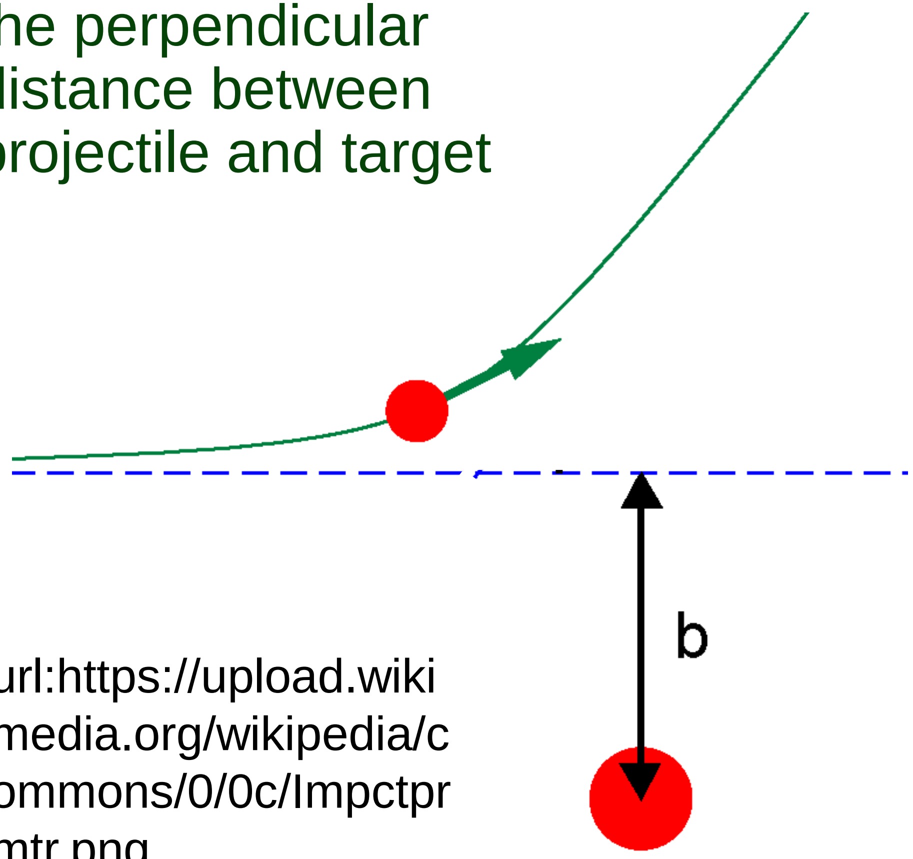
Impact parameter determination with Machine learning algorithm

Chun Yuen Tsang

Advisor: Professor Betty Tsang

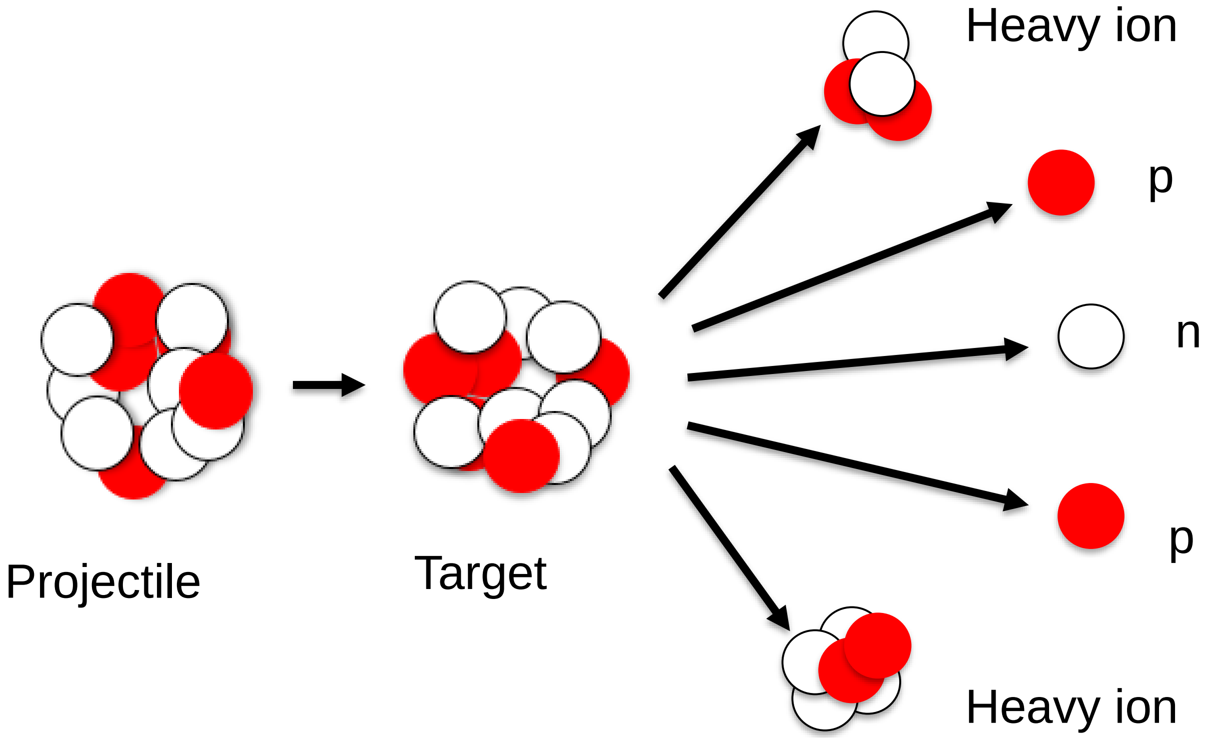
Terminology in this presentation

- Impact parameter b is the perpendicular distance between projectile and target



url:https://upload.wikimedia.org/wikipedia/commons/0/0c/Impctprmtr.png

- Multiplicity is the number of **CHARGED** fragments after nuclear reaction



- 4 charged particles (2 heavy ion, 2 protons, neutron doesn't count since it is neutral), so multiplicity = 4

The importance of impact parameter

- A lot of observables are sensitive to impact parameter.
- To compare data with theory, we need to know impact parameter of the detected events for the comparison to be valid.
- Traditionally, impact parameter is calculated from multiplicity through the following method:
  - Two major assumptions: Multiplicity increases monotonically with decreasing impact parameter, and event statistics follows geometrical distribution $d\sigma = 2\pi b db$
  - The formula is $b(M) = b_{max}\sqrt{\frac{N_{>M}}{N_{total}}}$
  - b_{max} is the measured cross-section, calculated from the ratio of beam particle frequency over reaction frequency and target thickness.

Machine learning (ML) algorithm

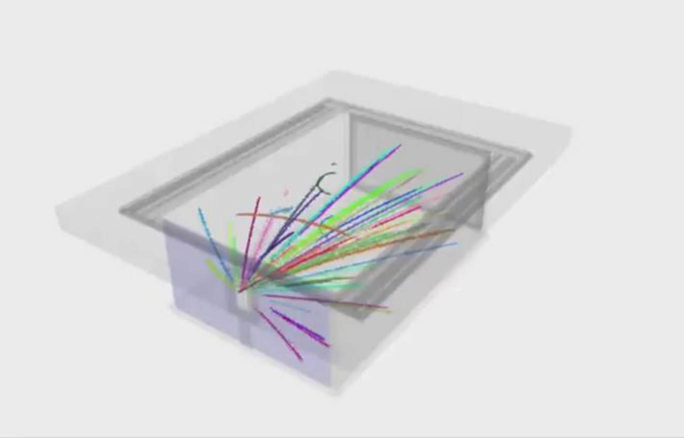
- A class of algorithms that figures out the correlation between input and output on its own with training data automatically without user intervention, and can be applied to data outside of the training data set.
- Examples: Handwritten character recognition, voice recognition, autopilot,…
- Need training data for ML algorithm to learn. It will figure out correlations between variables.

Previous endeavors on impact parameter determination with ML

- Yongjia Wang demonstrated the viability of using ML algorithms to predict impact parameter.
- Ref: https://iopscience.iop.org/article/10.1088/1361-6471/abb1f9
- The training data is generated from the Ultra Relativistic Quantum Molecular Dynamics (UrQMD) model. It is a model that simulates the dynamics of heavy-ion collision. The fragments it generated has been shown to agree with data at different beam energies.
- Specifically, the algorithm he used was LightGBM.
- Trained on UrQMD simulated data of Au + Au reaction with **PERFECT** detector.

With detector simulation

- Perfect detector is not realistic. Will the algorithm works with detector response?
- The detector to simulate: SπRIT detector, a time projection chamber (TPC). An experiment was performed with ^{132}Sn + ^{124}Sn at 270 MeV/A in 2016 at RIKEN, Japan, along with other Tin isotopic systems.
- Detector bias:
  1. Experimental is setup to take data from central collisions.
  2. Geometric efficiencies not 4 pi
  3. Track reconstructions are complicated so not every tracks can be reconstructed → Detector efficiencies.
  4. …



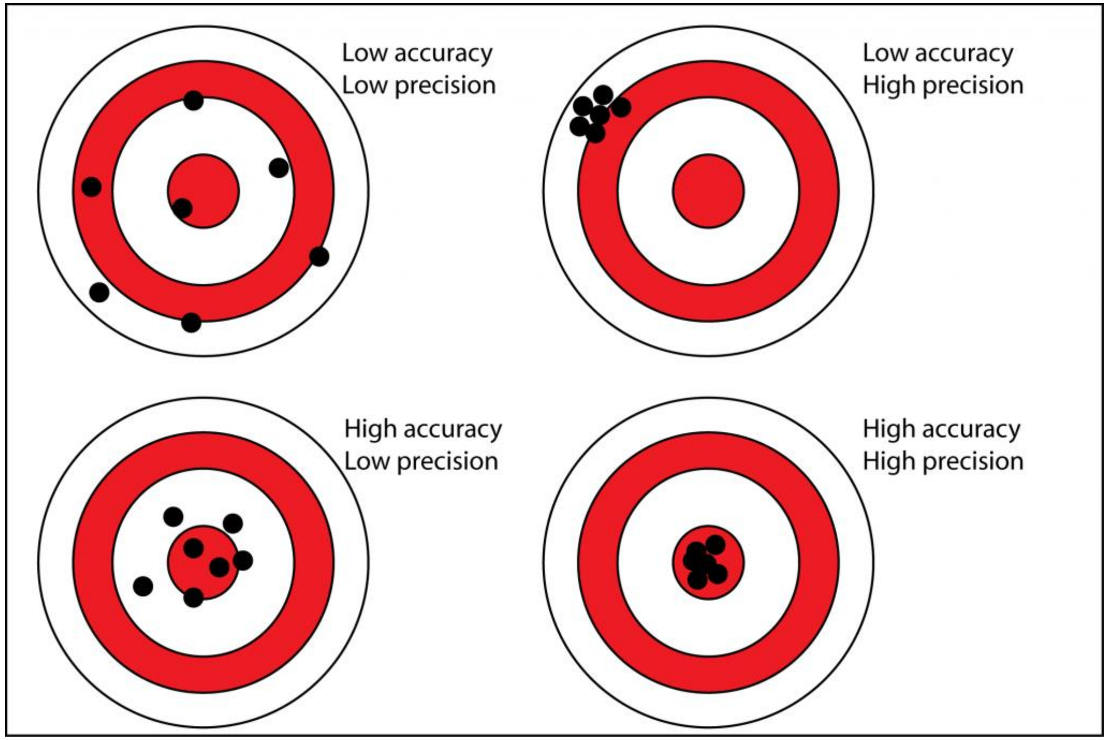
- To use the algorithm under a realistic experimental setting, we add detector response to UrQMD data with SpiRITROOT, simulation package for SpiRIT detector.

Input and output

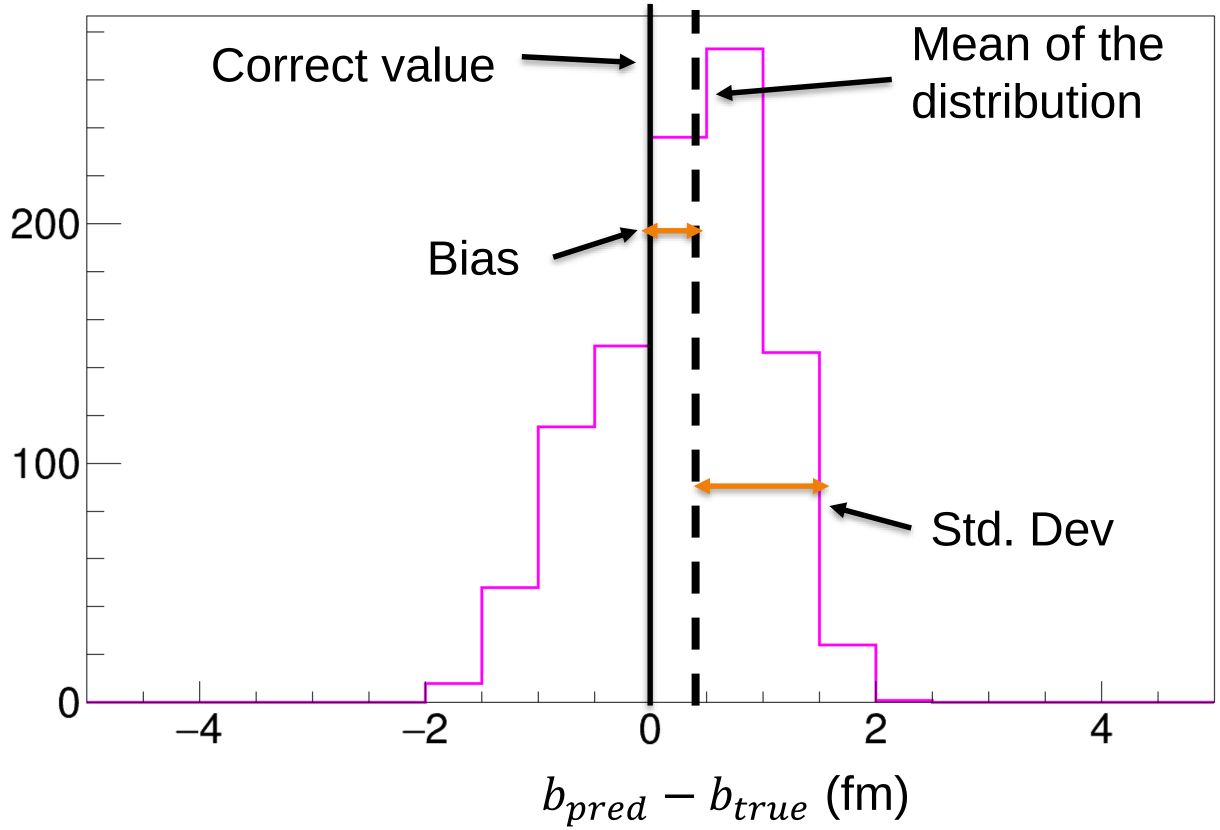
- 7 inputs (features):
  1. Transverse kinetic energy of hydrogen and helium isotopes.
  2. Ratio of total transverse to longitudinal kinetic energy.
  3. Total multiplicity of charged particles.
  4. Total number of hydrogen and helium isotopes.
  5. Averaged transverse momentum of hydrogen and helium isotopes
  6. Number of free protons at mid-rapidity |yz/ybeam|≤0.5
  7. Averaged transverse momentum of free protons at mid rapidity |yz/ybeam|≤0.5.
  - Not set in stone. We can use more/less observables if you desired.
- Output: impact parameter in fm
- Let’s train the algorithm on Sn + Sn at 270 MeV/A system instead of Au + Au.
- Impact parameter of training data set is distributed uniformly from 0 – 10 fm such that the algorithm is exposed to enough central events when it is trained.

NSCL

Quantify the uncertainty: Bias and Standard deviation



url: http://www.antarcticglaciers.org/glacial-geology/dating-glacial-sediments-2/precision-and-accuracy-glacial-geology/

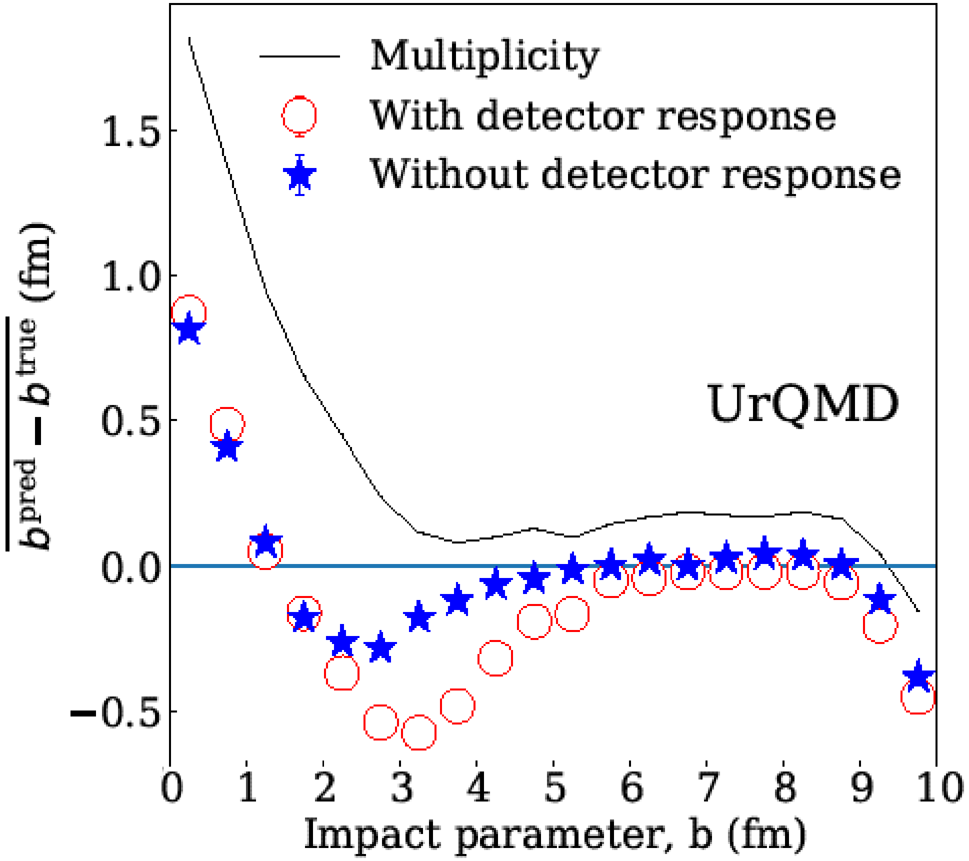


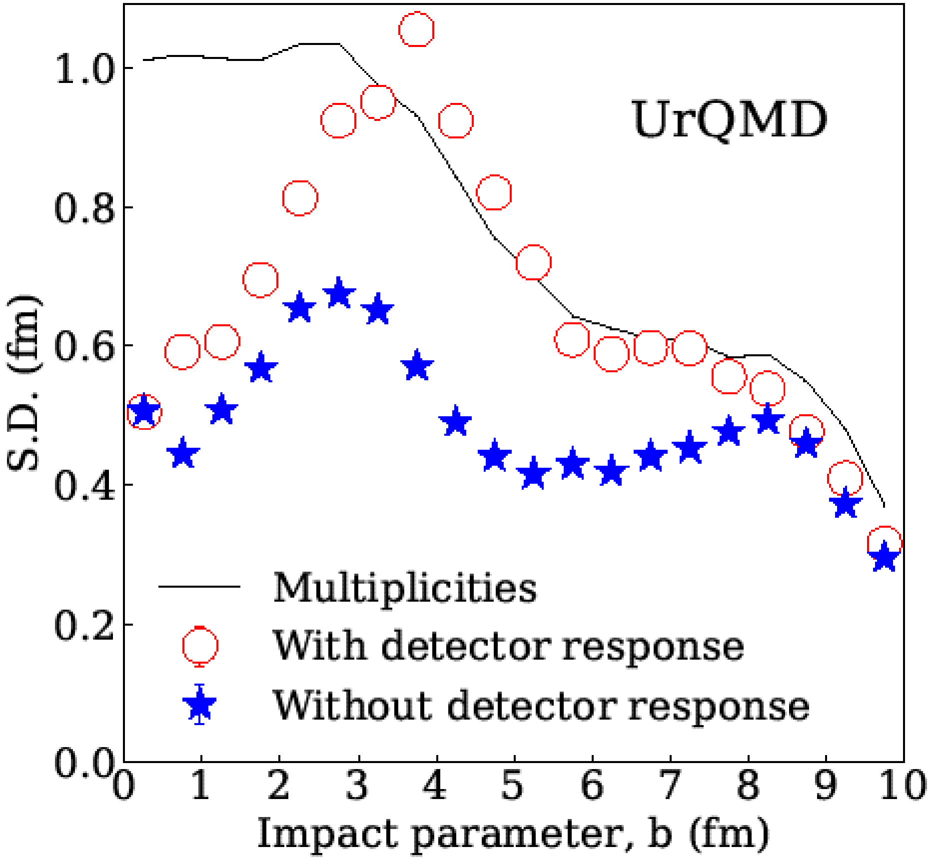
Bias corresponds to accuracy

Std corresponds to precision

LightGBM results

- Made predictions on **another** batch of UrQMD simulation (test data) with uniform b-distribution.





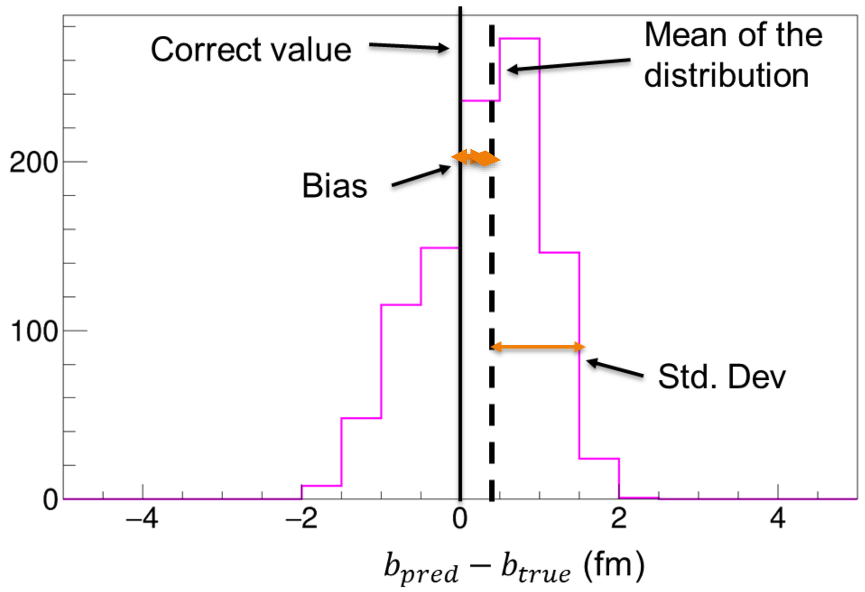
Model dependence of the result

- When trained with UrQMD events, the algorithm performs well on other batches of UrQMD simulations.
- But UrQMD is not 100% exactly reality, also different transport model predicts different results.
- What happens if the training data set and prediction data set come from different models? Will it introduce errors? How well are the prediction?
- That's what we are going to do, train algorithm on UrQMD events, make predictions on simulations of b = **3 fm events** from four other models: dcQMD, AMD, IQMD VUU and ImQMD

Model dependence after detector response

Without detector response	Bias (fm)	Std (fm)
AMD	1.1	1.0
DcQMD	-0.2	0.7
ImQMD	0.3	0.9
IQMD	0.4	0.8
Mean of absolute values	0.5	0.9
With detector response	Bias (fm)	Std (fm)
AMD	1.1	0.9
DcQMD	0.8	1.2
ImQMD	0.2	1.0
IQMD	-0.3	1.0
Mean of absolute values	0.6	1.0

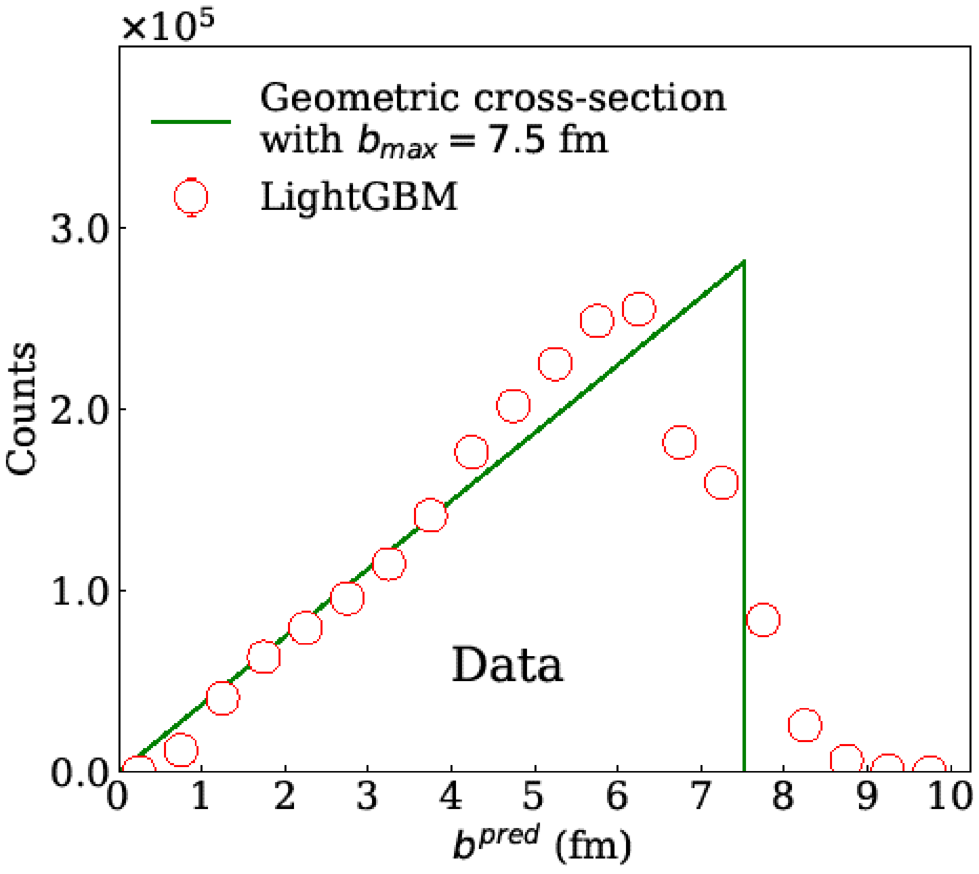
Illustration of Bias and Std. Dev.
Not actual distribution from the analysis.



The width of the distributions are wider.

Apply to real experimental data

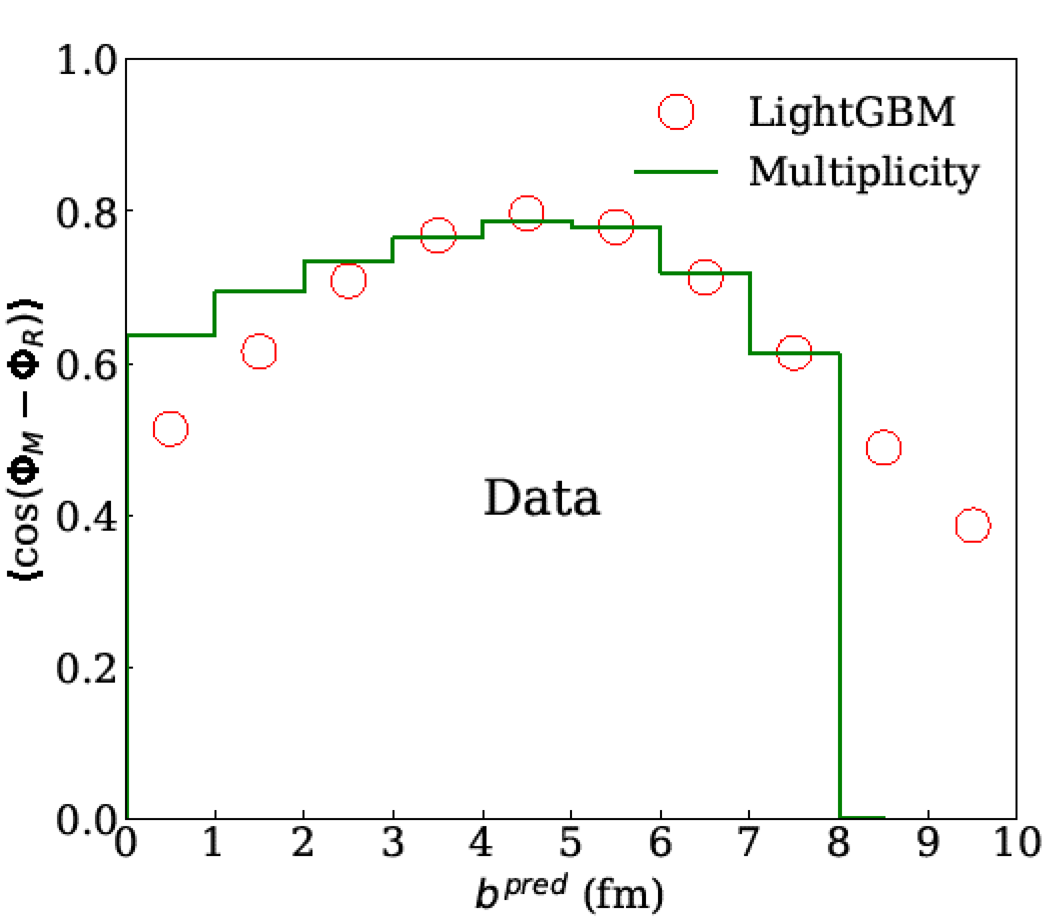
- There's no barrier to using the algorithm on experimental data, just substitute UrQMD observables with experimental values.
- Let's do it and inspect the result…
- Important notes:
  - Due to trigger conditions in SπRIT experiment, peripheral events are disproportionately discarded.
  - Also, the measured $b_{max} = 7.5$ fm is smaller than the real cross-section of Tin on Tin collision due to triggers.
  - Therefore at high impact parameter, distribution of recorded impact parameter is actually not triangular.
  - The tail of LightGBM actually agrees with our expectation better than the pure triangular distribution.
- To further inspect our predictions, we inspect an observable whose dependency on impact parameter is known.



Reaction plane resolution

- Reaction plane is reconstructed empirically,
  - The azimuth of the vector sum of transverse momentum vector is the reconstructed reaction plane.
- It relies on the fact that fragments preferably emits on reaction plane (collective flow).
- For central events, this asymmetry in fragment azimuth is less pronounced due to cylindrical symmetry arguments.
- Let Φ_M be the measured (reconstructed) reaction plane azimuth and Φ_R be the real reaction plane azimuth.
- With LightGBM, $\langle\cos(\Phi_M - \Phi_R)\rangle$ is closer to zero => Width of $\Phi_M - \Phi_R$ distribution is wider => Less pronounced cylindrical asymmetry.

Lower values correspond to worse reaction plane resolution



Summary

- ML algorithm trained on UrQMD data shows acceptable model dependence.
- ML algorithm can be trained to use in real experiment.
- Central data selected by ML algorithm shows smaller $<\cos(\Phi_M - \Phi_R)>$.

Acknowledgements

- **MSU/NSCL**
  - C. Anderson
  - J. Barney
  - G. Cerizza
  - J. Estee
  - G. Jhang
  - **W. G. Lynch***
  - J. Manfredi
  - P. Morfouace
  - C. Santamaria
  - H. Setiawan
  - R. Shane
  - S. Tangwancharoen
  - C. Y. Tsang
  - **M. B. Tsang***
  - Z. Zhang
- **Kyoto University**
  - M. Kaneko
  - **T. Murakami***
- **Texas A&M University**
  - A. B. McIntosh
  - S. J. Yennello
- **RIKEN**
  - H. Baba
  - N. Fukuda
  - T. Ichihara
  - N. Inabe
  - **T. Isobe***
  - D. Kameda
  - T. Kubo
  - Y. Nakai
  - M. Kurata-Nishimura
  - S. Nishimura
  - H. Otsu
  - H. Sato
  - H. Sakurai
  - H. Shimizu
  - H. Suzuki
  - F. Takeda
  - K. Yoneda
- **Korea University**
  - B. Hong
  - J.W. Lee
- **RISP, Korea**
  - H. Sang Lee
  - Y. Kim
  - Y.J. Kim
- **INP, Poland**
  - K. Pelczar
  - J. Lukasik
  - P. Pawlowski
  - P. Lasko
- **Tokyo Technical University**
  - Y. Kondo
  - T. Nakamura
- **Tsinghua University**
  - R. Wang
  - Z. Xiao
  - Y. Zhang
- **Rikkyo University**
  - K. Ieki
  - K. Kurita
  - J. Murata
- **Tohoku University**
  - T. Kobayashi
- **NeuLAND Collaboration**
  - L. Atar
  - T. Aumann
  - A. Horvat
  - D. M. Rossi
  - H. Scheit
  - H. Toernqvist
  - K. Boretzky
  - Y. Leifels
  - I. Gasparic

*SπRIT Spokespersons

This work is supported by the U.S. DOE under Grant Nos. DE-SC0014530, DE-NA0002923, US NSF Grant No. PHY-1565546, and the Japanese MEXT KAKENHI grant No. 24105004. Additional support from NSF award 1414979 (EAPSI).

U.S. DEPARTMENT OF ENERGY Office of Science

SπRIT

文部科学省 MEXT

SπRIT TPC Collaboration

U.S. DEPARTMENT OF ENERGY
Office of Science

SπRIT

文部科学省
MEXT