

**Particle Flow Studies  
with the Silicon Detector (SiD) at the  
International Linear Collider**

Calorimetry

The University of Iowa

Matthew J. Charles, Wolfgang F. Mader, Usha Mallik, and Niels T. Meyer

Contact Person: Usha Mallik  
E-mail: Usha.Mallik@uiowa.edu  
Phone: (319) 335-0499

Year 1: \$83,910  
Year 2: \$98,849

January 2005

# Particle Flow Studies with the Silicon Detector (SiD) at the International Linear Collider

Matthew J. Charles, Wolfgang F. Mader, Usha Mallik, and Niels T. Meyer  
(The University of Iowa)

January 2005

## 1 Introduction

In 2004, The University of Iowa group has concentrated their Linear Collider R&D activities on the simulation of the electromagnetic calorimeter (EMCal) for the Silicon Detector (SiD) at the International Linear Collider (ILC).

An excellent jet energy resolution is essential in order to achieve maximum sensitivity to the physics anticipated from the ILC. It is an important benchmark for the detector design in general and for the design of the calorimeter in particular, e.g. optimization of the granularity. A finely segmented detector should be matched by powerful reconstruction software for which a Particle Flow Algorithm (PFA<sup>1</sup>) is a promising approach.

The group, led by Usha Mallik, has been an integral part of simulation efforts at SLAC, working mainly on reconstruction software necessary for the proof of principle of a PFA. The milestones achieved in the year 2004 are discussed, focusing on the contributions from the Iowa group. The future plans for simulation efforts are laid out.

---

<sup>1</sup>also referred to as 'Energy Flow Algorithm'

## 2 The Importance of the EMCal in a PFA

Calorimetry plays an essential role in a PFA in the SiD design. The current goal of the Iowa group is the development of an algorithm to detect and trace minimum ionizing particles (MIPs) in the EMCal, and to reconstruct associated showers without assistance from the central tracking system. There are several benefits from this approach, three of which are discussed below.

1. Electromagnetic and hadronic showers can be separated by analyzing their shower shapes. However, having a longitudinally segmented calorimeter, the actual starting point of the shower and the presence or lack of a MIP-track associated with it can be used as an additional distinctive feature to increase the separation power.
2. For charged particles in the forward angular region, the number of hits in the central tracking system is limited by the detector geometry. The extrapolation of identified MIP-tracks to the last tracking layers can be used to assign unassociated tracker hits and thus improve the tracking efficiency.
3. Charged tracks produced in decays of long-lived particles like  $K_S^0$ ,  $\Lambda$ , or possible SUSY particles only leave a limited number of hits in the central tracking system. The identification of MIP-tracks can help to improve the detection efficiency in the same way as described above.

Efficient and precise reconstruction of MIP-tracks in the EMCal are important steps towards identification of showers from neutral hadrons and reconstruction of long-lived particles.

## 3 Personnel Involved

The SiD effort at the ILC is lead by Marty Breidenbach (SLAC). The University of Iowa group has been working closely with his group in the simulation effort for this detector design since 2003 and has made essential contributions towards the proof of principle of a PFA.

At present, Wolfgang Mader, a postdoc from The University of Iowa, is working full-time on this project together with the simulation software group at SLAC: Norman Graf, Ron Cassel and Tony Johnson.

Niels Meyer, a new postdoc who has just joined the Iowa group having completed his Ph.D. at the University of Hamburg, will work full-time on the ILC effort. Matthew Charles, a postdoc previously working on this project, is also rejoining the effort.

## 4 Simulation Overview

### 4.1 SiD Configuration

The simulation efforts are based on the SiD design which is described in detail elsewhere [1]. The EMCal, which is the main focus of the present effort of the Iowa group, consists (in the barrel region) of 30 layers of  $(0.5 \times 0.5) \text{ cm}^2$  silicon cells alternating with 0.25 cm thick tungsten layers. It extends from a radius of  $r_{\min} \approx 127 \text{ cm}$  to  $r_{\max} \approx 142 \text{ cm}$ , representing a total of 20 radiation lengths ( $X_0$ ) or 0.8 hadron interaction lengths ( $\lambda_I$ ). In combination with an effective PFA, the high granularity will help provide an excellent jet energy resolution.

### 4.2 Technical Details

In 2004, the detector simulation used at SLAC has been migrated from a GISMO [2] based simulation to a simulation based on GEANT4 [3]. The configuration of the detector is described using the XML language. The GEANT4 package is widely used in high energy physics and provides realistic description of the response of the individual detector components. The simulation packages previously developed by the Iowa group had to be migrated to work with the updated detector simulation.

The reconstruction and analysis software is written in JAVA, using the `hep.lcd` framework [4]. Simulated events used are stored in the SIO format [5]. The analysis is presently carried out using JAVA Analysis Studio version 3 [6] (migrating from version 2) and ROOT [7].

### 4.3 Event Samples

The SLAC group provided event samples to develop the necessary reconstruction tools and to test the overall performance of the SiD design. The University of Iowa group was involved in the testing of event samples and the debugging of the digitization process.

The event samples used in the simulation effort comprise two categories:

- Single particle samples:  
The detector response is simulated for (a) charged pions in a momentum range of  $1 \text{ GeV} < p_\pi < 10 \text{ GeV}$  and an angular range of  $4^\circ < \Theta < 176^\circ$ , and (b) events containing decays  $K_S^0 \rightarrow \pi^+ \pi^-$  of momentum  $10 \text{ GeV}$ . These samples are used to develop and test the algorithms described later in the document.
- $Z^0$  and  $t\bar{t}$  events:  
These physics events were used in order to test the algorithms under realistic

conditions, an important step towards developing an effective PFA. These tests will be extended to other event topologies.

#### 4.4 Future Developments

The SLAC group is working towards implementing a simulation package which will produce output in LCIO format [8]. LCIO is a general format agreed on by the ILC community in the US, Europe and Asia. It allows a direct comparison of the different detector designs and the various simulation efforts. A common format makes it possible to share reconstruction and analysis algorithms. These tools for the new framework are currently developed by the SLAC software group (their effort will be joined by The University of Iowa group).

### 5 Simulation Effort at The University of Iowa

In this section, the simulation effort of The University of Iowa group is explained in detail. The group has focused on the development of reconstruction software for the EMCal, an important ingredient of a PFA.

In the first step, the identification of MIP-tracks in the EMCal within the busy environment of multi-hadron final states was developed. This part has been finished in 2004 and was presented to the ILC community on several occasions [9, 10].

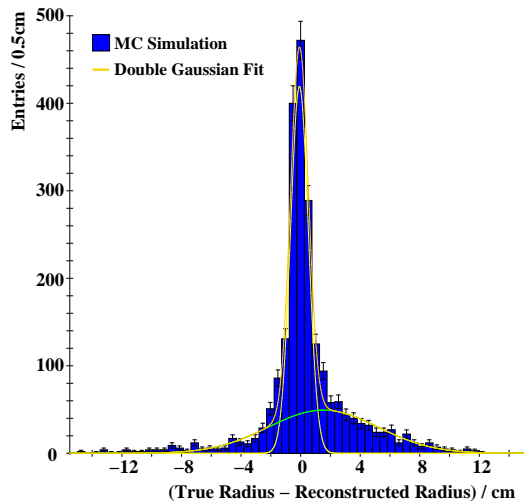
The next steps for the PFA are measurement and identification of electromagnetic and hadronic showers. The separation of contributions from charged and neutral hadrons is an important step. This is under development and the current status is presented.

#### 5.1 Identification and Reconstruction of MIP-Tracks

MIP-tracks leave a very characteristic signature of isolated energy deposits (occasionally accompanied by a close-by second energy deposit due to  $\delta$ -rays) in the EMCal. Starting from a seed point which is obtained from the track information in the central tracking system, they can be tracked with high efficiency as shown in previous studies by Matthew Charles in 2003.

In 2004, this algorithm has been verified using GEANT4 simulated events. Criteria for identifying MIP seeds independent of the central tracking system have been developed and tested. The algorithm has been further extended using pion samples and events containing  $Z^0$ -decays and  $t\bar{t}$  physics processes.

The efficiency of the algorithm exceeds 99% for pions that only leave a MIP-track in the EMCal. In the current detector design, however, about half of the charged hadrons start showering within the EMCal. For particles with transverse



**Figure 1:** Resolution of the radial position of the showering point using single pion samples. Superimposed is a fit with a double Gaussian, where the two individual contributions are shown separately. The resolution of the radial position as obtained from the core Gaussian is  $\pm 0.6$  cm corresponding to a reconstruction uncertainty of  $\pm 1$  layer in the EMCal. Physics processes contributing to the tails on the left and the right side of the distribution have been identified as  $\pi \rightarrow \mu\nu_\mu$  decays and scattering processes, respectively.

momentum in excess of 1 GeV (i.e. sufficient to reach the barrel EMCal) an overall reconstruction efficiency of about 90% was achieved.

The analysis has been further extended to the reconstruction of the showering point. A precision of  $\pm 0.6$  cm in the radial position has been achieved as illustrated in Figure 1, corresponding to  $\pm 1$  layer in the EMCal. The tails of the distribution extending to the left and the right side of the peak have been studied and identified as events containing  $\pi \rightarrow \mu\nu_\mu$  decays and scattering processes, respectively.

## 5.2 MIP-Track Fitting

### 5.2.1 Reconstruction of MIP-Segments

The seed for a MIP-candidate is initially defined as three isolated hits in four consecutive EMCal layers. In very busy events e.g. from  $t\bar{t}$  production, this approach leads to too many reconstructed MIP-track segments. This is partly due to the overlap of MIP-tracks with hadronic or electromagnetic showers from other particles.

Therefore, a two-pass algorithm has been developed. The algorithm is applied once, starting from the inner part of the detector, following the MIP-track as far as

isolated hits are found in successive outer layers. In a second pass, the algorithm starts from the outer part of the detector, following the MIP-track as long as isolated hits are found in successive inner layers of the EMCAL<sup>2</sup>. These MIP-segments are then matched using  $\chi^2$  criteria described below.

### 5.2.2 Association of MIP-Segments

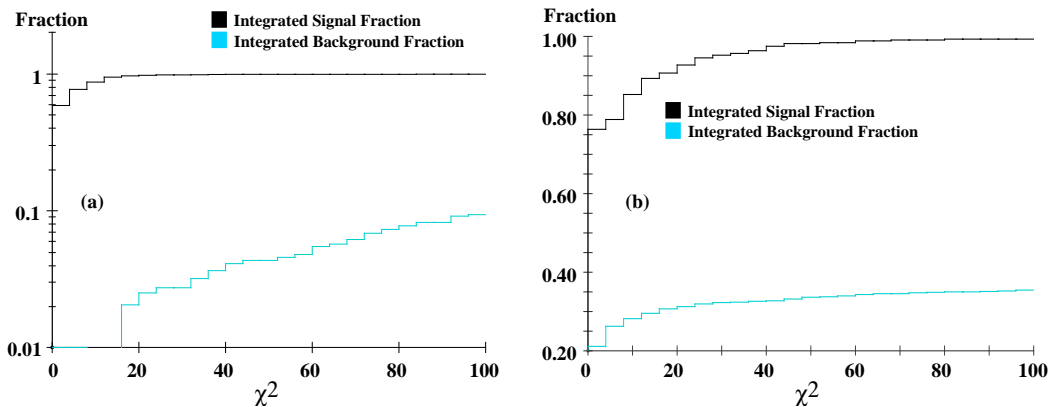
A five parameter helix fit is performed to the EMCAL hits of each MIP-segment found. For every pair of MIP-segments, a two-step  $\chi^2$  criterion is calculated using the fit parameters and the corresponding error matrices.

In the first step, only the track parameters ( $\kappa, \phi_0, d_0, \tan \lambda$ ) are used and the efficiency (upper curve) for matching MIP-segments and the background rate from random combinations (lower curve) is displayed in Figure 2(a). By retaining all combinations with  $\chi^2 < 40$ , an efficiency of  $\approx 100\%$  is achieved with a background fraction below 5%. The  $z$  component is not included in the first step, since its contribution exhibits artifacts due to the projective geometry of the EMCAL configuration used in the simulation (for example, when a MIP passing through the EMCAL in one tower of cells is scattered into a neighboring tower).

Instead, in a second step, a cut of  $\chi_z^2 < 80$  is used, where  $\chi_z^2$  is calculated from the  $z$  component of the helix parameters only, on those MIP-combinations that pass

---

<sup>2</sup>In order to account for detection inefficiencies, gaps of one layer are allowed.



**Figure 2:** Integrated  $\chi^2$  distributions as calculated using the helix parameters ( $\kappa, \phi_0, d_0, \tan \lambda$ ) (a) and  $z$  (b). The upper curves indicate the efficiency for associating matching MIP-segments, while the lower curves indicate the background rate from random MIP segment combinations. The background fraction in (b) is relative to that remaining after requiring  $\chi^2(\kappa, \phi_0, d_0, \tan \lambda) < 40$ .

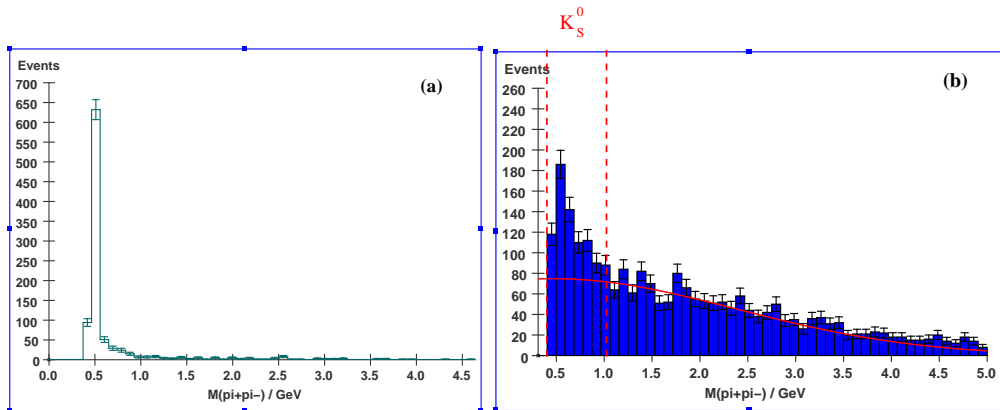
the first selection step just described. This rejects 2/3 of the remaining background while maintaining high efficiency as shown in Figure 2(b). Pairs of MIP-segments satisfying these criteria are then combined to a single MIP.

### 5.2.3 The Use of Tracking System Layers in Calorimeter Tracking

The reconstruction of particles like  $K_S^0$ ,  $\Lambda$ , or possible other long lived particles predicted from theories beyond the Standard Model is challenging in the current SiD design. In a typical  $Z^0$  event, a  $K_S^0$  with an average energy of 5 GeV has a mean decay length of about 25 cm. Therefore, the charged pions from  $K_S^0$  decays would not leave any hits in the vertex detector, and would therefore not be reconstructed by the current tracking algorithm.

The tracking of MIP particles in the EMCAL, established by The University of Iowa group (see Section 5.1), can help increase the reconstruction efficiency of long-lived particles. The helix fit to the EMCAL hits generated by MIP particles is extrapolated into the central tracking system. Using the extrapolated track, unassociated hits (i.e. hits which have not yet been used for track reconstruction by the standard tracking algorithm) are picked up with high efficiency. The extrapolation algorithm has a resolution of better than 1 cm for the last tracking layer, whereas the average distance between unassociated hits in  $Z^0$  events is of the order of 5 cm. Therefore, they can be associated to MIP-tracks with very high efficiency.

After picking up the hit from the last tracking layer, the track is refitted and the new track is then extrapolated to the second-to-last tracking layer. In case another



**Figure 3:** Invariant  $K_S^0$  mass distribution using MIP assisted tracking as described in the text. Plot (a) shows the invariant mass distribution as obtained from single  $K_S^0$  events, plot (b) shows the corresponding distribution as obtained from  $Z^0$  events. In (b) the  $K_S^0$  is sitting on top of combinatorial background.

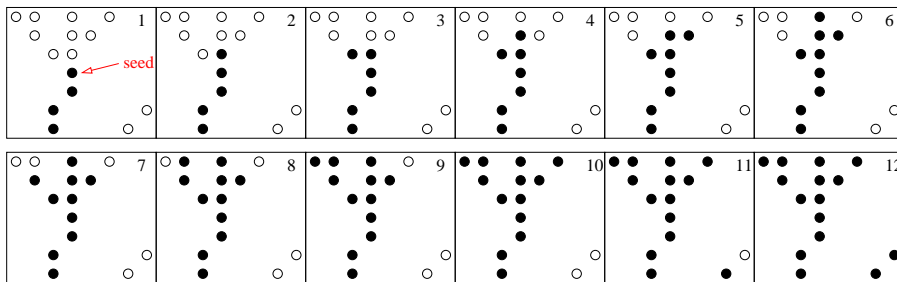
unassociated hit is found, it is added to the MIP hits and the track is refitted again.

In order to demonstrate the benefits from calorimeter tracking, pion tracks from neutral kaon decays were reconstructed as described above and fitted in pairs to a common vertex to form the  $K_S^0$  candidates. The results of the fits are illustrated in Figures 3 (a) and (b). Figure 3 (a) shows the  $K_S^0$  invariant mass for events from dedicated  $K_S^0$  event samples; a clear signal is obvious at the nominal  $K_S^0$  mass. The same algorithm is applied to  $Z^0$  events and the result is shown in Figure 3(b). While the  $K_S^0$  peak is reconstructed at the nominal mass, it has a broader mass distribution on top of a large combinatorial background.

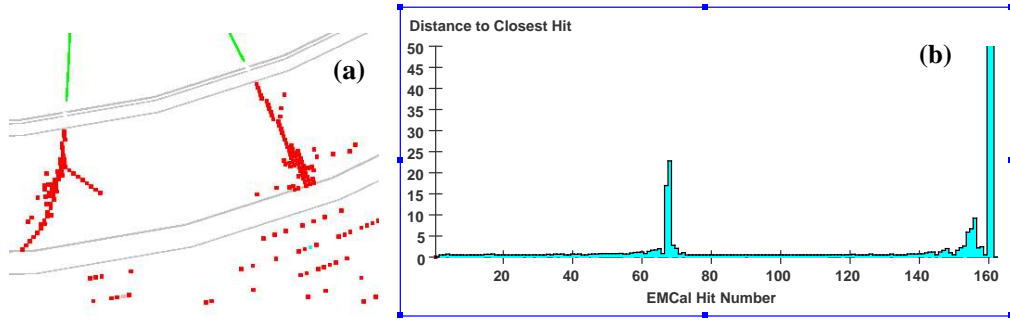
### 5.3 Reconstruction of Showers in the EMCal

Currently, the primary focus is on the reconstruction of showers from charged and neutral hadrons in the EMCal. To be able to study the hadronic showers in a clean environment, energy deposits from the electromagnetic showers are removed, presently using truth information. This procedure will be replaced by a more realistic algorithm and the Iowa group will participate with the SLAC group in this development.

The clustering algorithm currently being developed by the Iowa group is based on the Minimum Spanning Tree (MST) approach. In this approach, energy deposits are assigned to a particular cluster on an individual hit-by-hit basis as illustrated in Figure 4, where 12 consecutive steps of the clustering algorithm are displayed. To identify clusters from charged hadrons, the reconstructed shower points from the MIP-tracking algorithm and unassociated hits in the last tracking layer are used as cluster seeds. Starting from these seeds, the geometrical distance to all EMCal



**Figure 4:** Illustration of the Minimum Spanning Tree algorithm. Hits assigned to the cluster are shown as solid circles and unassigned hits as hollow circles. At each step the unassigned hits are scanned and the one which is closest to an assigned hit is added to the cluster. If the distance to the nearest hit is large, (# 11), then the hit is assigned to a new cluster.



**Figure 5:** Plot(a) shows the signature of a typical  $K_S^0 \rightarrow \pi^+\pi^-$  decay from a single  $K_S^0$  event. Plot (b) displays the distance to the closest hit not yet assigned to the cluster as explained in the text.

hits not yet associated to any cluster are calculated and the closest hit is added to the cluster. This procedure is repeated until all EMCal hits have been assigned to clusters. Currently, the hit assignment is based on the geometric distance between hits. In future, various weighting algorithms will be studied.

An example of the performance of the algorithm for a typical  $K_S^0$  event is shown in Figures 5 (a) and (b). The section of the detector where the pions from a  $K_S^0 \rightarrow \pi^+\pi^-$  decay interact with the EMCal is shown in Figure 5(a). Figure 5(b) shows the distance to the closest hit not yet assigned to the cluster. The spike around hit number 70, corresponding to a large geometrical distance, is used as a separation criterion between the two pions showers.

The current implementation of this algorithm uses the hit coordinates directly and is therefore largely independent of the detector geometry (e.g. the number of EMCal layers and cell sizes). Hence, with minimal changes it can also be used with other calorimeter configurations.

These studies have very recently been shown to the SiD simulation community [10]. The Iowa group was encouraged to pursue these studies and take a leading role in the MST-based clustering approach; we therefore expect to continue these studies.

## 6 Future Plans

Substantial progress has been made during the year 2004 towards the proof of principal of a PFA with important contributions from The University of Iowa group. However, essential ingredients of such an algorithm are still under development or missing. The Iowa group is playing a central role and is working in close collabora-

tion with other simulation groups that are covering different aspects of a PFA.

Our group is an integral part of the simulation effort at SLAC, working closely with the group of Marty Breidenbach, who is leading the SiD effort. The simulation tools developed by the Iowa group are aimed at the SiD design. We plan to intensify our efforts in this area. The achievements of our group are summarized in this section and the prospects for a future involvement, in both the development of a PFA and its evaluation studying benchmark physics cases, are laid out.

## 6.1 Software Development and SiD Design Studies

The Iowa group has made considerable progress in the development of reconstruction software for the EMCal. A very efficient MIP-tracking algorithm has been developed. A track fitting algorithm using only calorimeter hits has been developed; its benefits for particle tracking in general have been demonstrated by reconstructing long-lived particles with the use of the last layers of the central tracking system. We intend to carry out studies with other EMCal geometries and investigate their influence on the algorithms developed.

Reconstruction of hadronic showers in the EMCal has started and the first results are encouraging. We now intend to intensify these studies with the participation of Matthew Charles and Niels Meyer.

## 6.2 Study of Physics Processes

A study of the sensitivity of particular detector designs and reconstruction algorithms to physics observables as well as their impact on the accuracy of specific physics analyses is essential. We propose to participate in this effort, concentrating in particular on processes where the reconstruction software developed by the Iowa group can be applied, i.e. on processes leaving distinct signatures in the EMCal.

The Iowa group proposes to investigate the following benchmark processes where an excellent understanding of the calorimetry section is essential and which are of major physical interest at the ILC.

1. Higgs decays to photons,  $H \rightarrow \gamma\gamma$

The distinct signature of two isolated photons allows studies of the energy resolution of single high energetic particles from reconstruction of EMCal hits only. This decay is of special interest to exploit the Higgs-photon coupling, which is highly sensitive to heavy charged particles beyond the SM particle spectrum.

2. Higgs production in WW-fusion with decays  $H \rightarrow b\bar{b}$

The signature of two final state jets allows resolution studies of jet energy and jet-pair mass measurements. WW-fusion is the dominant Higgs production mechanism at high energies and probes the Higgs-W coupling. The latter is of special interest when deducing the total decay width in case of light Higgs bosons.

### 3. Double Higgs production ZHH

Events from this process may contain up to six jets, allowing to study the impact of calorimeter measurements on combination efficiencies. This process is essential to study the Higgs self coupling and to establish the Higgs mechanism as the origin of electroweak symmetry breaking.

## 6.3 Plans towards a Particle Flow Algorithm

The long-term goal of the Iowa group is to play a key role in the integration of the various components of a PFA, which takes full advantage of the natural interplay between all sub-detectors. Such an algorithm is best suited for optimization studies of the SiD design, e.g. detector geometry or calorimeter segmentation.

The University of Iowa group has been an integral part of the studies so far by providing an efficient MIP-tracking algorithm and working on general EMCAL reconstruction. These efforts will be intensified by Niels Meyer joining the group.

## 7 Resources and Budget

Wolfgang Mader has worked almost full time on simulations and algorithms for developing Particle Flow for SiD since he joined the group in March 2004. Niels Meyer, who graduated from University of Hamburg under the supervision of Professor Rolf-Dieter Heuer (at present Research Director of DESY, Germany), has just joined the group to start working full time on this project. Matthew Charles has completed an analysis on BABAR, and expects to rejoin the Particle Flow effort shortly. Two (equivalent) full time postdocs will be involved in the SiD Particle Flow project from The University of Iowa. A new graduate student will likely join the effort from late summer, after moving to the bay area.

The budget is listed below, primarily based on the support for one postdoc. The postdoc's salary is taken at \$46,000 with fringe benefits at 18% at The University of Iowa. The second postdoc's support is obtained by reallocating resources from the base grant support of the group for BABAR. Travel to the various ILC workshops for the two postdocs and for Usha Mallik is included for the year at \$12,000. The indirect costs at Iowa are charged at a rate of 26.6% for off-campus activities (the

postdocs all stay at SLAC where interaction is easier). A total of \$83,910 is thus requested to cover the FY05 cost.

The second year's budget is anticipated to be similar, with a \$2,000 increase in salary for the postdoc and partial support of \$8,000 for a graduate student. Tables 1 and 2 list the breakdown of the budgets.

**Table 1:** Budget for the First Year R&D. The total costs quoted include 26.6% for off-campus activities.

Item	Direct Cost (\$)	Total Cost (\$)
Salary and Fringe Benefits of Research Associate	$46,000 \times 1.18$	68,718
Travel to ILC Meetings (2 RAs and PI)	12,000	15,192
Total First Year R&D		83,910

**Table 2:** Budget for the Second Year R&D. The total costs quoted include 26.6% for off-campus activities.

Item	Direct Cost (\$)	Total Cost (\$)
Salary and Fringe Benefits of Research Associate	$48,000 \times 1.18$	71,706
Partial Support for a Graduate Student	$8,000 \times 1.18$	11,951
Travel to ILC Meetings (2 RAs and PI)	12,000	15,192
Total Second Year R&D		98,849

## References

- [1] SiD Homepage:  
<http://www-sid.slac.stanford.edu/>
- [2] W. B. Atwood, T. Burnett, R. Cailliau, D. R. Myers and K. M. Storr, "Gismo: An Object Oriented Program For High-Energy Physics Event Simulation And Reconstruction," *Int. J. Mod. Phys. C* **3**, 459 (1992), R. Cassell, "Full LCD detector simulation with GISMO," SLAC-REPRINT-2000-109 *Prepared for 5th International Linear Collider Workshop (LCWS 2000), Fermilab, Batavia, Illinois, 24-28 Oct 2000.*
- [3] S. Agostinelli *et al.* [GEANT4 Collaboration], *Nucl. Instrum. Meth. A* **506**, 250 (2003).

- [4] `hep.lcd` Homepage:  
<http://www-sldnt.slac.stanford.edu/jas/Documentation/lcd/>
- [5] <http://www-sldnt.slac.stanford.edu/nld/new/Docs/FileFormats/sio.pdf>
- [6] JAS webpage:  
<http://jas.freehep.org/jas3>
- [7] ROOT webpage:  
<http://root.cern.ch>
- [8] LCIO Homepage:  
<http://lcio.desy.de>
- [9] Linear Collider Simulation Workshop III, Jun-2004:  
<http://www.slac.stanford.edu/~mader/Talks/LCSIII/LCSIII.pdf>  
Victoria Linear Collider Workshop, Jul-2004:  
[http://www.linearcollider.ca:8080/lc/vic04/abstracts/detector/calorimetry/wolfgang\\_mader.pdf](http://www.linearcollider.ca:8080/lc/vic04/abstracts/detector/calorimetry/wolfgang_mader.pdf)  
ALCPG Calorimetry WG meetings:  
<http://nicadd.niu.edu/cdsagenda//fullAgenda.php?ida=a0450>
- [10] Argonne SiD PFA workshop, Jan-2005:  
<http://nicadd.niu.edu/cdsagenda//fullAgenda.php?ida=a057>

