

# MC-based feasibility study of new sampling calorimeter for measuring the $\gamma$ incident angle

YoungJun KIM and Jung Keun AHN

*Department of Physics, Korea University, Seoul 02841*

Junlee KIM\* and Eun-Joo KIM<sup>†</sup>

*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  $\gamma$  with energies ranging from hundred MeV to few GeV using a new sampling calorimeter. The sampling calorimeter consists of alternating 1-mm-thick lead plates producing the electromagnetic (EM) shower and strips of 5-mm-thick plastic scintillators measuring the energy deposit by the shower particles. The strips are arranged along the alternating horizontal and vertical directions to measure the transverse profile by combining both directions at a given longitudinal position. 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 ( $5X_0$ ) of the detector. The resolution does not depend on the incident angle up to 30 degrees, and largely depends on the incident energy, which can be expressed as  $0.2+1.1/\sqrt{E_\gamma}$ .

PACS numbers:

Keywords:

---

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

†E-mail: ejkim@jbnu.ac.kr

## I. MOTIVATION

The EM calorimeter has played an important role in experimental studies for the nuclear and particle physics. Various materials have been developed for better energy and timing resolution so far [1]. On the other side, the sampling calorimeter becomes popular in the large-scale high energy experiment mainly due 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 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 cylindrical calorimeter made of alternating lead plates and plastic scintillating plates is an example of the sampling calorimeter [2].

The layer structure of the sampling calorimeter allow one to measure the individual profile of the shower along the beam direction, and the incident angle can be estimated by correlating energy deposits on neighboring layers. The measurement of the incident angle largely benefits the background rejection. Of particular, The incident angle will be an important tool for the KOTO experiment [3].

Since the EM shower is evolving via stochastic processes, the incident angle would be reconstructed with a certain angular resolution. 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 to reconstruct the incident angle.

In sec II. , generic explanation of the EM shower and its evolution are described. We describe the reconstruction and angular resolution of the sampling calorimeter in sec III. . We summarize the study in sec IV. .

## II. ELECTROMAGNETIC SHOWER

The sampling calorimeter was designed as alternating 1-mm-thick lead plates for passive converters and 5-mm-thick strips of the plastic scintillator for active counters. Strips are alternatively aligned along  $x$  and  $y$  direction as shown in Fig. 1. The cross section of the lead plate and the scintillator is 525 mm  $\times$  525 mm and 15 mm  $\times$  525 mm, respectively.

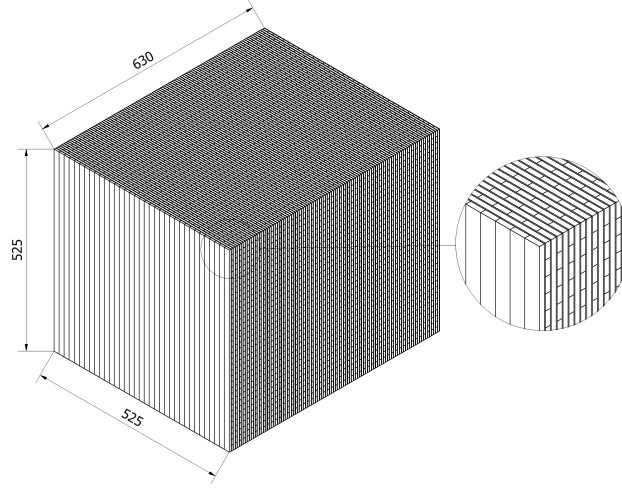


Fig. 1. Schematic view of the sampling calorimeter. It consists of 105 alternating lead plates and scintillator strips. Each scintillator layer consists of 35 scintillator strips. The cross section of each 1-mm-thick lead plate is 525 mm  $\times$  525 mm. The cross section of each 5-mm-thick scintillator strip is 525 mm  $\times$  15 mm.

The material for the scintillator strip is polyvinyltoluene. The total number of alternating layers is 105, which corresponds to  $20X_0$  to make it possible to absorb whole products from the EM shower ranging from 100 MeV to few GeV energy.

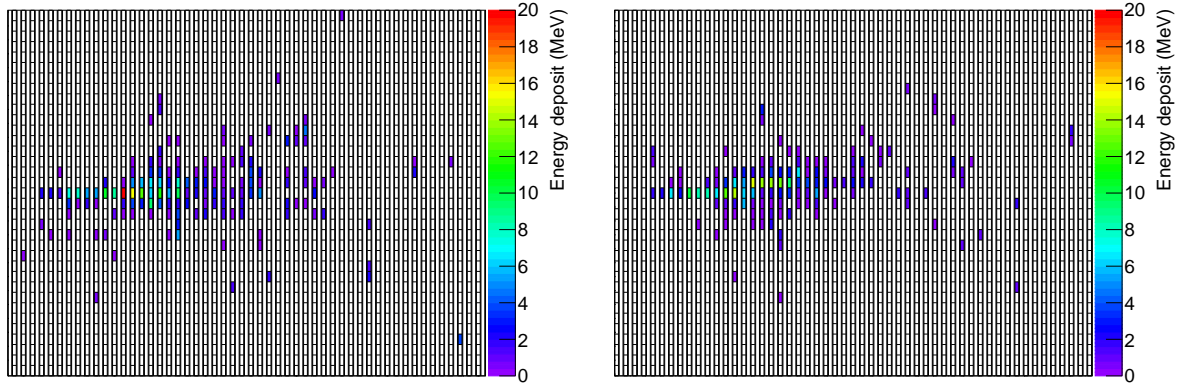


Fig. 2. An event display of the energy deposit to each scintillator strip in the  $x$ - $z$  plane (left) and the  $y$ - $z$  plane (right) for the 1 GeV  $\gamma$  perpendicularly entering to the detector ( $\theta = 0$ ).

The interaction of incident  $\gamma$  with the detector was simulated using the GEANT4 (ver. 4.10.06) with the standard EM sub-packages [4]. The surface of the detector is defined as  $z = 0$ , and the  $\gamma$  starts to interact with the detector. the incident angle of the  $\gamma$  is defined as

the polar angle ( $\theta$ ) along the  $z$  direction, where  $\theta = 0$  case denotes that the  $\gamma$  perpendicularly enters into the detector surface. 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  $\gamma$  with  $\theta = 0$ .

### III. ANGLE RECONSTRUCTION

The incident angle of the  $\gamma$  is reconstructed with the *XGBoost*, which is one of the popular machine learning toolkit providing a scalable tree boosting system [5]. 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 each scintillator strip and the target corresponds to the incident angle of the  $\gamma$ .

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 minimize this, the incident angle is uniformly generated in the range 0 to 50 degrees for incident polar angle ( $\theta$ ) and 0 to 360 degrees for incident azimuthal angle ( $\varphi$ ), which makes the *XGBoost* train all features we are interested in. The number of samples for the training is  $2 \times 10^5$  considering computing resources. To test the reconstruction of the incident angle, the incident angle is generated with the fixed  $\theta$ . Note that the incident energy is known.

Figure 3 shows reconstructed  $\theta$  for 1 GeV  $\gamma$  with  $\theta = 10$  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 distribution. The GG function can be expressed as

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

The variation of the GG function can be expressed as, then,  $\sigma^2 \equiv \alpha^2\Gamma(3/\beta)/\Gamma(1/\beta)$ . The angular resolution of the incident angle reconstruction is defined as  $\sigma$ .

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 hyperparam-

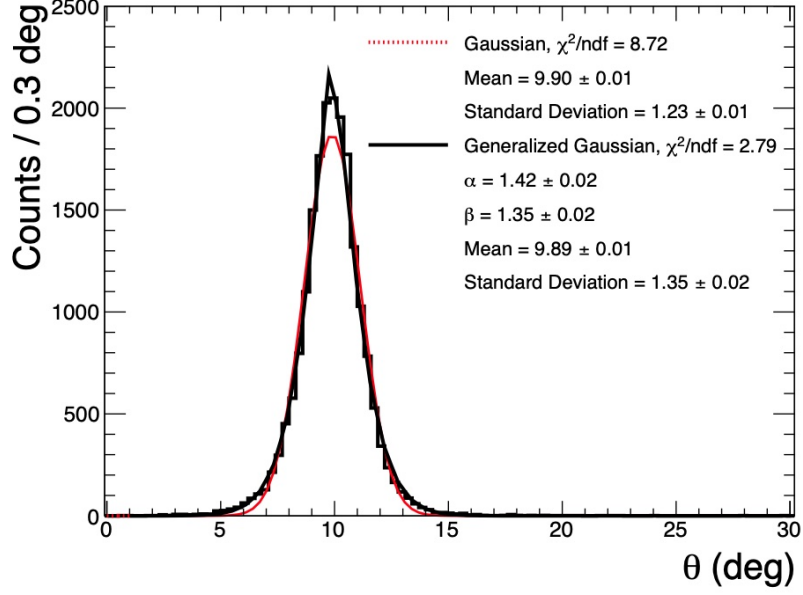


Fig. 3. Reconstructed  $\theta$  distribution for  $\theta = 10$  deg events. The distribution is fitted with the Gaussian function and the Generalized Gaussian function. The Generalized Gaussian function provides better description for both tails.

eter were executed. It is assumed that a set of hyperparameters providing the best angular resolution would be optimized. The test scans the evaluated angular resolution with different hyperparameters. Figure 4 shows the result of the test for N\_estimators and Max. depth. N\_estimators defines allowed maximal number of decision trees, and the Max. depth defines complexity of the structure of decision trees. The N\_estimators and Max. depth are determined to be 300 and 100, respectively. Similar tests are applied to different hyperparameters, and definitive values for each hyperparameter are shown in Tab 1.

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

The angular resolution was evaluated with the width of scintillator strips varying from 5

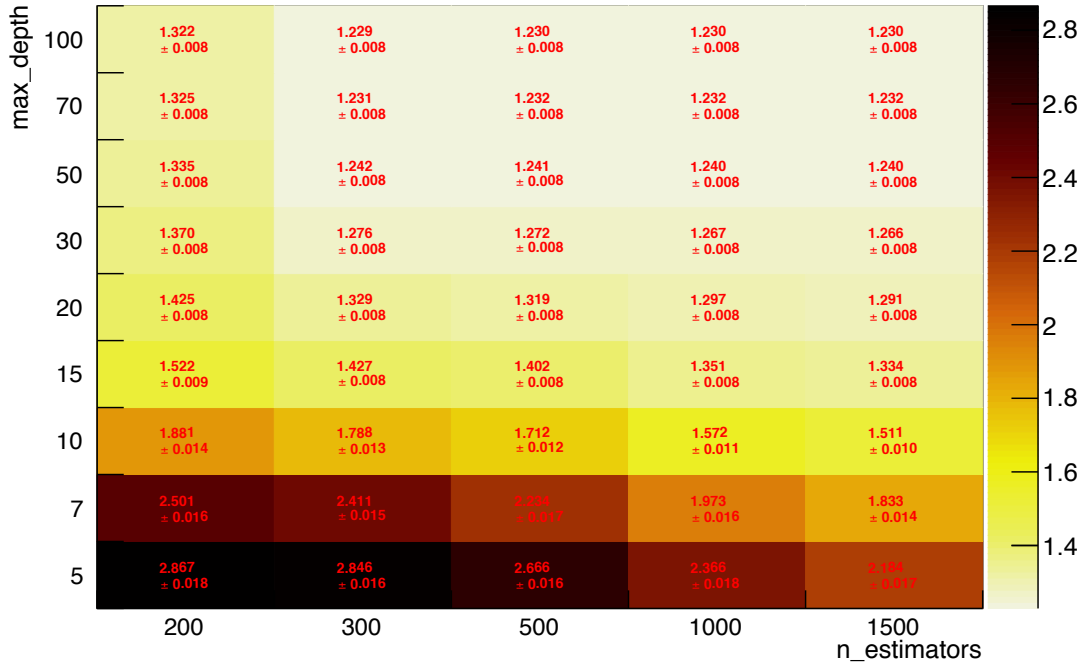


Fig. 4. The angular resolution for the varying N\_estimators and the varying Max. depth. The best angular resolution is obtained with N\_estimators=300 and Max. depth=100.

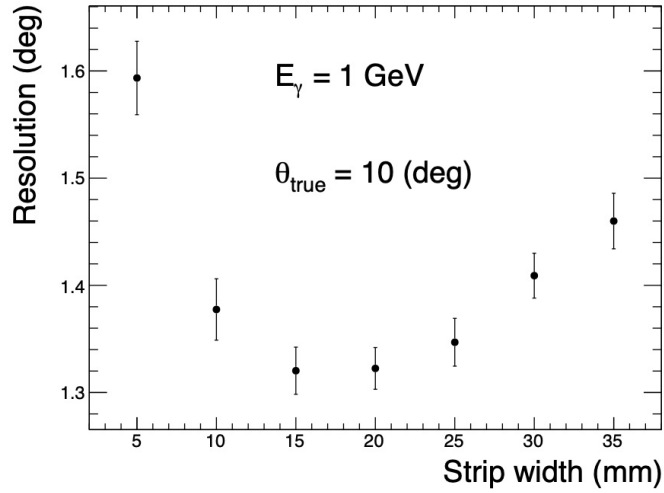


Fig. 5. The angular resolution as a function of the scintillator strip width for 1 GeV  $\gamma$  with  $\theta = 10$  deg. The angular resolution is optimized with 15- to 20-mm-wide strips.

mm to 35 mm. Figure 5 shows the angular resolution as a function of the width. 15-mm-wide strips provide 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 larger size of features, which causes a negative influence on the machine learning. Dedicated study is described in fig 5.

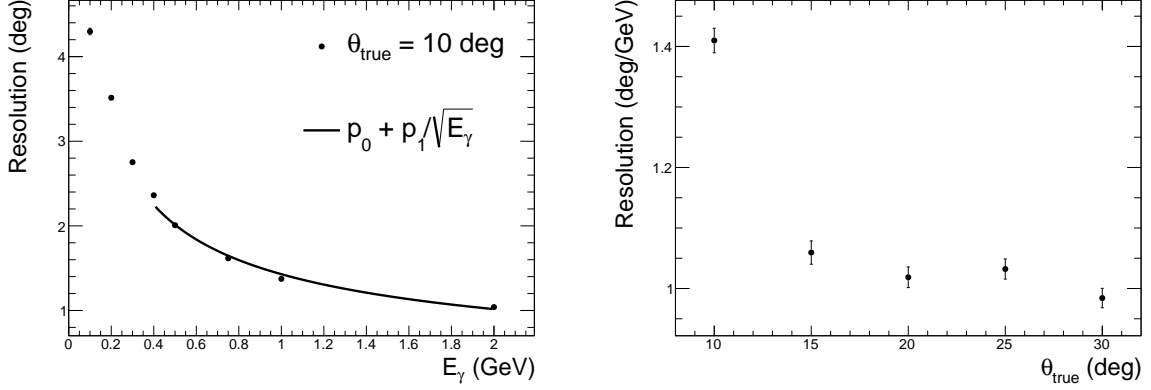


Fig. 6. The angular resolution as a function of the  $\gamma$  energy (left), and  $p_1$  for 1 GeV as a function of the incident angle (right).  $p_1$  is consistent down to  $\theta = 15$  deg and changing at  $\theta = 10$  deg, which is coming from the definition of  $\theta$

Figure 6 shows the angular resolution as a function of the incident  $\gamma$  energy for  $\theta = 10$  degrees on the left panel. The resolution is fitted with  $p_0 + p_1/\sqrt{E_\gamma(\text{GeV})}$ , where the  $p_0$  denotes the energy-independent contribution and is estimated to be 0.2, and the  $p_1$  denotes the energy-dependent contribution, mainly related with the development of the EM shower. The estimated  $p_1$  for different  $\theta$  can be seen on the right panel of fig. 6. They are deviating within  $\pm 0.05$  deg range from the 1.1 deg.

The reconstruction of the incident angle is tested with the front layer instead of the full layers considering the cost-effectiveness from the fact that 99% of the  $\gamma$  generates the EM shower in front of  $5X_0$ . The left panel of Figure 7 shows reconstructed angle using the front 24 layers of the detector, which corresponds to the  $4.6X_0$  for 1 GeV  $\gamma$ . Some  $\gamma$ s are failed to be reconstructed due to the lack of the channel having the energy deposit. 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



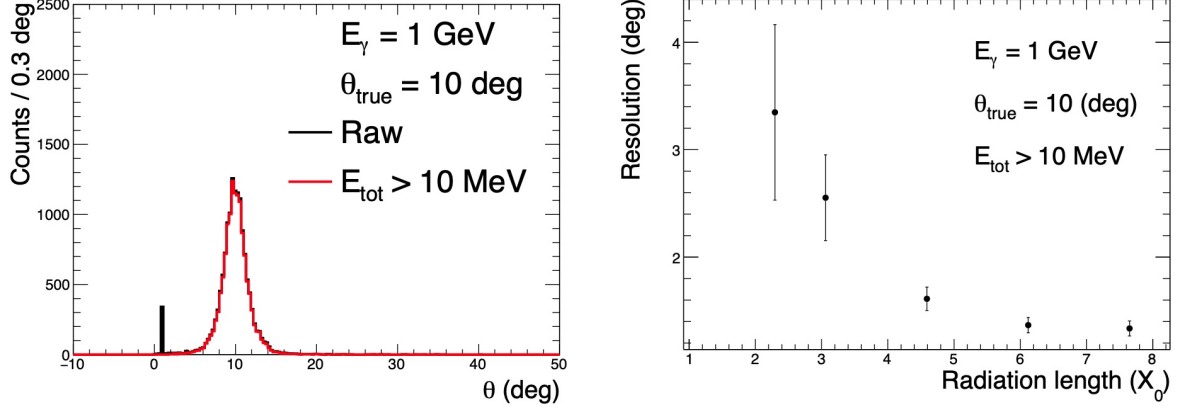


Fig. 7. Reconstructed  $\theta$  for 1 GeV  $\gamma$  with  $\theta = 10$  deg with (red) and without (black) the total energy ( $E_{\text{tot}}$ ) selection (left), and the angular resolution as a function of front layer depth used for the reconstruction (right). The inefficiency for  $4.6X_0$  is estimated 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.

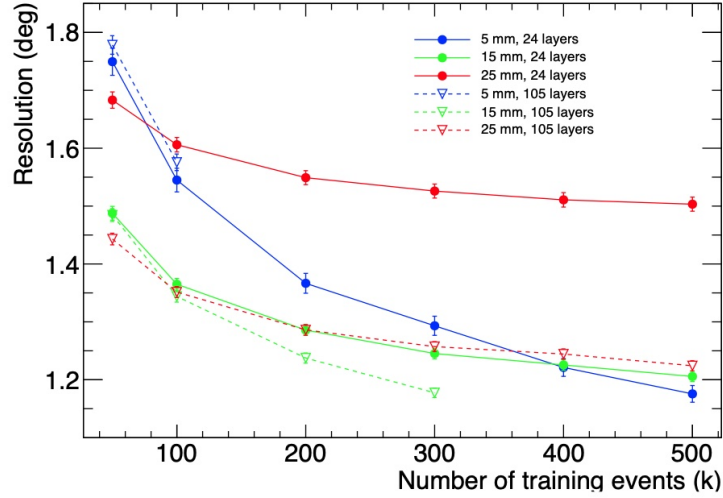


Fig. 8. The angular resolution for 1 GeV  $\gamma$  with  $\theta = 10$  deg as a function of the number of training samples with different strip widths using the first 24 layers or the full layers.

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 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  $2 \times 10^5$  events, 24 layers, and the 15-mm-wide strip is chosen for the training of the *XGBoost*, where the result is under 0.1 degree difference from the best setup.

#### IV. SUMMARY

We conduct a feasibility study to measure the incident angle of the  $\gamma$  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  $\gamma$  is reconstructed with the *XGBoost* using the energy deposited in each scintillator strip.

We optimize 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. We find that 15-mm-wide strips provide the best angular resolution, which can be expressed as  $0.2 + 1.1/\sqrt{E_\gamma}$  for different incident angles. We also find that using 24 front layers gives the angular resolution which is compatible with the angular resolution we get from the full layer.

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  $2 \times 10^5$  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] J.Comfort et al. (2006), URL [https://j-parc.jp/researcher/Hadron/en/pac\\_0606/pdf/p14-Yamanaka.pdf](https://j-parc.jp/researcher/Hadron/en/pac_0606/pdf/p14-Yamanaka.pdf).
- [4] GEANT4 EM sub-packages and processes (???), URL <https://geant4.web.cern.ch/node/1828>.
- [5] T. Chen and C. Guestrin, p. 785–794 (2016), URL <https://doi.org/10.1145/2939672.2939785>.