BESIII Electromagnetic Calorimeter Fast Simulation

Tong Liu[1], Qiaojia Ge[2], Wenxing Fang[1], Weidong Li[1], Rui Zhang[2], Ke Li[1]

[1] IHEP

[2] NJU

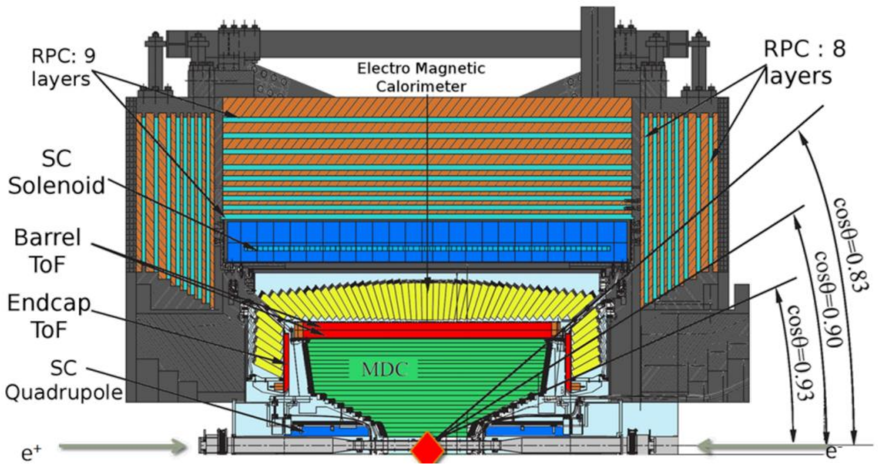
Outline

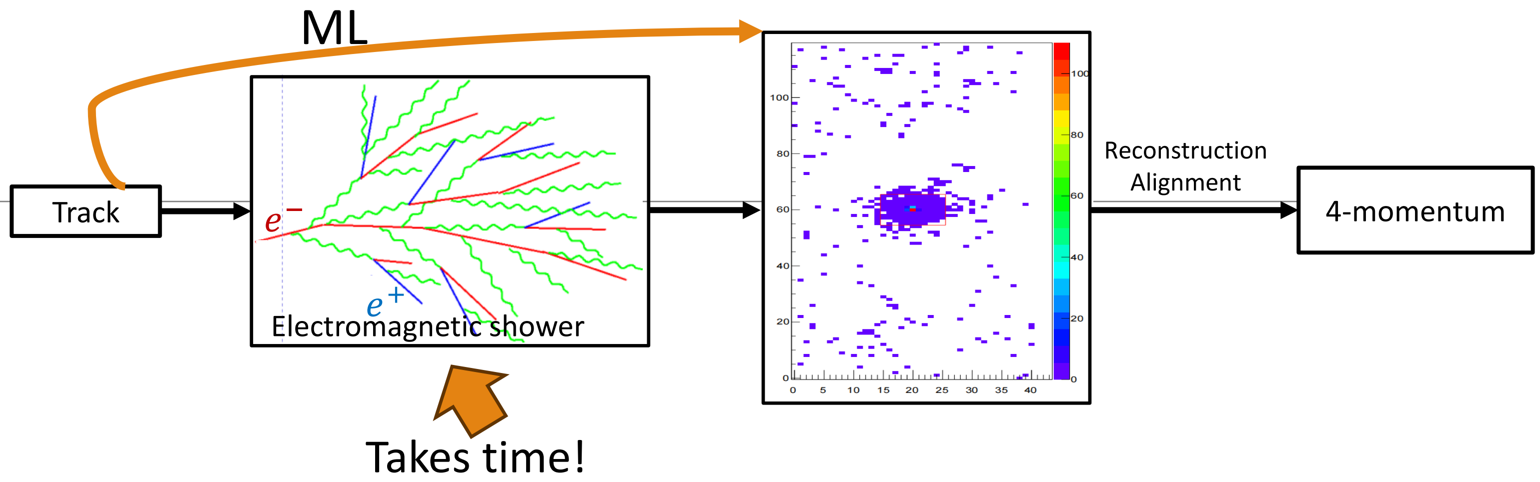
- Introduction & Training samples
- EMC simulation with GAN
- EMC simulation with diffusion
- Summary & Next

Introduction & Training samples

Introduction

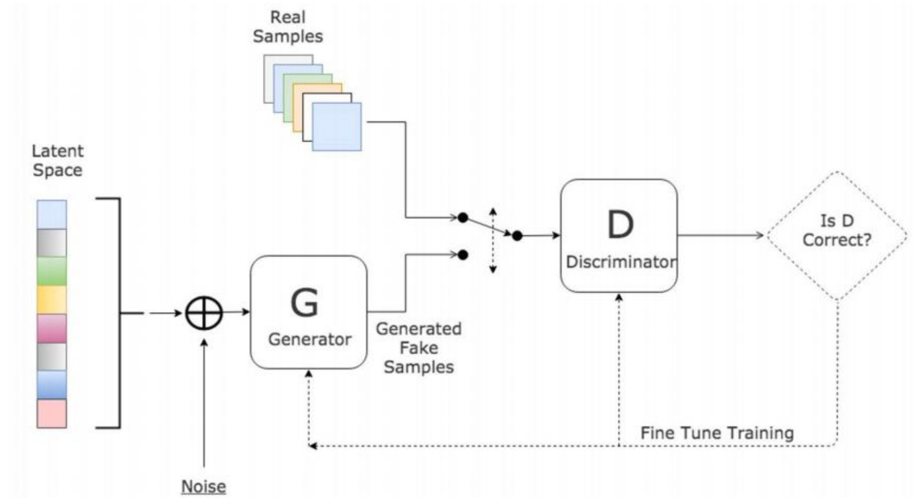
- With improved luminosity, MC with higher statistics is required
- MC simulation, especially for electromagnetic calorimeter (EMC), takes large CPU resources
  - Traditional: Geant4, gradually calculate the next state, complex due to secondary particles
  - ML: without Geant4, calculate hit map from input conditions
- EMC: 44 layer*120 crystals in barrel, we focus on barrel region firstly
  - To avoid energy leakage caused by the gap



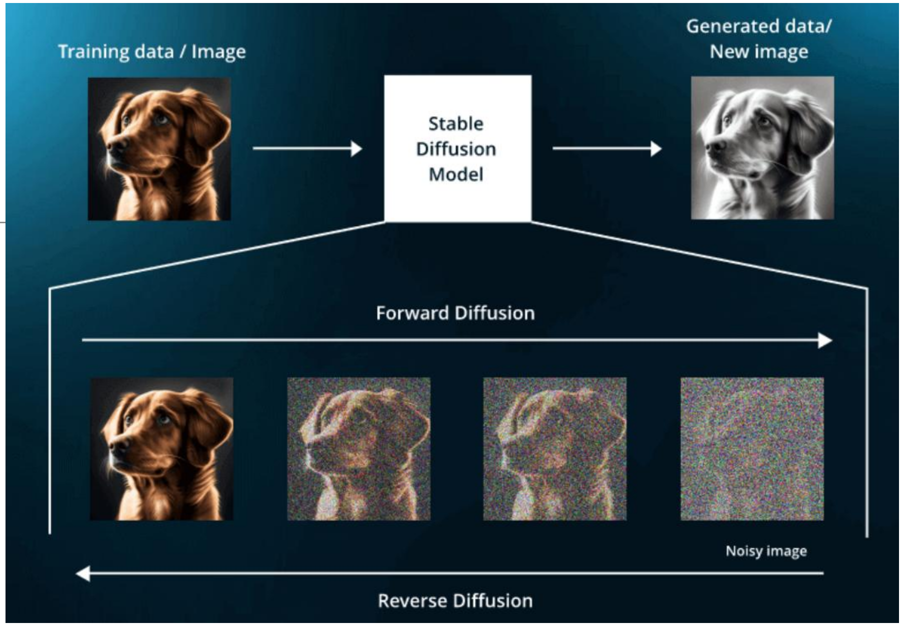


Introduction

- GAN and diffusion models are tested in this work
- Generative Adversarial Networks (GAN) arXiv:1406.2661
    - A discriminator tries to discriminate the real and fake data; a generator to produce fake data, tries to confuse discriminator
    - Train D and G alternately to improve performance
- Diffusion arxiv: 2006.11239
    - Add noise steply, train a model to do denoising



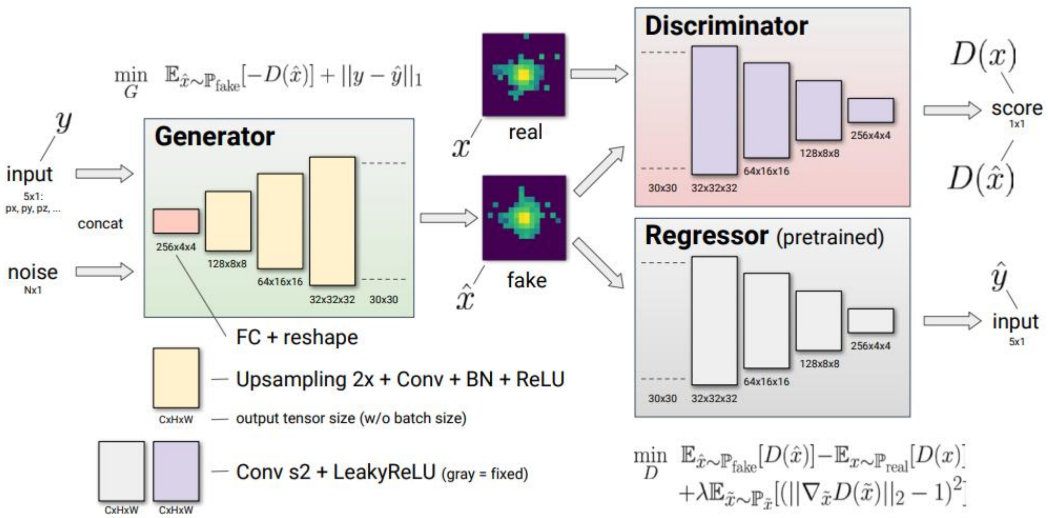
$$\min_G \max_D V(D,G) = \mathbb{E}_{x\sim p_{data}(x)}[\log D(x)] + \mathbb{E}_{z\sim p_z(z)}[\log(1 - D(G(z)))]$$



Introduction

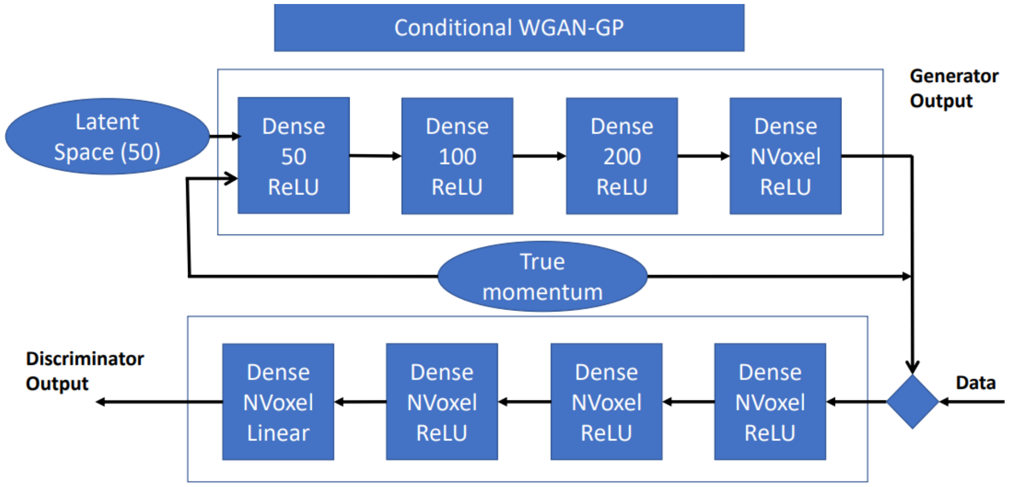
- ML, especially GAN, is used for simulation at LHCb and ATLAS
  - A pretrained regressor, for further constrain G result
  - Train ~ 100 GANs for different ATLAS detector regions
- BESIII is an ideal place to perform ML-simulation: simpler detector, smaller condition space

EPJ Web Conf. **214** 02034 (2019)



caloGAN @ LHCb

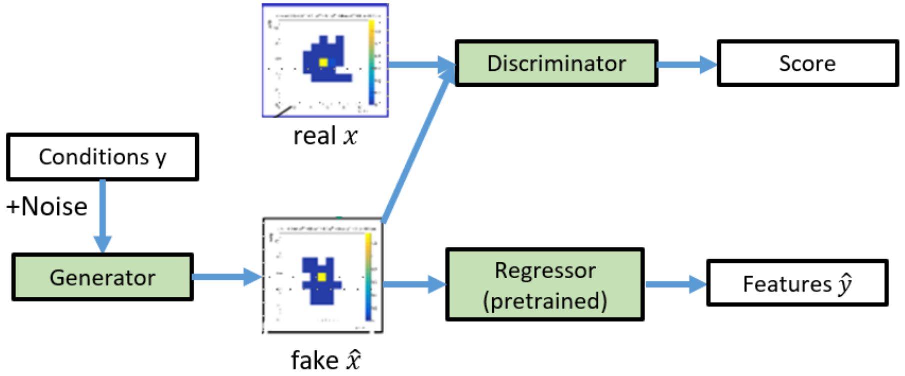
ACAT 2024

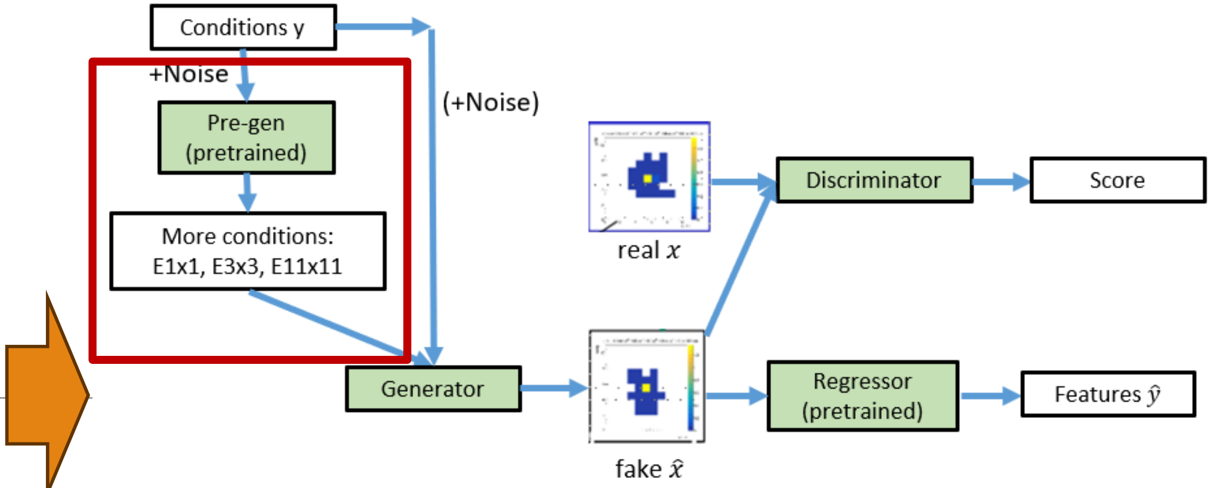


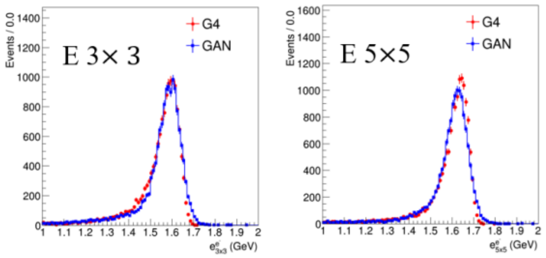
FastCaloGAN @ ATLAS

An Advanced Condition Generator

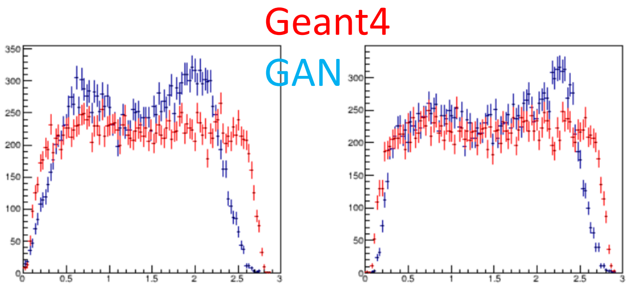
- BESIII tests the GAN simulation based on selected Bhabha ($e^+e^- \to e^+e^-$) events with caloGAN method
- However, the model does not work well in a larger condition space
- To cover the full condition space, a pretrained pre-gen model is integrated to improve GAN performance



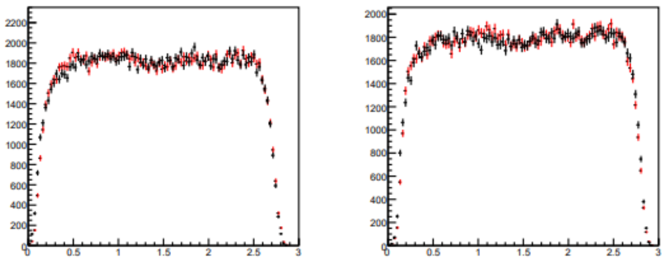




GAN for Bhabha events



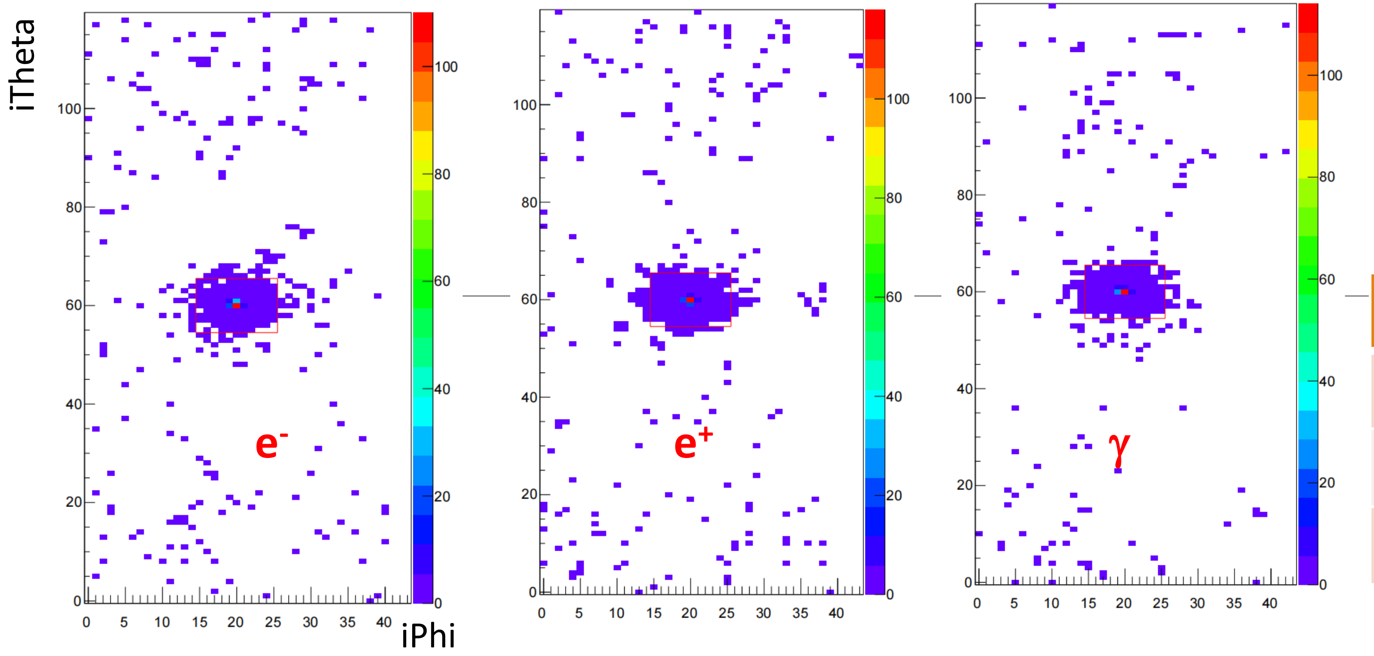
GAN in large condition space



Pre-gen + GAN

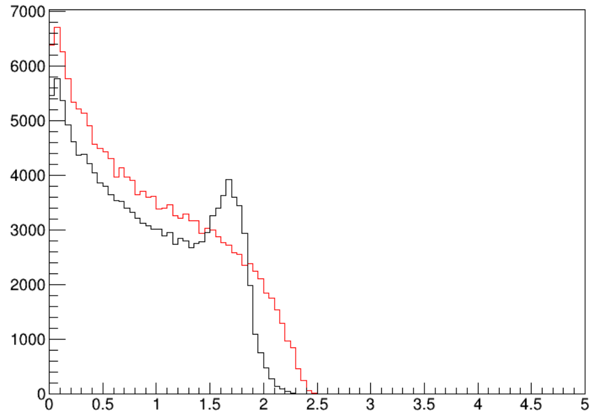
Samples

- We simulate ~1M single-track events for $e^+/e^-/\gamma$ as training set
  - With $0 < P < 3$ GeV$/c$, $0 < \theta < 2\pi$, $0 < \phi < 2\pi$
  - Single-momentum samples: $P = 0.5, 0.8, 1.2, \ldots 2.5$ GeV$/c$
  - Remove random trigger, randomness of IP, readout cutoff
- The 11x11 region contains nearly the entire detector response



EMC hit map in barrel region from 100 events

	E_{11x11} / E_{tot} (%)
e^+	99.7
e^-	99.8
γ	99.8

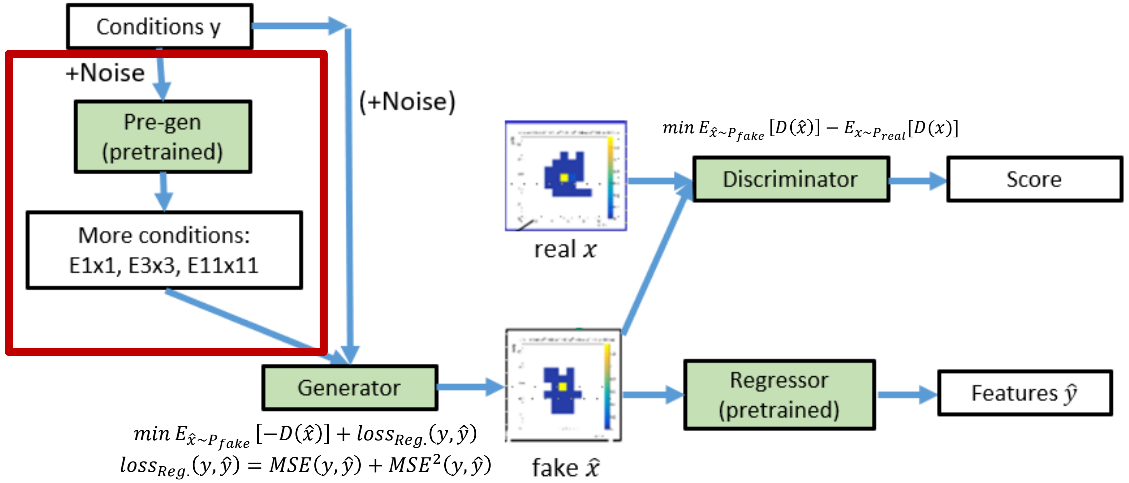


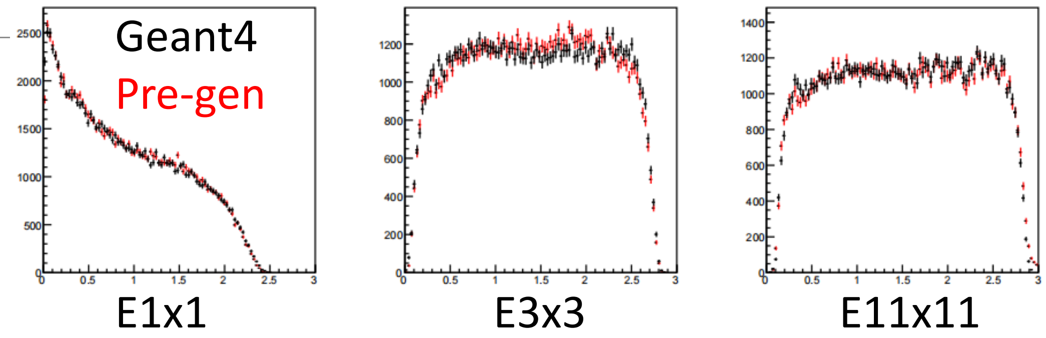
E1x1 w/ and w/o readout cutoff

EMC simulation with GAN

EMC simulation with ML - GAN

- Similar strategy as LHCb case: G + D + regressor (pretrained)
- Add a pre-gen model: an advanced condition generator
  - The full-space sample: the basic training
  - Single-momentum samples: compare the resolution after each 5 steps and apply an additional optimization step

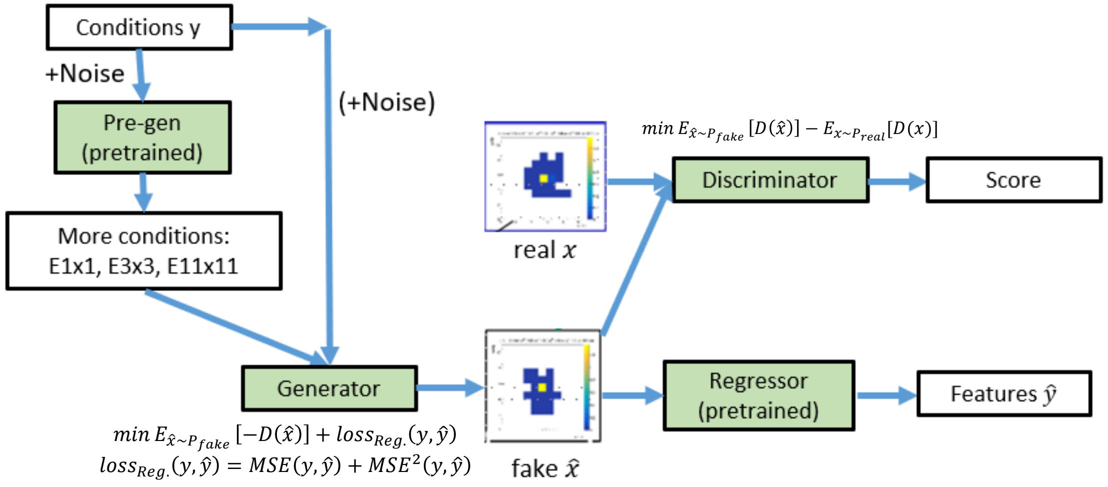


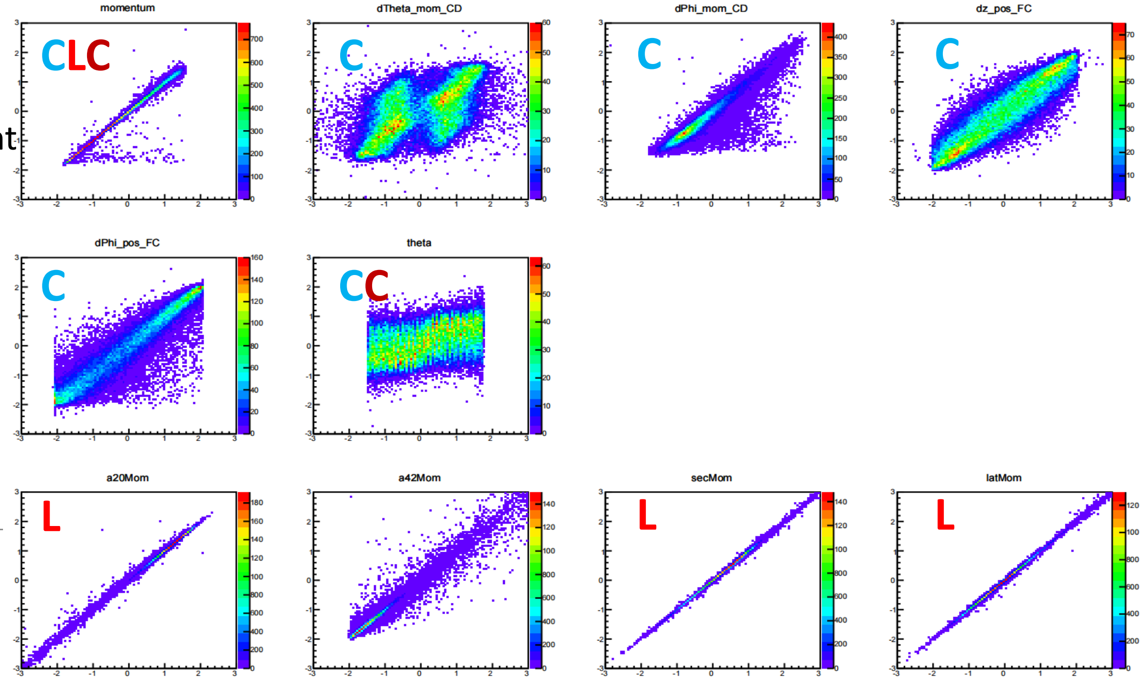


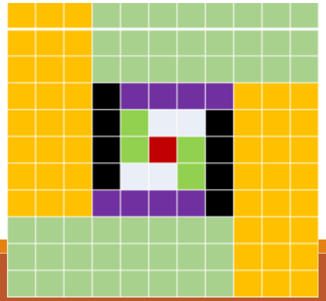
EMC simulation with ML - GAN

C: condition
L: for loss
C: condition for γ

- Similar strategy as LHCb case: G + D + regressor (pretrained)
- Use observables to constrain
  - Use regressor to reconstruct complex observables:
    - A20 moment, A42 moment, Second moment, lateral moment
  - Calculate deposition in regions for $loss_{reg}$:
    - E1x1,…E11x11
    - Re-binned deposition
- ~46k trainable parameters in generator







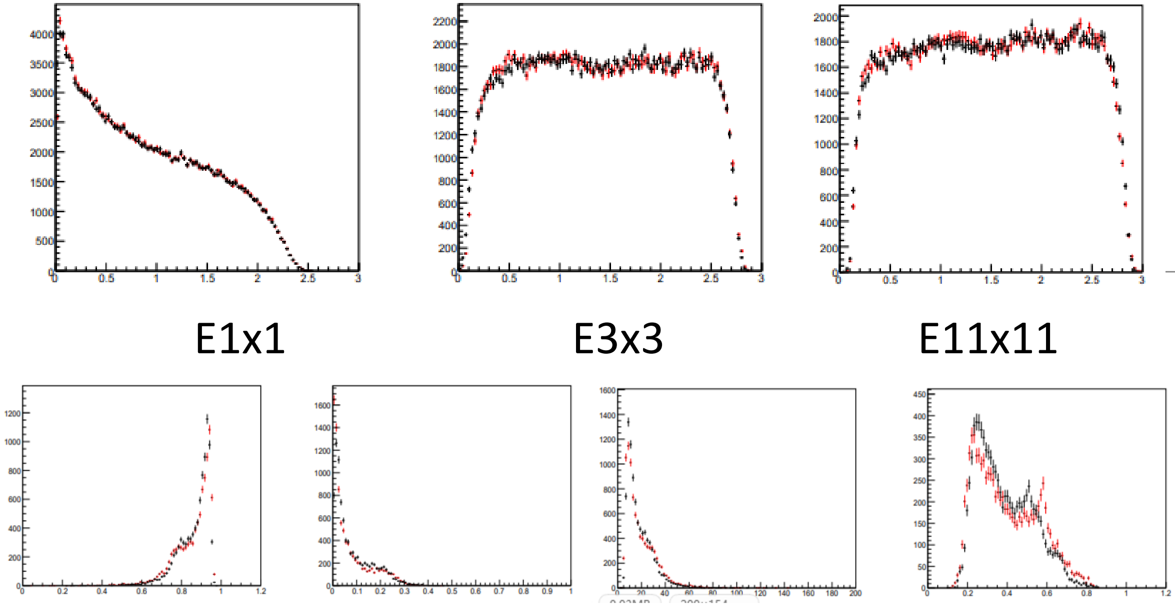
Re-binned deposition in 11x11 region

EMC simulation with ML – GAN for e⁺

- ML reachs comparable accuracy to Geant4
  - In the full condition space, and for the single-momentum sample
- 50k tracks ~ **2s**

Geant4
GAN

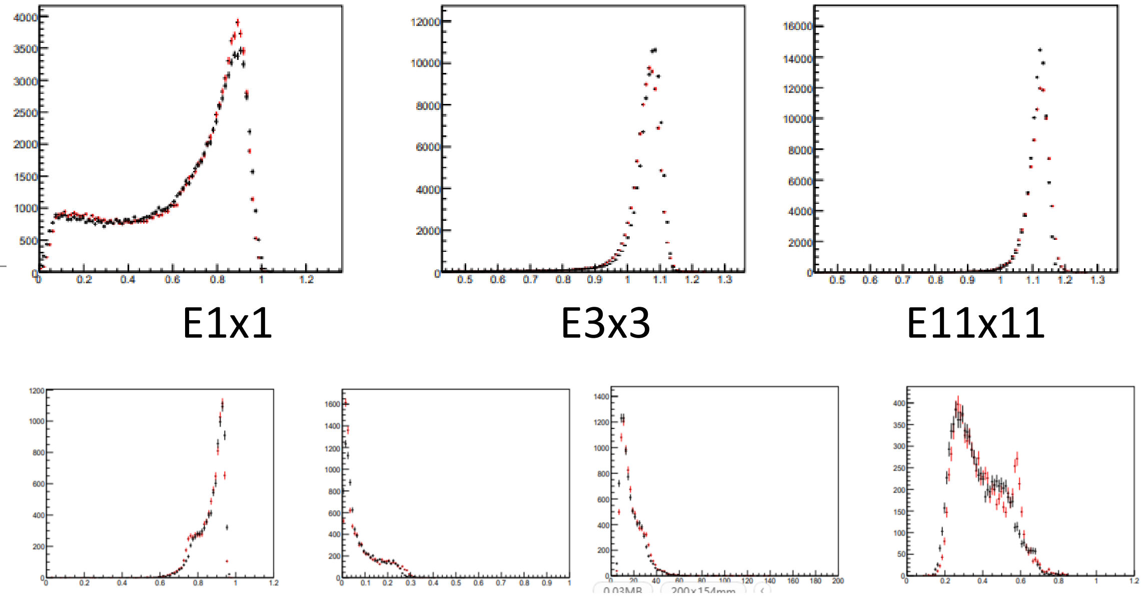
$\mathbf{0 < \boldsymbol{P} < 3\ GeV/\boldsymbol{c}}, 0 < \theta < 2\pi, 0 < \phi < 2\pi$



E1x1 E3x3 E11x11

a20Mom a42Mom secondMom latMom

$\boldsymbol{P} = \mathbf{1.2\ GeV/\boldsymbol{c}}, 0 < \theta < 2\pi, 0 < \phi < 2\pi$



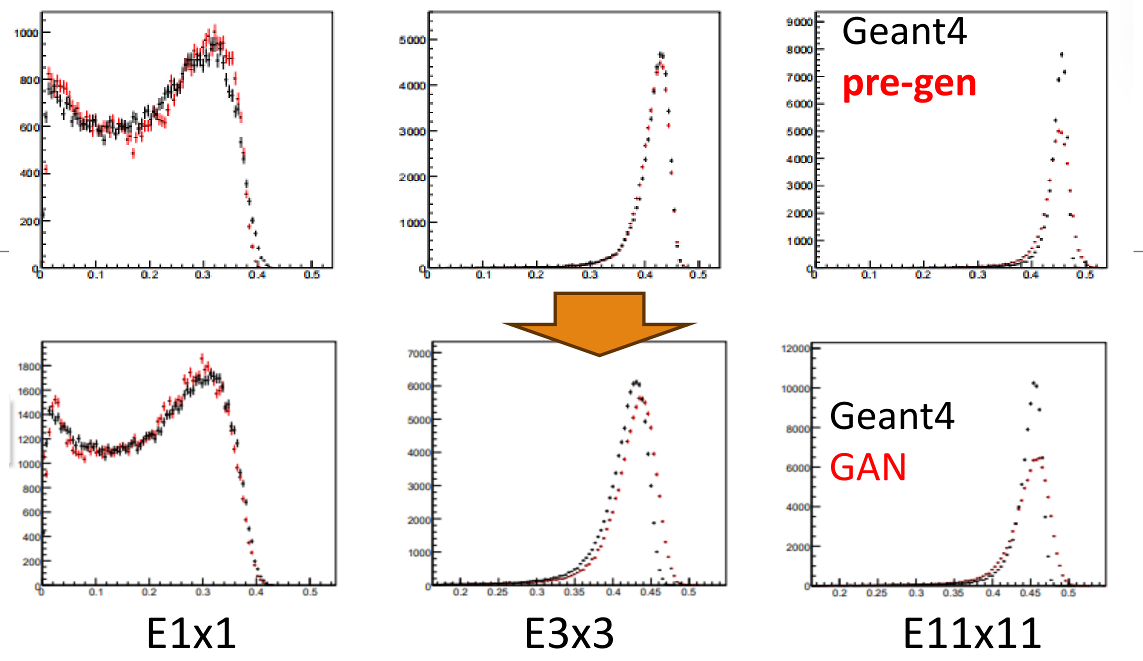
E1x1 E3x3 E11x11

a20Mom a42Mom secondMom latMom

EMC simulation with ML – GAN for e$^+$

- ML reach comparable accuracy to Geant4
- ~ 2s for 50k tracks
- However, the resolution for the lower momentum tracks is larger than Geant4
  - Seems correlated to the pre-gen quality: more refined training for the pre-gen model
  - Or increase the size of sample in lower momentum region

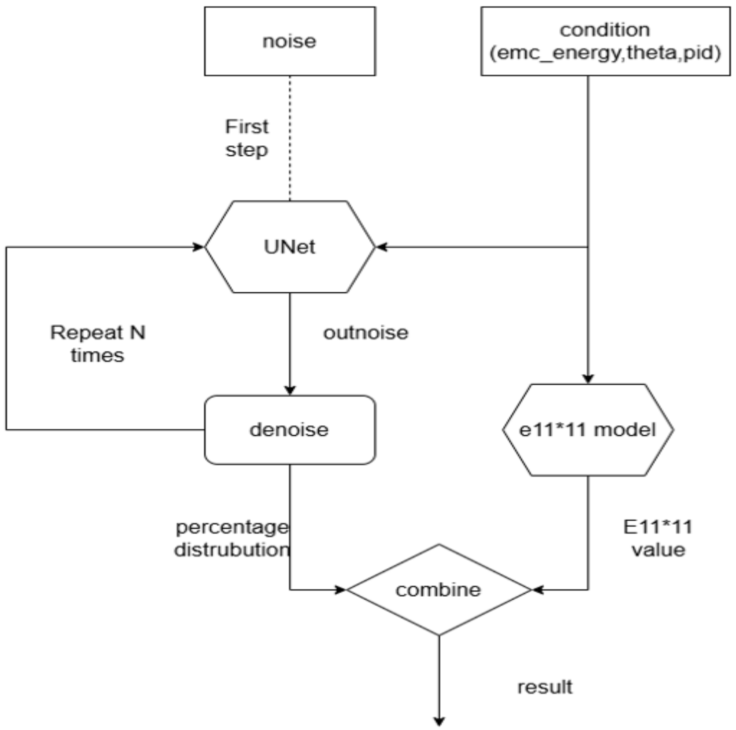
$$\boldsymbol{P = 0.5\ \mathrm{GeV}/c}, 0 < \theta < 2\pi, 0 < \phi < 2\pi$$

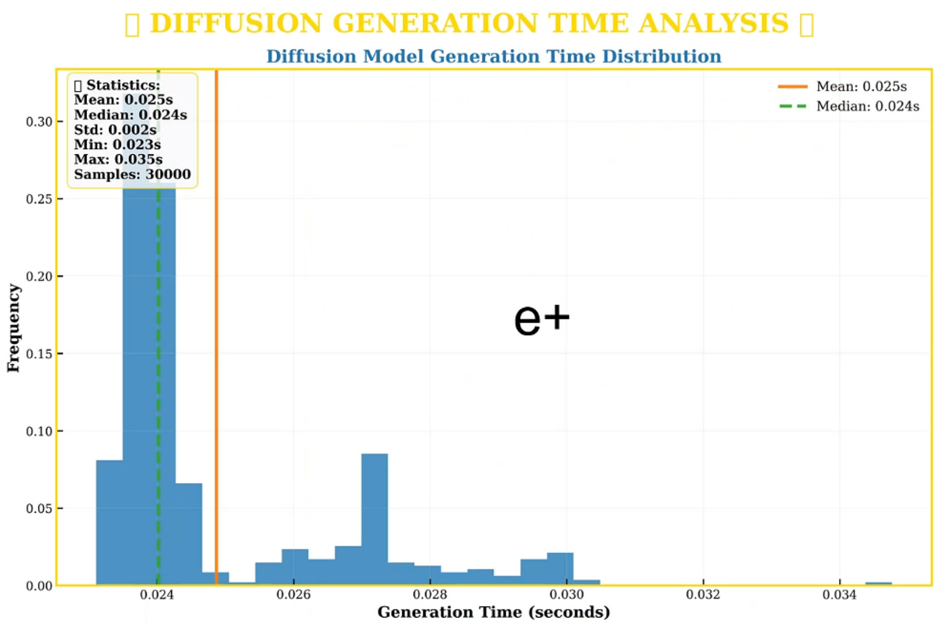


EMC simulation with diffusion

EMC simulation with ML – diffusion for e^+

- Turn to diffusion as a possible alternative approach
  - ~0.025 s per track

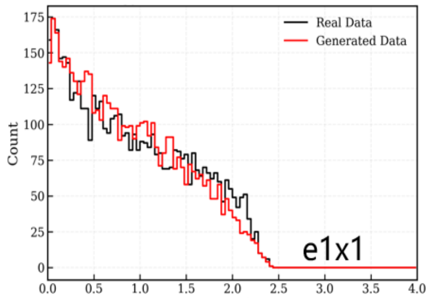


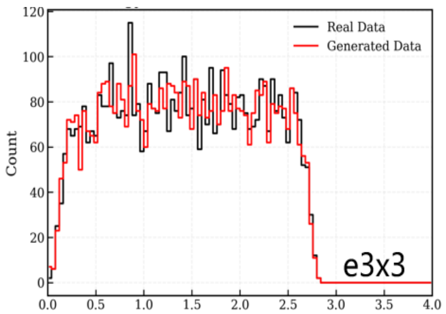


EMC simulation with ML – diffusion for e^+

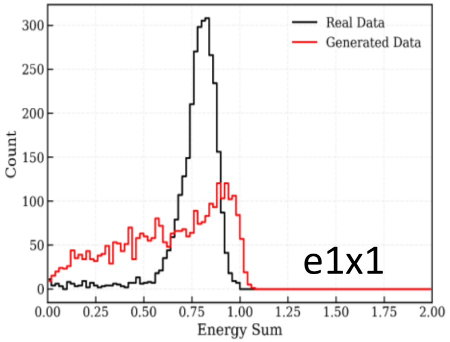
- Turn to diffusion as a possible alternative approach
  - Comparable accuracy to Geant4
  - The resolution needs to be improved

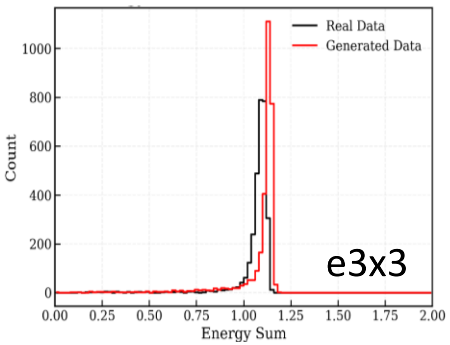
$\boldsymbol{0 < P < 3}$ **GeV**$/\boldsymbol{c}, 0 < \theta < 2\pi, 0 < \phi < 2\pi$





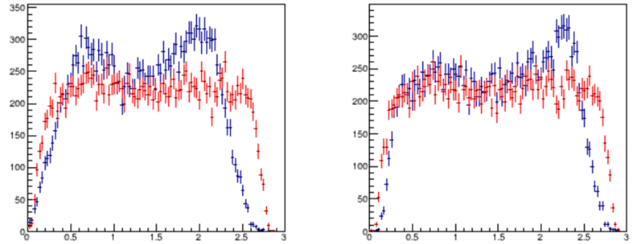
$\boldsymbol{P = 1.2}$ **GeV**$/\boldsymbol{c}, 0 < \theta < 2\pi, 0 < \phi < 2\pi$



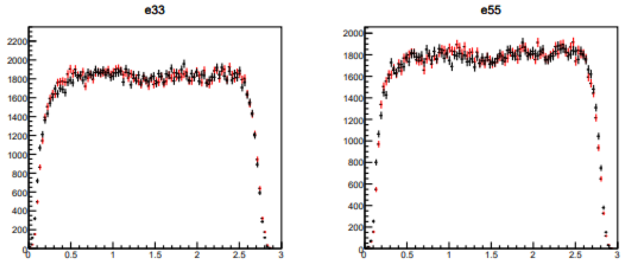


Summary & Next

- We use ML to speed up the simulation for electromagnetic calorimeter
- GAN: ~ 2s for 50k tracks
    - A pretrained pre-gen model improves GAN performance in the full condition space
- Diffusion: ~ 0.025s per track
    - Although slower than GAN, it is still faster than Geant4
- The preliminary model reach comparable accuracy to Geant4
- Next:
    - Conduct more refined training to optimize the accuracy
    - Increase sample in lower momentum region
    - Models for other tracks, specially, a pion shower results in a full-detector readout
    - Other similar experiments, like STCF



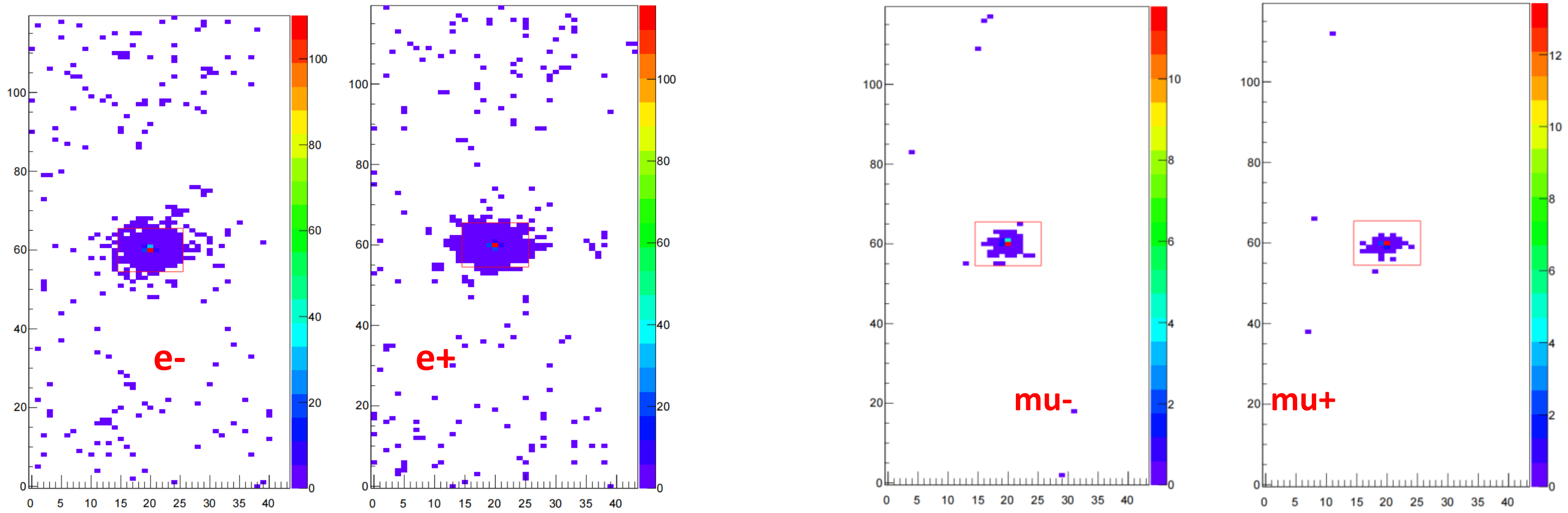
GAN w/o pre-gen



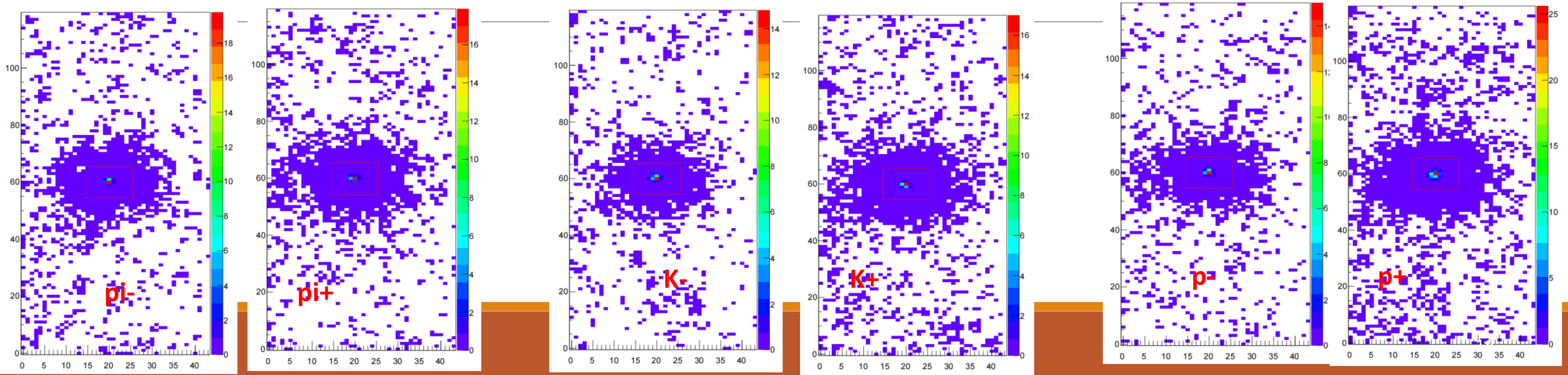
pre-gen + GAN

Readout region

- EMC hit map in barrel region from 100 events with P>1 GeV/c



	E11x11 / Etot (%)
ep/em	99.7/99.8
mup/mum	99.9/99.9
pip/pim	89.7/85.5
Kp/Km	89.1/86.0
Pp/pm	88.7/84.8

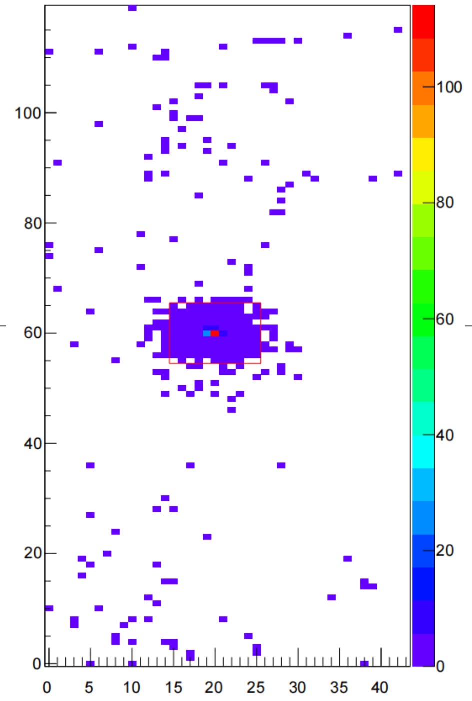


Readout region

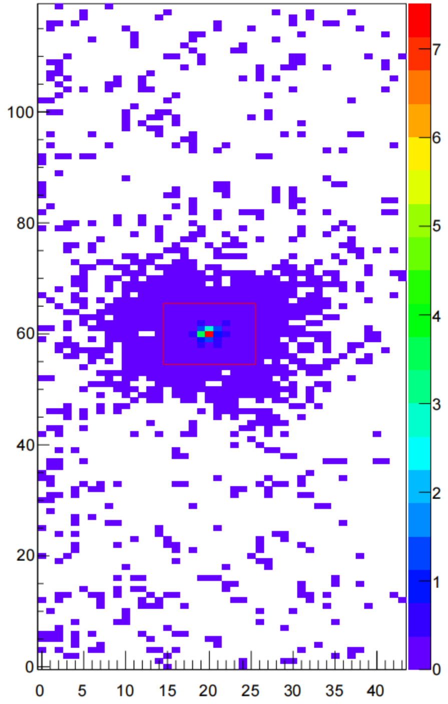
- EMC hit map in barrel region from 100 events with P>1 GeV/c

	E11x11 / Etot (%)
Gamma	99.8
n	87.7
nbar	91.6

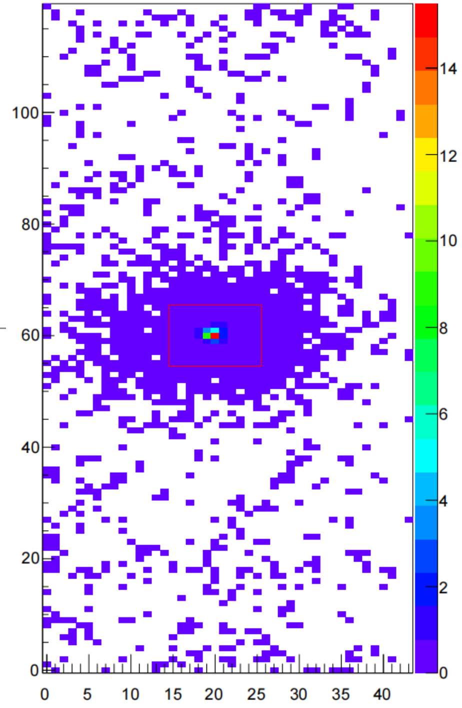
gamma



n



nbar



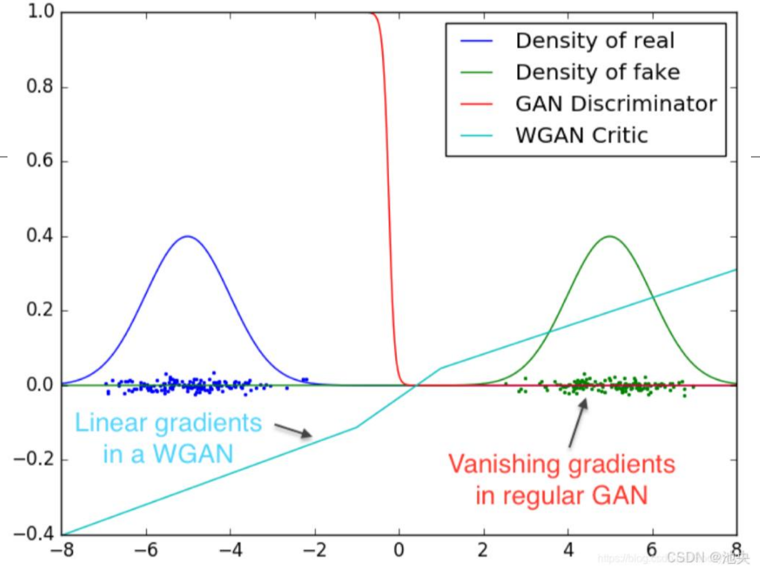
Introduction-wGAN

arXiv:1701.07875

- Wasserstein GAN, an improved GAN
  - To solve instability and gradient vanishing
    - Especially if two distributions are non-overlapping.
  - Replace Discriminator with a Critic
    - Replace JS divergence with the Wasserstein distance, print scores instead of probabilities
    - Lipschitz Constraint: via weight clipping or gradient penalty (wGAN-GP)
  - Loss for D(C) and G:
    - $L_{\text{critic}} = \mathbb{E}_{x \sim P_r}[f_w(x)] - \mathbb{E}_{z \sim P_z}[f_w(G(z))]$
    - $L_G = -\mathbb{E}_{z \sim P_z}[f_w(G(z))]$

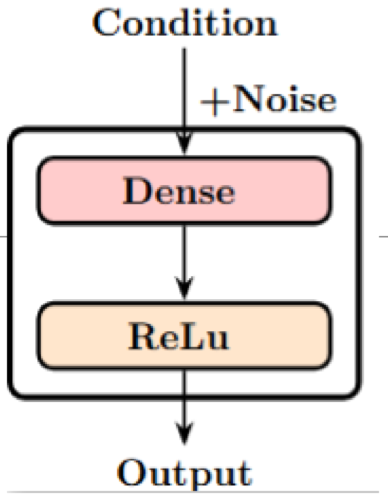
K-Lipschitz Constraint:

$$|f(x_1) - f(x_2)| \leq K|x_1 - x_2|$$

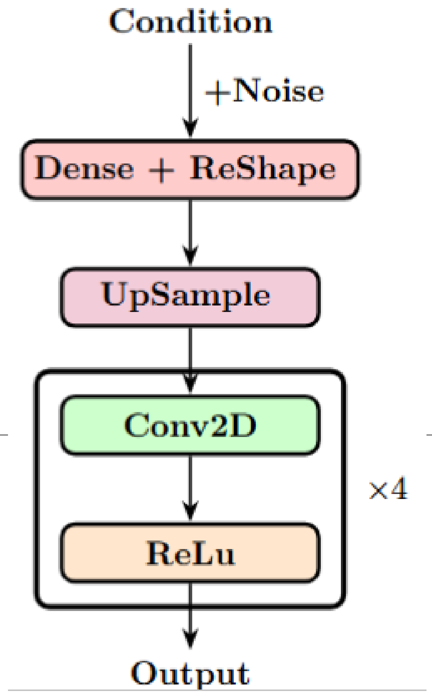


EMC simulation with ML - GAN

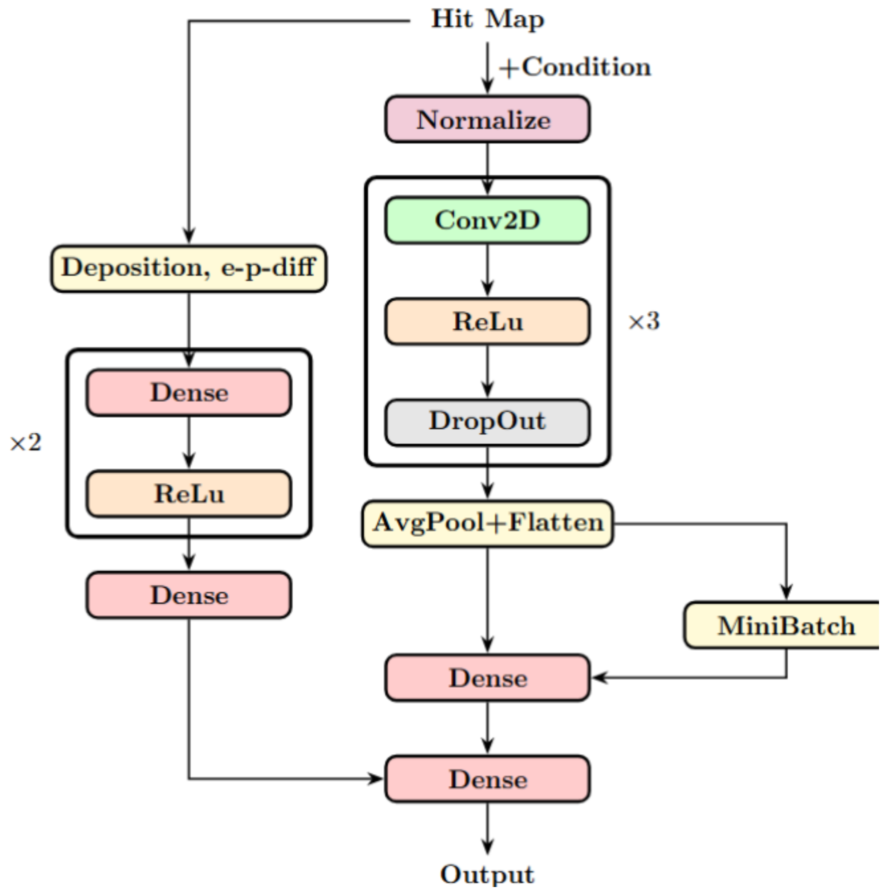
- Pre-gen, Generator, Discriminator



Pre-gen



Generator



Discriminator