STCF Electromagnetic Calorimeter Fast Simulation based on GAN

Xiao Yang[1], Tong Liu[2], Ke Li[2], Xiaoshuai Qin[1], Teng Li[1], Xingtao Huang[1]

[1]SDU

[2]IHEP

Outline

- Introduction
- Wasserstein GAN (WGAN)
- GAN for STCF ECAL Simulation
- Summary and outlook

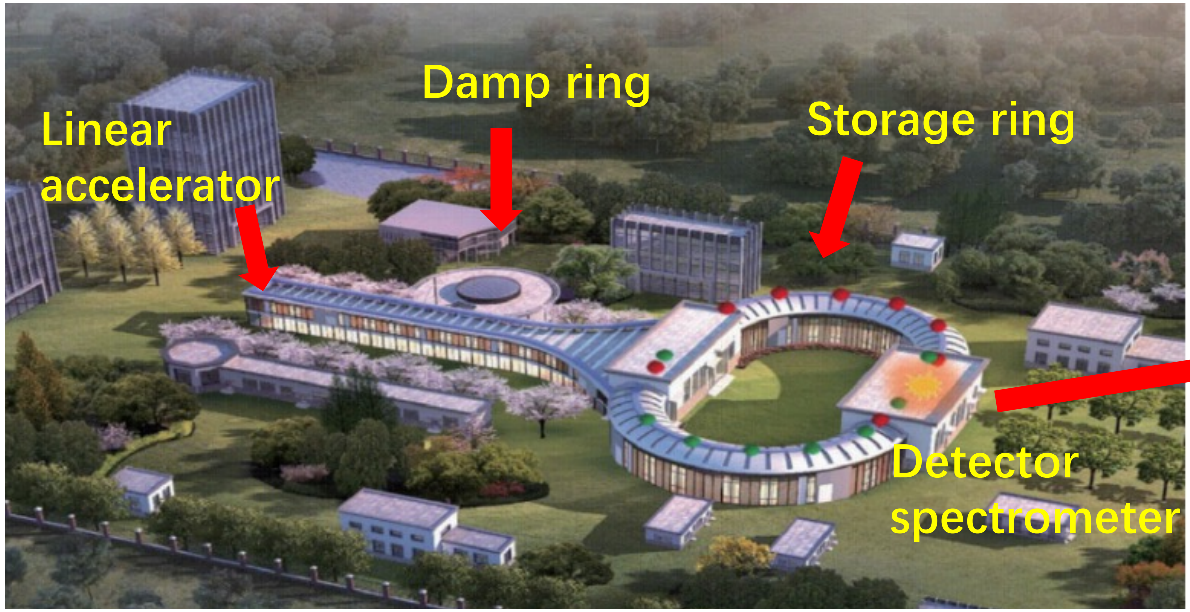
Introduction

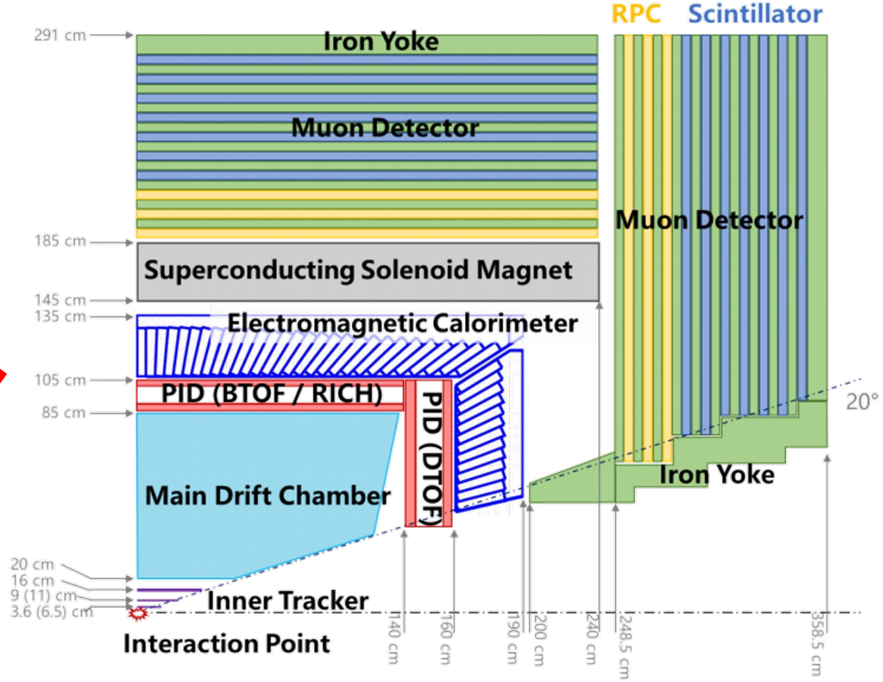
The super τ-charm facility (STCF):

- Parameters of STCF:
  - CMS： 2 ~ 7 GeV
  - Peak luminosity : $\geq 0.5\times10^{35}\mathrm{cm}^{-2}s^{-1}$

Approximately two orders of magnitude higher than BEPCII/BESIII, this poses a challenge for generating MC samples using offline software.

- Physics Objectives:
  - Hadron structure and QCD
  - CP Violation and Flavor Physics
  - New physics searching

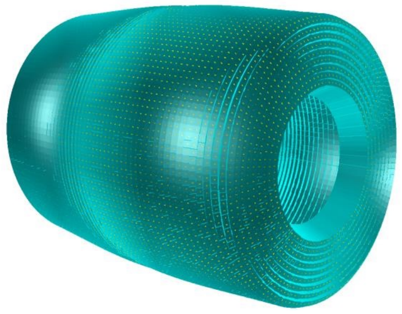




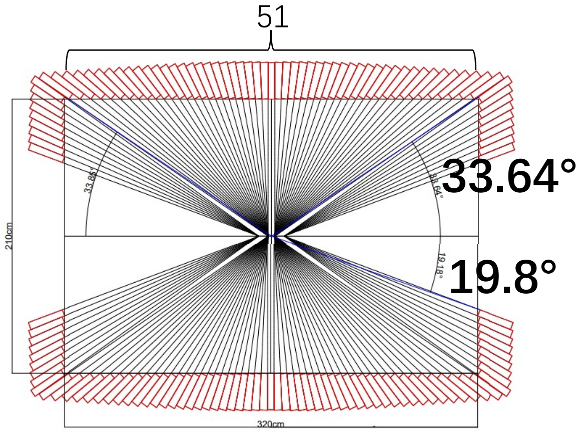
Introduction

Electromagnetic calorimeter (ECAL):

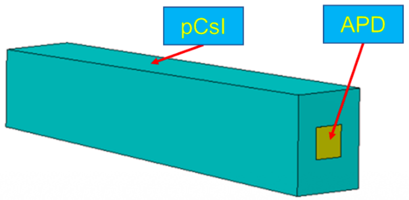
- Determine the energy, position and time information of photons and electrons
- Shaped like a round barrel, it is divided into a barrel part and two end cover areas
    - Barrel: 51 × 132 = 6732
    - Endcap: 3 × (85 + 102 + 136) × 2 = 1938
- Sensitive Unit: Pure CsI (pCsI) crystal
- Avalanche photodiode (APD)
- Convert the energy deposition in the ECAL into a two-dimensional pixel image of size 71 × 136

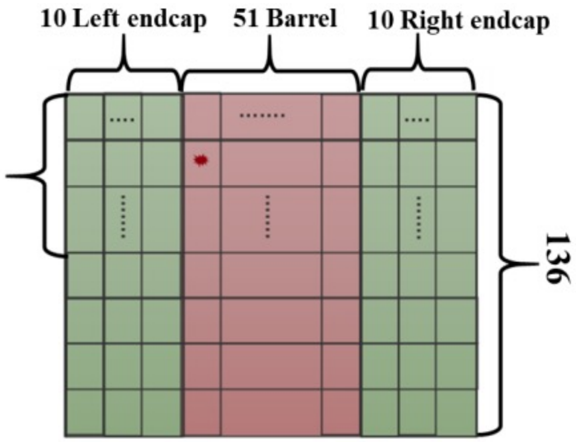


ECAL Geometric Model



ECAL layout design





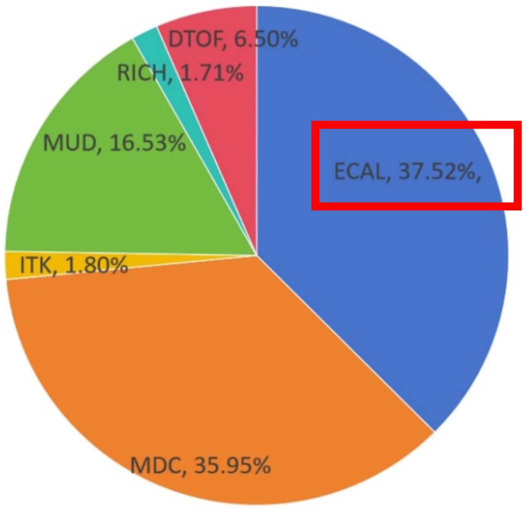
136*71 Energy deposition map

Introduction

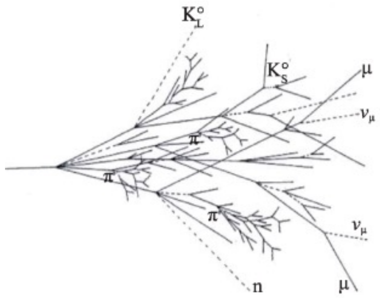
ECAL Simulation

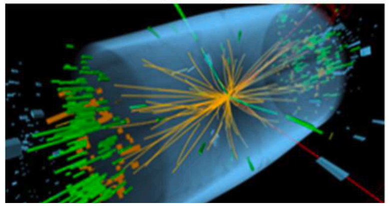
- Luminosity Increase
    - Computational cost becomes a bottleneck
    - ECAL simulation accounts for a large proportion of the computing resources
- Geant4 simulation
    - Electromagnetic showers involving a cascading process
    - Produce a large number of secondary particles
- Fast Simulation based on ML
    - Directly obtain the energy deposition from the incident particle information
    - Applying ML methods to ECAL fast simulation can significantly reduce

computational cost.



Full simulation time of each detector in Geant4





Introduction

Generative Models

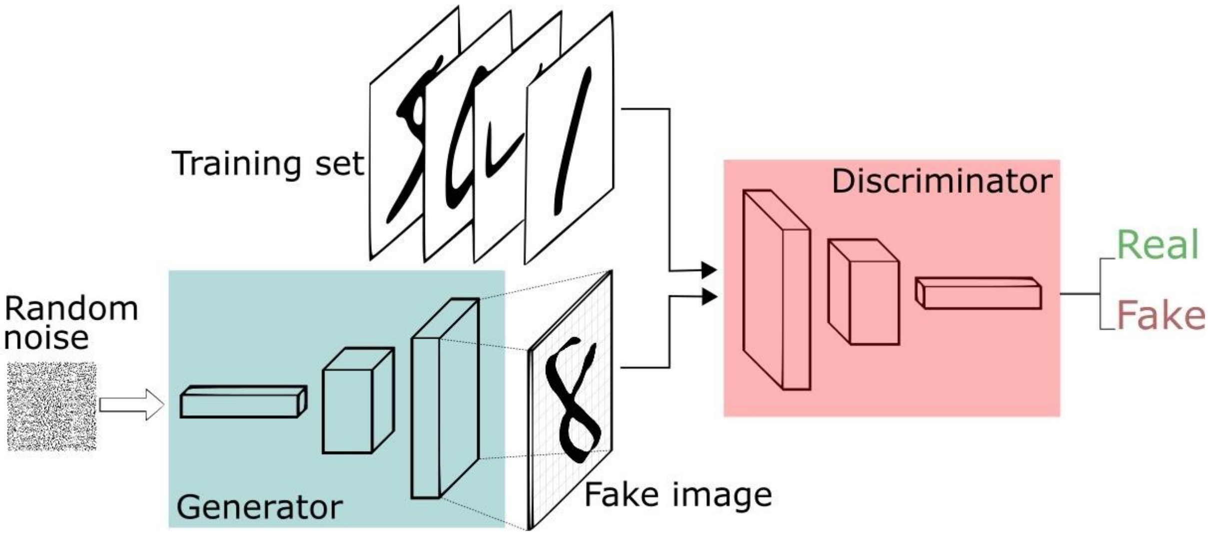
arXiv:2103.04922

Year	Model	Sample Speed	Strength	Weakness
2013	VAE	Fast	Stable training	Blurry outputs
2014	**GAN**	**Extremely Fast**	**Sharp images**	**Training instability**
2015	Normalizing Flow	Fast	Exact likelihood	Architectural constraints
2016	Autoregressive Models	Slow	Strong density modeling	Very slow sampling
2020	Diffusion	Slow	High sample quality	Expensive iterative sampling

- Using GANs can maximize the generation speed in STCF ECAL simulation while maintaining high accuracy.

Introduction

- Generative Adversarial Networks (GAN), a popular machine learning model arXiv:1406.2661
  - A generator to produce fake samples, tries to confuse discriminator
  - A discriminator tries to distinguish the real and fake data
  - Finally, the generator and discriminator reach a dynamic equilibrium.
  - A loss function measures the difference between the model predictions and the true data.
  - Loss for GAN: $\min_G \max_D V(D,G) = \mathbb{E}_{x\sim p_{data}(x)}[\log D(x)] + \mathbb{E}_{z\sim p_z(z)}[\log(1 - D(G(z)))]$



Introduction with WGAN

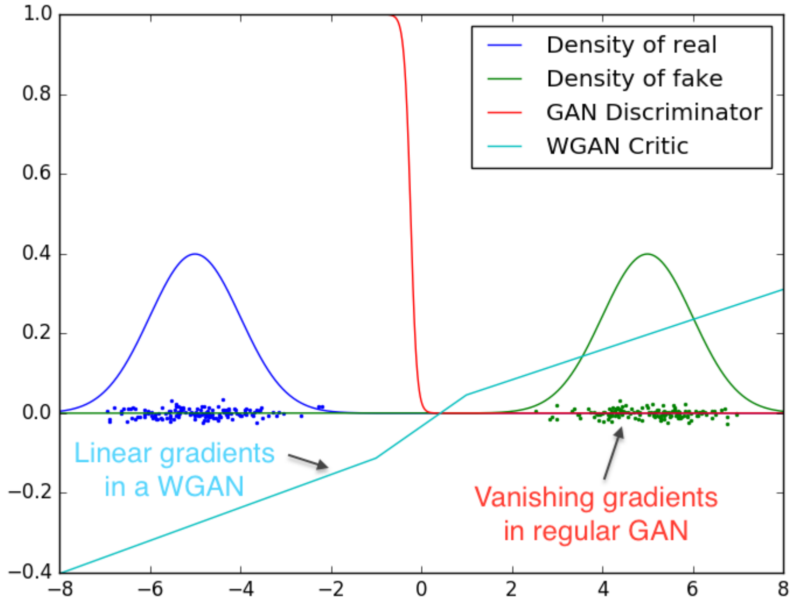
- Wasserstein GAN, an improved GAN model

arXiv:1701.07875

  - Binary crossentropy in Dloss may cause gradient vanishing in GAN
    - Especially if two distributions are non-overlapping
  - Redesign Dloss to aovid gradient vanishing
    - Replace JS divergence with the Wasserstein distance
    - Lipschitz Constraint: via weight clipping or gradient penalty (WGAN-GP)
  - Loss for WGAN

$$L_D = \mathbb{E}_{\tilde{x}\sim P_g}[D(\tilde{x})] - \mathbb{E}_{x\sim P_r}[D(x)] + \underbrace{\lambda \mathbb{E}_{\hat{x}\sim P_{\hat{x}}}[(\|\nabla_{\hat{x}} D(\hat{x})\|_2 - 1)^2]}_{\text{GP term}}$$

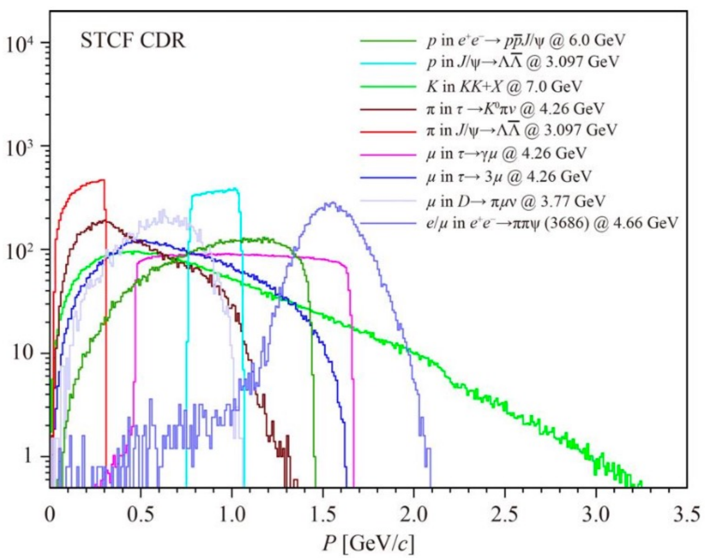
$$L_G = -\mathbb{E}_{\tilde{x}\sim P_g}[D(\tilde{x})]$$

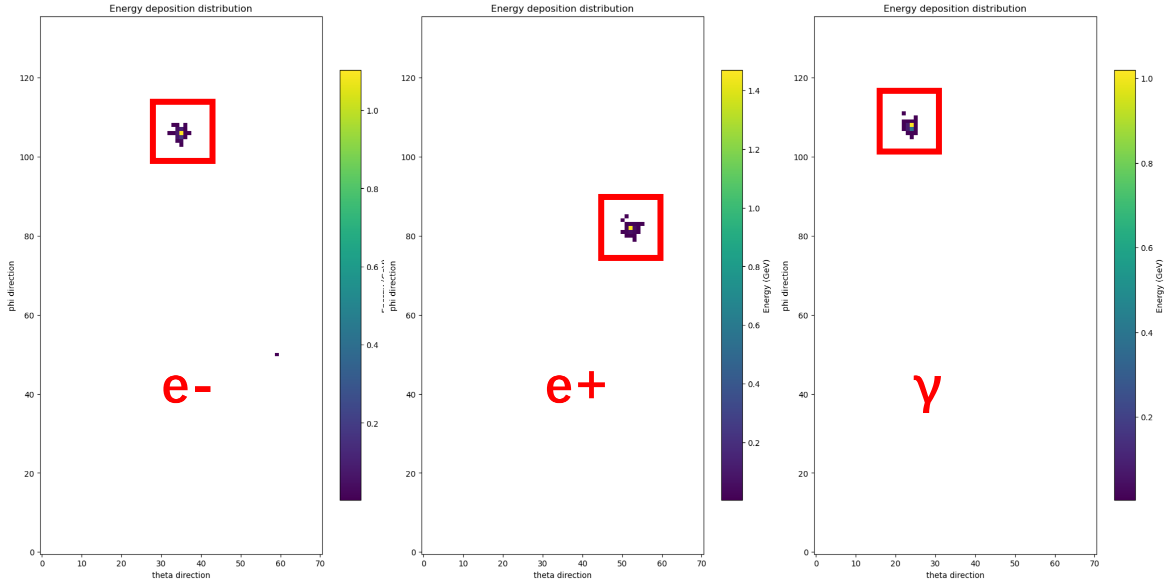


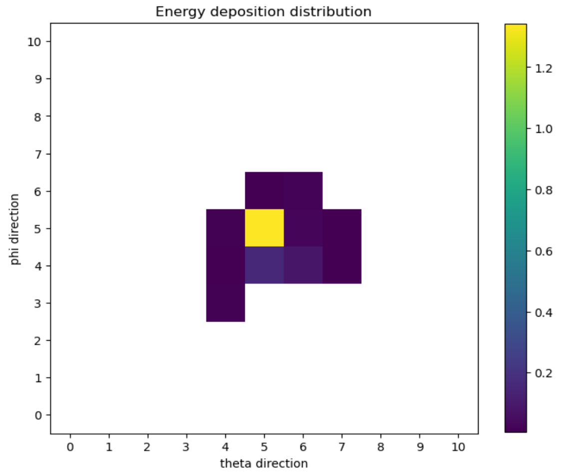
Training Samples

- Simulate about 1M events for e-/e+/γ as training set
  - With 0 < Mom < 2GeV, 20° < theta < 160°, 0 < phi < 360°
  - Single-momentum samples are used to check the performance of GAN.
- Smaller samples as test set and validation set
- Remove random trigger
- 11×11 region contains nearly the entire energy deposition as the input
- Events close to the detector boundary are removed to ensure complete 11×11 energy deposition maps

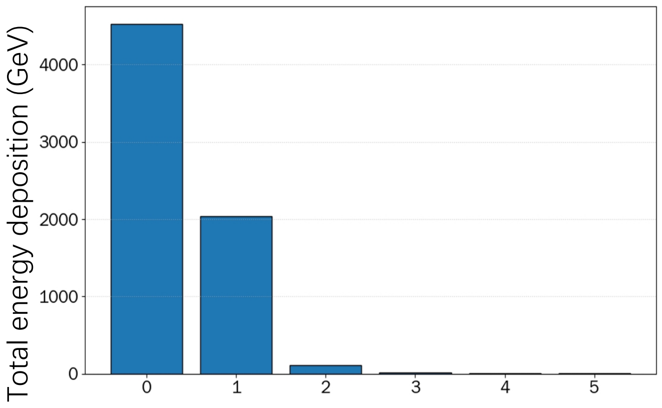
charged particle momentum







11×11 region Energy deposition map



The number of turns of the crystal from the center to the outside(0 to 5)

Model of GAN

Condition:

- Basic: Momentum-truth, inject direction, inject position, position_z
- GAN can generate deposition map conditioned on physical parameters such as particle momentum or position.

Training:

- G +D + Regressor (pretrained)

GNewLoss:

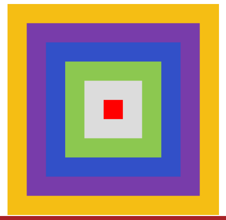
- Calculate deposition in regions:

  E1x1, … ,E11x11

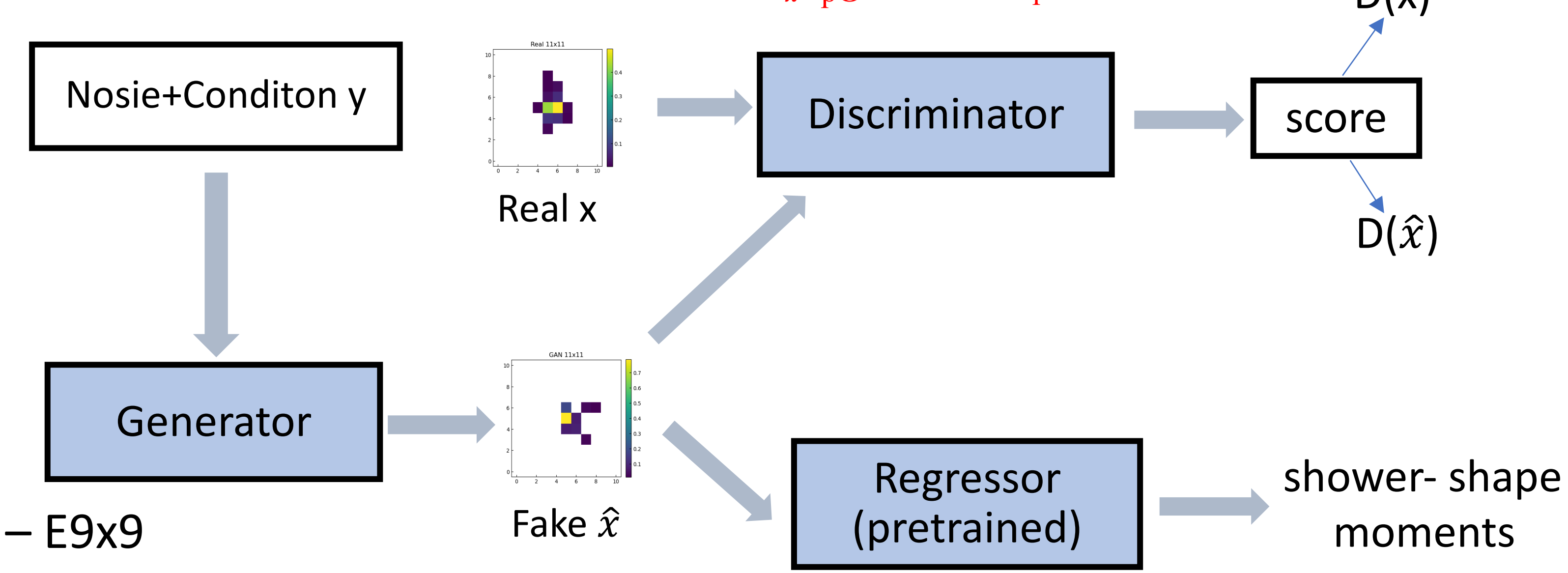
- Calculate deposition in core regions:

  E1x1, E3x3, E11x11

- Ring loss: E1x1, E3x3 - E1x1, ... , E11x11 – E9x9



$\mathrm{LossD} = \min E_{\hat{x}\sim pG}[D(\hat{x})] - E_{x\sim p}[D(x)] + GP$



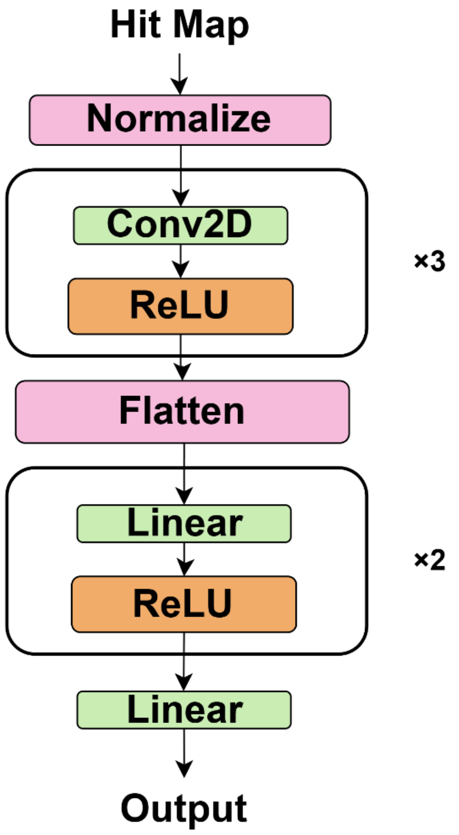
$\mathrm{LossG} = \min E_{\hat{x}\sim pG}[-D(\hat{x})] + \mathrm{loss}_{reg}(obs\ ) + \mathrm{GNewLoss}$

GAN for STCF ECAL Simulation - Regressor

Training strategy: G + D + regressor (pretrained)

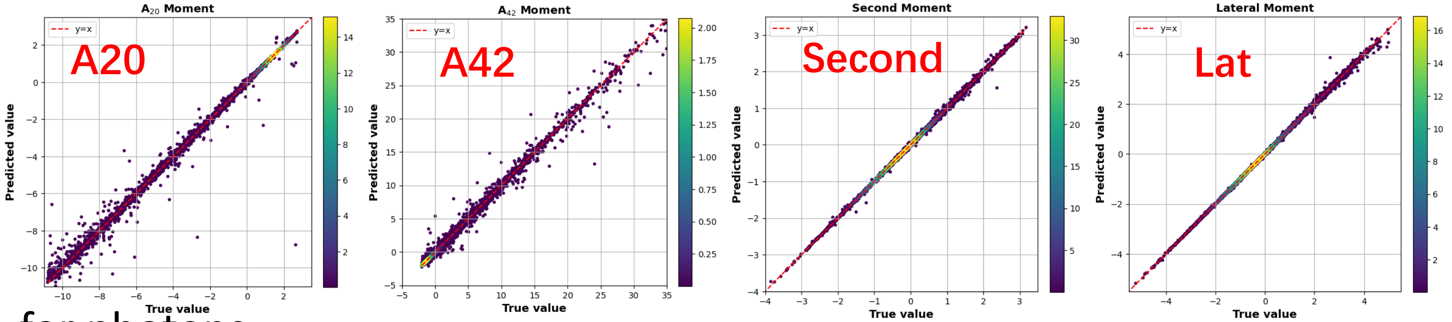
Use observables to constrain:

- Use regressor to reconstruct complex observables: A20 moment, A42 moment, Second moment, lateral moment ;

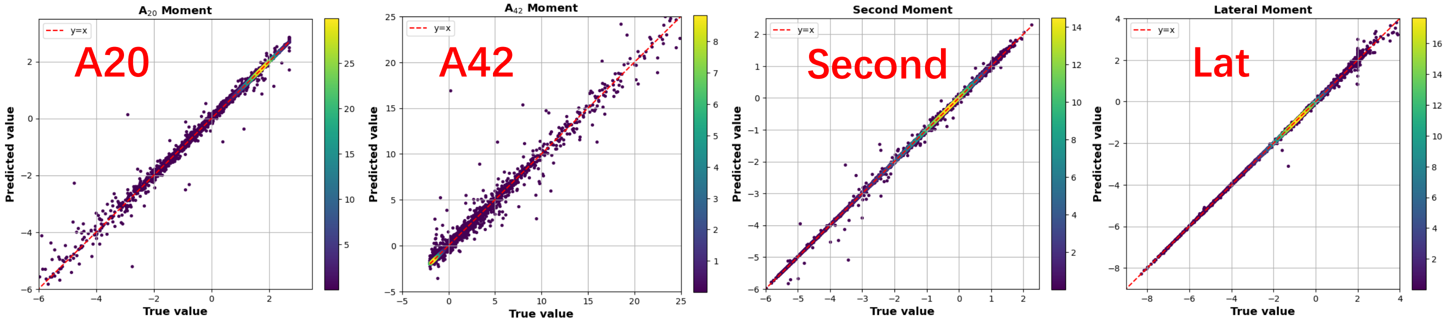


Regressor performance

for electrons:

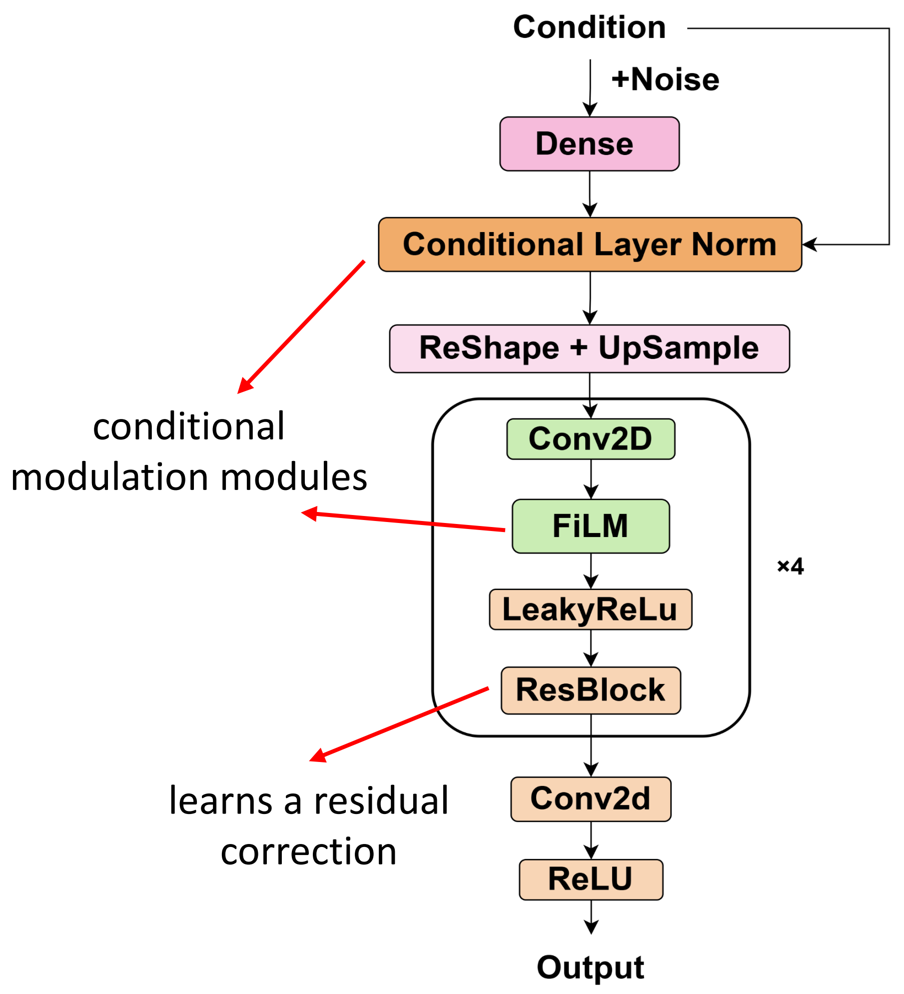


for photons:

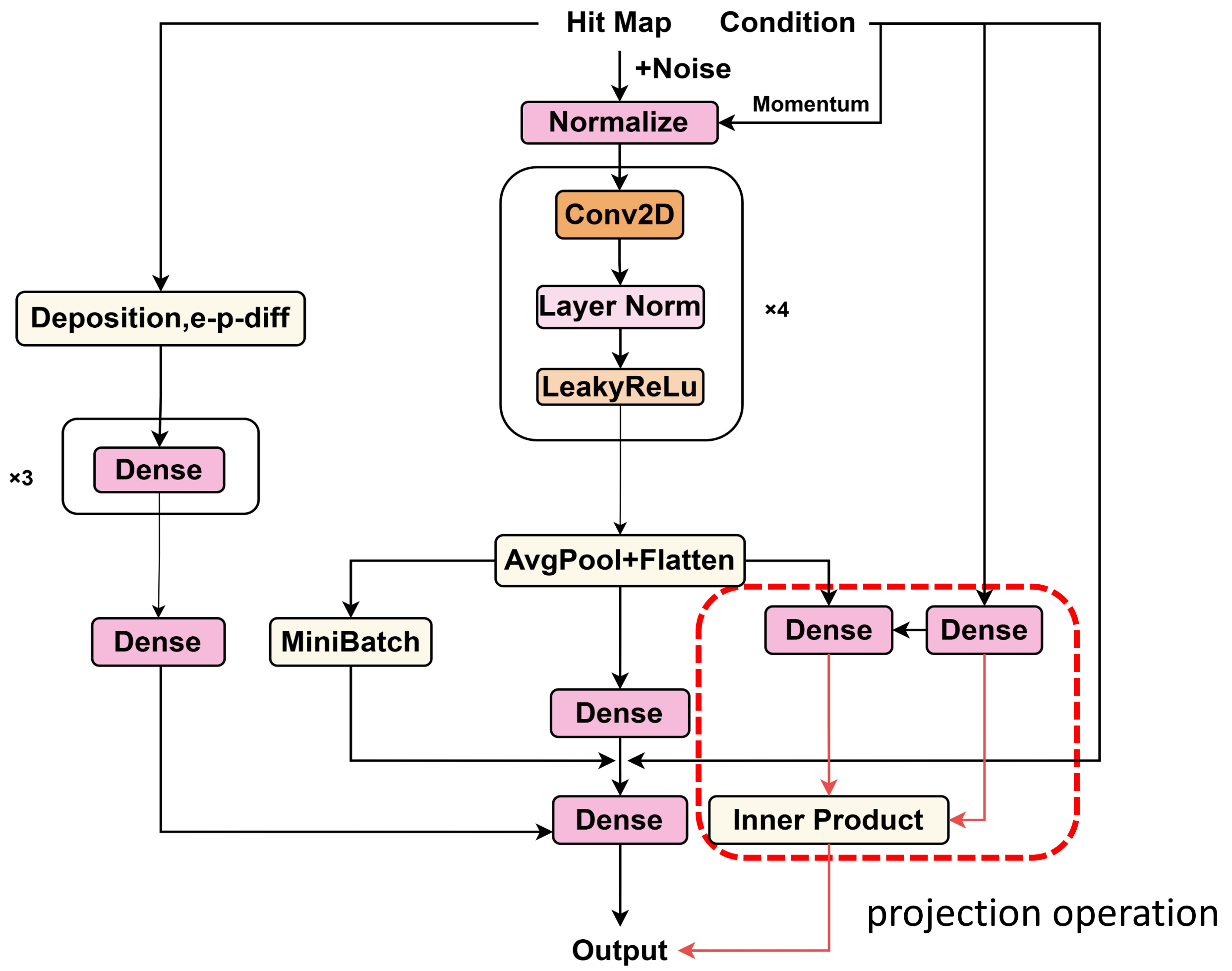


Framework of GAN

Generator, Discriminator: Convolutional Neural Networks(CNN) and MLP



Generator



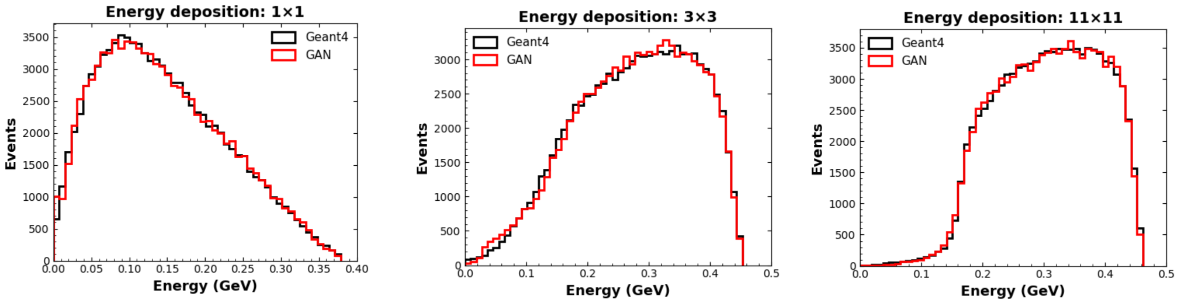
Discriminator

Performance on electron(full-space test set)

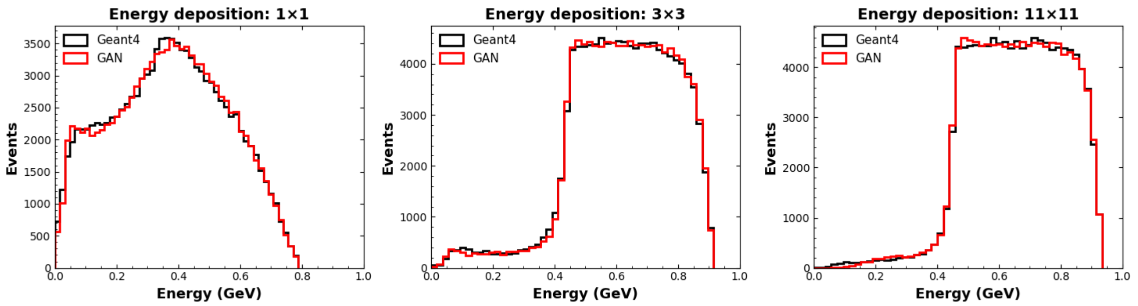
- Comparison of GAN simulation results and Geant4 results for electrons
- Including energy deposition in four different momentum intervals

$0 \leq P \leq 2$ GeV, $20° \leq \theta \leq 160°$, $0 \leq \phi \leq 360°$

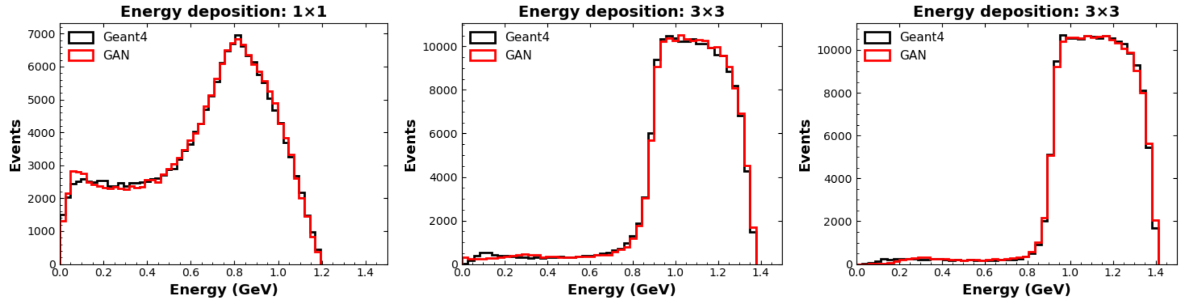
$0 \leq P \leq 0.5$ GeV



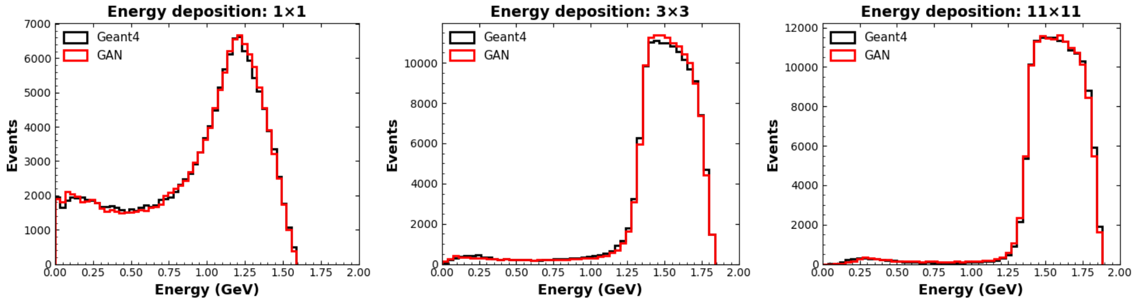
$0.5 \leq P \leq 1$ GeV



$1 \leq P \leq 1.5$ GeV

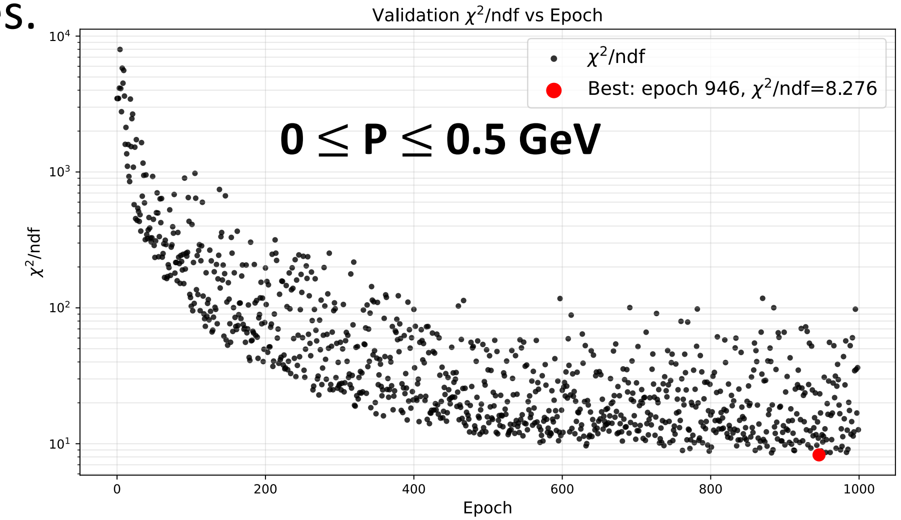


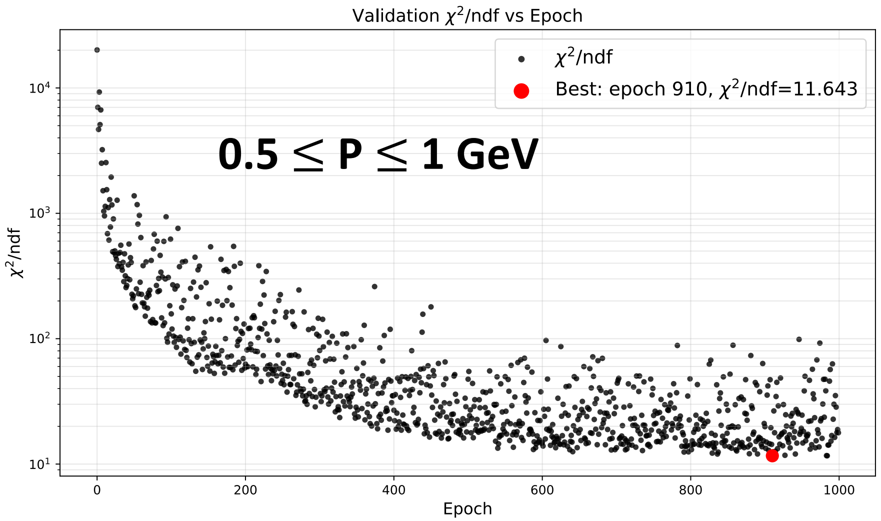
$1.5 \leq P \leq 2$ GeV

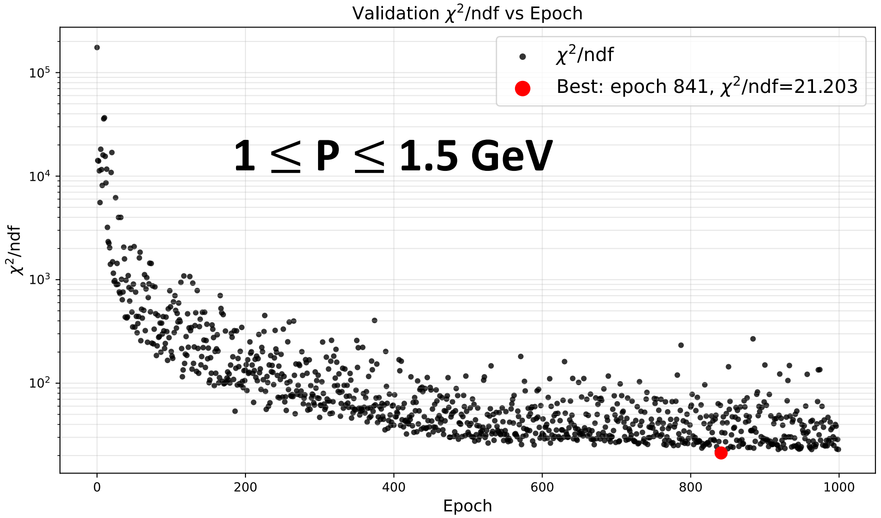


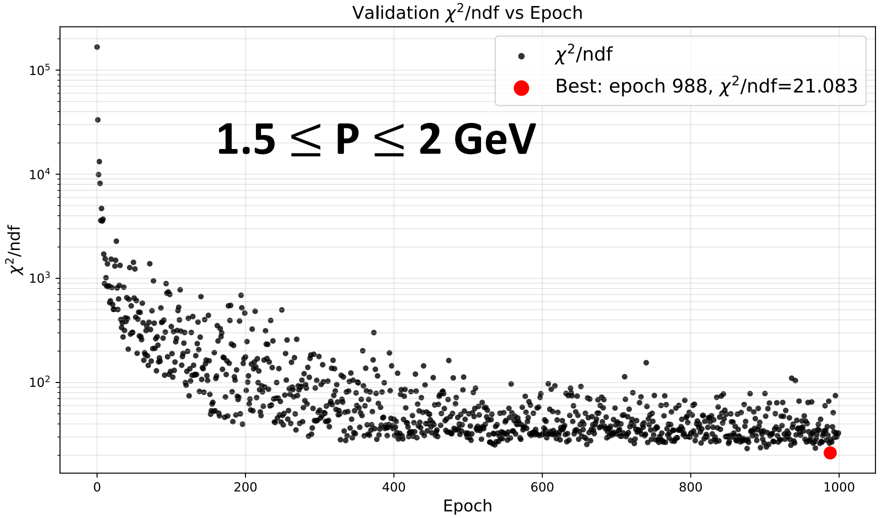
Best Model Selection

A weighted χ^2/NDF based on E1x1, E3x3, and E11x11 is used to evaluate the agreement between generated and real samples.









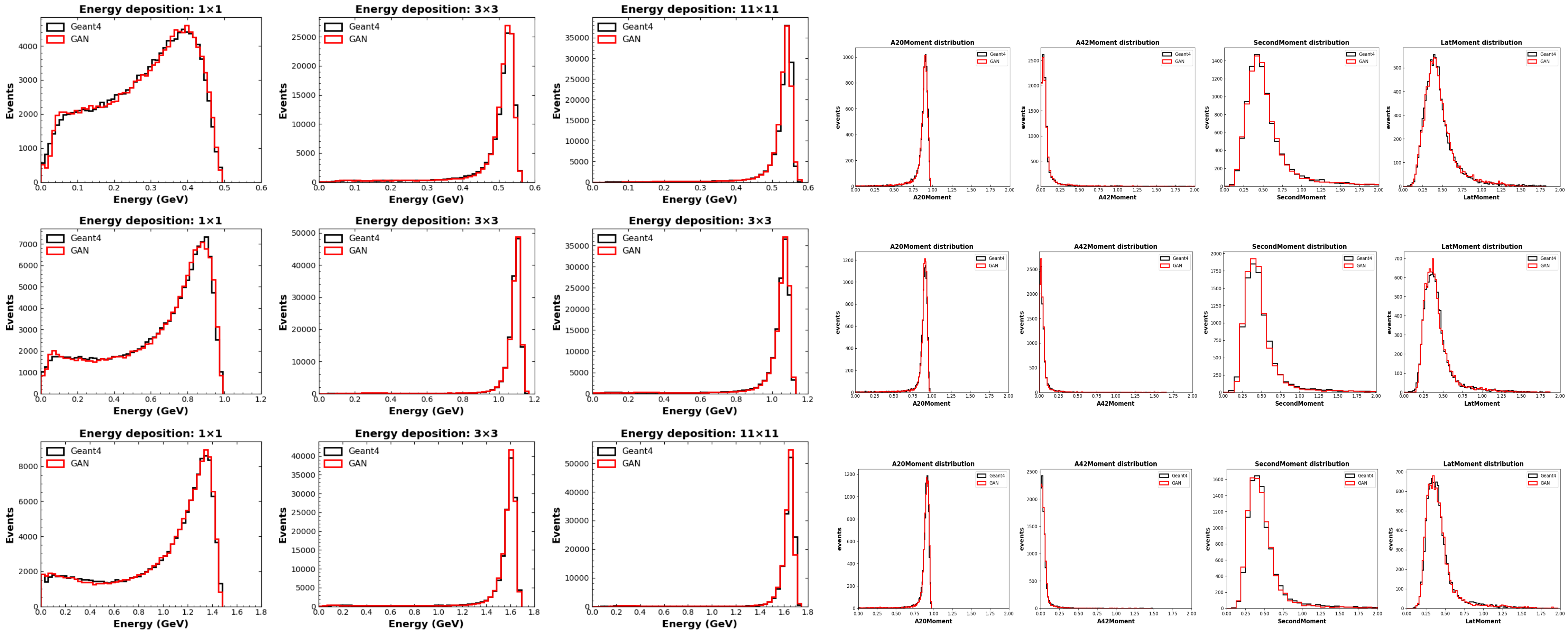
Performance on electron (single momentum test set)

- This comparison evaluates whether the GAN can correctly respond to different momentum conditions
- Including energy deposition in different regions and four complex observables

P = 0.6GeV

P = 1.2GeV

P = 1.8GeV



Boundry momentum points

P = 0.5GeV

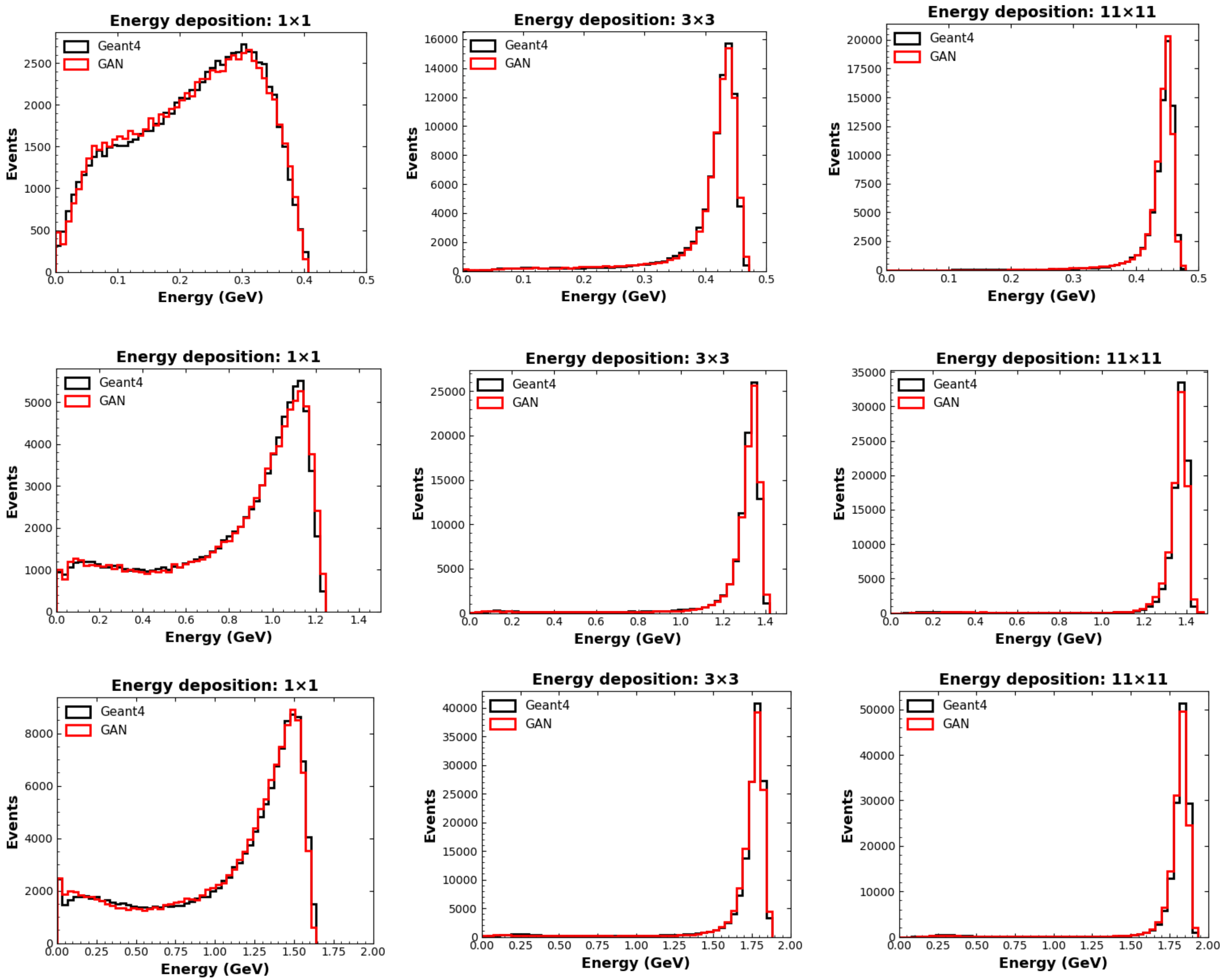
train 0.5-1GeV
with 0.45-1.05GeV dataset

P = 1.5GeV

train 1-1.5GeV
with 0.95-1.55GeV dataset

P = 2.0GeV

train 1.5-2GeV
with 1.45-2GeV dataset



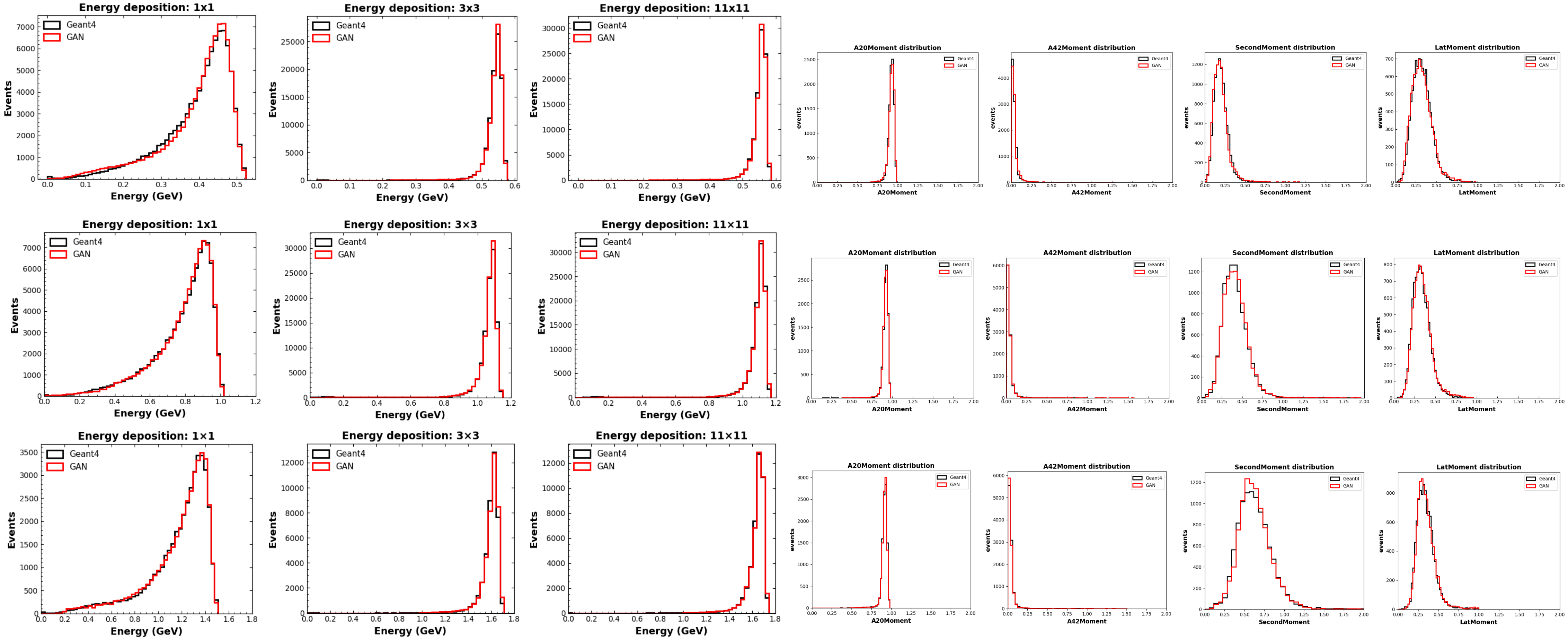
Performance on photon (single momentum test set)

- The model performs well in different momentum regions for photons

P = 0.6GeV

P = 1.2GeV

P = 1.8GeV



Simulation Speed Performance of GAN

- Multiple generation of 10^4 electron simulated samples for speed testing
- Comparison of average CPU time between two simulation methods for single particle ECAL simulation (Intel Xeon Gold 6338 @ 2.00 GHz)

simulation methods	energy	average time (ms)
Geant4	0-0.5GeV	32.41
GAN		0.15 (200x)
Geant4	0.5-1GeV	78.35
GAN		0.15 (500x)
Geant4	1to1.5GeV	122.05
GAN		0.15 (800x)
Geant4	1.5to2GeV	173.85
GAN		0.15 (1000x)

Overall, GAN simulation achieves a speedup of far more than 100 times compared with Geant4.

Summary & Outlook

- Calorimeter simulation is one of the most computationally intensive tasks in high-energy physics.
- Simulation for ECAL with cWGAN reaches comparable accuracy to Geant4 with orders-of-magnitude speedup.
- Four models are trained separately to improve the modeling performance in different momentum regions.
- Similar models can also be applied to other experiments, such as BESIII.
- Next to do:
  - Integration with Geant4: Simulation, Digitization and Reconstruction
  - Evaluate the model performance over the full physical process
  - Further optimize

Thanks for your attention!

Backup

Simulation Speed Performance of GAN

- Multiple generation of 10^4 photon simulated samples for speed testing
- Comparison of average CPU time between two simulation methods for single particle ECAL simulation (Intel Xeon Gold 6338 @ 2.00 GHz)

simulation methods	energy	average time (ms)
Geant4	0-0.5GeV	42.2
GAN		0.15 (200x)
Geant4	0.5-1GeV	45.2
GAN		0.15 (300x)
Geant4	1to1.5GeV	89.1
GAN		0.15 (500x)
Geant4	1.5to2GeV	90.3
GAN		0.15 (600x)

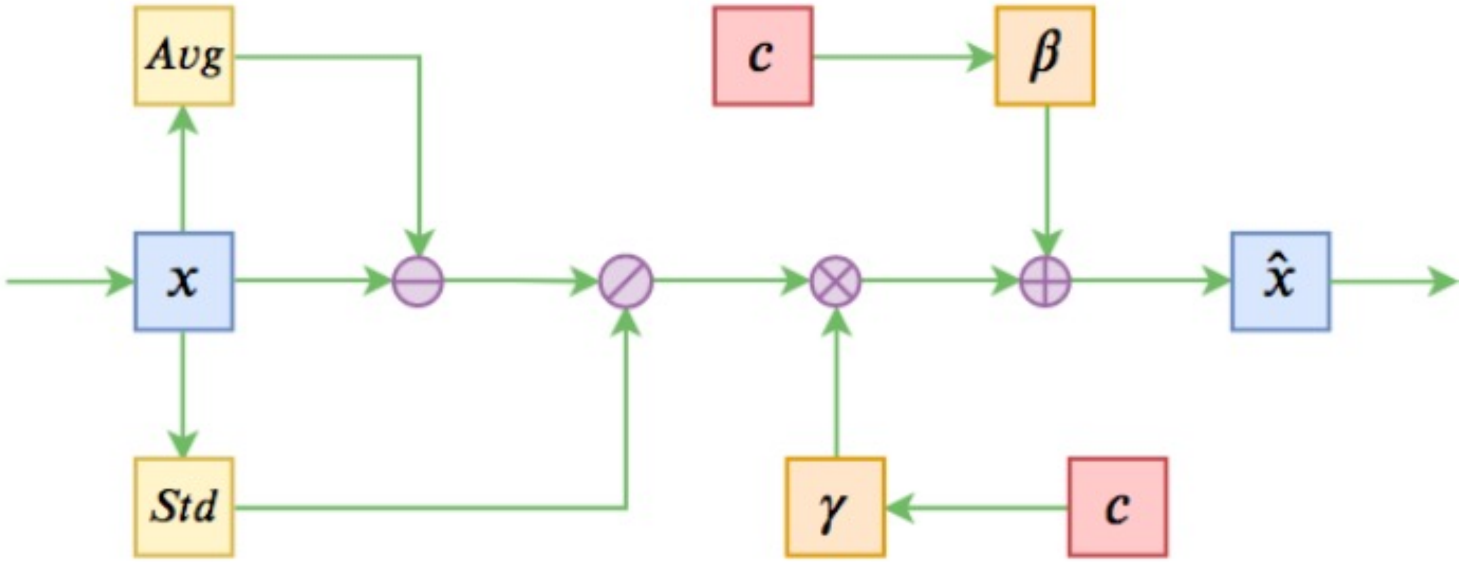
Overall, GAN simulation achieves a speedup of far more than 100 times compared with Geant4.

Conditional Layer Norm

Conditional Layer Normalization dynamically adjusts the normalization parameters according to the input condition, allowing the network to adapt its activations to the given condition.

$$LN(x) = \gamma \left(\frac{x-\mu}{\sigma}\right) + \beta$$

$$CLN(x) = \gamma' \left(\frac{x-\mu}{\sigma}\right) + \beta$$
$$\gamma' = \gamma + w_\gamma * c$$
$$\beta' = \beta + w_\gamma * c$$



cGAN with Projection D

In traditional GANs, the input to the discriminator is a sample x(e.g., an image). In the Projection Discriminator, in addition to the sample x, the conditional information c is mapped to the same space as the sample.

The projection operation can be achieved through inner product between the conditional vector c and the sample x. This projection typically allows the discriminator to effectively focus on the condition-related features in the sample space.

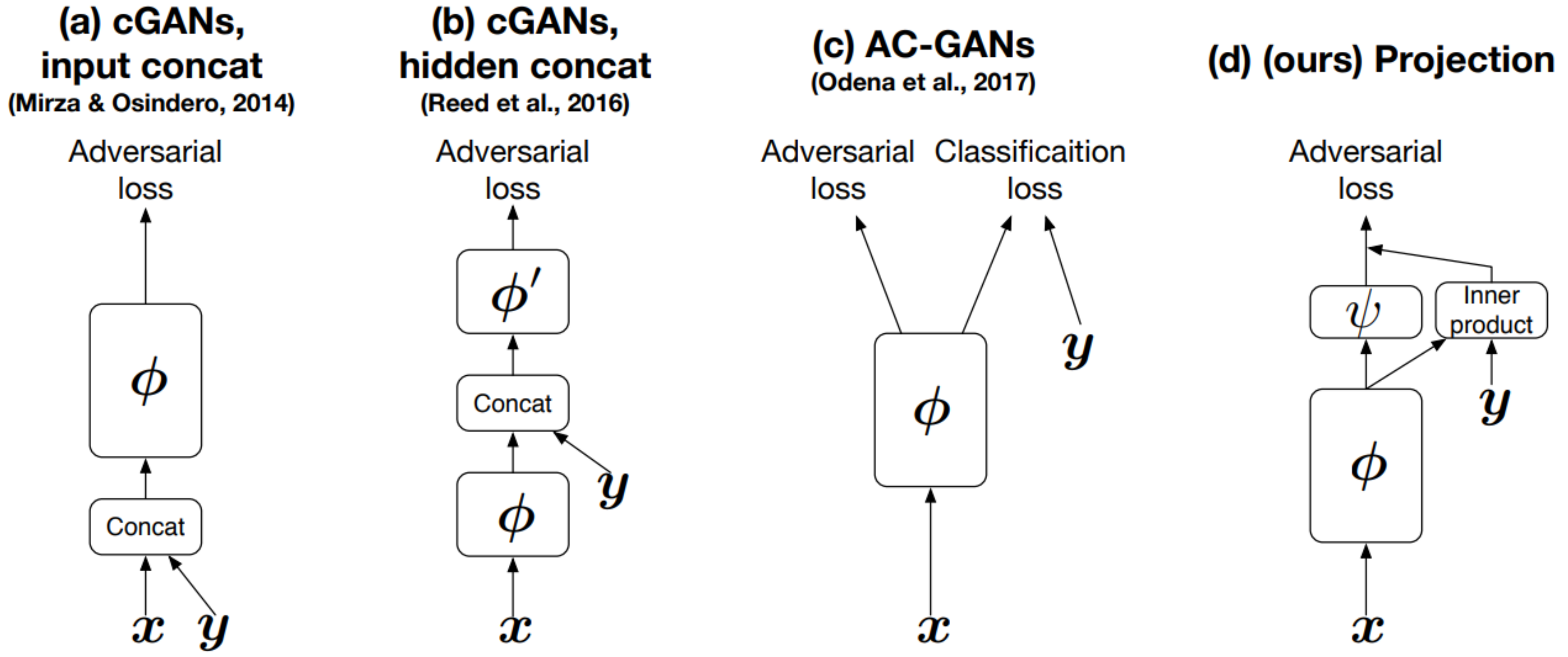
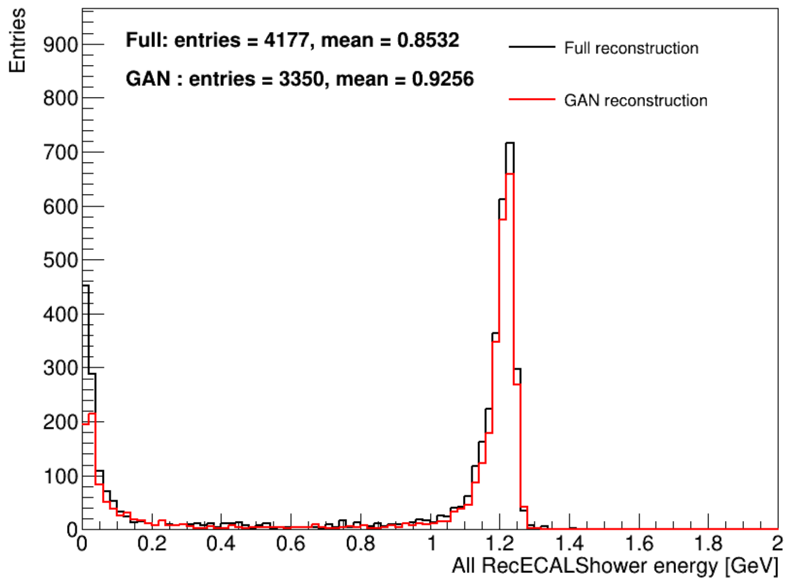
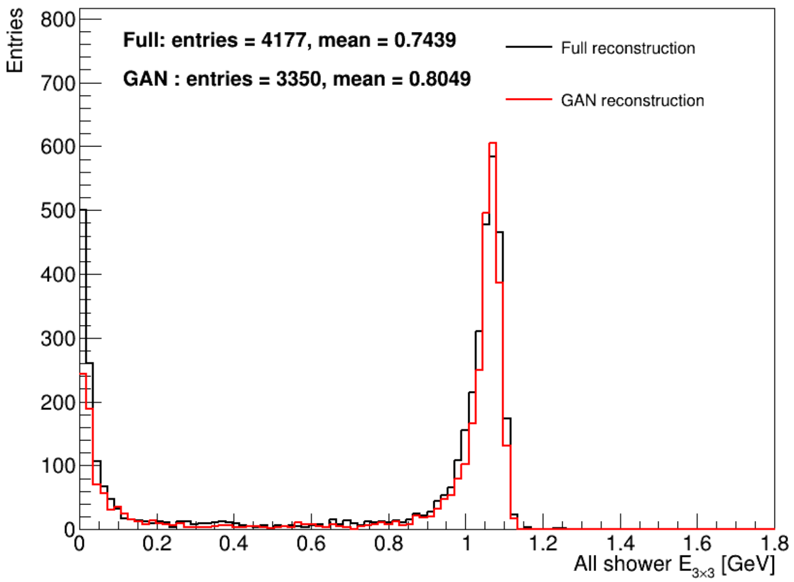
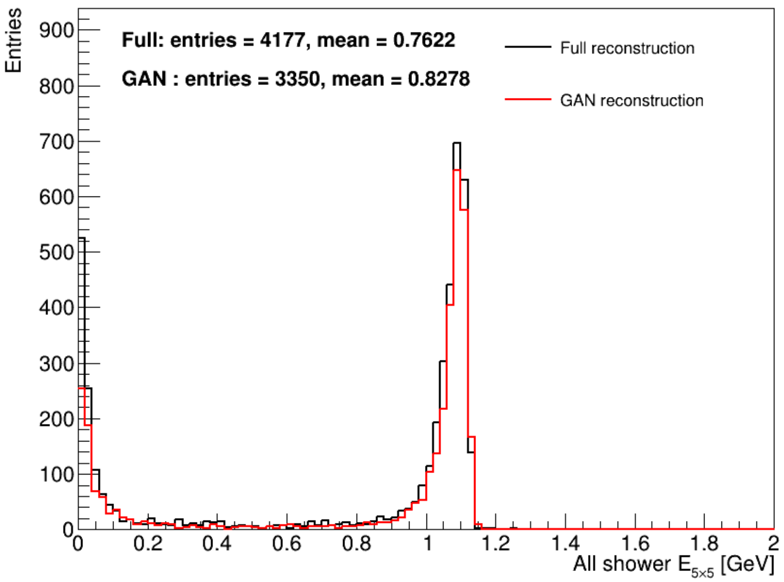


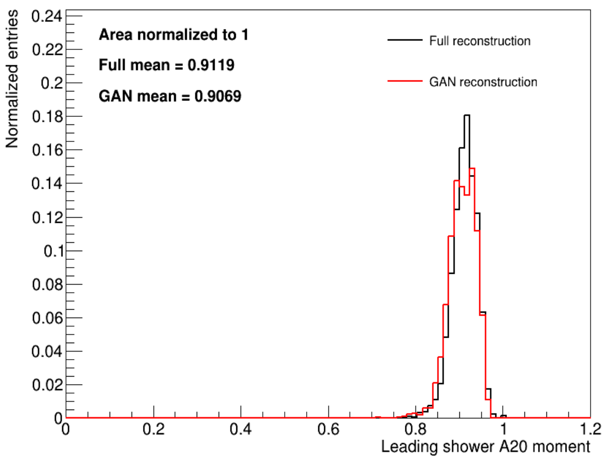
Figure 1: Discriminator models for conditional GANs

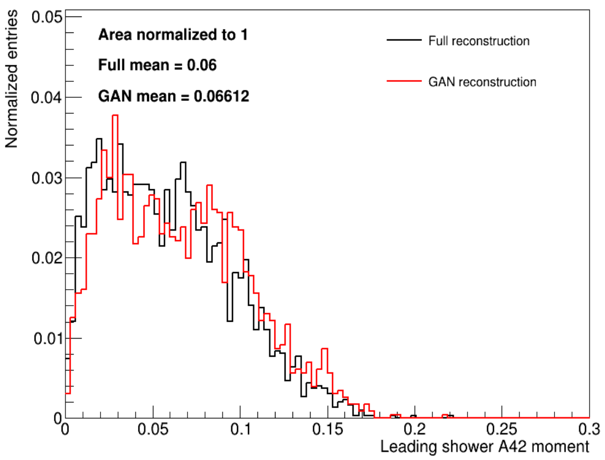
1.2GeV electrons Shower Test

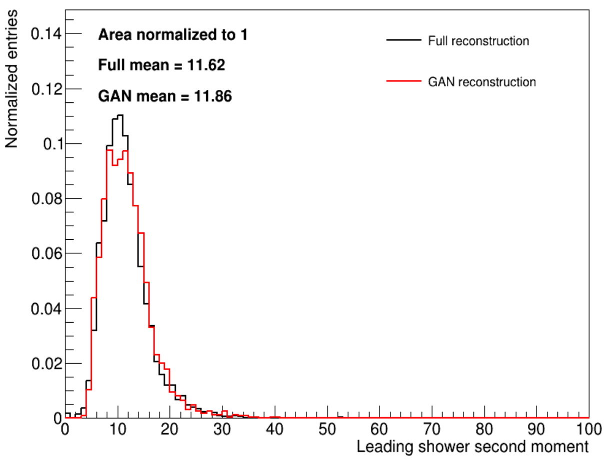


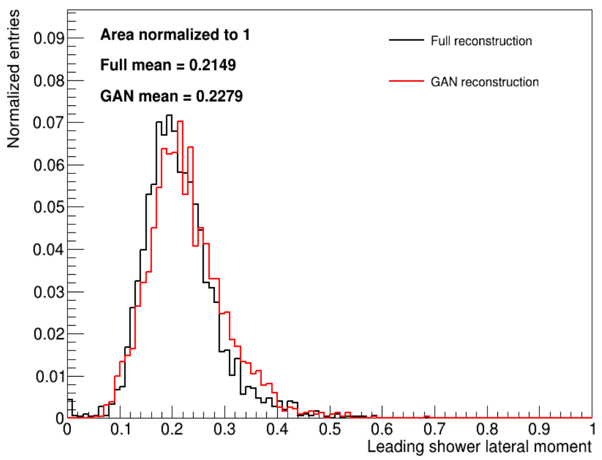




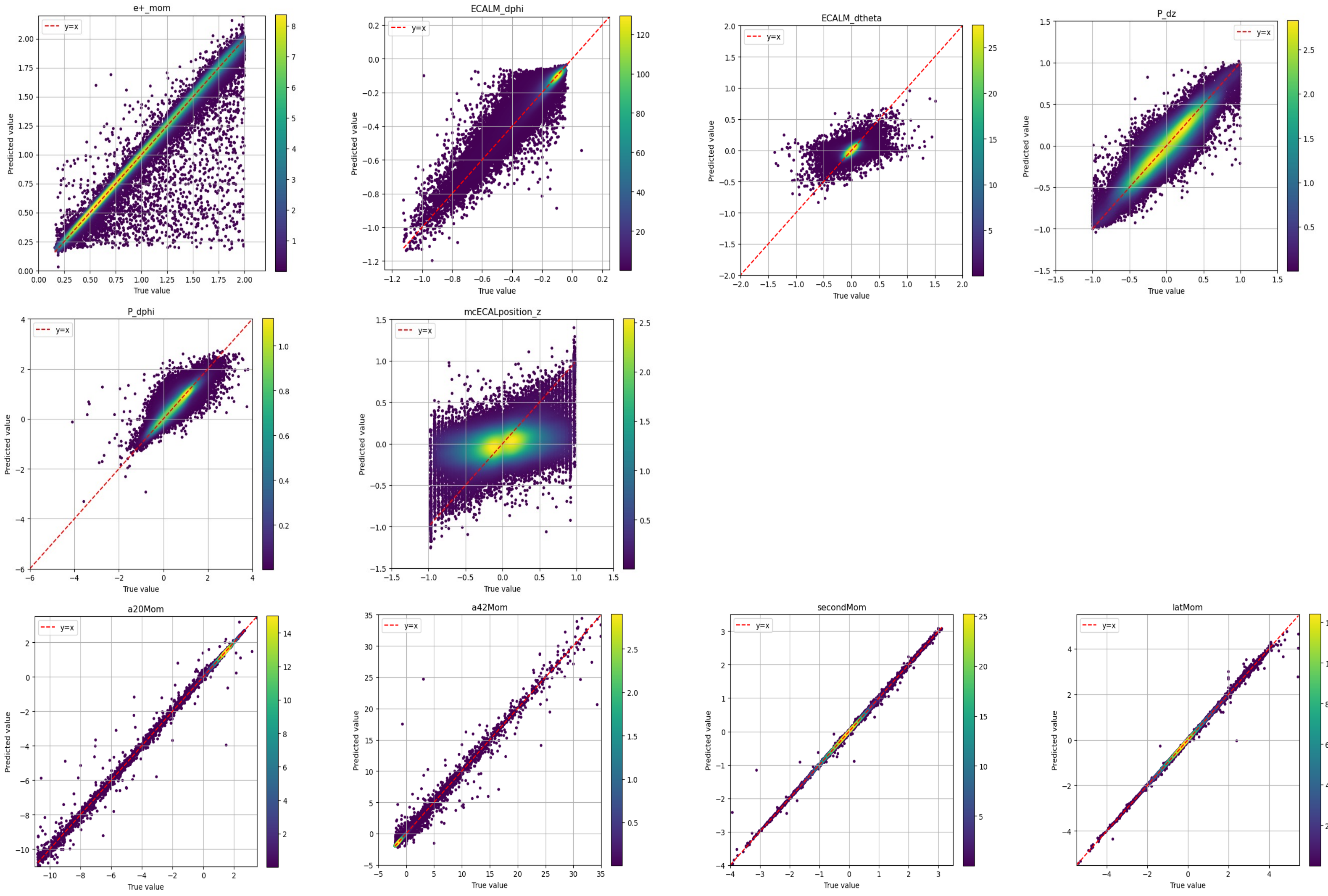




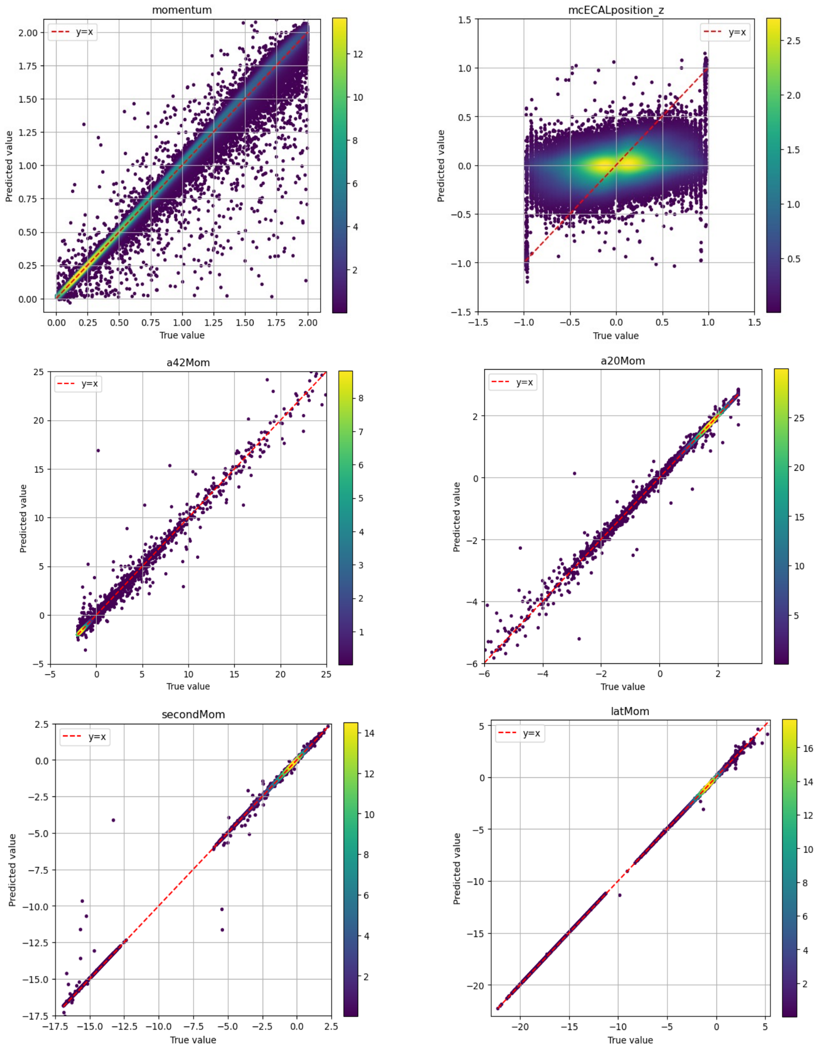




Regressor performance for e+ & γ



e+



γ