STCF ECAL Fast simulation with DCGAN

Yujie Chai, Teng Li, Xiaoshuai Qin, Xingtao Huang

Shandong University

The 6th International Workshop on Future Tau Charm Facilities

November 21, 2024

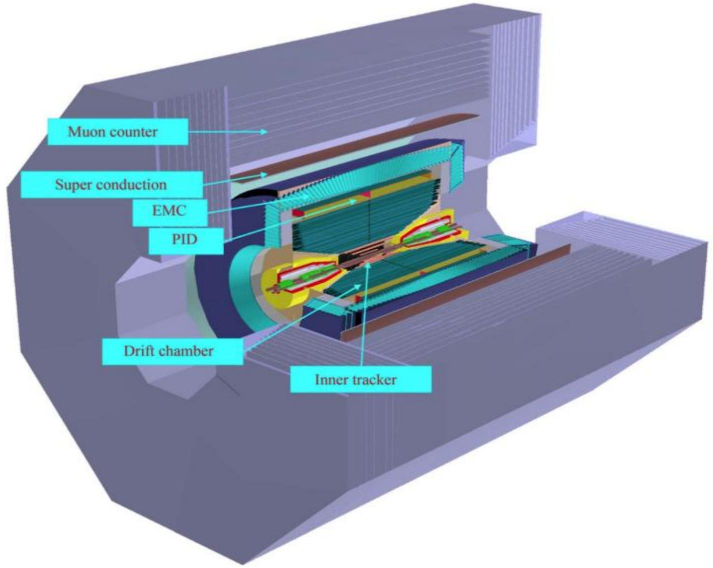
Outline

- Introduction and Motivation
- Methodology
- Preliminary Results
- Summary and Plans

Introduction

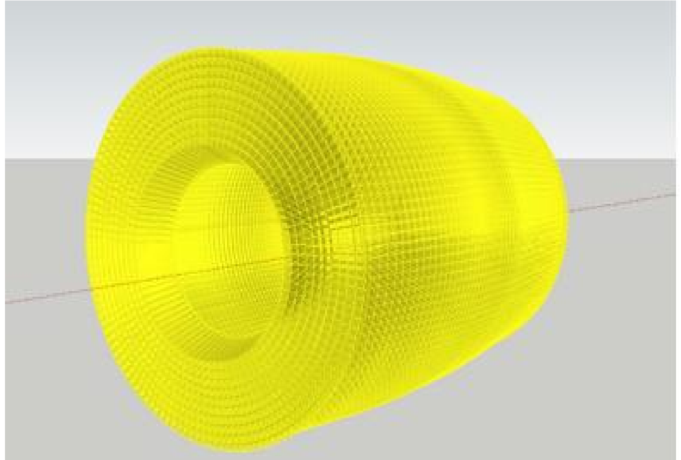
- The super τ-charm facility (STCF):
  - CMS： 2 ~ 7 GeV
  - Peak luminosity： $\geq 0.5 \times 10^{35} \mathrm{cm}^{-2} \mathrm{s}^{-1}$
  - Rich physics in the τ-charm energy region

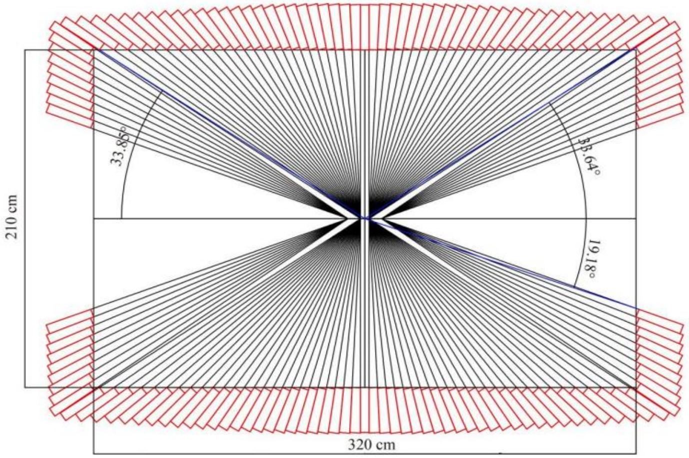
Data volume is approximately two orders of magnitude higher than BEPC II /BES III , this poses a challenge for MC production

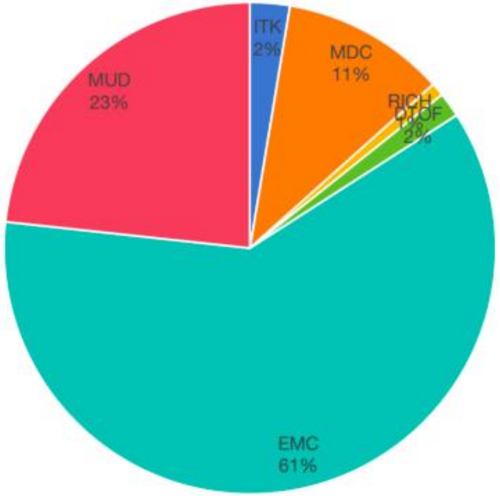


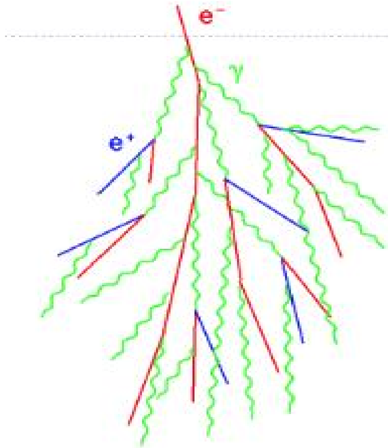
Introduction

- Electromagnetic calorimeter (ECAL):
  - 2D crystal calorimeter, shaped like a round barrel, divided into barrel part and two end cover areas
  - Barrel: 51 × 132 = 6732, Endcap: 3 × 85 + 102 + 136 = 969
- ECAL MC simulaton
  - Energy deposition in each crystal cell needs to be simulated
  - Simulation of electromagnetic shower requires significant amount of resources due to large number of secondary particles







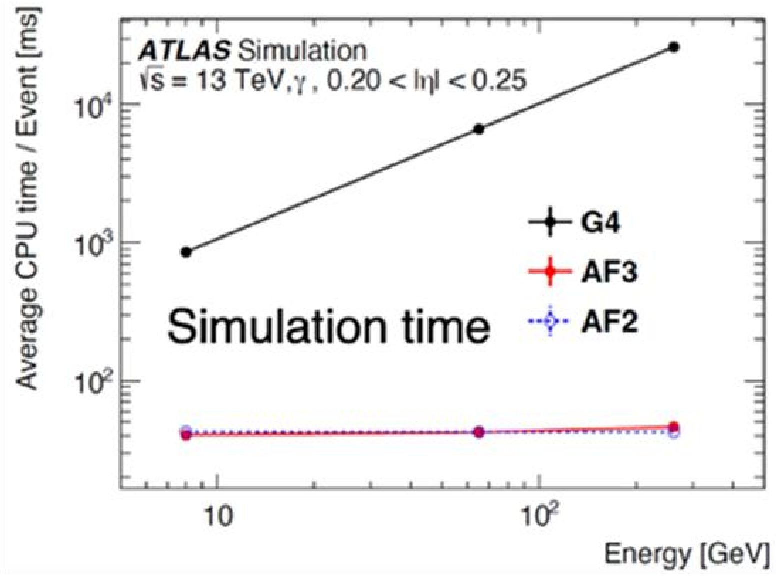


Geant4 simulation time consumption of each sub-detector

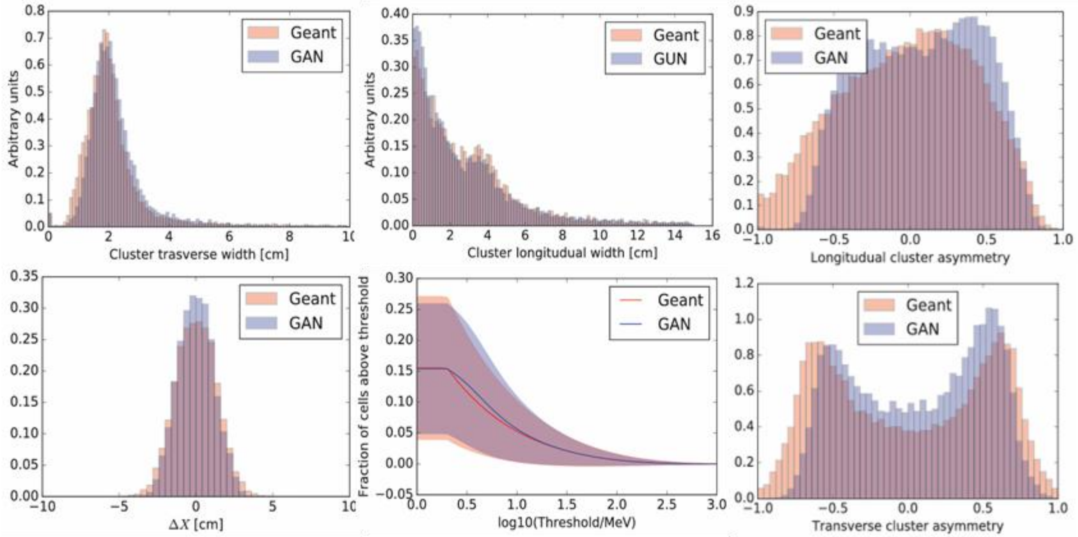
Parameterization or ML methods should be considered

Calorimeter Fast Simulation Based on Machine Learning

- Generative models are suitable for this kind of tasks (Generative Stochastic Network, Variational Auto-Econders, **Generative Adversarial Networks**, Diffusion models, ...)
  - Realistic generation of samples
  - Be able to generate complicated probability distributions using simple inputs
  - Work well with missing data
  - Extremely fast compared to full simulation



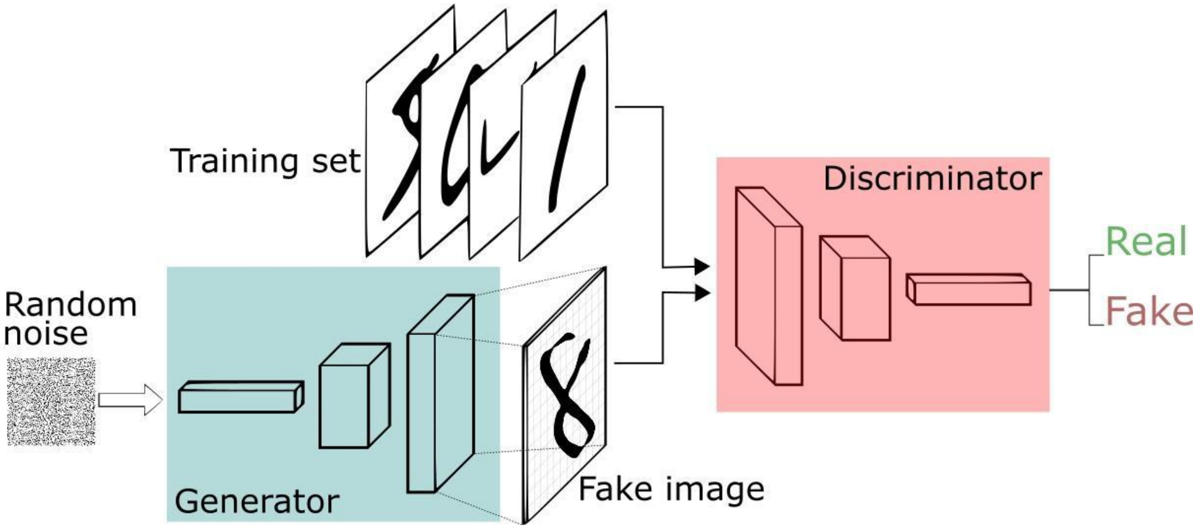
The AtlFast3 fast simulation for ATLAS based on GAN

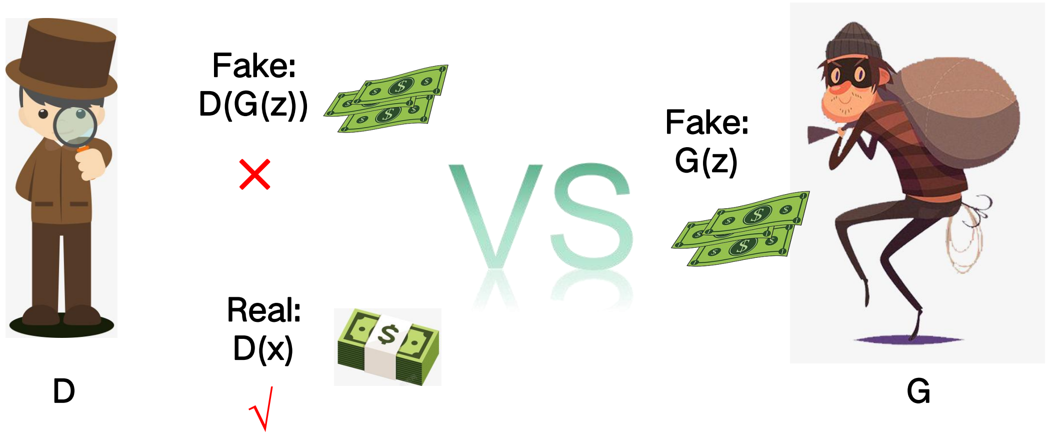


LHCb calorimeter fast simulation based on WGAN

Generative Adversarial Networks

- Adversarial: two ML models are trained simultaneously
  - Generator: captures the data distribution characteristics, and generate fake data
  - Discriminator: estimates the probability that a sample came from the training data rather than the generator
  - Training of the Generator aims to maximize the probability that Discriminator makes a mistake
  - Training of the Discriminator aims to minimize the mistake probability





$$\min_G \max_D V(D,G) = \mathbb{E}_{\boldsymbol{x}\sim p_{\text{data}}(\boldsymbol{x})}[\log D(\boldsymbol{x})] + \mathbb{E}_{\boldsymbol{z}\sim p_{\boldsymbol{z}}(\boldsymbol{z})}[\log(1 - D(G(\boldsymbol{z})))]$$

GAN is Quite Popular Internationally

Table 1 Calorimeter simulation with deep learning

No.	Model	Algorithm	Architecture	Condition	Output
1	LAGAN [21] (2017)	GAN	2D Locally connected	Particle type as discrete labels	$25 \times 25 =$ **625** cells
2	CALOGAN [22] (2017)	GAN	2D Locally connected	$E_P \sim U(1, 100)$ GeV	Layer1: 3×96 layer2: 12×12 layer3: $12 \times 6 =$ **504** cells
3	3DGAN initial prototype [17] (2018)	ACGAN	Conv3D	$E_P \sim U(2, 500)$ GeV	$25 \times 25 \times 25 =$ **15625** cells
4	ATLAS [23] (2018)	WGAN and VAE	Dense	$E_P \sim U(1, 260)$ GeV	Vector of **266** cells
5	LHCb [24] (2019)	WGAN	Conv2D	Five variables related to position and momentum	$30 \times 30 =$ **900** cells
6	HGCAL [25] (2019)	WGAN	Conv2D and locally connected	E_P and initial impact position (x, y)	Concatenation of 7 (12×15) layers = **1260** cells
7	3DGAN [19] (2019)	ACGAN	Conv3D	$E_P \sim U(2, 500)$ GeV and $\theta \sim U(60°, 120°)$	$51 \times 51 \times 25 =$ **65025** cells
8	DijetGAN [26] (2020)	WGAN	Conv2D		Vector of **7** jet variables
9	SARM [27] (2021)	Autoregressive models	Dense	data1: $P_T \sim [250, 300]$ GeV $I_{pixel} \sim [0, 276]$ data2: $P_T \sim [10, 20]$ GeV/c $I_{pixel} \sim [0, 172]$	Data1: $25 \times 25 =$ **625** cells data2: $32 \times 32 =$ **1024** cells
10	ILD [28] (2021)	GAN, WGAN and BIB-AE	Conv3D	$E_P \sim U(10, 100)$ GeV	$30 \times 30 \times 30 =$ **27000** cells
11	CaloFlow [29] (2021)	Normalizing flows	Dense	$E_P \sim U(1, 100)$ GeV	Layer1: 3×96 layer2: 12×12 layer3: $12 \times 6 =$ **504** cells

The sizes of simulated images in cells are emphasized in bold

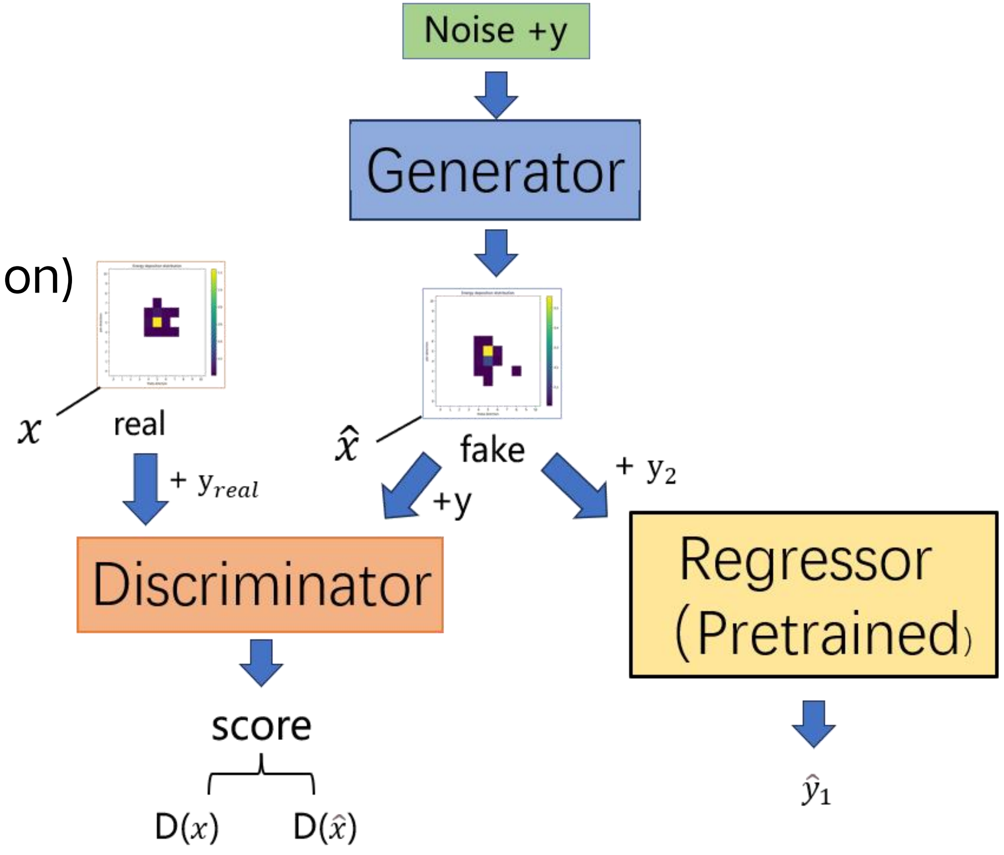
Eur. Phys. J. C (2022) 82:386

ECAL-GAN for STCF ECAL Fast Simulation

- The detector response (energy deposit) is converted into 2D images
  - 11*11 images, seeding crystal as the center pixel
  - All models (G, D and R) are basically deep CNNs
- ECAL-GAN should be made conditional
  - Extended to learn from a parameterized generator (able to simulate particles of certain momentum and position)
  - Some additional features (y) are made as inputs (4-momentum, incident angle, etc.)
- An additional Regressor is pre-trained to stabilize the training process
  - Predict the fake images y features, as an additional term of the Generator loss function

$$\min_G E_{\hat{x}\sim p(fake)} \log\big(1 - D(\hat{x})\big) + \boxed{\|y_1 - \hat{y}_1\|_1}$$

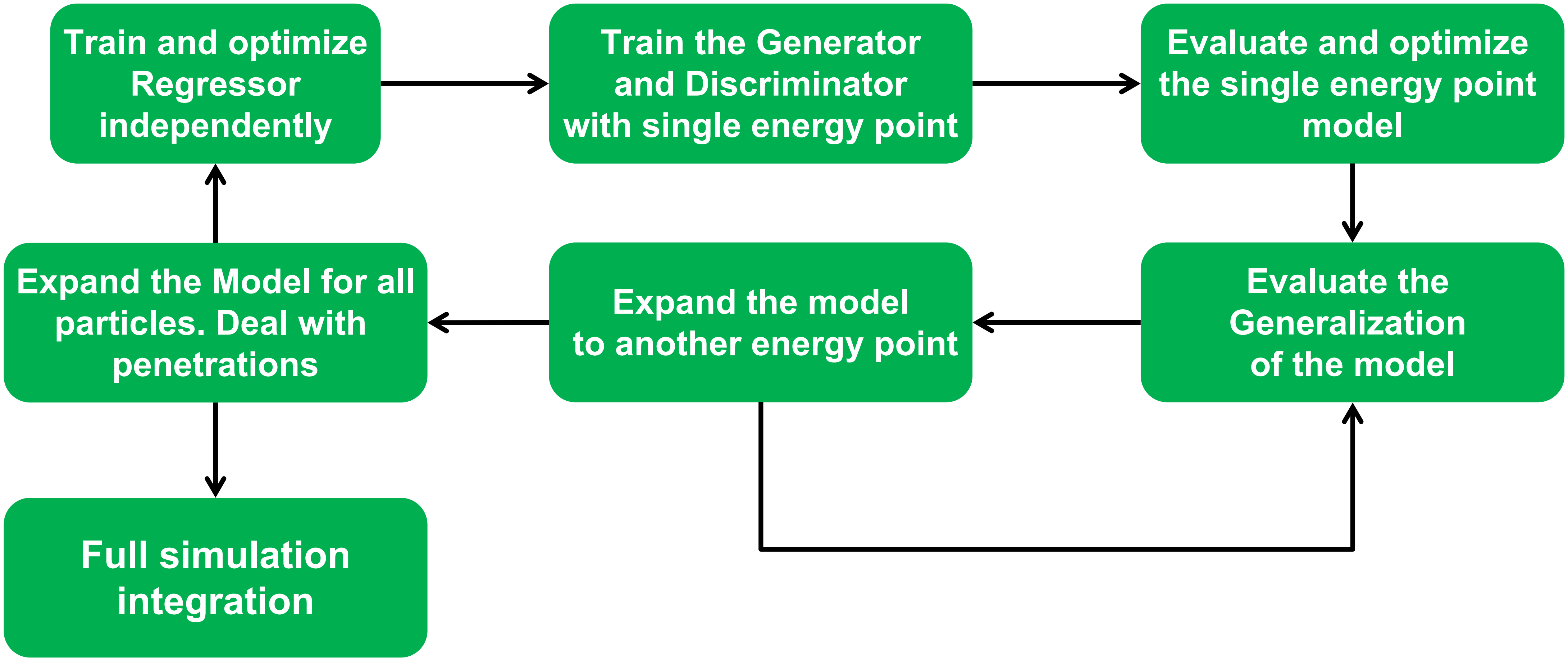
$$\max_D E_{x\sim p(data)} \log\big(D(x)\big) + E_{\hat{x}\sim p(fake)} \log\big(1 - D(\hat{x})\big)$$



STCF ECAL-GAN model

Training Strategy of ECAL-GAN

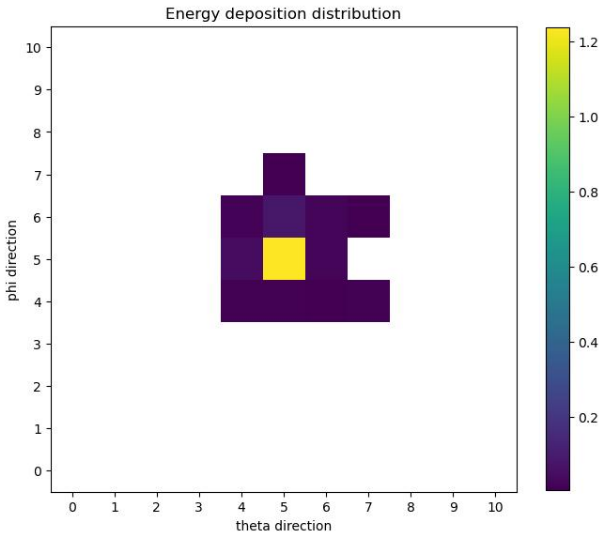
- The entire R&D of ECAL-GAN can be summarize as the following



Dataset

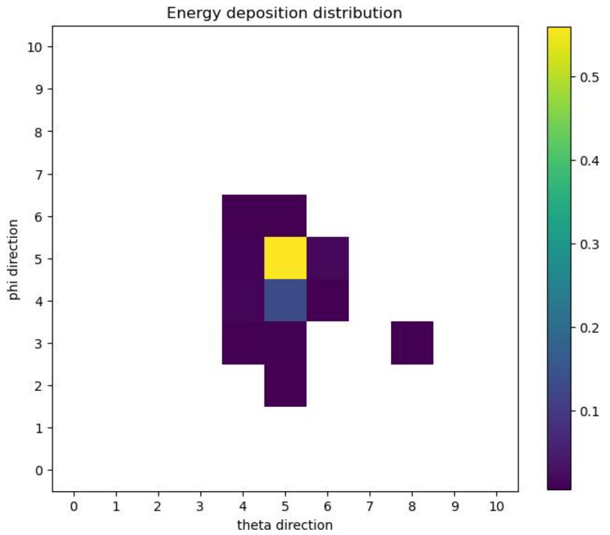
❖ Dataset for training the GAN model (including G, C and R)

- Data sample is generated using OSCAR 2.5.0
- Single photons generated using particle gun, with a unified distribution
  P ∈ (0， 2.0) Gev/c, $\theta \in (20°,160°)$, $\varphi \in (0°,360°)$
- Calorimeter response (energy depositions) are converted to images of 11x11 size.
- Each pixel is one crystal. The center pixel is the crystal where the particles hit ECAL



Mom=1.64318GeV, $\Delta\varphi^{Mom}$ = -1.8766°, $\Delta\theta^{Mom}$ = 0.266638°,

ΔZ^{Pos} = -0.632551mm, $\Delta\varphi^{Pos}$ = -0.52491°, ΔZ = 352.501mm



Mom= 0.883455GeV, $\Delta\varphi^{Mom}$ = -0.0952646°, $\Delta\theta^{Mom}$ = 0.268006°,

ΔZ^{Pos} = -1.60522mm, $\Delta\varphi^{Pos}$ = 1.26624°, ΔZ = 45.0363mm

Design and Training of the Regressor

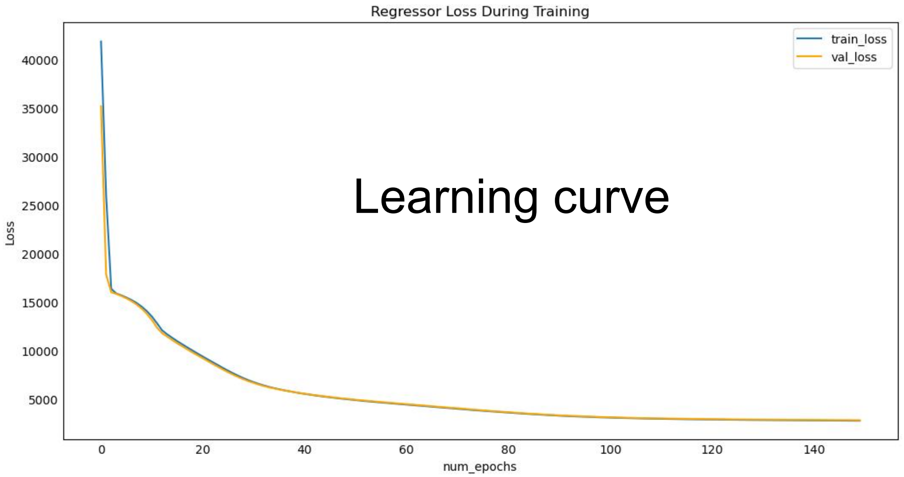
❖ Design and training of the regressor

Input(11*11) + y_2 (ΔZ^{Pos}, $\Delta\varphi^{Pos}$, ΔZ) → Regressor → Output: y_1 (Mom, $\Delta\varphi^{Mom}$, $\Delta\theta^{Mom}$)

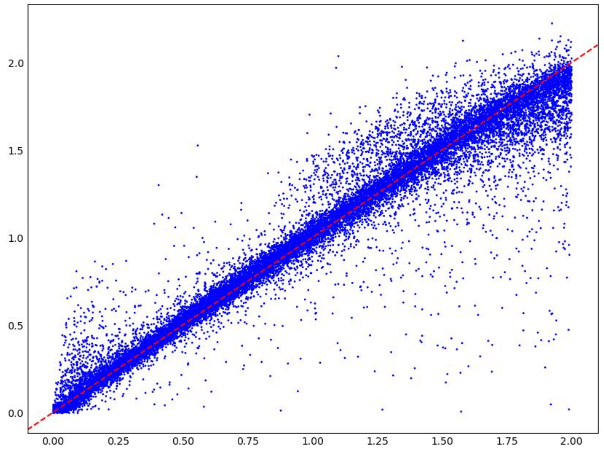
Regressor:
- Normalize
- conv2d(1,64,4,1,1)+BatchNorm2d+LeakyReLU
- conv2d(64,128,2,2,1)+BatchNorm2d+LeakyReLU
- Flatten+Linear(100)+ Linear(10) Linear(3)

Input:
- Image
- Generated additional features (hit position)

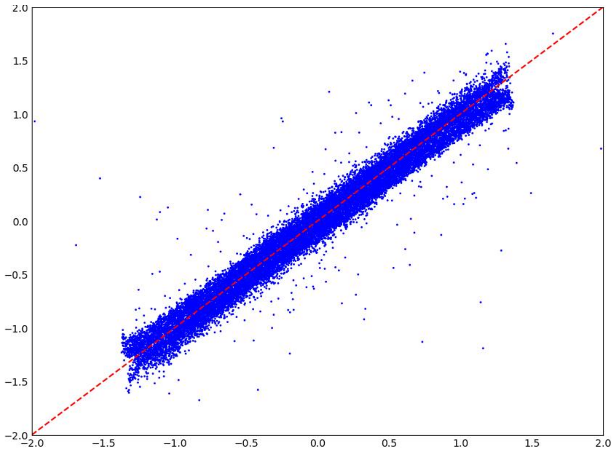
Output:
- Momentum and indicent angle



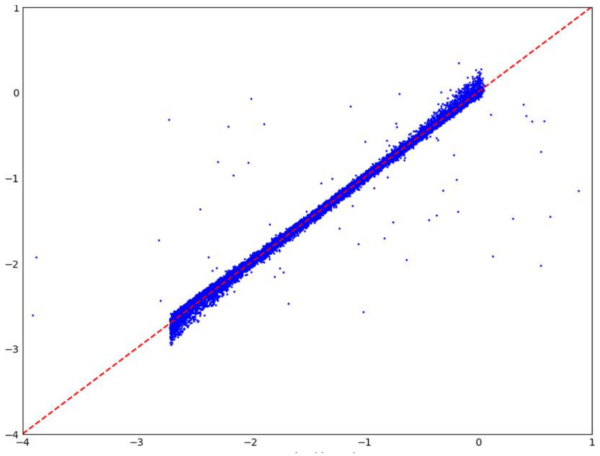
Predicted Values



True Momentum



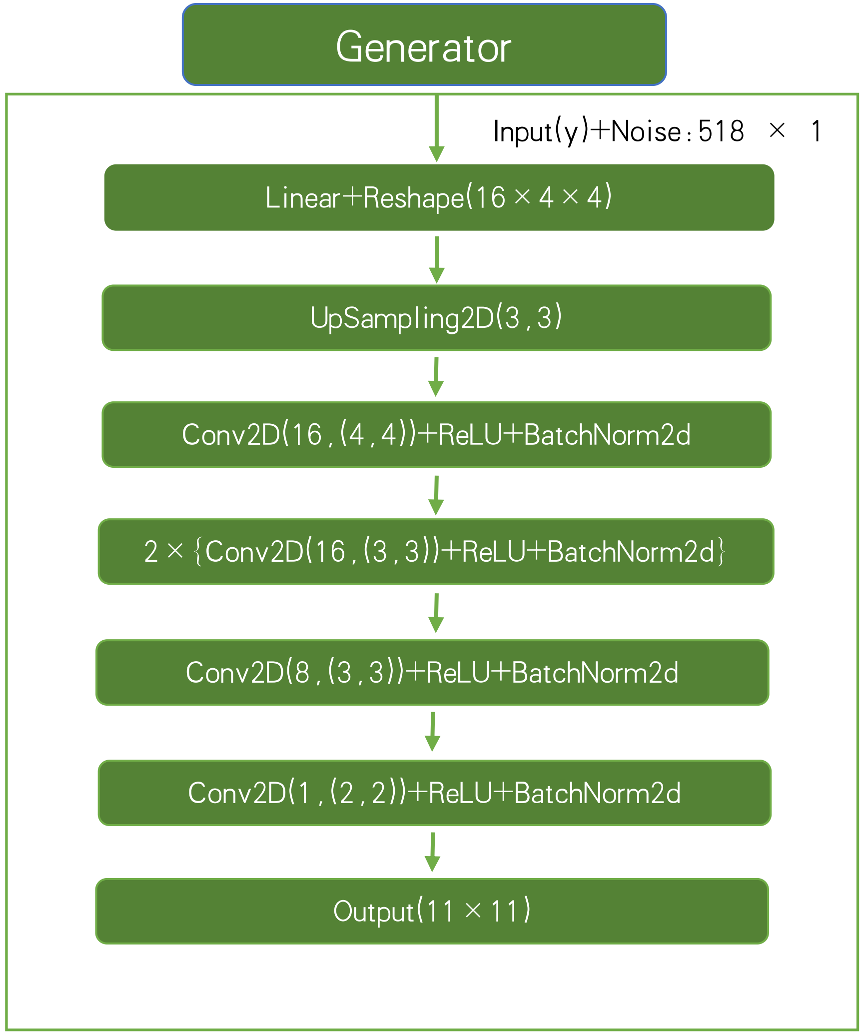
True delta theta



True delta phi

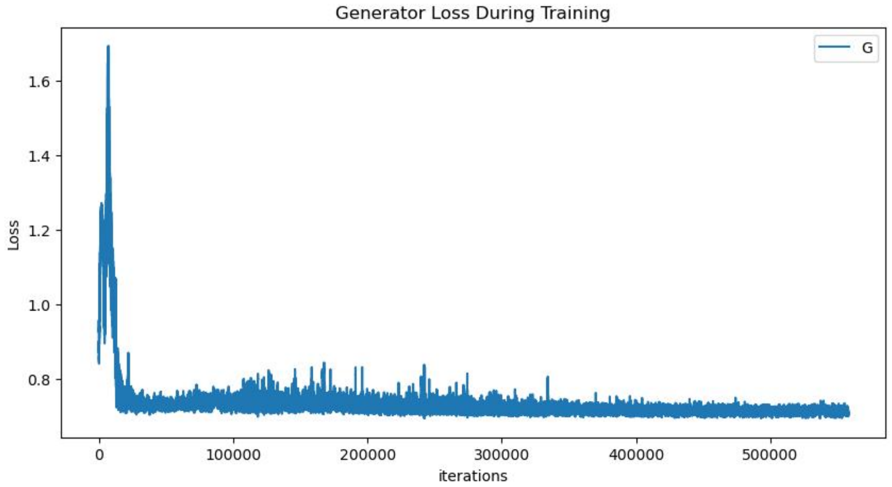
Design of the Generator and Discriminator

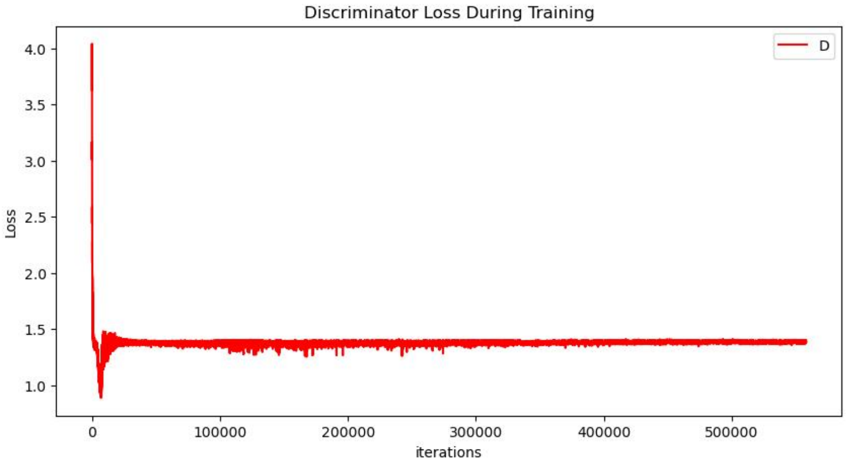
- Design of the generator and discriminator



Training of the Generator and Discriminator

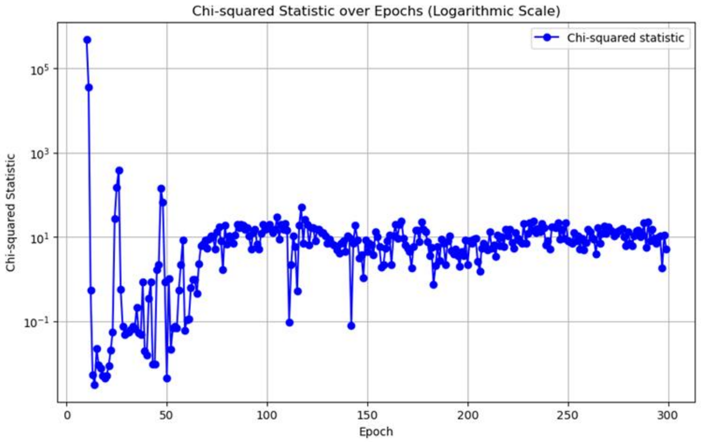
- The Generator and discriminator are trained simultaneously. We carefully tuned the loss function to make sure they converge roughly at the same speed





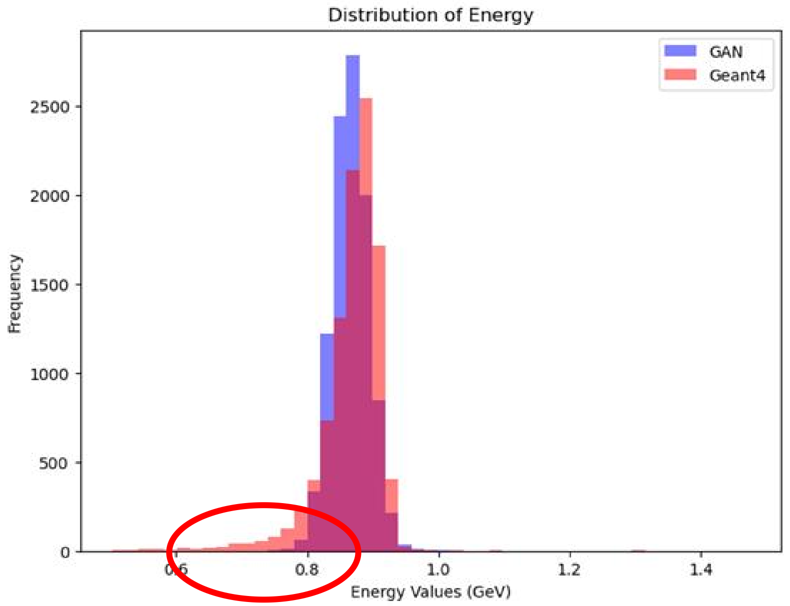
- The feature distribution difference between GAN and geant4 is used to monitor the training process

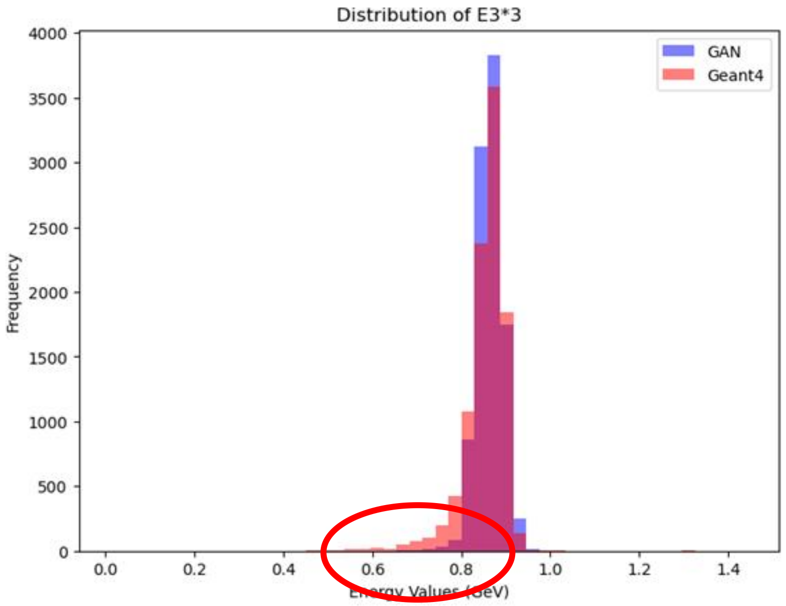
$$\chi^2 = \Sigma \frac{(O_i - E_i)^2}{E_i}$$

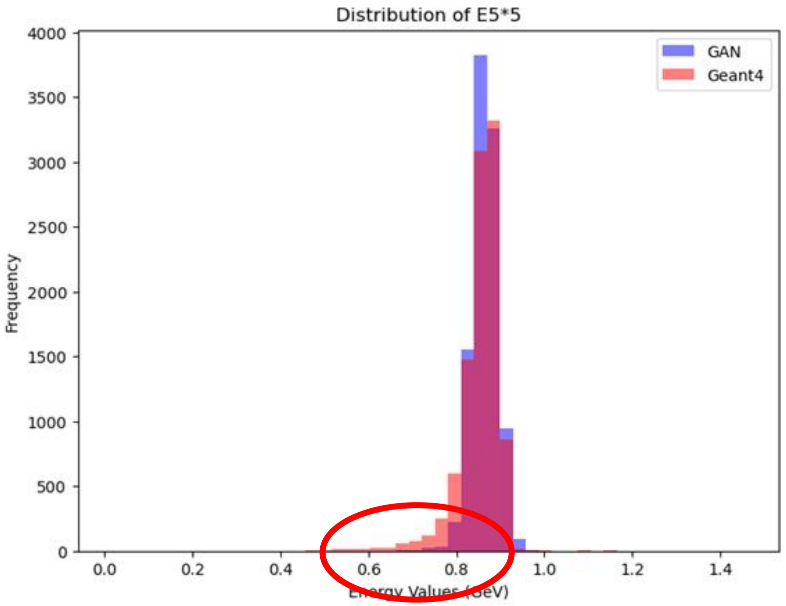


Preliminary Results I

- Step 1: train G and D with particles of single energy point
  - Previous experiences show that balancing generator and discriminator using the entire phase space data is diffcult
  - A 55W training set of 1GeV photons are used as the first step.
  - The model does not perform well for the part with lower energy deposition



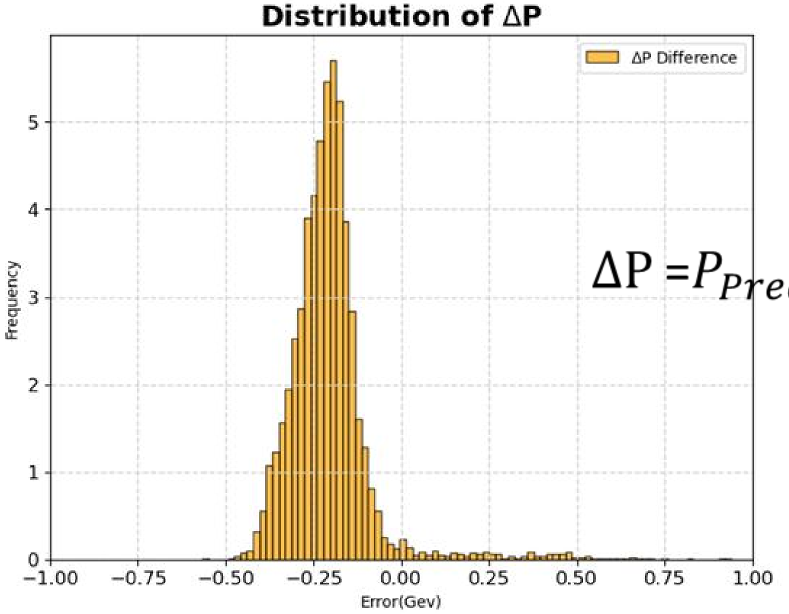




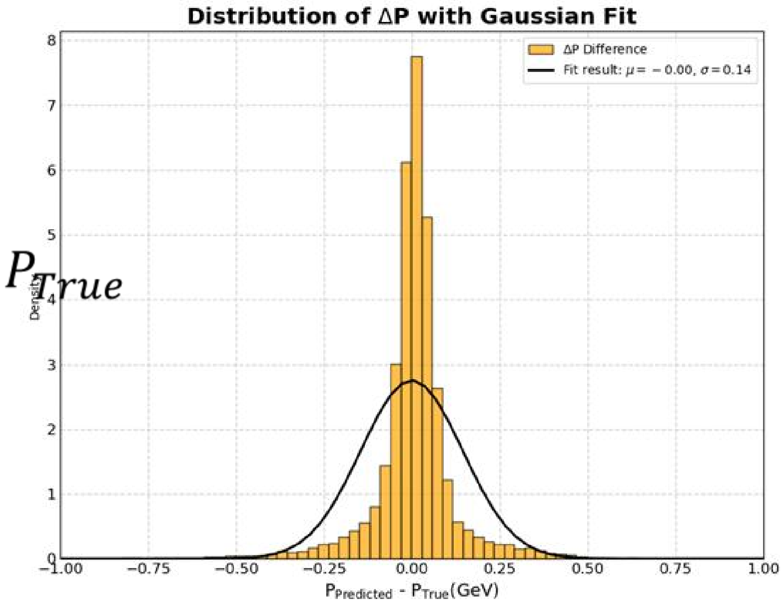
Improve the Performance

- Evaluate the performance of the Regressor when the deposited energy is low
  - The regressors well in the event of low deposition energy did not predict the momentum well

$$\Delta P = P_{Predicted} - P_{True}$$



Low-energy region

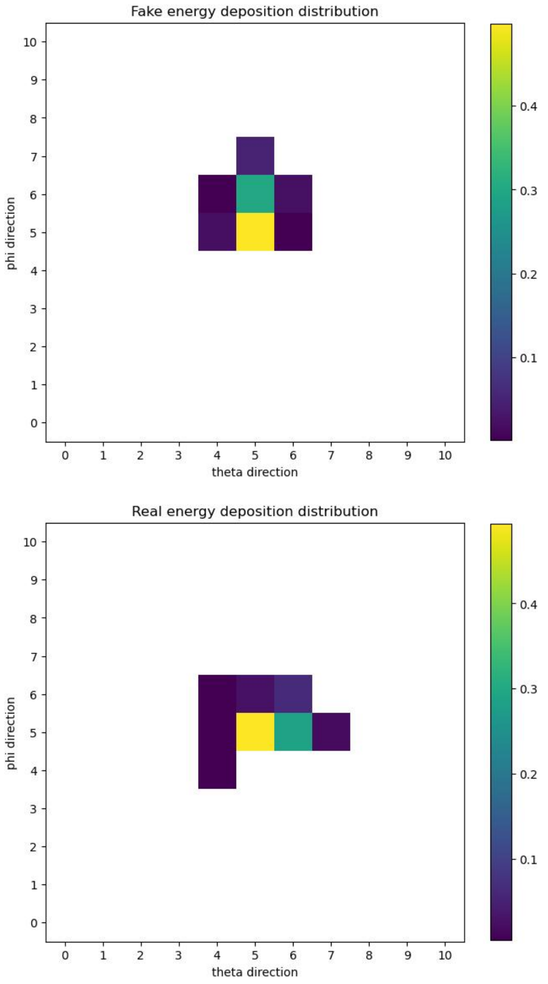


Full dataset

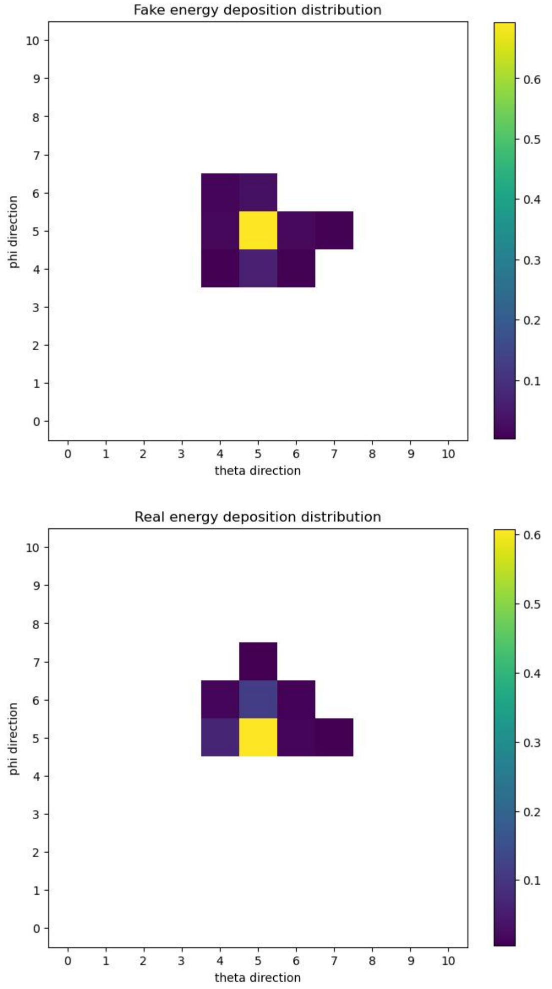
- Increasing events in the low energy region also improves the performance, by forcing the GAN model not to overlook them

Preliminary Results after Mitigation

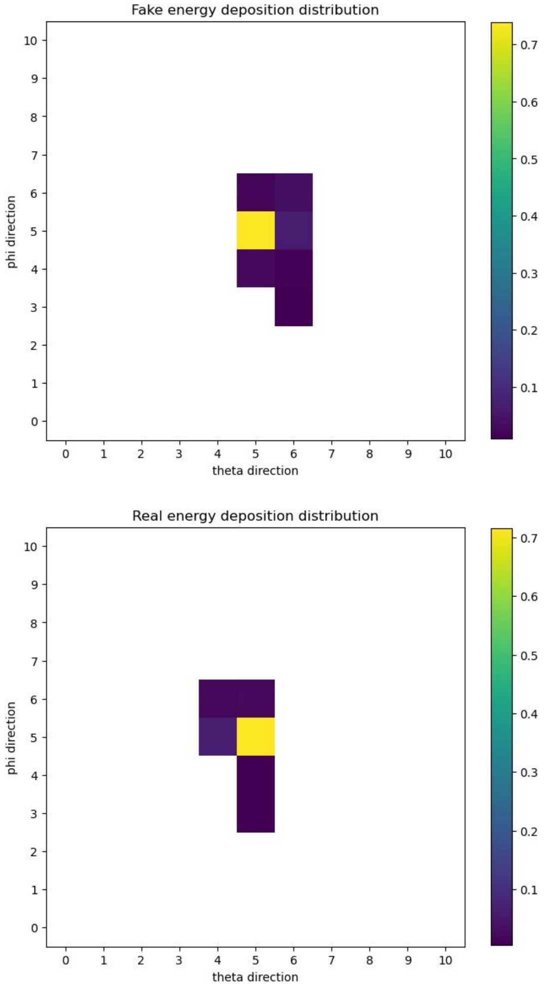
Mom=1.0GeV, $\Delta\varphi^{Mom}$ = -1.8766°, $\Delta\theta^{Mom}$ = 0.266638°,
ΔZ^{Pos} = -0.632551mm, $\Delta\varphi^{Pos}$ = -0.52491°, ΔZ = 352.501mm:



Mom=1.0GeV, $\Delta\varphi^{Mom}$ = -0.0221°, $\Delta\theta^{Mom}$ = 0.0111°,
ΔZ^{Pos} = -0.3303mm, $\Delta\varphi^{Pos}$ = -0.4320°, ΔZ = 0.6226mm:



Mom=1.0GeV, $\Delta\varphi^{Mom}$ = -0.0137°, $\Delta\theta^{Mom}$ = 0.0052°,
ΔZ^{Pos} = -0.3808mm, $\Delta\varphi^{Pos}$ = 0.0746°, ΔZ = -0.2961mm:

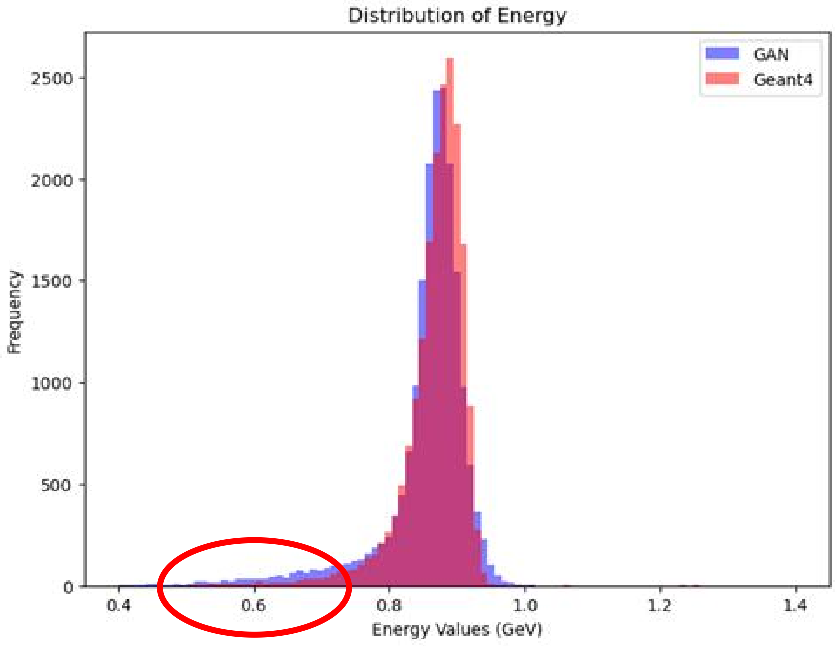


GAN

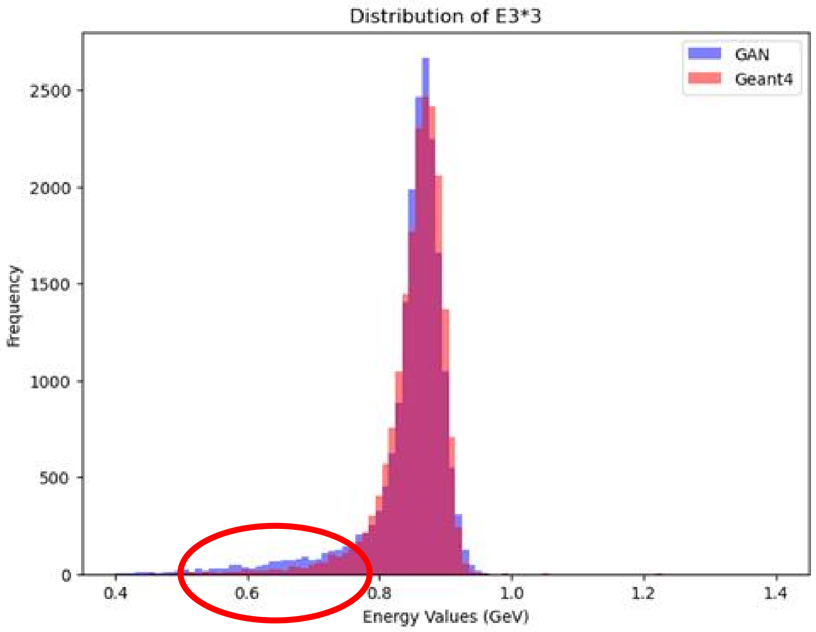
Geant4

Preliminary Results after Mitigation

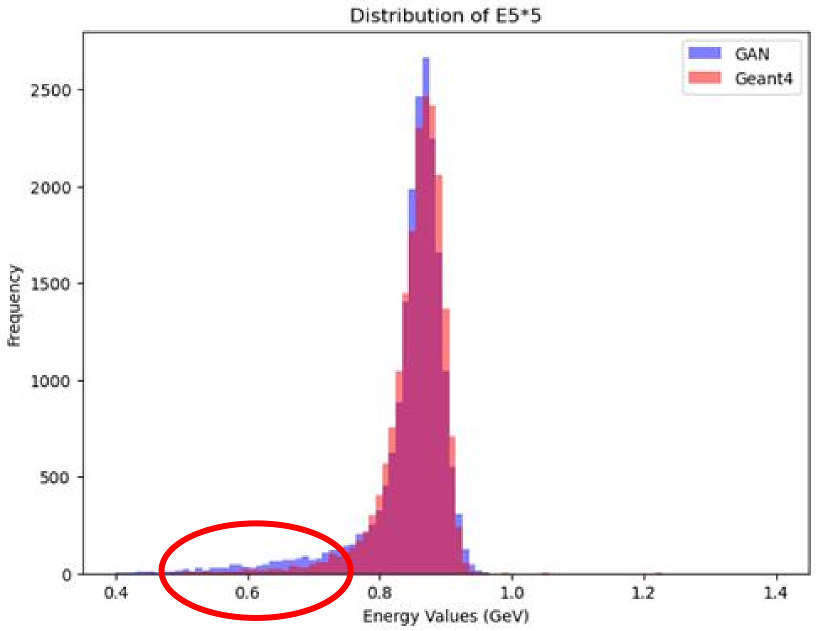
- Comparison of GAN model results and Geant4 results



E_{total}



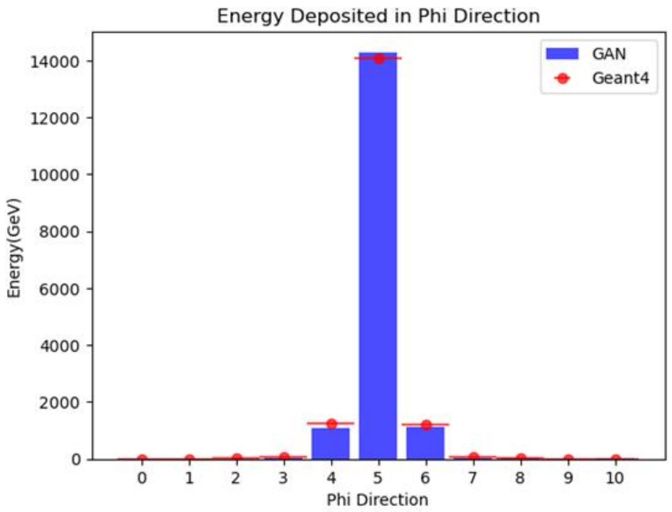
E_{3*3}

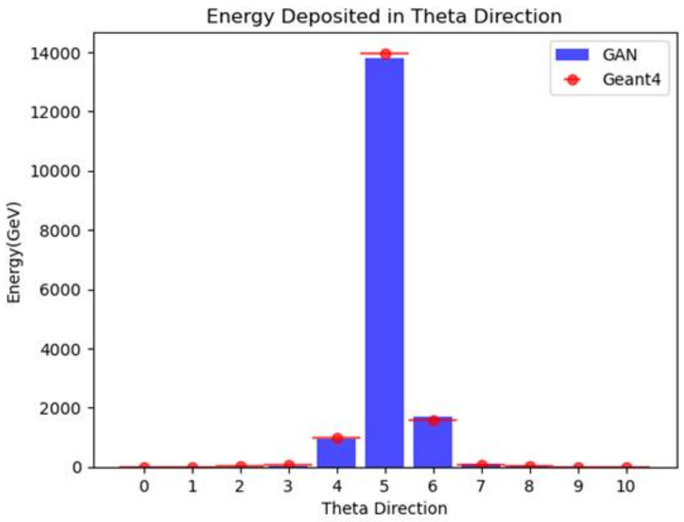


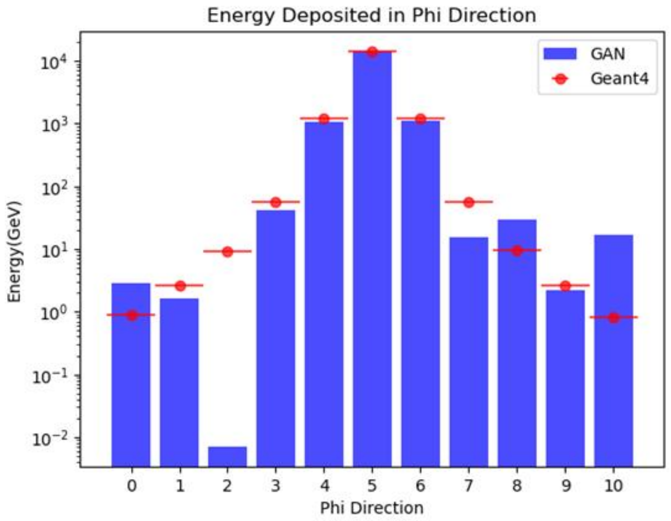
E_{5*5}

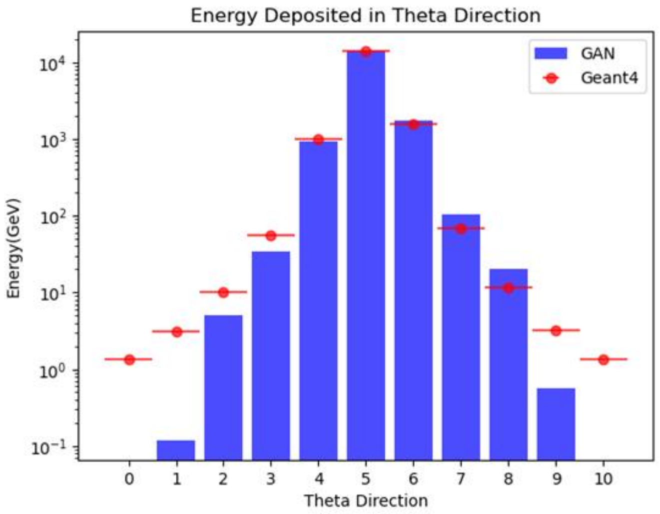
Preliminary Results

- Energy deposition distribution in the ϕ and Z directions









Summary

- To deal with the intensive MC data production requirement, a GAN-based fast simulation method is proposed
  - Based on experience (mostly) from LHC
- A fully functional fast simulation requires sophisticated R&D work
  - A preliminary ECAL-GAN model is developed, now capable of simulating single energy photons
  - A lot of work is ahead
    - Further evaluation model performance using reconstruction results
    - Expanding model for more energy points and particles
    - Dealing with punch through particles
    - Integration with Geant4
    - ...

Back up

02. Methodology

◆ The model:

- y ($y_1 + y_2$):
- y_1:
  - Momentum: the momentum of the particle.
  - $\Delta\phi^{Mom}$: the ϕ difference between the momentum of incoming particle and the direction of the crystal.
  - $\Delta\theta^{Mom}$: the θ difference between the momentum of incoming particle and the direction of the crystal.
- y_2:
  - ΔZ^{Pos}: the Z difference between the hit point of incoming particle and the z of front center of the crystal.
  - $\Delta\phi^{Pos}$: the ϕ difference between the hit point of incoming particle and the ϕ of front center of the crystal.
  - Z: the Z value of hit point.
- Loss: $\min_G E_{\hat{x}\sim p(fake)} \log\left(1 - D(\hat{x})\right) + \|y_1 - \hat{y}_1\|_1$

$$\max_D E_{x\sim p(data)} \log\left(D(x)\right) + E_{\hat{x}\sim p(fake)} \log\left(1 - D(\hat{x})\right)$$