particular emphasis on hadron identification – in particular, this will include conversion and expansion of an existing fast Particle ID package [3] and its integration into the JAS3[1]-based Linear Collider Detector (LCD) simulation package; (2) Investigation of the performance parameters required of a specialized hadron ID subsystem (if one is justified) and its impact on the overall experiment performance; (3) Liaison with detector concept groups and to re-establish and maintain an ILC PID web site [4].

This effort is not uniquely affiliated with any of the existing detector concept groups since the primary questions to be addressed in the proposal are generic to ILC physics. In addition, historically, the investigation of dedicated hadron ID subsystems has not been favored by current groups since the physics benefit for doing so is unclear and the effect on other aspects of detector performance is presumed to be negative. We will address both of these issues quantitatively in this activity.

Investigation of the LC physics justification for hadron ID will be done in collaboration with members of the various detector concepts groups and with the Linear Collider Flavour Identification collaboration in the UK [5]. For hadron ID hardware subsystem issues we will communicate closely with Stanford Linear Accelerator Center (SLAC) scientist Blair Ratcliff and members of his group. The software infrastructure tasks will be done in close collaboration with Anthony Johnson and Norman Graf (SLAC).

Broader Impact

The broadest impact of this proposal is, of course, as part of the vast intellectual endeavor of the International Linear Collider itself. However, this activity will make specific contributions to the broader impact of the project in several ways.

One example is the integration of research and training of the next generation of scientists. This project will continue a successful strategy used by the PI for several years, in which linear collider (LC) projects have been used to introduce new students (primarily graduate, but also some undergraduate) to high energy physics and scientific software design and implementation. Two recent PhD recipients (Mahalaxmi Krishnamurthy and Qinglin Zeng) received their initial training working with a prototype of the particle identification package proposed here. Both learned object oriented programming and the fundamentals of event generation and detector simulation on the LC tasks. They were both then able to move directly into analysis of a current experiment (BaBar). Recent B.S. (John Cairns) and M.S. (Sky Rolnick) recipients worked together to produce a complete package and though are no longer with the CSU HEP group, report that the experience has opened up opportunities for them on their career paths. As with

---

[1] Java Analysis Studio v.3, http://jas.freehep.org/jas3/

those previous projects, we will again train students in skills (software and web-based tutorial and information access) that will be valuable beyond the direct application to LC physics.

Another example of broader impact is the investigation of specialized hadron ID systems, which will involve collaboration with existing detector hardware projects. We anticipate that a demonstration of the need for such systems will spark interest and development in industry, particularly for the development of new photodetectors. The PI has already been approached by a small business (aPeak Inc., Boston, MA) about the possibility of a collaboration to develop silicon devices optimized for use with a compact Cerenkov detector (this would be an outgrowth of work currently being done with aPeak Inc. for LC muon systems).

Background

As described in the overview, the most pressing PID issue is the longstanding question of the need for hadron identification for high energy Linear Collider physics; this question is particularly acute for the SiD detector concept, which has silicon as the primary tracking device. This detector would lack even the basic hadron ID capabilities provided by gas-based trackers [2]. There have been a few limited efforts to address this issue: Mercadante and Yamamoto [6] have shown that the production of long-lived tau slepton pairs in a certain mass range may be detected with dE/dx in a gaseous tracking chamber; Wilson [7] has investigated the effect of hadron ID on neutral B meson tagging; and Soffer has considered the use of hadron ID for charm vertex tagging and R-parity and baryon number violating SUSY decays [8].

Most such studies have been done with crude event generator-level ID, partly due to the lack of tools in the U.S. group's standard simulation and reconstruction package. No compelling justification for hadron ID has been found, however, the investigation is clearly incomplete, in large part due to the lack of a sustained effort. As a practical matter, many of the associated issues have low priority in the individual detector concept groups, but taken together they represent an important part of any overall detector design optimization.

Results of Prior Support

In late summer 2003 we received supplemental funds from DoE to develop a JAS2-based package for PID fast simulation and reconstruction that would be easy to use, flexible, and provide tools to allow users to use PID information in their analyses. This work was a subset of a more comprehensive 2002 proposal to investigate the hadron ID issues, similar to that being proposed here.

The award of $30,000 was used primarily to provide salary and tuition for graduate student Sky Rolnick and partial salary for undergraduate student John Cairns (who was also a professional software developer at the time). The package was essentially completed, tested, documented and presented at the 2004 Linear Collider Workshop in Victoria. In this section we summarize the results of that activity.

Introduction

The Particle ID package consists of simulation, reconstruction, and analysis code specifically designed to simulate particle identification at the Linear Collider. It provides the infrastructure to add fast simulation for detector subsystems; provides functional examples of subsystems; and extends the existing event reconstruction class with particle ID information that can be used in user analysis code. It includes the capability to read several different data types, including LCD simulation files and StdHep data files directly. The package was designed to be detector independent so the user will be able to simulate a wide range of detectors either by creating a custom geometry file or choosing from several predefined geometry files. Other useful features of the PID package are: the ability to use multiple PID subsystems; run-time change of geometry parameters; set global or subsystem specific ID thresholds, such as log-likelihood differences; tools for calculating efficiencies and purities; and a class to combine the PID information derived from multiple subsystems (e.g. tracker dE/dx and calorimeter energy) into a single best-estimate parameter.

The current framework, available on the web [9], is a functional version of how particle ID could be implemented into the JAS v2.2.5 LCD structure. Future development is needed in order to bring the PID package into full functionality and, since the PID project was started, the core ILC software groups at SLAC and DESY have changed the base code to JAS3, which makes extensive use of the AIDA[2] system. A deliverable of this proposal is to convert the existing package to function with JAS3 and AIDA.

The particle ID package has been designed to be simple, flexible, and extensible for the end user. Our aim was to create a package that would allow users to simulate detector subsystems and include particle ID information into their analysis with minimal coding effort. A user can now add particle ID information to an existing driver file with as little as four lines of code. The package is flexible, in that it allows many of the detector parameters to be altered directly from the driver file during compile time and allows users to simulate several detector geometries simultaneously (e.g. l2, l2dirc, s2, s2dirc) for comparison in analysis.

A significant improvement over the original package was to restructure the code into separate fast simulation, fast reconstruction, and user analysis code modules, as illustrated in Figure 1. By separating out the simulation code from the reconstruction code, it is more portable and maintainable and provides the user with more flexibility in the way they choose to do analysis. In Figure 2, we provide an example of the basic JAS code to perform an analysis that includes PID information layered on top of the original MCFast code for fast simulation (track smearing) of generator level Monte Carlo events. After MCFast fills the LCDEvent, MCFastPID performs the equivalent fast simulation for PID parameters (track energy loss, dE/dx, by default). This is followed by a PIDRecon module that takes subdetector specific information such as dE/dx or Cerenkov angles and converts it to particle likelihoods based on models of expected values for different particle hypotheses. The PIDCombiner module uses PID information from all available subsystems and combines it into a single best-estimate for particle type and provides a numerical

---

[2] Abstract Interfaces for Data Analysis, http://aida.freehep.org/

value for the confidence in the ID assignment. This information is made available to the user analysis code (PIDAnalyzer). If desired, a final module PIDEfficiencyPurity will calculate and print the efficiency and purity matrices for the analysis.

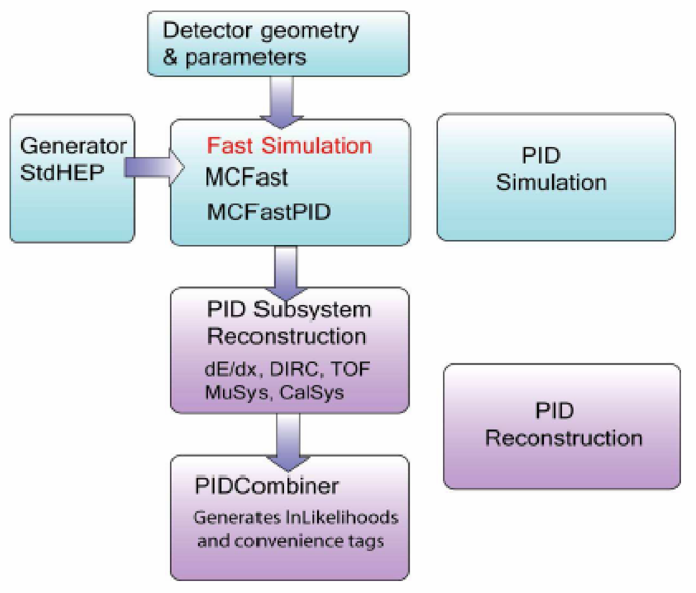
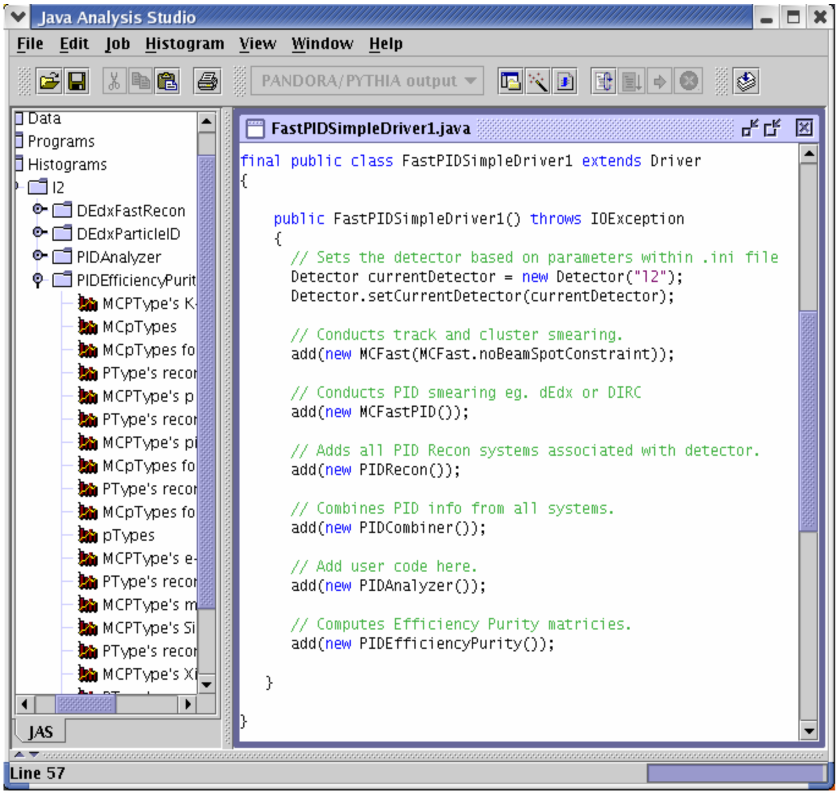


Figure 1: Program flow for particle ID software package, and the separation of simulation from reconstruction.



```
final public class FastPIDSimpleDriver1 extends Driver
{

   public FastPIDSimpleDriver1() throws IOException
   {
     // Sets the detector based on parameters within .ini file
     Detector currentDetector = new Detector("l2");
     Detector.setCurrentDetector(currentDetector);

     // Conducts track and cluster smearing.
     add(new MCFast(MCFast.noBeamSpotConstraint));

     // Conducts PID smearing eg. dEdx or DIRC
     add(new MCFastPID());

     // Adds all PID Recon systems associated with detector.
     add(new PIDRecon());

     // Combines PID info from all systems.
     add(new PIDCombiner());

     // Add user code here.
     add(new PIDAnalyzer());

     // Computes Efficiency Purity matricies.
     add(new PIDEfficiencyPurity());

   }

}
```

Figure 2: Program flow for particle ID software package, and the separation of simulation from reconstruction.

Software Design and Infrastructure

To allow the user to control all the parameters of simulation and reconstruction from a single driver file, without having to modify the source code, we have developed several control classes. The abstract class AbsPIDSystem can be extended to allow the user to define a detector and build all the necessary PID components based on the detector design. An instantiated PIDSystem defines a set of contributing systems and a detector name and provides the functionality to add, *post facto*, subsystems not implemented in the original design. The use of these classes allows a user to create a set of detector subsystems and add the output of the subsystem simulation or reconstruction to the event reconstruction data.

The detector geometry for PID subsystems is no longer hard-coded into the software, as in earlier versions, but rather is provided through data files consistent with other LCD detector geometry files. New parameters have been added to the geometry files that are specific to the particle ID module, and provision has been made for others to be added in future simulation modules. Several of these parameters can be modified at run time through convenience methods, allowing the users to loop over various subsystem combinations such as dE/dx alone, then dE/dx plus Time-of-Flight.

The PIDInfo class is the primary class through which the user extracts and stores particle ID information for an event. It contains all relevant particle ID information such as: lnLikelihood differences, lists of contributing subsystems, reconstructed best-ID (“goodness”) parameters, and the various convenience tags (*isaKaon*, *notaPion* etc.). The information stored in PIDInfo can be accessed through the ReconstructedParticle class after all Particle ID information has been combined using the PIDCombiner. The ReconstructedParticle class was developed in collaboration with SLAC LCD detector group. The user can use this information to create their own selection algorithms and likelihood cuts. Subsystems can be enabled or disabled at run time by the user, which is useful for studying PID algorithms.

The user has the ability to incorporate their own analysis routines and PID subsystems into the structure. This is done by the standard JAS procedure of extending a Driver (or AbstractProcessor) and including a process method to make the LCDEvent available. A feature of the package is that the user can choose either default PID selection parameters, such as minimum ID thresholds (e.g. 2-sigma pi-K separation) for the various subsystems, or set thresholds themselves. In order for users to simulate subsystems not currently implemented into the package they must create classes that extend these base classes, as well as modify the detector.ini files to include detector geometry specific parameters. Example simulation classes are provided, including a module SmearDEdx that has user-selectable models of track energy loss in gas-based detectors (e.g. Sternheimer model and the Yamamoto et al. parameterization [10]). There is also an example module for fast simulation of a specialized hadron ID system SmearDIRC (developed by Wilson). Developing additional modules based on these examples is straightforward.

By default the driver module MCFastPID will check for available subsystems through the detector.ini file, but additional “smearing” drivers can be passed to MCFastPID directly by the user. Figure 3, shows the code structure of the fast tracker simulation and illustrates one aspect of the kinds of modern code design tools used for the project.

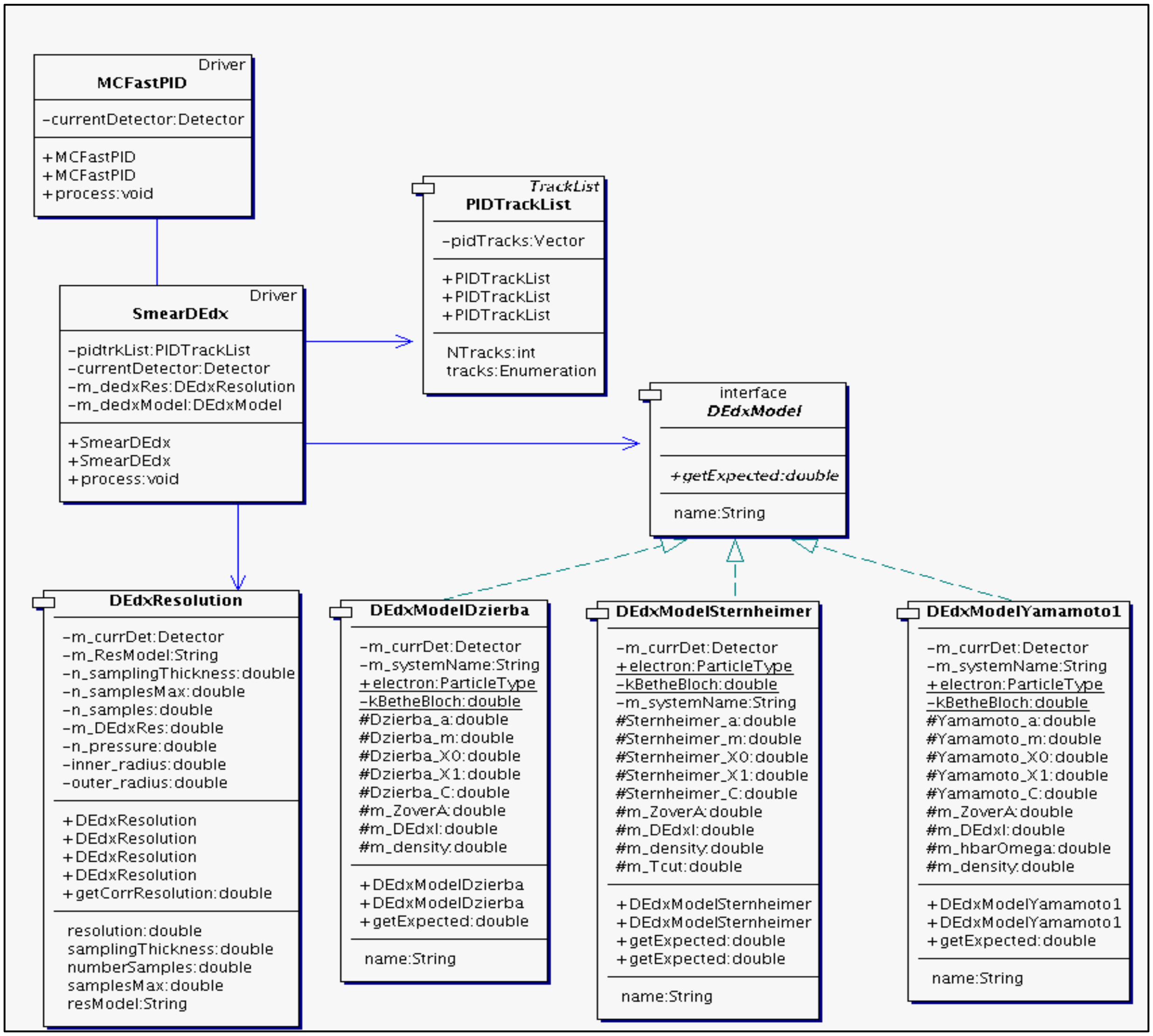


Figure 3: Code structure for fast simulation of energy loss in a tracker – the user may instantiate one of several energy loss models that have been implemented (Dzierba, Sternheimer, Yamamoto), or add their own.

Example uses and results of the PID Package

In this section we illustrate the functionality of the package by presenting a representative selection of histograms that may be created. They are produced from a dataset of ILC Monte Carlo top quarks events generated at 500 GeV center of mass energy (except for Figure 7, which came from a Z-Higgs dataset).

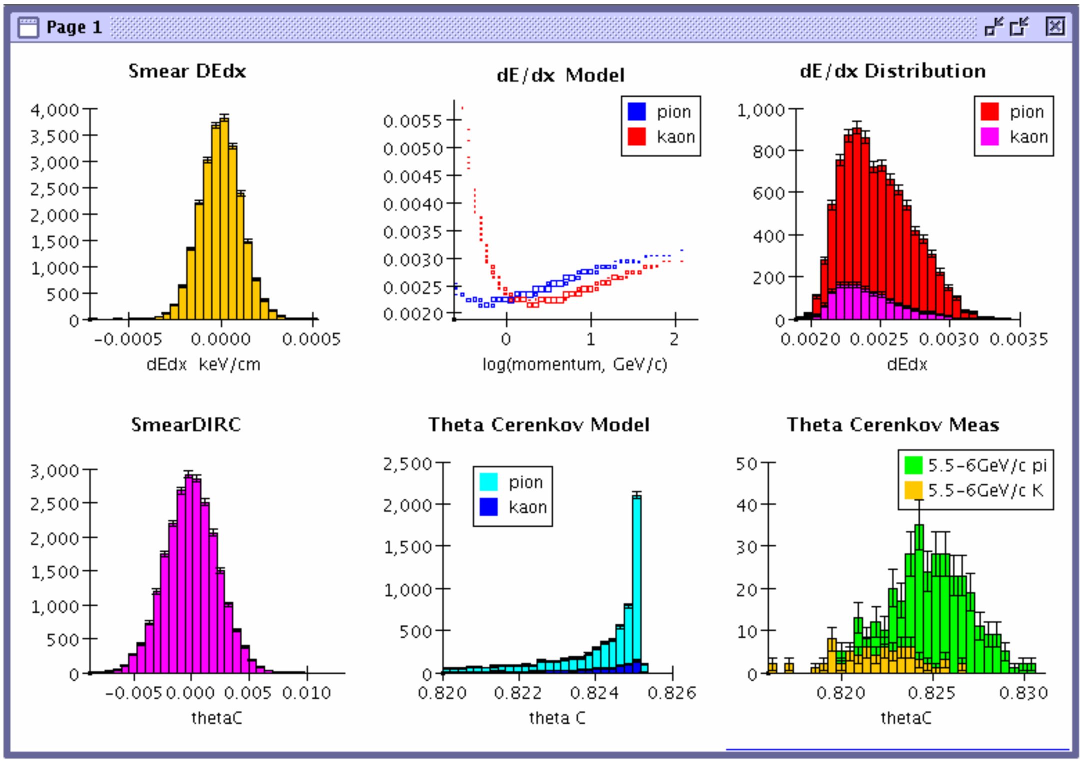


Figure 4: Example distributions from the fast PID simulation. Top row – gas-based tracker energy loss resolution, energy loss as a function of momentum, and comparison of energy loss for pions and kaons. Bottom row – a compact Cerenkov system (DIRC) fast simulation of the Cerenkov angle resolution, value for pions and kaons for the entire spectrum, and comparison for pions and kaons in a limited momentum range.

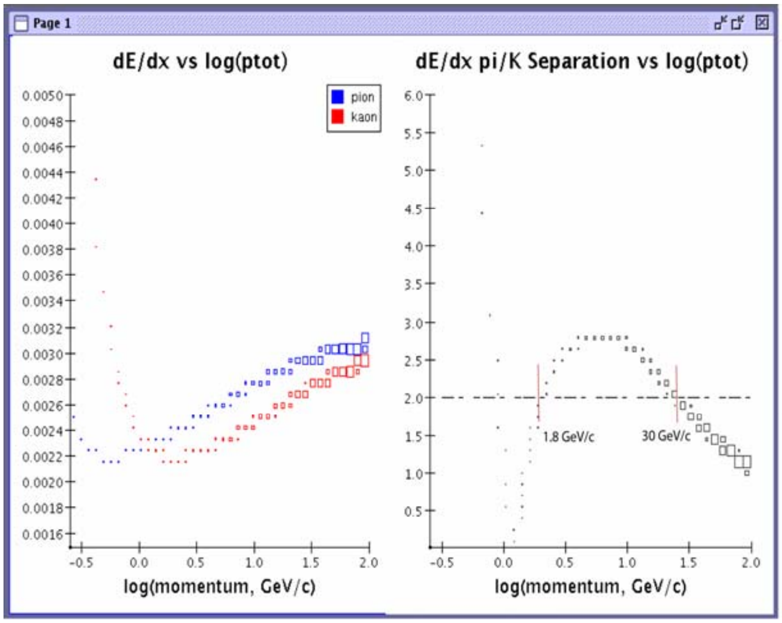


Figure 5: Energy loss (left) and separation in units of detector resolution sigma (right) for pions and kaons as a function of momentum in a gas-based tracker with energy loss resolution of 4.5%.

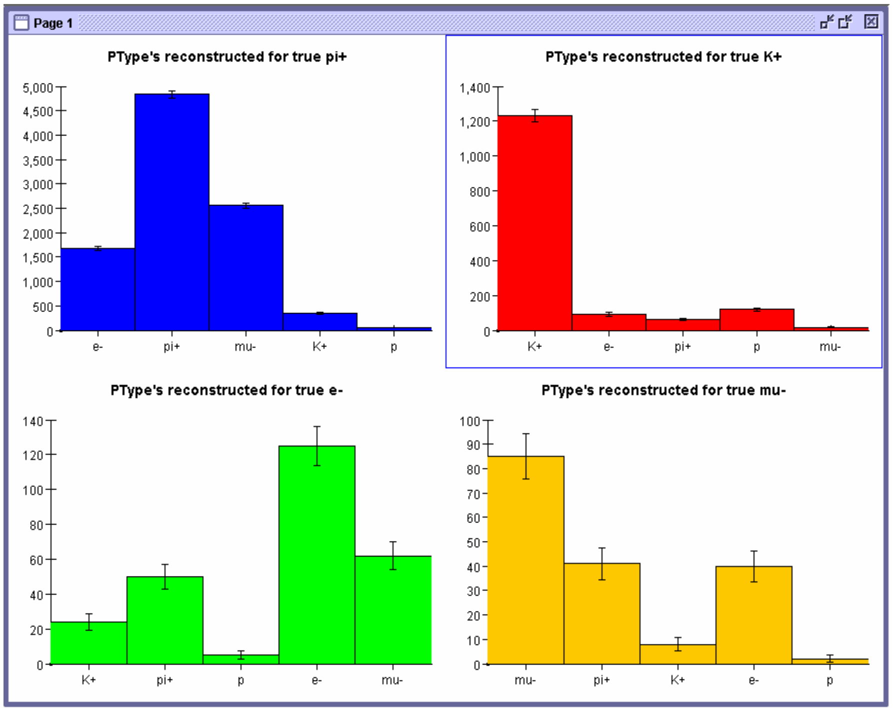


Figure 6: Reconstructed particle types (“PTypes”) from fast simulation. The histograms show the particle identity assigned by the reconstruction code based on user criteria for ID “goodness” (likelihood differences) for true pions (top left), kaons (top right), electrons (bottom left), and muons (bottom right).

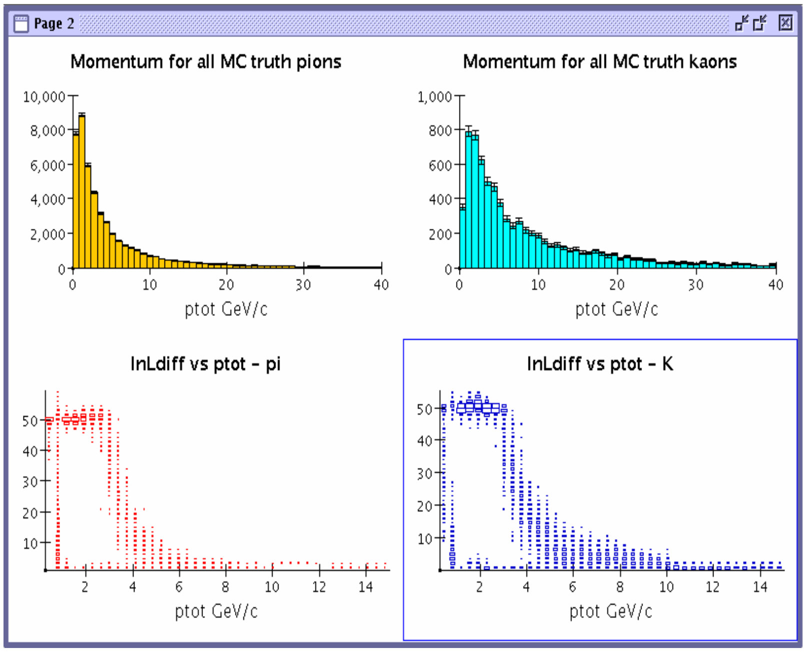
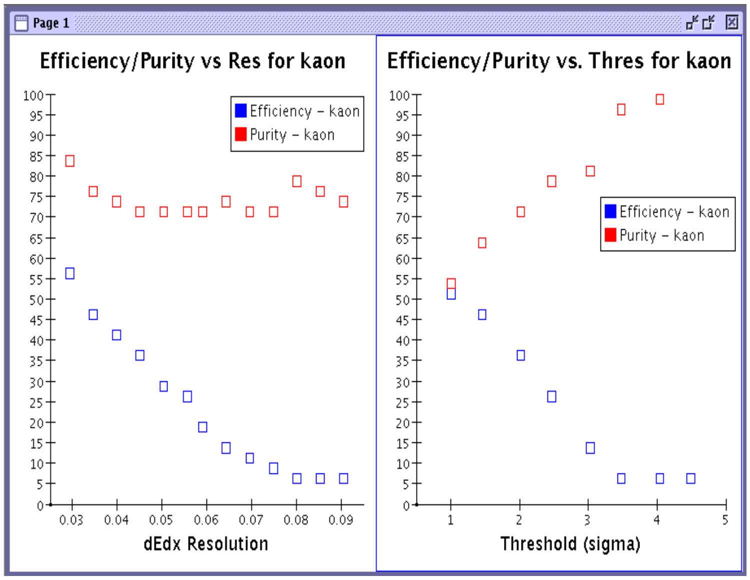


Figure 7: Momentum distributions (top) and reconstructed log-likelihood difference (bottom) between the best ID hypothesis and the next best hypothesis for true pions (left) and kaons (right). The plateau at a lnLikelihood difference of 50 (~10 sigma separation) is imposed by the package as a practical cutoff.


Efficiency/Purity vs Res for kaon

Efficiency/Purity vs. Thres for kaon

Figure 8: Efficiency and purity for kaons as a function of the achievable energy loss resolution in a gas-based tracker (left) and the efficiency and purity for kaons as a function of the user set separation threshold (in Gaussian sigma). These plots illustrate the flexibility of the package to allow the user to loop over detector performance parameters to investigate the effect of design decisions on PID performance.

Physics Requirements for Particle ID

To ensure the success of the Linear Collider program, the detector capabilities needed to address different physics scenarios should be well-understood before significant resources are spent on detector R&D. For example, detector designs should not be optimized assuming that hadron ID is not required before there has been a thorough study of the physics that may be lost due to this assumption. Though preliminary studies found no obvious need for hadron ID, it is clear that more time and thought should be invested to understand these questions. In some cases, improvements to our previous work are obvious. For example, Soffer’s study of the use of proton ID to detect R_p and baryon number violating neutralino decays should be extended to the lower-background center-of-mass energies below the t-tbar threshold, and repeated with different SUSY parameters. Wilson’s b-tagging and single particle ID studies should be extended to higher energies and integrated with the work of the Linear Collider Flavor Identification collaboration. Similarly, the significance of other PID requirements, such as low-momentum lepton ID, must be determined in coordination with detector subsystem and physics working groups.

In parallel with the task of identifying physics processes that might benefit from hadron ID, there should not only be an evaluation of the hadron ID potential of gas and silicon-based tracking systems, but also of specialized detectors, such as scintillator time-of-flight or quartz-based Cerenkov ring-imaging devices. The state-of-the-art in the technology used in such systems has advanced considerably in the last few years so that compact PID systems are more feasible than it appeared in the past.

The detrimental effect of introducing an additional subsystem into a detector needs to be investigated carefully. A study on photon resolution degradation in the calorimeter due to additional material was been performed by G. Bower [11] gave reason for optimism, but it did not include the effect on particle flow algorithms that are now an essential component of several calorimeter designs. A set of benchmark physics processes is needed to allow a quantitative comparison of the loss or gain associated with different technology choices.

Facilities, Equipment and Other Resources

The High Energy Physics group at Colorado State University maintains a cluster of more than 20 LINUX computers and data servers providing 2 Tb of RAID storage. The HEP laboratory is equipped with a standard data acquisition and other equipment needed for photodetector testing. In 2006 the group will receive university funds to update both the computing facility and photodetector applications lab.

Year 1 (2006/07) Project Activities and Deliverables

TASK I. SIMULATION

We will: convert the existing fast JAS v2.2.5 PID simulation and reconstruction package to allow its integration with existing core code-base (and continue to adapt it as the implementation of the core packages are refined by the central software groups at SLAC and elsewhere); refine existing subsystem simulation and reconstruction as needed (e.g. updated DIRC simulation based on feedback from the Ratcliff group); add new basic fast simulations for other PID systems, such as Time-of-Flight, as needed.

TASK II. USE OF HADRON ID FOR LC PHYSICS

We will: perform a broader study of the physics justification for hadron identification – this will include extensions to our previous studies to other energies, and a broader range of physics channels; help to generate a list of benchmark physics processes for the physics working groups to use for quantitative comparisons of the capabilities and negative effects of particle ID technologies.

TASK III. PID COORDINATION ACROSS SUBSYSTEMS

We will: work with detector concept groups to develop a definition of the software interface and infrastructure issues related to heavy particle identification; re-establish and maintain a web site for communication of particle ID developments.

TASK IV. PLANNING FOR SPECIALIZED HADRON ID SUBSYSTEM

We will: do first stage planning for a potential hadron ID subsystem R&D program to be started in Year 2 if such a system is indicated by the preliminary results of tasks I and II.

Deliverables will include reports on each of these categories at ILC workshops or other meetings.

Year 2 (2007/08) Project Activities and Deliverables

Task I. Continuation of tasks I, II, and III of year 1, but to include final recommendations of the PID investigation to the Global Design team.

Task II. If there is sufficient indication from Year 1 of the need for a specialized hadron ID subsystem, such as a next generation DIRC or TOF system, a significant R&D program should be started in Year 2. The specific program will depend on the outcome of Year 1 planning (Task IV).

Budget and Budget Justification

Table 1: Project budget. Year 2 entries are only estimates, and do not include M&S and equipment (see text for further explanation).

	FY2006/07 (k$)	*FY2007/08 (k$) Estimates*
1 FTE post doctoral researcher (12 mths)	46.0	47.8
fringe@20.3%	9.3	9.7
Graduate student (12 mths incl. fringe)	18.6	19.3
fringe@3.6%	0.7	0.7
Technical support*	4.2	19.6
Domestic Travel:	5.4	6.6
International Travel	5.0	5.2
Equipment	0	tbd
M&S	0	tbd
Tuition (2 semesters non-resident)	15.0	16.2
Total direct costs	104.2	125.1
Indirect cost @ 46% (excl. tuition)	41.0	50.1
Total:	**145.2**	**175.2**

*Technical support is provided at an hourly rate of $53 in AY2006 and estimated with 4% inflation in AY2007. This charge includes personnel salary fringe rate of 20.3%.

The major expenses in both years 1 and 2 are for personnel. The primary impediment to progress on the hadron ID issues has been the lack of full-time effort at the level of a PhD. Providing support for a graduate student in addition will not only increase the number of topics that can be addressed but also provides the opportunity for training future personnel and give additional support for performing this activity in a university setting. The tuition amount is for a full-time non-resident graduate student.

The technical support request is for project engineer and HEP lab manager David Warner. In year 1, we request 80 hours, which we estimate is the time needed for Warner to assist in the planning for an R&D program the following year. This time includes one trip to SLAC to consult with SLAC collaborators. In the second year, we estimate the need for two months of effort from Warner, however, this request and that for M&S and equipment for such an R&D effort will be revisited in a revised Year 2 proposal based on the outcome of the year 1 studies.

Travel support in year 1 is estimated at 6 domestic trips at $900/trip. These include 2 trips each to ILC workshops for the PI and post doc, 1 such trip each for the student and project engineer. An additional trip is budgeted in year for the engineer. Two international trips at $2500/trip are budgeted to allow participation of the PI and post doc in ILC workshops in Europe or Asia.

Fringe benefits account for less than 10% of the request. The indirect cost (IC) rate is 46%; no IC is applied to tuition. Inflation at 4% has been used in the estimated of Year 2 costs.

References & Related papers and talks by the proposers

[1] Wilson, R.J., “Report from the Particle ID/Muon”, Santa Cruz Workshop 2002.

[2] Wilson, R.J., “Hadron ID in the S2 Detector”, Santa Cruz Workshop 2002.

[3] Rolnick, S., Wilson, R.J., “PID Software for Linear Collider Detector Studies”, Linear Collider Workshop, Victoria, July 2004.

[4] http://hep45.hep.colostate.edu/~wilson/flc/pid/pidmain.html

[5] http://hepwww.rl.ac.uk/lcfi/

[6] Mercadante, P.G. and Yamamoto, H., “Analyses of Long Lived Slepton NLSP in GMSB Model at Linear Collider,” in 1999 Sitges meeting proceedings: *Physics and Experiments with Future Linear Colliders*, edited by E. Fernandez and A. Pacheco, Universitat Autonoma de Barcleona/IFAE Publications, 2000.

[7] Wilson, R.J., “Some Thoughts on Hadron Identification with Linear Collider Detectors,” in 1999 Sitges meeting proceedings, see ref. [3]; Wilson, R.J., “B^0 Tagging with Kaons,” in the FNAL meeting proceedings: *Physics and Experiments with Future Linear Colliders*, edited by A. Para.

[8] Soffer, A., SNOWMASS-2001-E3039, APS/DPF/DPB Summer Study on the Future of Particle Physics (Snowmass 2001), Snowmass, Colorado, 2001; Soffer, A., “A Look at R_p and Baryon Number Violation Using Hadron ID” – Santa Cruz Workshop 2002.

[9] http://hep45.hep.colostate.edu/~rolnick/pid.html

[10] Yamamoto, H., “dE/dx Particle Identification for Collider Detectors,” 1999 Sitges meeting proceedings, ref. [6]; Hauschild, M., “Particle Identification using a TPC,” in the FNAL proceedings: *Physics and Experiments with Future Linear Colliders*, edited by A. Para.

[11] Bower, G., “The Effect of a DIRC on EMCal Resolution”, presented by Wilson at the Santa Cruz Workshop 2002.