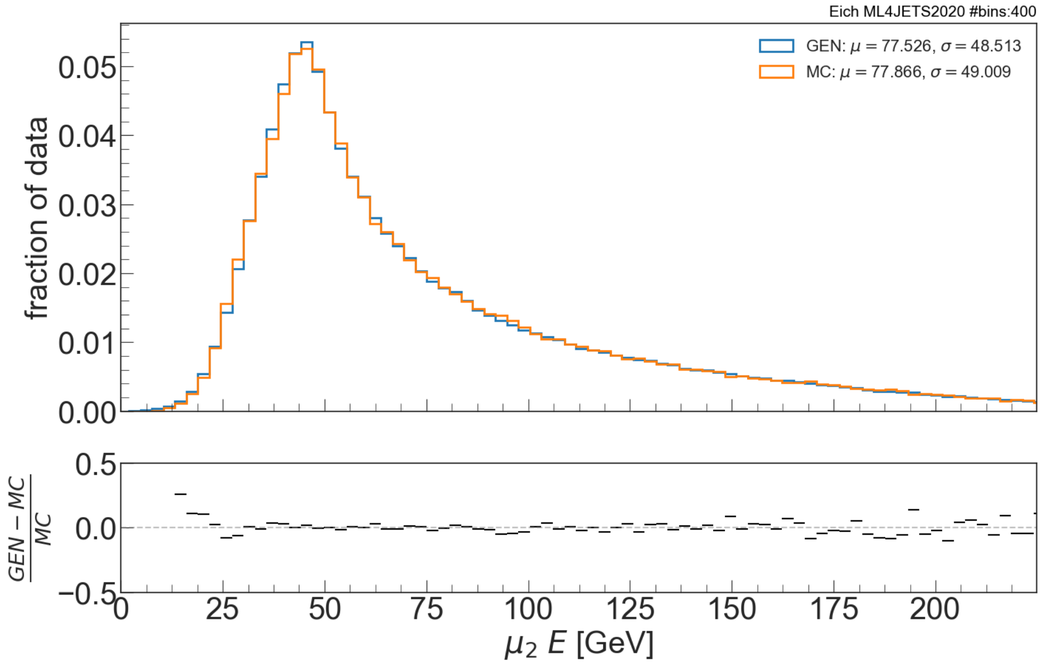
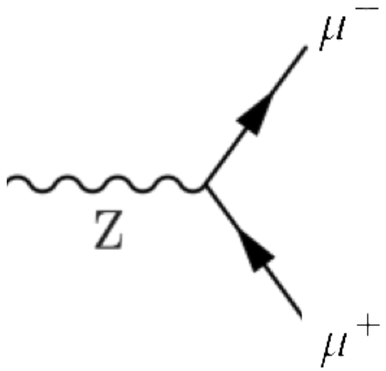
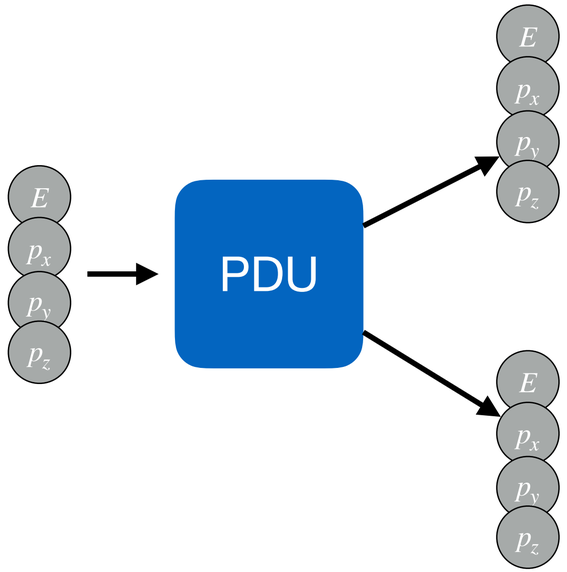






A generator cell for LHC event GANs: Particle Decay Unit

Niclas Eich, Martin Erdmann, Benjamin Fischer
RWTH Aachen University

ML4Jets2020

15th January 2020

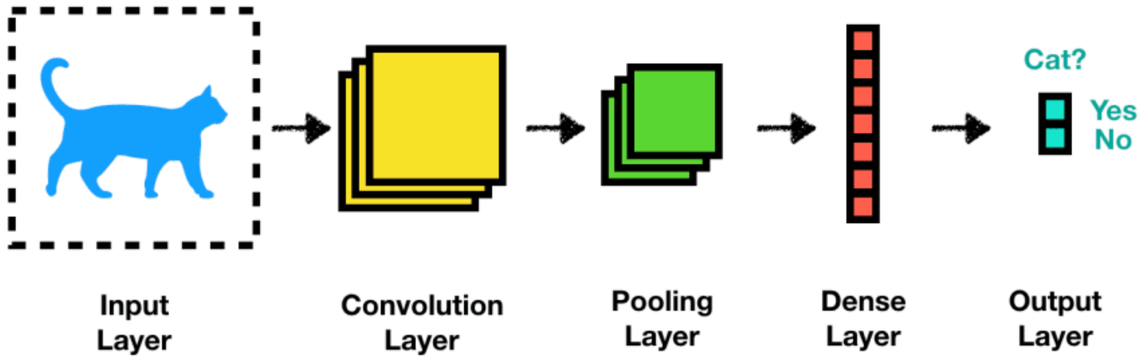
RWTH AACHEN UNIVERSITY

Our Goal

Computer Vision (CV)

Convolutions are the fundamental operations of image recognition

- Accelerated CV in deep learning
- Many successful applications in physics as well

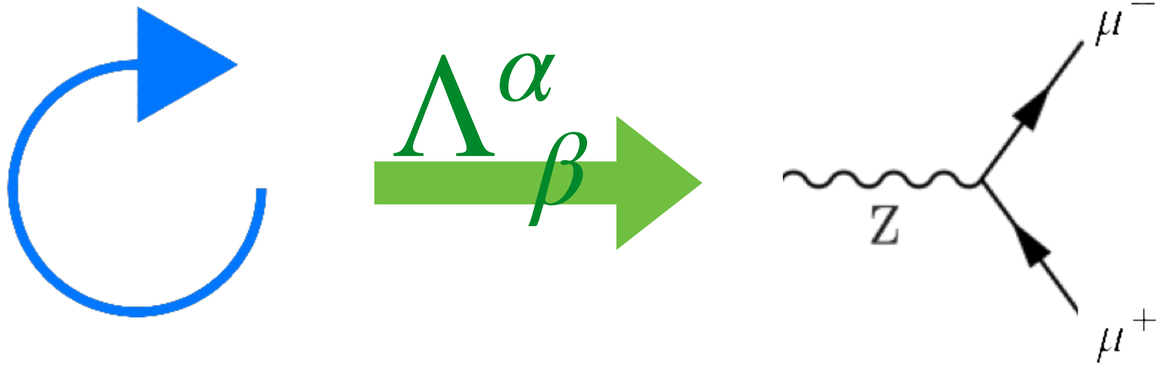


Physics

The **Lorentz group**, **boosts** and **rotations**, are the fundamental transformations of particle physics

➜ Can a generating network be built with these transformations?

➜ Study this at a simple process $Z \to \mu^+\mu^-$



Generation of a two-body decay

1. **Rest frame** of decaying particle

$$Z = \begin{pmatrix} E \\ p_x \\ p_y \\ p_z \end{pmatrix} \rightarrow \begin{pmatrix} M \\ 0 \\ 0 \\ 0 \end{pmatrix}$$

2. **Energy**- and **momentum-conservation** (back to back)

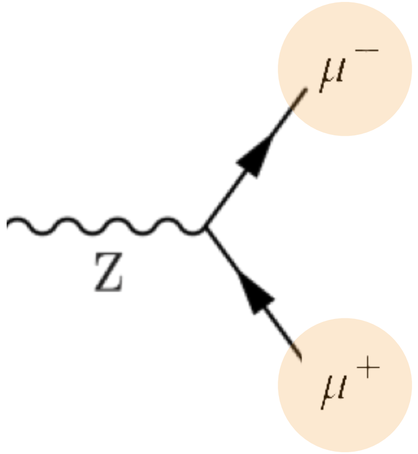
$$p_1 = \frac{\sqrt{\left(M^2 + m_1^2 - m_2^2\right)^2 - 4M^2 m_1^2}}{2M} \qquad l_i = \begin{pmatrix} \sqrt{m_i^2 + p_i^2} \\ \pm\vec{p_i} \end{pmatrix}$$

3. Pick **random orientation** isotropically

$$l_i = \begin{pmatrix} E_i \\ cos(\theta)sin(\phi)p_i \\ sin(\theta)sin(\phi)p_i \\ cos(\phi)p_i \end{pmatrix}$$

4. **Boost** into lab frame of mother particle

$$l_i^{\alpha'} = \Lambda_{\alpha}^{\alpha'} l_i^{\alpha}$$



Generation of a two-body decay

Niclas Eich - ML4Jets2020

1. **Rest frame** of decaying particle

$$Z = \begin{pmatrix} E \\ p_x \\ p_y \\ p_z \end{pmatrix} \rightarrow \begin{pmatrix} M \\ 0 \\ 0 \\ 0 \end{pmatrix}$$

2. **Energy**- and **momentum-conservation** (back to back)

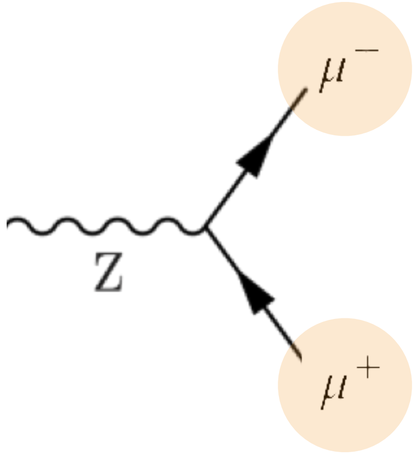
$$p_1 = \frac{\sqrt{\left(M^2 + m_1^2 - m_2^2\right)^2 - 4M^2 m_1^2}}{2M} \qquad l_i = \begin{pmatrix} \sqrt{m_i^2 + p_i^2} \\ \pm\vec{p_i} \end{pmatrix}$$

3. Pick **random orientation** isotropically

$$l_i = \begin{pmatrix} E_i \\ cos(\theta)sin(\phi)p_i \\ sin(\theta)sin(\phi)p_i \\ cos(\phi)p_i \end{pmatrix}$$

4. **Boost** into lab frame of mother particle

$$l_i^{\alpha'} = \Lambda_{\alpha}^{\alpha'} l_i^{\alpha}$$

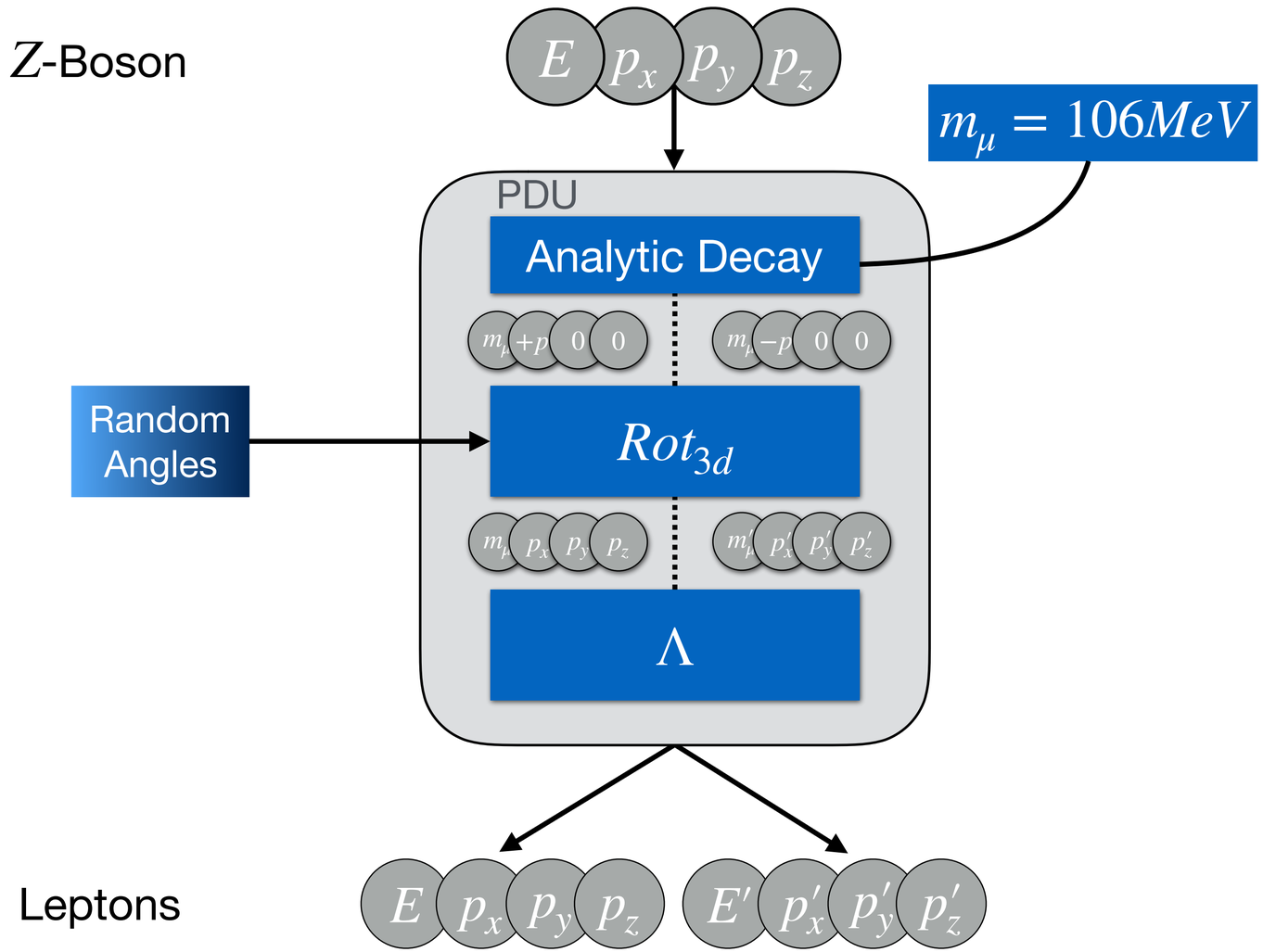


Embed this recipe inside a neural network

4 Particle Decay Unit (PDU)

Full Generator:

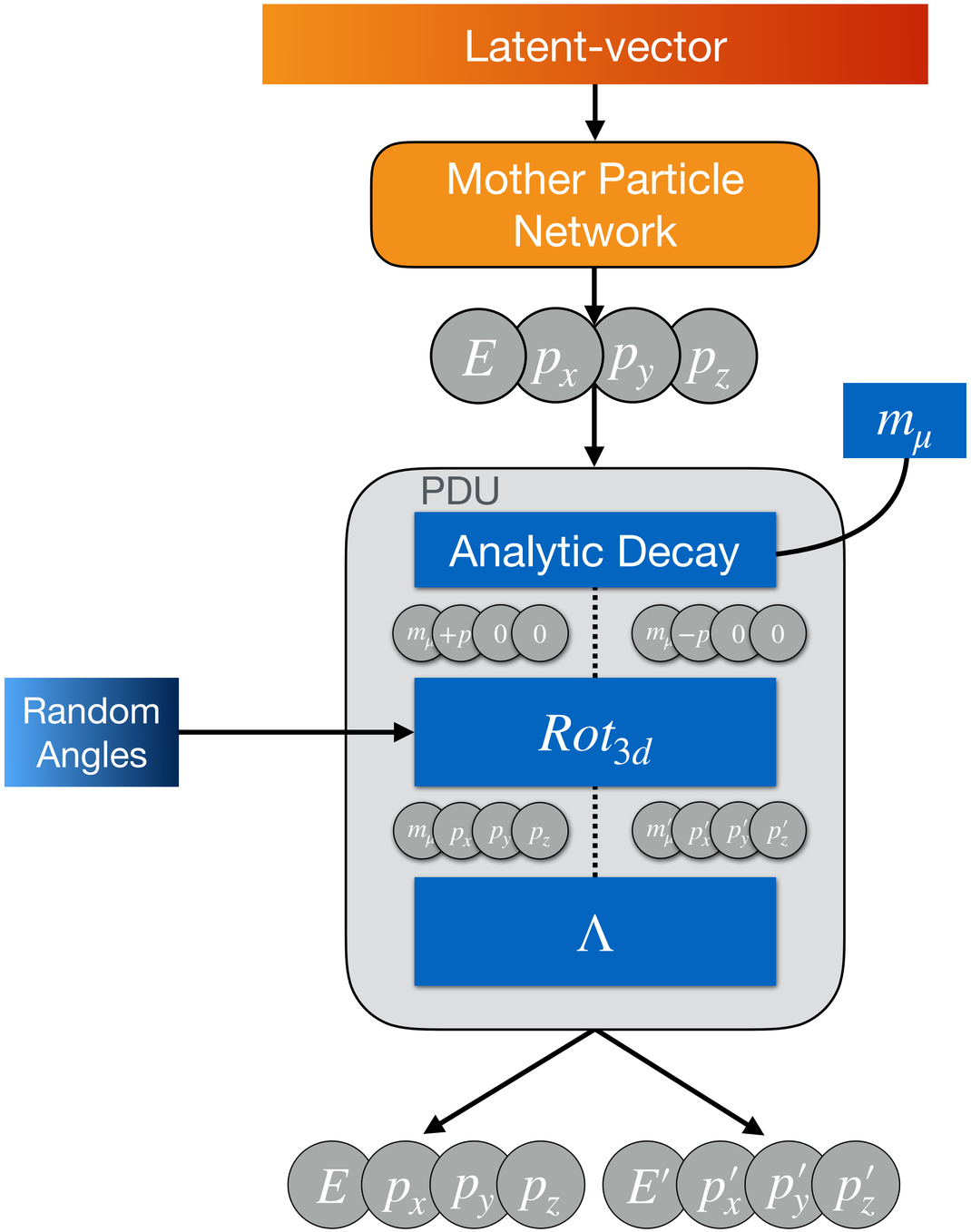
- ☐ Variance
- ☐ Detector acceptance
- ☐ Detector effects & higher orders



5 Mother particle modelling

Full Generator:

- ✓ Variance
- ☐ Detector acceptance
- ☐ Detector effects & higher orders



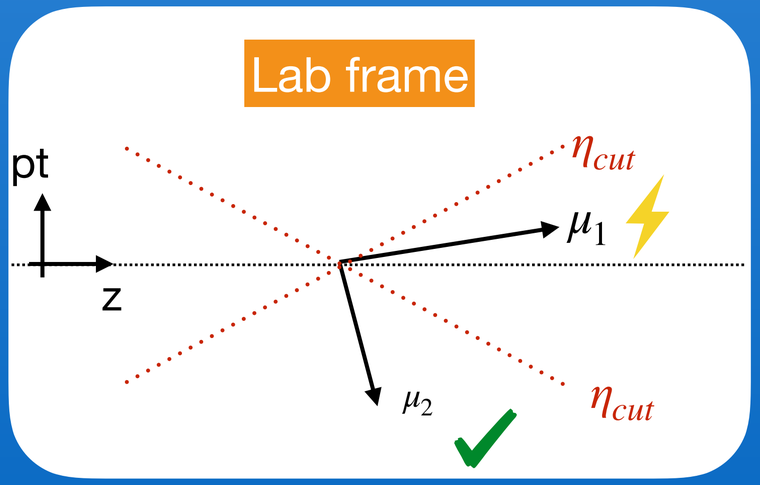
RWTH AACHEN UNIVERSITY

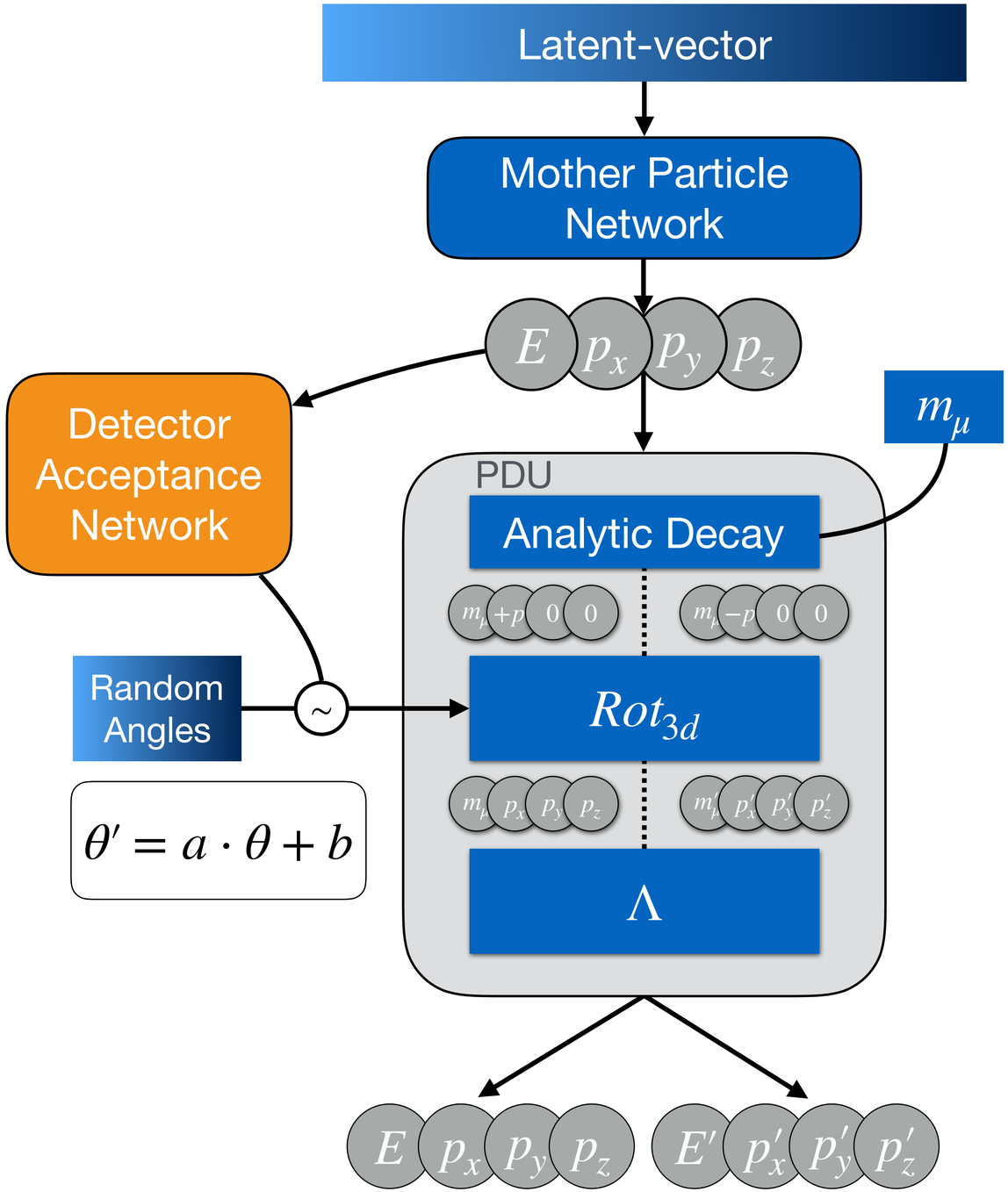
Parametrised Rotation

Full Generator:

- ✓ Variance
- ✓ Detector acceptance
- ☐ Detector effects & higher orders

Decay products need to be in detector acceptance





4-vector refinement

Full Generator:

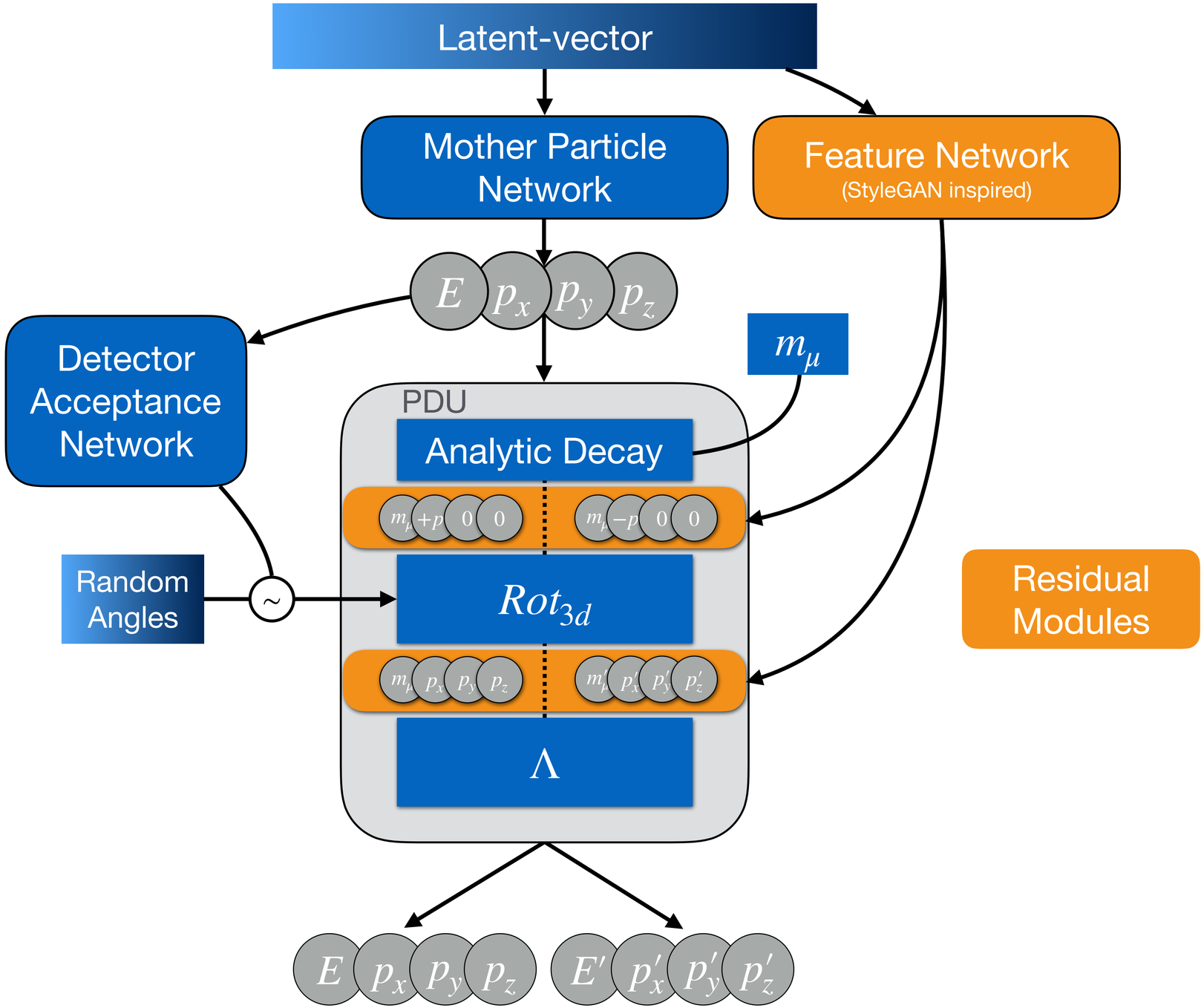
- ✓ Variance
- ✓ Detector acceptance
- ✓ Detector effects & higher orders

Feature Network preprocesses noise [arXiv:1812.04948]

Additional residual modules further model effects on the 4-vectors

$x' = x + f_{Res}(x)$

➜ only small changes



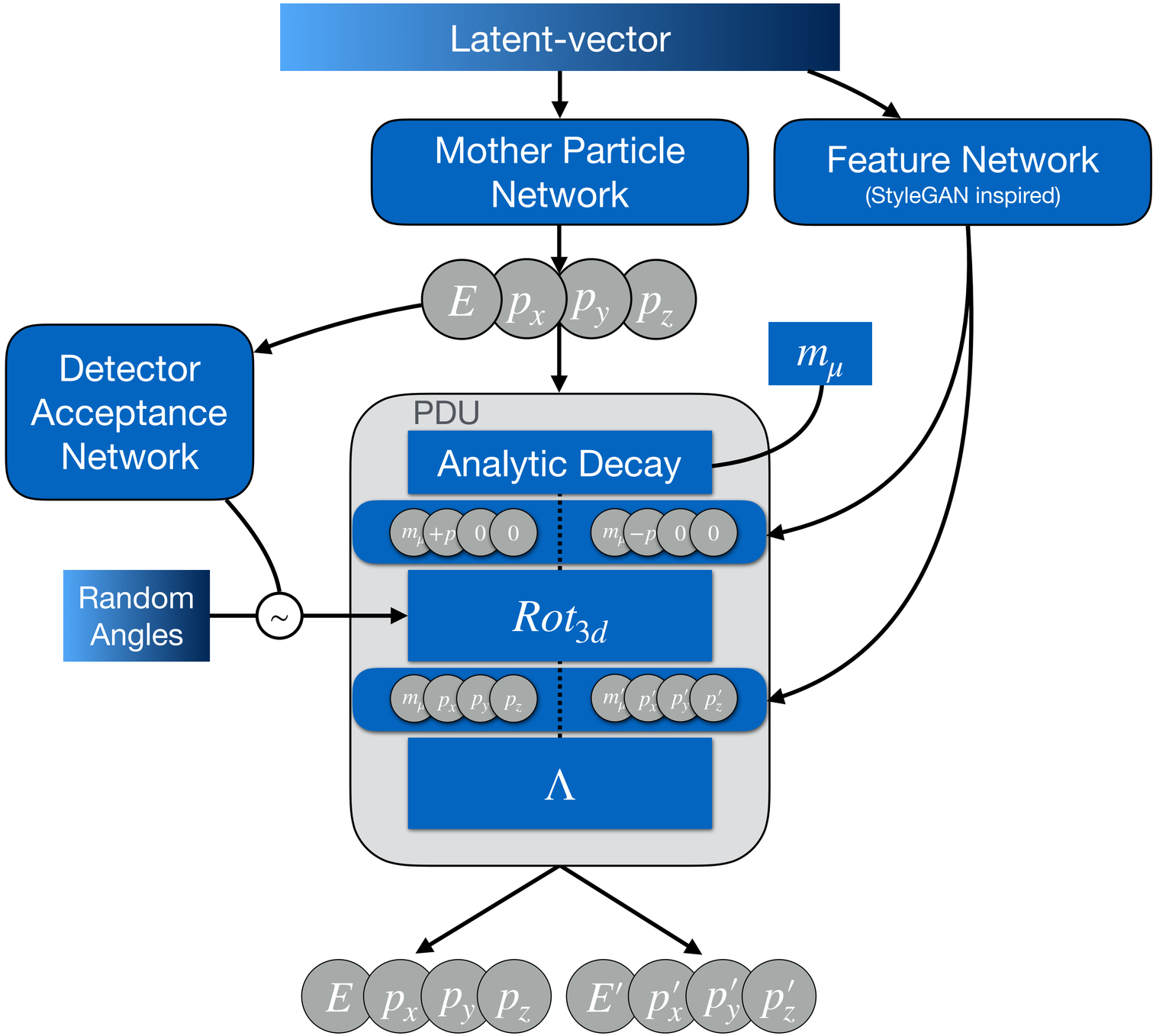
RWTH AACHEN UNIVERSITY

Generator Network

Full Generator:

- ✓ Variance
- ✓ Detector acceptance
- ✓ Detector effects & higher orders

Every component has a physical interpretation to it



RWTH AACHEN UNIVERSITY

WGAN Training

WGAN Setup

Wasserstein Loss

$$\mathrm{EMD}\left(P_r, P_\theta\right) = \sup_{\|f\|_{L\leq 1}} \mathbb{E}_{x\sim P_r} f(x) - \mathbb{E}_{x\sim P_\theta} f(x)$$

Gradient penalty for Lipschitz-constraint [arXiv:1704.00028]

Schedule:

1/10 Generator/Critic updates

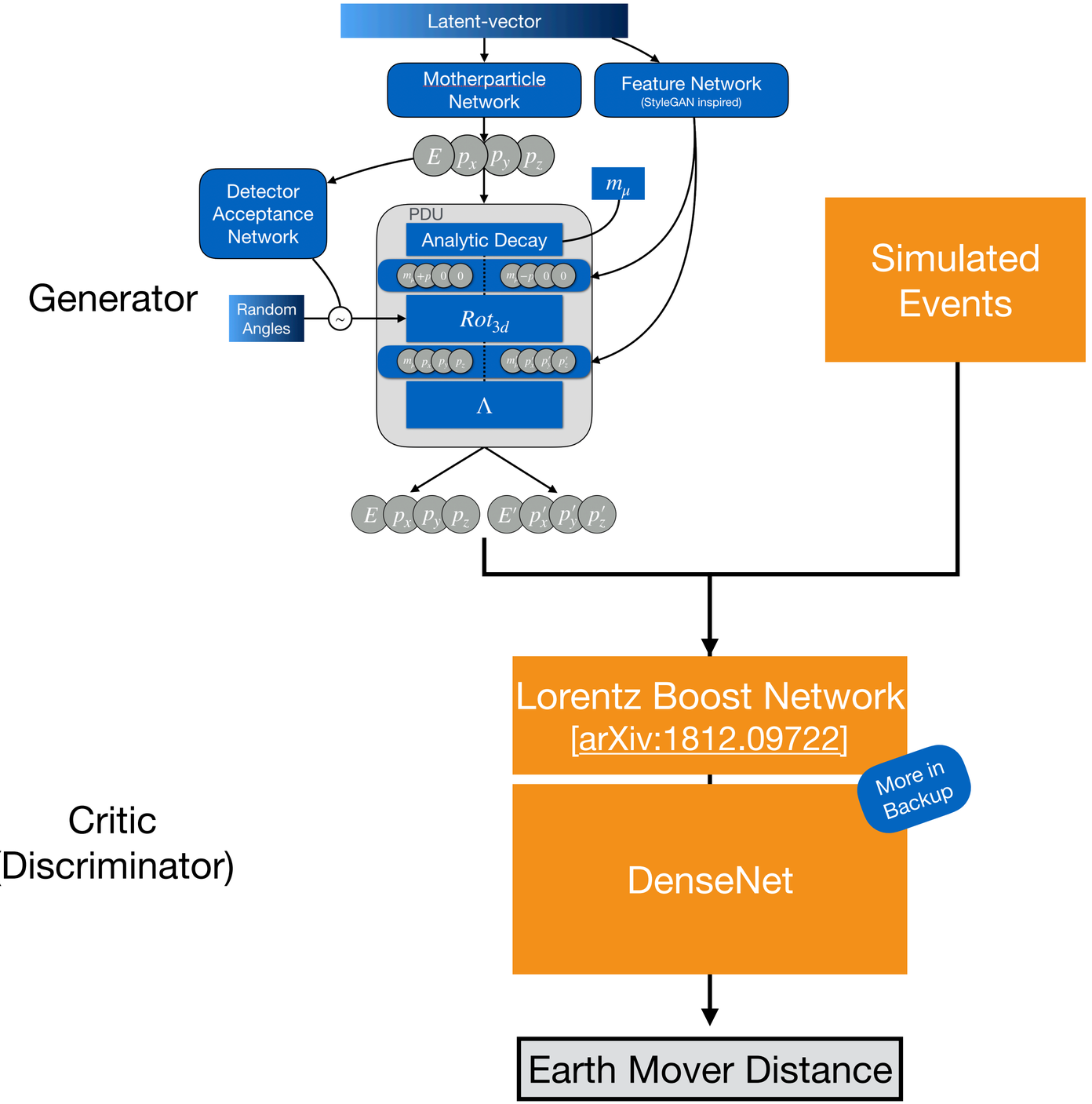
Simulation:

$Z \to \mu^+\mu^-$

450k events

MadGraph5_aMC@NLO

Delphes



10 Critic Network

Pre-Stage:

Lorentz Boost Network

[arXiv:1812.09722]

Can combine particles to mother particle candidate

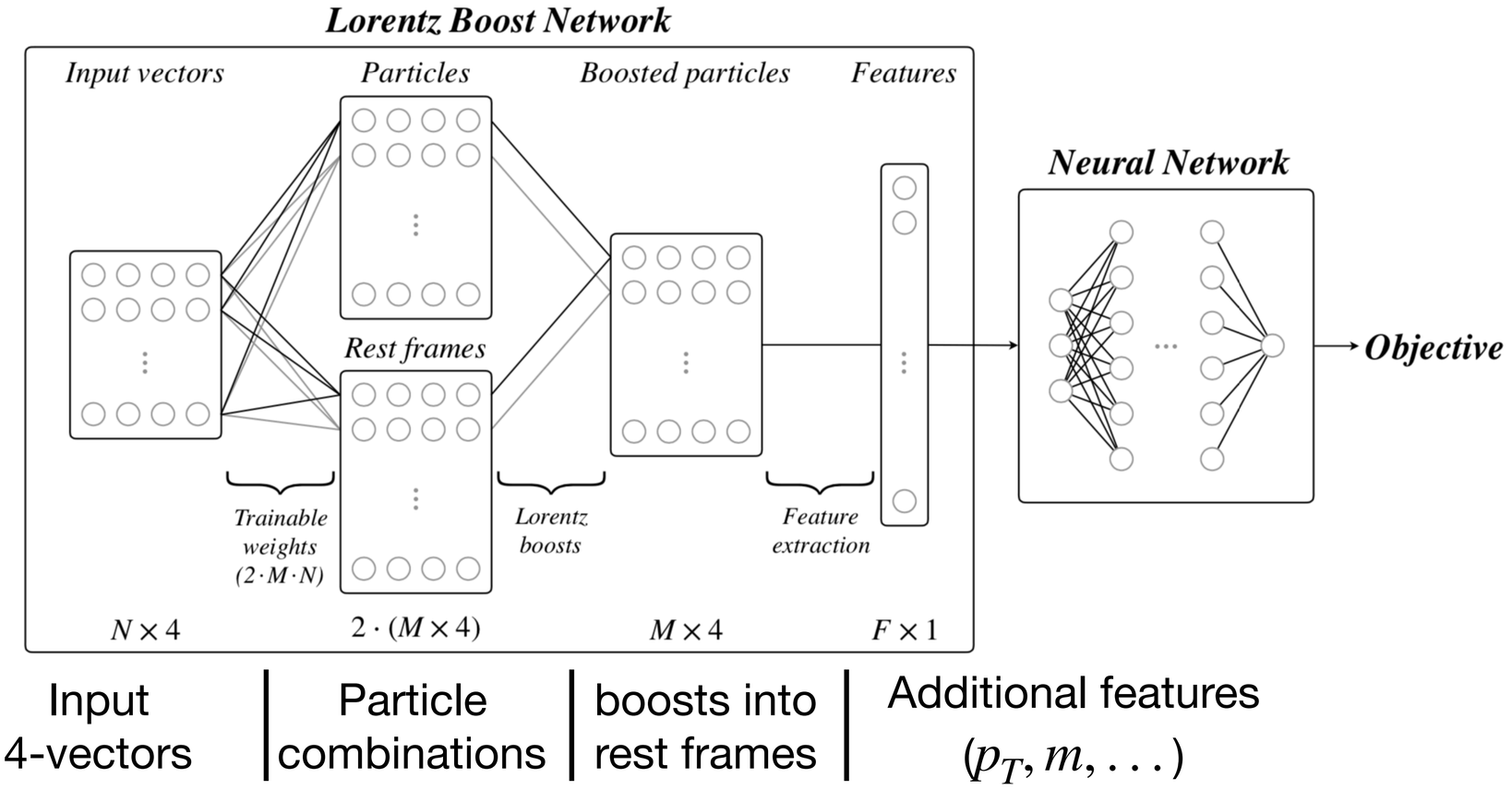
$$1.0 \cdot \vec{p_1} + 1.0 \cdot \vec{p_2} + 0.0 \cdot \vec{p_i} = \vec{p_\alpha} \rightarrow \vec{p'_\alpha}$$

DNN-Part:

DenseNet Architecture (shortcuts)

3×2 Layers, 280 Nodes,

leaky ReLU activation

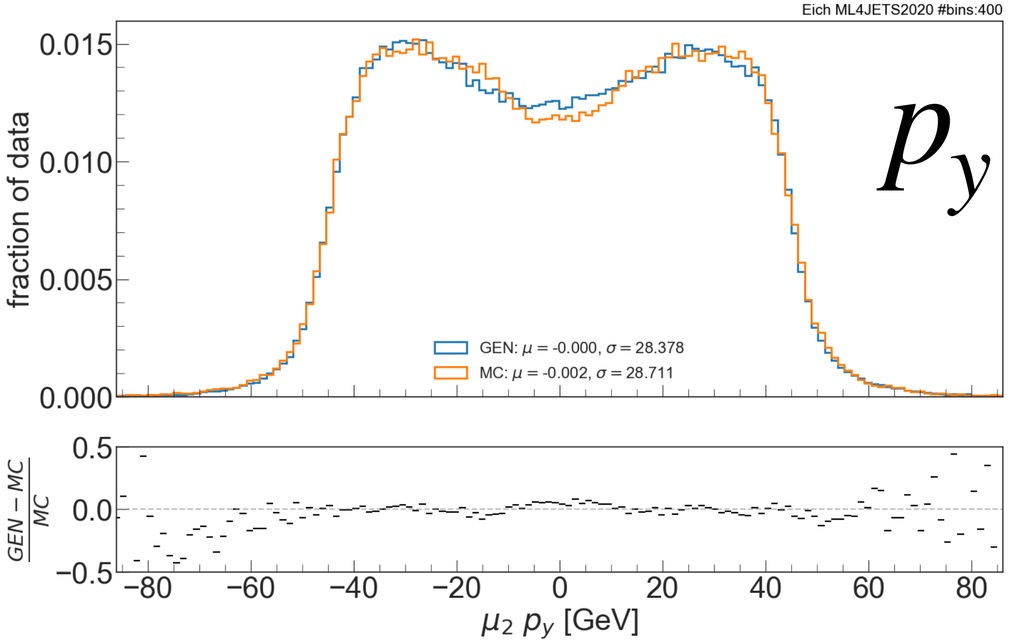


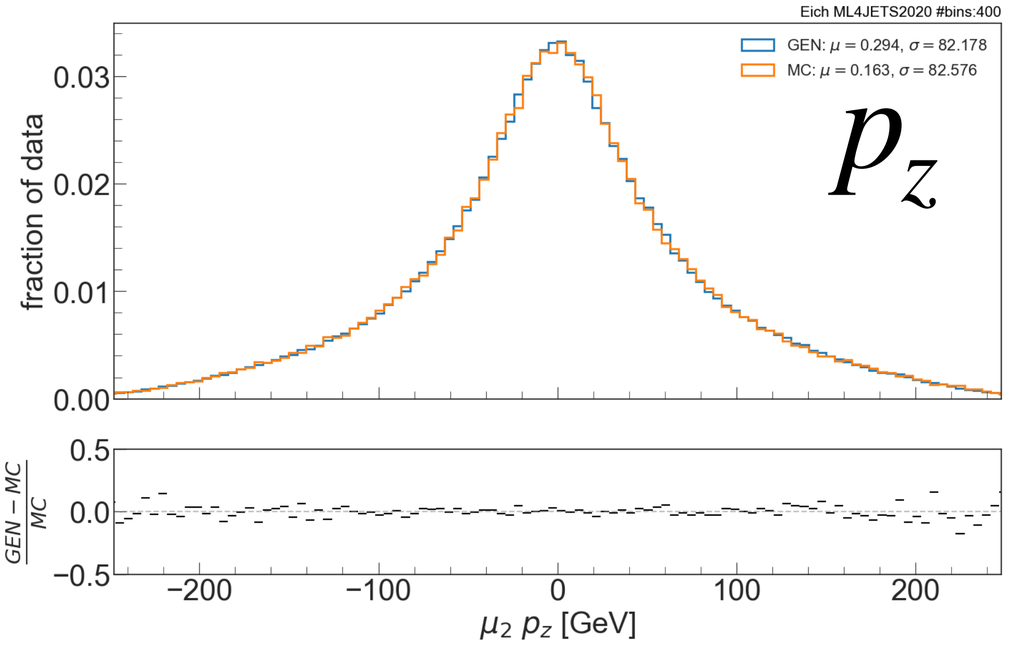
Input 4-vectors	Particle combinations	boosts into rest frames	Additional features $(p_T, m, \ldots)$

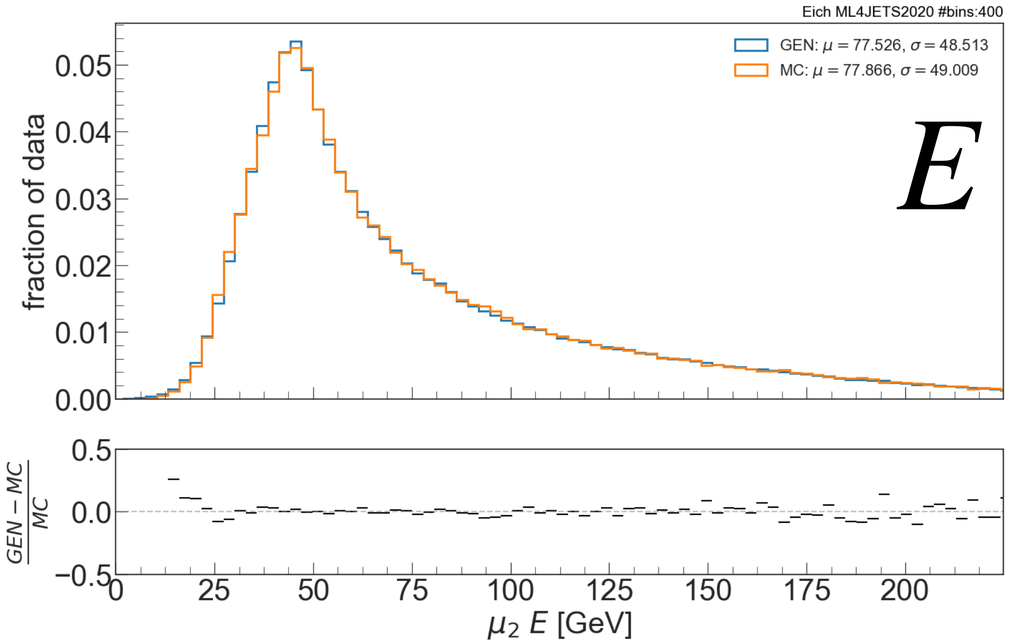
RWTH AACHEN UNIVERSITY

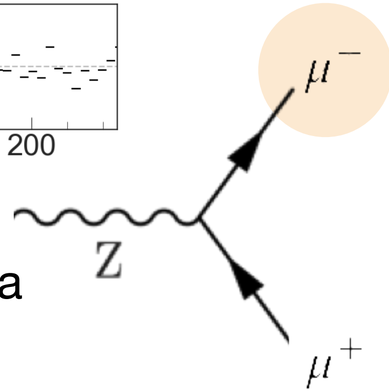
Basic Kinematics

Niclas Eich - ML4Jets2020



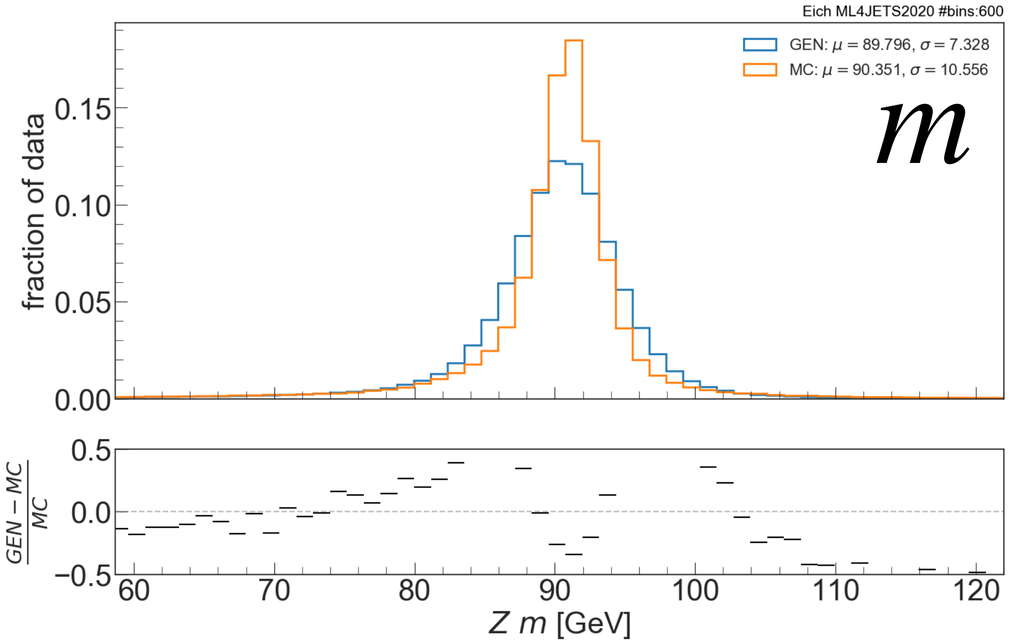


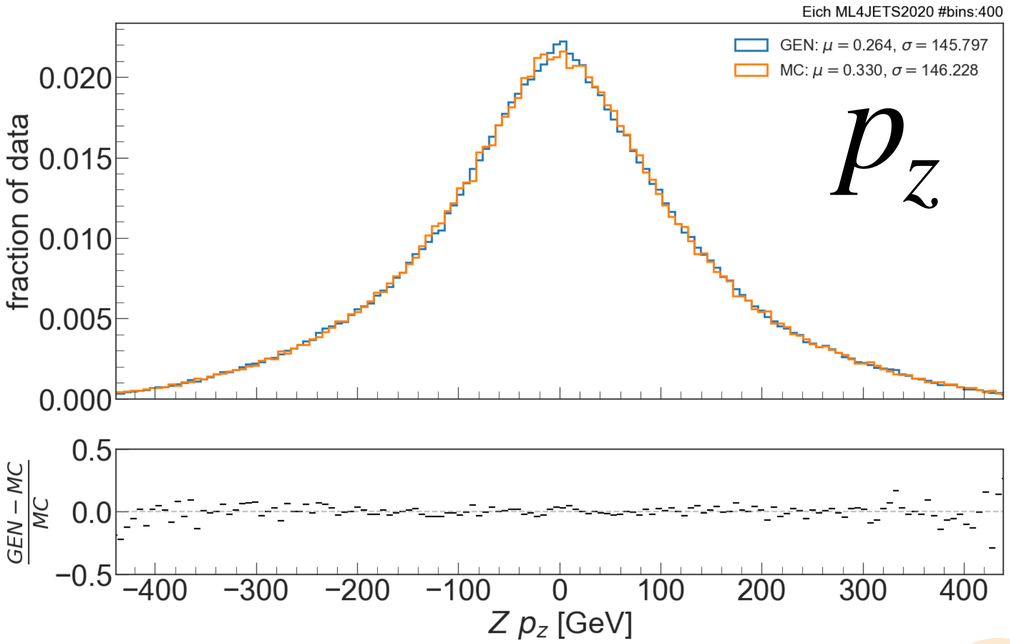


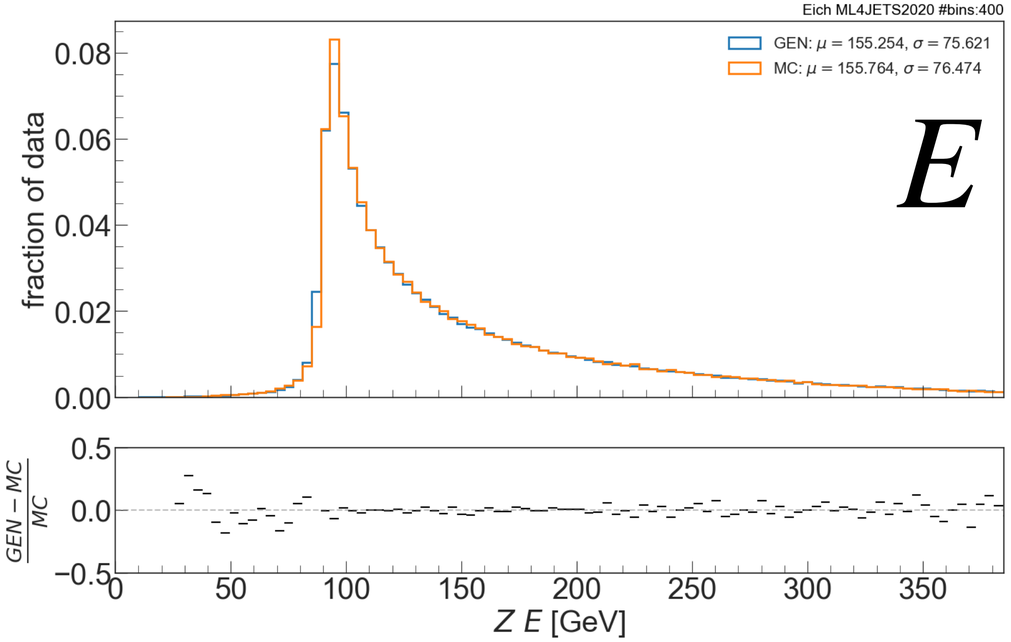


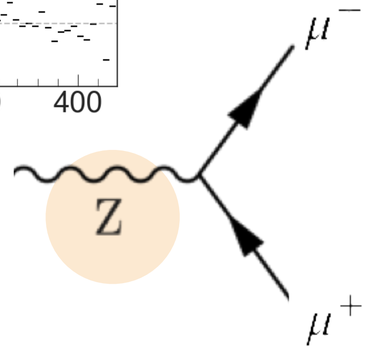
- Good agreement of generated data to simulation
- Almost independently of high/low occupancy regions

Mother Particle Candidate









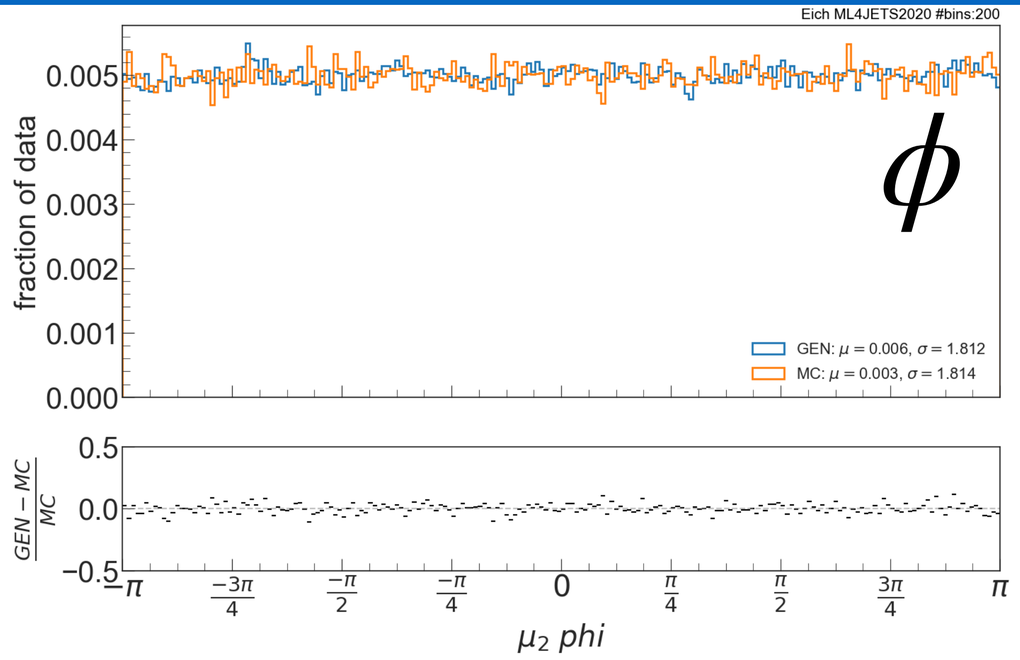
- Add muons to reconstruct Z

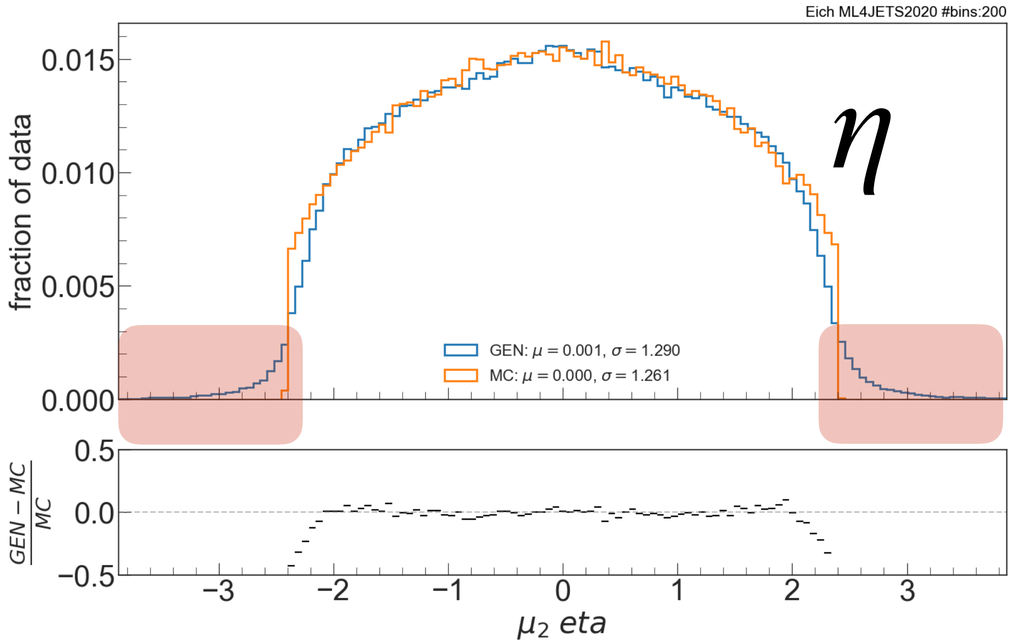
  $\overrightarrow{\mu_1} + \overrightarrow{\mu_2} = \overrightarrow{Z}_{Candidate}$

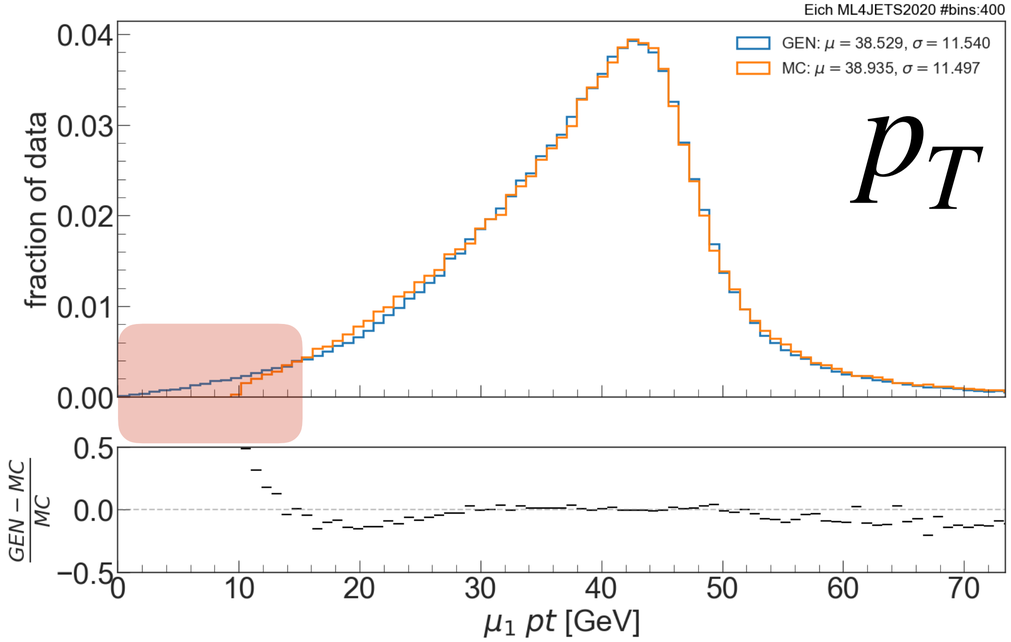
- Mass too much spread
- E and p_z good agreement

Detector Acceptance & Isotropy

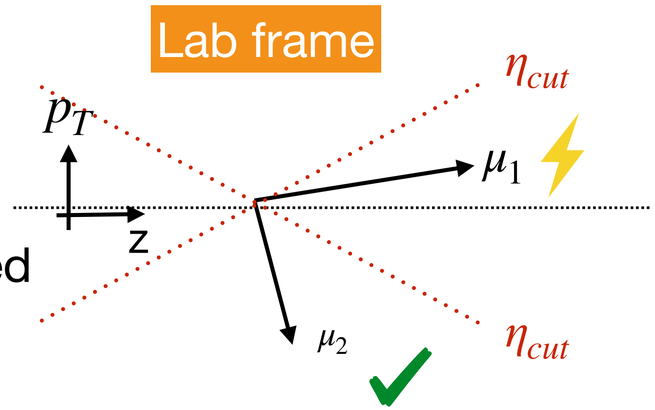
Niclas Eich - ML4Jets2020







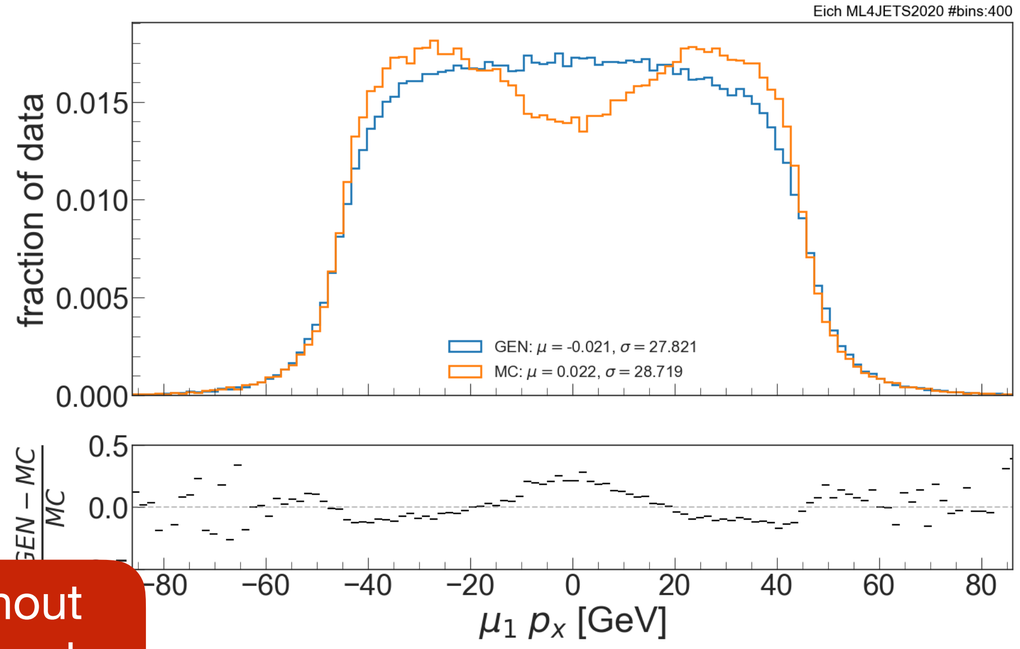
- Decay stays isotropic after residual modules (ϕ-Distribution)
- p_T and η are not modelled perfectly

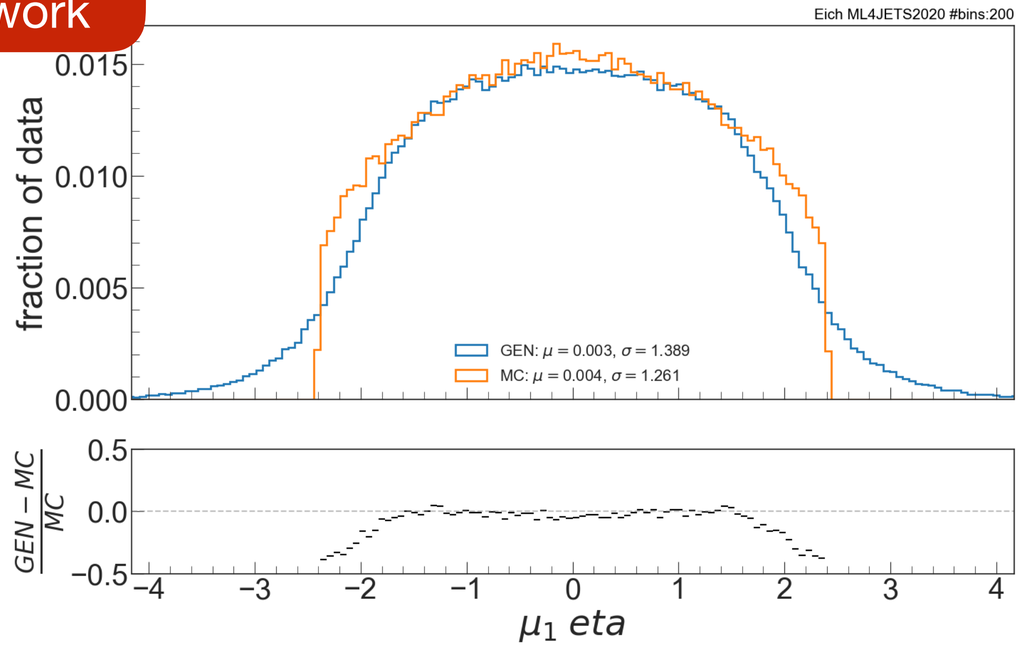


Detector Acceptance Network: Effects

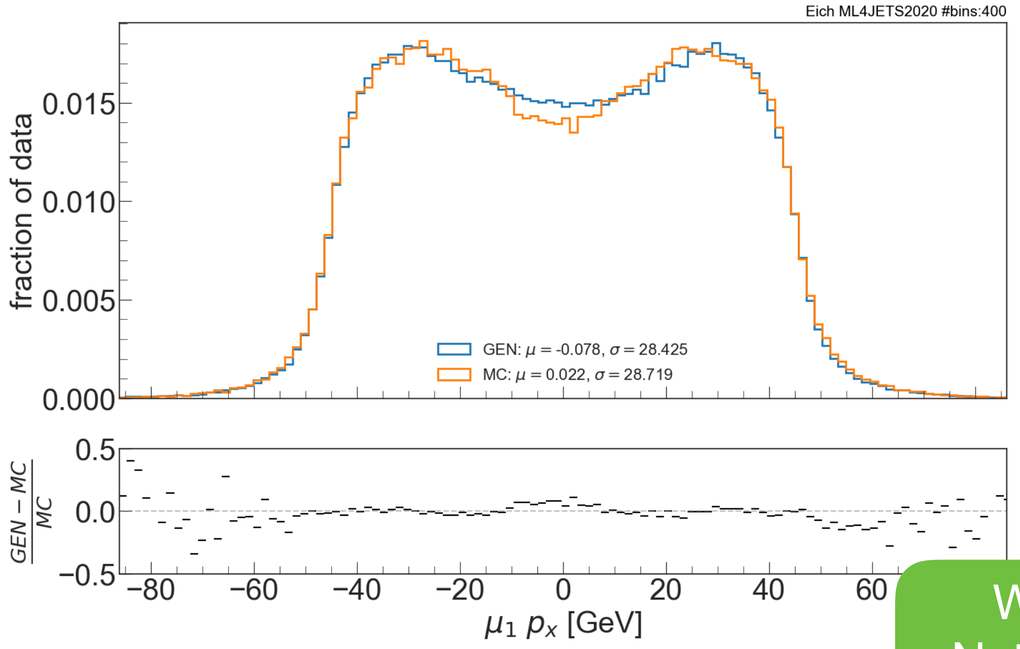
Niclas Eich - ML4Jets2020

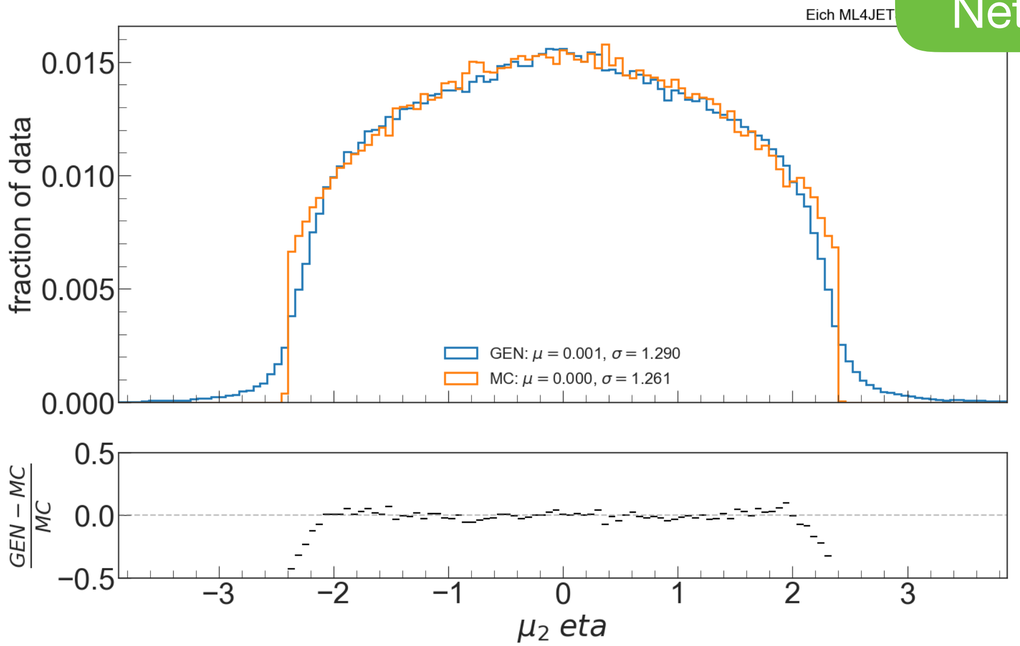
Without Network





With Network





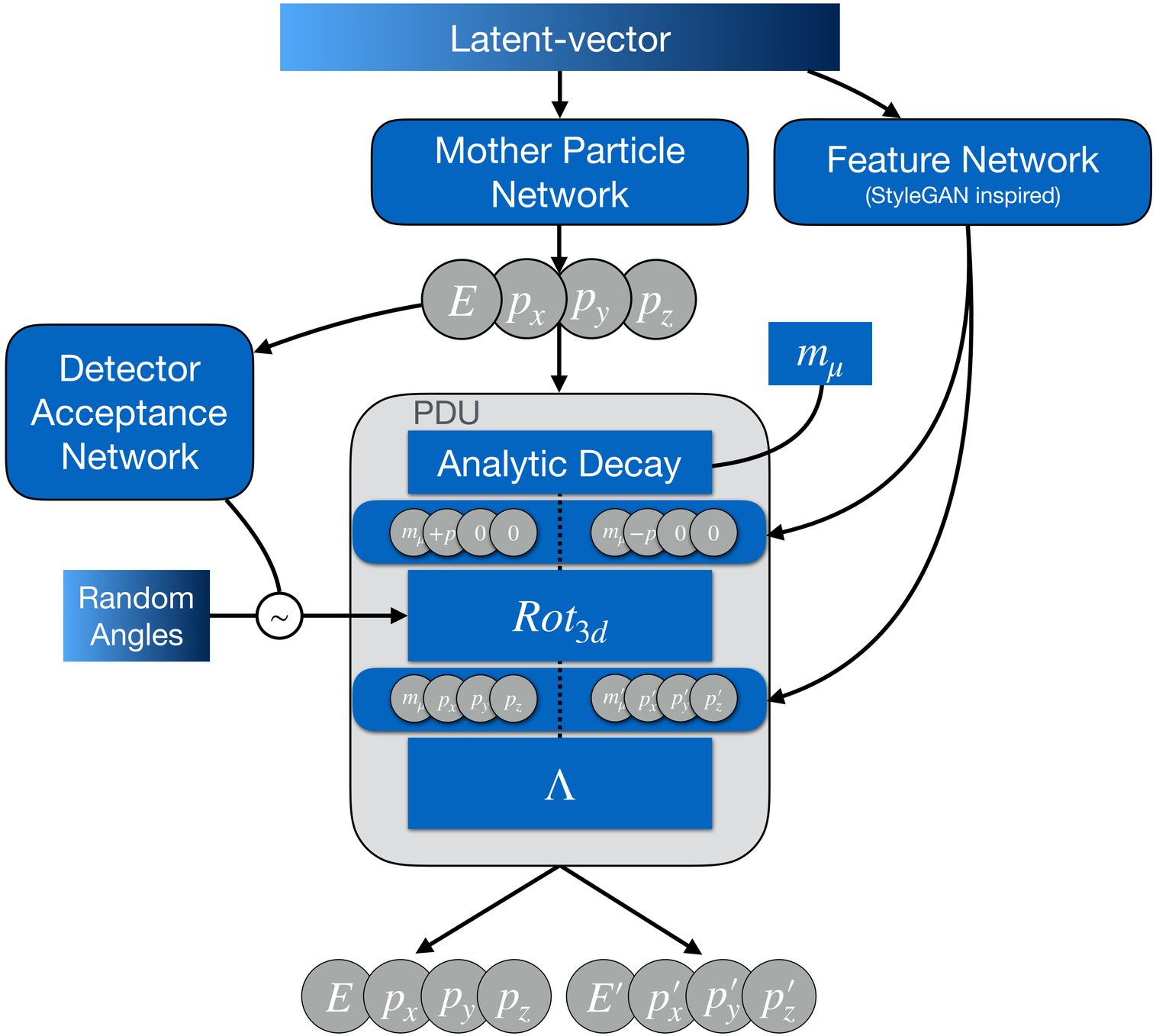
15 Network Structure

Network Size:

- Mother particle Network
  - 3 Layers
  - 600 Nodes
  - SeLU Activation
- Feature-Network
  - 3×3 Layers densely connected
  - 280 Nodes
  - ReLU Activation
- Detector Acceptance Network
  - 6 Layers
  - 280 Nodes
  - ReLU Activation
- Residual Modules
  - 2×3 Layers added on 4-vectors
  - 280 Nodes
  - ReLU Activation

Total:

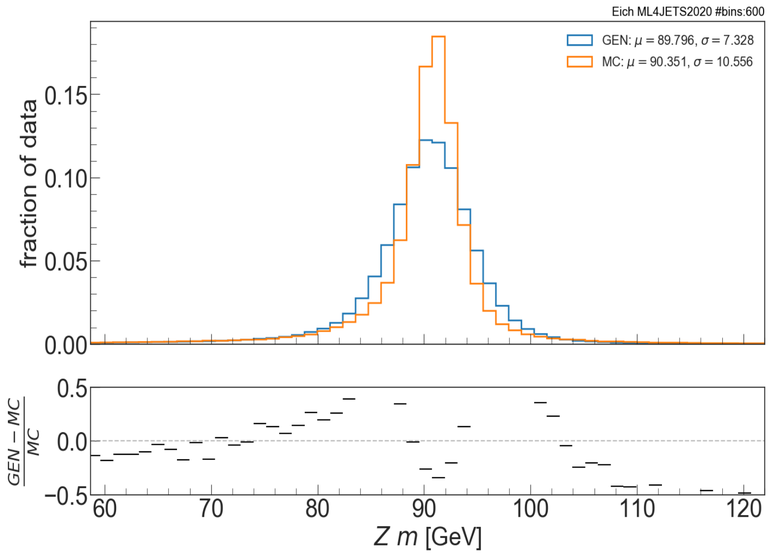
- 0.5M - **3.3M** Params
- 6 hours Training
  (GTX 1080)

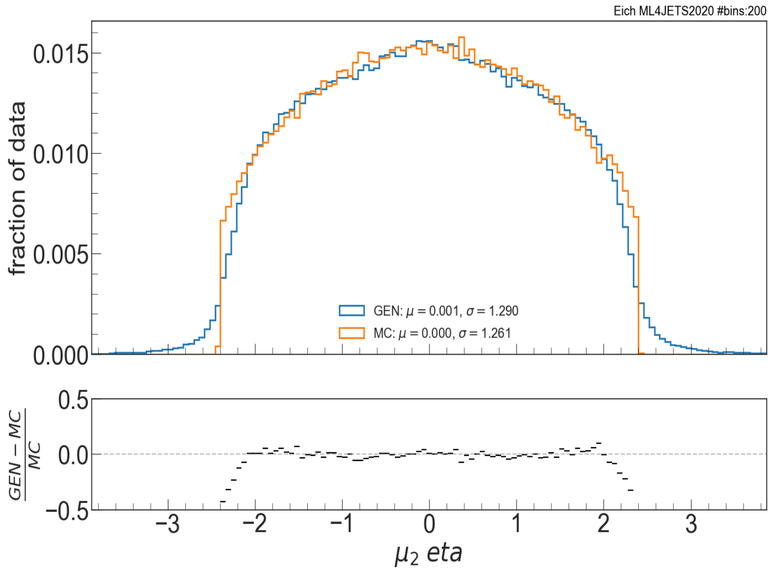


Challenges

Niclas Eich - ML4Jets2020

- Mass modelling

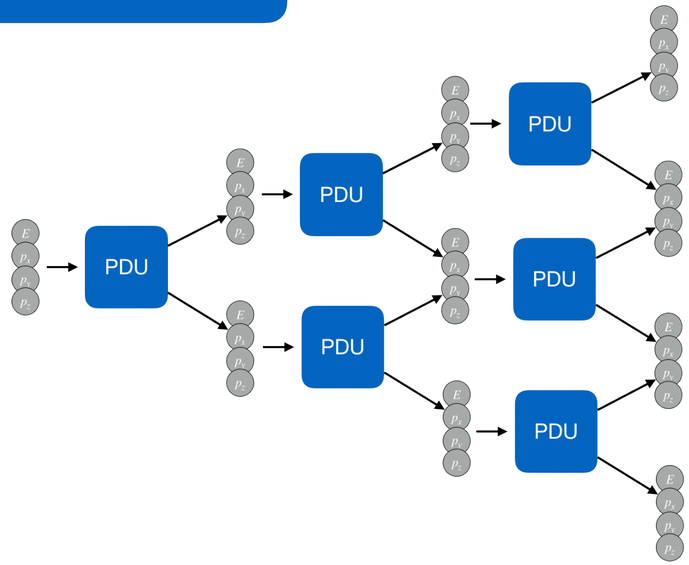




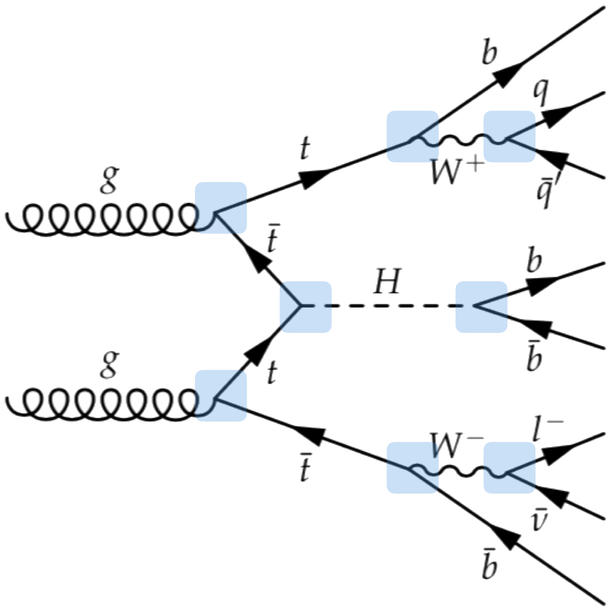
- Detector Acceptance, p_T/η cuts

- Free floating masses inside the PDU

$$m_\mu = f(\dots)$$



- Jets and more complex Feynman diagrams
  - $Z \to b\bar{b}$
  - QCD modelling



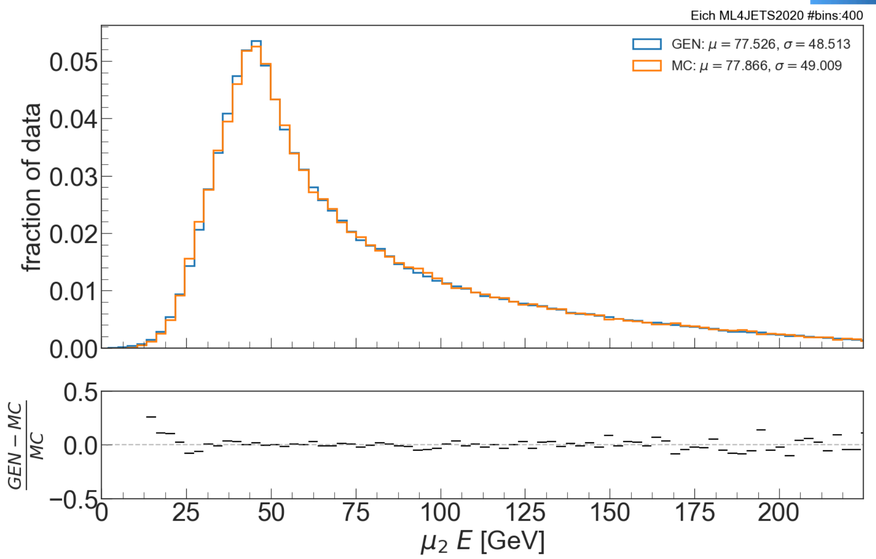
- Application on full simulation

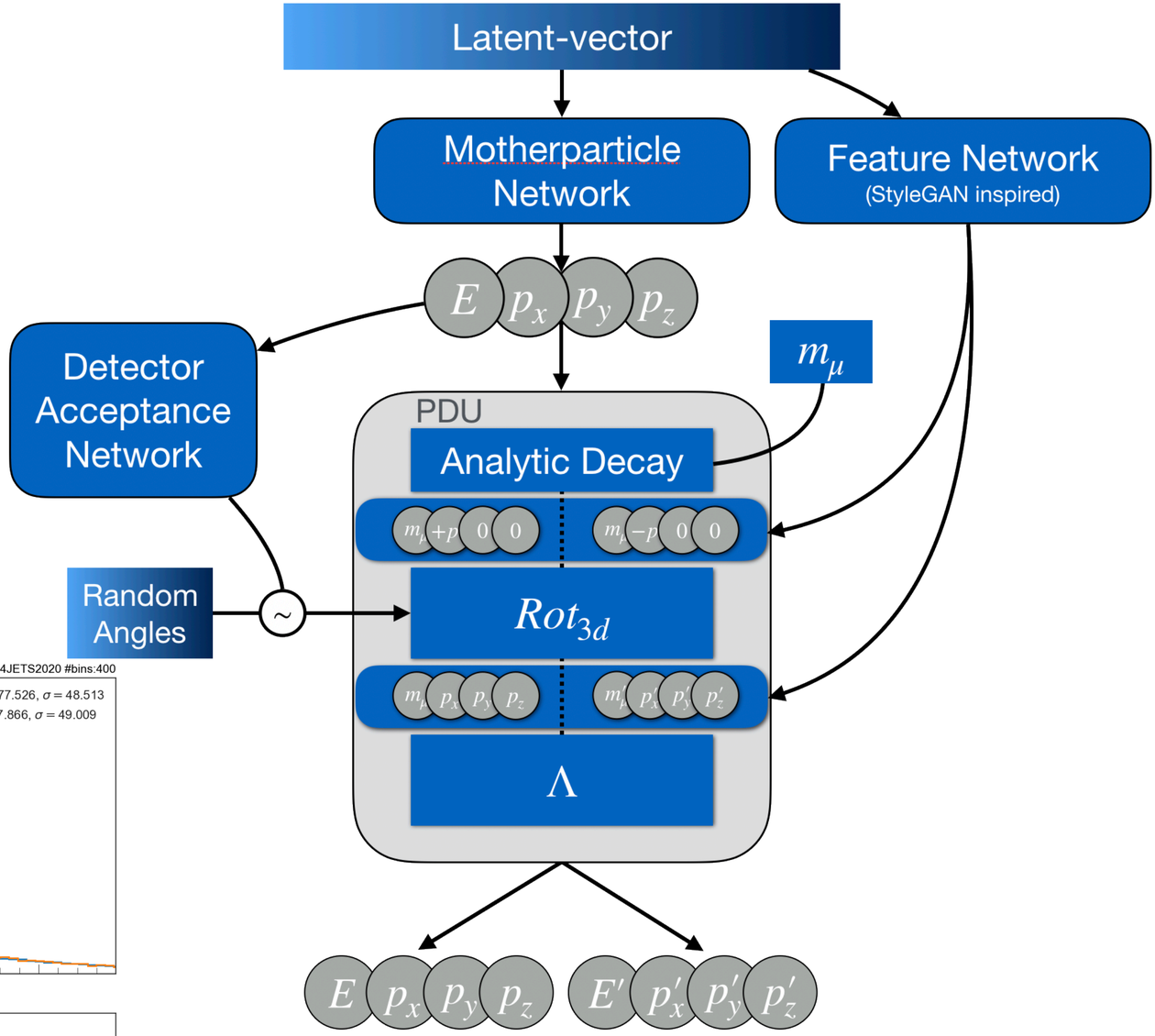
Summary

- Built Particle Decay Unit with:
  - Lorentz Boosts
  - Rotations

- Generator components have physical interpretation
  - No black-box

- Generated decay particles are in good agreement with simulation.
  Control plots of reconstructed mother particle and detector acceptance still to be improved
  (m, p_T, η)

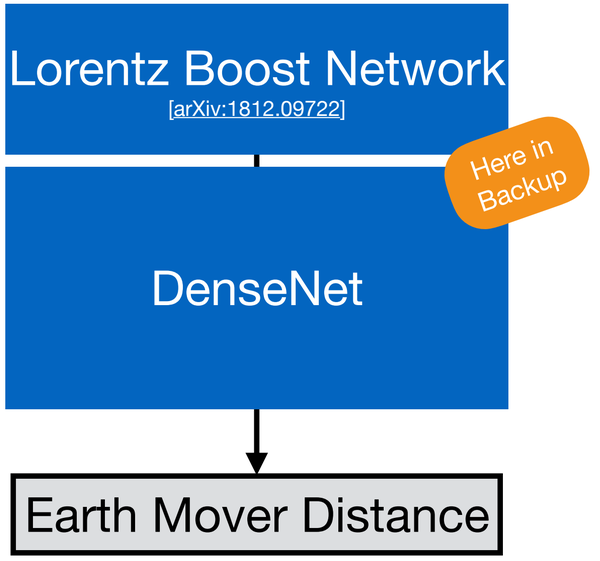




Backup

Critic Setup - Generation Time

- Additional computed features on 4-vecs: η, ϕ, m
- Pre-stage: Lorentz Boost Network [arXiv:1812.09722]
  - Autonomously combines particles and boosts them into learned rest frames
  - Especially useful for processes with more final state particles
- Main-Body
  - DenseNet architecture fully connected [arXiv:1608.06993]
    - 3×2 Layers, 280 Nodes, leaky ReLU activation
  - Minibatch-discrimination like layer [arXiv:1606.03498]
    - Computes mean and variance across batch as features



Generation Time

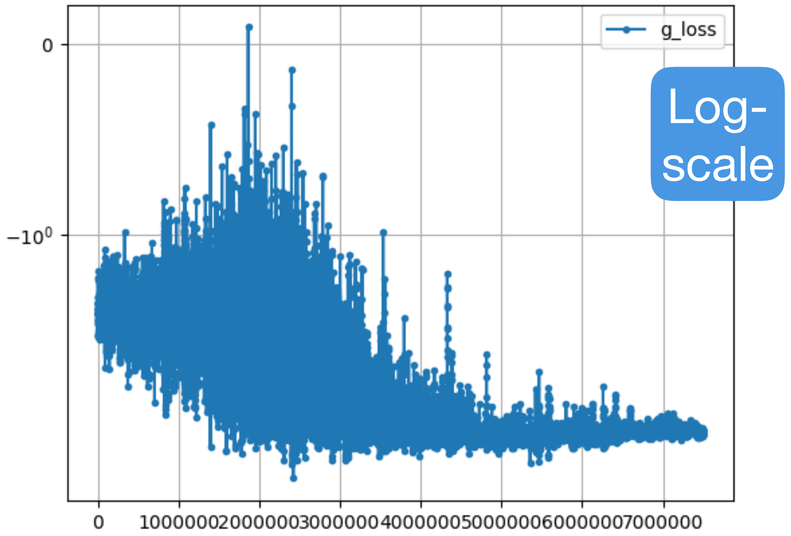
- **~100k Events/s** at batch-size ~1000
- GTX 1080

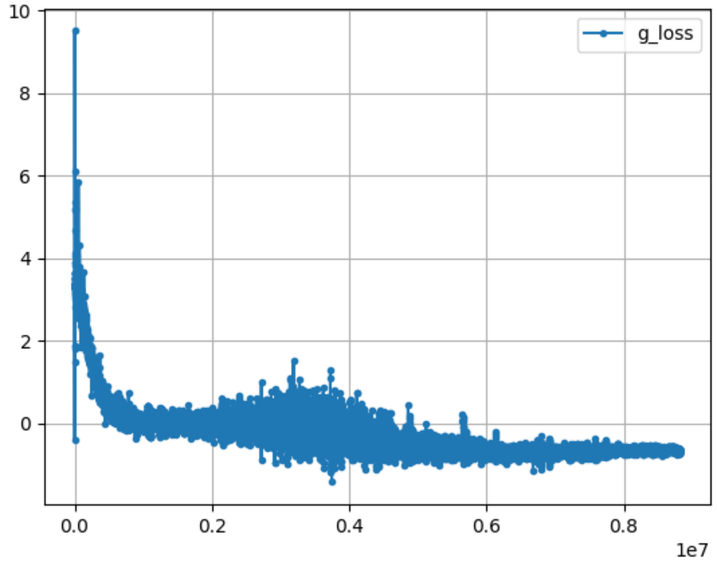
Training	Generator	Critic
Optimizer	Adam	Adam
Beta1/2	0.5/0.9	0.5/0.9
Learning Rate	$5 \cdot 10^{-4}$	$5 \cdot 10^{-4}$
LR-Decay	10^{-2}/1500 epochs	
Trained epochs	2000	
Gradient Penalty	1.0	

Training Loss Plots - no improvement in longer Training

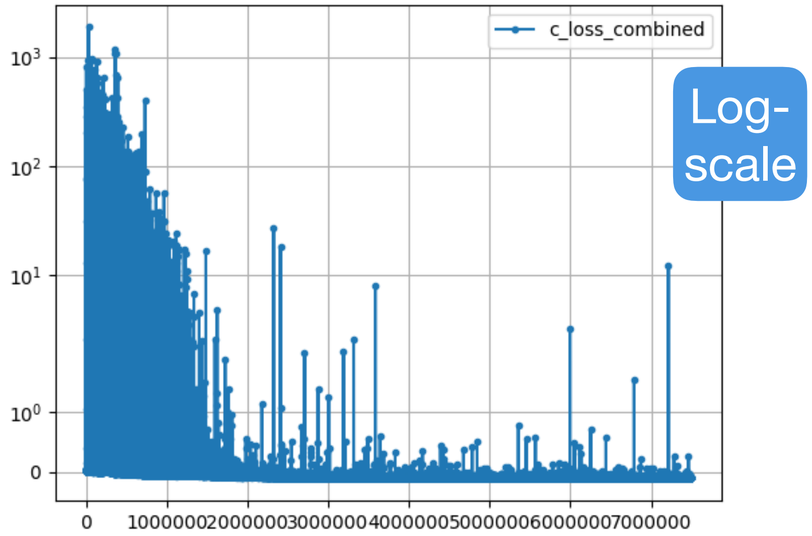
Niclas Eich - ML4Jets2020

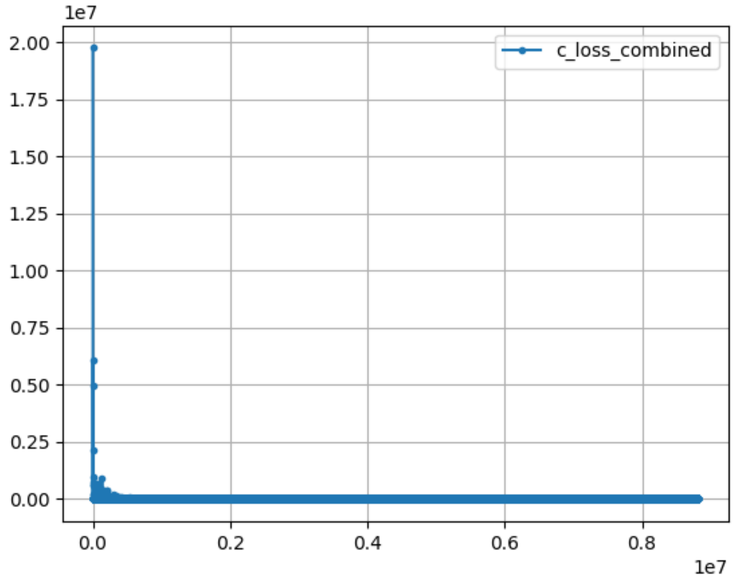
Generator



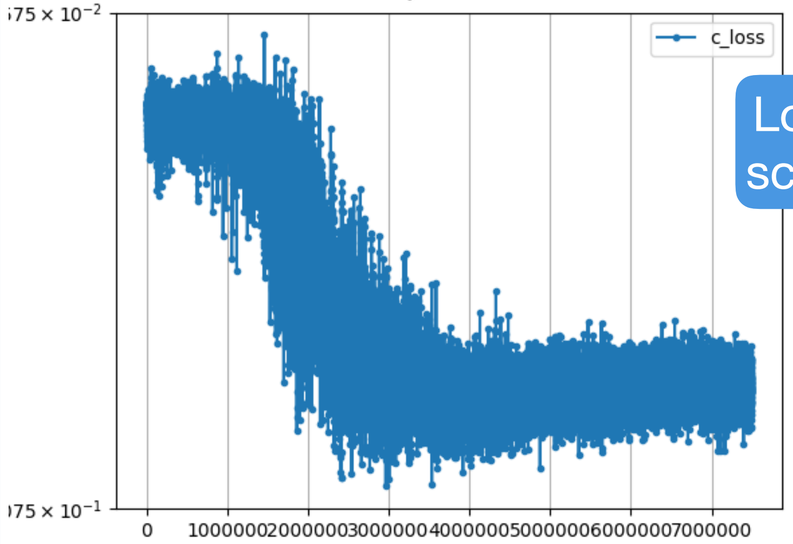


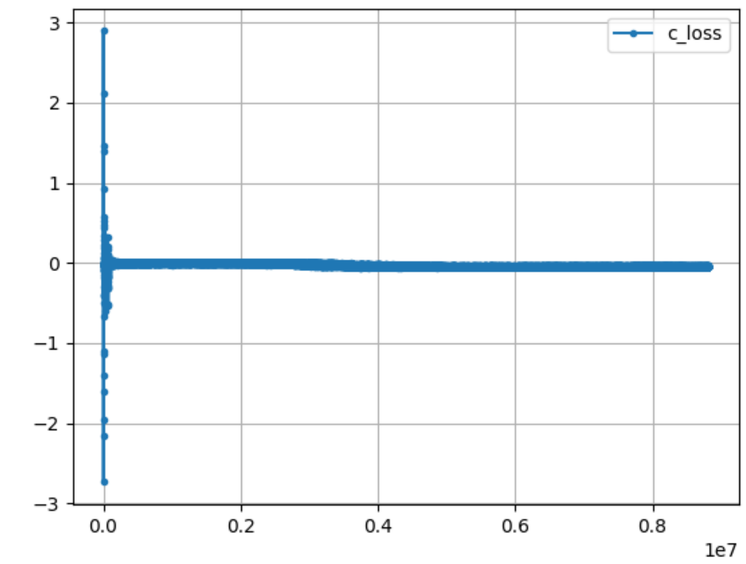
Full Critic



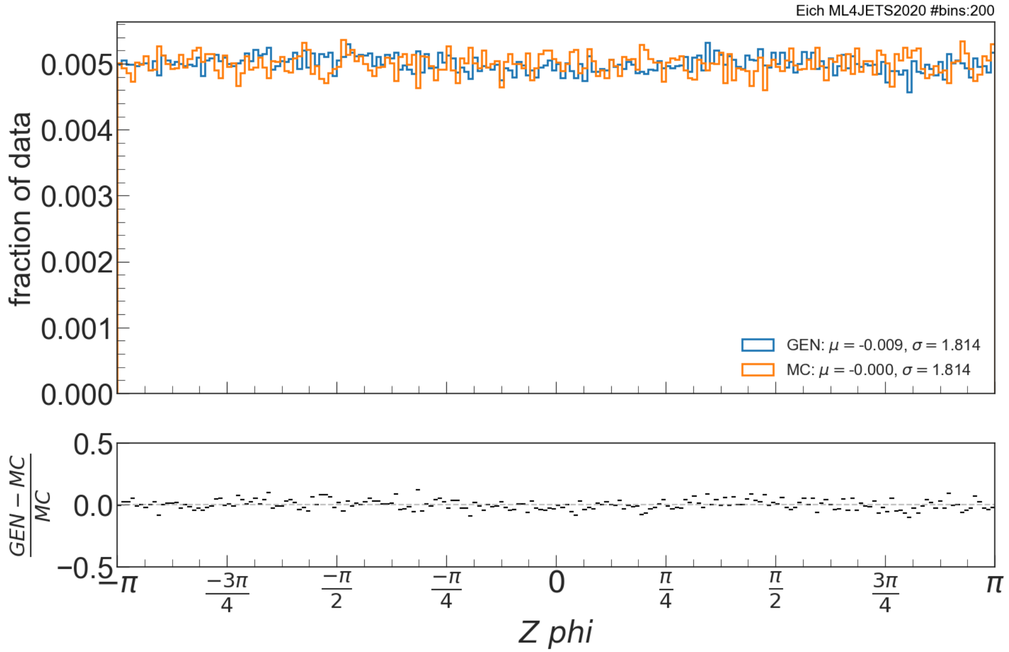


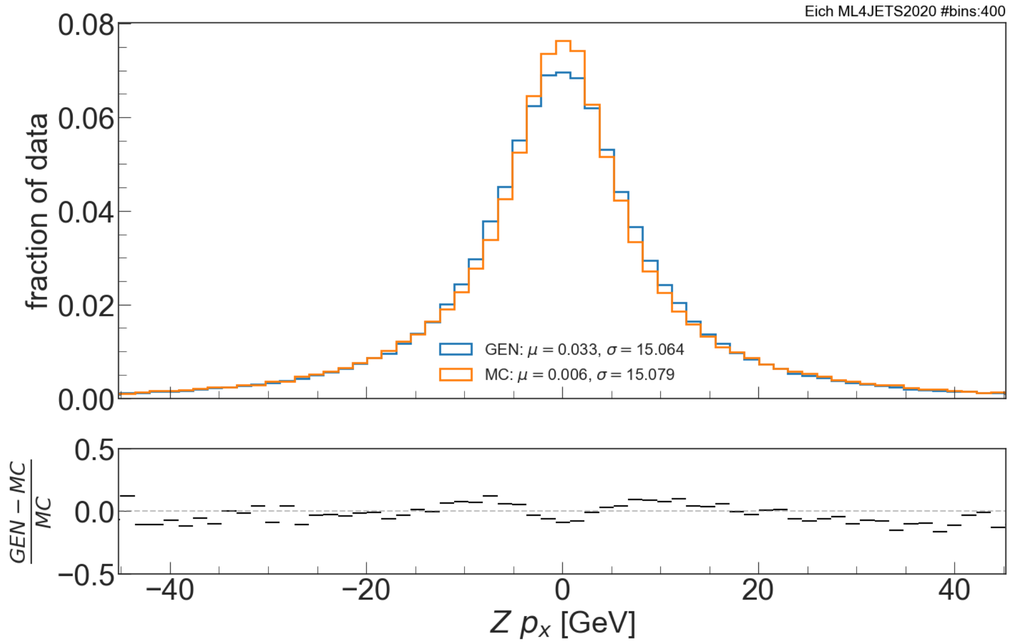
Critic Only EMD

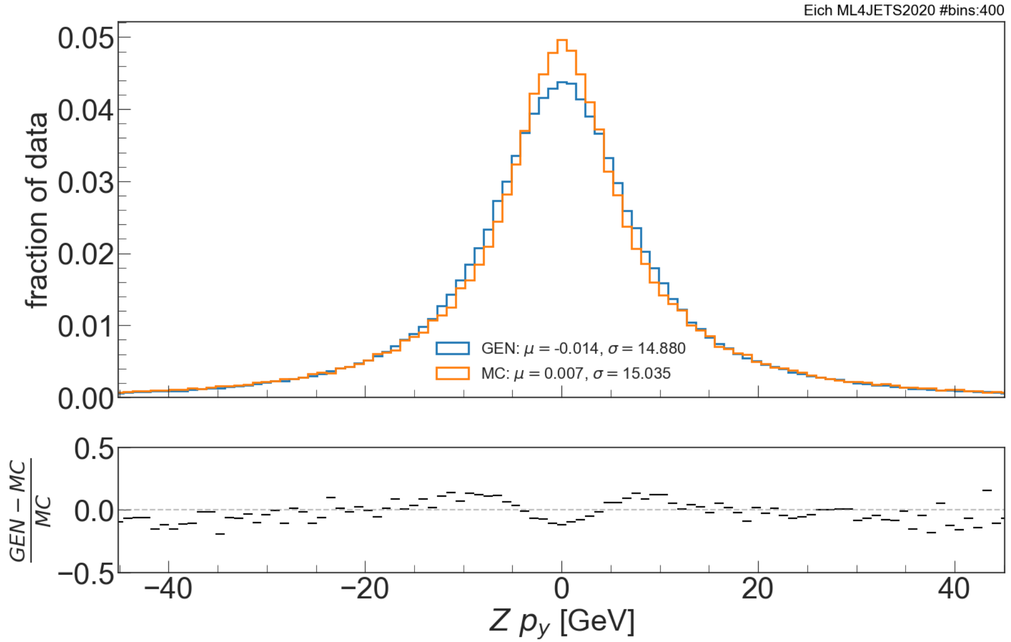


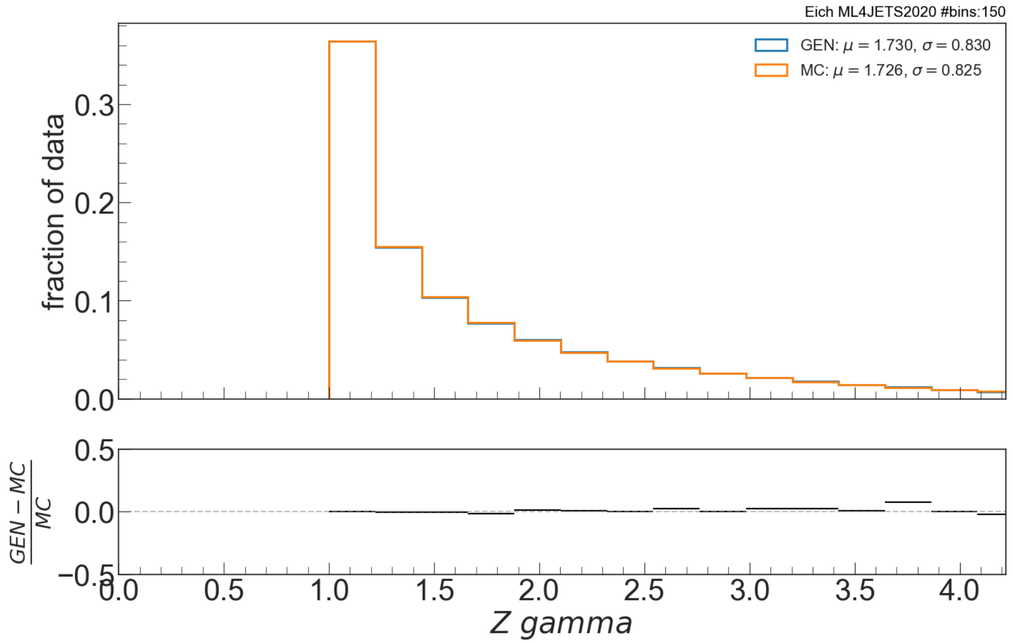


$Z \rightarrow \mu^+\mu^-$ - more distributions - ϕ, p_x, p_y, γ

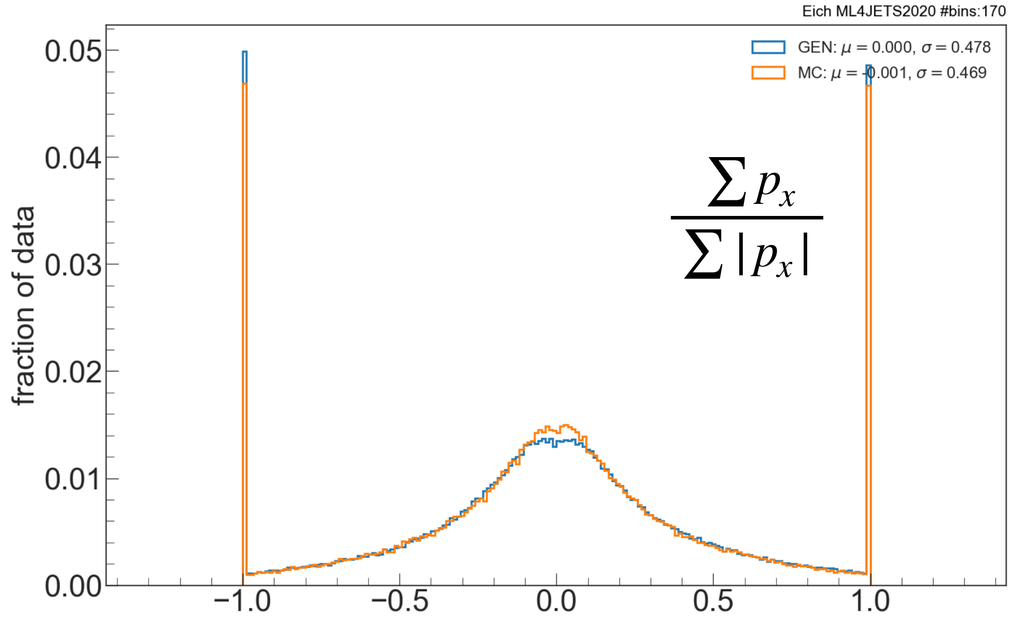


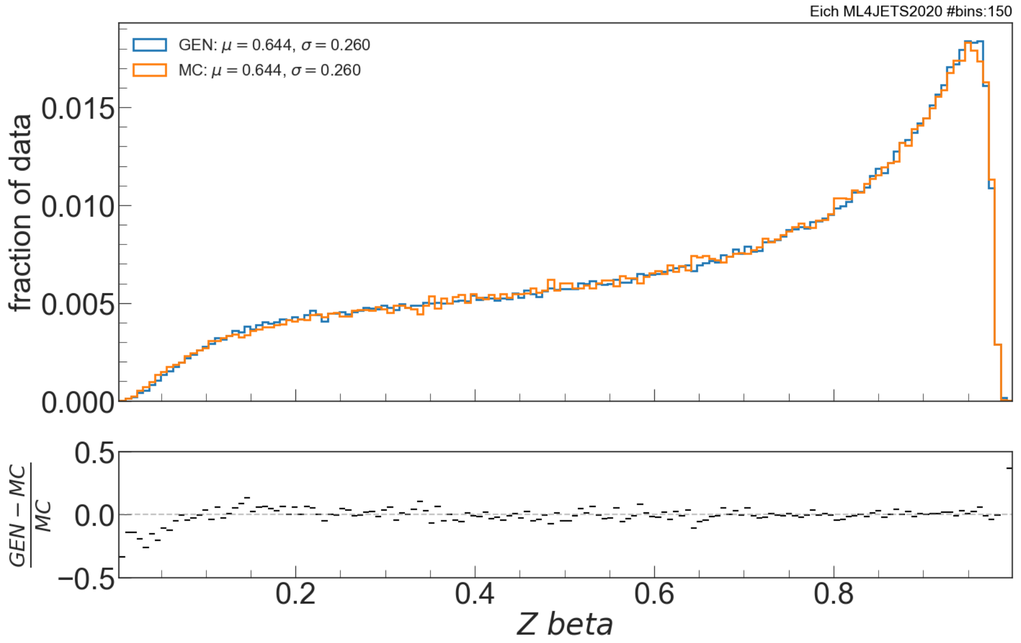


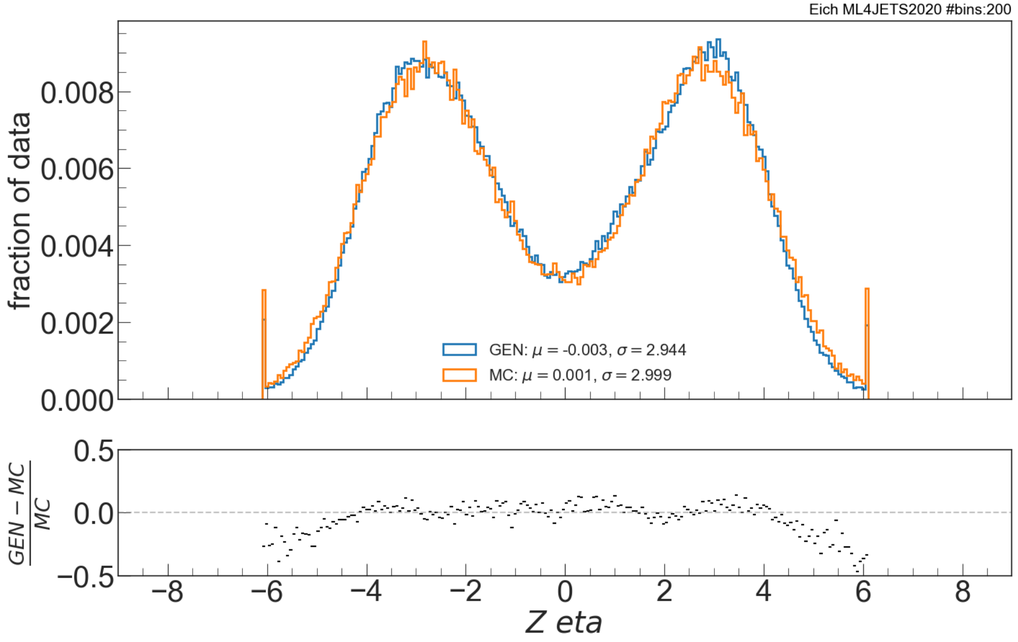


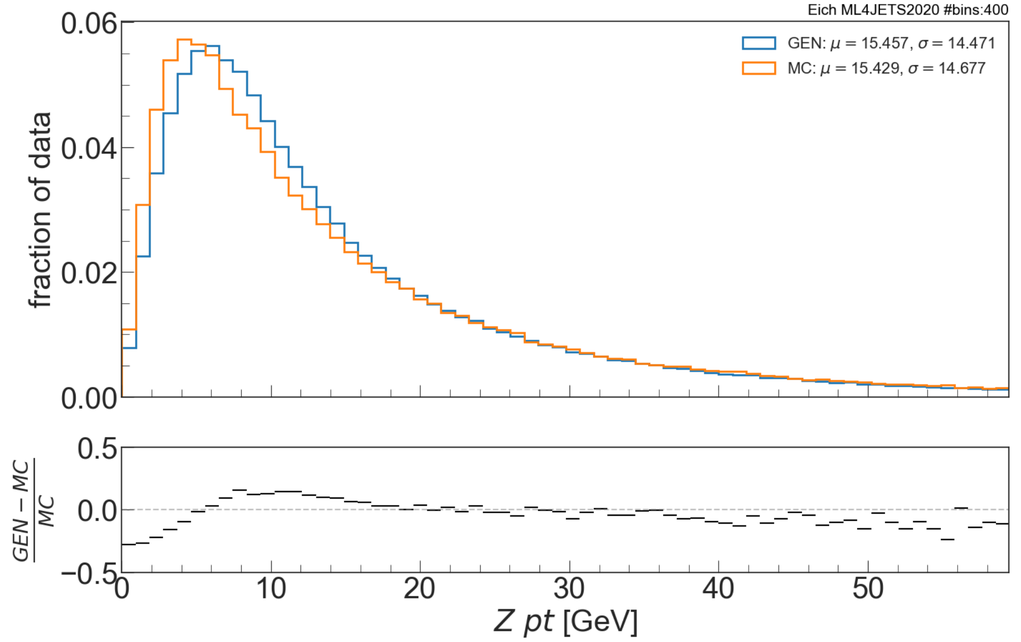


$Z \rightarrow \mu^+\mu^-$ - more distributions $\Delta p_x, \beta, \eta, p_T$





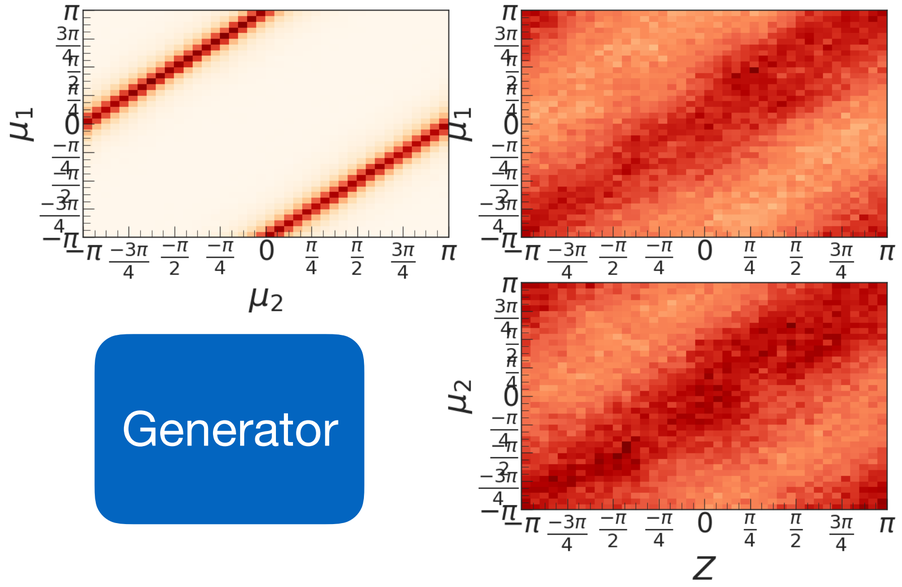




$Z \to \mu^+\mu^-$ - 2d histograms ϕ, η

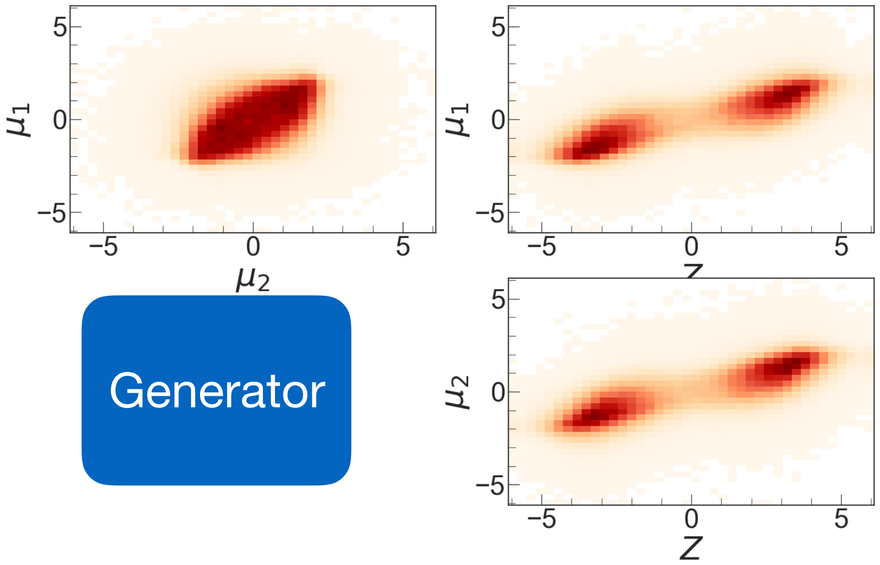
Niclas Eich - ML4Jets2020

phi - 2d histograms



Generator

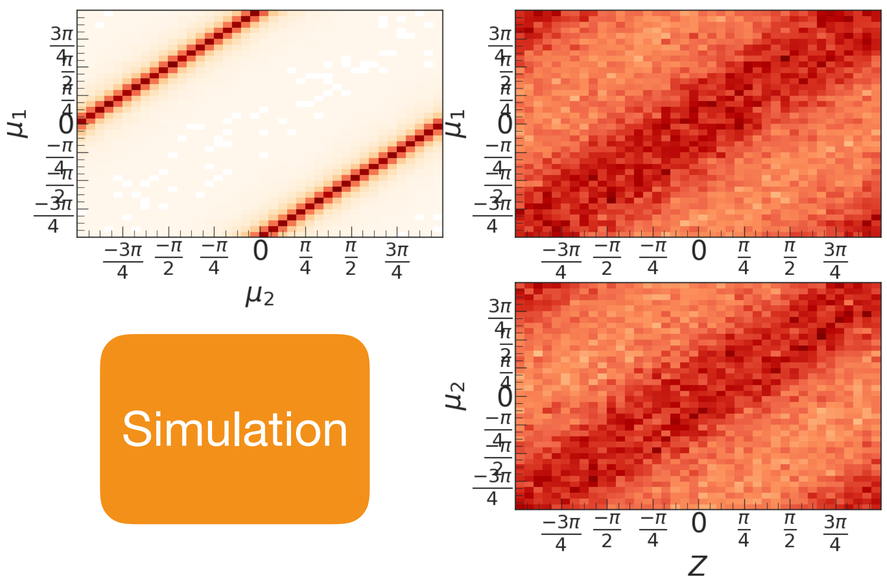
eta - 2d histograms



Generator

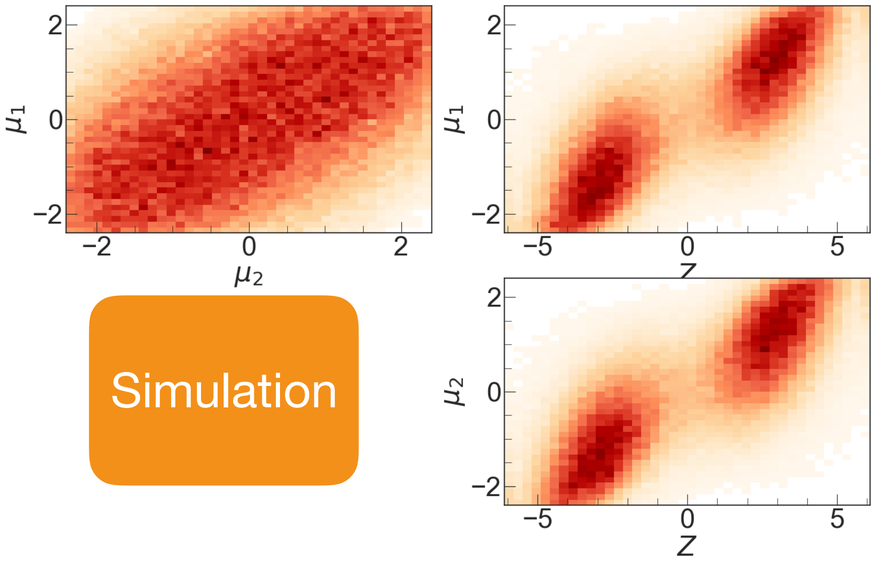
Plot axis scaled differently!

phi - 2d histograms



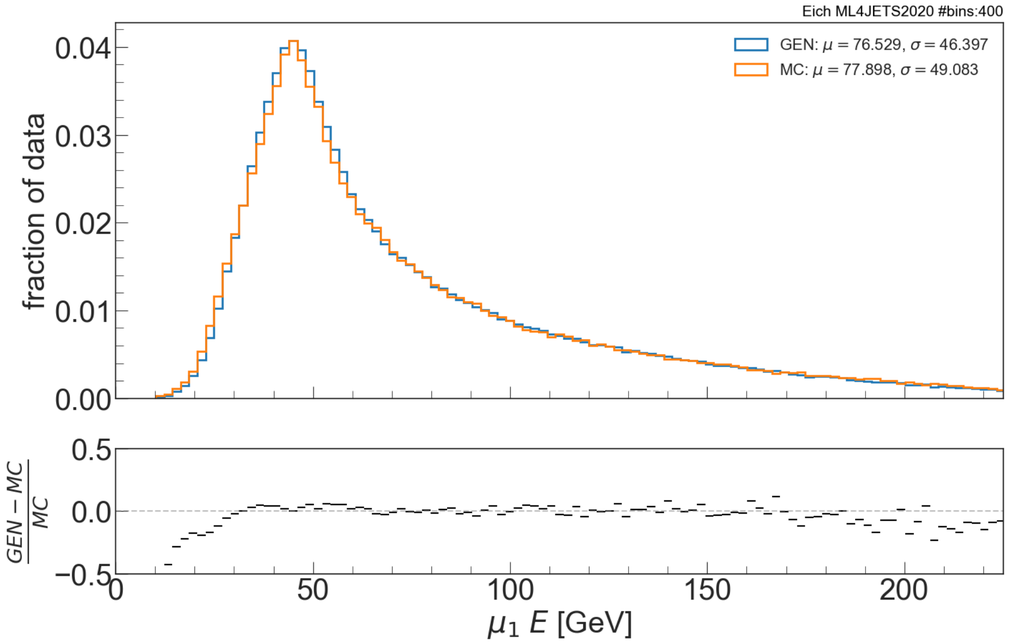
Simulation

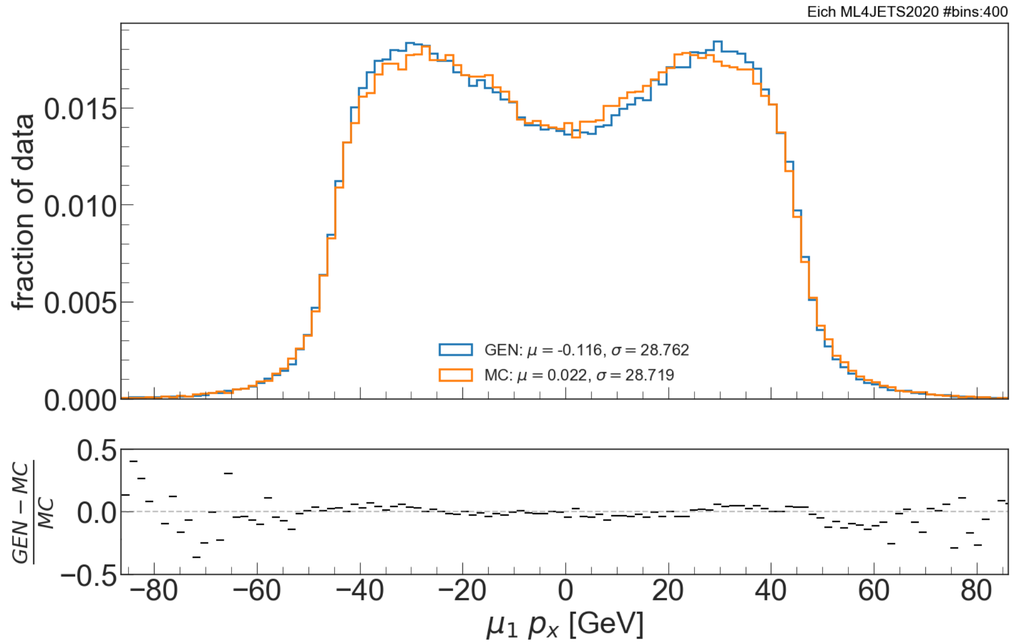
eta - 2d histograms

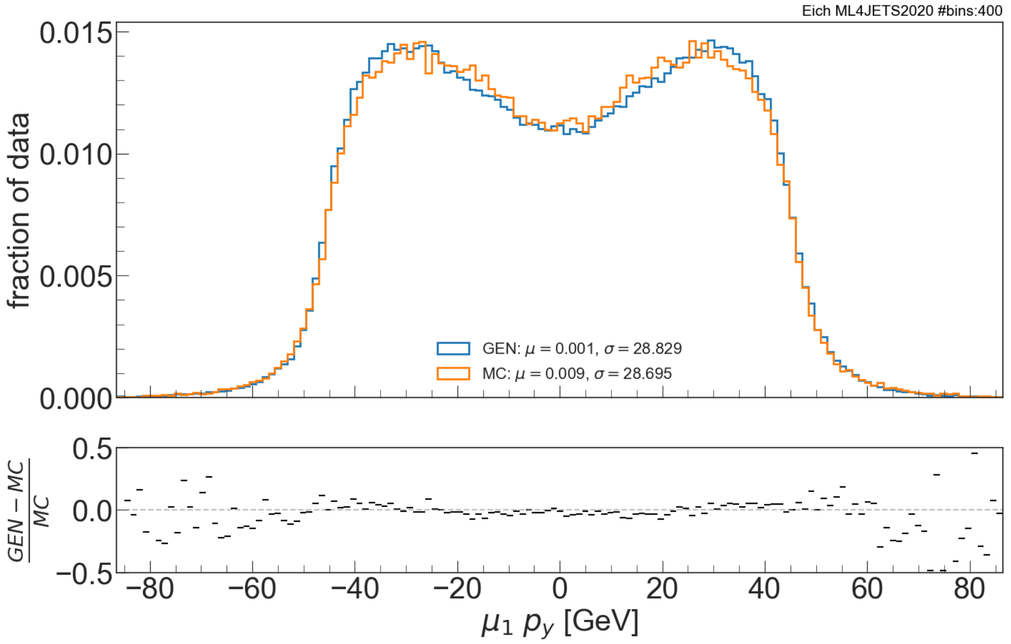


Simulation

$Z \rightarrow \mu^+\mu^-$ - applied η, p_T-cuts after generation

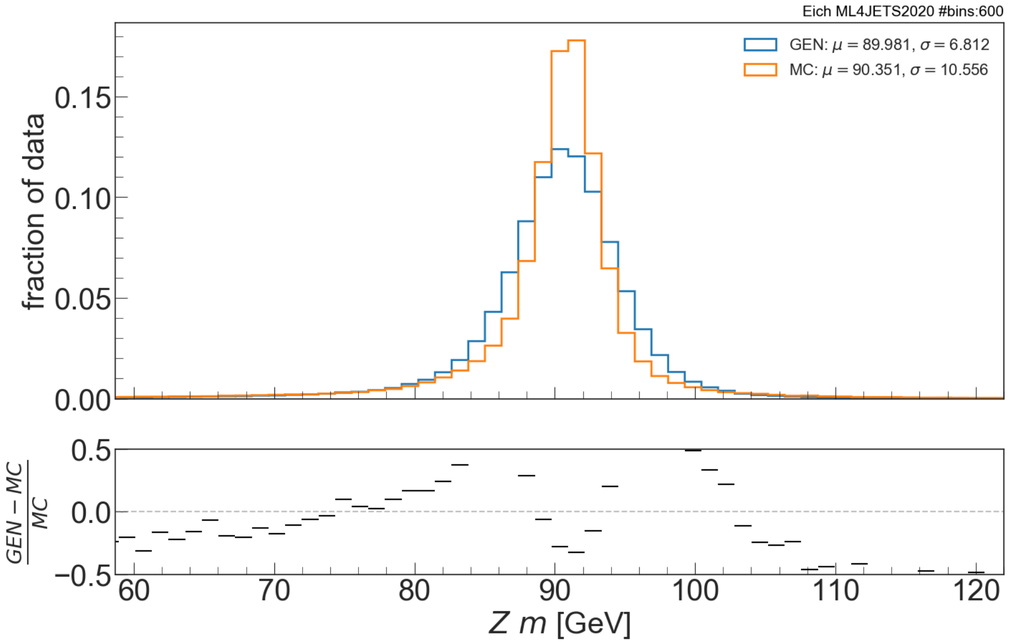


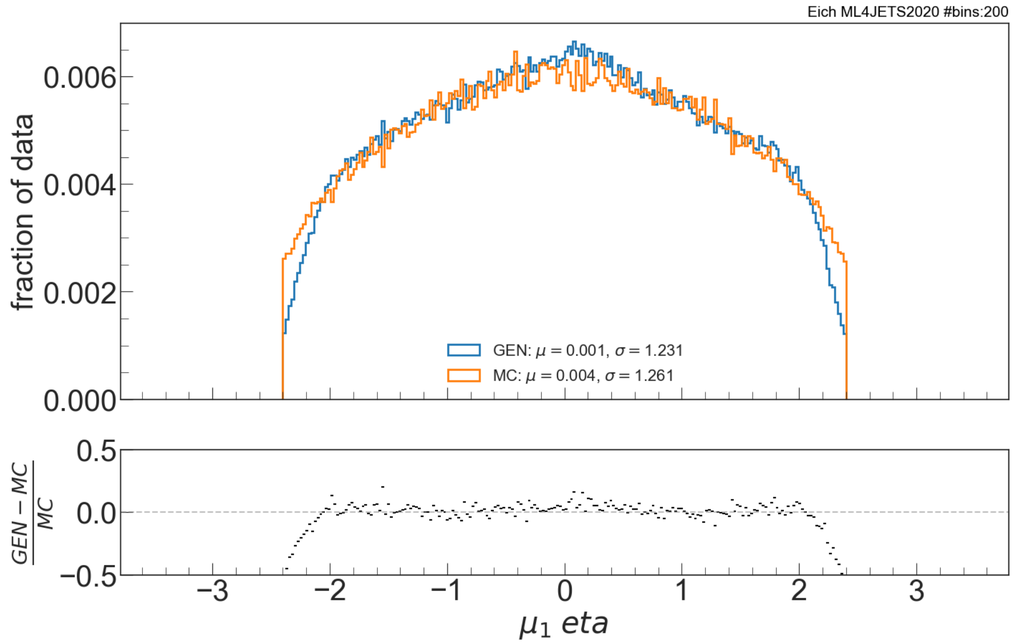


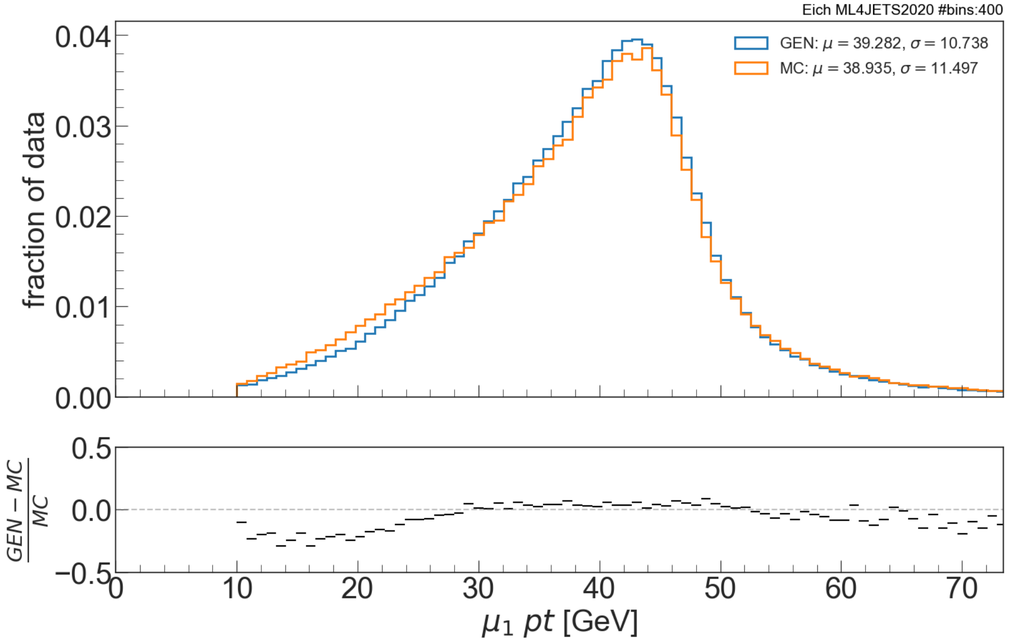


- Cuts improve transverse momenta p_x, p_y of the final state particles
- Rising slope of E-distribution gets worse
- Only losing < 1% events

$Z \rightarrow \mu^+\mu^-$ - applied η, p_T-cuts after generation

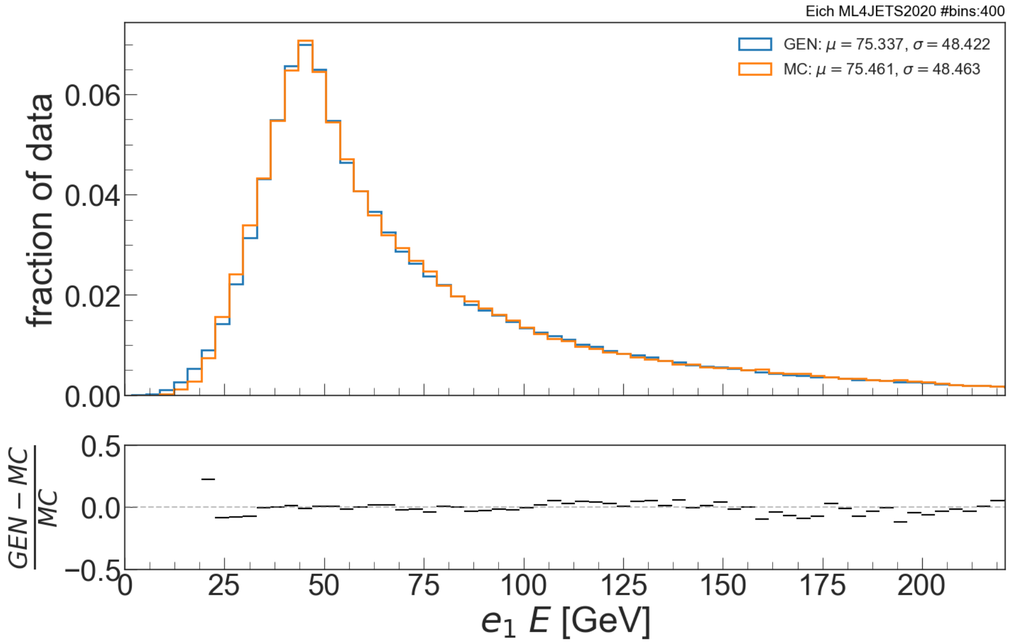


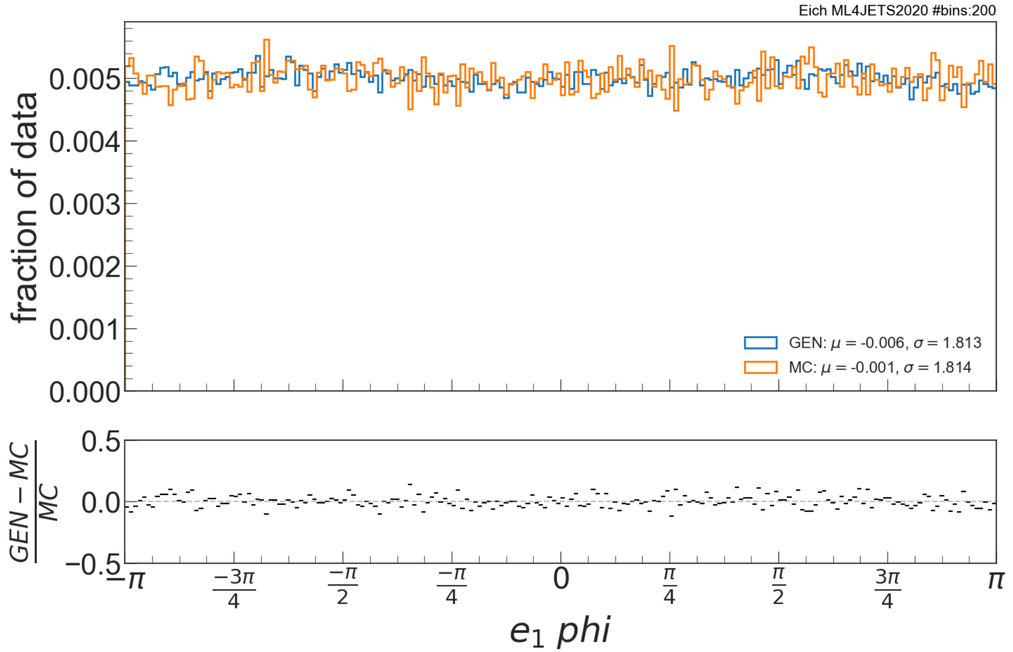


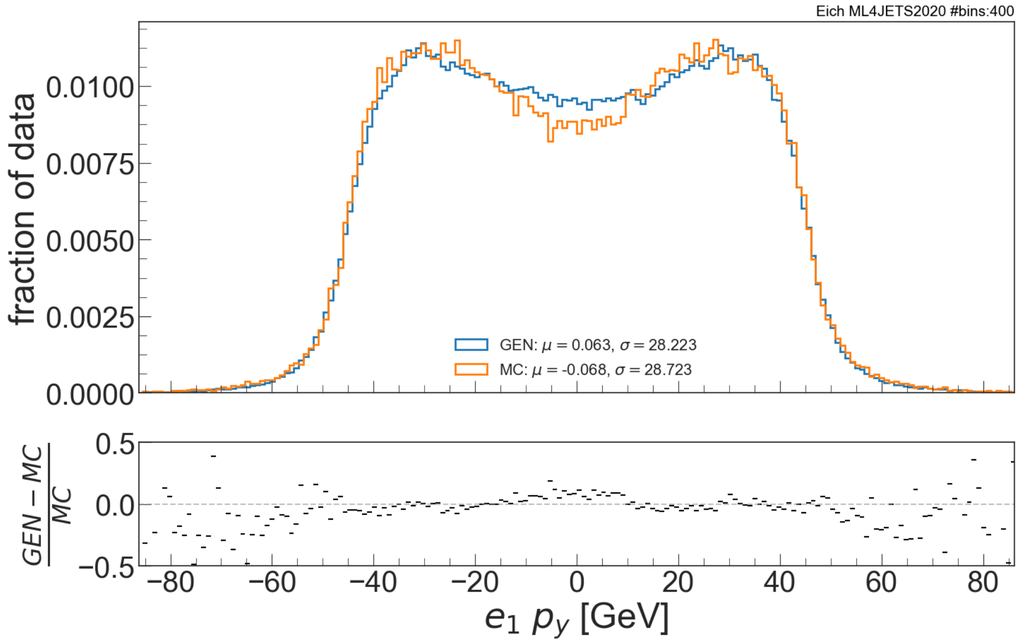


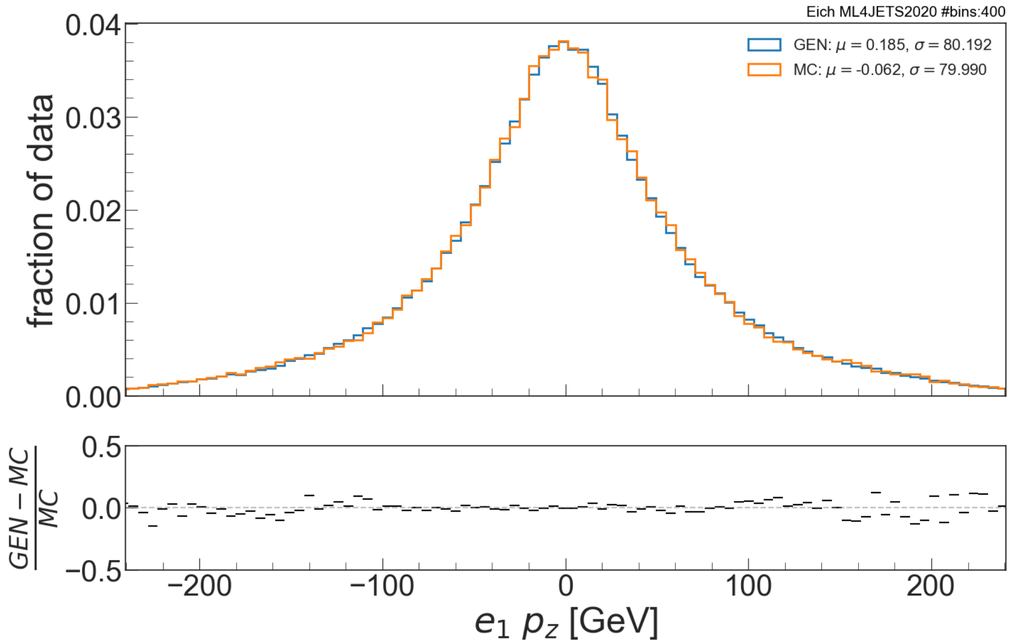
- Z mass does not improve
- Shape of p_T and η not in same agreement with simulation as other distributions

$Z \rightarrow e^+e^-$ - basic kinematics

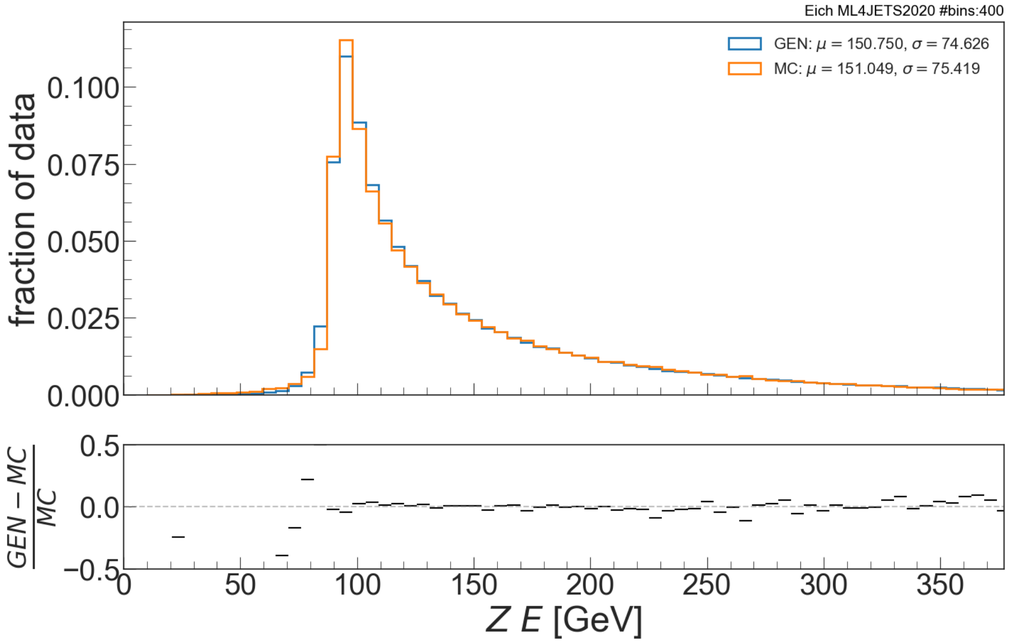
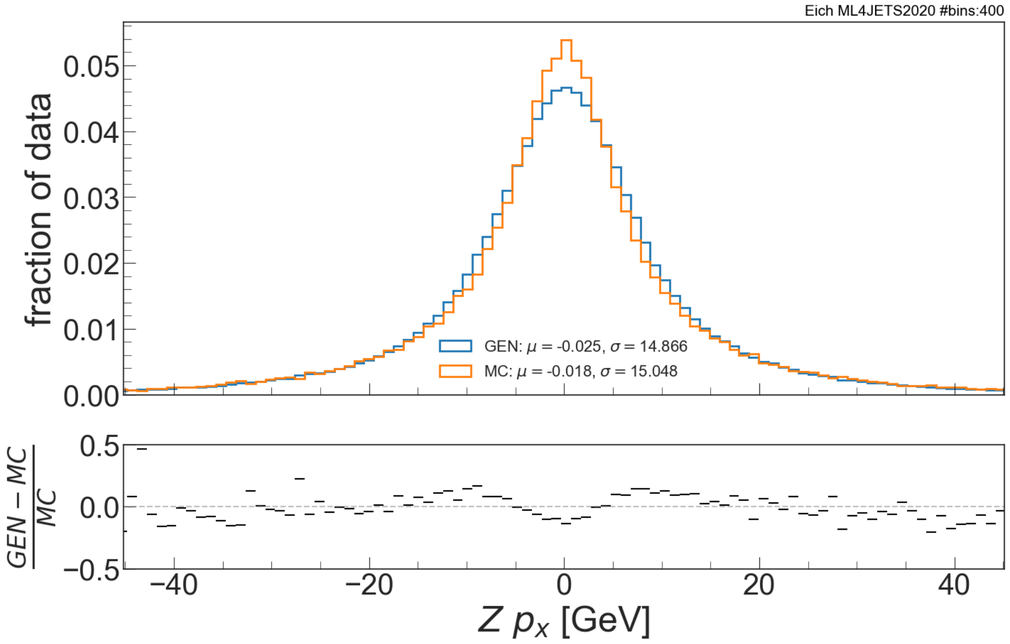
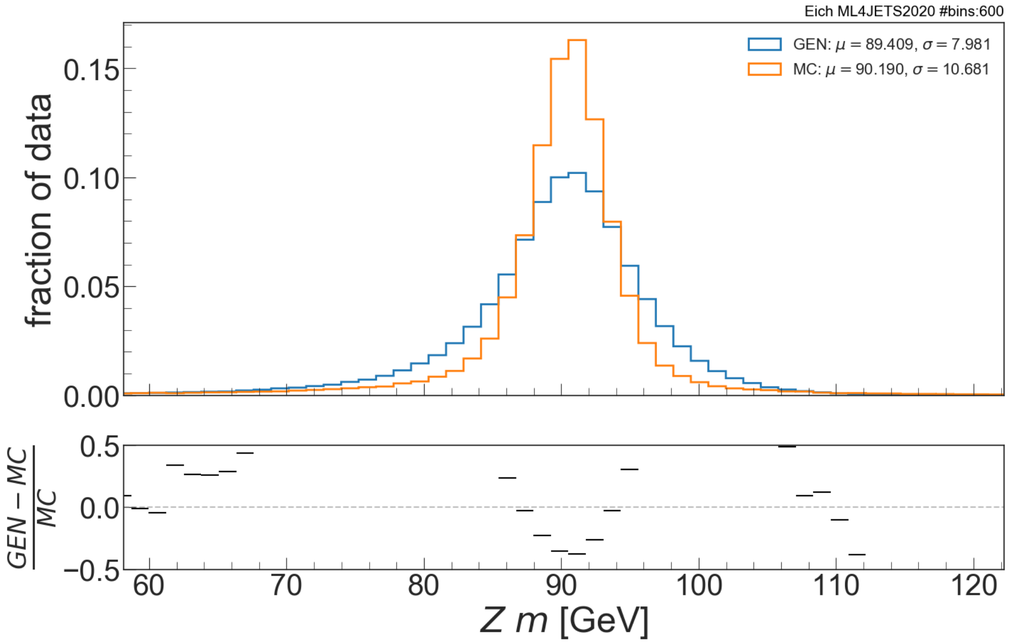
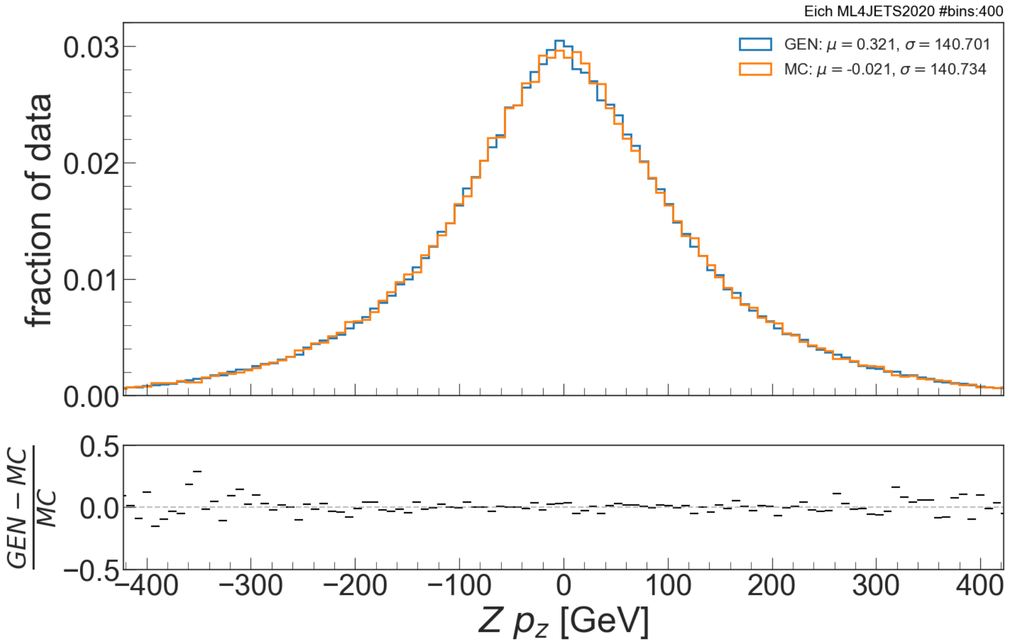








$Z \rightarrow e^+e^-$ - Z-candidate

$$\overrightarrow{e_1} + \overrightarrow{e_2} = \overrightarrow{Z}_{Candidate}$$

Dataset has very similar behaviour to $Z \rightarrow \mu^+\mu^-$