Innovations in ILC detector design using a particle flow algorithm approach

Stephen R Magill
Argonne National Laboratory, Argonne, IL, USA
E-mail: srm@anl.gov

New Journal of Physics 9 (2007) 409
Received 26 July 2007
Published 14 November 2007
Online at http://www.njp.org/
doi:10.1088/1367-2630/9/11/409

Abstract. The International Linear Collider (ILC) is a future e⁺e⁻ collider that will produce particles with masses up to the design center-of-mass (CM) energy of 500 GeV. The ILC complements the Large Hadron Collider (LHC) which, although colliding protons at 14 TeV in the CM, will be luminosity-limited to particle production with masses up to ~1–2 TeV. At the ILC, interesting cross-sections are small, but there are no backgrounds from underlying events, so masses should be able to be measured by hadronic decays to dijets (~80% BR) as well as in leptonic decay modes. The precise measurement of jets will require major detector innovations, in particular to the calorimeter, which will be optimized to reconstruct final state particle 4-vectors—called the particle flow algorithm approach to jet reconstruction.
1. Introduction

The future e⁺e⁻ International Linear Collider (ILC) will be built to produce and study particles with masses up to its design center-of-mass (CM) energy of 500 GeV [1]. This will complement and enhance the discovery of new physics processes and elementary particles anticipated at the Large Hadron Collider (LHC) which is soon to be commissioned. At the ILC, particles are produced in a relatively background-free environment, so detectors will be able to make precision measurements of not only particle masses, but of branching ratios and coupling parameters as well—opening up the domain of expected new physics in the electro-weak scaling regime. For example, it is expected that Higgs particle production, branching ratios and coupling to other fundamental objects will comprise much of the physics program at the ILC. For the ILC detectors to make precision measurements, full use of the hadronic decay modes of vector bosons and Higgs particles will be necessary. This means that a dijet mass resolution of unprecedented precision must be obtained so that these decays can give unambiguous results. A detector designed for these precision measurements will rely on optimal uses of all of the subcomponents with 4π acceptance, including the full tracking and vertex systems, and the electromagnetic calorimeter (ECAL) and the hadron calorimeter (HCAL). It is envisioned that a detector designed to reconstruct the 4-vectors of individual final state particles produced in an e⁺e⁻ collision will enable precise jet energy and position measurements to be made. Reconstruction of individual particles to use as inputs to jet algorithms can reduce or even eliminate a source of systematic uncertainty in the jet measurement—that caused by the correction of detector jets to the particle level. This approach to detector design has been labeled particle flow with the full reconstruction package called a particle flow algorithm (PFA). In this paper, implications of the PFA approach on ILC detector design will be discussed.
2. Physics at the ILC

The apparent success of the standard model (SM), along with recent precision measurements by, e.g., LEP2 [2] and the Tevatron [3], all provide indirect evidence that the final piece of the SM, the Higgs boson, is light ($\lesssim 200$ GeV). Even if new physics such as SUSY should appear, it is likely that there would be at least one light Higgs boson [4]. At a 500 GeV CM energy $e^+e^-$ LC, a light Higgs would be produced in association with vector gauge bosons, e.g. in the process $e^+e^- \rightarrow ZH$. Figure 1 shows a prediction of this and other relevant cross-sections at the ILC as a function of $e^+e^-$ CM energy—most of these involving multi-jet final states [5].

For precision measurements to be made of associated Higgs production, hadronic decay modes of the final states must be included in addition to the leptonic modes, taking advantage of all decays. This requires that the dijet mass resolution of an $e^+e^-$ LC detector be better than the best achieved so far. For a figure-of-merit, identification of vector bosons on an event-by-event basis sets the scale for dijet mass resolution. Production of vector bosons in processes like $e^+e^- \rightarrow WWX$ or $ZZX$ without beam constraints thus demands a dijet mass resolution of $\sim 3$ GeV in order to cleanly identify separately W and Z particles. Usually, the dijet mass resolution is dominated by the calorimeter jet energy resolution, which follows a $1/\sqrt{E_{jj}}$ dependence, where $E_{jj}$ is the sum of the jet energies, so the dijet mass resolution requirement for separation of W and Z bosons is sometimes expressed as a requirement on the jet energy resolution of $\sim 30%/\sqrt{E_{jj}}$ for $e^+e^- \rightarrow$ dijets at the $Z$ pole—a factor of about 2 better than the best current detector performance. As $e^+e^-$ CM energy increases to 500 GeV and beyond, the average jet energy generally increases as well. For example, in 4-jet events at 500 GeV CM
energy from say, $e^+e^- \rightarrow ZH$, the average jet energy is $\sim 120$ GeV. While the energy resolution requirement in terms of $\sigma/E$ remains constant at $\sim 3$–4%, the requirement in terms of $1/\sqrt{E_{jj}}$ relaxes to $\sim 48%/\sqrt{E_{jj}}$. The ability to separate $W$ and $Z$ bosons particle-by-particle using dijet masses results in increased use of available luminosity for physics analyses—the key to the ILC detector as a precision instrument. For example, if the dijet mass resolution, $\sigma/M$, improves from 6 to 3%, the useable luminosity increases by $\sim 40%$. To achieve this precision in real dijet mass measurements, the correction of the jet energy from detector level to final state particle level must be kept as small as possible. This means that the correspondence between reconstructed detector object and final state particle be as close as possible. Such reconstruction procedures have been denoted PFAs. A PFA is not just an analysis package applied to data—as will be shown; the use of this type of jet reconstruction also affects the detector design itself.

3. PFAs

A PFA is a full analysis package that uses the subcomponents of a detector in an optimal way to reconstruct the complete set of final state particles in an $e^+e^-$ interaction. The particles found are then used as inputs to a standard jet algorithm to reconstruct the 4-vectors of fundamental quarks and gluons. To achieve the best dijet mass resolution, it is important to measure each particle type in an optimal way, e.g. photon energy and position in the ECAL and charged particles in the tracking system. In the calorimeter, photons in the ECAL must be separated from early-interacting hadrons, and since neutral hadrons must be measured by the calorimeter only, in the optimization of subcomponents, both the ECAL and HCAL must be built to enable separation of these showers from those of charged hadrons and photons. This is both the biggest challenge to PFAs and the biggest advantage—successful separation of charged and neutral hadron showers in the calorimeter allows the precise charged hadron momentum measurements made in the tracker to be used unambiguously as inputs to a jet algorithm.

Figure 2 shows jet energy distributions carried by the components of a typical jet, divided into charged hadrons, photons and neutral hadrons. Not included are neutrinos and muons which have a small effect on the relative fractions shown.

Typical average fractions are: $\sim 62\%$ for charged hadrons, $\sim 25\%$ for photons and $\sim 13\%$ for neutral hadrons. These fractions are roughly constant as a function of physics process and jet energy, but are really dominated by fluctuations on an event-by-event basis. From this plot, the advantage of a PFA for jet reconstruction is clear—if the charged particle momentum is measured in a tracker instead of a calorimeter, the majority of the jet energy is measured with relatively infinite precision compared with current calorimeter-based jet energy measurements. For a typical ECAL with a resolution of $20%/\sqrt{E}$ or better for photons, the dominant contribution to the jet energy resolution in a perfect PFA is that due to the measurement of the $13\%$ of the energy that is carried by the neutral hadrons. For example, for a 100 GeV jet, on average, 62 GeV is measured with negligible contribution to the overall resolution of the jet energy since it is measured in the tracker; 25 GeV comes in the form of photons, which contribute $\sim 1$ GeV to the overall jet energy resolution for a typical ECAL energy resolution given by $\sigma/E \sim 20%/\sqrt{E}$; the 13 GeV of neutral hadron energy contributes $\sim 2$ GeV to the overall jet energy resolution for a typical HCAL energy resolution of $\sigma/E \sim 60%/\sqrt{E}$.

A typical PFA will consist of various steps optimized to eventually isolate the final state particles produced in an event. Charged particles are tracks—4-vectors with optimized mass assumptions. Photons are primarily ECAL cell clusters with the possibility to include
Figure 2. Typical 120 GeV jet energy distribution fractions carried by charged hadrons (red boxes), photons (blue dots) and neutral hadrons (black triangles).

HCAL cells if the ECAL is too thin for the energy deposit of a particular photon. Charged and neutral hadrons are also calorimeter cell clusters—mainly being distinguished by the presence or absence of a pointing track and/or a minimum ionizing particle (mip) trace. Since the electromagnetic energy deposits from photons can be analytically described, neutral hadron/photon cluster separation can be achieved by a simple $\chi^2$ test on some relevant cluster property such as longitudinal energy distribution. A full PFA will be characterized by the order and specific types of cluster algorithms and hit collection analysis algorithms optimized to produce the final state particle list with the least amount of confusion or overlap of hits and/or clusters. Since hadron energy deposits are not analytic functions and are dominated by fluctuations in energy deposition that produce many fragmented clusters per particle, the biggest challenge to PFAs is the correct association of cluster fragments to either a charged particle track or other neutral hadron cluster fragments. Mistakes made in hit/cluster associations to particles in an event affect the overall performance of the PFA, ultimately leading to degradation in the dijet mass resolution. The contribution to overall dijet mass resolution from these association mistakes in the application of a PFA has come to be known as ‘confusion’. The optimization of a PFA is the minimization of this so-called ‘confusion term’, where all particle identification (ID) mistakes, energy under-counting and double-counting have been included into one term in the resolution determination.

4. PFA impact on detector design

A detector at the ILC will be designed as a general-purpose particle physics detector with full $4\pi$ coverage of tracking, calorimetry and muon detection—capable of measuring the momenta of charged particles and reconstructing jets in multi-jet events with unprecedented precision. Most likely, it will be constructed in the usual barrel and endcap geometry typical of detectors.
at symmetric colliders. It will require a few Tesla solenoidal $B$-field for high momentum tracking precision and to contain (inside the beampipe) the large number of low-energy electron–positron pairs produced by beam–beam interactions. Furthermore, this magnet must also have a large bore to accommodate not only the tracking system, but the entire calorimeter as well, since material in front of the calorimeter should be minimized. The use and assumption of PFA event reconstruction emphasizes precision measurements of particle properties, instead of energy measurements in detector towers, so the design of individual subcomponents in some cases will be much different than in a more traditional detector design approach. The most important impact of PFAs on detector concepts is on the calorimeter design and, in particular, on the design of the HCAL. The following sections break down the specific requirements imposed by the use of PFAs on various detector subcomponents. Where trade-offs in parameters are indicated in the optimization of the detector design, detailed costs, although important considerations in the ultimate detector design, are not considered in this paper.

4.1. PFA requirements for $B$-field

Due to the large number of electron–positron pairs produced by processes related to beam–beam interactions, the central solenoidal $B$-field must be fairly large—of the order of a few Tesla. For example, the minimum radius for placement of the first layer of the vertex detector is $\sim 1$ cm at the IP, increasing to $\sim 3$ cm at a point 50 cm downstream of the interaction point (IP) for a 3 Tesla field [5, 6]. This requirement determines the minimum $B$-field required to use the vertex detector in physics analyses. But also, a strong, few tesla solenoidal $B$-field is required to make precise measurements of charged particle momenta, especially if the tracking volume is minimized so that the entire calorimeter can be kept inside the solenoid. For jet measurements at the ILC, charged particle momenta of up to $\sim 50$ GeV $c^{-1}$ must be measured in the tracker aiming for a resolution in $1/p_t$ of $\sim 10^{-5}$ GeV$^{-1}$ [5, 7]. Also, as the $B$-field strength increases, the volume necessary to separate charged particles from neutrals and positive from negative charges decreases. For cost reasons, it is desirable to keep the magnet bore as small as possible since the figure of merit in the cost of solenoidal $B$-fields is typically taken to be proportional to $BR^2$ (or $B^2R^2$) where $R$ is the radius of the bore. The entire calorimeter must be placed inside the solenoid windings since the amount of material necessary for such a strong magnet would degrade the calorimeter performance both in energy measurement and in shower reconstruction. Clearly, this is an area of trade-off and optimization of the ILC detector in which the PFA approach plays an important role—the implications of this on calorimeter design will be discussed later. With trade-offs in the solenoidal volume and $B$-field strength, most ILC designs assume a central tracking radius of $\sim 1–2$ m with field strengths of 3–5 Tesla roughly inversely proportional to the radii.

To summarize, the positive features of the high-field solenoid magnet requirements are (a) containment of low-energy electron–positron pairs from beam–beam interactions, (b) separation of particles in the smallest possible volume and (c) very precise momentum measurement of high-energy charged particles. However, as mentioned, there are also some negative effects. For example, a non-negligible fraction of low $p_t$ charged tracks are bent into helices which spiral off to the endcaps of the detector. These must be reconstructed and measured correctly to keep the PFA performance optimal. Also, while PFA performance wants to push the ECAL as far away from the IP as possible to enhance particle separation, the solenoid cost rises quadratically with $R$. 

New Journal of Physics 9 (2007) 409 (http://www.njp.org/)
4.2. PFA requirements for tracking

Isolated charged particles, in particular high-energy leptons, occur in many anticipated new physics processes. The momentum resolution goal for the tracking system of $\sim 10^{-5}$ GeV$^{-1}$ in $1/p_t$ is driven by requirements on the mass resolution of dileptons from Z bosons in associated Higgs particle production and also for charged particles from $\tau$ decays. Figure 3 shows the tracking performance goals as a function of momentum for several central tracking options being considered for an ILC detector [8].

In jets, the momentum of charged particles is low, easing the high momentum $p_t$ resolution requirement; but 2-track separation becomes important, requiring good $\phi$ (azimuth) and/or $\theta$ (polar angle) resolution. PFA jet reconstruction will require extrapolation of tracks to the calorimeter, also with good 2-track separation and especially with good $\phi$ angular resolution. Options include a pixel-based vertex detector coupled with a gaseous time-projection chamber (TPC) for three-dimensional (3D) space point measurements or silicon strip central tracking emphasizing the angle $\phi$ resolution. Also, in the presence of a strong central $B$-field, reconstruction of low momentum tracks which spiral into the forward region is important for PFA jet reconstruction, again requiring 2-track separation performance and angular resolution.

4.3. PFA requirements for calorimetry

For the calorimeter, emphasis is placed on shower reconstruction rather than on energy resolution. This means that the granularity (transverse cell size) and segmentation (longitudinal layer readout) both must be large—in other words, very small cell volumes resulting in close
to ∼100 million readout channels in the combined ECAL and HCAL, barrel and endcaps. It is important to note here that shower reconstruction in the calorimeter requires both longitudinal and transverse information for the showers to be reconstructed in all three spatial dimensions. Separation of shower clusters in the calorimeter by the PFA approach requires 3D reconstruction of the particle energy deposits—not just the typical 2D (transverse only) approach of using calorimeter readout towers. This favors sandwich calorimeter construction in order to have maximum flexibility in the longitudinal as well as transverse readout.

The entire calorimeter should be placed inside the solenoid to take advantage of the $B$-field to separate charged particle showers, spreading out the jet. Also, the front face of the calorimeter should be as far from the IP as possible to spread out the neutral particles (photons and neutral hadrons). Since the $B$-field volume should be kept as small as possible to reduce magnet costs, these requirements mean that the calorimeter absorber material should be dense in order to include as many nuclear interaction lengths ($\lambda_i$) as is necessary to contain hadron showers in the smallest possible volume.

The PFA approach to detector design involves using the PFA to evaluate calorimeters in various detector models, understanding the dependence on trade-offs between $B$-field strength, distance from IP to calorimeter, calorimeter granularity and segmentation, and choices of technology for calorimeter absorber material and active media. The PFA itself must be robust and perform well enough to be sensitive to variations in these detector parameters and not be dominated by its own inherent confusion.

4.3.1. Specific requirements for the ECAL. The primary purpose of the ECAL is to reconstruct individual photon showers and measure the photon energy in the event. Photons (on average 2 GeV in $e^+e^-$ at 500 GeV CM) can be contained in a thickness of ∼20 $X_0$ in a section of calorimeter with fine enough longitudinal sampling ($\lesssim 1 X_0$ per layer) to get an adequate energy measurement such that the contribution of the total photon energy resolution in an event is smaller than the contribution of the neutral hadron resolution. In order to avoid overlap with other particles in the event and to separate close-by photons from $\pi^0$ decays, the transverse size of a photon shower should be managed to be as small as possible. This size is determined by the Moliere radius, $r_M$, which is inversely proportional to the density of the absorber material in the calorimeter. Also, since in a layered calorimeter, the readout gaps are composed of relatively low density material, the effective $r_M$ is increased with the size of the readout gap—therefore, minimizing this gap is essential to keep $r_M$ small. A photon shower, being electromagnetic in nature, can be described analytically with well-defined correlations in energy deposition to all directions from the shower core. Therefore, fine sampling is needed for these showers in both transverse and longitudinal directions. Thus, for single photons, the transverse granularity should be such that the cells adequately sample the energy in $r_M$—cell size of order less than ∼0.5 $r_M$, while for separating photons from $\pi^0$ decay even smaller cell sizes are preferred. Figure 4 shows a $\pi^0 \rightarrow \gamma\gamma$ decay with small opening angle in a highly granular ECAL (W absorber, Si readout).

Additionally, a PFA ECAL should be built to maximize the longitudinal separation between photon and hadron showers. A photon typically begins showering in the first $X_0$ of the ECAL, whereas hadrons typically traverse up to $1 \lambda_i$ before showering. Keeping this separation as large as possible can be accomplished by maximizing the ratio of nuclear interaction length to radiation length, $\lambda_i/X_0$, in the absorber material of the ECAL. Materials (non-magnetic) that maximize this ratio are dense—for example, W, Au, Pt, Pb and U. This ratio is ∼3 times...
Figure 4. $\pi^0 \rightarrow \gamma \gamma$ decay in a highly granular W/Si ECAL.

larger in these materials than in other typical calorimeter absorber materials, e.g. Fe (stainless steel) or Cu.

So, for an ECAL designed for PFA jet reconstruction, layered construction is required containing a dense absorber and thin readout gaps. For the ILC detector concepts, the preferred selection is W absorber with silicon readout pixels [5, 9]. Typically ~30 layers of W/Si is chosen, with the first 20 layers consisting of 1/4 to 1/3 $X_0$ sampling and the last 10 layers doubling the absorber thickness—resulting in a total ECAL absorber depth of ~21 $X_0$. The effective $r_M$ is ~1.4 cm in this design. The silicon pixels are 4 mm squares or hexagonal with approximately same area. Photon energy resolutions of ~15–20%/√E can be obtained with this design, keeping the contribution to the dijet mass resolution from photons smaller than that from neutral hadrons—a PFA goal.

4.3.2. Specific requirements for the HCAL. The ultimate goal and biggest challenge of the calorimeter design for the ILC is the measurement of neutral hadron energy, involving both the ECAL and HCAL. Even though the fraction of neutral hadron energy in a typical jet is smaller than either the charged hadron or photon component, its fluctuations and measurement errors dominate the jet energy resolution and the final dijet mass resolution. For perfect PFA jet reconstruction, the HCAL ultimately becomes a neutral hadron detector since hits from charged hadrons are removed and photons are contained in the ECAL. Shower separation performance obtained with high cell granularity and segmentation is important for the HCAL, especially since it is confined inside the solenoid. To keep the number of $\lambda_1$ large enough to contain neutral hadron showers, yet keep the overall detector size and $B$-field volume small, dense absorbers should be used. There is also evidence from simulation that the hadron shower size, as determined by cell hit distributions, is smaller if a dense absorber is used (e.g. W), than if a less dense material like Fe (stainless steel) is employed. Figure 5 shows the rms of the energy distribution for a pion as a function of the cone size used to define the shower in a W/scintillator calorimeter compared with one using Fe (stainless steel) as the absorber.

From this simulation, pion showers are more spread out in a Fe absorber calorimeter than in the one using W for the absorber. Keeping hadron showers as small as possible should help in
the association of calorimeter hit cells to charged hadron tracks and in the separation of shower fragments due to charged and neutral hadrons—key essential parts of a PFA.

The active medium of an ILC HCAL is an area of critical and intense R&D which will result in an optimized technology choice for this detector in a PFA-based jet reconstruction model. Several detector technologies are under consideration for the HCAL—gaseous options including resistive plate chambers (RPCs) [10], gaseous electron multipliers (GEMs) [11] and Micromegas active layers and scintillator [12]. Absorber materials under investigation include Fe (stainless steel), W and Pb. Since cell volumes must be kept small for optimal PFA performance, and since hadron showers are typically spread out in a pattern with low hit density, digital readout of hit cells may perform as well as needed for jet reconstruction purposes. For sampling calorimeters, the energy deposited in an active layer is proportional to the number of charged particles crossing that layer. So, if cell sizes are small, and the number of particles crossing an active layer is close to one, then just counting each hit cell (digital readout) should be approximately equivalent to measuring an energy in that cell (analog readout). The number of hit cells is proportional to the energy, so counting hits gives an energy measurement. Typical hit distributions for digitally measured particle showers of a constant energy are Gaussian in shape. For the same particle showers measured in the normal analog fashion, effects of Landau fluctuations and the actual trajectory of the particle through the layer act to introduce a high-side tail in the Gaussian distribution, leading to a degraded resolution parameter. Figure 6 shows a $5 \text{ GeV} K_L^0$ particle measured by both analog and digital methods.

For the analog case, $\sigma$/mean is $\sim 22\%$, while for the digital method, it is $\sim 17\%$ and more nearly Gaussian in shape. For typical energies of hadrons in jets at the 500 GeV CM energy ILC, the resolution of digitally measured hadron showers is slightly better than the analog method. This persists even to low energies since the Landau and path length effects become even more dominant in the analog method. At high energies, the analog method eventually becomes better
Figure 6. 5 GeV $K^0_L$ energy measured with (left) analog readout and (right) digital readout methods.

since the relative contribution of Landau and path length effects is reduced and the digital method suffers from saturation effects. Using a semi-digital approach employing multiple hit thresholds may extend the energy range in which the digital method performs better than analog. In gaseous calorimeters, the digital approach is necessary since the visible energy is very small. In scintillator, the analog and both pure-digital and semi-digital approaches are being studied. Of course, for any readout method, keeping the thickness of the active layers in the HCAL as thin as possible is important for the same reason as a dense absorber is preferred—to keep the overall $B$-field volume small.

As mentioned above, for perfect PFA, the HCAL becomes a neutral hadron detector. Traditional neutron detectors have been constructed using scintillator readout—a neutron can elastically scatter a proton in the scintillator active media which then becomes visible in the readout. In gaseous calorimeters, the active media density is orders of magnitude less than scintillator and typical gases do not contain much hydrogen, so very little neutron/proton scattering occurs. In simulation, neutron showers in RPCs typically have factors of 2–5 fewer hits than their scintillator counterparts. This has a detrimental effect on the ultimate energy resolution for neutral hadrons, but for PFA applications, this may be offset by making shower separation easier and more efficient. These effects are currently under study in simulation.

4.4. PFA requirements for muon detectors

The main purpose of a muon detector at the ILC is to identify and to provide a trigger for events with muons in the final state. Muons in the final state signal the presence of heavy particles and therefore are key components in the detection of new and/or exotic physics processes. Of secondary importance so far, is the use of these chambers as a backing calorimeter for hadron showers that penetrate the entire ECAL, HCAL and solenoid. Studies are under way to determine the need for this capability at the ILC [13]. Detector cost optimization may be
obtained by keeping the HCAL thinner than would normally be optimal, recovering some of the lost energy in the muon system.

For PFA purposes, detectors outside the HCAL and solenoid are probably of limited use. If charged hadron showers leak out of the HCAL, no information is lost since the momentum of the reconstructed particle track is used. It is only the loss of energy from a high-energy neutral hadron that affects the PFA result; however, as shown above, these particles are typically low energy and, in any case, make up only \( \sim 13\% \) of the jet energy. So far, PFA optimization has included only the detector components inside the solenoid.

5. Detector models for the ILC

Based on the specific requirements as outlined above, several detector models have been conceptualized for the ILC which rely on PFAs for jet reconstruction [14]. These models have been denoted as SiD [http://www-sid.slac.stanford.edu], LDC [http://www.ilcldc.org] and GLD [http://ilcphys.kek.jp/gld/]. The most significant difference in these models is the distance from the IP to the face of the ECAL and, consequently, the type of tracker that fills this volume. Figure 7 shows quadrants from each of these detectors on roughly the same scale.

The calorimeters in these models have in common high granularity and segmentation, but differ in the choice of absorber and readout technologies. There are many variables affecting the detector performance in the PFA approach which must be studied systematically with an optimized PFA to determine their ultimate impact on the final detector design. This is especially important for the calorimeter components. For example, it may be possible to trade off the strength of the solenoidal \( B \)-field with the size of the central tracking volume. However, one...
must remember that $\sim 40\%$ of the jet energy is composed of neutral particles whose separation is not affected by the magnet, depending only on distance from the IP. Also, the size of an HCAL cell (transverse granularity and longitudinal segmentation) can be optimized for different IP-to-HCAL distances. These effects and others are currently under study, which will lead to an optimized detector design for the ILC using the PFA approach.

6. PFA development

In this section, a particular development of a full PFA is described [15]. It is designed in a modular way, with the goal of optimizing each step in the PFA for the particle type being sought. Since different particle species display very different showering characteristics in a sampling calorimeter, this method of modular optimization is seen as crucial to the development of an optimal PFA.

The starting point for any successful PFA is a high purity sample of calorimeter objects. The most basic of these objects is an individual calorimeter readout cell. A calorimeter cell must be small enough so that the single particle purity is close to 100%—only one particle contributes to the energy deposit in a calorimeter cell. Figure 8 illustrates this purity for a detector model with a highly granular and segmented calorimeter.

In this 4-jet event, essentially all of the HCAL cells contain energy from only one particle, while a few of the ECAL hits include overlapping particle depositions—in most cases due to two close-by photons from a $\pi^0$ decay. With cell purity such as this, it is in principle possible to perfectly separate the showers in the calorimeter—the goal of a PFA. Figure 9 shows an $e^+e^- \rightarrow \text{t}\overline{\text{t}}$bar event which forms six jets, with perfect particles shown as green and yellow lines and the perfect calorimeter clusters associated with the particles. This detector model has a W/Si ECAL and an HCAL with W absorber and RPC active media.

One way to take advantage of this high purity is to start with tracks found in the tracking detector, extrapolating them to the calorimeter and linking single calorimeter hits layer-by-layer until evidence of an interaction is found. In this way, the mip tail of a charged hadron shower can be found without diluting the purity of the calorimeter objects. In practice, this is done by assigning a density value (in a single calorimeter layer) to each calorimeter hit cell centered
in a suitable window. When cells with low density (∼1) are close enough to the extrapolated track, they are included in a ‘mip cluster’ for that track. These mip clusters are built up by extrapolating layer-by-layer into the calorimeter. The first layer in which there is no close hit match or the density of a matched hit is large is called the interaction layer since this signifies the end of the mip tail and the beginning of a hadronic shower. Mip clusters are found in this way with very high purity, so they do not adversely affect the remaining steps in the PFA since almost no confusion is introduced at this step. The calorimeter cells matched to tracks in this step are removed from further analysis and the interaction layer and mip cluster are linked to the tracks. Figure 10 shows the PFA run on an $e^+e^- \rightarrow \text{t}\bar{\text{t}}$ event at 500 GeV CM energy with (perfect) tracks and associated mip clusters found by this algorithm. Again, the detector model includes a W/Si ECAL and a W/RPC HCAL.

At this point, having removed the calorimeter mip hits associated with found tracks, the first application of cell clustering is done. The object is to cluster only hit cells in the ECAL using a cluster algorithm optimized for finding photon energy deposits. Since clustering is not perfect, this introduces impurity into the sample of calorimeter objects—at this point, a clustering algorithm must be chosen to efficiently identify photons and measure their energy while ensuring that the cluster purity is high. For this particular PFA, a combination of two cluster algorithms results in photon ID with both high purity and efficiency. First, a cluster algorithm with high efficiency for photons is used to define the overall photon shower. For example, a fixed cone clusterer can be used here—tuned to single photons in simulation. These clusters must be separated from any extrapolated tracks. Photon showers, being electromagnetic in nature, can be described analytically since energy deposition in these showers is correlated with both the longitudinal and transverse extent of the shower. The individual calorimeter cells included in each cone cluster are then clustered with a nearest-neighbor clusterer. This step defines the core of the electromagnetic shower with high purity. It is this shower core that is best described by the correlation of deposited energy to longitudinal extent of the shower, so a Hessian matrix (H-matrix) is constructed to test these correlations in the nearest-neighbor
cores [16]. Using only the longitudinal H-matrix test means that close-by photons from $\pi^0$ decay cannot be distinguished, but by introducing the transverse correlation test, it should be possible to separate these photon showers—this is currently under study. This combined clustering and H-matrix approach results in photon ID with both high purity and efficiency, for photons with energy greater than $\sim 500$ MeV. Figure 11 shows the photons in an event found by this method. The green lines represent all neutral particles, both photons and neutral hadrons. In this case, one photon cluster is in the final state.

**Figure 10.** PFA results for perfect tracks (yellow curves) and associated mip clusters.

**Figure 11.** Photon cluster as found by PFA and perfect photon (green line) from Monte Carlo.
For lower energy photons, a different approach must be used, since these photons have very few calorimeter cell hits. At the present time, shower clusters with too few hits to be tested with the H-matrix are classified as photons if the cluster energy is less than 1 GeV and the cluster begins within the first 3 $X_0$ of the ECAL. Clusters passing these tests are then defined as reconstructed photons, ready for inclusion into a jet algorithm. Calorimeter hits used in these clusters are removed from further consideration.

The next step starts with the extrapolated tracks and their associated mip clusters as found previously. A cluster algorithm is now needed to group hits from the showers associated with charged hadrons. Again, high purity is more important in this step than is high efficiency, since the shower hits are removed and do not contribute to the charged particle momentum measurement. As in the photon shower core, a very good cluster algorithm with high purity is the nearest-neighbor clusterer. This type of algorithm also has the attractive property of being able to follow branches of showers without including separated close-by fragments from neutral hadrons. Associating the shower core with high purity to the charged track is the most important part of this step and of the entire PFA. Shower fragments will be evaluated later when efficient reconstruction of the neutral hadron component of the event is defined. Shower pieces are added to the track in an iterative fashion, testing the ratio of calorimeter energy of the combined cluster to the track momentum, $E/p$, after each shower addition—this ensures that a large energy shower made up of contributions from both charged and neutral hadrons does not get removed from further consideration. When $E/p$ is much too large, the last shower can be reclustered to try to separate the correct portion from any other contributions and $E/p$ recalculated. The goal of this step is to remove calorimeter hits associated with high purity to the charged hadrons in the event. Figure 12 shows the clusters matched to tracks including the mip traces found earlier.

Lastly, the remaining showers and shower fragments are evaluated and characterized as either belonging to neutral hadrons—in which case they can be merged into particle clusters, or as belonging to a charged particle—in which case they are not used. At this stage of the PFA, the leftover calorimeter hits are primarily due to showering of neutral hadrons—a high purity

Figure 12. Clusters associated with perfect charged particles (yellow curves) in PFA.
sample of hits and clusters has been removed by association with tracks and a high purity and efficiency collection of clusters has been defined as photons. Fragments are evaluated by their proximity to other shower fragments or to a track extrapolation, the number of hits in a fragment, etc. Also, cluster pointing using internal cluster axes is also under study as a useful discriminant. At this point, neutral hadrons are defined by their shower clusters and energy as measured in the calorimeter. A mass assumption can be used, but no distinguishing characteristic useful for separating neutrons from $K^0_L$ mesons has been identified yet. Figure 13 shows the neutral hadrons as defined by the PFA.

At this last stage of the completed PFA, a list of reconstructed particles is now available for inclusion into a jet algorithm—charged particle momenta from the tracker, photon 4-vectors from the ECAL, and neutral hadron energy measurements with a mass assumption from the ECAL and HCAL. Figure 14 shows this event with the PFA objects and reconstructed particles displayed.

7. PFA results for jets

As mentioned above, the detector performance goal made possible by the use of PFAs is to be able to achieve dijet mass resolutions of $\sim 3\%$, allowing the separation of $W$ and $Z$ vector bosons on an event-by-event basis. Using the PFA described in the last section, this goal can be tested with $e^+e^- \rightarrow ZZ$ events where one $Z$ decays to $\nu\bar{\nu}$ and the other to a pair of jets. The PFA and detector performance can be tested without including ambiguities related to wrong jet–jet combinations, but with jet energies of $\sim 120$ GeV which is characteristic of 4-jet events at 500 GeV CM energy. Figure 15 shows dijet results using reconstructed particles from a PFA as input to the kT jet algorithm compared with the perfect dijets for this event.

More complicated final states can also be reconstructed with PFAs. The process $e^+e^- \rightarrow t\bar{t}$ can produce six jets in the final state. Figure 16 shows the event previously shown in figure 9 with six jets formed from PFA reconstructed particles.
Figure 14. All PFA objects and corresponding reconstructed particles for this event.

Figure 15. Left: Two jets formed from PFA reconstructed particles in a ZZ final state where one Z decays to jets. Right: Perfect PFA jets for this event.

Dijet mass results from PFAs on $e^+e^-$ events at 500 GeV CM energy are now being obtained with resolutions close to the goal of $\sim 3–4\%$ as indicated above. Improvements in individual PFA particle ID algorithms as well as in the overall PFA performance are producing results with mass resolutions that are dominated by the calorimeter energy resolution instead of the confusion contribution resulting from PFA mistakes.

The following plots illustrate the performance of the PFA for $e^+e^- \rightarrow Z$ at 91 GeV CM energy (Z-pole events). Figure 17 shows the overall performance of the PFA for measurement of both the perfect and real PFA dijet masses.

For the modular PFA, individual reconstructed particles as defined by the real PFA can be compared with those for a perfect PFA. Figure 18 shows the difference in the number of
Figure 16. Six jets (color-coded particle traces and calorimeter clusters) formed from PFA reconstructed particles in a ttbar final state.

Figure 17. Dijet mass distribution for $e^+e^- \rightarrow Z$ events at 91 GeV CM energy comparing real PFA results (blue dots) with the perfect PFA results (green histogram).

The present photon finder used in the PFA is inefficient for identifying low-energy photons ($<500$ MeV)—this is reflected in the difference plots both in the number of photons and in photons per event and the total photon energy per event comparing real PFA particles with perfect particles.

New Journal of Physics 9 (2007) 409 (http://www.njp.org/)
Figure 18. Difference between real PFA and perfect PFA number of photons per event (left) and photon energy sum per event (right).

the energy sum per event. Figure 19 shows the difference in number of neutral hadron clusters per event and the total neutral hadron energy per event comparing the real PFA result with the perfect PFA.

In this algorithm, clustering is done in the ECAL and HCAL separately, so a typical hadron has a minimum of at least two clusters per particle. Along with the fragments associated with a typical neutral hadron, the difference in number of clusters shows a positive result when comparing real clusters with the perfect particle (single cluster). Finally, figure 20 shows the difference between the PFA dijet mass and the perfect PFA result per event.

A fit to this distribution of current work in progress yields a resolution of $\sim 4.4$ GeV, still $\sim 35\%$ higher than the goal needed to separate $W$ and $Z$ vector bosons in $e^+e^-$ events; but a factor of $\sim 2$ better than current non-PFA results. With the modular PFA approach, additional improvements can be obtained by modifying individual particle algorithms for best single particle 4-vector reconstruction.

With some small additional improvements, PFAs can be used for detector design optimization of the size of detector components, technology choices for calorimeter readout, material choices for calorimeter absorber layers and many other parameters. Events such as $e^+e^- \rightarrow ZZ$ and ttbar, along with other final state benchmark processes, are being studied with PFAs in order to optimize the various detector concepts. Trade-off relationships between several detector design parameters are expected to be uncovered, and there are already indications that some variables are related. For example, it appears that there is a stronger dependence of PFA performance on the size of the central tracking volume (specifically, the radius from IP to the face of the ECAL) than on the strength of the solenoidal $B$-field [17]. Also, first indications are that for fixed inner radius, there is a minimum transverse size of HCAL cells that gives adequate PFA performance—smaller sizes do not improve the performance of the PFA [17].
Early indications are that an ILC detector optimized for jet reconstruction using PFAs must be large with somewhat relaxed requirements on the transverse HCAL cell granularity, with the size of the inner tracking volume taking precedence over the strength of the $B$-field.

### 8. Importance of test beam validation

In this paper, as in the worldwide development of the PFAs, only simulations have been used to understand the PFA approach and also to begin to optimize detector design parameters. This is necessary because the goal of PFA analysis is to optimize jet reconstruction. For calorimeters in particular, in the past, designs have been optimized for best single particle energy resolution in calorimeter readout towers. Simulation results for single particle energy resolution were easily verified in test beams for large-scale prototypes of calorimeter modules. However, the PFA approach requires optimal particle reconstruction in the calorimeter, not optimal energy resolution, with the goal being to obtain the best dijet mass resolution possible. An optimal ‘test beam’ for PFA validation would be a source of high-energy (QCD) jets with known ID and energy and angular distributions of all particles in the jet. Of course, this is impossible to obtain, so instead, test beams will be used to try to verify single particle showering characteristics in various calorimeter prototypes. It is hoped that shower shapes and hit multiplicities in both analog and digital calorimeter prototypes can be used to choose an optimal hadron shower model to use in event simulations. In general, data do not exist for calorimeters with the granularity and segmentation required for PFAs in an ILC detector, so placement of such a calorimeter prototype in a test beam is a high priority for the ILC detector R&D community. Ultimately, $e^+e^-$ event
simulations must be used to design an ILC detector—with a vigorous and well-designed test beam program, it is anticipated that individual particle showering can be modeled accurately enough to use PFAs in these designs \[18\].

9. Summary

A detector designed for the future ILC must be capable of dijet mass measurements made with unprecedented precision. A novel way to achieve the required precision is to define jets using a detector designed to reconstruct in an optimal way individual final state particles—the particle flow approach. The dijet mass resolution using this approach in detector simulations is, in principle, good enough to separate vector bosons on an event-by-event basis, thus allowing full use of the hadronic decay modes and therefore all of the available luminosity—a requirement for precision measurements at the ILC. PFAs are currently under development and are, themselves, being optimized to produce the required mass resolution performance for these precise measurements. Using a modular approach, individual particle reconstruction and ID can be obtained with the resulting contribution to the full PFA mass resolution from mistakes being smaller than the intrinsic calorimeter resolution for hadrons. PFAs can then be used to optimize many detector parameters, including the strength of the $B$-field, the size of the tracking volume, the ECAL granularity, and most importantly, the HCAL design—with the ultimate goal being an optimal ILC detector for precision jet measurements.
Acknowledgment

Argonne National Laboratory’s work was supported by the US Department of Energy, Office of Science, under contract DE-AC02-06CH11357.

References

[1] *International Linear Collider Reference Design Report 2007* online at http://www.linearcollider.org/cms/

[2] Barate R et al 2003 CERN-EP-2003–011 Phys. Lett. B 565 61–75

[3] TEVNPH Working Group 2006 CDF Note 8384/D0 Note 5227

[4] LEP Working Group for Higgs Boson Searches 2005 LHWG-Note 2005–01 Abazov V M et al 2005 FERMILAB-PUB-05/058-E Phys. Rev. Lett. 95 151801

[5] American Linear Collider Working Group 2001 *Linear Collider Physics—Resource Book for Snowmass 2001* FERMILAB-PUB-01/058-E

[6] Damerell C J S 2005 Vertex detectors and the Linear Collider *Proc. 2005 International Linear Collider Workshop—Stanford USA (March 2005)* vol 1, SLAC R-772

[7] Hewett J (ed) *Proc. International Linear Collider Workshop—Stanford, USA (March 2005)* vol 2, SLAC R-772

[8] Schumm B 2005 *Proc. 2005 International Linear Collider Physics and Detector Workshop (Snowmass Village, CO, August 2005)* online at http://www.slac.stanford.edu/econf/C0508141

[9] White A P 2005 *Proc. 2005 International Linear Collider Workshop—Stanford, USA (March 2005)* vol 1, SLAC R-772

[10] White A P 2005 *Proc. 2005 International Linear Collider Workshop—Stanford, USA (March 2005)* vol 2, SLAC R-772

[11] Repond J 2006 *Proc. XII Int. Conf. on Calorimetry in High Energy Physics (Chicago, IL, June 2006)*

[12] Xia L 2005 *Proc. 2005 International Linear Collider Workshop—Stanford, USA (March 2005)* vol 2, SLAC R-772

[13] Tabarelli T 2005 *Proc. 2005 International Linear Collider Workshop—Stanford, USA (March 2005)* vol 2, SLAC R-772

[14] White A P and Yu J 2005 *Proc. 2005 International Linear Collider Physics and Detector Workshop (Snowmass Village, CO, August 2005)* online at http://www.slac.stanford.edu/econf/C0508141

[15] Takeshita T 2006 *Proc. XII Int. Conf. on Calorimetry in High Energy Physics (Chicago, IL, June 2006)*

[16] Sefkow F 2005 *Proc. 2005 International Linear Collider Workshop—Stanford, USA (March 2005)* vol 2, SLAC R-772

[17] Chakraborty D 2005 *Proc. 2005 International Linear Collider Workshop—Stanford, USA (March 2005)* vol 2, SLAC R-772

[18] *International Linear Collider Reference Design Report 2007* vol 4 Detectors online at http://www.linearcollider.org/cms/

[19] Behnke T 2005 *Proc. 2005 International Linear Collider Workshop—Stanford, USA (March 2005)* vol 1, SLAC R-772

[20] Graf N et al 2006 *Proc. XII Int. Conf. on Calorimetry in High Energy Physics (Chicago, IL, June 2006)*

[21] Graf N 2007 private communication

[22] Thomson M 2007 online at http://events.lal.in2p3.fr/conferences/ILCSoftware/

[23] World-wide ILC Test Beam Working Group 2004 online at http://www-lc.fnal.gov/lc_testbeams/wwlctb_working_group.pdf

New Journal of Physics 9 (2007) 409 (http://www.njp.org/)