Artificial neural networks application in estimating the impact parameter in heavy ion collision using the microchannel plate detector data

Authors: Galaktionov K.A., Roudnev V.A., Valiev F.F., Feofilov G.A.
Speaker: Galaktionov K.A.

St. Petersburg University

6th International Conference on Particle Physics and Astrophysics
https://indico.particle.mephi.ru/event/275/
29.11.2022 - 02.12.2022

Microchannel plate detectors[1]

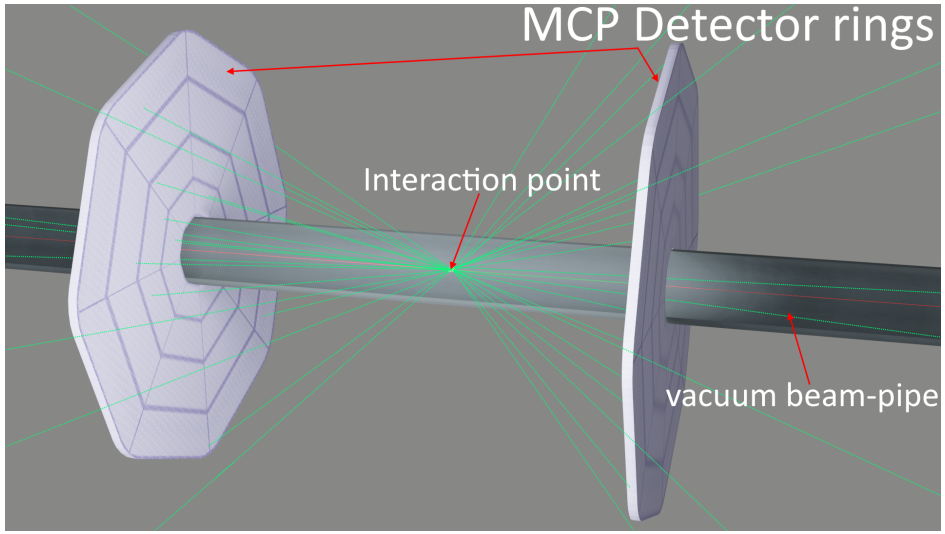
Some features of these detectors:

- Variability in size
- Registration of charged particle hits
- Time of flight resolution $\approx$ 50 - 100 ps

Data preparation:

- MC generated information about charged particles, produced in heavy ions collision
- Event-wise information about number of charged particles and polar angles (θ) of emitted particles was used to calculate their spatial distribution
- Charged particles momenta were used to calculate time of flight of particles

Configuration № 1 (Big rings)



Configuration № 2 (Small rings)

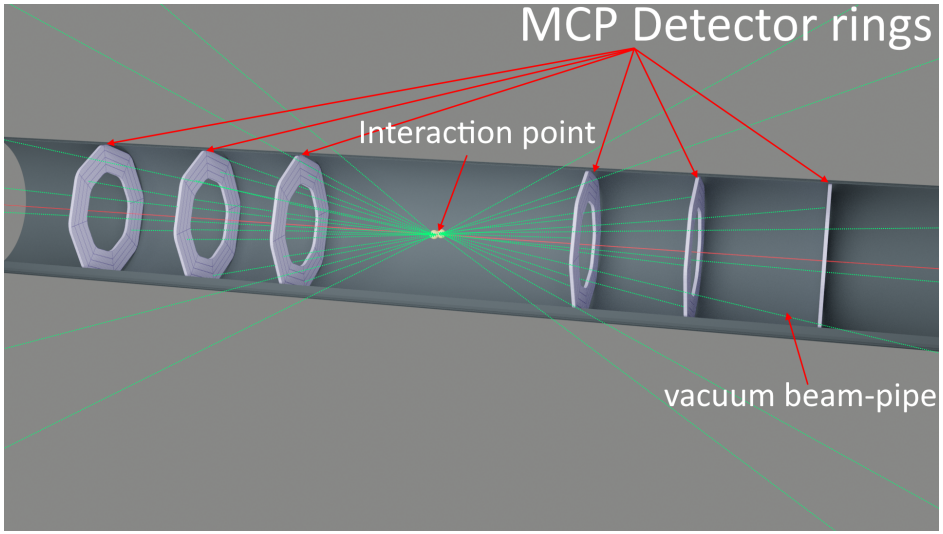


Fig. 1 Scheme of modeled detector configurations (not to scale). (left) - inside vacuum beam-pipe, three pairs of small rings ($d = 3$ cm, $D = 5$ cm), (right) - outside the beam-pipe in thin-wall vacuum chambers, one pair of big rings ($d = 5$ cm, $D = 50$ cm).

[1] A.A. Baldin et al. "Fast beam–beam collisions monitor for experiments at NICA". In: (). DOI: https://doi.org/10.1016/j.nima.2019.04.108.

The problem of impact parameter estimation

Statistical approach

Using statistical event features such as number of detected hits (N_{ch}) and mean angle of hits (Θ_{ch}).

Time-of-flight approach

Using information about every particle hit, including time of flight of particle and detector cell.

Scheme of computational experiments:

1. The QGSM model of Au+Au collisions ($\sqrt{s_{NN}} = 11\text{GeV}$) is used as a source data. (Dataset has 200000 events)
2. Dataset has *bdb* weighted distribution of impact parameter (i.e., the number of events with an impact parameter b being proportional to b)
3. Spacial and temporal data for the detector hits is generated according to the detector configuration.
4. The detector data is used for the neural network training.

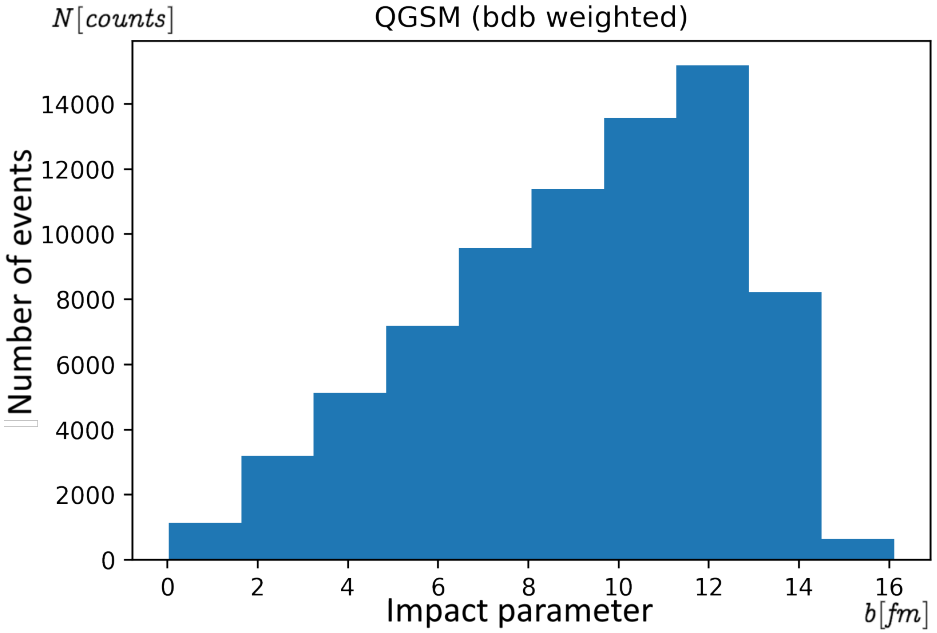


Fig. 2 Impact parameter distribution (by QGSM MC event generator).

Artificial neural networks (ANNs) were used to solve the problem of estimation of impact parameter.

Configuration № 1 (Big rings). Statistical approach. Event features.

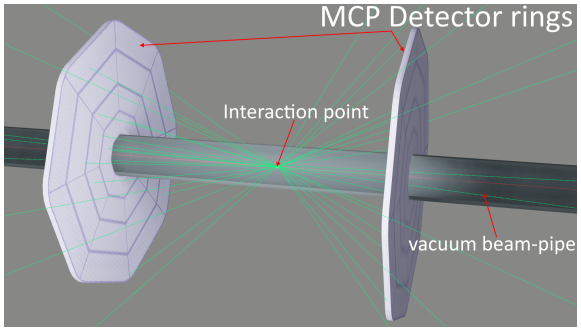


Fig. 3 Scheme of the configuration

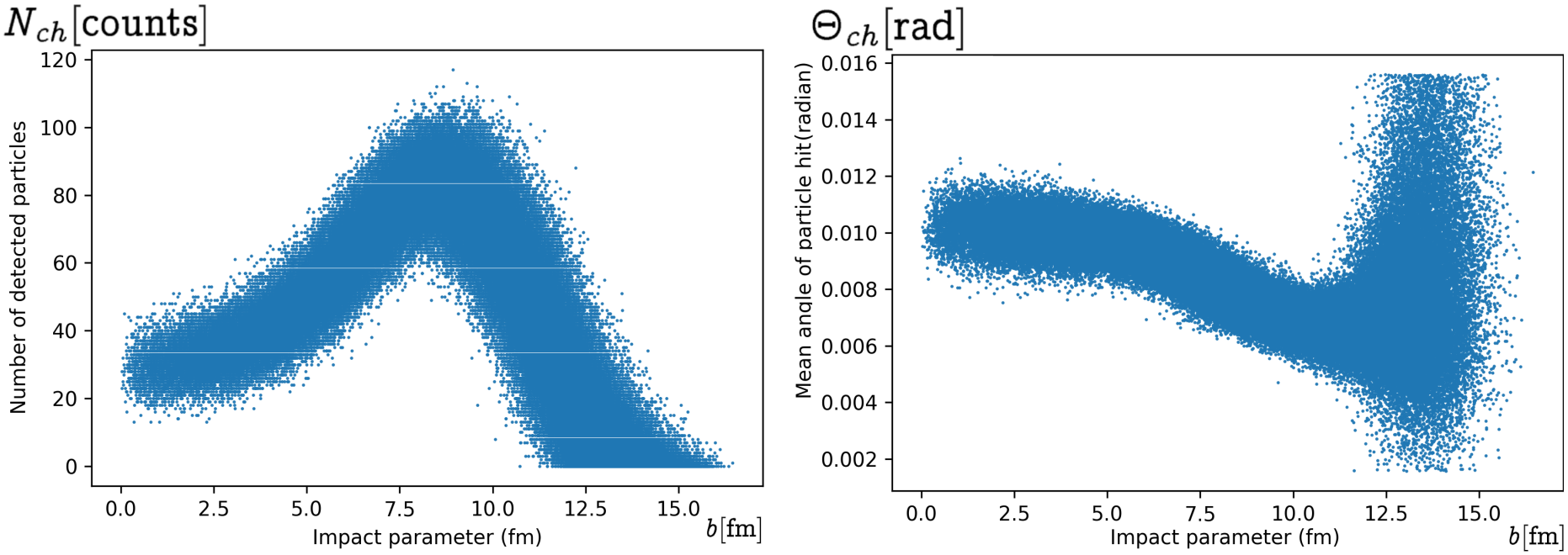


Fig. 4 Event features dependence on the impact parameter of an event. (left) - number of registered hits, (right) - mean angle of particle hits (first moment of distribution)

Configuration № 1 (Big rings). Statistical approach. Regression results.

The goal of the neural network was to estimate the value of the impact parameter of the event, by minimizing mean squared error on training set. Accuracy metrics: MSE - mean squared error, MAE - mean absolute error.

Reached accuracy: $\sqrt{\mathrm{MSE}} = 0.77$ fm, $\mathrm{MAE} = 0.58$ fm

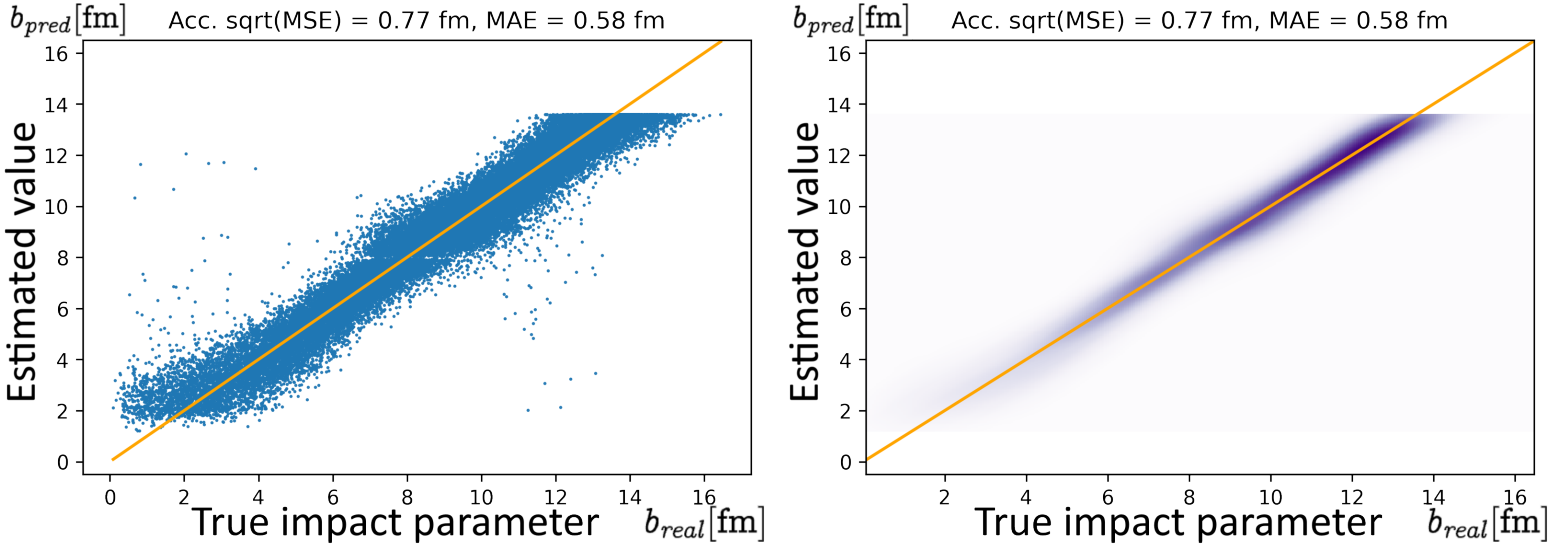


Fig. 5 Dependence of the evaluated impact parameter on the true value. (left) - scatter plot, where each dot represents one event from test set, (right) - density of event-dots.

Configuration № 1 (Big rings). Statistical approach. Classification results.

The goal of the neural network was to divide all events into two classes by the value of impact parameter by minimizing number of misclassified events on training set.
Class 1 - impact parameter below threshold, Class 2 - above. Accuracy metrics: Accuracy - percentage of correctly identified events.

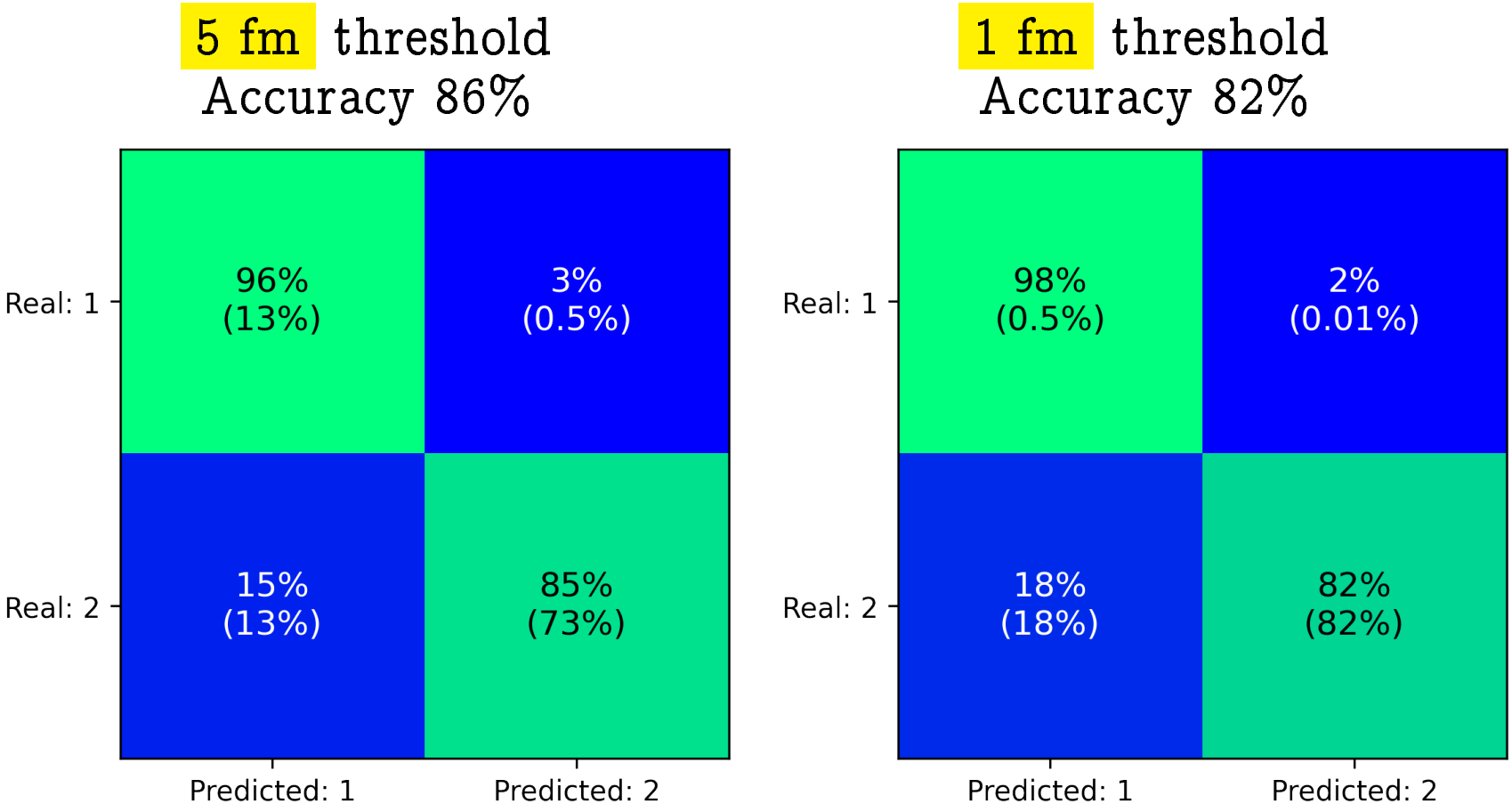


Fig. 6 Confusion matrices. Value in brackets - normalized to the number of events in test set, value outside of brackets - normalized to the number of events in real class.

Configuration № 2 (Small rings). Statistical approach. Event features.

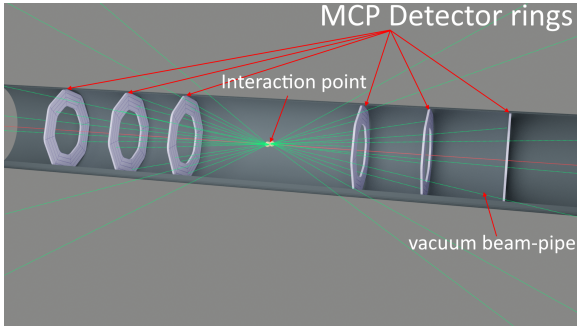


Fig. 7 Scheme of the configuration

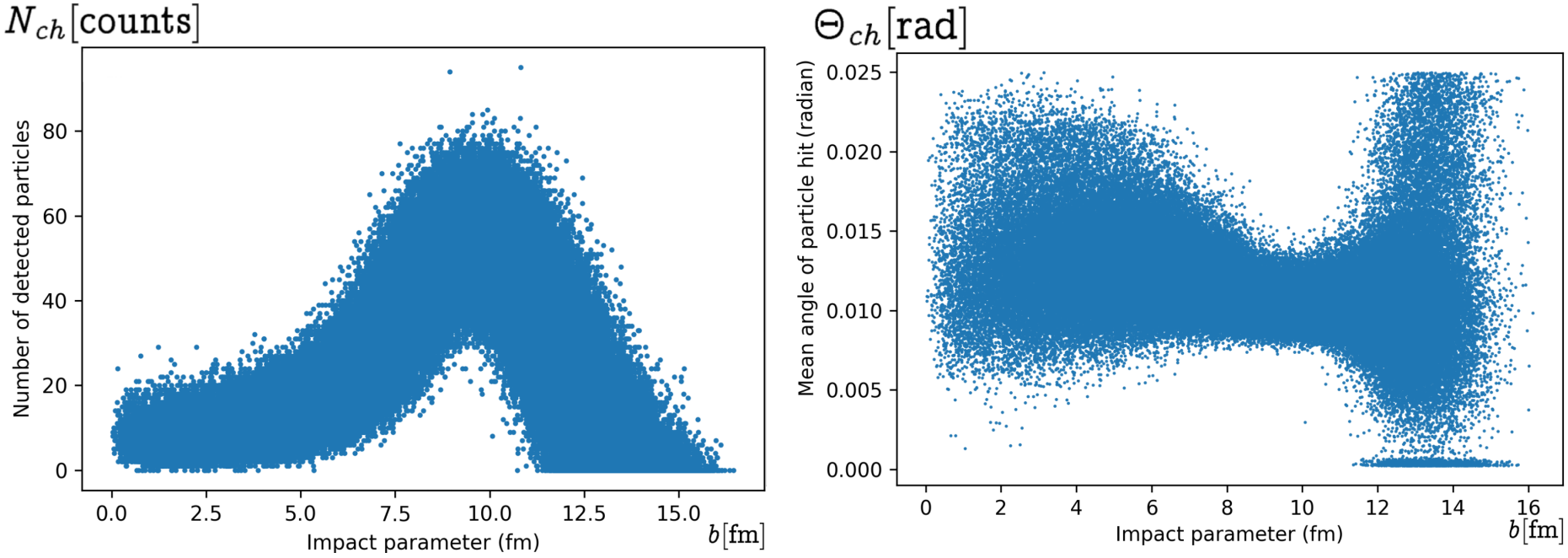
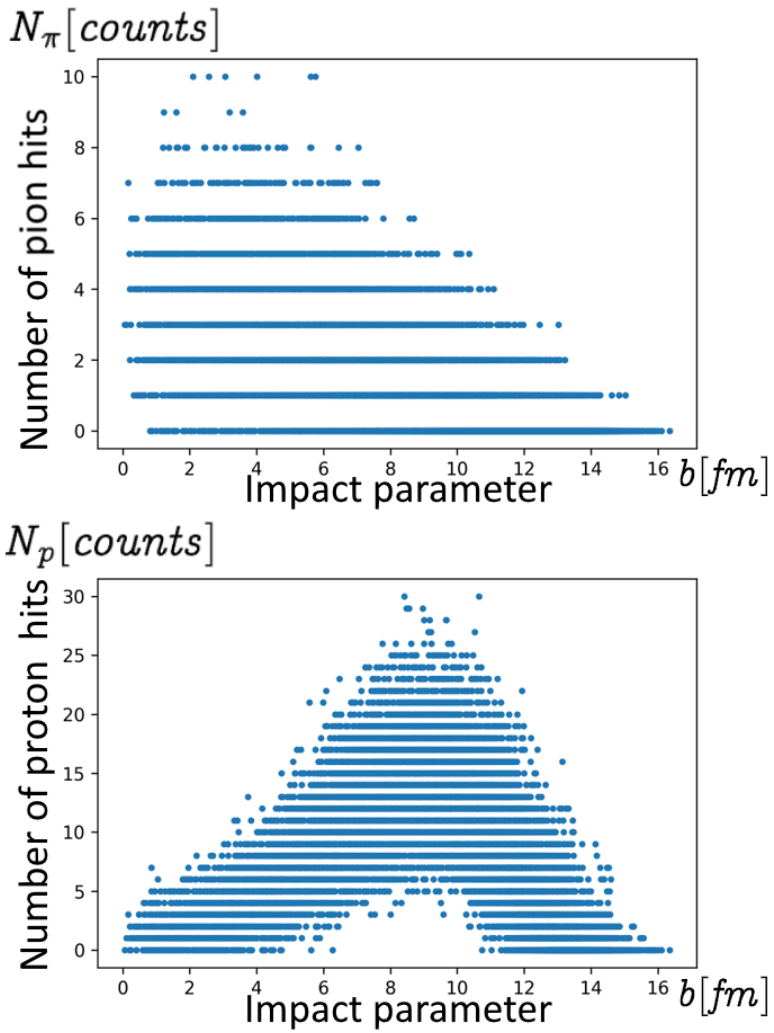
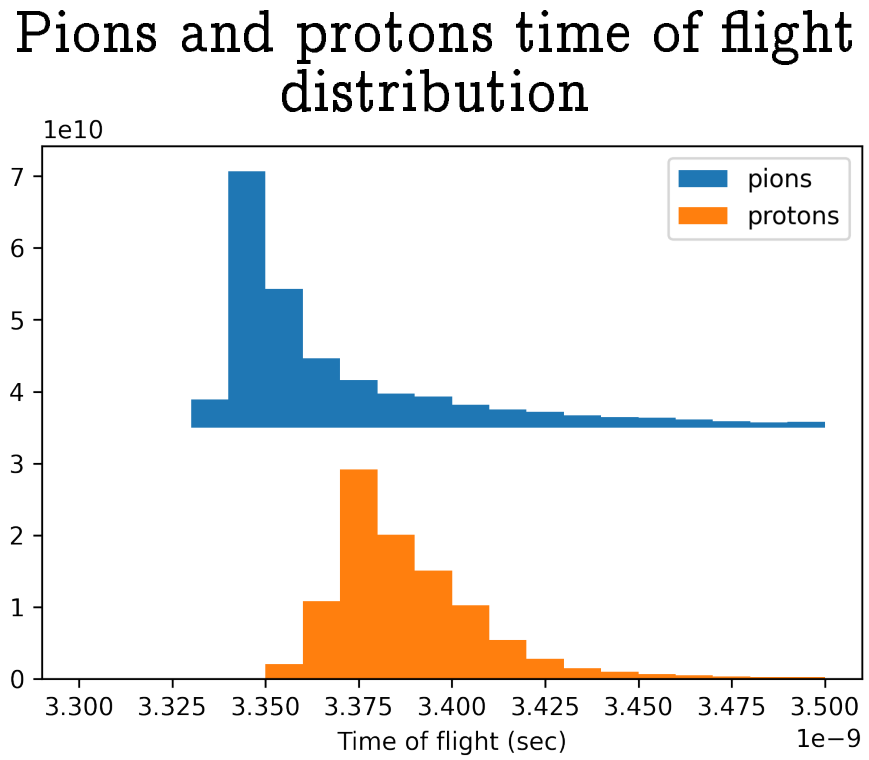


Fig. 8 Event features dependence on the impact parameter of an event. (left) - number of registered hits, (right) - mean angle of particle hits (first moment of distribution)

Configuration № 2 (Small rings). Time of flight approach.





$$\nu = \frac{1}{t - t_{0i}} \qquad (1)$$

Fig. 9 (left) - dependence of number of registered pions and protons (most part of the particles) on the impact parameter of an event (right) - pions and protons time-of-flight distribution (at 1 m distance) and transformation formula, where t - time-of-flight, t_{0i} - average time of flight of pions on i-th detector

Configuration № 2 (Small rings). Time of flight approach. Regression results.

The goal of the neural network was to estimate the value of the impact parameter of the event, by minimizing mean squared error on training set. Accuracy metrics: MSE - mean squared error, MAE - mean absolute error.

Reached accuracy: $\sqrt{\mathrm{MSE}} = 1.86$ fm, $\mathrm{MAE} = 1.19$ fm

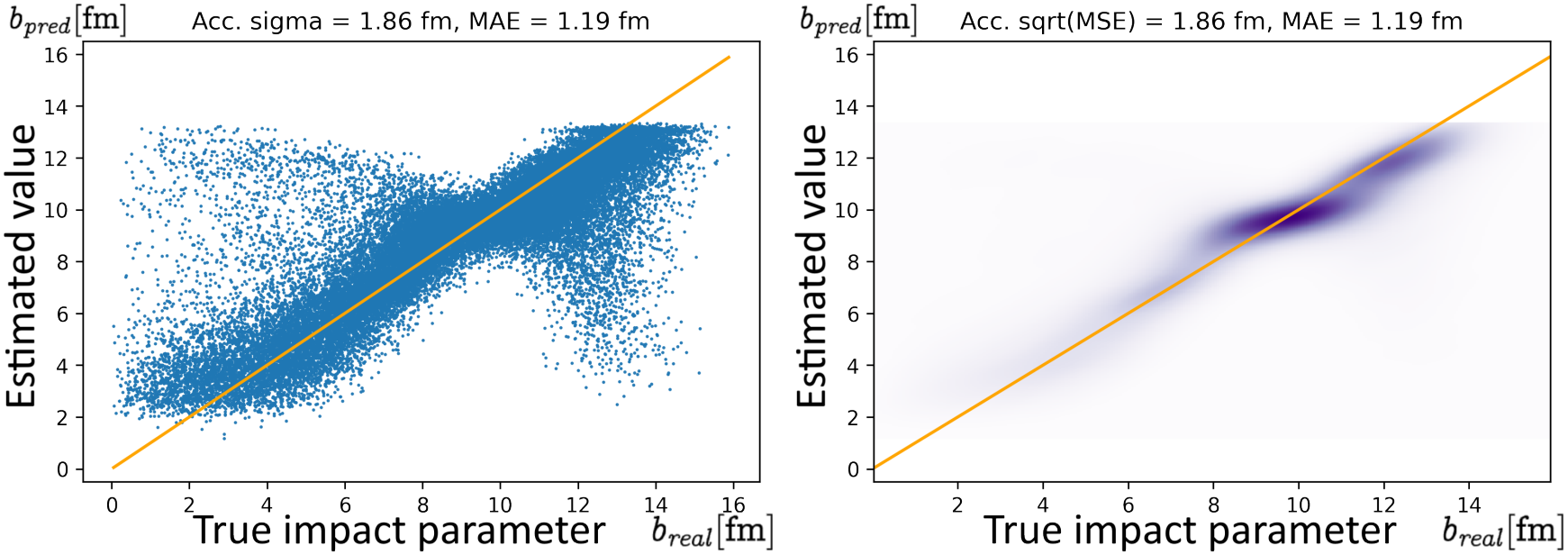


Fig. 10 Dependence of the evaluated impact parameter on the true value. (left) - scatter plot, where each dot represents one event from test set, (right) - density of event-dots.

Configuration № 2 (Small rings). Time of flight approach. Classification results.

The goal of the neural network was to divide all events into two classes by the value of impact parameter by minimizing number of misclassified events on training
Class 1 - impact parameter below threshold, Class 2 - above. Accuracy metrics: Accuracy - percentage of correctly identified events.

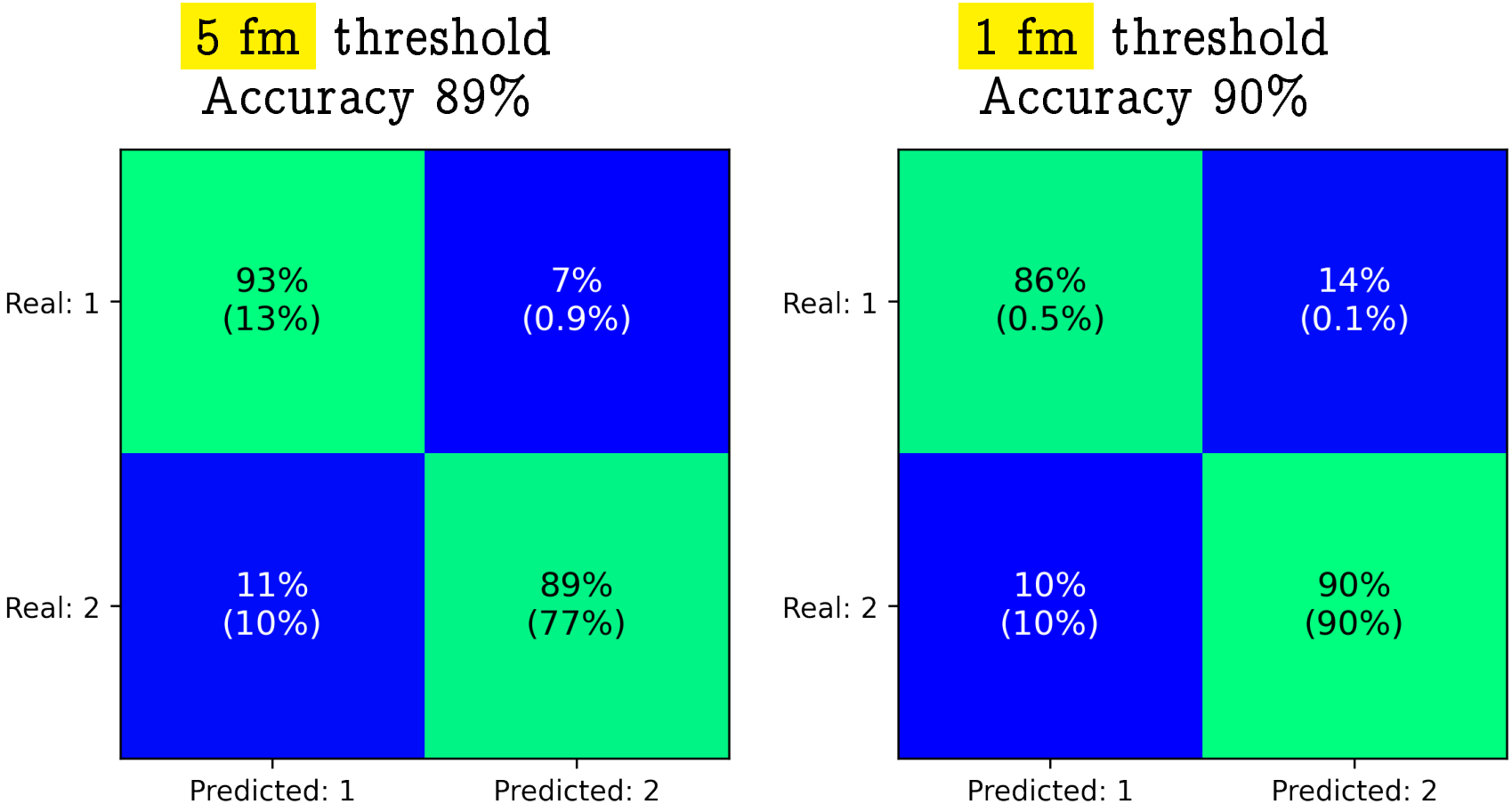


Fig. 11 Confusion matrices. Value in brackets - normalized to the number of events in test set, value outside of brackets - normalized to the number of events in real class.

Configuration № 1 (Big rings). Time of flight approach. Regression results.

The goal of the neural network was to estimate the value of the impact parameter of the event, by minimizing mean squared error on training set. Accuracy metrics: MSE - mean squared error, MAE - mean absolute error.

Reached accuracy: $\sqrt{\mathrm{MSE}} = 0.69$ fm, $\mathrm{MAE} = 0.54$ fm

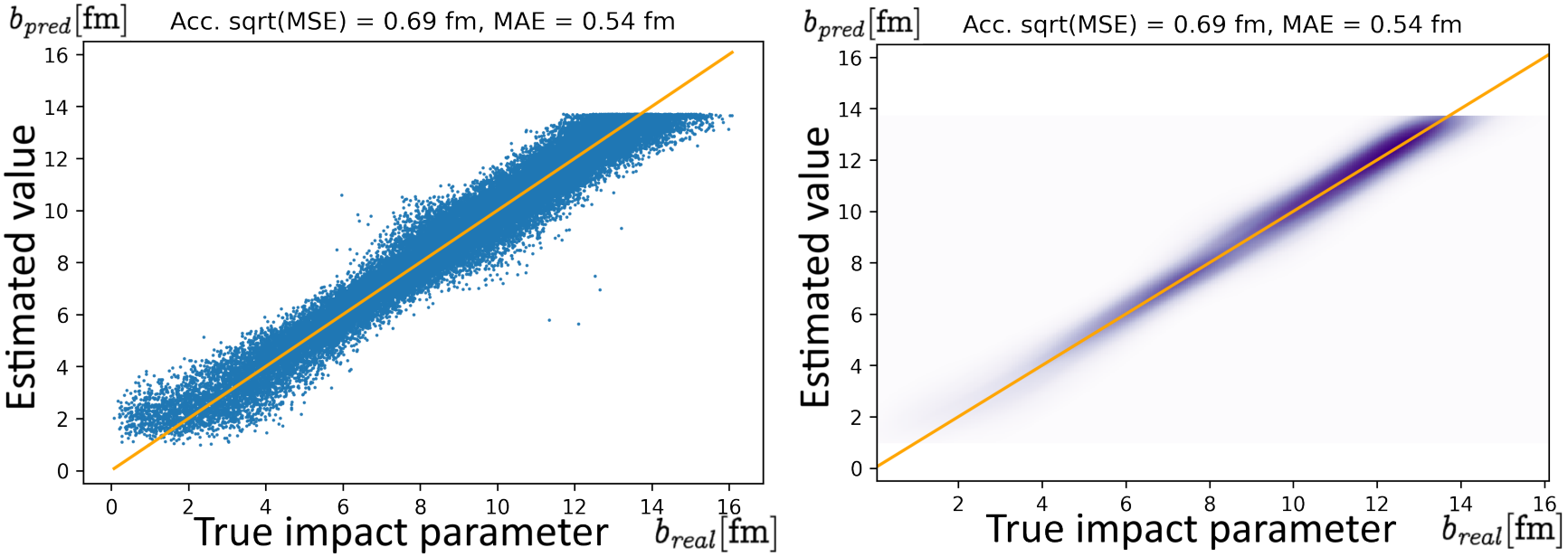


Fig. 12 Dependence of the evaluated impact parameter on the true value. (left) - scatter plot, where each dot represents one event from test set, (right) - density of event-dots.

Configuration № 1 (Big rings). Time of flight approach. Classification results.

The goal of the neural network was to divide all events into two classes by the value of impact parameter by minimizing number of misclassified events on training set.
Class 1 - impact parameter below threshold, Class 2 - above. Accuracy metrics: Accuracy - percentage of correctly identified events.

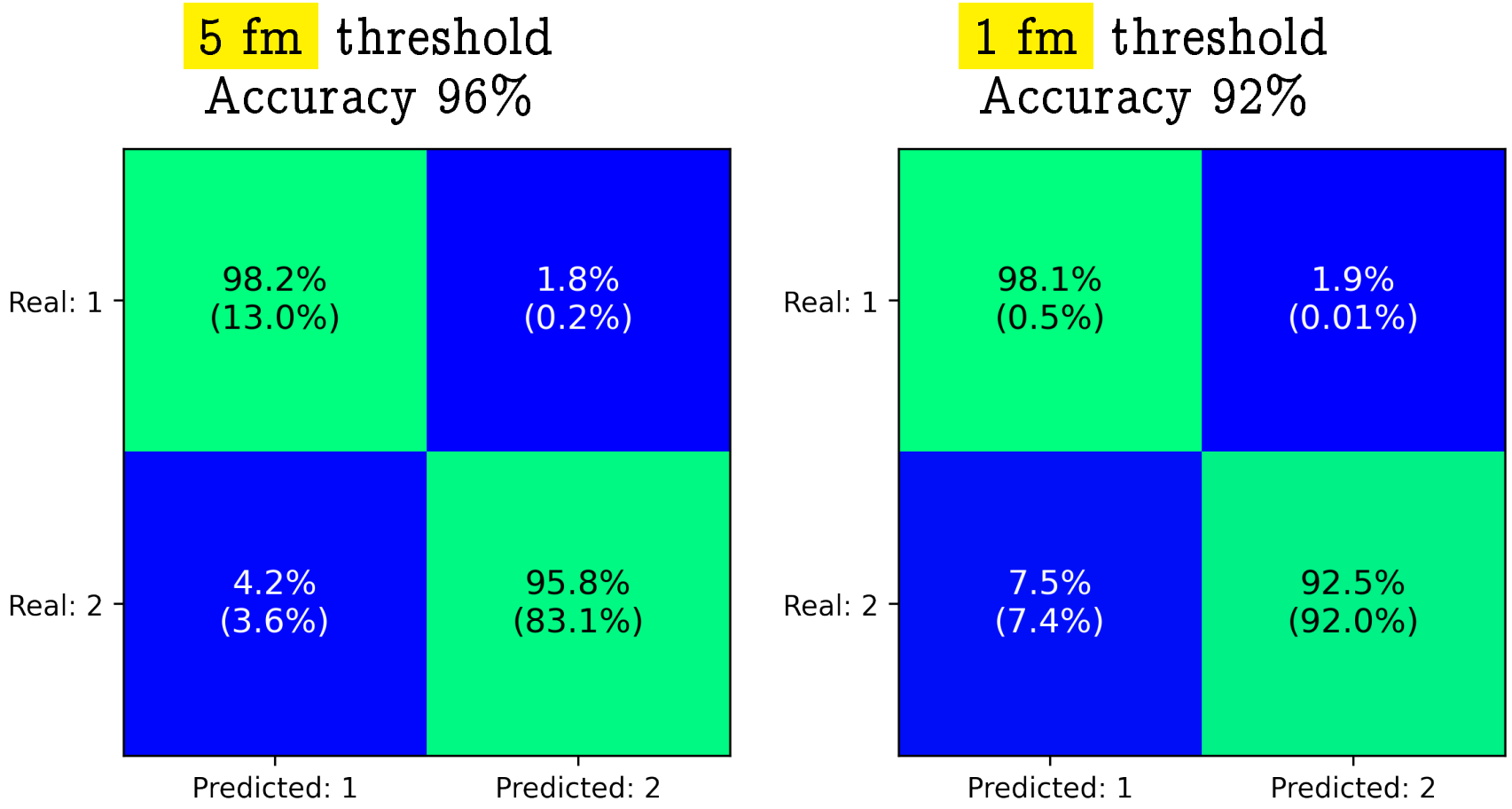


Fig. 13 Confusion matrices. Value in brackets - normalized to the number of events in test set, value outside of brackets - normalized to the number of events in real class.

Conclusion

We developed the machine learning technique of estimation of impact parameter of a single event. Both statistical and time of flight approach were tested, and showed good results. It was shown that the developed technique can benefit from usage of information about time of flight of particles.

Advantages of big detectors geometry:

+ Good results can be achieved using only statistical data about particle hits
+ Statistical data require much small neural network and computational resources

Advantages of small detectors geometry:

+ Similar results of classification problem detecting less particles hits

Future research plans:

- Confirm results on datasets obtained from different MC event generators

Backup slides: Bibliography

Baldin, A.A. et al. "Fast beam–beam collisions monitor for experiments at NICA". In: (). DOI: https://doi.org/10.1016/j.nima.2019.04.108.

Karpushkin, N et al. "Application of Machine Learning methods for centrality determination in heavy ion reactions at the BM@N and MPD@NICA". In: *Journal of Physics: Conference Series* 1690 (Dec. 2020), p. 012121. DOI: 10.1088/1742-6596/1690/1/012121. URL: https://doi.org/10.1088/1742-6596/1690/1/012121.

Li, Fupeng et al. "Application of artificial intelligence in the determination of impact parameter in heavy-ion collisions at intermediate energies". In: *Journal of Physics G: Nuclear and Particle Physics* 47.11 (Oct. 2020), p. 115104. DOI: 10.1088/1361-6471/abb1f9. URL: https://doi.org/10.1088/1361-6471/abb1f9.

Li, Fupeng et al. "Application of machine learning in the determination of impact parameter in the ^{132}Sn+^{124}Sn system". In: *Physical Review C* 104.3 (Sept. 2021). DOI: 10.1103/physrevc.104.034608. URL: https://doi.org/10.1103%2Fphysrevc.104.034608.

Overall comparison table

Detector type	Small rings detector	Big rings detector
Regression, Statistics $\sqrt{MSE}$ fm)	-	0.77
Regression, Time of flight $\sqrt{MSE}$ fm)	1.86	0.69
5 fm classification, Stats, true positive	-	96 %
5 fm classification, Stats, true negative	-	85 %
1 fm classification, Stats, true positive	-	98 %
1 fm classification, Stats, true negative	-	82 %
5 fm classification, Time, true positive	93 %	98 %
5 fm classification, Time, true negative	89 %	96 %
1 fm classification, Time, true positive	86 %	98 %
1 fm classification, Time, true negative	90 %	92 %

Computational resources

Evaluation was performed by calculating the amount of floating point multiplications needed for the work of algorithm.

Big detectors geometry, statistical approach:

- 300 - 400 floating point multiplications
- Preprocessing: number of hits and mean angle
- 2 x 352 cells

Small detectors geometry, time of flight approach:

- 10000 - 80000 floating point multiplications
- Preprocessing: time-of-flight evaluation
- 6 x 32 cells

All values are approximate and require fine tuning.

Artificial neural networks (ANNs)

ANN - an example of supervised learning.
Formula describing a dense layer of a neural network.

$$y = \theta(x * A^T + b) \tag{1}$$

Formula describing a convolutional layer of a neural network.

$$out(N_i, C_{out_j}) = \theta(bias(C_{out_j}) + \sum_{k=0}^{C_{in}-1} weight(C_{out_j}, k) \star input(N_i, k)) \tag{2}$$

Where: y, out - outputs of layer; x, $input$ - inputs of layer; A^T - transpose of a matrix of weights; $weight$ - convolution kernel; b, bias - biases of layer, $\theta(x)$ - activation function.

Configuration № 1 (Big rings). Statistical approach. Regression results. Distributed coordinate of primary vertex.

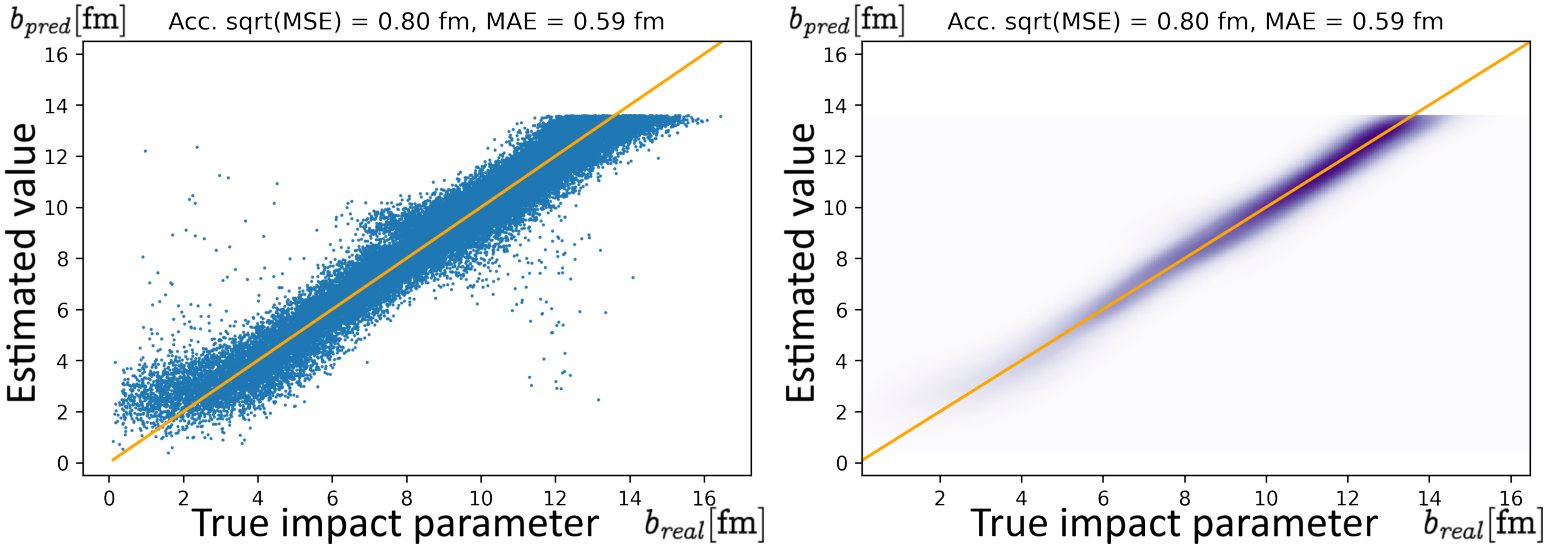


Fig. 14 Dependence of the evaluated impact parameter on the true value. Coordinate of primary vertex is taken randomly from a normal distribution $Z \sim \mathcal{N}(\mu, \sigma^2); \mu = 0; \sigma = 15$ cm.

Configuration № 1 (Big rings). Statistical approach. Classification results. Distributed coordinate of primary vertex.

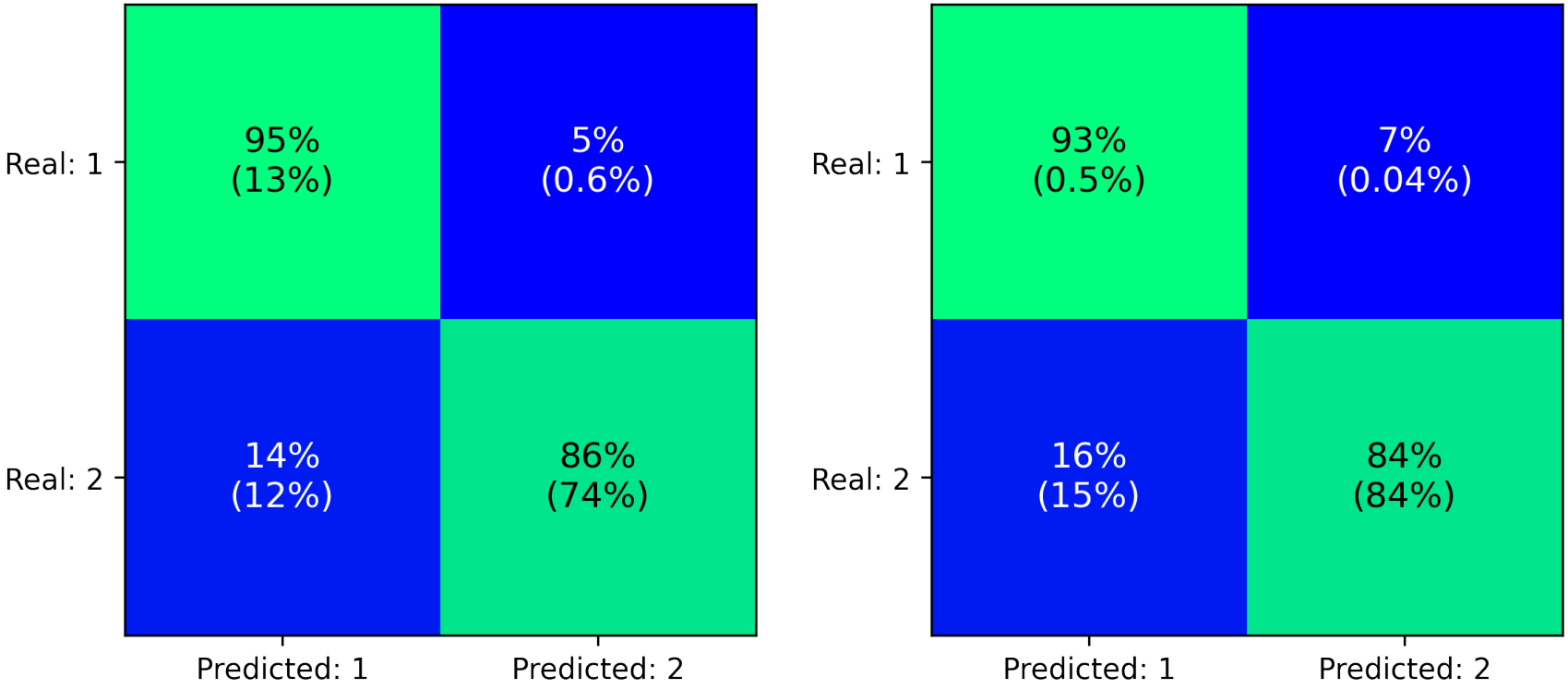


Fig. 15 Confusion matrices, coordinate of primary vertex is taken randomly from a normal distribution $Z \sim \mathcal{N}(\mu, \sigma^2); \mu = 0; \sigma = 15$ cm: (left) - threshold = 5 fm. Overall accuracy reaches 87 %, (right) - threshold = 1 fm. Overall accuracy reaches 84 %

Configuration № 2 (Small rings). Time of flight approach. Regression results. Fixed and distributed coordinate of primary vertex.

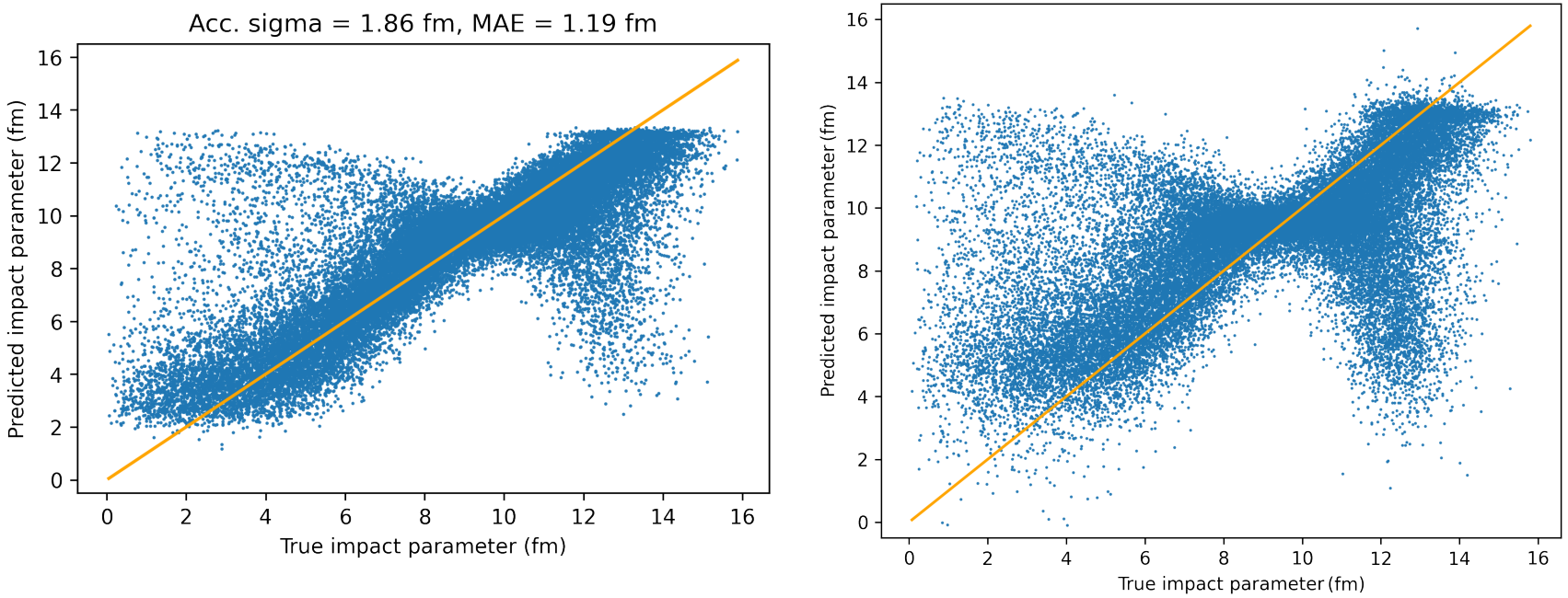


Fig. 16 Dependence of the evaluated impact parameter on the true value. (left) - fixed coordinate of primary vertex, the result is $\sqrt{MSE} = 1.86$fm, (right) - distributed coordinate of primary vertex, taken randomly from a normal distribution $Z \sim \mathcal{N}(\mu, \sigma^2); \mu = 0; \sigma = 15$ cm, the result is $\sqrt{MSE} = 2.4$fm