~~MC-based feasibility study~~ Simulation Study of ~~new sampling calorimeter for measuring the γ incident angle~~ Angular Resolution of a New Electromagnetic Sampling Calorimeter

YoungJun Kim and Jung Keun Ahn

Department of Physics, Korea University, Seoul 02841

Junlee Kim* and Eun-Joo Kim†

Department of Physics education, Jeonbuk National University, Jeonju 54896

GeiYoub Lim

IPNS/KEK Tsukuba, Japan 305-0801

Abstract

We present studies on the detector configuration to measure the incident angle of the γ with energies ranging from hundred MeV to few GeVusing a new sampling calorimeter. The sampling report simulation results on the angular resolution of electromagnetic (EM) sampling calorimeters with photons being in the range of 100 MeV to 2 GeV. A simulation model of the EM calorimeter consists of alternating layers of a 1-mm-thick lead plates producing the electromagnetic (EM) shower and strips of Pb plate and a 5-mm-thick plastic scintillators measuring the energy deposit by the shower particles. The strips are arranged along the alternating scintillator plate. The scintillator plates are alternatively segmented in horizontal and vertical directions to measure the transverse profile by combining both directions at a given longitudinal positionstrips. In this paper, we used the GEANT4 to simulate energy deposit to the individual strips of the calorimeter, and the *XGBoost* to reconstruct the incident angle from energy deposits.

The angular resolution weakly depends on the width of the strips up to 15 mm and becomes worse with the increasing width. It is found that the incident angle is reconstructed with the compatible angular resolution using front 5 radiation length study, we obtained energy deposit in individual strips using Geant4 simulations and reconstructed incident photon angles using a XGBoost model with boosted decision trees. Angular resolutions are studied in terms of the strip width. The 15-mm-wide strips provide the best angular resolution of $1.23 \pm 0.01^\circ$. Incident photon directions can be well reconstructed in the front part ($5X_0$) of the detector. The resolution does not depend on the incident angle up to EM calorimeter. The angular resolution is almost constant for small angles less than 30degrees, and largely depends on the incident energy, which °, and varies significantly with the incident photon energy. The energy dependence can be expressed as $0.2+1.1/\sqrt{E_\gamma}\sigma_\theta$ = xxx $\oplus$ yyy $/\sqrt{E_\gamma}$.

PACS numbers:

Keywords: Electromagnetic Calorimeter, Geant4, XGBoost, KOTO

*E-mail: junlee.kim@cern.ch

†E-mail: ejkim@jbnu.ac.kr

I. MOTIVATION

The ~~EM calorimeter has played~~ electromagnetic (EM) calorimeter plays an important role in experimental studies for the nuclear and particle physics. Various materials have been developed for better energy and timing ~~resolution~~ resolutions so far [1]. EM calorimeters can be divided into two groups: sampling calorimeters and homogeneous calorimeter. Sampling calorimeters consist of alternating layers of an absorber generating EM showers and an active medium providing signals. On the other ~~side,~~ hand, homogeneous calorimeters are built of only a single type of material that facilitates both EM showers and signal generations. Today, the sampling calorimeter becomes ~~popular in the~~ very popular in large-scale high energy ~~experiment mainly due~~ experiments, owing to its cost-effectiveness. ~~The sampling calorimeter consists of alternating passive converters that generate the EM shower and active counters that measure the energy deposit. Since the energy deposit in the passive convert is not measured, the fluctuation of energy deposits in the converter and the counter, called sampling fluctuation , determines~~ In sampling calorimeters the energy deposited in the active medium fluctuates because the active layers are interleaved with absorber layers. This so-called sampling fluctuation dominates the energy resolution~~of the calorimeter.~~ .

The sampling fluctuation can be ~~optimized by combining the converter and the counter with a specific portion. A 5-m-long~~ improved by optimizing the thickness ratio of the absorber to the active medium. A 3-m-long cylindrical calorimeter made of alternating ~~lead plates~~ layers of Pb and plastic scintillating plates is an example of the sampling calorimeter ~~[2]~~ [3].

The ~~layer~~ segmented structure of the ~~sampling calorimeter allow one to measure the individual profile of the~~ active medium facilitates the measurement of lateral distributions of the EM shower along the ~~beam direction, and the incident angle can be estimated by correlating energy deposits on neighboring layers~~incident photon direction. The photon direction can also be deduced from the information on the EM shower shape. The measurement of ~~the incident angle largely benefits the background rejection. Of particular, The incident angle will be an important tool for the~~ photon directions can help predict photon production vertices, which is crucial for the background suppression in KOTO experiment ~~[4]~~ [5].

~~Since~~ Stochastic behaviors of the EM shower ~~is evolving via stochastic processes, the~~

~~incident angle would be reconstructed with a certain angular resolution~~ development limit to the angular resolution of the photon direction measurement. Energy deposits ~~from random processes are used to reconstruct the incident angle with the machine learning, which provides better angular resolution. The evolution of the EM shower is generated with the GEANT4, and the *XGBoost* was used~~ in each active layer of a sampling calorimeter fluctuate event by event so that the reconstruction of photon directions necessitates the longitudinal and lateral segmentation of the calorimeter.

This paper describes a machine learning approach based on XGBoost (XGB) model [6] with boosted decision trees in order to reconstruct the ~~incident angle.~~

~~In sec II. , generic explanation of the EM shower and its evolution are described. We describe the reconstruction and~~ directions of incident photons entering a sampling calorimeter with alternating layers of a Pb plate and plastic scintillator strips in horizontal and vertical directions. We generated EM showers in a sampling calorimeter using Geant4 simulations. We report on the detector configuration providing the acceptable angular resolution, and report on the angular resolution of the EM sampling calorimeter in ~~sec III. . We summarize the study in sec IV.~~ terms of the incident angle and energy.

II. ~~ELECTROMAGNETIC SHOWER~~ELECTROMAGNETIC SHOWER SIMULATION

~~The sampling calorimeter was designed as alternating~~ Simulation models of a sampling calorimeter were constructed as a block consisting of alternating layers of a 1-mm-thick ~~lead plates for passive converters and~~ Pb absorber and a 5-mm-thick ~~strips of the plastic scintillatorfor active counters. Strips are alternatively aligned along x and y direction~~ polyvinyltoluene-based plastic scintillator. The plastic scintillator is segmented in 15-mm-wide strips, which is alternatively oriented in vertical and horizontal directions as shown in Fig. 1. The ~~cross section of the lead plate and the scintillator is~~ sampling calorimeter model has a cross-section of 525~~mm~~ × 525~~mm and 15 mm × 525 mm, respectively. The material for the scintillator strip is polyvinyltoluene. The total number of alternating layers is 105, which corresponds to~~ mm^2 and accommodates 105 alternating layers of 630 mm length ($20X_0$~~to make it possible to absorb whole products from the EM shower ranging from~~), which is sufficiently long to absorb full photon energy in the range

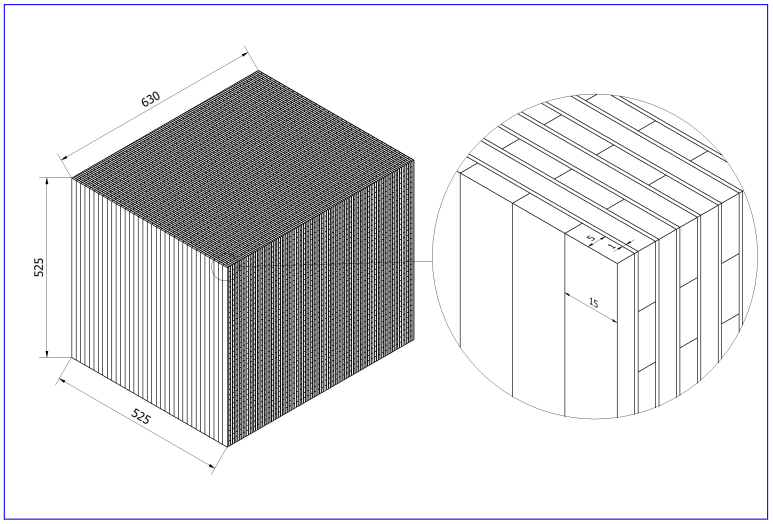


Fig. 1. Schematic ~~view~~ of the sampling calorimeter ~~. It consists~~ model consisting of 105 alternating ~~lead plates~~ layers of a Pb plate and a segmented scintillator plate in 35 strips. Each scintillator ~~layer consists of 35 scintillator strips. The cross section of each 1-mm-thick lead~~ plate is ~~525 mm × 525 mm~~oriented alternatively in horizontal and vertical directions. ~~The cross section of each 5-mm-tihck scintillator strip is 525 mm × 15 mm~~See text for details.

of 100~~MeV to few GeVenergy~~ MeV to 2 GeV.

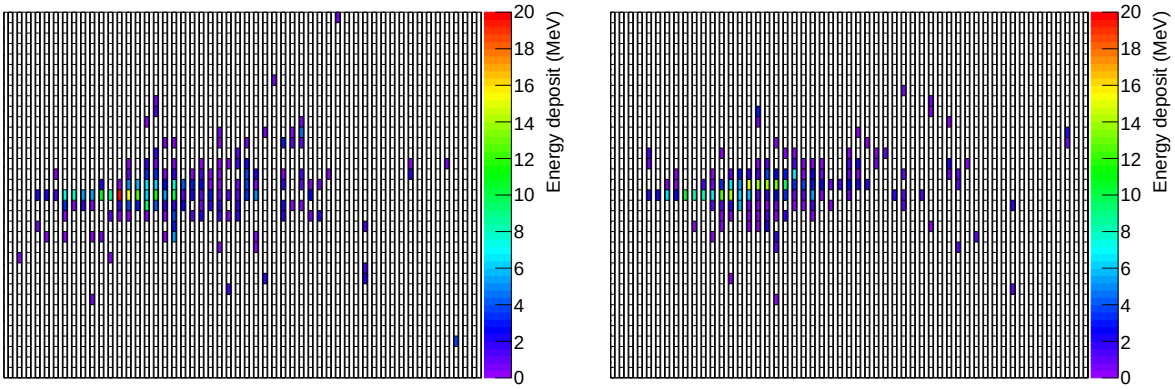


Fig. 2. ~~An event~~ Event display of ~~the~~ simulated energy deposit ~~to each scintillator strip in the x–z plane (left) and the y–z plane (right)~~ patterns for ~~the 1 GeV γ perpendicularly~~ a 1-GeV photon entering ~~to~~ the ~~detector~~ calorimeter ($\theta = 0$) in (a) xz- and (b) yz-planes.

The ~~interaction of incident γ with the detector~~ detector response to incident photons was simulated using ~~the GEANT4~~ Geant4 (ver. 4.10.06) with the standard EM sub-packages [7]. The ~~surface of the detector is defined as $z = 0$, and the γ starts to interact with the detector.~~

~~the incident angle of the γ~~ beam direction defines the z-axis. The photon direction is defined as ~~the~~ polar angle (θ) ~~along~~ with respect to the z~~direction, where $\theta = 0$ case denotes that the γ perpendicularly enters into the detector surface~~-axis. Figure 2 ~~shows the event display illustrating the energy deposit on the each scintillator strip in both of x–z and y–z planes for the 1 GeV γ with $\theta = 0$~~illustrates simulated energy deposit patterns in each strip for a 1-GeV photon at a normal incidence in xz- and yz-planes. Each segmented region in Fig. 2 represents each channel.

III. ~~ANGLE RECONSTRUCTION~~RECONSTRUCTION OF INCIDENCE ANGLES

~~The incident angle of the γ is reconstructed with the *XGBoost*, which is one of the popular machine learning toolkit providing a scalable tree boosting system~~Incidence angles of photons are reconstructed using the XGB model that is a scalable machine learning system for tree boosting [6]. ~~The machine learning is required to be trained. The feature, characteristic of a phenomenon, and the target, output to be predicted, are essential inputs for the training. The machine learning correlates each feature with the corresponding target for different features. Trained machine learning takes measured features and gives the target based on correlations which are made during the training. In this paper, the feature corresponds to energy deposits on~~ A machine learning model maps a set of data inputs, known as features, to target variables. In this study, the XGB model maps a dataset of energy deposits in each scintillator strip ~~and the target corresponds to the incident angle of the γ.~~

~~If the feature for the training is biased, the prediction of the *XGBoost* would be correspondingly biased as the *XGBoost* reliably believes there are no biases in all features given by a user~~to an incidence angle of a photon. Training data were carefully prepared using Geant4 simulation such that the input datasets are representative of a detector response of real sampling calorimeters. To minimize ~~this, the incident angle is uniformly generated in the range~~ data bias, the incidence angles were uniformly generated at the detector surface in the angular range of 0~~to~~ $< \theta <$ 50~~degrees for incident polar angle (θ)~~ $^\circ$ and 0~~to~~ $< \varphi <$ 360~~degrees for incident azimuthal angle~~($^\circ$, where φ ~~), which makes the *XGBoost* train all features we are interested in~~denotes azimuthal angle. Such a wide angular coverage provides good training

datasets for the XGB. The number of ~~samples for the training is 2×10^5 considering~~ training samples is 10^5 considering limited computing resources. To test the reconstruction of ~~the incident angle, the incident angle is generated with the fixed~~ incidence angles, we generated photons at a fixed incidence angle θ. Note that the incident energy is known.

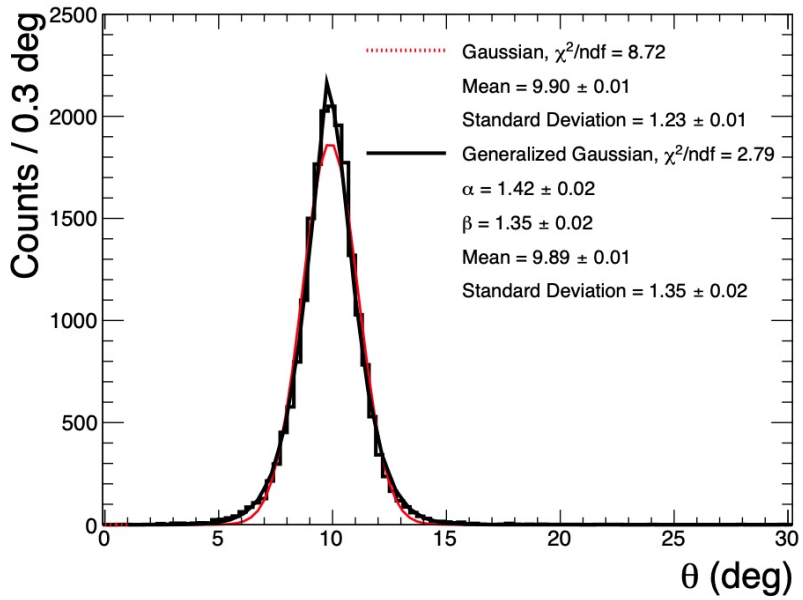


Fig. 3. Reconstructed ~~θ~~ polar angle distribution for 1-GeV photons generated at $\theta = 10$ ~~deg events~~°. The distribution is fitted with the Gaussian ~~function~~ and the Generalized Gaussian function. The ~~Generalized Gaussian function provides~~ later gives a better ~~description for both tails~~result.

Figure 3 ~~shows reconstructed~~ represents a distribution of reconstructed incidence angles θ for ~~1 GeV γ with~~ 1-GeV photons generated at $\theta = 10$ ~~degrees~~°. The distribution is fitted with the Gaussian function and the Generalized Gaussian (GG) function. ~~GG function provides better description for both tails of the~~ We tested two functions, Gaussian and GG to describe the reconstructed angle distribution. The GG function~~can be expressed as~~

$$f(x;\mu,\alpha,\beta) = \frac{\beta}{2\alpha\Gamma(1/\beta)} e^{-(|x-\mu|/\alpha)^\beta}$$

, also known as the generalized error distribution, is expressed as

$$f(x;\mu,\alpha,\beta) = \frac{\beta}{2\alpha\Gamma(1/\beta)} e^{-(|x-\mu|/\alpha)^\beta}, \quad (1)$$

~~The variation of the GG function can be expressed as, then,~~ where μ denotes a mean value. Parameters α and β determine the scale and shape of the distribution, respectively. Variance of the GG function is given by $\sigma^2 \equiv \alpha^2\Gamma(3/\beta)/\Gamma(1/\beta)$. The angular resolution of the ~~incident~~ incidence angle reconstruction is defined as σ.

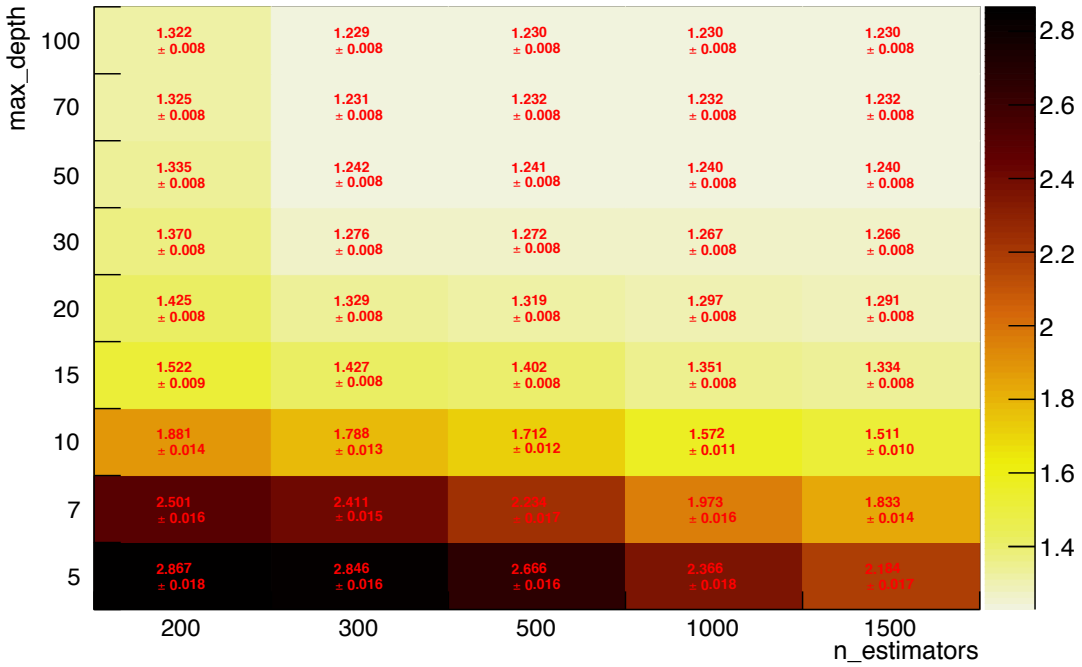


Fig. 4. ~~The angular resolution for the varying~~ Angular resolutions are displayed in terms of combination of N_estimators and ~~the varying~~ Max. depth. The best angular resolution is obtained with N_estimators = 300 and Max. depth = 100.

~~As the reconstruction of the incident angle depends on hyperparameters of the *XGBoost*, which controls the details of training processes, dedicated tests to optimize the hyperparameter were executed. It is assumed that a set of hyperperameters providing the best angular resolutionwould be optimized~~We studied a correlation between hyperparameters of the XGB model and the angular resolution. The test scans the evaluated angular resolution with different hyperparameters. Figure 4 shows the ~~result of the test for~~ test results with N_estimators and Max. depth. N_estimators defines ~~allowed maximal~~ the maximum allowed number of decision trees to be developed, and the Max. depth defines the complexity of the structure of decision trees. The best hyperparameter combination was searched for in terms of the angular resolution. As a results, N_estimators and Max. depth are ~~determined~~

~~to be~~ set to 300 and 100, respectively. Similar tests ~~are applied to different hyperparameters, and definitive values for each hyperparameter are shown~~ were also performed for different hyperparameters. The best values for other hyperparameters are displayed in Tab 1. Other hyperparameters tuned in this study are Subsample, Learning rate, and Gamma. Subsample controls the fraction of total event samples for each boosting procedure, and Learning rate weights a decision tree to be added current model. Lastly, Gamma regulates the evaluation of each decision tree.

Table 1. Hyperparameters of the *XGBoost*

Parameter	Function	Default value	Used value
N_estimators	The number of decision trees	N.A.	300
Max. depth	Possible maximum depth of tree structure	6	100
Subsample	Fraction of total data used for a single decision	1	0.8
Learning rate	Step length for calculation	0.3	0.02
Gamma	Requirement on minimum loss function	0	0

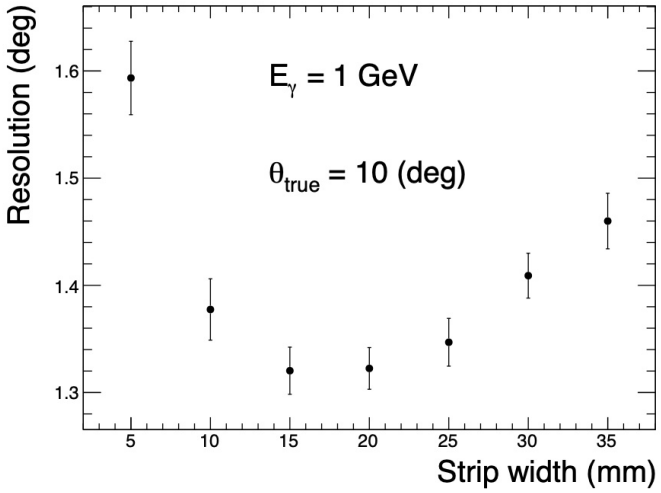
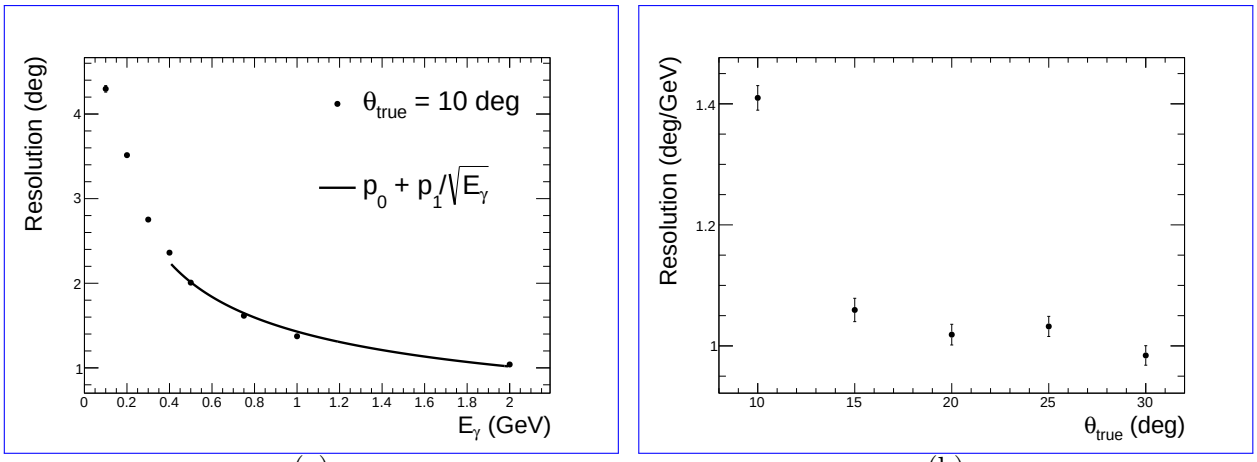


Fig. 5. ~~The angular resolution as a function~~ Angular resolutions are deduced in terms of the ~~scintillator~~ strip width for ~~1 GeV γ with~~ 1-GeV photons at θ = 10 ~~deg. The angular resolution is optimized with 15- to 20-mm-wide strips.~~°

~~The angular resolution was evaluated with the width of scintillator strips varying~~ Angular resolutions were deduced with varying strip widths from 5 mm to 35 mm as shown in Fig. 5.

~~Figure 5shows the angular resolution as a function of the width.~~ The 15-mm-wide strips ~~provide~~ yield the best angular resolution ~~. The width longer than 15 mm suffers from the dilution of the EM shower, resulting in worse resolution. On the other hand, the width shorter than 15 mm results in~~ of 1.23 ± 0.01°. The shorter the strip width is, the larger size of features, which ~~causes a negative influence on the~~ can influence negatively in machine learning. ~~Dedicated study is described in fig 5~~For longer strips, the EM shower information can hardly be obtained.

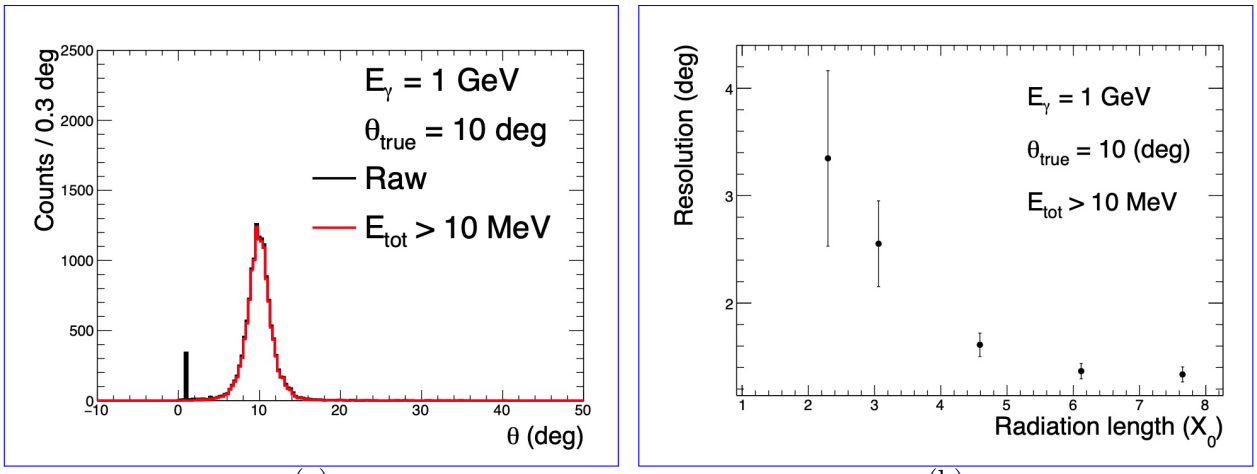


(a) (b)

Fig. 6. ~~The~~ Deduced angular ~~resolution~~ resolutions as ~~a function~~ functions of ~~the γ energy (left), and p_1 for 1 GeV as a function of the incident angle (right)~~ . p_1 ~~is consistent down to~~ photon energy for $\theta =$ ~~15 deg~~ 10° and ~~changing at $\theta =$~~ (b) polar angle for $E_\gamma =$ ~~101 deg, which is coming from the definition of θ~~ GeV.

Figure 6 (a) shows the angular resolution as a function of the incident ~~γ energy~~ photon energy ($E_\gamma =$) for $\theta = 10$ ~~degrees on the left panel. The resolution is fitted with $p_0 + p_1/\sqrt{E_\gamma(GeV)}$, where the~~ °. Angular resolutions are fitted with a function of $p_0 \oplus p_1/\sqrt{E_\gamma(GeV)}$, where p_0 ~~denotes the~~ represents an energy-independent contribution and is estimated to be ~~0.2, and the~~ xxx, and p_1 denotes the energy-dependent ~~contribution~~parameter, mainly related with the development of the EM shower. ~~The estimated p_1 for different θ can be seen on the right panel of fig~~is estimated to be xxx for $\theta > 15°$ and xxx° ± xxx° for $\theta = 10°$ in Fig. 6 (b). The discrepancy between them is attributed to the followings: polar angle to be always positive and poor angular resolutions for low E_γ photons. When μ and α get comparable each other in Eq 1, the left tail of the distribution would extend to negative area. This largely affects the case $\mu = 10°$ and not

the case $\mu > 15°$ considering that the angular resolutions for $E_\gamma = 0.1$ GeV is estimated to be 4°. ~~They are deviating within ± 0.05 deg range from the 1.1 deg.~~



(a) (b)

Fig. 7. (a) Reconstructed polar angles (θ) for ~~1 GeV γ with~~ 1-GeV photons at $\theta = 10°$. Red histogram represents the polar angle distribution by requiring $E_{\rm tot} >$ ~~deg with~~ 10 MeV. (~~red~~b) ~~and without (black) the total energy ($E_{\rm tot}$) selection (left), and the angular~~ Angular resolution as a function of ~~front layer depth used for the reconstruction~~ radiation length (~~right~~X_0). ~~The inefficiency for $4.6X_0$ is estimated to be 5.8% with 10 MeV selection. The large error bars less than $4.6X_0$ is due to insufficient fitting because the reconstructed angular distributions deviated from the General Gaussian function.~~

~~The reconstruction of the incident angle is tested~~ Furthermore, incidence angles are reconstructed with the front ~~layer instead of the full layers considering the cost-effectiveness from the fact that~~ layers of X_0 in which 99% of ~~the γ generates the EM shower in front of $5X_0$. The left panel of~~ incident photons generate EM showers. Figure 7 ~~shows reconstructed angle using the~~ (a) shows reconstructed angles using front 24 layers of the detector, which corresponds to ~~the~~ $4.6X_0$ for ~~1 GeV γ. Some γs are~~ 1-GeV photons. A fraction of photons failed to be reconstructed due to the lack of ~~the channel having the energy deposit~~active layers. The failed events are represented as a delta function near 0, and such events are removed by requiring the total energy deposit to be larger than 10 MeV. The angular resolution with the front layer is estimated to be XXX while possessing inefficiency estimated to be 5.8%.

Figure 8 shows the angular resolution with different numbers of training samples for different detector configurations. All configurations show a decrease in the angular resolution

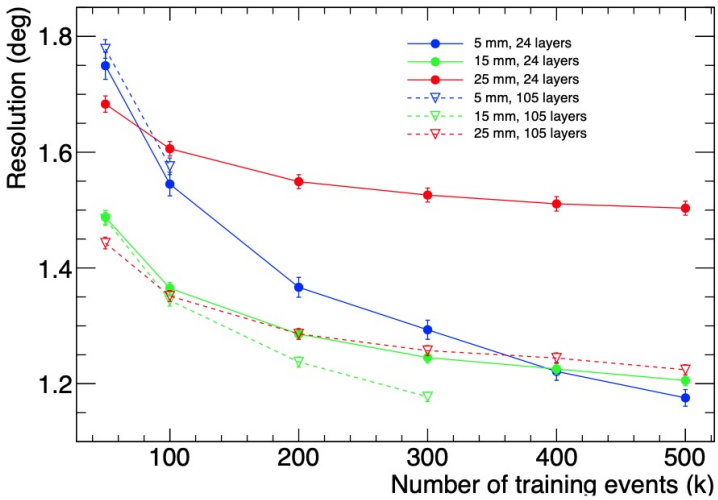


Fig. 8. The angular resolution for ~~1 GeV γ with~~ 1-GeV photons at $\theta = 10$ ~~deg~~ ° as a function of the number of training samples with different strip widths using the first 24 layers or the full layers.

with increasing training samples, which means that enhanced statistics for the training results in better reconstruction. The angular resolution with 5-mm-wide scintillator strips ~~is more rapidly decreasing~~ decreases more rapidly than others. The number of features for the 5-mm-wide strip is much larger than others, and the number of training samples is much required to correlate larger features. Considering limited computing resources, ~~The setup with 2×10^5~~ the setup with 10^5 events, 24 layers, and the 15-mm-wide strip ~~is chosen for the training of~~ was chosen for training the *XGBoost* ~~, where the result is under~~ model. The training result differs by only 0.1~~degree difference~~ ° from the best setup will full layers.

IV. SUMMARY

We ~~conduct a feasibility study to measure the incident angle of the γ with the finely segmented sampling calorimeter. With alternating 1-mm-thick lead plates and 5-mm-thick plastic scintillator strips, the EM shower is largely generated in lead sheets, and is measured by scintillators.The incident angle the γ is reconstructed with the~~ report simulation results on the incidece angle reconstruction of EM sampling calorimeters with photons being in the range of 100 MeV to 2 GeV. EM shwers are simulated in the new samplinng calorimeter using Geant4. We utilize the *XGBoost* ~~using the energy deposited in each scintillator strip~~ model

to reconstruct the photon incident angle by correlating all energy deposits in each strip with the incidence angle all events together.

We ~~optimize~~ tune the hyperparameter of the *XGBoost* ~~, and optimize the detector configuration using the optimized *XGBoost*. We scan the hyperparameter in the 5-dimensional space and find the set providing the best angular resolution for different detector configurations~~to reconstruct the incidence angle with the best angular resolution, and study the angular resolution in terms of the detector configuration. We find that 15-mm-wide strips provide the best angular resolution, which can be expressed as 0.2~~+~~ $\oplus$ 1.1~~/$\sqrt{E_\gamma}$~~ /$\sqrt{E_\gamma}$ for different incident angles. We also ~~find~~ conclude that using 24 front layers gives the angular resolution which is ~~compatible~~ comparable with the angular resolution we get from the full layer.

~~The~~ We study that the angular resolution is changing with the variation of the training of the *XGBoost* and the detector configuration. As the strip width decreases, more channels of scintillator strips are required, and the training of the *XGBoost* gets affected due to the increased feature size. This is quenched with the larger number of training samples, which qualitatively improves the training of the *XGBoost*. On the other hand, the longer strip width provides the worse angular resolution as the position resolution of the EM shower gets worse. We conclude that 15-mm-wide strips provide the favorable angular resolution with 24 front layers and ~~2×10^5~~ 10^5 training samples.

ACKNOWLEDGMENTS

REFERENCES

[1] R.Wigmans, Prog. in Part. and Nucl. Phys. **103**, 109 (2018).

[2] Y.Tajima et al., Nucl. Instrum. Meth. A **592**, 261 (2008).

[3] R. Murayama et al., Nucl. Instrum. Meth. A **953**, 163255 (2020).

[4] J.Comfort et al. (2006), URL https://j-parc.jp/researcher/Hadron/en/pac_0606/pdf/p14-Yamanaka.pdf.

[5] T. Yamanaka and for the KOTO Collaboration (KOTO), PTEP **2012**, 02B006 (2012).

[6] T. Chen and C. Guestrin (2016), 1603.02754.

[7] S. Agostinelli et al., Nuclear Instruments and Methods in Physics Research Section A **506**, 250 (2003), ISSN 0168-9002.