GAN for prediction of direct photons in longitudinally polarized proton-proton collisions at energy $\sqrt{s} = 27$ GeV

A. A. Lobanov[1] Ya. A. Berdnikov[1] D. O. Kotov[1]

[1]Peter the Great St.Petersburg Polytechniv University

ICPPA, 23 october 2024

POLYTECH
Peter the Great
St.Petersburg Polytechnic
University

Direct photons. DLSA

Study of direct photons is a part of NICA SPD physical program.

Direct photons are photons which are born mainly in $q\overline{q} \to g\gamma$ and $q(\overline{q})g \to q(\overline{q})\gamma$ (80%).

DLSA (double longitudinal spin asymmetry) are defined as:

$$A_{LL} = \frac{d\Delta\sigma}{d\sigma},$$

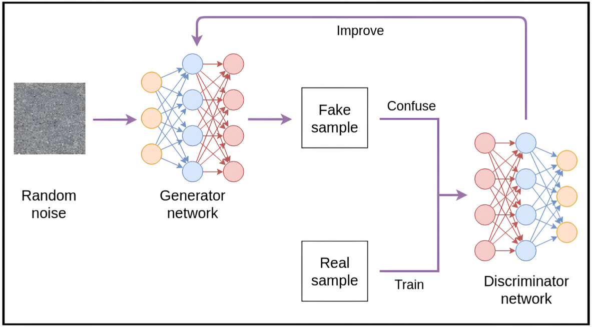
where σ ($\Delta\sigma$) – direct photons production cross-section for unpolarized (polarized) protons.

$$(\Delta)\sigma = \sum_{a,b=q\overline{q},g} \int dx_a dx_b (\Delta) f_a(x_a, \mu^2)(\Delta) f_b(x_b, \mu^2) d(\Delta)\sigma_{ab}^{\gamma},$$

where $(\Delta)\sigma_{ab}^{\gamma}$ – direct photon production (polarized) cross-section at the partonic level, $(\Delta)f_{a(b)}(x_{a(b)}, \mu^2)$ – parton distribution functions $a(b)$ in (polarized) proton, $x_{a(b)} = \frac{2p_T}{\sqrt{s}}$ – momentum fraction of colliding protons at an energy in the COM frame $\sqrt{s}$, μ – hard scale (taken in $p_T/2 < \mu < 2p_T$).

GAN

GAN – one type of generative neural network. It has the following architecture:



For loss function was used least squares:

$$\min_D V_{\mathsf{LSGAN}}(D) = \frac{1}{2}\mathbb{E}_{\mathbf{x}\sim p_{\mathsf{data}(\mathbf{x})}}\left[(D(\mathbf{x})-1)^2\right] + \frac{1}{2}\mathbb{E}_{\mathbf{z}\sim p_{\mathbf{z}}(\mathbf{z})}\left[(D(G(\mathbf{z})))^2\right]$$

$$\min_G V_{\mathsf{LSGAN}}(G) = \frac{1}{2}\mathbb{E}_{\mathbf{z}\sim p_{\mathbf{z}}(\mathbf{z})}\left[(D(G(\mathbf{z}))-1)^2\right]$$

PYTHIA8 settings

PYTHIA8 was used for modeling direct photons production in unpolarized/polarized proton-proton collision at energy $\sqrt{s} = 27$ GeV.

Setting below was used:

- Beam:eCM = 27;
- PromptPhoton: qg2qgamma = on;
- PromptPhoton: qqbar2ggamma = on;
- MultipartonInteraction: pT0Ref = 2.2;
- PDF:pSet = LHAPDF6:NNPDFpol11_100 for polarized PDF;
- PDF:pSet = LHAPDF6:NNPDF31_nlo_as_0118 for unpolarized PDF;

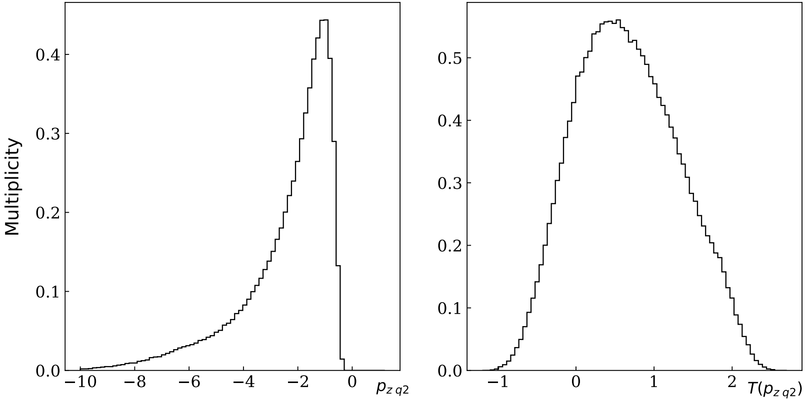
100’000 events for both unpolarized and longitudinally polarized proton-proton collisions were generated.

Transformed features

Features were choosen:

- For direct photons: p_x, p_y, p_z
- For partons: $p_{z\ q1}$, $p_{z\ q2}$

But for better work of GAN $T(p_{z\ 1}) = \ln(p_{z\ q1})$ and $T(p_{z\ q2}) = \ln(-p_{z\ q2})$ were used.



Generator and discriminator

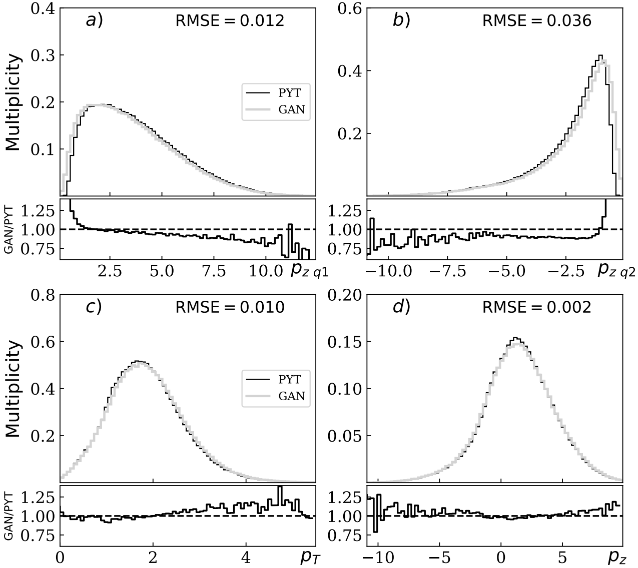
Generator is a feedforward neural network:

- Input layer – 128 noise vector and flags for polarized or unpolarized event, first and second particle ids;
- Number of hidden layers – 7;
- Number of nodes – 512;
- Activation function – LeakyReLU with $\alpha = 0.2$ ($f(x) = \max(0, \alpha x)$);
- Output layer – 5 nodes for every features: $T(p_{z\ q1}), T(p_{z\ q2}), p_x, p_y, p_z$.

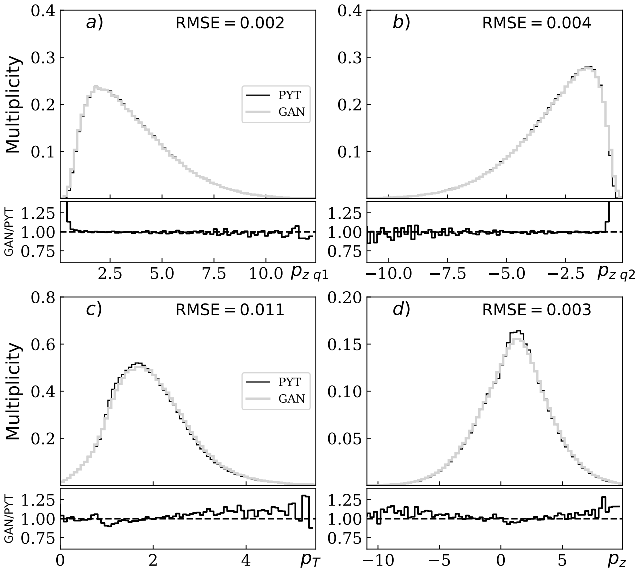
Discriminator has same architecture with differencies below:

- Input layer – 5 nodes (generator output);
- For every hidden layer Spectral normalization is applied;
- Output layer – 1 node with linear activation function.

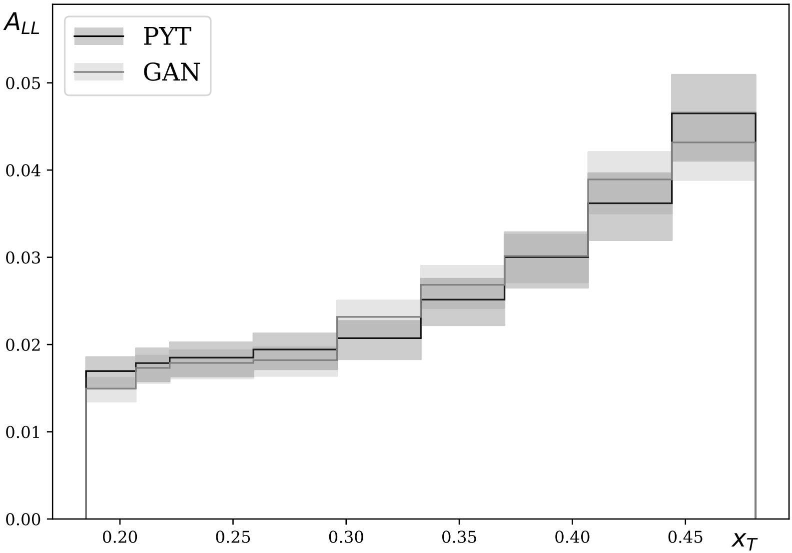
Results. Unpolarized $p + p$ collisions



Results. Polarized $p^{\rightarrow} + p^{\rightarrow}$ collisions



Results. DLSA



Conclusions

- GAN model can predict direct photon features with hight precision in both polarized and unpolarized proton-proton colisions;
- GAN learns not just features distribution, but learns their relations too, that shows DLSA predictions;

We acknowledge support from Russian Ministry of Education and Science. State assignment for fundamental research (code FSEG-2024-0033).

Thanks for listening

Backup slides

Training parameters

- epochs: 1000;
- batch size: 500;
- gradient optimizer: RMSProp;
- learning rates $2 \cdot 10^{-4}$ and 10^{-4} for generator and discriminator.

Network parameters determination

For determination of all hyperparameters (generator, discriminator, training parameters) library Optuna was used with PTE algorithm (Parzen tree estimators).

For comparasion of hyperparameters set MMD with Gaussian kernel was used.
MMD (maximum mean discrepancy) distance is determinied as:

$$\mathrm{MMD}^2(P, Q) = \mathbb{E}_{X,X'\sim P}k(X, X') + \mathbb{E}_{Y,Y'\sim Q}k(Y, Y') - 2\mathbb{E}_{X\sim P,Y\sim Q}k(X, Y).$$