The ATLAS detector as a muon fixed-target experiment: using generative models to simulate muonic force carriers

Tigran Ramazyan[a,*], Sofia Vallecorsa[b], Enrique Kajomovitz[c]

[a] *National Research University Higher School of Economics, Moscow, Russia*
[b] *CERN, Switzerland*
[c] *Physics Department, Technion – Israel Institute of Technology, Haifa 3200003, Israel*

Abstract

Greater luminosities of future Large Hadron Collider runs will demand an unprecedented number of event simulations. Computationally that would be an extremely demanding task. Hence new approaches for such undertakes are required. In recent years the usage of generative models for event simulation has become a common practice. A new experiment on the measurement of the g-Factor of the anomalistic magnetic moment of the muon will be conducted on FASER. We present a Generative Adversarial Network for fast event simulation for the experiment.

1. Introduction

In High Energy Physics (HEP) the use of Monte Carlo (MC) tools such as GEANT-based detector simulators are an established method to describe experimental characteristics. This approach has been practised successfully over the last decades [1, 2, 3, 4, 5, 6]. Lately more frequently HEP MC simulations have statistical accuracy issues because of high CPU cost of generating [7] and memory cost of storing [8] MC events. Over the last years generative models such as Generative Adversarial Networks (GANs) [9] and Variational Auto-Encoders (VAEs) [10] have started to replace MC for event reconstruction. Deep generative models do not suffer the resource limitations of MC methods and are significantly faster.

This report describes an implementation of a Conditional GAN architecture for generating analysis-specific datasets for a new experiment on the measurement of the g-Factor of the anomalistic magnetic moment of the muon that will be conducted as part of the FASER experiment.

*Corresponding author

Email addresses: ramazyant@gmail.com (Tigran Ramazyan), sofia.vallecorsa@cern.ch (Sofia Vallecorsa), enrique.kajomovitz@gmail.com (Enrique Kajomovitz)

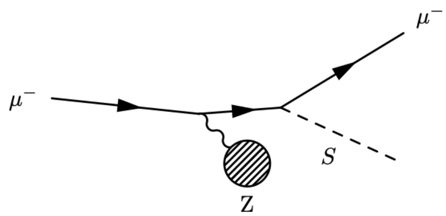


Figure 1: Dark bremsstrahlung signal process for simplified models with invisibly decaying scalar (left) and vector (right) forces that couple predominantly to muons. In both cases, a relativistic muon beam is incident on a fixed target and scatters coherently off a nucleus to produce the new particle as initial- or final-state radiation.

2. Related Work

Generative Adversarial Networks [9] have been tested for LHC applications to simulate directly high-level experiment features [11, 12, 13, 14, 15, 16, 17]. Specifically, GANs have been used for LHC applications in numerous ways: calorimeters [12, 18], Cherenkov detectors [19], etc.

3. Experiment Description

The experiment can be described by the diagrams in Fig. 1 from [20]. The simplest interpretation would be the following. An incoming muon collides with a nucleus. The collision yields an outgoing muon and a Muonic Force Carrier (MFC).

4. Data

We consider the collision presented in the Section 3, generated using MadGraph [21] event generator for 7 different values (see Fig. 2) for the energy of the incoming muon: 50 Gev (see Fig. 3), 75 Gev, 100 Gev, 125 Gev, 150 Gev, 175 Gev, 200 Gev.

An event is represented by 6 values: the corresponding to the outgoing muon and the MFC energies, transverse momenta and pseudorapidities. Also we condition the GAN on the energy of the incoming muon.

Multiple test runs have shown the model capability to reproduce the original although non-Gaussian shaped distributions of energies. Also shapes of the original distributions of pseudorapidities are Gaussian-like. Hence only the transverse momenta distributions have been transformed. The following transformations were reasoned by the idea to make the shapes of transverse momenta distributions more Gaussian-like and multiple test runs (Fig. 4, Fig. 5).

$$\textbf{Muon: } P_T^{scaled} = 100 \cdot P_T^{0.2}; \ \textbf{MFC: } P_T^{scaled} = 100 \cdot P_T^{0.124}$$

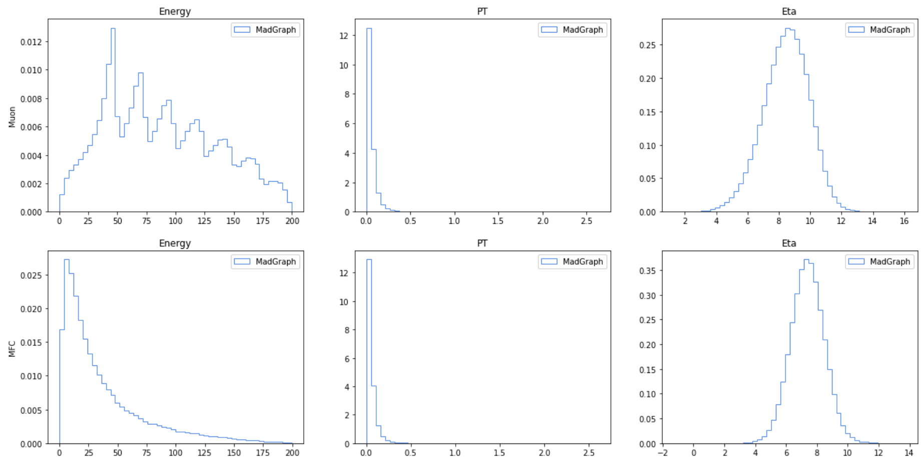


Figure 2: Histograms of the full dataset, i.e. all the seven datasets in one.

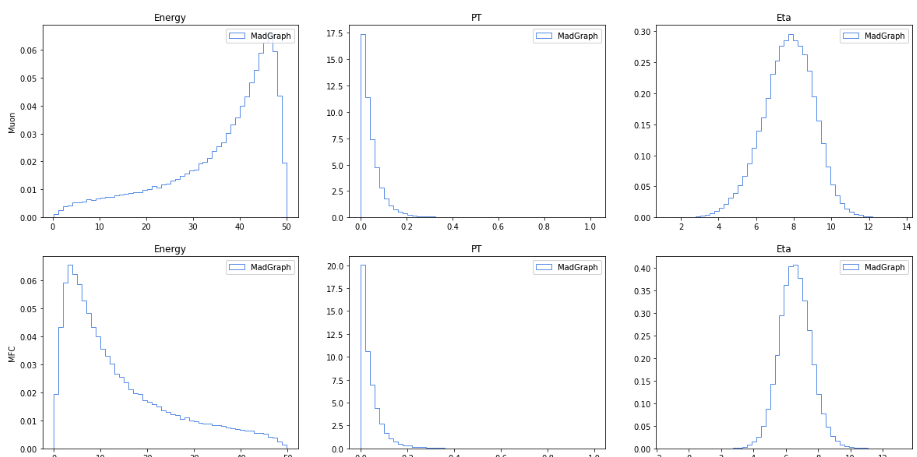


Figure 3: Histograms of the 50 GeV dataset, i.e. the energy of the incoming muon was 50 GeV.

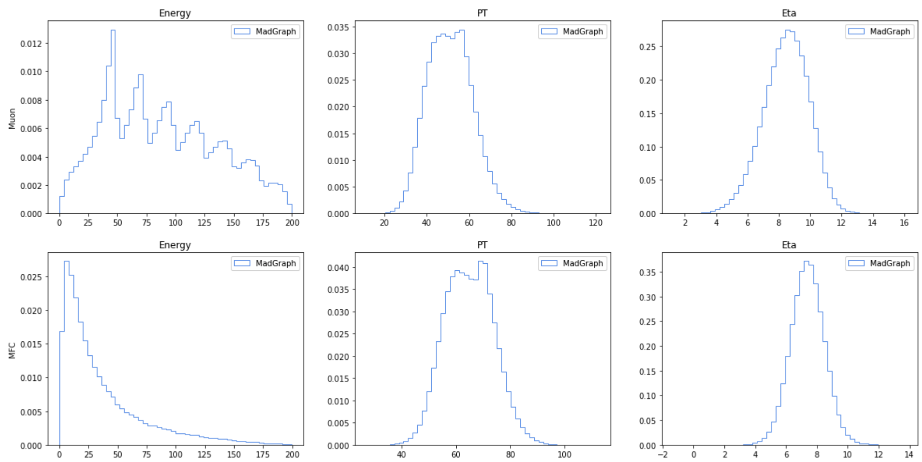


Figure 4: Histograms of the scaled full dataset, i.e. all the seven datasets in one.

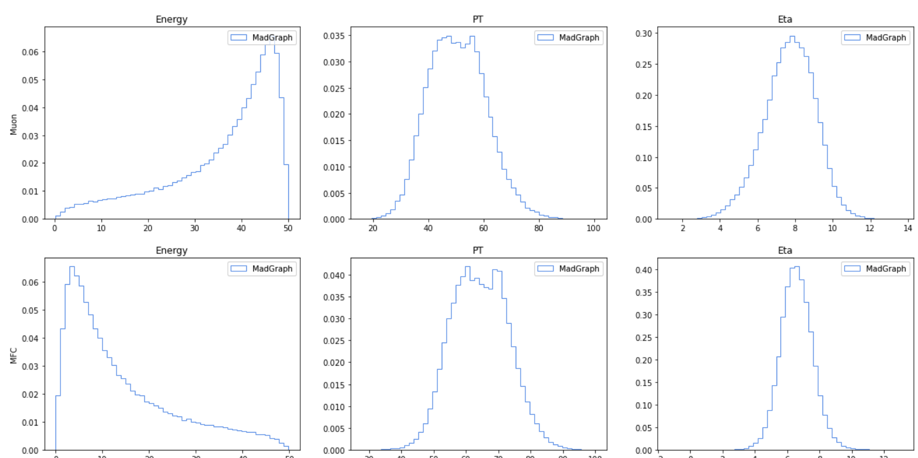
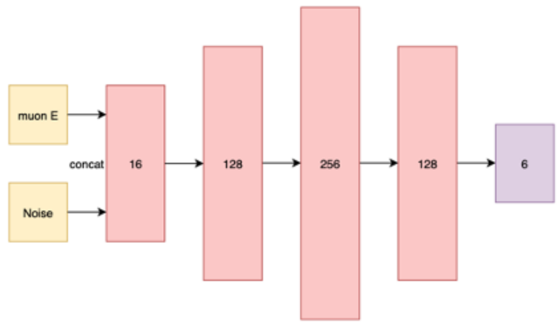
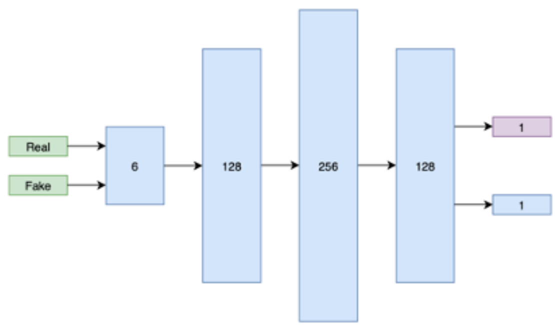


Figure 5: Histograms of the non-scaled 50 GeV dataset, i.e. the energy of the incoming muon was 50 GeV.



(a) Generator



(b) Discriminator

Figure 6: MFC GAN Architecture.

5. Model Architecture

The model architecture is a Conditional Generative Adversarial Network (CGAN) (see Fig. 6). The Generator network consists of 5 fully connected layers with 16, 128, 256, 128 and 6 neurons. Hidden layer neurons are activated by a sigmoid function. Normalisation layers are placed in between layers. The Generator is getting as input a noise vector drawn from a Gaussian distribution centred at 0 with unit variance, i.e. N(0, 1), concatenated with the condition. The condition is presented by the energy of the incoming muon. The Generator is trained using the Binary Cross Entropy loss, i.e.

$$L_G = -\frac{1}{2}\mathbb{E}_{\hat{x}\sim\mathbb{P}_{fake}}\log(D(\hat{x})),\ \text{where } \hat{x} = G(e, z), z \sim \mathcal{N}(0, 1).$$

The Discriminator network consists of 6 fully connected layers with 6, 128, 256, 128 neurons and two output layers with 1 neuron. Hidden layer neurons are activated by a sigmoid function. So is one of the output layers. The Discriminator gets as input either original data or fake data, i.e. created by the Generated. One of the output layers is used to compute the Binary Cross Entropy loss, while the other output layer is used to build a regression on the condition, i.e. the Discriminator is trained using the following loss function:

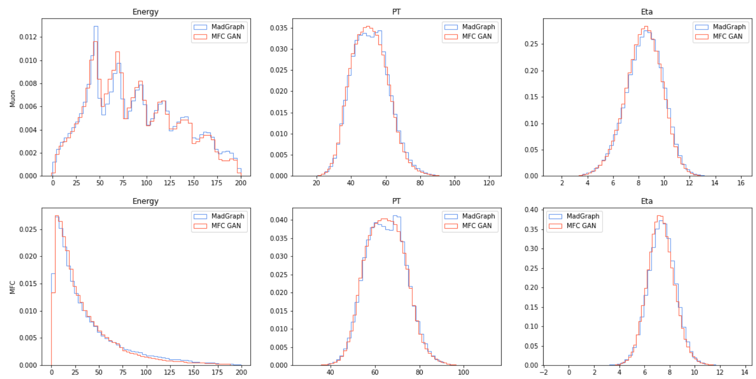


Figure 7: Full dataset results.

$$L_D = -\frac{1}{2}\mathbb{E}_{x\sim\mathbb{P}_{real}}\log(D(x)) - \frac{1}{2}\mathbb{E}_{x\sim\mathbb{P}_{fake}}\log(1-D(\hat{x})) + \lambda\cdot(\mathrm{MSE}(e,\hat{e}) + (\sigma(e)-\sigma(\hat{e}))^2).$$

We have also added a standard deviation term of the energy outputted by the Discriminator to the loss function. Empirically this has proven to increase the performance of the model.

We also ran experiments with a Wasserstein CGAN architecture. The loss functions remained constant for the whole training process.

Additionally, conditioning on the transverse momentum and/or the pseudorapidity of the incoming muon might be tested in an attempt to increase the performance further.

6. Results

All networks were implemented in PyTorch [22]. The training was carried on using the RMSprop optimizer. The parameters of the model were tuned by running a thorough grid search for the following parameters: learning rate, auxiliary loss weight, number of epochs, number of hidden layers, size of hidden layers. The best set of parameters were then fine tuned by minimising the Hellinger distance between the original data and the generated one. The optimal set of parameters yielded the result presented in Fig. 7. The parameters suit even for generation of a single dataset, e.g. Fig. 8.

One of the key aspects of the projects was testing the interpolation capabilities of the model. As an interpolation test we trained the model one 6 out of the 7 datasets, leaving out one in between, and then made a prediction for it. Fig. 9 shows the results of such test. The model has shown decent performance interpolating among the data it was trained on.

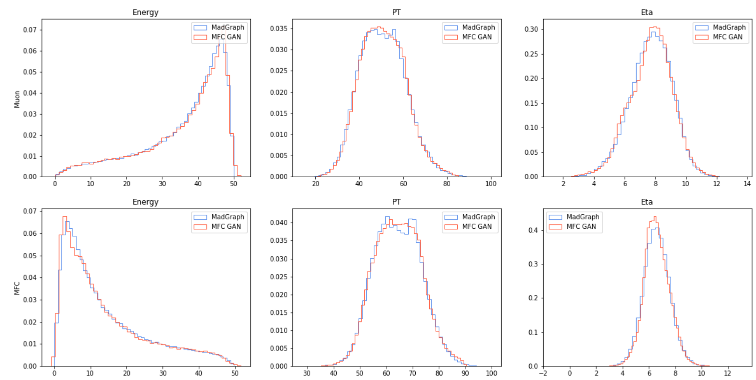


Figure 8: 50 GeV dataset results.

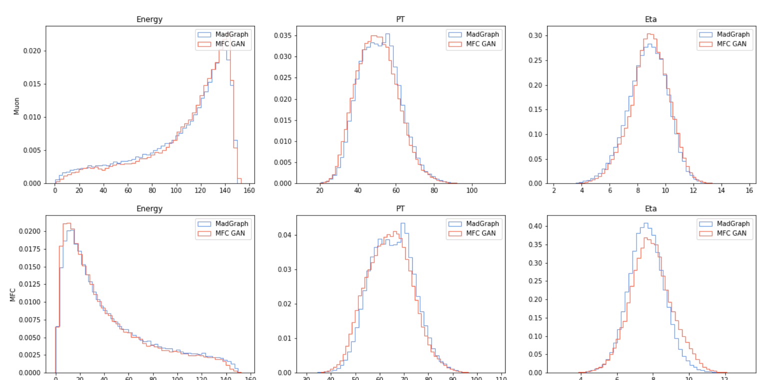


Figure 9: Interpolation test results.

7. Conclusion

We presented a fully working Conditional Generative Adversarial Network for event simulation for a new experiment on the measurement of the g-Factor of the anomalistic magnetic moment of the muon that will be conducted on FASER. The model has shown strong results and will presumably be used during the next run of the LHC.

References

[1] W. F. Fry, D. Haidt, Calculation of the neutron-induced background in the Gargamelle neutral current search, CERN Yellow Reports: Monographs, CERN, Geneva, 1975. doi:10.5170/CERN-1975-001.
URL http://cds.cern.ch/record/185979

[2] G. Aad, T. Abajyan, B. Abbott, J. Abdallah, S. Abdel Khalek, A. Abdelalim, O. Abdinov, R. Aben, B. Abi, M. Abolins, et al., Observation of a new particle in the search for the standard model higgs boson with the atlas detector at the lhc, Physics Letters B 716 (1) (2012) 1–29. doi:10.1016/j.physletb.2012.08.020.
URL http://dx.doi.org/10.1016/j.physletb.2012.08.020

[3] S. Chatrchyan, V. Khachatryan, A. Sirunyan, A. Tumasyan, W. Adam, E. Aguilo, T. Bergauer, M. Dragicevic, J. Erö, C. Fabjan, et al., Observation of a new boson at a mass of 125 gev with the cms experiment at the lhc, Physics Letters B 716 (1) (2012) 30–61. doi:10.1016/j.physletb.2012.08.021.
URL http://dx.doi.org/10.1016/j.physletb.2012.08.021

[4] T. Sjöstrand, S. Ask, J. R. Christiansen, R. Corke, N. Desai, P. Ilten, S. Mrenna, S. Prestel, C. O. Rasmussen, P. Z. Skands, An introduction to pythia 8.2, Computer Physics Communications 191 (2015) 159–177. doi:10.1016/j.cpc.2015.01.024.
URL http://dx.doi.org/10.1016/j.cpc.2015.01.024

[5] J. de Favereau, C. Delaere, P. Demin, A. Giammanco, V. Lemaître, A. Mertens, M. Selvaggi, Delphes 3: a modular framework for fast simulation of a generic collider experiment, Journal of High Energy Physics 2014 (2). doi:10.1007/jhep02(2014)057.
URL http://dx.doi.org/10.1007/JHEP02(2014)057

[6] S. Agostinelli, et al., GEANT4–a simulation toolkit, Nucl. Instrum. Meth. A 506 (2003) 250–303. doi:10.1016/S0168-9002(03)01368-8.

[7] A. M. Sirunyan, A. Tumasyan, W. Adam, F. Ambrogi, E. Asilar, T. Bergauer, J. Brandstetter, E. Brondolin, M. Dragicevic, et al., Measurement of the cross section for top quark pair production in association

with a w or z boson in proton-proton collisions at

$$\sqrt{s} = 13$$

tev, Journal of High Energy Physics 2018 (8). doi:10.1007/jhep08(2018)011.
URL http://dx.doi.org/10.1007/JHEP08(2018)011

[8] A. Sirunyan, A. Tumasyan, W. Adam, F. Ambrogi, E. Asilar, T. Bergauer, J. Brandstetter, M. Dragicevic, J. Erö, A. Escalante Del Valle, et al., Measurement of the top quark polarization and tt^{-} spin correlations using dilepton final states in proton-proton collisions at s=13tev, Physical Review D 100 (7). doi:10.1103/physrevd.100.072002.
URL http://dx.doi.org/10.1103/PhysRevD.100.072002

[9] I. J. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, Y. Bengio, Generative adversarial networks (2014). arXiv:1406.2661.

[10] D. P. Kingma, M. Welling, Auto-encoding variational bayes (2014). arXiv:1312.6114.

[11] M. Paganini, L. de Oliveira, B. Nachman, Accelerating science with generative adversarial networks: An application to 3d particle showers in multilayer calorimeters, Physical Review Letters 120 (4). doi:10.1103/physrevlett.120.042003.
URL http://dx.doi.org/10.1103/PhysRevLett.120.042003

[12] M. Paganini, L. de Oliveira, B. Nachman, Calogan: Simulating 3d high energy particle showers in multilayer electromagnetic calorimeters with generative adversarial networks, Physical Review D 97 (1). doi:10.1103/physrevd.97.014021.
URL http://dx.doi.org/10.1103/PhysRevD.97.014021

[13] L. de Oliveira, M. Paganini, B. Nachman, Learning particle physics by example: Location-aware generative adversarial networks for physics synthesis, Computing and Software for Big Science 1 (1). doi:10.1007/s41781-017-0004-6.
URL http://dx.doi.org/10.1007/s41781-017-0004-6

[14] P. Musella, F. Pandolfi, Fast and accurate simulation of particle detectors using generative adversarial networks, Computing and Software for Big Science 2 (1). doi:10.1007/s41781-018-0015-y.
URL http://dx.doi.org/10.1007/s41781-018-0015-y

[15] B. Hashemi, N. Amin, K. Datta, D. Olivito, M. Pierini, Lhc analysis-specific datasets with generative adversarial networks (2019). arXiv:1901.05282.

[16] S. Otten, S. Caron, W. de Swart, M. van Beekveld, L. Hendriks, C. van Leeuwen, D. Podareanu, R. R. de Austri, R. Verheyen, Event generation and statistical sampling for physics with deep generative models and a density information buffer (2021). `arXiv:1901.00875`.

[17] R. Di Sipio, M. F. Giannelli, S. K. Haghighat, S. Palazzo, Dijetgan: a generative-adversarial network approach for the simulation of qcd dijet events at the lhc, Journal of High Energy Physics 2019 (8). `doi:10.1007/jhep08(2019)110`.
URL `http://dx.doi.org/10.1007/JHEP08(2019)110`

[18] M. Erdmann, J. Glombitza, T. Quast, Precise simulation of electromagnetic calorimeter showers using a wasserstein generative adversarial network, Computing and Software for Big Science 3 (1). `doi:10.1007/s41781-018-0019-7`.
URL `http://dx.doi.org/10.1007/s41781-018-0019-7`

[19] A. Maevskiy, D. Derkach, N. Kazeev, A. Ustyuzhanin, M. Artemev, L. Anderlini, Fast data-driven simulation of cherenkov detectors using generative adversarial networks, Journal of Physics: Conference Series 1525 (2020) 012097. `doi:10.1088/1742-6596/1525/1/012097`.
URL `http://dx.doi.org/10.1088/1742-6596/1525/1/012097`

[20] Y. Kahn, G. Krnjaic, N. Tran, A. Whitbeck, M3: a new muon missing momentum experiment to probe (g 2) and dark matter at fermilab, Journal of High Energy Physics 2018 (9). `doi:10.1007/jhep09(2018)153`.
URL `http://dx.doi.org/10.1007/JHEP09(2018)153`

[21] J. Alwall, M. Herquet, F. Maltoni, O. Mattelaer, T. Stelzer, Madgraph 5: going beyond, Journal of High Energy Physics 2011 (6). `doi:10.1007/jhep06(2011)128`.
URL `http://dx.doi.org/10.1007/JHEP06(2011)128`

[22] A. Paszke, S. Gross, F. Massa, A. Lerer, J. Bradbury, G. Chanan, T. Killeen, Z. Lin, N. Gimelshein, L. Antiga, A. Desmaison, A. Köpf, E. Yang, Z. DeVito, M. Raison, A. Tejani, S. Chilamkurthy, B. Steiner, L. Fang, J. Bai, S. Chintala, Pytorch: An imperative style, high-performance deep learning library (2019). `arXiv:1912.01703`.