Study of Tracking and Flavor Tagging with FPCCD Vertex Detector

"Talk presented at the International Workshop on Future Linear Colliders (LCWS13), Tokyo, Japan, 11-15 November 2013."

Tatsuya Mori1 and Daisuke Kamai1, Akira Miyamoto2, Yasuhiro Sugimoto2, Akimasa Ishikawa1, Taikan Suehara3, Eriko Kato1, Hitoshi Yamamoto1
Tohoku University1, KEK2, Kyushu University3

One of the major physics goals at the ILC is the precise measurement of the Higgs coupling constants to b-quarks and c-quarks. To achieve this measurement, we need a high-performance vertex detector leading to precise flavor tagging. For this purpose, we are developing the Fine Pixel CCD (FPCCD) vertex detector. In this paper, we will report on the development status of FPCCDTrackFinder, a new track finder improving tracking efficiency, especially in the low \( p_T \) region, and an evaluation result of the flavor tagging performance with FPCCDTrackFinder in the FPCCD vertex detector.

I. INTRODUCTION

A. Role of Vertex Detector in ILC

One of the major physics goals at the ILC is the precise measurement of the Higgs coupling constants to b-quarks and c-quarks, which is a critical test of Higgs and Yukawa interactions in the Standard Model (Figure 1) [1]. To achieve this measurement, we must identify Higgs decays into \( b\bar{b} \), \( c\bar{c} \), or \( q\bar{q} \) as precisely as possible, where \( q \) is a u-, d-, or s-quark. This identification is called flavor tagging. One of the most discriminating variables for the flavor tagging is the difference between the secondary vertex position and the Higgs decay at the IP; typically, the proper decay length of b-hadron is \( 400 \sim 500 \mu m \), and that of c-hadron is \( 100 \sim 300 \mu m \). Thus when the vertex resolution is improved, flavor tagging will be improved. The vertex resolution is determined by the resolution of the impact parameter \( d_0 \) (in this study, we use the definition of track parameters in \([2]\)), which in the vertex detector at the ILC is required to satisfy

\[
\sigma_{d_0} = 5 \mu m \oplus 10 \mu m \cdot \frac{\text{GeV}}{c} \cdot \frac{p \beta \sin^{3/2} \theta}{10^{10}}.
\]  

In addition, the vertex detector is also required to keep a low occupancy of less than 2 \( \sim 3\% \). The dominant background in the vertex detector at the ILC is from \( e^+e^- \) pairs generated from beamstrahlung. The pair-BGs have relatively low momentum, so they curl many times through layers and increase pixel occupancy. To achieve these requirements and improve the flavor tagging, we are researching and developing the Fine Pixel CCD (FPCCD) Vertex Detector. The FPCCD Vertex Detector is an optional detector for ILD detector concept \([4]\), and the performance was studied using the ILD software framework \([5]\).

B. FPCCD’s Features

FPCCD is a fully depleted silicon pixel sensor with very small pixel size \([3]\). As shown in Table \([1]\) the pixel size on layer-0 and layer-1 is 5 \( \mu m \), and that on the outer layers are 10 \( \mu m \). The sensitive and total thickness of this sensor is 15 \( \mu m \) and 50 \( \mu m \), respectively. The number of pixels of the FPCCD vertex detector is around 400 million. These features are expected to lead to low occupancy and good vertex resolution.

A notable feature of the very small pixels is that FPCCD can make clusters of hit pixels. As shown in Figure \([2]\) when a particle goes through a layer, it deposits energy into a few pixels in the layer. Those hit

* moritatsu@epx.phys.tohoku.ac.jp

FIG. 1. Expected measurement precision of Higgs coupling constant in ILC.
TABLE I. Geometry of the FPCCD VXD.

| layer (from IP [mm]) | pixel size [$\mu m^2$] |
|----------------------|------------------------|
| 0                    | 16                     |
| 1                    | 18                     |
| 2                    | 37                     |
| 3                    | 39                     |
| 4                    | 58                     |
| 5                    | 60                     |

pixels can be regarded as a cluster. The cluster shape depends on the way of traversing a layer, so the shape is useful for extrapolation in tracking and discrimination between pair-BG clusters and signal clusters.

Furthermore, by taking the weighted average of energy deposits of each pixel belonging to a cluster, we can reconstruct the hit point more precisely.

![FIG. 2. Explanation of Cluster. The red boxes denote pixels hit by particles. The pixels adjacent to each other are regarded as a cluster.](image)

The readout happens between trains, so the hit pixel data from all bunches in a train are accumulated in the FPCCD. An advantage of this procedure is that we can ignore beam-induced RF noise, while it has the disadvantage that tracking is challenging due to so many hit pixels.

C. Occupancy and Impact Parameter Resolution

We evaluated the pixel occupancy of the FPCCD vertex detector from pair-BGs using the ILC TDR beam parameters [4], and the result is shown in Table II. The occupancy on layer-0 for collision energies up to 500 GeV is relatively low, but at 1000 GeV it is too high. The solutions for 1000 GeV are still being studied.

We also evaluated the impact parameter resolution of the FPCCD vertex detector. Since the FPCCD has very good position resolution, the impact parameter resolution is expected to be good. The result of the study of the impact parameter resolution of the FPCCD is shown in Figure 3 [3]. For comparison, the configuration of the ILD VXD used in the Detailed Baseline Design (DBD) [6] is also evaluated. The position resolution of the DBD ILD VXD and the FPCCD VXD is shown in Table III. The FPCCD satisfies the requirement of Equation II and gives around 1 $\mu m$ $d_0$ resolution in the high momentum region.

| $E_{CM}$ [GeV] | # of bunch crossings per train | Pixel Occupancy [%] |
|----------------|-------------------------------|---------------------|
| 250            | 1312                          | 0.561               |
| 350            | 1312                          | 0.702               |
| 500            | 1312                          | 1.244               |
| 1000           | 2450                          | 12.752              |

### TABLE II. Pixel occupancy of the FPCCD vertex detector on layer-0 from pair-BGs at the several center of mass energy.

| layer | Position Resolution [$\mu m$] |
|-------|--------------------------------|
| DBD   | FPCCD                         |
| 0     | 2.8                           | 1.4                |
| 1     | 6.0                           | 1.4                |
| 2     | 4.0                           | 2.8                |
| 3     | 4.0                           | 2.8                |
| 4     | 4.0                           | 2.8                |
| 5     | 4.0                           | 2.8                |

### TABLE III. Position Resolution of the DBD ILD VXD and the FPCCD VXD in each layer.

![FIG. 3. Impact parameter resolution of the DBD ILD VXD (solid line) and the FPCCD VXD (dotted line) and the requirement of Equation II (long dashes). We assume the baseline position resolution given in Table III.](image)

II. FPCCDTRACKFINDER, A NEW TRACKING ALGORITHM

For studying the flavor tagging performance using FPCCD, the tracking efficiency has been evaluated. The
ILD tracking algorithm used in the DBD [6] was unable to efficiently reconstruct low $p_T$ tracks (less than around 1.7 GeV/c) though there were enough VXD hits. We have developed FPCCDTrackFinder, a new tracking algorithm for reconstructing low $p_T$ tracks with high efficiency. In this section, firstly, the DBD ILD tracking algorithm will be introduced briefly, and the result of evaluating the tracking efficiency will be shown. Secondly, the difference between FPCCDTrackFinder and the DBD ILD tracking is explained. Thirdly, the result of evaluating the tracking efficiency of FPCCDTrackFinder is shown. Finally, the amount CPU time and the memory consumption are mentioned.

A. The DBD ILD Tracking

The strategy of the DBD ILD tracking is shown in Figure 4. In the first phase, the SiliconTracking processor reconstructs tracks using only VXD, SIT, and FTD hits, and CLUPATRA processor reconstructs tracks using only TPC hits. We refer to those tracks as "silicon track" and "TPC track". In the second phase, one silicon track and one TPC track are combined into one track, named full track. This full track is used for vertexing.

In this paper the tracking efficiency is defined by

\[
\text{Tracking Efficiency} \equiv \frac{\# \text{ of Good Track originated from Good Particle}}{\# \text{ of Good Particle}} \quad (2)
\]

where "Good Track" is a reconstructed track with VXD hits $\geq 5$ and track purity $> 0.75$, and "Good Particle" is a Monte Carlo particle making VXD hits $\geq 6$ and SIT hits $\geq 4$. Track purity is defined by

\[
\text{track purity} \equiv \frac{\# \text{ of true hits belong to the track}}{\# \text{ of hits belong to the track}} \quad (3)
\]

where true hits mean the hits originated from a particle corresponding to the track.

The sample used here is $t\bar{t} \rightarrow 6\text{jets}$ at 350 GeV, and the number of events is 1000. The results of evaluating the tracking efficiency with the DBD ILD tracking are shown in Figure 5 (efficiency of silicon track) and Figure II A (efficiency of full track). As shown in Figure 5 and Figure II A, the tracking efficiency of silicon and full tracks is around 98 and 99% in $p_T > 1.7$ GeV/c, but decreases below $p_T = 1.7$ GeV/c in both cases of the DBD ILD VXD and the FPCCD VXD. The reason of slight increase from 98% (silicon track) to 99% (full track) is that after combining a silicon and a TPC track into a full track, hits unused for tracking at this stage are used to add further full tracks. However, in spite of this process, the efficiency of finding a full track deteriorates at the same $p_T$ as for silicon tracks. This means that the efficiency of finding a silicon track contributes to the full track reconstruction efficiency, because the number of silicon tracks decreases at $p_T < 1.7$ GeV/c. Thus, to improve tracking efficiency in the low $p_T$ region, we need to improve the silicon tracking. For this purpose, FPCCDTrackFinder, a new tracking algorithm, has been developed.

B. Differences between FPCCDTrackFinder and SiliconTracking of the DBD ILD Tracking

FPCCDTrackFinder is based on the DBD ILD tracking, so we shall explain how FPCCDTrackFinder works by comparing SiliconTracking of the DBD ILD tracking.

Firstly, SiliconTracking generates track seeds (Figure 7). To simplify the graphical representation, the sectional view of only the VXD is shown in Figure 7 but actually the SIT is also used in SiliconTracking. At the beginning,
as shown in Figure 7, we divide the whole region into 80 pieces (namely 4.5° per piece). Then, if there are 3 hits in each of 3 determined layers in a piece, simple helical fitting is applied and selected as a track seed, if the fit succeeds.

Combinations of the 3 determined layers are as follows.

\[(8 \, 6 \, 5) \quad (8 \, 6 \, 4) \quad (8 \, 6 \, 3) \quad (8 \, 6 \, 2) \quad (8 \, 5 \, 3) \quad (8 \, 5 \, 2) \]
\[(8 \, 4 \, 3) \quad (8 \, 4 \, 2) \quad (6 \, 5 \, 3) \quad (6 \, 5 \, 2) \quad (6 \, 4 \, 3) \quad (6 \, 4 \, 2) \]
\[(6 \, 3 \, 1) \quad (6 \, 3 \, 0) \quad (6 \, 2 \, 1) \quad (6 \, 2 \, 0) \quad (5 \, 3 \, 1) \quad (5 \, 3 \, 0) \]
\[(5 \, 2 \, 1) \quad (5 \, 2 \, 0) \quad (4 \, 3 \, 1) \quad (4 \, 3 \, 0) \quad (4 \, 2 \, 1) \quad (4 \, 2 \, 0) \]

where 8 and 6 denote the outer and inner layer of the SIT, and numbers from 5 to 0 denote VXD layers.

After the track seeding process, SiliconTracking extrapolates tracks from track seeds as shown in Figure 8. Like the track seeding process, the extrapolation process begins with the division of the whole region into 80 pieces. Next, if at least one track seed exists in one of the 80 pieces, the following processes will be done with respect to the track seed.

1. On a layer in the piece (namely, in the red line shown in Figure 8), we try to find the hit closest to the track seed.

2. If the distance between the hit and the track seed is less than a given threshold, after adding the hit into the track seed, we apply simple helical fitting to the combined track seed and the track is renewed if the fit succeeds.

3. We redo (1) and (2) with respect to the other layers.

That is the basic way that track seeds grow up to be well-reconstructed tracks.

However, there are two shortcomings in the track seeding process. One is that the width of a piece, 4.5°, is too narrow to generate a low \( p_T \) track seed. In addition, the way of dividing the whole region leads to tracks crossing the region boundary not being able to generate their track seeds. The reason why widening pieces is not preferred is that the number of ghost track seeds and CPU time increase. The other is that there are many combinations of the 3 layers. This creates many ghost track seeds and consumes a lot of CPU time.

Furthermore, there are two known shortcomings of the extrapolation process. One is that SiliconTracking uses simple helical fitting for extrapolation. This fitting doesn’t consider multiple scattering and energy loss effect. Low \( p_T \) tracks tend to be affected relatively significantly by these effect, so \( \chi^2/ndf \) from this fitting tends to become improperly high. The third is that the search window for extrapolation is not flexible. Since the search region does not depend on the processed track seed but on the piece width, some hits for a low \( p_T \) track seed cannot exist in the search region because low \( p_T \) tracks tend to cross the region boundary.

For overcoming these shortcomings, FPCCDTrackFinder generates track seeds as shown in Figure 9. Firstly, FPCCDTrackFinder chooses one of the hits on the outer layer among given 3 layers for creating track seeds. Secondly, FPCCDTrackFinder creates all candidate track seeds whose hits on the middle and inner layers exist in a search window calculated in the following way.

1. We draw on the plane perpendicular to the uniform magnetic field two hypothetical tracks generated at the IP, going through the hit on the outer layer, and having \( p_T = 0.18 \text{ GeV/c} \) (by default) and +1 and
FIG. 9. The graphical representation of track seeding process in FPCCDTrackFinder. Crosses denote the VXD hits. Crosses surrounded by red circles in the outer layer are used to determine a wide enough search region to catch track seeds with $p_T > 0.18$ GeV/c (by default). Dotted curved lines denote tracks passing through the IP with $p_T = 0.18$ GeV/c. Red dashed lines denote search windows for generating track seeds, and their length is determined by intersections between layers and the dotted curved lines. Blue curved lines denote the track seeds generated by this process.

-1 electric charge respectively.

(2) The two points of intersection between the middle layer among given 3 layers for creating track seeds and the two tracks are regarded as the end points of the search window. The same goes for the inner layer.

As compared to SiliconTracking, FPCCDTrackFinder hardly fails to reconstruct low $p_T$ track seeds due to the crossings of the region boundary.

In addition, the combinations of 3 layers are reduced from the ones for DBD ILD tracking to

$$(8 \ 6 \ 5) \ (8 \ 6 \ 4) \ (8 \ 5 \ 4) \ (6 \ 5 \ 4) \ (5 \ 4 \ 3)$$

This reduction of combinations is appropriate because the above search window can catch almost all track seeds with $p_T > 0.18$ GeV/c Notice that the inner layers of the VXD, 0 to 2-layer, are not used in order not to increase the number of ghost track seeds and the CPU time, because the hit density of these layers is relatively high.

After track seeding process, FPCCDTrackFinder extrapolates tracks from track seeds as shown in Figure 10. Instead of simple helical fitting, FPCCDTrackFinder uses a Kalman Filter, a fitter considering multiple scattering and energy loss effects. Thus, although the computational budget increases, a Kalman Filter can output $\chi^2/ndf$ more accurately than simple helical fitting for low $p_T$ tracks.

In addition, FPCCDTrackFinder does not divide the whole region into 80 pieces, but instead determines the width of the search window by considering track parameters and the errors of the processed track seed. FIG. 10. The graphical representation of the extrapolation process in FPCCDTrackFinder. FPCCDTrackFinder determines the width of the search window (namely, red dashed line) by considering track parameters and the errors of the processed track seed.

C. Tracking Efficiency of FPCCDTrackFinder

The setup for the evaluation of the tracking efficiency is the same as described in section II A. Comparisons of the DBD ILD tracking and FPCCDTrackFinder with FPCCD VXD are shown in Figure 11 (efficiency of reconstructing silicon tracks) and Figure 12 (efficiency of reconstructing full tracks). As shown in Figure 11 and Figure 12, FPCCDTrackFinder improves the efficiency of both silicon and full tracks to $\sim 99\%$ above $p_T = 0.6$ GeV/c, as compared to 97% with the DBD ILD tracking. In addition, dependency of the efficiency on $\cos \theta$ is shown in Figure 13 (silicon tracks) and Figure 14 (full tracks), where tracks with $|p| > 1$ GeV/c are evaluated. These Figures show that FPCCDTrackFinder improves the tracking efficiency of both silicon and full track to $\sim 99\%$ within $|\cos \theta| = 0.9$, as compared to 96.5% with the DBD ILD tracking. The reason that the efficiency decreases in $|\cos \theta| > 0.9$ is that the acceptance of the SIT is $|\cos \theta| < 0.9$, so the efficiency in $|\cos \theta| > 0.9$ is not considered in this evaluation.

Next, the result of the evaluation of the tracking efficiency with pair-BG at 350 GeV is shown in Figure 15 (silicon track) and Figure 16 (full track). As shown in Figure 15 and Figure 16, the tracking efficiency holds $\sim 99\%$ above $p_T = 0.6$ GeV/c regardless of the presence
FIG. 11. The tracking efficiency vs. $p_T$ of silicon tracks with the FPCCD VXD and the DBD ILD tracking (black crosses) and FPCCDTrackFinder (red crosses).

FIG. 12. The tracking efficiency vs. $p_T$ of full tracks with the FPCCD VXD and the DBD ILD tracking (black crosses) and FPCCDTrackFinder (red crosses).

FIG. 13. The tracking efficiency vs. $p_T$ of silicon tracks ($|p| > 1$ GeV/c) with the FPCCD VXD and the DBD ILD tracking (black crosses) and FPCCDTrackFinder (red crosses).

FIG. 14. The tracking efficiency vs. $p_T$ of full tracks ($|p| > 1$ GeV/c) with the FPCCD VXD and the DBD ILD tracking (black crosses) and FPCCDTrackFinder (red crosses).

III. PERFORMANCE EVALUATION OF FLAVOR TAGGING

In this section, the result of the evaluation of the flavor tagging performance is shown.

A. Setup for the Evaluation

The flavor tagging process is implemented by LCFIPlus [8]. The sample for the evaluation is $Z \rightarrow b\bar{b}, c\bar{c}, q\bar{q}$ at 91.2 GeV, with 25000 events for the testing, and 25000 events for the training. LCFIPlus gives us the information of b-jet, c-jet, and q-jet likelihood for the reconstructed jets. The b-tag efficiency and the b-tag purity are defined by

$$b\text{-tag efficiency} = \frac{\# \text{ of b-tagged } b\text{-jet}}{\# \text{ of all } b\text{-jet}}$$

$$b\text{-tag purity} = \frac{\# \text{ of b-tagged } b\text{-jet}}{\# \text{ of all } b\text{-tagged } b\text{-jets}}$$

The same goes for the c-tag and the q-tag. The value of the purity depends on the branching ratio of $Z \rightarrow$...
b\bar{b}, c\bar{c}, q\bar{q}, so in this paper, we assume \[ BR(Z \to b\bar{b}) = 0.1512 \quad (6) \\
BR(Z \to c\bar{c}) = 0.1203 \quad (7) \\
BR(Z \to q\bar{q}) = 0.428 \quad (8) \]

In this evaluation, we do not include pair-BG.

### B. Flavor Tagging with FPCCDTrackFinder

In Figure 19, the red line shows the b-tag performance, and the blue line shows the c-tag performance. The solid line shows the performance with the DBD ILD VXD and the DBD ILD tracking, the dotted line using the DBD ILD VXD and FPCCDTrackFinder, and the dashed line using the FPCCD VXD and FPCCDTrackFinder. By comparing the solid line with the dotted line, we can see that FPCCDTrackFinder improves the c-tag performance (for example, the c-tag efficiency improves by 2.5% in c-tag purity 70%). Furthermore, by comparing the dotted line with the dashed line, we can see that FPCCD VXD improves b-tag and c-tag performance (for example, b-tag efficiency improves by 2% in b-tag purity 90%, and the c-tag efficiency improves by 4% in c-tag purity 70%).

As a result, FPCCDTrackFinder and the FPCCD VXD improve the flavor tagging performance without considering pair-BG.

### IV. SUMMARY

We have presented FPCCDTrackFinder, a new tracking algorithm. FPCCDTrackFinder improves the tracking efficiency to ~ 99% in $p_T > 0.6$ GeV/c and $|\cos \theta| < 0.9$. FPCCDTrackFinder has enabled the tracking with
FPCCD VXD in the presence of pair-BG for the first time, and achieves a tracking efficiency of \(\sim 99\%\) in \(p_T > 0.6\ \text{GeV}/c\) and \(|\cos \theta| < 0.9\) regardless of the presence of pair-BG. We have also evaluated the performance of flavor tagging in some cases. FPCCDTrackFinder improves flavor tagging performance; c-tag efficiency increases by 2.5% for a c-tag purity of 70%. The FPCCD VXD also improves flavor tagging performance; b-tag efficiency increases by 2% for a b-tag purity of 90%, and the c-tag efficiency increases by 4% for a c-tag purity of 70%.

V. ACKNOWLEDGMENTS

The authors would like to thank all the members of the ILC physics subgroup and the ILD group for useful discussions, Dr. Jan Fridolf Strube for the detailed checks for this paper, and the JSPS grant-in-aid for specially promoted research No. 23000002.

[1] ILC Technical Design Report Volume 2 - Physics, http://www.linearcollider.org/ILC/Publications/Technical-Design-Report, 2012.
[2] Thomas Krämer, Track Parameters in LCIO, http://www-flc.desy.de/lcnotes/notes/LC-DET-2006-004.pdf, 2006.
[3] ILC Technical Design Report Volume 4 - Detectors, http://www.linearcollider.org/ILC/Publications/Technical-Design-Report, 2012.
[4] ILC Technical Design Report Volume 1 - Executive Summary, http://www.linearcollider.org/ILC/Publications/Technical-Design-Report, 2012.
[5] ILCSOft, http://ilcssoft.desy.de/portal
[6] DBD documents, http://ilcild.org/documents/dbd-documents
[7] Keisuke Fujii, The ACF-A Sim-J Group, "Extended Kalman Filter", http://www-jlc.kek.jp/subg/offl/kaltest/doc/ReferenceManual.pdf
[8] LCFIPlus, https://confluence.slac.stanford.edu/display/ilc/LCFIPlus
[9] Particle Data Group, particle listings, GAUGE AND HIGGS BOSONS, http://pdg.lbl.gov/2013/listings/rpp2013-list-z-boson.pdf