Symmetry aware generation of two-staged particle decays in high-energy physics

Niclas Eich, Martin Erdmann, Benjamin Fischer, Niklas Zimmermann

RWTH Aachen University, Physics Insitute III A

E-mail: `niclas.eich@rwth-aachen.de`

Abstract. We present a specialised layer for generative modeling of LHC events with generative adversarial networks. We use Lorentz boosts, rotations, momentum and energy conservation to build a network cell generating a 2-body particle decay. This cell is stacked consecutively in order to model two staged decays, respecting the symmetries across the decay chain. We allow for modifications of the resulting four-vectors in order to model higher order and detector effects. We give an evaluation of the generator quality in a Higgs decay into two Z bosons, further decaying into a muon pair each.

1. Introduction

The High Luminosity LHC will increase the rate of proton-proton collisions by a factor of $\sim$10. This results in a large increase in data and brings many challenges with it, one being the need for a sufficient amount of simulated events for analyses. The standard tool for these are Monte Carlo (MC) methods which are computationally expensive. Recently, machine learning methods such as Generative Adversarial Networks (GANs) [1] produced remarkable results in the context of natural image generation as well as pattern generation in fundamental research. GANs are being explored in various ways for their potential in future simulations [2].

2. Particle Decay Unit (PDU)

2.1. Kinematic Components

The Particle Decay Unit [3] is a specialised machine learning layer for Deep Neural Networks that models a two-body particle decay. Its input is the four-vector of the parent particle and its output are the two four-vectors of the decay products.We combine machine learning tools with analytic computations and transformations to ensure physics constraints and fundamental symmetries are preserved inside the generator network, as shown in figure 1.

(a) Boost into rest frame of the incoming particle by the Lorentz transformation Λ
(b) Analytic computation of decay particle momenta and energy
(c) Random orientation in 3D by rotation
(d) Boost decay particles into detector frame by the inverse Lorentz transformation Λ^{-1}

The momenta of the decay products

$$p_{1,2} = \frac{\sqrt{\left(M^2+m_{1,2}^2-m_{2,1}^2\right)^2-4M^2m_{1,2}^2}}{2M} \tag{1}$$

are computed inside the parent particle rest frame, respecting energy and momentum conservation. The system is at rest, thus its four-vector is $P = (M, \vec{0})^T$ with M being the mass of the parent particle and $m_{1,2}$ refer to the masses of the decay particles. At this stage, the resulting four-vectors are $(\sqrt{m_{1,2}^2 + \vec{p}_{1,2}^{\,2}}, \vec{p}_{1,2})^T$ with the three-momentum vector not having a specific orientation, yet. To fulfill the isotropic behaviour of such decays, a rotation is applied by multiplying the momentum vectors with a randomly sampled 3D rotation matrix.

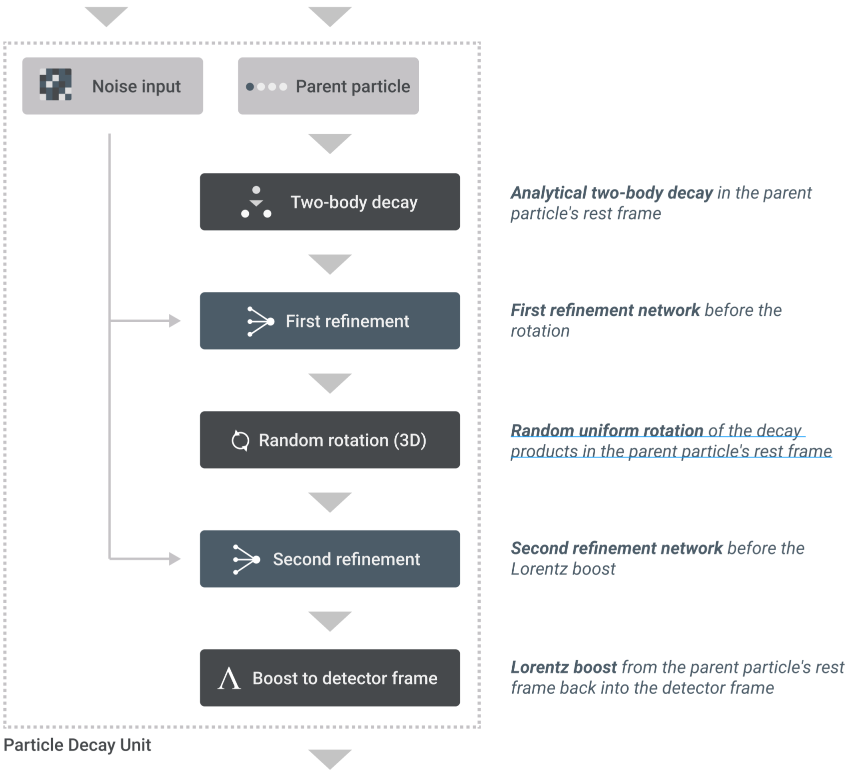

Figure 1. Sketch of the particle decay unit (PDU) structure. Analytic components like computation of the two-body decay momenta, Lorentz boost and rotation are used together with neural network layers [4].

2.2. Refinement Layers

The analytic transformations yield an algorithm that models a particle decay with the only random component being the angles of the rotation. In order to model more complex effects, such as spin-correlation or final-state radiation, residual layers are used that preserve the mass of the four-vectors but change their momenta. These residual layers receive the particle four-vectors, as well as noise as secondary input to account for natural variance. With the additional noise input, the refinement blocks have the capability to change the four-vectors of the decay particles event-by-event.

3. Model Architecture

3.1. Generator

The generator network is motivated by the decay chain $H \to ZZ \to 4\mu$. Each step is modeled by a dedicated substructure and works on the four-vectors without changing their masses.

- H four-vector: parent particle network
- $H \to ZZ$: first PDU
- $2\times(Z \to \mu\mu)$: second and third PDU

This results in a tree-like structure, where the outputs of the first PDU, representing the Z-bosons, are the inputs of the second and third PDUs, as shown in figure 2.

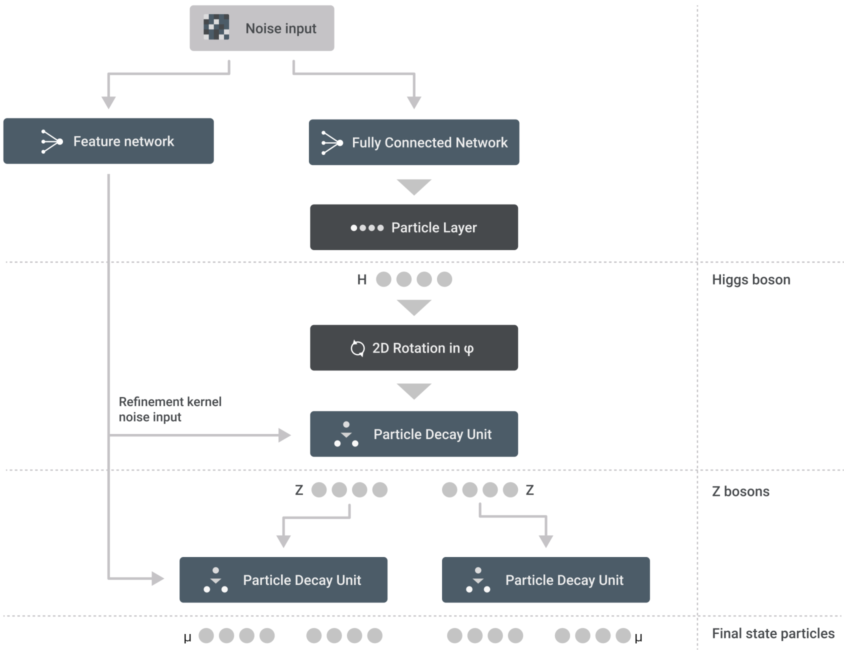

Figure 2. Sketch of the generator structure. The PDUs are chained after another to model the structure of the desired decay-chain [4].

Parent Particle Network The initial four-vector of the Higgs is constructed by a fully connected network block. Its input is a noise vector, sampling from a 30-dimensional Gaussian distribution. The four-vector is constructed by having four output nodes to the network, one for m, p_x, p_y and p_z. The energy is then computed by the energy-momentum relation. In order to enforce a symmetry in the transverse plane of the detector, x- and y-axis, a random rotation in this plane (azimuth) is applied to the four-vector.

Noise Preprocessing Generic GANs build up samples by passing the input noise through the network. The StyleGAN [5] architecture has shown noise injection at different points in the network as another suitable concept. In this work, a dedicated preprocessing network of densely connected layers is used, which feeds the PDUs to provide additional variance.

Particle Decay Units The PDUs are used to model the decay of parent particles into their decay particles. For the $H \to ZZ$ decay, the first PDU after the parent particle network is used.

Because the mass of the Z bosons typically follows a Breit-Wigner-like distribution with a mean value around 91 GeV, two output nodes of the noise preprocessing network are used to generate masses for the decay. The four-vectors of the Z bosons are then passed into the following two PDUs, both modeling the $Z \to \mu\mu$ decay separately for each new parent particle. For the Z decay PDUs, the mass of the muons are set to the fixed value of 0.105 GeV.

3.2. Discriminator

The discriminator network is a densely connected network. As input, it receives the four-vectors of the final state muons together with the reconstructed four-vectors of the intermediate particle resonances. The resonances are reconstructed by adding the respective muon-pair four-vectors. Additional inputs are the angular properties ϕ, η and the mass m.

3.3. Training

As advancement to the original GAN [1], several different training methods exist that proved to be more stable and achieve more realistic samples. Plehn *et. al.* [6] used a regularised GAN loss

$$\begin{aligned} L_{\text{Regularised}} &= L_{GAN} - \tfrac{\gamma}{2}\Omega_{JS}\left(P_r, P_\theta; D\right) \\ \Omega_{JS} &= \mathbb{E}_{x\sim P_r}\left[(1-D(x))^2\|\nabla\phi(x)\|^2\right] + \mathbb{E}_{x\sim P_\theta}\left[(D(x)^2\,\|\nabla\phi(x)\|^2\right], \end{aligned} \tag{2}$$

which they found to be suitable for particle physics applications. Furthermore, to model the sharp mass distributions of the resonances more accurately, an additional regularisation term, called *Maximum Mean Discrepancy* [6] is deployed

$$L_{\text{MMD}} = L + \left\langle k\left(x, x'\right)\right\rangle_{x,x'\sim P_r} + \left\langle k\left(y, y'\right)\right\rangle_{y,y'\sim P_\theta} - 2\langle k(x,y)\rangle_{x\sim P_r, y\sim P_\theta}, \tag{3}$$

where the kernels are chosen as Breit-Wigner functions $k_{\text{BW}}\left(x_1, x_2\right) = \frac{\sigma^2}{\left(x_1-x_2\right)^2+\sigma^2}$. This additional loss is used for each of the three bosons, H, Z_1 and Z_2 to improve the quality of the mass distributions.

4. Results

The quality of the generated samples is evaluated by comparing the resulting distributions with simulations that are obtained using the Madgraph5_aMC@nlo generator [7] and the DELPHES detector simulation [8]. In total $\mathcal{O}(10^5)$ events are evaluated.

Figure 3 shows the p_T and p_z distribution for one of the muons. For the p_T distribution, the generated samples match the MC events and the first and second moments within a few percent. The momentum along z-axis shows a peak that is a little broader with a standard deviation that is ∼10% larger than the one of the target distribution.

The resonances of the Z bosons and the H boson are reconstructed to investigate correlation between the particles. Figure 4 shows the mass distributions of the H and one of the Z bosons respectively. With the additional MMD-loss, the mass distributions match the expectations well. Normal GANs or Wasserstein GANs did not achive such a precision in our tests.

5. Conclusion

We presented and evaluated a two-stage particle decay using the Particle Decay Unit (PDU) as a building block. This Generative Adversarial Network setup combines a machine learning approach with analytical calculations ensuring physics constraints. The generator network used for the $H \to ZZ \to 4\mu$ decay incorporates three PDUs and has an internal structure that represents the components of the decay chain. With the Lorentz boosts and rotations, the network produces distributions that match the target Monte Carlo simulations rather well. The resonance masses are also modeled with reasonable accuracy, which we obtained when including the *Maximum Mean Discrepancy* loss.

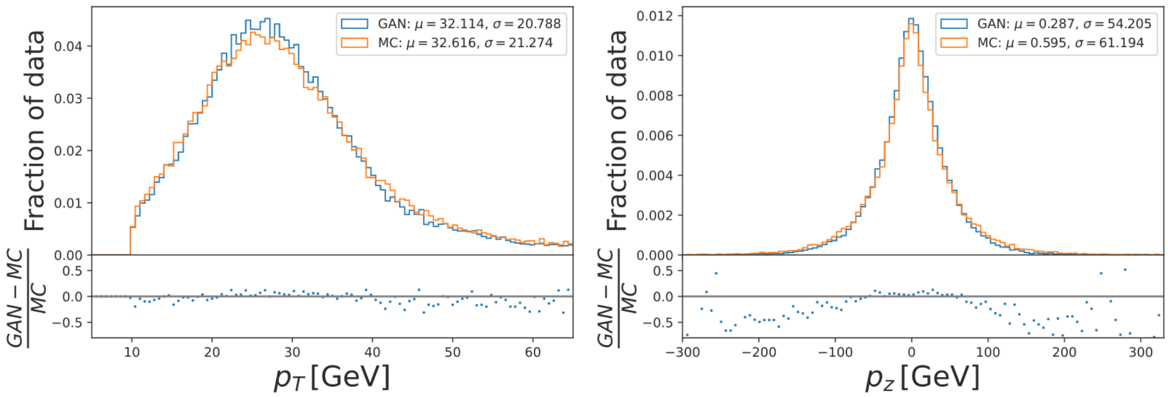

Figure 3. p_T and p_z for one of the final state muons.[4]

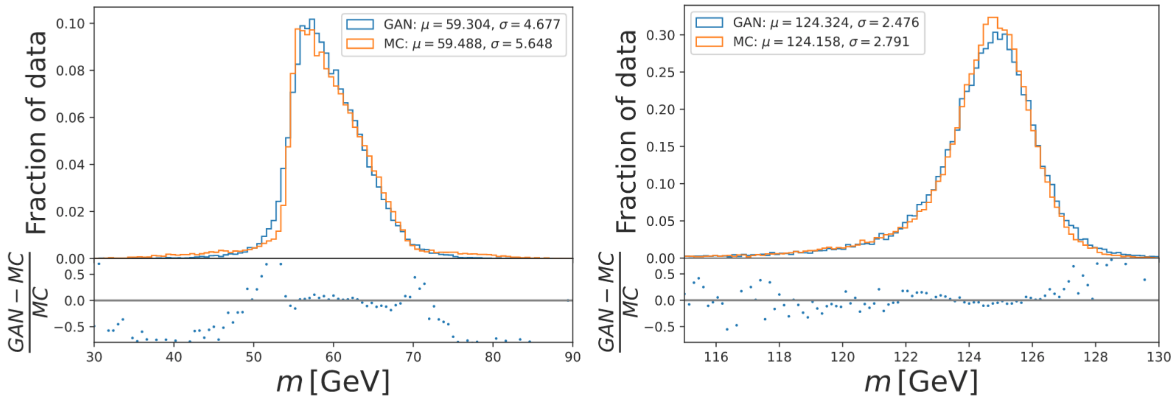

Figure 4. The mass distributions of the reconstructed Z (left) and H (right) resonances.[4]

References

[1] Ian J. Goodfellow, Jean Pouget-Abadie, Mehdi Mirza, Bing Xu, David Warde-Farley, Sherjil Ozair, Aaron Courville, Yoshua Bengio *Generative Adversarial Networks arXiv:1406.2661*

[2] Matthew Feickert and Benjamin Nachman,A Living Review of Machine Learning for Particle Physics *arXiv:2102.02770*

[3] Niclas Eich 2020 *Symmetry Aware Generation of Particle Decays using Generative Adversarial Networks in High-Energy Physics* Master thesis, RWTH Aachen University `https://www.institut3a.physik.rwth-aachen.de/global/show_document.asp?id=aaaaaaaaaovhdpg`

[4] Niklas Zimmermann 2021 *Generation of Two-Stage Particle Decays using Symmetry-Aware Generative Adversarial Networks at the CMS Experiment* Bachelor thesis, RWTH Aachen University `https://www.institut3a.physik.rwth-aachen.de/global/show_document.asp?id=aaaaaaaabevdlxr`

[5] Tero Karras, Samuli Laine, Timo Aila 2018 *A Style-Based Generator Architecture for Generative Adversarial Networks CoRR* **abs/1812.04948**

[6] Anja Butter and Tilman Plehn and Ramon Winterhalder 2019 *How to GAN LHC Events SciPost Phys.* **7** 75

[7] J. Alwall, R. Frederix, S. Frixione, V. Hirschi, F. Maltoni, O. Mattelaer, H.-S. Shao, T. Stelzer, P. Torrielli, M. Zaro 2014 *Journal of High Energy Physics* **7** *79*

[8] J. de Favereau, C. Delaere, P. Demin, A. Giammanco, V. Lemaître, A. Mertens, and M. Selvaggi 2014 *Journal of High Energy Physics* **2** *57*